Capital Accumulation and Growth:

A New Look at the Empirical Evidence

Steve Bond — Nuffield College, Oxford and IFS

Asli Leblebicioglu — Boston College

Fabio Schiantarelli — Boston College and IZA

December 2, 2004

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 steady state. These results are found using pooled annual data for a large panel of countries, using pooled data for non-overlapping five-year periods, or allowing for heterogeneity across countries in regression coefficients. They are robust to model specifications and estimation methods. The evidence that investment has a long-run effect on growth rates is consistent with the main implication of certain endogenous growth models, such as the AK model.

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, H. Pesaran, L. Serven, J. Temple and seminar participants at the World Bank, Inter-American Development Bank, Boston University, Brandeis University, European University Institute, Ente Einaudi, and the European Meeting of the Econometric Society for helpful comments. We thank L. Gaul for excellent research assistance.

1 Introduction

An influential view has emerged which suggests that investment in physical capital is relatively unimportant in explaining economic growth. This is perhaps epitomized in the title of Easterly and Levine's (2001) review of the recent empirical literature: "it's not factor accumulation". More precisely, they conclude that "the data do not provide strong support for the contention that factor accumulation ignites faster growth in output per worker".[1]

Our analysis of annual data for some 98 countries in the period 1960-98 points to a quite different conclusion. Not only do we find that a higher share of investment in GDP predicts a higher *level* of output per worker in the steady state, 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 steady state. The long-run effect on growth rates is quantitatively substantial, as well as statistically significant. This evidence is consistent with the main implication of certain endogenous growth models, such as the AK model.

One key to our analysis is that our empirical models allow the long-run growth rate in each country to depend on the share of output 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 regard the suggestion that capital accumulation plays only a minor role in economic growth to be, at best, premature.

This issue is of such fundamental importance that it has naturally received considerable attention in previous research. Unlike much of the literature focused on issues of convergence,

---

[1] The role of capital accumulation is similarly downplayed in Easterly (2001). Econometric support is provided by Blomstrom, Lipsey and Zejan (1996) and Jones (1995).

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 considered Granger-causality between investment and growth, 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. Our preferred results allow for heterogeneity across countries in all regression coefficients, following the approach of Pesaran and Smith (1995) and Lee, Pesaran and Smith (1997). However our main results on the role of capital investment are also found when we pool annual data for all countries, or indeed when we consider a panel based on five-year periods, as suggested by Islam (1995) and Caselli, Esquivel and Lefort (1996).

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 steady state, while allowing also 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 results for panel specifications, in which the data for all countries is pooled. Section 6 presents the results based on individual time series models for each country. Section 7 concludes.

---

[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.

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, Δy_{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', η_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 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 β 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

---

[3] See also De Long and Summers (1991, 1993), who emphasise the role played by equipment investment; and Beaudry, Collard and Green (2002), who suggest that the effect of investment has become larger in more recent periods.

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 explanatory variable, 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 θ 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 β on the change in investment again indicates whether investment affects the (log) level of output per worker along the steady state growth path.

Both time series and panel studies have typically reported estimates of θ that are insignificantly different from zero in this type of model, suggesting that investment does not Granger-cause growth.[4] For example, Jones (1995) finds no effect of investment on the long-run growth rate, using annual data for individual OECD countries in the post-war period. Similarly Blomstrom, Lipsey and Zejan (1996), using pooled data for non-overlapping five-year periods and a

---

[4] Most specifications in the Granger-causality literature would replace the current change Δx_{it} in (4) with lagged differences of the share of investment in output, but this distinction is inessential for the main points we emphasize here.

large sample of developed and less developed countries, find that investment does not Granger-cause growth, while growth does Granger-cause investment.[5] While there are some exceptions in the more recent literature, such as Li (2002), this econometric support for the view that higher investment does not predict faster growth appears to have been influential.[6]

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 ε_{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 estimate of α will typically be downwards, so this approach will underestimate the degree of persistence in growth rates.[7] The bias in the estimates of β and θ is harder to sign *a priori*, but may be sufficiently serious to warrant further investigation. Similar biases will be present if the shocks ε_{it} contain a serially uncorrelated or indeed any stationary component. The only case in which least squares estimators could yield consistent estimates of the parameters in (4) would be when 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

---

[5] 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 and negative Granger-causality running from saving to growth.

[6] 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 of those with poorly educated populations.

