# Capital Accumulation and Growth:

# A New Look at the Empirical Evidence

Steve Bond*

Nuffield College, Oxford and IFS

Asli Leblebicioglu

North Carolina State University

Fabio Schiantarelli

Boston College and IZA

30 July 2007

**Abstract**

We present evidence that an increase in investment as a share of GDP predicts a higher growth rate of output per worker, not only temporarily, but also in the long run. These results are found using pooled annual data for a large panel of countries, using pooled data for non-overlapping five-year periods, allowing for heterogeneity across countries in regression coefficients, and allowing for cross-section dependence. They are robust to model specifications and estimation methods, particularly for our full sample and for the sub-sample of non-OECD countries. The evidence that investment has a long-run effect on growth rates is consistent with the main implication of some endogenous growth models.

**Key Words:** Growth, Capital Accumulation, Investment

**JEL Classification:** C23, E22, O40

**Acknowledgement:** Our interest in this subject was stimulated by B. Easterly's excellent book on economic growth. We thank S. Ardagna, K. Baum, P. Beaudry, A. Lewbel, N. Loayza, M.H. Pesaran, J. Seater, L. Serven, J. Temple and seminar participants at the World Bank, Inter-American Development Bank, Boston University, Brandeis University, Cambridge University, European University Institute, Ente Einaudi, the European Meeting of the Econometric Society and the African Econometrics Society for helpful comments. We thank L. Gaul for excellent research assistance, and the ESRC for financial support under award RES-156-25-0006.

*Corresponding author: Nuffield College, Oxford, OX1 1NF, UK; tel: +44 1865 278500; fax: +44 1865 278621; email: steve.bond@nuffield.ox.ac.uk

# 1 Introduction

An influential view has emerged which suggests that investment in physical capital is relatively unimportant in explaining differences in long-run economic growth rates. More precisely, while there is substantial empirical evidence that capital accumulation affects the long-run level of income per capita, previous studies, such as those by Jones (1995) and Blomstrom, Lipsey and Zejan (1996), do not provide econometric support for a significant effect of capital accumulation on long-run growth rates. Easterly and Levine (2001), in their influential review of the recent empirical literature, conclude that "the data do not provide strong support for the contention that factor accumulation ignites faster growth in output per worker".[1]

In this paper we ask whether we have gone too far in de-emphasizing the role of capital accumulation as a determinant of long-run growth. Our analysis of annual data for up to 94 countries in the period 1960-2000 suggests that the answer may be yes. Not only do we find that a higher share of investment in GDP predicts a higher *level* of output per worker in the long run, we also find that an increase in the share of investment predicts a higher *growth rate* of output per worker, both in the short run and, more importantly, in the long run. The long-run effect on growth rates is quantitatively substantial, as well as statistically significant. For our full sample of countries, and for the sub-sample of non-OECD countries, this effect is found for every specification we consider. For the smaller sub-sample of OECD countries, the result is not quite so robust to alternative specifications.

The overall evidence is inconsistent with the predictions of the Solow and Swan growth model in which diminishing returns to capital make the effects of an increase in capital accumulation temporary, and the steady state growth rate is determined by the exogenous rate of technological progress. It is instead consistent with the main implication of a range of endogenous growth models; for example, AK-type models in which broadening the definition of

[1] The role of capital accumulation is similarly downplayed in Easterly (2001).

capital to include human capital, or allowing for learning by doing, counteracts the tendency to diminishing returns in the aggregate (Romer, 1986); and Schumpeterian-type models in which capital is an input to the production technology for innovations (Howitt and Aghion, 1998).

One key to our analysis is that our empirical specifications allow the long-run growth rate in each country to depend on the share of income that is invested. Other important factors are that we analyze time series data for a large sample of countries, and we allow for some estimation issues that may have been neglected in earlier studies. We do not conclude that only investment matters. Indeed, we stress the importance of heterogeneity across countries, that may well reflect differences in economic policies and institutions. We do however question the empirical evidence that has led to the conclusion that capital accumulation is of only minor importance for long-run economic growth, and regard that conclusion to be, at best, premature.

This issue is of such fundamental importance that it has naturally received considerable attention in previous empirical research. Unlike much of the literature that has focused on issues of convergence, we estimate specifications in which higher investment is allowed to have a permanent effect on the growth rate, and not only a temporary effect during the transition to a new steady state growth path. Unlike some of the previous studies that have focused on testing instead AK models, we allow for the fact that an empirical model of the growth rate is unlikely to have a serially uncorrelated error process. Since the growth rate is the change in the log of output per worker, any transient shocks to the (log) level of output per worker will introduce a moving average error component when we model the growth rate.[2] Importantly, we also allow for unobserved country-specific factors that could result in both high levels of investment and high rates of growth, and for the likely endogeneity of the current investment share.

[2]Such serial correlation would be absent only if *all* shocks to the (log) level of the process are of a random walk nature, so that their first-differences are innovations.

Our basic results are obtained by estimating pooled models using a yearly panel of observations for a large number of countries. However they are robust to using a panel based on non-overlapping five-year periods, as suggested by Islam (1995) and Caselli, Esquivel and Lefort (1996). The mean effect of investment on long-run growth is also found to be positive and significant, and broadly similar in magnitude, when we allow all coefficients in our models to vary across countries, using the mean group approach of Pesaran and Smith (1995) and Lee, Pesaran and Smith (1997). For all these specifications, similar results are found for sub-samples of OECD and non-OECD countries. When we further allow for cross-country dependence in the shocks to the income process, using the multi-factor error structure proposed by Pesaran (2006), we find a positive and significant mean effect of investment on long-run growth rates for our full sample and for the sub-sample of non-OECD countries, though not for the smaller sub-sample of OECD countries.

Section 2 provides a brief review of previous related research, highlighting some restrictive features of earlier specifications that will be relaxed in the models we estimate. Section 3 outlines our specifications, which allow investment to affect both the level and the growth rate of output per worker in the long run, while also allowing for business cycle dynamics. Section 4 describes our data set and the time series properties of the main variables we use. Section 5 presents the econometric results, while Section 6 concludes.

## 2 Related Literature

An important branch of the recent empirical literature on economic growth estimates specifications based on variants of the (augmented) Solow model, in which the long-run growth rate of output per worker is determined by technical progress, which is taken to be exogenous. The standard model used to evaluate this framework and to study the issue of (beta) convergence is

derived from the transition dynamics to the steady state growth path, as suggested by Mankiw, Romer and Weil (1992). These models relate growth to investment, but condition on the initial level of output per worker. As a result, consistent with the underlying Solow framework, they do not allow investment to influence the steady state growth rate.

A typical specification has the form

$$\Delta y_{it} = (\alpha - 1) y_{i,t-1} + \beta x_{it} + \gamma t + \eta_i + \varepsilon_{it} \tag{1}$$

where $y_{it}$ denotes the logarithm of output per worker in country i at time t, $\Delta y_{it}$ is the growth rate of output per worker between time t-1 and t, and $x_{it}$ denotes the logarithm of the share of investment in output. Additional explanatory variables related to population growth, human capital or other factors may be included, but they do not change the essence of the points we make here. The time trend allows for a common rate of steady state growth, and the country-specific intercept (or 'fixed effect', $\eta_i$) allows for variation across countries in initial conditions, or other unobserved factors that affect the level of the country's steady state growth path. The residual reflects the influence of shocks that affect the (log) level of output per worker.

Cross-section studies generally focus on average growth rates measured over long periods of time, and relate these to average investment shares measured over the same period. Panel studies use repeated observations over shorter time periods, commonly five-year averages. In cross-section applications, the intercept cannot be allowed to be specific to individual countries, and the coefficient on the trend is not identified. In panel applications it is possible to allow for heterogeneous intercepts, and the coefficient on the trend is separately identified. The inclusion of time dummies rather than a simple linear trend allows for a more general evolution of total factor productivity (TFP), but still restricts TFP growth to be common across countries and independent of investment.

To clarify these points, we first rewrite equation (1) in autoregressive-distributed lag form

as

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \gamma t + \eta_i + \varepsilon_{it}. \tag{2}$$

This is a dynamic model for the *level* of $y_{it}$, provided $\alpha \neq 1$. If we consider a steady state in which there are no shocks, the share of investment takes the constant value $x_{it} = x_i$ and output per worker grows at the common rate $g$, so that $y_{it} = y_{i,t-1} + g$, we obtain

$$y_{it} = \left(\frac{\beta}{1-\alpha}\right) x_i + \left(\frac{\gamma}{1-\alpha}\right) t + \frac{\eta_i - \alpha g}{1-\alpha}. \tag{3}$$

This confirms that the steady state growth rate implied by this model, $g = \gamma/(1-\alpha)$, is indeed common to all countries, and does not depend on the level of investment. A permanent increase in investment predicts a higher level of output per worker along the steady state growth path, but affects growth only during the transition to the new steady state.

Both cross-section and panel studies have reported evidence that the coefficient $\beta$ on measures of investment in this type of specification is positive and significantly different from zero. Examples of the former include Mankiw, Romer and Weil (1992), Levine and Renelt (1992) and Barro and Sala-I-Martin (1995); examples of the latter include Caselli, Esquivel and Lefort (1996) and Bond, Hoeffler and Temple (1999).[3] This suggests that investment affects the level of output per worker in the steady state, but does not address the question of whether investment affects the growth rate of output per worker in the long run. This is not surprising, given that these specifications are derived from the (augmented) Solow growth model in which the steady state growth rate is taken to be exogenous.

A different branch of the empirical growth literature has focused on testing the (extended) AK model, and on Granger-causality between investment and growth rates. If we take first-differences of equation (2) and introduce a lagged level of investment term as an additional

[3] See also De Long and Summers (1991, 1993), who emphasize the role played by equipment investment, particularly for their sub-samples of developing countries; and Beaudry, Collard and Green (2002), who suggest that the effect of investment has become larger in more recent periods.

explanatory variable (this last modification is crucial), we obtain

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \beta \Delta x_{it} + \theta x_{i,t-1} + \gamma + \Delta \varepsilon_{it}. \tag{4}$$

This is indeed a dynamic model for the *growth rate* of $y_{it}$, and the coefficient $\theta$ on the lagged level of investment tests whether a higher level of investment predicts a faster growth rate in the long run. If we now consider a steady state in which the share of investment takes the constant value $x_{it} = x_i$ and output per worker grows at the country-specific rate $g_i$, we obtain

$$g_i = \frac{\gamma}{1-\alpha} + \left(\frac{\theta}{1-\alpha}\right) x_i.$$

This confirms that the steady state growth rates implied by this model are heterogeneous, and depend positively on the share of investment in output, if $\theta > 0$. The coefficient $\beta$ on the change in investment again indicates whether investment affects the (log) level of output per worker along the steady state growth path. If we replace $\Delta x_{it}$ with $\Delta x_{it-1}$, equation (4) provides the basis for a Granger-causality type test.

Both time series and panel studies have typically reported estimates of $\theta$ that are insignificantly different from zero in this type of model, suggesting that investment does not Granger-cause growth. For example, Jones (1995) finds no effect of investment on the long-run growth rate, either using annual data for individual OECD countries in the post-war period, or pooling these models across countries. Similarly Blomstrom, Lipsey and Zejan (1996), using pooled data for non-overlapping five-year periods and a large sample of developed and less developed countries, find that investment does not Granger-cause growth, while growth does Granger-cause investment. Attanasio, Picci and Scorcu (2001) find that investment Granger-causes growth, but with a negative sign.[4] While there are some exceptions in the more recent

[4] Other papers that consider Granger-causality have focused on savings rather than investment. Carroll and Weil (1994) present panel data evidence for non-overlapping five-year periods, both for OECD countries and for a wider sample, that Granger-causality runs from growth to saving, but not vice versa. If anything, the savings rate is negatively related to future growth. Attanasio, Picci and Scorcu (2001) find significant

literature, such as Li (2002), this econometric support for the view that higher investment does not predict faster growth appears to have been influential.[5]

Our derivation of this dynamic model for the growth rate suggests a potentially important problem with the least squares estimation procedures that have typically been used to obtain these results. Suppose that the shocks $\varepsilon_{it}$ to the (log) level of output per worker, introduced in equation (1), are serially uncorrelated. The error term $\Delta\varepsilon_{it}$ in the growth equation (4) will then be serially correlated. Moreover this model includes the lagged growth rate $\Delta y_{i,t-1}$ as an explanatory variable. This lagged dependent variable will then necessarily be correlated with the lagged shock $\varepsilon_{i,t-1}$ that appears in $\Delta\varepsilon_{it}$, rendering Ordinary Least Squares estimates (or Within estimates) of the parameters of interest biased and inconsistent. The bias in the least squares estimates of $\alpha$ will typically be downwards, so this approach will underestimate the degree of persistence in growth rates.[6] The bias in the estimates of $\beta$ and $\theta$ is harder to sign *a priori*, but may be sufficiently serious to warrant further investigation. Similar biases will be present if the shocks $\varepsilon_{it}$ contain a serially uncorrelated component, or indeed any stationary component. The only case in which least squares estimators could yield consistent estimates of the parameters in (4) would be if these shocks to the (log) level of output per worker follow a random walk, so that their first-difference is an innovation, orthogonal to $\Delta y_{i,t-1}$. In specifications where current investment is included, as in (4), consistency of least squares estimates would further require that current investment is uncorrelated with $\Delta\varepsilon_{it}$. The

and negative Granger-causality running from saving to growth. A negative association between saving and growth is consistent with a Ramsey exogenous growth model with optimizing consumers, given plausible values for the degree of risk aversion (Carroll and Weil, 1994). Aghion, Comin and Howitt (2006) find a significant positive effect of lagged saving on growth of GDP per worker for poor countries, but not for rich countries, in a specification that controls for the initial level of GDP per worker.

[5] A different way to use panel data to evaluate growth theories has been proposed in Evans (1998), based on the cointegration properties of income series for different countries. He concludes that exogenous growth theories may characterize the experiences of countries with well educated populations, but not those of countries with lower levels of education.

[6] Both Easterly and Levine (2001) and Attanasio, Picci and Scorcu (2001) emphasize the apparent lack of persistence in growth rates.

approach that we develop in this paper allows the long-run effects of investment on the level and growth rate of income to be estimated consistently under much weaker assumptions.

A recent paper by Bernanke and Gurkaynak (2001) studies the cross-section correlation between average investment shares and average growth rates in output per worker or TFP, calculated over long periods of time. In contrast to the conclusion from the Granger-causality literature, they report a significant positive correlation in both cases. While their approach avoids the estimation problem noted above, there are of course reasons to be cautious about inferring any causation from a cross-section correlation. In particular there may be unobserved country-specific factors, such as economic, political and legal institutions, that favour both high investment and fast growth. This could account for the positive cross-section correlation even in the absence of any causal link running from investment to growth. Moreover, since we observe growth in actual and not in steady state output per worker, the average investment rate may be correlated with the difference between the two.

The approach we develop in the next section can be motivated either as a way of obtaining consistent estimates of the 'growth effect' $\theta/(1-\alpha)$ and the 'level effect' $\beta/(1-\alpha)$ in dynamic growth models of the kind illustrated in equation (4); or alternatively as an extension of the test of the Solow model proposed by Bernanke and Gurkaynak (2001) to a dynamic panel data context, which allows us to control for unobserved country-specific factors that affect both investment and growth, as well as to allow for the likely endogeneity of current investment measures.[7] Our instrumental variable approach allows for correlation between the error term in the growth equation and all the explanatory variables (dated t or t-1). We further allow for unobserved heterogeneity in growth rates (i.e. $\gamma$ becomes $\gamma_i$ in (4)) and for heterogeneity in all the slope parameters (i.e. $\alpha, \beta, \theta$ become $\alpha_i, \beta_i, \theta_i$).

[7] That is, the likely correlation between the current shock $\varepsilon_{it}$ and the current investment share $x_{it}$ in models like (1) or (4).

## 3 Model Specification

In this section we describe the econometric models that we use to examine the relationship between the growth of output per worker and investment in physical capital. The spirit of the exercise is to start from a specification that is general enough to encompass the predictions of different theories. We will estimate these models using both annual data and data for non-overlapping five-year periods; and for the annual data we will allow all slope parameters, as well as intercepts, to vary across countries. Derivations of the basic specifications for the pooled annual data are presented below.

Denote with $y_{it}$ the logarithm of GDP per worker, and with $x_{it}$ the logarithm of the investment to GDP ratio. We assume that the behavior of $y_{it}$ can be represented by the following $ADL(p,p)$ (Autoregressive-Distributed Lag) model:

$$y_{it} = c_{it} + \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + ... + \alpha_p y_{i,t-p} + \beta_0 x_{it} + \beta_1 x_{i,t-1} + ... + \beta_p x_{i,t-p} + \varepsilon_{it} \qquad (5)$$

where $\varepsilon_{it}$ is a mean zero, serially uncorrelated shock which we initially assume to be independent across countries, conditional on $c_{it}$. Here $c_{it}$ is a non-stationary process that determines the behavior of the growth rate of $y_{it}$ in the long run, and different processes for $c_{it}$ will lead to different econometric specifications. This framework nests simpler dynamic specifications like equation (2) that have been used to evaluate the Solow growth model, and in that context the $c_{it}$ process would reflect the growth of total factor productivity. We allow for richer dynamics in our empirical models based on annual data to control for business cycle influences.

For convenience, (5) may be reparameterized in differences and levels form:

$$\begin{aligned} \Delta y_{it} \;=\; & c_{it} + (\alpha_1 - 1)\Delta y_{i,t-1} + (\alpha_2 + \alpha_1 - 1)\Delta y_{i,t-2} + ... \\ & + (\alpha_p + \alpha_{p-1} + ... + \alpha_1 - 1) y_{i,t-p} + \beta_0 \Delta x_{it} + (\beta_1 + \beta_0)\Delta x_{i,t-1} + ... \\ & + (\beta_p + \beta_{p-1} + ... + \beta_0) x_{i,t-p} + \varepsilon_{it}. \end{aligned} \qquad (6)$$

Or, simply redefining the coefficients:

$$\begin{aligned}\Delta y_{it} &= c_{it} + \pi_1 \Delta y_{i,t-1} + \pi_2 \Delta y_{i,t-2} + ... + \pi_p y_{i,t-p} \\ &\quad + \phi_0 \Delta x_{it} + \phi_1 \Delta x_{i,t-1} + ... + \phi_p x_{i,t-p} + \varepsilon_{it}.\end{aligned} \tag{7}$$

Note that, like equation (5), equation (7) is still a dynamic model for the (log) level of output per worker, provided $\pi_p \neq 0$.[8] In particular, the error term in (7) will be serially uncorrelated if the shocks entering the process in (5) are serially uncorrelated. Given $c_{it}$, the ratio $-\phi_p/\pi_p$ measures the long-run effect of the (log) investment share on the (log) level of income per worker.[9]

The crucial issue for this paper is how $c_{it}$ is allowed to depend on investment. We will experiment with three specifications of the process for $c_{it}$ that differ principally in the degree of unobserved heterogeneity they allow for in the model, and that have different implications for the time series properties of the variables. The first and simplest option is to assume that $c_{it}$ evolves according to:

$$c_{it} = c_{i,t-1} + \theta_0 + \theta_1 x_{it} + e_t. \tag{8}$$

$c_{it}$ is therefore a non-stationary stochastic process. Its change is assumed to depend directly on the current (log) share of investment, which we assume to be a stationary stochastic process. Here $e_t$ is a (productivity) innovation that is common to all countries. In the steady state (i.e. when $x_{it} = x_{it-1} = x_i$ for all periods and $\varepsilon_{it}$ and $e_t$ are set permanently to their expected value of zero) equations (5) and (8) imply:

$$g_i = y^*_{it} - y^*_{i,t-1} = \frac{\theta_0 + \theta_1 x_i}{(1 - \alpha_1 - \alpha_2 - ... - \alpha_p)}$$

where $g_i$ is the steady state growth rate and $y^*_{it}$ denotes output per worker on the steady state growth path. If $\theta_1 = 0$, the growth in $c_{it}$ and hence in steady state output per worker is

[8] Or, equivalently, provided $\sum_{i=1}^{p} \alpha_i \neq 1$ in (5).

[9] In the notation of (5), we have $-\phi_p/\pi_p = \sum_{i=0}^{p} \beta_i / (1 - \sum_{i=1}^{p} \alpha_i)$.

common to all countries. If $\theta_1 > 0$, the growth in $c_{it}$ varies across countries with their (log) investment shares.

Note that solving (8) backward we obtain:

$$c_{it} = c_{i0} + \theta_0 t + \theta_1 \sum_{s=1}^{t} x_{is} + \sum_{s=1}^{t} e_s = c_{i0} + \theta_1 \sum_{s=1}^{t} x_{is} + D_t \tag{9}$$

where $D_t = \theta_0 t + \sum_{s=1}^{t} e_s$ is a time effect that is common to all countries.

Substituting (9) in (7) gives:

$$\begin{aligned} \Delta y_{it} &= c_{i0} + D_t + \pi_1 \Delta y_{i,t-1} + \pi_2 \Delta y_{i,t-2} + ... + \pi_p y_{i,t-p} \\ &\quad + \phi_0 \Delta x_{it} + \phi_1 \Delta x_{i,t-1} + ... + \phi_p x_{i,t-p} + \theta_1 \sum_{s=1}^{t} x_{is} + \varepsilon_{it}. \end{aligned} \tag{10}$$

Equation (10) is a dynamic model for the log level of income per worker, with country effects ($c_{i0}$), time effects ($D_t$) and a new term in $bsx_{it} = \sum_{s=1}^{t} x_{is}$ , the backward sum of (log) investment shares. This backward sum variable captures the idea that, at each point in time, the level of output per worker reflects the history of the country's investment up to that point. Since we can replace the term $bsx_{it}$ by $\left(\frac{1}{t}\sum_{s=1}^{t} x_{is}\right) t$, simply by multiplying and dividing by $t$, another way to think about the basic implication of this model is that now there is a country-specific trend in the process for $y_{it}$, whose coefficient depends on the average of investment shares over past periods.[10] The lags of $\Delta y_{it}$ and $\Delta x_{it}$ are included to control for fluctuations at business cycle frequencies. $c_{i0}$ reflects time-invariant, country-specific influences on the steady state level of output per worker, while transient idiosyncratic shocks to the level of output per worker are reflected in $\varepsilon_{it}$. As before, $-\phi_p/\pi_p$ measures the effect of investment on the (log) level of output per worker along the steady state growth path. Here $-\theta_1/\pi_p$ measures the effect of investment on the steady state growth rate of output per worker.

[10] In an earlier version of this paper we also considered a specification in which $c_{it}$ depends on the mean of the log of the investment share in each country over the entire period ($\bar{x}_i$), i.e. $c_{it} = c_{i,t-1} + \gamma_0 + \gamma_1 \bar{x}_i + e_t$. In this case the coefficient on the trend depends on the overall mean of investment and not on the rolling mean calculated for periods up to time $t$. The qualitative results were similar to those obtained for the backward sum model summarized by equation (10).

With regard to the time series properties of the model, the backward sum of the investment shares, $bsx_{it} = \sum_{s=1}^{t} x_{is}$ is necessarily integrated of order one ($I(1)$), if $x_{it}$ is stationary. The process for $c_{it}$ in (8) then implies that $y_{it}$ is also $I(1)$, even after allowing for common $I(1)$ time effects $D_t$. The specific process for $c_{it}$ in (8) further implies that income and the backward sum of investment shares should be cointegrated after eliminating the common time effects $D_t$, although this pairwise cointegration property depends on the implicit assumption that the investment rate $x_{it}$ is the only country-level stationary influence on growth rates $\Delta c_{it}$.

The test of whether capital accumulation affects the growth rate of output per worker in the long run is here simply a test of the restriction $-\theta_1/\pi_p = 0$. Evidence that $\theta_1$ equals zero would be consistent with the Solow growth model, in which the steady state growth rate of output per worker is given by purely exogenous technological progress. Our approach extends the test of the Solow model proposed by Bernanke and Gurkaynak (2001) in a cross-sectional setting to a dynamic panel context. The advantage of the panel approach is that it allows us to address the endogeneity issues that naturally arise when investigating the effect of investment on growth, and that cannot be addressed satisfactorily in a single cross-section.

In some endogenous growth models capital accumulation should affect the steady state growth rate, so we would expect $-\theta_1/\pi_p$ to be significantly different from zero. For example, in AK models, broadening the definition of capital to include human capital, or allowing for learning by doing, counteracts the tendency to diminishing returns to capital accumulation present in the Solow model (see, for example, Romer (1986)). Similarly in Schumpeterian models with endogenous innovation, investment has a permanent effect on growth rates if capital is used as an input in the production of innovations (see, for example, Howitt and Aghion (1998)).

Even if there is no effect of investment on the long-run growth rate, equation (10) allows for a long-run effect of investment on the level of output per worker, captured by $-\phi_p/\pi_p$. Putting

it differently, given an initial level of income, the out of steady state growth rate depends on the investment rate, as predicted by the Solow growth model.