[7] Interestingly the apparent lack of persistence in growth rates is another reason suggested by Easterly and Levine (2001) for their conclusion that capital accumulation cannot be a major influence on growth rates.

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 conclusion that a higher level of investment does not predict a higher long-run rate of growth appears to be based on these assumptions, and may therefore be premature.

A recent paper by Bernanke and Gurkaynak (2001) studies the cross-section correlation between investment shares and 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 is immune to 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 actual and not 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' θ and the 'level effect' β 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.[8] The most general specification we estimate will satisfy both objectives, although we will show that our main empirical results are also found in a range of more restrictive models.

---

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

3 Model Specifications

In this section we describe the models 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 general enough to encompass the predictions of different theories. Initially we will estimate these models for a panel of countries, using pooled data at the annual frequency. We will then present panel specifications using data for non-overlapping five-year periods. Finally, we will estimate individual time series models for each country, using annual data. 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 (or, in an alternative specification, the level) of the investment to GDP ratio. Assume that the behavior of y_{it} is represented by the following $ADL(p,p)$ (Autoregressive-Distributed Lag) model:

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

where ε_{it} is a mean zero, serially uncorrelated shock assumed to be independent across countries, conditional on c_{it}. We assume that c_{it} is a non-stationary process that determines the behavior of the growth rate of y_{it} in the steady state. This model nests simpler dynamic specifications like equation (2) that have been used to evaluate the Solow growth model, and in this context the c_{it} process would reflect the growth of total factor productivity.[9] We allow for richer dynamics in our empirical models based on annual data to control for business cycle influences.

It is useful first to express all variables as deviations from the average value calculated across all countries in the same time period. Taking this average of all the variables in (5) we obtain:

$$\overline{y}_{.t} = \overline{c}_{.t} + \alpha_1 \overline{y}_{.,t-1} + ... + \alpha_p \overline{y}_{.,t-p} + \beta_0 \overline{x}_{.,t} + \beta_1 \overline{x}_{.,t-1} + ... + \beta_p \overline{x}_{.,t-p} + \overline{\varepsilon}_{.t} \tag{6}$$

---

[9] Lee, Pesaran and Smith (1997) show that a similar specification can be obtained from a version of the Solow model that explicitly incorporates stochastic TFP shocks.

where, for example, $\overline{y}_{.t} = \frac{1}{N}\sum_{i=1}^{N} y_{it}$. Subtracting (6) from (5), we get:

$$\tilde{y}_{it} = \tilde{c}_{it} + \alpha_1\tilde{y}_{i,t-1} + \alpha_2\tilde{y}_{i,t-2} + ... + \alpha_p\tilde{y}_{i,t-p} + \beta_0\tilde{x}_{it} + \beta_1\tilde{x}_{i,t-1} + ... + \beta_p\tilde{x}_{i,t-p} + \tilde{\varepsilon}_{it} \tag{7}$$

where the variables with tildes denote deviations from these year-specific means, i.e. $\tilde{y}_{it} = y_{it} - \overline{y}_{.t}$, etc. .

We will experiment with different specifications of the process for c_{it} that embody different assumptions on how exactly investment affects the steady state growth rate of output per worker. One option is to assume that c_{it} evolves according to:

$$c_{it} = c_{i,t-1} + \gamma_0 + \gamma_1\bar{x}_{i.} + e_t. \tag{8}$$

This allows for a time-invariant drift $(\gamma_0 + \gamma_1\bar{x}_{i.})$ that varies across countries, and depends on the country's average investment share, $\bar{x}_{i.} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. The component e_t represents, instead, a common technological shock or other macro shock whose effect is common across countries. Taking deviations from year-specific means removes the common components $\gamma_0 + e_t$, giving:

$$\tilde{c}_{it} = \tilde{c}_{i,t-1} + \gamma_1\overline{\overline{x}}_i \tag{9}$$

where $\overline{\overline{x}}_{i.} = \bar{x}_{i.} - \frac{1}{N}\sum_{j=1}^{N}\bar{x}_{j.} = \frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{it}$.