We can also estimate this model in first-differences, which becomes more important if we cannot rely on cointegration between the two $I(1)$ series $y_{it}$ and $bsx_{it}$. Taking first-differences of equation (5) and substituting for $\Delta c_{it}$ from equation (8), the model becomes:

$$\Delta y_{it} = \theta_0 + e_t + \alpha_1 \Delta y_{i,t-1} + ... + \alpha_p \Delta y_{i,t-p} + \beta_0 \Delta x_{it} + ... + \beta_p \Delta x_{i,t-p} + \theta_1 x_{it} + \Delta \varepsilon_{it}. \quad (11)$$

In this case the growth rate of output per worker is expressed as a distributed lag of itself and a distributed lag of first-differences of the investment share, with an additional term in the (log) level of the investment share. Notice that this is now a model for the growth rate rather than for the level of output per worker, and the error term here reflects first-differences of the shocks to the level of output per worker that enter equation (5). Moreover in this form all the variables in the empirical specification are stationary, provided that the (log) share of investment in output is stationary. An equation of this form has been estimated by various authors focused on testing the AK model (see, for instance, Jones (1995), Blomstrom, Lipsey and Zejan (1996), Li (2002) and Madsen (2002)). Again, a rejection of the hypothesis $\theta_1/(1 - \alpha_1 - ... - \alpha_p) = 0$ suggests a long-run effect of investment on growth, consistent with an endogenous growth approach.

Notice that, if there are serially uncorrelated shocks $\varepsilon_{it}$ that affect the (log) level of output per worker, as in (5), the error term in (11) has an MA(1) structure that makes it necessarily correlated with the lagged dependent variable $\Delta y_{i,t-1}$. More generally, the error term in the first-differenced specification would only be serially uncorrelated if the idiosyncratic shocks $\varepsilon_{it}$ follow a random walk. Otherwise least squares estimates of the parameters in (11) will be biased and inconsistent. Consistent estimates may be obtained, provided the shocks $\varepsilon_{it}$ are serially uncorrelated, by using lagged values of endogenous variables from periods t-2 and earlier as instrumental variables; and/or by using as instruments current values of exogenous

variables that are uncorrelated with both $\varepsilon_{it}$ and $\varepsilon_{i,t-1}$. We will explore the importance of these potential biases in such first-differenced specifications, and the availability of valid and informative instruments, in our empirical analysis.

In these specifications based on (8), there are no other country-specific influences on the growth of $c_{it}$ besides investment. As a result, for instance, there is no time-invariant country-specific component of the error term in the first-differenced equation (11), since the country-specific term $c_{i0}$ that affects the steady state level of output per worker has been eliminated by differencing. However, in this context, we can also allow for a time-invariant country-specific drift term, $d_i$, to enter the process for $c_{it}$, i.e.:

$$c_{it} = c_{i,t-1} + d_i + \theta_0 + \theta_1 x_{it} + e_t. \tag{12}$$

The corresponding dynamic model for the growth rate of output per worker, analagous to equation (11), is given by:

$$\Delta y_{it} = \theta_0 + d_i + e_t + \alpha_1 \Delta y_{i,t-1} + ... + \alpha_p \Delta y_{i,t-p} + \beta_0 \Delta x_{it} + ... + \beta_p \Delta x_{i,t-p} + \theta_1 x_{it} + \Delta \varepsilon_{it} \tag{13}$$

and does include country-specific intercepts ($d_i$). This generalization is potentially important as it allows for unobserved country-specific factors, such as the quality of institutions, that may affect both investment shares and steady state growth rates. The model in first-differences can still be estimated consistently by instrumental variables, provided a long time series is available, since country dummies can simply be included to allow for the heterogeneity in long-run growth rates reflected in $d_i$.

The corresponding extension to the dynamic model for the level of output per worker in equation (10) would introduce a set of unrestricted country-specific linear trends, $d_i t$, in addition to the backward sum of investment variable. The process for $c_{it}$ in (12) thus implies that the $I(1)$ output per worker and backward sum of investment variables ($y_{it}$ and $bsx_{it}$) are cointegrated after eliminating both common time effects, $D_t$, and country-specific deterministic

trends, $d_i t$.

Finally, again in the context of the model in first-differences, we can further allow for the possibility that the $c_{it}$ process contains an idiosyncratic random walk component. This gives our most general specification of the process for $c_{it}$ as:

$$c_{it} = c_{i,t-1} + d_i + \theta_0 + \theta_1 x_{it} + e_t + v_{it} \tag{14}$$

where the serially uncorrelated $v_{it}$ reflect 'permanent' shocks to the logarithm of output per worker that are here assumed to be independent across countries, conditional on $e_t$.[11] This most general specification for $c_{it}$ then has the implication that the idiosyncratic component of the error term in the reparameterized level equation (10) is $\varepsilon_{it} + \sum_{s=1}^{t} v_{is}$. In this case, the $I(1)$ variables $y_{it}$ and $bsx_{it}$ are not cointegrated, even conditional on common time effects and country-specific trends. As a result we cannot rely on cointegration methods to estimate parameters consistently in models in levels like (10). However this model can still be estimated consistently in first-differenced form. The dynamic model for the growth rate of output per worker becomes:

$$\Delta y_{it} = \theta_0 + d_i + e_t + \alpha_1 \Delta y_{i,t-1} + ... + \alpha_p \Delta y_{i,t-p} + \beta_0 \Delta x_{it} + ... + \beta_p \Delta x_{i,t-p} + \theta_1 x_{it} + v_{it} + \Delta \varepsilon_{it}. \tag{15}$$

After controlling for country dummies, instrumental variables estimates of (15) can allow for both transient shocks ($\varepsilon_{it}$) and permanent shocks ($v_{it}$) to the (log) level of output per worker in each country. While the two observed $I(1)$ series $y_{it}$ and $bsx_{it}$ may not be cointegrated, we can identify a coherent relationship between the two $I(0)$ series (i.e. log investment shares and growth rates of income per worker).

The various specifications outlined above will constitute the basis of our empirical analysis. Before presenting the econometric results, however, we first describe the data sources and the time series characteristics of the data used in estimation.

[11] Recall that the $e_t$ reflect 'permanent' shocks that are common to all countries.

# 4 The Data: Sources and Time Series Properties

The data for estimating the basic model comes from the Penn World Table 6.1 (PWT 6.1) data set. Among many other variables, this data set includes the series for GDP per worker, the share of the total gross investment in GDP (both in real terms), and population. The national accounting variables are measured in constant international dollars. Our data set contains 94 countries and the annual data covers the period 1960-2000.[12] Even though the PWT 6.1 data set has information on more countries, we include only 94 of them in order to have a minimum of 30 annual observations on these basic series for each country. Following Mankiw, Romer and Weil (1992), we also excluded countries for which oil production is the dominant industry. The reasoning is that the standard growth theories cannot be applied to the data from those countries since a large fraction of their GDP depends on natural resources. The excluded oil producers are Bahrain, Gabon, Iran, Iraq, Kuwait, Oman, Saudi Arabia, the United Arab Emirates and Venezuela. We also excluded Botswana, because of the dominant role of mining (diamonds). Finally, we excluded Nicaragua and Chad because recorded gross investment was negative in some years.

The number of countries in our sample drops to 75 when we use instrumental variables estimators, due to missing data for some of the countries on additional variables that are included in the instrument set. We will use as additional instruments appropriately lagged values of the inflation rate (measured using the GDP deflator), and of government spending and trade (import plus exports), both measured as a percentage of GDP. All these variables are taken from the World Bank World Development Indicators (WDI) 2005. For some robustness checks we also use measures of human capital accumulation obtained from Barro and Lee (2000).

We now briefly discuss the time series properties of our main variables, the log of output

[12] See the Data Appendix for a list of the countries and summary statistics.

per worker and the log of the investment share. To eliminate common (to all countries) $I(1)$ trending components, we investigate the time series properties of $\widetilde{y}_{it} = y_{it} - \overline{y}_t$ and $\widetilde{x}_{it} = x_{it} - \overline{x}_t$, where (e.g.) $\overline{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$. We first run the augmented Dickey-Fuller (ADF) test for both the level and the first-difference of these variables, separately for each of the 94 countries, using the annual data and allowing for 4 lags.[13]

Table 1A gives the number of countries for which the null hypothesis of a unit root is rejected for each series. The country-by-country ADF test cannot reject in almost all cases the presence of a unit root for both the log of GDP per worker ($\widetilde{y}$) and for the log of the investment share ($\widetilde{x}$). When we apply the ADF test to the first-differences of these variables, we are able to reject at the 5% level the non-stationarity of the growth rate of GDP per worker for 35 countries when we do not allow for a trend, and for 23 countries when we allow for a trend. The number of countries for which this test rejects the non-stationarity of the first-difference of the log investment share is larger: 49 without the trend and 31 with a trend. These results suggest that the levels of both these variables are $I(1)$, and may even be $I(2)$. However these tests are known to have low power for distinguishing highly persistent, yet stationary processes from unit root processes. Moreover, there is the danger of misinterpreting structural breaks in trend stationary processes with the presence of a unit root.

We also consider a more powerful test for unit roots in heterogeneous panels, the $W_{tbar}$ test proposed by Im, Pesaran and Shin (2003). This statistic is based on an appropriately standardized average of the individual ADF statistics, and has a standard normal limiting distribution. Results are reported in Table 1B.[14] For the log of the investment share we can reject the null hypothesis of non-stationarity if we rely on the version of the test that does not include a (country-specific) trend, but not if we consider the version with trends. We suspect that

[13] The ADF tests have also been run using different lag lengths. The results are very similar to those presented here.

[14] These tests are based on the balanced panel of 87 countries for which we have data for all 41 years.

the former version is more appropriate for the investment share, which cannot grow (or fall) without bounds and so should have neither a deterministic nor a stochastic trend in the long run. Neither version of the test rejects the null of non-stationarity for the log level of GDP per worker, while both reject this null for the growth rate. These results suggest that the log level of output per worker may be an $I(1)$ variable, even when measured in the form of deviations from common year-specific means.

Finally we consider cointegration between the log of GDP per worker ($\widetilde{y}$) and the backward sum of investment shares variable ($bs\widetilde{x}$), on the assumption that these are both $I(1)$ variables. Table 1C reports results for the group and panel ADF statistics of Pedroni (1995, 1999), both with and without country-specific deterministic trends, which test the null hypothesis of no cointegration.[15] These tests provide conflicting evidence, with the group ADF test suggesting that the two series are cointegrated when we allow for the presence of trends, but with the panel ADF test not rejecting the null of no cointegration.

Formally these time series properties are consistent with one of the processes for $c_{it}$ represented by equations (12) or (14), which suggest that in deviations from year-specific means, the log of GDP per worker and the backward sum of investment shares are $I(1)$ variables, that are pairwise cointegrated when we allow for country-specific trends in the case of (12) and not cointegrated in the case of (14). They are inconsistent with our most restrictive specification (see (8)) in which investment is the only source of variation across countries in long-run growth rates.[16] However there are sufficient reservations about the power and reliability of these tests for unit roots and cointegration that we prefer not to rely too heavily on these results. Instead we will present empirical results for the range of specifications introduced in the previous section, and focus on econometric results that are appropriate in each case given the time series

[15] These tests are also based on the balanced panel of 87 countries for which we have data for all 41 years.

[16] In that case we should find cointegration between the two series without conditioning on country-specific trends.

implications of the corresponding processes for $c_{it}$.

# 5 Econometric Results

We first present the results obtained when the dynamic models for the log level of GDP per worker or its first-difference (equations (10) and (11)) are estimated using pooled annual data. In all cases we allow for unobserved country-specific factors that affect the log level of GDP per worker. In some specifications we will also control for unobserved country-specific influences on growth rates, as well as for observed time-varying policy variables that may have a direct effect on growth rates. We will briefly discuss the robustness of these pooled annual results to changes in the functional form, to the measure of investment used, and to the inclusion of additional population growth variables. Subsequent sections present results using data for non-overlapping five-year periods, and mean group results that allow for heterogeneous parameters across countries and for cross-section dependence.

## 5.1 Basic Panel Results on Yearly Data

We estimate dynamic models for the log level of income per worker using the parameterization in equation (10) and testing for a long-run effect of investment on growth rates by including the backward sum of the log of investment shares, $bsx_{it}$ (backward sum model for short, from now on). For these models estimated on yearly data we have experimented with different lag lengths. Since the results are qualitatively quite similar, we will focus on the results obtained when $p$ is set to four, allowing for rather rich business cycle dynamics. All specifications include year dummies to control for common time effects.

The first two columns of Table 2 present the Within estimates of the model based on our most restrictive specification for $c_{it}$, and the Within estimates of an alternative specification

that excludes the contemporaneous $\Delta x_{it}$ and further uses $bsx_{i,t-1}$ instead of $bsx_{it}$. The process for $c_{it}$ in (8) that underlies the specification reported here implies that $y_{it}$ is non-stationary and cointegrated with $bsx_{it}$. In this case the Within estimates of the long-run growth effect, $-\theta_1/\pi_4$, are superconsistent, whether or not the specification includes contemporaneous investment variables. Moreover, Pesaran and Shin (1999) show that standard normal asymptotic inference will be valid for this ratio of coefficients on the cointegrated $I(1)$ variables, and the standard error can be calculated using the delta method.[17] These specifications allow for unobserved time-invariant country-specific factors to affect the steady state log level of GDP per worker, but do not allow for unobserved heterogeneity in long-run growth rates.

In both specifications there is strong evidence that investment has a significant positive effect on both the level and the steady state growth rate of output per worker, captured respectively by $-\phi_4/\pi_4$ and $-\theta_1/\pi_4$.[18] The hypothesis that the long-run effect on the growth rate is equal to zero can be rejected at the 1% significance level in both cases.[19]

The last two columns of Table 2 report Within estimates for more general specifications in which we include separate linear trend terms as explanatory variables for each country. The process for $c_{it}$ that underlies this specification is (12), indicating that $y_{it}$ and $bsx_{it}$ are cointegrated in the presence of country-specific trends. These models thus allow for unobserved time-invariant country-specific factors to affect long-run growth rates.

The inclusion of these country trends has a striking effect on the coefficient ($\pi_4$) on the lagged level of log GDP per worker ($y_{i,t-4}$), suggesting more rapid conditional convergence than found in the specifications that do not allow for unobserved heterogeneity in long-run growth

[17] For the pooled annual data, we rely on the length of the time series to justify the consistency of these Within estimates in the presence of lagged dependent variables.

[18] Note also that in all the specifications reported in Tables 2, the coefficient $\pi_4$ on the lagged level of log output per worker is negative and significantly different from zero, consistent with the presence of conditional convergence.

[19] We have reported in square brackets the marginal significance level of the tests that these ratios of estimated coefficients are equal to zero (obtained using the standard delta method).

rates.[20] The statistical significance of the coefficients on the backward sum of investment variables and the corresponding long-run growth effects ($-\theta_1/\pi_4$) is weaker in these specifications, which is not surprising given the collinearity between the trending backward sum variables and the included linear trends. Nevertheless the hypothesis that investment has no long-run effect on growth can still be rejected at around the 2.6% significance level in the specification reported in the final column, which includes only the lagged values of the investment variables.

The size of these estimated effects of investment on long-run growth rates is also quite large. For example, using the Within results without country trends, presented in the first column of Table 2, a country that has an investment share equal to the median of the sample distribution (14.04%) has an annual growth rate of output per worker that is 0.84 percentage points higher than a country with an investment share at the lower quartile of the sample distribution (8.43%). The difference in growth rates between a country at the upper quartile (22.38%) and one at the lower quartile is estimated to be 1.60 percentage points.

As we explained in section 3, we can also estimate the model containing the backward sum of investment shares in its first-differenced form (see (11)). This formulation is interesting because it has been used in some of the papers aimed at testing the AK model, such as Jones (1995). Moreover, now all the variables entering the estimated equation are stationary. In (11), the effect of investment on the steady state growth rate is captured by the coefficient on the level of log investment share ($x_{it}$). If there are no country-specific factors besides investment in the equation determining the evolution of $c_{it}$ (see (8)), then no country-specific 'fixed effects' should be present in the first-differenced model. If instead there is unobserved heterogeneity across countries in long-run growth rates, as in (12) or (14), then such country-specific effects are present and they can be controlled for in estimation.

As we discussed earlier, consistent estimation may also require an Instrumental Variables

[20] Note that in these specifications the form of this conditional convergence is towards a country-specific long-run growth path for log income per worker whose level and slope are allowed to vary across countries.

(IV) approach if there are serially uncorrelated shocks ($\varepsilon_{it}$) to the (log) level of output per worker, generating an MA(1) error component ($\Delta\varepsilon_{it}$) when the model is estimated in first-differences. Table 3 reports estimates of dynamic models for the growth rate of GDP per worker, using the parameterization in equation (11). Here our main interest is in the IV results and consequently, for reasons discussed below, we focus on more parsimonious dynamic specifications. For comparison, the first two columns report OLS and IV estimates of a more general specification, which includes current and lagged investment variables and two lags of the growth rate.

Consider first the IV estimates of this specification, reported in the second column of Table 3. This specification does not include country-specific intercepts and so does not allow for unobserved heterogeneity in growth rates. The instruments used are dated t-2 and earlier, to allow for the suspected MA(1) error structure. Specifically we use lagged observations dated from t-2 to t-6 on the log of income per worker ($y_{it}$) and the log investment share ($x_{it}$), and lagged observations dated t-2 and t-3 on a set of additional instruments (inflation, trade as a share of GDP, and government spending as a share of GDP). As expected, we find strong evidence of negative first-order serial correlation in the residuals from this specification. There is no evidence of higher order serial correlation, which is consistent with the use of endogenous variables dated t-2 in the instrument set. The validity of these instruments is not rejected by the Sargan-Hansen test of over-identifying restrictions.[21] The validity of our additional external instruments will be assessed further in Table 4 below.

The sum of the coefficients on the two lagged growth rates is around 0.7, giving an estimate of $(\alpha_1 + \alpha_2 - 1)$ of -0.3. This estimate also suggests much faster conditional convergence than the corresponding coefficients on $y_{i,t-4}$ in the first two columns of Table 2, but still indicates a considerable degree of persistence in growth rates. In contrast, OLS estimates of the same

[21] Similar results are found for all the IV specifications reported in Table 3.

specification reported in the first column of Table 3 suggest very little persistence in growth rates. As explained in section 3, we should expect OLS estimates of the coefficient on the lagged dependent variable ($\Delta y_{i,t-1}$) to be biased downwards if the error term in the first-differenced model has an MA(1) component ($\Delta \varepsilon_{it}$). Our IV results cast doubt on the claim in some earlier papers that growth rates display very little persistence.[22]

More importantly, these IV estimates of the model in first-differences also suggest a significant positive long-run effect of investment on growth rates. This effect is significant at the 1% level and has a magnitude similar to that suggested by the first two columns in Table 2. For instance, going from the lower quartile to the median of the sample distribution of investment rates is predicted to raise the growth rate by 1.14 percentage points. The OLS estimates of this model also suggest a significant and quantitatively similar long-run effect of investment on growth in our sample of 75 countries.[23]

One potential concern with our IV estimates of this specification is that identification may be weak. Although the difference between the IV and OLS estimates of the coefficients on the lagged dependent variables reported in Table 3 is reassuring, formally the null hypothesis of under-identification is not rejected by the Anderson (1984) test of canonical correlations.[24] To address this concern, the third column of Table 3 reports IV estimates for a more parsimonious specification that omits the insignificant current-dated first-difference of the investment share ($\Delta x_{it}$). This is sufficient for the Anderson test to now strongly reject the null hypothesis of under-identification in this specification. However this has little effect on the results discussed previously, and in particular the effect of investment on long-run growth rates remains signifi-

[22] See, for example, Jones (1995), Attanasio, Picci and Scorcu (2001) and Easterly and Levine (2001).

[23] Jones (1995) reported no long-run effect of investment on growth rates based on OLS estimates of a similar specification for a sample of OECD countries. We discuss our results for the OECD sub-sample in Table 7 below.

[24] See Hall et al. (1996). The idea is that for all parameters in the equation to be identified, all the canonical correlations between the matrix of explanatory variables and the matrix of instruments must be significantly different from zero. The null hypothesis of the test is that the smallest canonical correlation is zero.

cant at the 1% level. One difference is that we obtain a more precise estimate of the positive effect of investment on the long-run level of log GDP per worker in this more parsimonious specification.

The fourth column of Table 3 additionally includes country-specific intercepts in this dynamic model of growth rates, so allowing for (time-invariant) unobserved heterogeneity across countries in long-run growth. We continue to find a significant long-run effect of investment on growth rates, and now find a much more significant long-run effect of investment on income levels. There is conflicting evidence about the importance of allowing for unobserved heterogeneity in growth rates in this specification. A Wald test of the joint significance of the additional 74 intercept dummies indicates that they can be safely omitted from the model. However a Hausman test suggests that their omission has a statistically significant impact on the estimated coefficient on the lagged dependent variable ($\Delta y_{i,t-1}$). The convergence parameter ($\alpha_1 + \alpha_2 - 1$) falls to around -0.5 when we include a full set of country dummies in this first-differenced specification. Most importantly, the significance of the coefficient on the level of the log investment share ($x_{it}$), and the associated long-run effect of investment on growth rates, are found to be robust to allowing for other unobserved country-specific factors to influence long-run growth rates. The last two columns of Table 3 show that very similar results are obtained using an alternative specification that uses the lagged investment share in place of the current investment share.[25] In particular, the growth effect remains significant at conventional levels and of very similar magnitude to that obtained in the model with contemporaneous information.

[25] Similar results were also found using OLS estimates of specifications that omit all explanatory variables dated t and t-1, and relate current growth to lagged growth dated t-2 and earlier and to the level and first-difference of the investment share dated t-2 and earlier.

### 5.2 Controlling for Additional Observed Explanatory Variables

The IV estimates of the growth models with unobserved country-specific effects presented in Table 3 control for time-invariant unobserved heterogeneity, but there may be additional observed explanatory variables that have an effect on short-run or long-run growth, even controlling for the rate of capital accumulation. In particular, it is possible that the variables we have used as additional instruments (government spending, openness to trade, and inflation) may have a direct effect on growth, in addition to being informative instruments for investment. The test of over-identifying restrictions does not suggest that this is a problem, but there may be concerns about its power.

In Table 4 we report the results for three additional specifications in which the level and first-difference of these variables are added to the specification reported in the penultimate column of Table 3.[26] In each case, the estimated coefficients on both the level ($z_{i,t-1}$) and the first-difference ($\Delta z_{i,t-1}$) of these additional explanatory variables are individually and jointly insignificant, while the estimated coefficients on both the level ($x_{i,t-1}$) and the first-difference ($\Delta x_{i,t-1}$) of the log investment share remain highly significant. The null hypothesis of under-identification is strongly rejected for each of these models.[27] The long-run effect of investment on growth rates suggested by our results in Table 3 is thus found to be robust to the inclusion of these additional explanatory variables. The implication is that any effect of these additional variables on long-run growth rates, or indeed on the level of GDP per worker, is adequately summarized through their effect on investment shares.

[26] Similar results were found when we allow for country-specific effects, or when we allow for influences from current levels or first-differences of these additional variables.

[27] This was not the case for more general specifications in which two or more of these additional variables were both included.

### 5.3 Other Extensions: Functional Form and Variable Definition

In this sub-section we discuss a set of additional experiments to investigate whether our results are robust to various alternative specifications and extensions of the model. We first re-estimate the model reported in the penultimate column of Table 3, using the level of the investment share and not its logarithm as an explanatory variable. We then use the logarithm of fixed investment, as opposed to total investment, and finally we include a measure of population growth as an additional explanatory variable. These results are reported in Table 5.

The results are robust to changes in the functional form. When the level of the investment share is used as a regressor, we obtain qualitatively very similar results to those reported previously using its logarithm. In particular, the estimated coefficients on both the level and the first-difference of the investment share remain significant at the 5% level. The long-run growth effect remains significant at the 1% level, while the long-run level effect remains significant at around the 10% level. Previous papers that focus on testing the AK model have used the level of the investment share, as suggested by this theory, while those that focus on testing the (augmented) Solow model have generally used its log.

When we use the log of fixed investment as a share of GDP instead of the log of total investment as a share of GDP, we lose 419 observations from earlier years for which the fixed investment data is not available. The results for this specification are less satisfactory, in that we find marginally significant evidence of both second-order serial correlation in the residuals and instrument invalidity when the log of total investment is replaced by the log of fixed investment.[28] The estimated coefficient on the lagged dependent variable is much lower in this specification. Nevertheless we still find a highly significant long-run effect of fixed investment on growth rates, and we also find a highly significant long-run effect of fixed investment on

[28] We have confirmed that this is primarily due to the measure of investment used and not to the change in the sample size that results here.

income levels.

The final column of Table 5 considers a population growth measure as an additional explanatory variable. The measure used here is the log of the population growth rate plus a common capital depreciation rate of 3% and a common growth rate of 2%, as in Mankiw, Romer and Weil (1992). We allow for effects from both the level and the first-difference of this measure, but find neither term to be significant in our first-differenced specification. Our estimate of the long-run effect of investment on growth rates remains highly significant. The significance of the level effect increases marginally.