The implied relationship between investment and growth in the steady state (i.e. when $\tilde{x}_{it} = \tilde{x}_{i,t-1} = \tilde{x}_i$ for all periods and $\tilde{\varepsilon}_{it}$ is set to its expected value of zero) is easily derived by taking first-differences of (7) and using the expression for $\Delta\tilde{c}_{it}$ in (9):

$$\widetilde{g}_i = \tilde{y}_{it}^{*} - \tilde{y}_{i,t-1}^{*} = \frac{\gamma_1\overline{\overline{x}}_i}{(1 - \alpha_1 - \alpha_2 - ... - \alpha_p)}$$

where $\widetilde{g}_i$ is the steady state growth rate and the superscript $*$ denotes steady state values of the variables. The last equation shows that the steady state growth rate of output per worker depends on the country-specific average share of investment in GDP (all expressed as deviations from year-specific means). Alternatively the steady state growth rate of actual output per worker for a country with a high share of investment will tend to be higher than the average for all countries if the parameter γ_1 is strictly positive.

Note that, solving (9) backward we obtain:

$$\tilde{c}_{it} = \tilde{c}_{i0} + \gamma_1 \overline{\overline{x}}_{i.} t. \tag{10}$$

Substituting (10) in (7) we obtain the representation:

$$\begin{aligned} \tilde{y}_{it} \quad = \quad & \tilde{c}_{i0} + \gamma_1 \overline{\overline{x}}_{i.} t + \ \alpha_1 \tilde{y}_{i,t-1} + \alpha_2 \tilde{y}_{i,t-2} + ... + \alpha_p \tilde{y}_{i,t-p} \\ & +\beta_0 \tilde{x}_{it} + \beta_1 \tilde{x}_{i,t-1} + ... + \beta_p \tilde{x}_{i,t-p} + \tilde{\varepsilon}_{it}. \end{aligned} \tag{11}$$

For convenience, we can reparameterize (11) as:

$$\begin{aligned} \Delta \tilde{y}_{it} \quad = \quad & \gamma_1 \overline{\overline{x}}_{i.} t + (\alpha_1 - 1) \Delta \tilde{y}_{i,t-1} + (\alpha_2 + \alpha_1 - 1) \Delta \tilde{y}_{i,t-2} + ... \\ & (\alpha_p + \alpha_{p-1} + ... + \alpha_1 - 1) \tilde{y}_{i,t-p} + \beta_0 \Delta \tilde{x}_{it} + (\beta_1 + \beta_0) \Delta \tilde{x}_{i,t-1} \\ & +... + (\beta_p + \beta_{p-1} + ... + \beta_0) \tilde{x}_{i,t-p} + \tilde{c}_{i0} + \tilde{\varepsilon}_{it}. \end{aligned} \tag{12}$$

Or, simply redefining the coefficients:

$$\begin{aligned} \Delta \tilde{y}_{it} \quad = \quad & \gamma_1 \overline{\overline{x}}_{i.} t + \pi_1 \Delta \tilde{y}_{i,t-1} + \pi_2 \Delta \tilde{y}_{i,t-2} + ... + \pi_p \tilde{y}_{i,t-p} \\ & +\phi_0 \Delta \tilde{x}_{it} + \phi_1 \Delta \tilde{x}_{i,t-1} + ... + \phi_p \tilde{x}_{i,t-p} + \tilde{c}_{i0} + \tilde{\varepsilon}_{it}. \end{aligned} \tag{13}$$

Note that, like equation (11), equation (13) is still a dynamic model for the (log) level of output per worker, provided $\pi_p \neq 0$. In particular, the error term in (13) will be serially uncorrelated if the shocks entering the process in (5) are serially uncorrelated. We are thus modelling the growth rate of output (in deviation form) in terms of its lags, an initial level of output, a distributed lag of the investment share, and an interaction between a trend and the country-specific average investment share. A general common trend process has also been controlled for by taking deviations of all variables from their year-specific means. The lags of Δy_{it} and Δx_{it} are included to control for fluctuations at business cycle frequencies. $\tilde{c}_{i0}$ 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 $\widetilde{\varepsilon}_{it}$.

The test of whether capital accumulation affects the growth rate of output per worker in the steady state is here simply a test of $\gamma_1/\pi_p = 0$. Evidence that γ_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 to a dynamic panel context the test of the Solow model proposed by Bernanke and Gurkaynak (2001) in a cross-sectional setting. The advantage of the panel approach is that it allows one to address the endogeneity issues that naturally arise when one investigates the effect of investment on growth, and that cannot be satisfactorily addressed in a single cross-section.

In the AK growth model, capital accumulation should affect the steady state growth rate, so we would expect γ_1/π_p to be significantly different from zero. In the standard AK model, there should be no transitional dynamics, while the model in (13) allows for transitional dynamics in approaching the steady state. However, extensions of the AK model imply transitional dynamics. For instance, adding to the AK production function a component with the standard neoclassical characteristics generates conditional convergence.[10] A slightly different type of conditional convergence can be obtained also from an endogenous growth model with technological diffusion.[11]

Even if there is no effect on the steady state growth rate, equation (13) allows for a steady state effect of investment on the level of output per worker, captured by the negative of ϕ_p/π_p. Putting it differently, given an initial level of income, the out of steady state growth rate depends upon the investment rate, as predicted by the Solow growth model.

The process for c_{it} in (8) implies that if the share of investment in output ($\tilde{x}_{it}$) is stationary,

---

[10] See, for instance, Barro and Sala-I-Martin (1995) p. 161 *et seq.*

[11] See, for instance, Barro and Sala-I-Martin (1995), chapter 8. For this last class of models, e_t in (8) may capture the evolution of output per worker in the leading country.

then the log of output per worker ($\widetilde{y}_{it}$) should be trend stationary, when expressed in deviations from year-specific means across all countries.[12] A different way to let investment influence the steady state growth rate, which allows for a unit root in the log of output per worker that is not common to all countries, is to assume that c_{it} in (5) evolves according to:

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

Here we allow the change in c_{it} to depend directly on the current share of investment, which we assume to be a stationary stochastic process. In deviation form this becomes:

$$\tilde{c}_{it} = \tilde{c}_{i,t-1} + \theta_1 \tilde{x}_{it}. \tag{15}$$

It follows that:

$$\tilde{c}_{it} = \tilde{c}_{i0} + \theta_1 \sum_{s=1}^{t} \tilde{x}_{is} \tag{16}$$

and:

$$\begin{aligned} \Delta\tilde{y}_{it} \quad = \quad & \theta_1 \sum_{s=1}^{t} \tilde{x}_{is} + \pi_1 \Delta\tilde{y}_{i,t-1} + \pi_2 \Delta\tilde{y}_{i,t-2} + ... + \pi_p \tilde{y}_{i,t-p} \\ & + \phi_0 \Delta\tilde{x}_{it} + \phi_1 \Delta\tilde{x}_{i,t-1} + ... + \phi_p \tilde{x}_{i,t-p} + \tilde{c}_{i0} + \tilde{\varepsilon}_{it}. \end{aligned} \tag{17}$$

In this specification the interaction term $\overline{\widetilde{x}}_{i.}t$ in (13) has been replaced by the backward sum of the investment shares, $bs\widetilde{x}_{it} = \sum_{s=1}^{t} \tilde{x}_{is}$. This term is clearly integrated of order one ($I(1)$), if $\tilde{x}_{it}$ is stationary. The process for c_{it} in (14) then implies that $\widetilde{y}_{it}$ is also $I(1)$ and cointegrated with the backward sum of investment shares.

The 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. Another way to think about the basic implication of this model is that now the coefficient of the country-specific time trend in the process for $\tilde{y}_{it}$ depends on the average of *past* investment shares (i.e.

---

[12] This can be seen from equation (13), in which the only non-stationary influence on the level of $\widetilde{y}_{it}$ is the deterministic trend. The untransformed series (y_{it}) will however be integrated of order one ($I(1)$), since these contain the 'permanent' shocks (e_t) that are common to all countries.

we can replace the term $bs\widetilde{x}_{it}$ by $\left(\frac{1}{t}\sum_{s=1}^{t}\tilde{x}_{is}\right)t$, simply by multiplying and dividing by t). As in the previous model, the steady state growth rate will depend on the (constant) level of the investment rate. This formulation may be more convenient for estimation because, when a long time series is available, we can appeal to results in the cointegration literature to address the issue of endogeneity of current investment.

There is an alternative representation of this last model that is also useful for estimation purposes. Taking first-differences of equation (7) and substituting for $\Delta\tilde{c}_{it}$ from equation (15), the model becomes:

$$\begin{aligned} \Delta\tilde{y}_{it} &= \alpha_1\Delta\tilde{y}_{i,t-1} + \alpha_2\Delta\tilde{y}_{i,t-2} + ... + \alpha_p\Delta\tilde{y}_{i,t-p} \\ &\quad +\theta_1\tilde{x}_{it} + \beta_0\Delta\tilde{x}_{it} + \beta_1\Delta\tilde{x}_{i,t-1} + ... + \beta_p\Delta\tilde{x}_{i,t-p} + \Delta\tilde{\varepsilon}_{it}. \end{aligned} \tag{18}$$

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 (7). Moreover in this form all the variables in the empirical specification are stationary, provided that the share of investment in output is stationary. An equation of this form, albeit not necessarily in deviation 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 $\theta_1/\pi_p = 0$ suggests a long-run effect of investment on growth, consistent with an endogenous growth approach.