### 5.4 Estimation using Five-Year Averages

In this sub-section we present estimates of a basic specification estimated using data on average growth rates and investment shares for non-overlapping five-year periods. Taking such five-year averages is one way to smooth out fluctuations at the business cycle frequency, and has been used previously in the empirical growth literature by, for example, Islam (1995) and Caselli, Esquivel and Lefort (1996). This leaves us with data for eight five-year periods. One advantage of this approach is that we can consider robustness to the inclusion of schooling variables that for many countries are only available at five-year intervals.

This kind of specification has been used previously by Blomstrom, Lipsey, and Zejan (1996) to test for the effect of capital accumulation on long-run growth. They report OLS estimates for specifications that include only the lagged value of the investment share, both with and without country dummies. They find that the investment share does not Granger-cause growth. However neither the OLS nor the Within estimator are appropriate for models that contain lagged dependent variables and other endogenous variables as regressors, when the number of time periods used is very small. We apply the first-differenced Instrumental

Variables estimators suggested in Arellano and Bond (1991) to this data set.[29]

Having experimented with different lag structures, we focus on the results for the following simple dynamic specification (cf. (10)):

$$\Delta_5 y_{it} = c_{i0} + D_t + (\alpha_1 - 1) y_{i,t-5} + \beta_0 x_{it}^A + \theta_1 \sum_{s=1}^{t/5} x_{i,5s}^A + \varepsilon_{it} \tag{16}$$

for $t = 5, 10, ..., 40$, where $\Delta_5 y_{it} = y_{it} - y_{i,t-5}$, and $x_{it}^A$ denotes the average value of the log of the investment share over this five-year period. The backward sum term is constructed by cumulating these five-year averages. Taking first-differences to eliminate the country-specific intercepts then gives:

$$\Delta_5 y_{it} - \Delta_5 y_{i,t-5} = \Delta_5 D_t + (\alpha_1 - 1) \Delta_5 y_{i,t-5} + \beta_0 \Delta_5 x_{it}^A + \theta_1 x_{it}^A + \Delta_5 \varepsilon_{it}. \tag{17}$$

We assume here that, after first-differencing, there is no country-specific effect remaining in the error term. This approach does not control for unobserved heterogeneity in growth rates, which in this context would require estimation of the model in second-differences rather than in first-differences.[30]

The first two columns of Table 6 report the results for this first-differenced specification. The columns denoted by GMM1 report the one-step estimator in Arellano and Bond (1991), while those denoted by GMM2 report the two-step estimator, with standard errors that use the finite-sample correction proposed by Windmeijer (2005). The instrument set includes the logs of GDP per worker and the average investment share, each lagged 2, 3 and 4 (five-year) periods, as well as the lagged values of averaged inflation, government spending and trade variables from the same periods. As expected, there is evidence of first-order serial correlation in the first-differenced residuals, while the hypothesis of no second-order serial correlation

[29] Caselli, Esquivel and Lefort (1996) use these estimators with five-year panels to study the issue of conditional convergence.

[30] We did experiment with second-differenced specifications, using instruments that remain valid in the presence of an MA(2) error process. Not surprisingly, these estimates were imprecise and are not reported in detail. The effect of investment on the long-run growth rate was found to be positive and significant at the 10% level.

cannot be rejected. The Sargan-Hansen test of over-identifying restrictions for the GMM estimates does not reject the validity of our instruments and suggests that there is no gross form of mis-specification.

For both estimators, the t-test for the coefficient $\theta_1$ on the log level of the average investment share suggests a highly significant positive effect of investment on the long-run growth rate. The t-test for the coefficient $\beta_0$ on the log first-difference of the average investment share suggests a marginally significant positive effect of investment on the long-run level of income. The implied annual speed of convergence, $\lambda$, is between 0.0926 and 0.0935, which is very similar to that reported by Caselli, Esquivel and Lefort (1996), using similar GMM estimators for data on five-year periods.[31]

In the second two columns we report results for an alternative specification that excludes contemporaneous information, i.e. replacing the terms $\Delta_5 x_{it}^A$ and $x_{it}^A$ in (17) by $\Delta_5 x_{i,t-5}^A$ and $x_{i,t-5}^A$ respectively. In this case the growth effect is significant only at the 10% level and the level effect becomes insignificant. The decreased significance of the growth effect here is not surprising, since lagging the *five-year average* investment share variable means trying to identify the effect using much more distant observations on the investment variable than is implied by simply dropping the current *annual* investment share in our earlier specifications based on annual data. The strategy of using contemporaneous information and addressing the endogeneity of the current investment share using Instrumental Variables seems to be a more appealing strategy in this context.

One advantage of the five-year average specification is that it allows us to control for human capital measures, using the Barro and Lee (2000) data set, that are not available annually for many countries. The last two columns of Table 6 report results for the specification with contemporaneous information that additionally includes the current first-difference and level

[31] $\lambda$ is calculated from $-(1-\alpha_1) = -(1-e^{-5\lambda})$. See, for instance, Caselli et al. (1996), equation (10). Their estimate was 0.128.

of the average years of schooling in the population aged 15 and above; this allows for the level of schooling to affect the level of income per worker, consistent with the augmented Solow growth model. Lagged levels of this measure are also included in the instrument set. The inclusion of this schooling variable reduces the available sample to 67 countries.

The coefficients on this schooling measure are insignificant in this first-differenced specification.[32] The inclusion of this variable does not change the significance of the long-run effect of investment on growth rates, and strengthens the significance of the long-run effect of investment on income levels. Conventional tests would suggest omitting the schooling variable from the model.

### 5.5 Results for Sub-samples

Table 7 reports annual panel results separately for sub-samples of 20 OECD countries and 74 non-OECD countries, in the case of the backward-sum model, or 55 non-OECD countries in the case of the first-differenced model (where we require data on the additional instruments). For the non-OECD countries, we report results for the specification of the model in levels as reported in the last column of Table 2, and for the specification of the model in first-differences as reported in the last column of Table 3. Not surprisingly, since the non-OECD countries form the majority of our full sample, these results are similar to those obtained for the full sample. Both specifications indicate highly significant long-run effects of investment on both income levels and growth rates.

For the smaller sample of OECD countries the results were found to be sensitive to the timing of the explanatory variables, and we report results for the specification of the model in levels as reported in the third column of Table 2, and for the specification of the model in

[32] Similar results were found using the percentage of the working age population with secondary schooling as an alternative measure; and in specifications where the level of these schooling measures was omitted from the first-differenced equations.

first-differences as reported in the fourth column of Table 3; these specifications also allow for unobserved heterogeneity in growth rates, but include current-dated values of the explanatory variables.[33] For this specification, the long-run effect of investment on growth rates is found to be positive and significant at the 5% level using our IV estimates of the model in first-differences, and to be positive and significant at the 10% level using the OLS estimates of the model in levels.[34] The long-run effect of investment on income levels is positive in both cases, but only weakly significant for the model in first-differences.

Our results for the OECD sub-sample are weaker (in terms of statistical significance) and more sensitive to timing than our results for the (larger) non-OECD sub-sample, although the point estimates of the growth effect tend to be larger for the OECD countries. Moreover, we can easily accept the restriction that the coefficients (and long-run effects) are common to the two sub-samples.

The finding of a significant long-run growth effect in our IV estimates of the model in first-differences for the sub-sample of OECD countries also presents an interesting contrast to the results reported in Jones (1995). Jones estimated a similar specification for a sample of OECD countries but considered only the OLS estimates. As we noted in our discussion of the first column of Table 3, this is likely to result in biased estimates. When we estimate this specification in first-differences for the OECD sub-sample using OLS, the estimate of $\alpha_1+\alpha_2-1$ falls from -0.35 to -0.7, indicating much less persistence in growth rates. More importantly, for the OECD sub-sample, the OLS estimate of the long-run effect of investment on growth rates falls to 0.006 and is insignificantly different from zero (p-value = 0.37); while the OLS estimate of the long-run effect on income levels increases to 0.26 and becomes significant at

[33] Similar results to those reported in Table 7 were found using this specification for the non-OECD countries. The long-run effects of investment on income levels and growth rates were estimated less precisely for the OECD countries when the current-dated explanatory variables were omitted from the specification.

[34] Recall that valid inference here relies on cointegration between the two $I(1)$ income and backward sum variables around a country-specific trend.

the 1% level. These OLS results are broadly consistent with those reported in Jones (1995). The contrast with our preferred IV results suggests that, for the OECD sub-sample, allowing appropriately for the MA(1) error component in dynamic growth equations is important not only for estimating the degree of persistence in growth rates, but also for identifying the effect of investment on long-run growth rates.

### 5.6 Results for Individual Countries and Mean Group Estimates

In this sub-section we re-estimate the various models that we have developed, using annual time series data for individual countries. The main objective is to see whether our conclusion that investment has a positive effect on long-run growth rates is also supported by the evidence when we allow all the parameters to differ across countries. The trade-off here is clear: on the one hand, by imposing equality of coefficients across countries, as we have done in the pooled models, we gain in efficiency. On the other hand, if these restrictions are invalid, the inferences we draw from the panel estimates may be misleading. For instance, what we have called 'the' growth effect would be an inconsistent estimate of the average effect (across countries) of investment on long-run growth rates. In this case, relying on time series regressions for individual countries may provide a more accurate picture of the average effect of investment on long-run growth, and of its dispersion across countries.[35] The usefulness of this exercise is enhanced by the fact that formal tests of the validity of pooling tend to reject the restriction of identical slope coefficients across countries, even at the 1% level.

We initially measure each variable as deviations from overall year-specific means in order to control for common time effects. If the slope coefficients are identical across countries, this exactly controls for the presence of common factors that affect growth in all countries in the

[35] See Pesaran and Smith (1995) for a discussion of the biases that can result from not recognizing the heterogeneity of slope coefficients in dynamic panel data models. See also Lee, Pesaran, and Smith (1997) for an application to the issue of convergence.

same way, and is equivalent to the inclusion of year dummies in the pooled models we presented in Tables 2 and 3. If the slope coefficients differ across countries, this procedure is only an approximate way to deal with such common influences.

A summary of the results for both the backward sum model (with country-specific trends) and the parsimonious first-differenced model (with country-specific intercepts) is reported in Table 8. We report the mean and median value across countries of the main coefficients, and of the long-run growth and level effects. Due to the presence of a few outliers in the individual coefficient estimates that distort significantly the unweighted means, we report here robust estimates of the mean, together with its standard error.[36] This is a robust variant of what Pesaran and Smith (1995) call the Mean Group estimator.

We report the results obtained when contemporaneous information is used. In the backward sum model with country-specific trends, the effect of investment on the steady state growth rate of output is captured by $-\theta_{1i}/\pi_{4i}$. The robust estimate of its mean value is positive and highly significant, and similar in magnitude to the median. The effect of investment on the steady state (log) level of output is given by $-\phi_{4i}/\pi_{4i}$. The robust estimate of its mean is also positive and significant. The size of the growth effect here is larger than that obtained for the same specification when the data are pooled (see Table 2, third column).

In the first-differenced model with country-specific intercepts, the growth effect is represented by $-\theta_{1i}/(\alpha_{1i}+\alpha_{2i}-1)$ and the level effect by $-(\beta_{0i}+\beta_{1i})/(\alpha_{1i}+\alpha_{2i}-1)$. Again the average growth and level effects are highly significant. The magnitude of the growth effect is very similar to the mean group estimate for the backward sum model, and smaller than that for the differenced model estimated on the pooled data (see Table 3, fourth column). Its magnitude is again economically significant: if the investment rate increases from the lower

[36] The robust estimate of the mean is obtained using the *rreg* comand in Stata. The *rreg* command performs a robust regression, based initially on Huber weights and then on biweights. When no explanatory variables are specified, *rreg* produces a robust estimate of the mean. For details, see Hamilton (1991).

quartile to the median, the growth rate of output increases by 1.07 percentage points.

Table 9 reports these results separately for the sub-samples of OECD and non-OECD countries (i.e. the robust means and medians of the parameters from the time series models for the individual countries are calculated separately for these sub-samples). Here the growth and level effects are statistically significant for both sub-samples, and are found to be larger for the OECD countries.