A crucial point is that, if there are serially uncorrelated shocks ε_{it} that affect the (log) level of output per worker, as in (5), the error term in (18) has an MA(1) structure that makes it necessarily correlated with the lagged dependent variable $\Delta\tilde{y}_{i,t-1}$. More generally, the error term in the first-differenced specification would only be serially uncorrelated if the idiosyncratic shocks ε_{it} follow a random walk. Otherwise least squares estimates of the parameters in (18) will be biased and inconsistent. Consistent estimates may be obtained, provided the shocks ε_{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 ε_{it} and $\varepsilon_{i,t-1}$. We will explore the importance of these potential biases in such first-differenced specifications in our empirical analysis.

In all the specifications presented so far, 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 (18), 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.:

$$c_{it} = c_{i,t-1} + d_i + \theta_0 + \theta_1 x_{it} + e_t. \tag{19}$$

In this case we introduce a time-invariant country-specific component $(\widetilde{d_i})$ into the error term in equation (18). This generalization is particularly 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 corresponding extension to the dynamic model for the level of output per worker in equation (17) would introduce a set of unrestricted country-specific linear trends, in addition to the backward sum of investment variable. The process for c_{it} in (19) thus implies that the $I(1)$ output per worker and backward sum of investment variables $(\widetilde{y}_{it}$ and $bs\widetilde{x}_{it})$ are cointegrated after eliminating a country-specific deterministic trend.

This extension is not possible in the specifications based on (8). Introducing separate linear trends for each country would not allow the identification of the coefficient γ_1 on the interaction term in (13). Likewise if we first-difference this model, we cannot identify γ_1

if we allow for a time-invariant country-specific component in the resulting error term. In a similar spirit, we could investigate whether the country-specific coefficients on the trends in an extended version of (13) depend not only upon average investment shares, but also upon additional time-invariant country characteristics measuring, for example, institutional quality, policy choices and human capital. However it should be stressed that this approach is less general than the one discussed in the preceding paragraph, in that it cannot control for *unobserved* country-specific influences on steady state growth rates that are correlated with average investment shares.

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

$$c_{it} = c_{i,t-1} + d_i + \theta_0 + \theta_1 x_{it} + e_t + v_{it} \tag{20}$$

where the serially uncorrelated v_{it} reflect 'permanent' shocks to the logarithm of output per worker that are independent across countries, conditional on e_t.[13]

Taking deviations from year-specific means, and substituting the resulting specification for $\Delta\widetilde{c}_{it}$ into equation (7) in first-differences, gives our most general empirical model as:

$$\begin{aligned} \Delta\tilde{y}_{it} \quad = \quad & \alpha_1\Delta\tilde{y}_{i,t-1} + \alpha_2\Delta\tilde{y}_{i,t-2} + ... + \alpha_p\Delta\tilde{y}_{i,t-p} \\ & +\theta_1\tilde{x}_{it} + \beta_0\Delta\tilde{x}_{it} + \beta_1\Delta\tilde{x}_{i,t-1} + ... + \beta_p\Delta\tilde{x}_{i,t-p} + \widetilde{d}_i + \widetilde{v}_{it} + \Delta\tilde{\varepsilon}_{it}. \end{aligned} \tag{21}$$

This most general specification for c_{it} then has the implication that the $I(1)$ variables $\widetilde{y}_{it}$ and $bs\widetilde{x}_{it}$ are not cointegrated. However the model can still be estimated consistently in this first-differenced form. Provided a long time series is available, country-dummies can simply be included to allow for the heterogeneity in long-run growth rates reflected in $\widetilde{d}_i$. Instrumental

---

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

variables estimates of (21) can then allow for both transient shocks (ε_{it}) and permanent shocks (v_{it}) to the (log) level of output per worker in each country.

Notice that while the presence of transient shocks (ε_{it}) leads to biases when Ordinary Least Squares or Within estimators are used for models in first-differences like (21), the presence of country-specific permanent shocks (v_{it}) would introduce a random walk component into the error term of models in levels like (13) or (17). The resulting serial correlation would again result in biased estimates of coefficients, particularly those on lagged dependent variables, and in this context it is not clear that valid instrumental variables would be available to obtain consistent parameter estimates.

The various specifications presented above will constitute the basis of our empirical analysis. We will present both pooled estimates and estimates that allow for heterogeneity of all coefficients across countries. Moreover, we will generalize the models in some cases by allowing for additional regressors, such as the population growth rate, or proxies for human capital. Before presenting the econometric results, however, we will describe the data sources and the time series characteristics of the data used in estimation.

4 The Data: Sources and Time Series Properties

The data for estimating the basic model comes from the Penn World Table 6.0 (PWT 6.0) 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 98 countries and the annual data covers the period 1960-1998.[14] Even though the PWT 6.0 data set has information on more countries, we include only 98 of them in order to have a balanced panel data set. Following Mankiw, Romer and Weil (1992), we also excluded the countries

---

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

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.[15]

The number of countries in our sample drops to 76 when we use instrumental variable estimators, due to the lack of observations 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 as a percentage of GDP. All these variables are taken from the World Bank World Development Indicators (WDI) 2002. 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 per worker and the log of the investment share. 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 98 countries, allowing for 4 lags.[16] Both variables are measured as deviations from their year-specific cross-country mean values, in order to control for common trends. This is also the form in which the variables enter in all our regression models.

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 33 countries when we do not allow for a trend, and for 21 countries when we allow for a trend. The number of countries for which this test rejects the non-stationarity of the first-difference

---

[15] The excluded oil producers are Bahrain, Gabon, Iran, Iraq, Kuwait, Oman, Saudi Arabia, and the United Arab Emirates.

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

of the log investment share is larger: 46 without the trend and 26 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 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. 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 trend, but not if we consider the version with a trend. We suspect that 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 in deviations from year-specific means form.

Finally we consider cointegration between the log of GDP per worker ($\widetilde{y}$) and the backward sum of investment shares variable ($\widetilde{bsx}$), 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 a deterministic trend, which test the null hypothesis of no cointegration. These results suggest that these two variables are cointegrated when we allow for country-specific deterministic trends, but not otherwise.

Formally these time series properties are consistent with the process for c_{it} represented by equation (19), which suggests that in deviations from year-specific means, the log of GDP per worker should be an I(1) variable that is cointegrated with the backward sum of investment shares when we allow for country-specific trends. 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 a range of specifications introduced in the previous section, and focus on econometric results that are appropriate in each case given the time series implications of the corresponding processes for c_{it}.

5 Panel Results

We first present the results obtained when the dynamic models for the log level of GDP per worker, equations (13) and (17), are estimated using pooled annual data. We will also estimate the latter model in first-differenced form (equation (18) or (21)), which allows us to control for unobserved country-specific influences on growth rates. In addition, we will present GMM estimates of this differenced model obtained using pooled data for non-overlapping five-year periods. Finally in this section we will assess the robustness of the pooled annual results to changes in the functional form, to the measure of investment used, and to the inclusion of additional population growth and human capital variables suggested by the (augmented) Solow growth model.

5.1 Basic Results on Yearly Data

For the 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. We first present results obtained using the country-specific mean of the logarithm of the investment share ($\overline{\overline{x}}_{i.}$) interacted with a time trend as a regressor, followed by those obtained using the backward sum variable ($bs\widetilde{x}_{it}$). This last model will also be estimated in its first-differenced form on annual data, and also using pooled data for non-overlapping five-year periods.

In Table 2 we report the results obtained from estimating equation (13), in which the long-run growth rate of output per worker depends on the country's overall mean share of investment in GDP. Table 3 reports estimates of equation (17), in which the long-run growth rate depends on the mean of current and past shares of investment. In both tables we present the Within estimates of the basic model, and the Within estimates of an alternative specification that excludes the contemporaneous $\Delta\tilde{x}_{it}$ (and further uses $bs\widetilde{x}_{i,t-1}$ instead of $bs\widetilde{x}_{it}$ in Table 3).[17] In addition, we present Instrumental Variables estimates. In this last case our sample drops from 98 countries to 76 because of data availability. All these specifications allow for unobserved time-invariant country-specific factors to affect the steady state log level of output per worker, but do not allow for unobserved heterogeneity in long-run growth rates.