### 5.6.1 Cross-section dependence

All the specifications we have considered in the preceding sections have maintained that the time-varying shocks to the income processes are either common to all countries (e.g. $D_t$ in (10)) or $e_t$ in (11)) or independent across countries (e.g. $\varepsilon_{it}$ in (10) or $\Delta\varepsilon_{it}$ in (11)). Pesaran (2006) has recently suggested a simple procedure to allow for a form of cross-section dependence with a multiplicative factor structure. Analogously to allowing each country to have its own slope coefficients on the observed explanatory variables, such as investment shares or lagged growth rates, this allows each country to have its own slope coefficient on the unobserved common factor. So, for example, the term $+D_t$ in (10) can be replaced by a term of the form $+\phi_i D_t$. This form of cross-country dependence can be allowed for by including as additional regressors, in the time series models for each country, the mean across countries for each year of the dependent variable and each of the explanatory variables. The time series $\overline{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$ and $\overline{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it}$ are thus used as additional regressors, where $y_{it}$ denotes the dependent variable and $x_{it}$ denotes the vector of explanatory variables. The mean group estimator is then obtained in the usual way, by taking the (robust) mean across countries of the coefficients in the country time series regressions on the variables of interest.[37]

[37] Pesaran (2007) suggests tests for unit roots in heterogeneous panels with cross-section dependence of this form. These gave some surprising results in our sample. For example, allowing for country-specific trends, we could reject the null hypothesis of a unit root for the log of GDP per worker; but we could not reject the null hypothesis of a unit root for the log of investment as a share of GDP (whether or not we allowed for country-

Table 10 reports the full sample results for our first-differenced model, first with a small set of controls for cross-section dependence (the cross-country means of $\Delta y_{it,}$ $\Delta x_{it}$ and $x_{it}$ only) and then with the full set of controls for cross-section dependence (the cross-country means of $\Delta y_{it}, \Delta y_{it-1}, \Delta y_{it-2,}$ $\Delta x_{it}, \Delta x_{it-1}$ and $x_{it}$). The robust mean of the long-run growth effects is statistically significant and quantitatively similar in both cases, and slightly larger than the corresponding estimate in Table 8 that does not allow for cross-section dependence. The robust mean of the long-run level effect is significantly different from zero only in the specification with the full set of controls, and slightly smaller than the corresponding estimate in Table 8.

Table 11 reports the (robust) mean group estimates from these specifications, separately for the sub-samples of OECD and non-OECD countries. Here we include the year means calculated for the full sample as additional controls in the individual country regressions for all countries. However similar results were obtained when the additional year means included as controls in the country regressions for each sub-sample were calculated using the data only for that sub-sample of countries, or when year means for the OECD and non-OECD sub-samples were both included.

The results for the non-OECD countries again indicate that the robust mean of the long-run growth effects is positive and significant, and rather larger in magnitude than the corresponding estimate in Table 9. However the results for the smaller sample of OECD countries do not identify a significant long-run growth effect when we consider this group of countries in isolation and use this approach to allow for parameter heterogeneity and cross-section dependence. This may in part reflect the limited variation in country-level investment rates that remains within the OECD sub-sample once we condition on the world (or OECD) average investment rate, although this is not the whole story since we continue to identify a significant long-run level effect for the OECD sub-sample.

specific trends). These results were also more sensitive to the choice of lag length than the Im, Pesaran and Shin (2003) tests reported in Table 1B.

# 6 Conclusions

In contrast to the suggestion that capital accumulation plays a minor role in economic growth, we find that the share of investment in GDP has a large and statistically significant effect, not only on the level of output per worker, but more importantly on its long-run growth rate. We find these results using pooled annual data for a large sample of countries, using pooled data for five-year periods, and using the average effects estimated from time series models for individual countries. They are robust to controlling for unobserved country-specific effects, not only on the steady state level of output per worker, but also on the long-run growth rates. The cross-section correlation between investment shares and average growth rates reported by Bernanke and Gurkaynak (2001) is thus found to be robust to controlling for unobserved heterogeneity in growth rates. A permanent increase in the share of GDP devoted to investment predicts not just a higher level of output per worker, but also a faster growth rate in the long run. These findings are particularly robust for the sub-sample of non-OECD countries, and somewhat less so for the smaller sample of OECD countries. In particular, when we allow for heterogeneity across countries in all parameters of our models, and for cross-section dependence using the multi-factor error structure suggested by Pesaran (2006), we still find a significant mean effect of investment on long-run growth rates for the non-OECD countries, though not for the OECD countries. This echoes differences in the growth process between countries at different stages of development reported previously by Evans (1998), comparing results for countries with different levels of education, and by De Long and Summers (1993) and Aghion, Comin and Howitt (2006), comparing results for countries with different income levels.

These findings are inconsistent with the predictions of the Solow and Swan growth model. They are instead consistent with the predictions of some endogenous growth models; for example, AK-type models (Romer, 1986), and Schumpeterian-type models in which capital is used as an input in the production of innovations (Howitt and Aghion, 1998).

It should be stressed that our evidence does not rule out a very important role for many other factors in the growth process, such as the quality of economic, political and legal institutions, and the quality of macroeconomic and microeconomic government policies, including those related to education and research. Such factors may play a key role in determining either capital accumulation, or the impact of investment on growth, in addition to affecting growth directly, at a given level of investment, in ways that may be captured by the country effects and time effects in our specifications. We recognize that measured variation in the share of investment in GDP may, to some extent, act as a proxy for unmeasured country-level time series variation in some of these factors. At least we have shown that measured investment is an informative proxy. In the absence of convincing annual data for a large set of countries on the quality of institutions and policies, it will remain challenging to make much stronger claims about the identification of the true causal determinants of economic growth.

In future work we plan to explore the heterogeneity in the effects of investment suggested by our analysis of models for individual countries. Are there systematic differences by region? Are there observable characteristics that explain why investment seems to have a greater impact on growth in some countries than in others? How are these related to factors that explain differences across countries in investment levels?[38] This heterogeneity is consistent with the view that differences in policies and institutions and the incentives they generate can have a tremendous impact on economic growth. Additional evidence on these issues will enhance our understanding of the determinants of growth and of the channels through which they operate.

[38] See Hall and Jones (1999) for cross-sectional evidence on the relationship between capital accumulation, on the one hand, and institutions and government policies, on the other.

## References


[1] Aghion, P., D. Comin, and P. Howitt , "When Does Domestic Saving Matter for Economic Growth?", Harvard University, mimeo, 2006.

[2] Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, New York, Wiley, 1984.

[3] Arellano, M and S.R. Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 1991, 58, 277-297.

[4] Attanasio, O., L. Picci, and A.E. Scorcu, "Saving, Growth, and Investment: A Macroeconomic Analysis Using A Panel of Countries", *The Review of Economics and Statistics*, 2000, 82(2), 182-211.

[5] Barro, R.J. and X. Sala-I-Martin, *Economic Growth*, 1995, McGraw-Hill, New York.

[6] Barro, R.J. and J.W. Lee, "International Data on Educational Attainment: Updates and Implications", Center for International Development WP # 42, April 2000.

[7] Beaudry, P., F. Collard and D. Green, "Decomposing the Twin Peaks: A Study of the Changing World Distribution of Output per Worker", University of British Columbia, mimeo, 2002.

[8] Bernanke, B.S. and R.F. Gurkaynak, "Is Growth Exogenous? Taking Mankiw, Romer and Weil Seriously", NBER Macroeconomics Annual, 2001, The MIT Press, Cambridge, Massachusetts.

[9] Blomstrom, M., R.E. Lipsey and M. Zejan, "Is Fixed Investment Key to Economic Growth?", *Quarterly Journal of Economics*, 1996, 111(1), 269-276.

[10] Bond, S.R., A. Hoeffler and J. Temple, "GMM Estimation of Empirical Growth Models", 2001, CEPR Discussion Paper 3048.

[11] Carroll, C.D. and D.N. Weil, "Saving and Growth: A Reinterpretation," *Carnegie-Rochester Series on Public Policy,* 1994, 40, 133-192.

[12] Caselli, F., G. Esquivel and F. Lefort, "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, 1996, 1, 363-389.

[13] De Long, J.B. and L.H. Summers, "Equipment Investment and Economic Growth", *Quarterly Journal of Economics*, 1991, 106(2), 445-502.

[14] De Long, J.B. and L.H. Summers, "How Strongly Do Developing Economies Benefit From Equipment Investment?", *Journal of Monetary Economics*, 1993, 32, 395-415.

[15] Easterly, W., *The Elusive Quest for Growth*, 2001, The MIT Press, Cambridge, Massachusetts.

[16] Easterly, W. and R. Levine, "What Have We Learned From a Decade of Empirical Research on Growth? It's Not Factor Accumulation: Stylized Facts and Growth Models", *World Bank Econ Review,* 2001, 15,177-219.

[17] Evans, P., "Using Panel Data to Evaluate Growth Theories", *International Economic Review*, 1998, 39(2), 295-306.

[18] Hall, A.R., G.D. Rudebush, D.W. Wilcox, "Judging Instrument relevance in Instrumental variable Estimation", *International Economic Review*, 1996, 37(2).

[19] Hall, R.E. and C.I. Jones, "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, 1999, 114(1), 83-116.

[20] Hamilton, L.C., "How Robust is Robust Regression", *Stata Technical Bulletin,* 1991, 2. 21-26.

[21] Howitt, P. and P. Aghion (1998), "Capital Accumulation and Innovation as Complementary Factors in Long-run Growth", *Journal of Economc Growth*, 3, 111-130.

[22] Im, K.S., M.H. Pesaran, and Y. Shin, "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 2003, 115, 53-74.

[23] Islam, N., "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, 1995, 110, 1127-1170.

[24] Jones, C.I., "Time Series Tests of Endogenous Growth Models", *Quarterly Journal of Economics,* 1995, 110(2), 495-525.

[25] Lee, K., M.H. Pesaran and R. Smith, "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model", *Journal of Applied Econometrics,* 1997, 12, 357-392.

[26] Levine, R. and D. Renelt, "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review*, 1992, 82(4), 942-963.

[27] Li, D., "Is the AK Model Still Alive? The Long-Run Relation Between Growth And Investment Re-examined", *Canadian Journal of Economics*, 2002, 35(1), 93-113.

[28] Madsen, J.B., "The Causality Between Investment and Economic Growth", *Economic Letters,* 2002, 74, 157-163.

[29] Mankiw, N.G., D. Romer and D.N. Weil, "A Contribution to the Empirics of Growth", *Quarterly Journal of Economics,* 1992, 107(2), 407-437.

[30] Pedroni, P., "Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the Purchasing Power Parity Hypothesis", *Econometric Theory*, 2004, 20, 597-625

[31] Pedroni, P., "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", *Oxford Bulletin of Economics and Statistics*, 1999, 61:4, 653-670.

[32] Pesaran, M.H., "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", *Econometrica,* 2006, 74 (4), 967-1012.

[33] Pesaran, M.H., "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence", *Journal of Applied Econometrics,* 2007, 22, 265-312.

[34] Pesaran, M.H. and R. Smith, "Estimating Long-Run Relationships From Dynamic Heterogeneous Panels", *Journal of Econometrics*, 1995, 68, 79-113.

[35] Pesaran, M.H. and Y. Shin, "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis", in S. Strom (ed), *Econometrics and Economic Thoery in the 20th Century: The Ragnar Frisch Centennial Symposium*, 1999, Cambridge University Press, Cambridge, UK.

[36] Romer, P., "Increasing Returns and Long-run Growth", *Journal of Political Economy*, 1986, 94, 1002-1037 .

[37] Windmeijer, F., "A Finite Sample Correction for the Variance of Linear Two-Step GMM Estimators", *Journal of Econometrics,* 2005, 126(1), 25-51.

Table 1A: Augmented Dickey Fuller Tests

| | | Without trend | With trend |
|---|---|---|---|
| $\tilde{y}$ | 5% | 4 | 3 |
| | 10% | 4 | 2 |
| $\Delta\tilde{y}$ | 5% | 35 | 23 |
| | 10% | 19 | 15 |
| $\tilde{x}$ | 5% | 6 | 9 |
| | 10% | 9 | 1 |
| $\Delta\tilde{x}$ | 5% | 49 | 31 |
| | 10% | 17 | 12 |

- The figures indicate the number of countries for which the null hypothesis of a unit root is rejected at the corresponding marginal significance level. Common time means are eliminated from the series.
- The total number of countries is 94, and the sample period covers 1960-2000.
- The number of lags used in the test is 4.

Table 1B: Im-Pesaran-Shin Tests for Unit Roots in Heterogeneous Panels

| | Without trend | With trend |
|---|---|---|
| $\tilde{y}$ | 5.152 | 2.360 |
| $\Delta\tilde{y}$ | -14.470 | -12.154 |
| $\tilde{x}$ | -3.264 | -0.527 |
| $\Delta\tilde{x}$ | -18.010 | -14.178 |

- Reported figures are the W statistic in Im, Pesaran and Shin (2003). Under the null hypothesis of a unit root (where the alternative is one sided), the test statistic is asymptotically distributed as a standard normal.
- To have a balanced panel, information on 87 countries is used for the tests.
- The number of lags used in the test is 4.

Table 1C: Pedroni Tests of Cointegration Between $\tilde{y}$ and $bs\,\tilde{x}$

| | Without trend | With trend |
|---|---|---|
| Panel ADF | 2.606 | -1.328 |
| Group ADF | 0.916 | -4.528 |

- See Pedroni (1995, 1999). Under the null hypothesis of no cointegration (where the alternative is one sided), the test statistic is asymptotically distributed as a standard normal.
- The test is applied after removing the common time means in both series.
- To have a balanced panel, information on 87 countries is used in this test.
- The number of lags used in the test is 4.

Table 2: Annual Panel Regressions: Models in Levels Using the Backward-Sum of Investment Shares

| | Within | Within (no cont. info.) | Within With country trends | Within With country trends (no cont. info) |
|---|---|---|---|---|
| $\Delta y_{it-1}$ | 0.0258 (0.83) | 0.0180 (0.58) | -0.0938 (-3.14) | -0.1019 (-3.47) |
| $\Delta y_{it-2}$ | -0.0543 (-2.07) | -0.0576 (-2.16) | -0.1694 (-6.40) | -0.1736 (-6.40) |
| $\Delta y_{it-3}$ | -0.0524 (-1.99) | -0.0558 (-2.07) | -0.1700 (-6.44) | -0.1740 (-6.53) |
| $y_{it-4}$ | -0.0563 (-8.38) | -0.0578 (-8.37) | -0.1968 (-13.47) | -0.1990 (-13.38) |
| $\Delta x_{it}$ | -0.0273 (-2.46) | | -0.0268 (-2.59) | |
| $\Delta x_{it-1}$ | 0.0301 (4.24) | 0.0361 (5.14) | 0.0285 (4.06) | 0.0370 (5.36) |
| $\Delta x_{it-2}$ | 0.0161 (2.44) | 0.0216 (3.25) | 0.0172 (2.48) | 0.0249 (3.65) |
| $\Delta x_{it-3}$ | 0.0297 (4.13) | 0.0359 (4.83) | 0.0302 (3.96) | 0.0386 (5.15) |
| $x_{it-4}$ | 0.0243 (5.53) | 0.0300 (6.70) | 0.0306 (4.74) | 0.0384 (6.24) |
| $bsx_{it}$ | 0.0009 (4.60) | | 0.0012 (1.25) | |
| $bsx_{it-1}$ | | 0.0010 (4.84) | | 0.0020 (2.15) |
| Growth effect: $-\frac{\theta_1}{\pi_4}$ | 0.0164 [0.0000] | 0.0181 [0.0000] | 0.0060 [0.2028] | 0.0103 [0.0258] |
| Level effect: $-\frac{\phi_4}{\pi_4}$ | 0.4311 [0.0000] | 0.5182 [0.0000] | 0.1554 [0.0000] | 0.1930 [0.0000] |
| $R^2$ | 0.2165 | 0.2109 | 0.3060] | 0.3013 |
| Test of first order serial correlation | -0.41 [0.6785] | -0.53 [0.5988] | 0.19 [0.9232] | 0.05 [0.9600] |
| Test of joint significance of country trends | | | 0.0000 | 0.0000 |
| Number of countries | 94 | 94 | 94 | 94 |
| Number of observations | 3466 | 3466 | 3466 | 3466 |

Notes:

- t-ratios are in parentheses, and the p-value of the significance test of the long run effects and the p-value of the -serial correlation tests are in square brackets.
- Full set of year dummies are included in estimation to control for common time effects.

Table 3: Annual Panel Regressions: Models in First-Differences

| | OLS | IV | Parsimonious IV | Parsimonious IV with country effects | Parsimonious IV No cont. info | Parsimonious IV No cont. info with country effects |
|---|---|---|---|---|---|---|
| $\Delta y_{it-1}$ | 0.1121<br>( 3.08) | 0.7247<br>(4.84) | 0.7298<br>(4.94) | 0.6125<br>(4.04) | 0.7270<br>(4.92) | 0.6388<br>(4.48) |
| $\Delta y_{it-2}$ | 0.0232<br>( 0.80) | -0.0544<br>(-1.23) | -0.0525<br>(-1.19) | -0.0727<br>(-2.02) | -0.0533<br>(-1.21) | -0.0742<br>(-2.03) |
| $x_{it}$ | 0.0134<br>( 7.91) | 0.0074<br>(2.66) | 0.0075<br>(2.67) | 0.0224<br>(1.78) | | |
| $x_{it-1}$ | | | | | 0.0075<br>(2.68) | 0.0189<br>(1.72) |
| $\Delta x_{it}$ | -0.0332<br>( -3.16) | -0.0163<br>(-0.22) | | | | |
| $\Delta x_{it-1}$ | 0.0260<br>( 3.66) | 0.0851<br>(1.61) | 0.0764<br>(2.35) | 0.0947<br>(2.74) | 0.0804<br>(2.44) | 0.0992<br>(2.63) |
| Growth effect | 0.0155<br>[0.0000] | 0.0224<br>[0.0070] | 0.0232<br>[0.0041] | 0.0488<br>[0.0072] | 0.0228<br>[0.0036] | 0.0433<br>[0.0127] |
| Level effect | -0.0083<br>[0.5841] | 0.2086<br>[0.2735] | 0.2368<br>[0.1264] | 0.2058<br>[0.0015] | 0.2464<br>[0.1123] | 0.2277<br>[0.0019] |
| Test of first order serial correlation | | -3.44<br>[0.0006] | -3.79<br>[0.0002] | -3.27<br>[0.0011] | -3.73<br>[0.0002] | -3.77<br>[0.0002] |
| Test of second order serial correlation | | -0.79<br>[0.4272] | -0.61<br>[0.5386] | -0.22<br>[0.8255] | -0.69<br>[0.4894] | -0.44<br>[0.6609] |
| Anderson Test of Under-identification (Canonical Corr) | | 0.1096 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Test of o.i.d. restrictions | | 0.3823 | 0.4351 | 0.4155 | 0.4322 | 0.3973 |
| Test of the joint significance of the country effects | | | | 1.0000 | | 1.0000 |
| Number of countries | 75 | 75 | 75 | 75 | 75 | 75 |
| Number of observations | 2566 | 2566 | 2566 | 2566 | 2566 | 2566 |

Notes:

- t-ratios are in parentheses. The p-value of the significance test of the long run effects and the p-value of the serial correlation tests in Arellano and Bond (1991) are in square brackets.
- Full set of year dummies are included in estimation to control for common time effects.
- In all the IV specifications $x_{it}$ ($x_{it-1}$), $\Delta x_{it}$, $\Delta x_{it-1}$ and $\Delta y_{it-1}$ are instrumented. The instrument set consists of lags 2-6 of $x_{it}$ and $y_{it}$, lags 2 and 3 of log of annual inflation plus one, log of trade (sum of exports and imports) as a percentage of GDP, and log of government spending as a percentage of GDP.
- The covariance matrix in all estimations allows for heteroskedasticity and MA(1) errors.
- p-values are reported for the Anderson test of under-identification, the Sargan-Hansen test of over-identifying restrictions, and the joint significance test for country-specific intercepts.

Table 4: First-Differenced Models with Additional Variables

| | Government Spending & Investment | Trade & Investment | Inflation& Investment |
|---|---|---|---|
| $\Delta y_{it-1}$ | 0.6526<br>(3.15) | 0.7583<br>(4.83) | 0.7149<br>(4.39) |
| $\Delta y_{it-2}$ | -0.0422<br>(-0.91) | -0.0618<br>(-1.30) | -0.0513<br>(-1.15) |
| $z_{it-1}$ | -0.0076<br>(-1.38) | 0.0009<br>(0.34) | -0.0014<br>(-0.14) |
| $\Delta z_{it-1}$ | -0.0657<br>(-0.62) | 0.0450<br>(0.49) | -0.0092<br>(-0.30) |
| $x_{it-1}$ | 0.0101<br>(2.73) | 0.0072<br>(2.51) | 0.0075<br>(2.65) |
| $\Delta x_{it-1}$ | 0.0716<br>(2.07) | 0.0827<br>(2.44) | 0.0773<br>(2.28) |
| Growth effect | 0.0258<br>[0.0173] | 0.0238<br>[0.0065] | 0.0222<br>[0.0044] |
| Level effect | 0.1837<br>[0.2137] | 0.2724<br>[0.1337] | 0.2298<br>[0.1472] |
| Test of joint significance of $z_{it-1}$ & $\Delta z_{it-1}$ | 0.2867 | 0.8854 | 0.9252 |
| Test of joint significance of $x_{it-1}$ & $\Delta x_{it-1}$ | 0.0072 | 0.0113 | 0.0096 |
| Test of first order serial correlation | -2.55<br>[0.0108] | -3.69<br>[0.0002] | -3.48<br>[0.0005] |
| Test of second order serial correlation | -0.45<br>[0.6515] | -0.67<br>[0.5013] | -0.79<br>[0.4286] |
| Anderson Test of Under-Identification | 0.0090 | 0.0000 | 0.0000 |
| Test of o.i.d. restrictions | 0.3872 | 0.3016 | 0.2759 |
| Number of countries | 75 | 75 | 75 |
| Number of observations | 2562 | 2564 | 2564 |

Notes:

- $z_{it}$ denotes the additional variables (the log of government spending over GDP, the log of imports plus exports over GDP, and the log of one plus the inflation rate).
- Full set of year dummies are included in estimation to control for common time effects.
- In all the specifications $x_{it-1}$, $\Delta x_{it-1}$ , $z_{it-1}$, $\Delta z_{it-1}$ and $\Delta y_{it-1}$ are instrumented. See Table 3 for a list of instruments.
- t-ratios are in parentheses, and the p-value of the significance test of the long-run effects and the p-value of the serial correlation tests are in square brackets.
- The covariance matrix in all estimations allows for heteroskedasticity and MA(1) errors.
- p-values are reported for the Anderson test of under-identification, the Sargan-Hansen test of over-identifying restrictions, and the joint significance tests.

Table 5: Robustness Checks

| | Using level of investment instead of log of investment | Using the log of fixed investment instead of log of total investment | Including population growth |
|---|---|---|---|
| $\Delta y_{it-1}$ | 0.7081<br>(5.11) | 0.3453<br>(1.86) | 0.6479<br>(4.26) |
| $\Delta y_{it-2}$ | -0.0502<br>(-1.16) | -0.0359<br>(-0.82) | -0.0284<br>(-0.65) |
| $x_{it-1}$ | 0.0005<br>(2.56) | 0.0325<br>(3.13) | 0.0080<br>(2.06) |
| $\Delta x_{it-1}$ | 0.0047<br>(2.17) | 0.1278<br>(2.41) | 0.0714<br>(2.15) |
| $z_{it}$ | | | 0.0009<br>(0.80) |
| $\Delta z_{it}$ | | | 0.138<br>(-0.08) |
| Growth effect | 0.0014<br>[0.0034] | 0.0471<br>[0.0000] | 0.0209<br>[0.00073] |
| Level effect | 0.0136<br>[0.1142] | 0.1851<br>[0.0039] | 0.1878<br>[0.0988] |
| Test of first order serial correlation | -3.88<br>[0.0001] | -1.27<br>[0.2029] | -3.49<br>[0.0005] |
| Test of second order serial correlation | -0.51<br>[0.6120] | -1.94<br>[0.0526] | -0.30<br>[0.9753] |
| Anderson Test of Under-Identification | 0.0000 | 0.0004 | 0.0077 |
| Test of o.i.d. restrictions | 0.4258 | 0.0834 | 0.1724 |
| Number of countries | 75 | 75 | 75 |
| Number of observations | 2566 | 2147 | 2559 |

Notes:

- In all three columns the first-differenced model specification (without country effects) is estimated.
- Full set of year dummies are included in estimation to control for common time effects.
- $z_{it}$ denotes population growth.
- In all the specifications $x_{it-1}$, $\Delta x_{it-1}$ and $\Delta y_{it-1}$ are instrumented. In the last specification $z_{it}$, $\Delta z_{it}$ are also instrumented. See Table 3 for a list of instruments.
- t-ratios are in parentheses, and the p-value of the significance test of the long run effects and the p-value of the serial correlation tests are in square brackets.
- The covariance matrix in all estimations allows for heteroskedasticity and MA(1) errors.
- p-values are reported for the Anderson test of under-identification, and the Sargan-Hansen test of over-identifying restrictions.