The process for c_{it} in (8) that underlies the specification reported in Table 2 implies that $\widetilde{y}_{it}$ is trend stationary, with a trend that varies across countries with average investment shares. In this case the possible endogeneity of current investment could result in serious biases for least squares estimators, and our Instrumental Variables estimates may be more appropriate. Standard inference procedures can however be applied. In Table 2, $\Delta\tilde{x}_{it}$ is instrumented using lagged values of log investment and output per worker, plus lags of inflation (measured by the log of one plus the inflation rate), and the logs of trade and government spending (both expressed as percentages of GDP). Note that the interaction term between the time trend and the country-specific mean investment rate, $\bar{\widetilde{x}}_i$, is not instrumented in this specification.

The process for c_{it} in (14) that underlies the specification reported in Table 3 implies that $\widetilde{y}_{it}$ is non-stationary and cointegrated with $bs\widetilde{x}_{it}$. In this case the Within estimates of the long-run growth effect, $-\theta_1/\pi_4$, are superconsistent, whether or not the specification includes

---

[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.

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. For comparison we also report the Instrumental Variables estimates, in which we instrument both $\Delta\tilde{x}_{it}$ and $bs\widetilde{x}_{it}$, adding $bs\widetilde{x}_{i,t-1}$ to the instrument list.

In all these cases 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 $-\gamma_1/\pi_4$ in Table 2, and by $-\phi_4/\pi_4$ and $-\theta_1/\pi_4$ in Table 3.[18] The hypothesis that the long-run effect on the growth rate is equal to zero can be rejected at the 1% significance level, using either the Instrumental Variables results in Table 2 or the Within results in Table 3.[19]

The size of this estimated effect of investment on long-run growth is also quite large. For instance, using the Within results presented in Table 3, a country that has an investment share equal to the third quartile of the sample distribution (21.61%) has an annual growth rate of output per worker that is 0.92 percentage points higher than a country that has the median investment share (13.47%). The difference between a country at the third quartile and one at the first quartile (7.93%) is estimated to be 1.95 percentage points.

The standard test of over-identifying restrictions does not suggest any gross mis-specification for our Instrumental Variables specifications. Similarly, the tests of serial correlation proposed by Arellano and Bond (1991) do not suggest the presence of either first-order or second-order serial correlation in the shocks ($\tilde{\varepsilon}_{it}$) to the log level of output per worker, given the dynamics included in these models.

As we explained in section 3, we can also estimate the model containing the backward

---

[18] Note also that in all the specifications reported in Tables 2 and 3, the coefficient π_4 on the lagged level of log output per worker is negative, 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).

sum of investment shares in its first-differenced form (see (18)). 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 (18), the effect of investment on the steady state growth rate is captured by the coefficient on the level of investment. If there are no country-specific factors besides investment in the equation determining the evolution of c_{it} (see (14)), 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, then such country-specific effects are present, as in (21), and they can be controlled for in estimation. As we discussed earlier, consistent estimation may also require an Instrumental Variables approach if there are serially uncorrelated shocks (ε_{it}) to the (log) level of output per worker.

In the first two columns of Table 4 we report the OLS results (with and without contemporaneous information). These do not control for either MA(1) errors or for country effects, and are presented for comparison purposes. Even if there are no country-specific effects in the first-differenced model, OLS will yield inconsistent estimates in this autoregressive specification if the errors in the corresponding levels equations are serially uncorrelated, which implies an MA(1) error structure in the differenced equations. For our sample of annual data for 98 countries, this approach nevertheless suggests a significant positive coefficient on the level of investment or, in the specification with no contemporaneous information, that investment Granger-causes growth. The tests of serial correlation indicate no problem with the residuals. However these are based on parameter estimates that may be seriously biased, and we would treat the findings in these columns with considerable caution.

In the third column we report an Instrumental Variables estimator for this specification, using instruments dated t-2 and earlier to allow for the possible MA(1) structure of the error term in the first-differenced model. The results are quite different, consistent with our suspicion that OLS estimates of this dynamic model for growth rates are likely to be seriously biased. The sum of the estimated coefficients on the lagged dependent variables is close to one and the long-run effects are poorly determined. However these IV results will also be biased if there is unobserved heterogeneity in steady state growth rates. In the fourth column we allow for this case by controlling for unobserved country-specific effects in the first-differenced equations. The significance of the coefficient on the level of investment increases, but we also note that this dynamic model in first-differences appears to be heavily over-parameterised. In the final column we report our preferred IV results for a more parsimonious dynamic specification, in which both the growth and level effects are significant at the around the 1% level.[20]

It is useful to compare these preferred results, which allow for both the MA(1) error structure and for unobserved heterogeneity in growth rates, to the basic OLS results presented in column 1 or 2 of Table 4. The serial correlation tests for our preferred specification indicate the expected MA(1) form of serial correlation, implied by the presence of transient shocks to the log level of GDP per worker. This source of bias appears to dominate that due to unobserved country-specific effects, resulting in an OLS estimate of the coefficient on the lagged dependent variable that is severely biased downwards. Our preferred estimates suggest a much higher degree of persistence in growth rates than has typically been reported.