Table 6: Models Using Five-Year Averages

| | GMM1 | GMM2 | GMM1 no cont. info. | GMM2 no cont. info | GMM1 | GMM2 |
|---|---|---|---|---|---|---|
| $\Delta_5 y_{i,t-5}$ | -0.3735 (-4.17) | -0.3706 (-4.00) | -0.3971 (-3.53) | -0.3976 (-3.60) | -0.3956 (-4.57) | -0.4881 (-3.82) |
| $x_{it}^A$ | 0.0310 (3.03) | 0.0315 (3.38) | | | 0.0340 (3.56) | 0.0370 (3.36) |
| $\Delta_5 x_{it}^A$ | 0.0879 (1.87) | 0.0840 (1.79) | | | 0.1763 (4.22) | 0.1568 (3.36) |
| $x_{i,t-5}^A$ | | | 0.0200 (1.66) | 0.0205 (1.75) | | |
| $\Delta_5 x_{i,t-5}^A$ | | | -0.0112 (-0.23) | -.0194 (-0.41) | | |
| $z_{it}$ | | | | | 0.0024 (0.14) | -0.0120 (-0.65) |
| $\Delta_5 z_{it}$ | | | | | -0.0283 (-0.17) | -0.1509 (-0.75) |
| Growth Effect | 0.0166 | 0.0170 | 0.0100 | 0.0104 | 0.0171 | 0.0152 |
| Level Effect | 0.2354 | 0.2266 | -0.0283 | -0.0487 | 0.4456 | 0.3213 |
| Test of joint significance of education | | | | | 0.6615 | 0.7483 |
| Test of first order serial correlation | -4.23 [0.000] | -3.62 [0.000] | -3.61 [0.000] | -3.09 [0.002] | -3.79 [0.000] | -2.14 [0.033] |
| Test of second order serial correlation | -0.25 [0.799] | -0.27 [0.784] | -0.12 [0.906] | -0.10 [0.921] | 0.07 [0.943] | 0.22 [0.823] |
| Test of o.i.d restrictions | 0.707 | 0.510 | 0.455 | 0.267 | 1.000 | 0.223 |
| Number of countries | 75 | 75 | 75 | 75 | 67 | 67 |
| Number of observations | 446 | 446 | 446 | 446 | 353 | 353 |

Notes:

- Investment ( $x_{it}^A$ ) is the average of the log investment share over each five-year period. The growth rate of GDP per worker ( $\Delta_5 y_{it}$ ) is defined over the same five-year period.
- In columns 5 and 6, $z_{it}$ denotes average years of schooling in the population aged 15 and above.
- Full set of period dummies are included in estimation to control for common time effects.
- The instrument set consists of lags 2, 3 and 4 of investment and log GDP per worker and lags 2, 3 and 4 of log of inflation, trade and government spending. In the last two columns lags 2, 3 and 4 of average years of schooling are added to the instrument set.
- GMM1 corresponds to the one-step estimator with robust standard errors in Stata (xtabond2 command), and GMM2 corresponds to the two-step estimator with corrected standard errors (Windmeijer (2005) correction).
- The p-value of the Sargan-Hansen test of over-identifying restrictions is reported.

Table 7: Annual Panel Regressions for OECD and Non-OECD Countries

| Backward-Sum Model | | | First-Differenced Model | | |
|---|---|---|---|---|---|
| | Country Trends, OECD | No cont. info., Country Trends, Non-OECD | | Country Constants, OECD | No cont. info., Country Constants, Non-OECD |
| $\theta_1$ | 0.0026<br>(1.54) | 0.0024<br>(2.45) | $\theta_1$ | 0.0293<br>(1.70) | 0.0197<br>(1.68) |
| $\phi_4$ | 0.0220<br>(1.71) | 0.0374<br>(5.87) | $\beta_1$ | 0.0582<br>(1.02) | 0.1033<br>(2.62) |
| $\pi_4$ | -0.0898<br>(-5.37) | -0.2063<br>(-12.98) | | -0.3597<br>[0.0067] | -0.4056<br>[0.0094] |
| Growth Effect | 0.0294<br>[0.0863] | 0.0115<br>[0.0109] | $\alpha_1 + \alpha_2 - 1$<br>Growth Effect | 0.0816<br>[0.0455] | 0.0485<br>[0.0133] |
| Level Effect | 0.2450<br>[0.0927] | 0.1815<br>[0.0000] | Level Effect | 0.1617<br>[0.2075] | 0.2547<br>[0.0047] |
| Test of first order serial correlation | 1.24<br>[0.2138] | 0.10<br>[0.9176] | Test of first order serial correlation | -4.14<br>[0.0000] | -3.65<br>[0.0003] |
| Test of second order serial correlation | | | Test of second order serial correlation | -1.28<br>[0.2012] | -0.34<br>[0.7344] |
| Anderson Test of Under-Identification | | | Anderson Test of Under-Identification | 0.0000 | 0.0015 |
| Test of o.i.d. restrictions | | | Test of o.i.d. restrictions | 0.7457 | 0.5106 |
| Test of joint significance of country trends | 0.0000 | 0.0000 | Test of joint significance of country effects | 0.9894 | 1.0000 |
| Number of Countries | 20 | 74 | Number of Countries | 20 | 55 |
| Number of Observations | 740 | 2726 | Number of Observations | 697 | 1869 |

Notes

- See notes for Table 3.

Table 8: Mean Group Estimates

| Backward-Sum Model with country-specific trends | | | First-Differenced Model with country-specific intercepts | | |
|---|---|---|---|---|---|
| $\theta_1$ | mean | 0.0080<br>(5.71) | $\theta_1$ | mean | 0.0116<br>(2.70) |
| | median | 0.0071 | | median | 0.0092 |
| $\phi_4$ | mean | 0.0414<br>(6.47) | $\beta_0 + \beta_1$ | mean | 0.0722<br>(6.39) |
| | median | 0.0526 | | median | 0.546 |
| $\pi_4$ | mean | -0.3380<br>(-20.36) | $\alpha_1 + \alpha_2 - 1$ | mean | -0.7319<br>(-21.28) |
| | median | -0.3300 | | median | -0.7091 |
| Growth effect | mean | 0.0233<br>(3.64) | Growth effect | mean | 0.0210<br>(3.39) |
| | median | 0.0239 | | median | 0.0219 |
| Level effect | mean | 0.1238<br>(4.31) | Level effect | mean | 0.0955<br>(5.25) |
| | median | 0.1341 | | median | 0.1014 |

Notes:

- The number of countries is 94 for the model in levels, and 75 for the model in first-differences.
- The models are estimated using demeaned data.
- Robust means are reported; t-ratios are in parentheses.
- The backward-sum model is estimated using OLS, and the first-differenced model is estimated using instrumental variables. See Table 3 for information about instruments used.

Table 9: Mean Group Estimates for OECD and non-OECD countries

| Backward-Sum Model with country-specific trends | | | First-Differenced Model with country-specific intercepts | | |
|---|---|---|---|---|---|
| Growth effect OECD | Mean | 0.0511<br>(5.38) | Growth effect OECD | Mean | 0.0301<br>(3.46) |
| | Median | 0.0619 | | Median | 0.0449 |
| Level effect OECD | Mean | 0.2948<br>(4.27) | Level effect OECD | Mean | 0.2132<br>(7.81) |
| | Median | 0.1985 | | Median | 0.2197 |
| Growth effect Non-OECD | Mean | 0.0150<br>(2.11) | Growth effect Non-OECD | Mean | 0.0176<br>(2.26) |
| | Median | 0.0177 | | Median | 0.0166 |
| Level effect Non-OECD | Mean | 0.0812<br>(2.69) | Level effect Non-OECD | Mean | 0.0458<br>(2.02) |
| | Median | 0.335 | | Median | 0.0361 |

Notes:

- See Table 8 for notes.
- Variables are demeaned using full sample means.

Table 10: Mean Group Estimates (First-Differenced Model), Allowing for Cross-Section Dependence

| Including the means of $\Delta y_{it}$, $\Delta x_{it}$ and $x_{it}$ | | | Including the means of $\Delta y_{it}$, $\Delta y_{it-1}$, $\Delta y_{it-2}$, $\Delta x_{it}$, $\Delta x_{it-1}$ and $x_{it}$ | | |
|---|---|---|---|---|---|
| $\theta_1$ | mean | 0.0176<br>(3.38) | $\theta_1$ | mean | 0.0164<br>(3.04) |
| | median | 0.0234 | | median | 0.0207 |
| $\beta_0 + \beta_1$ | mean | 0.0465<br>(4.39) | $\beta_0 + \beta_1$ | mean | 0.0470<br>(4.23) |
| | median | 0.0381 | | median | 0.0385 |
| $\alpha_1 + \alpha_2 - 1$ | mean | -0.7852<br>(-25.91) | $\alpha_1 + \alpha_2 - 1$ | mean | -0.7992<br>(-25.29) |
| | median | -0.8001 | | median | -0.8091 |
| Growth effect | mean | 0.0260<br>(2.34) | Growth effect | mean | 0.0288<br>(2.25) |
| | median | 0.0294 | | median | 0.0295 |
| Level effect | mean | 0.0681<br>(1.26) | Level effect | mean | 0.0734<br>(4.10) |
| | median | 0.0542 | | median | 0.0664 |

Notes:

- See Table 8 for notes.
- The year means are calculated for the full sample.

Table 11: Mean Group Estimates (First-Differenced Model) for OECD and non-OECD Countries, Allowing for Cross-Section Dependence

| Including the means of $\Delta y_{it}$, $\Delta x_{it}$ and $x_{it}$ | | | Including the means of $\Delta y_{it}$, $\Delta y_{it-1}$, $\Delta y_{it-2}$, $\Delta x_{it}$, $\Delta x_{it-1}$ and $x_{it}$ | | |
|---|---|---|---|---|---|
| Growth effect OECD | Mean | 0.0097<br>(0.53) | Growth effect OECD | Mean | 0.0109<br>(0.57) |
| | Median | 0.0211 | | Median | 0.0217 |
| Level effect OECD | Mean | 0.1977<br>(5.29) | Level effect OECD | Mean | 0.2041<br>(4.85) |
| | Median | 0.2489 | | Median | 0.2167 |
| Growth effect Non-OECD | Mean | 0.0313<br>(2.30) | Growth effect Non-OECD | Mean | 0.0360<br>(2.12) |
| | Median | 0.0324 | | Median | 0.0408 |
| Level effect Non-OECD | Mean | 0.264<br>(1.43) | Level effect Non-OECD | Mean | 0.0209<br>(1.15) |
| | Median | 0.169 | | Median | -0.0037 |

Notes:

- See Table 8 for notes.
- The year means are calculated for the full sample.

## DATA APPENDIX

### LIST OF COUNTRIES

| | | | |
|---|---|---|---|
| Argentina | El Salvador | Madagascar | Singapore* |
| Australia | Ethiopia* | Malawi | South Africa |
| Austria | Finland | Malaysia | Spain |
| Bangladesh | France | Mali | Sri Lanka |
| Belgium | Gambia* | Mauritania | Sweden |
| Benin | Ghana | Mauritius* | Switzerland |
| Bolivia | Greece | Mexico | Syria |
| Brazil | Guatemala | Morocco | Taiwan* |
| Burkina Faso | Guinea* | Mozambique* | Tanzania* |
| Burundi | Guinea-Bissau* | Namibia* | Thailand |
| Cameroon | Guyana | Nepal* | Togo |
| Canada | Honduras | Netherlands | Trinidad and Tobago |
| Cape Verde* | Hong Kong | Niger | Turkey |
| Central African Rep. | Iceland | Nigeria | Uganda* |
| Chile | India | Norway | United Kingdom |
| China | Indonesia | Pakistan | United States of America |
| Colombia | Ireland | Panama* | Uruguay |
| Comoros* | Israel | Papua New Guinea | Zambia |
| Congo, Republic of | Italy | Paraguay | Zimbabwe* |
| Costa Rica | Jamaica | Peru | |
| Cote d'Ivoire | Japan | Philippines | |
| Denmark | Jordan* | Romania* | |
| Dominican Republic | Kenya | Rwanda | |
| Ecuador | Korea, Republic of | Senegal | |
| Egypt | Luxemburg | Seychelles* | |

Note: Countries that are marked with stars are excluded from the 75-country panel.

### DESCRIPTIVE STATISTICS

| | | Growth rate of output per worker | Investment share |
|---|---|---|---|
| FULL SAMPLE | Mean | 0.0175 | 15.8980 |
| | Standard Deviation | 0.0609 | 9.6339 |
| | First Quartile | -0.0078 | 8.4333 |
| | Second Quartile | 0.0199 | 14.0424 |
| | Third Quartile | 0.0465 | 22.3752 |
| | Min | -0.5430 | 0.7141 |
| | Max | 0.4611 | 101.1316 |
| | Number of observations | 3748 | 3842 |

- Data Source: Penn World Table 6.1
- The investment share is measured as total investment as a percentage of GDP.
- The sample covers the 1960-2000 period.