While this IV estimate in the first-differenced model is much *higher* than the corresponding OLS estimates, it also suggests a much *faster* ‘speed of convergence’ than that estimated in the levels models reported in Tables 2 and 3.[21] This points to the potential importance of unobserved heterogeneity across countries in steady state growth rates, and is one of the motivations for our exploration of time series models for individual countries in section 6, which

---

[20] Similar results were obtained for alternative, more or less restrictive, versions of this specification. Notice that more informative instruments are available when we omit the insignificant longer lags from the equation, and this allows more precise estimates of the parameters to be obtained.

[21] The interpretation of the ‘speed of convergence’ becomes more subtle in models that allow for differences across countries in long-run growth rates. We estimate relatively fast adjustment to each country’s steady state growth path, but these growth paths may themselves be diverging.

allow all parameters to vary across countries. Nevertheless we have shown that the significant positive effect of investment on long-run growth rates, reported in Tables 2 and 3, is robust to controlling for this unobserved heterogeneity in growth rates, at least in the context of the pooled annual panel data models considered in this section.

5.2 Differenced Model on Five-Year Averages

In this section we present estimates of the first-differenced model, using data on average growth rates and investment shares for non-overlapping five-year periods. Taking such 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 seven five-year periods.

This kind of specification has been used by Blomstrom, Lipsey, and Zejan (1996) to test for the effect of capital accumulation on steady state growth. They report OLS estimates, both with and without country dummies. However neither the OLS nor the Within estimator are appropriate for models in first-differences that contain lagged dependent variables and other endogenous variables as regressors, when the number of time periods used is very small. We apply the Instrumental Variables estimators suggested in Arellano and Bond (1991) to this data set.[22]

Having experimented with different lag structures, we focus on the results for the following simple specification:

$$\Delta_5 \tilde{y}_{it} = \alpha_1 \Delta_5 \tilde{y}_{i,t-5} + \theta_1 \tilde{x}^A_{it} + \beta_0 \Delta_5 \tilde{x}^A_{it} + \Delta_5 \tilde{\varepsilon}_{it} \tag{22}$$

where $\Delta_5 \tilde{y}_{it} = \tilde{y}_{it} - \tilde{y}_{i,t-5}$, and $\tilde{x}^A_{it}$ denotes the average value of the log of the investment share

---

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

over this five-year period. 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. We did experiment with second-differenced specifications, using instruments that remain valid in the presence of an MA(2) error process. The 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.

Table 5 reports the results for the first-differenced specification. The columns denoted by IV1 report the one-step estimator in Arellano and Bond (1991), while those denoted by GMM report the two-step estimator, with standard errors that use the finite-sample correction proposed by Windmeijer (2000). In the 'small' instrument set we include only the logs of GDP per worker and the investment share, each lagged 2, 3 and 4 times. The 'augmented' instrument set also includes lags 2, 3 and 4 of the policy variables, which again results in a smaller sample of countries. As expected, there is evidence of first-order serial correlation in the first-differenced residuals, while the hypothesis of no second-order serial correlation cannot be rejected. The Sargan-Hansen test of over-identifying restrictions for the GMM estimates suggests that our instruments are valid and that there is no gross form of mis-specification.

For all the estimators and for all the instrument sets, the asymptotic t test on the coefficient θ_1 on the log level of the average investment share suggests a significant positive effect of investment on the steady state growth rate. The long-run effect of investment on the level of output per worker is also highly significant when we use the augmented instrument set. The implied annual speed of convergence, λ, is between 0.104 and 0.118, which is very similar to that reported by Caselli, Esquivel and Lefort (1996), using similar GMM estimators for data on five-year periods.[23]

---

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

One advantage of the five-year average specification is that it allows us to control for measures of the percentage of the working age population in secondary schooling, or measures of the average years of schooling, which are not available annually for many countries. All the results reported in this section were robust to including both the level and the first-difference of these schooling variables in extended specifications, and these human capital terms were found to be insignificant at conventional levels.

5.3 Robustness and Extensions

In this section we present a set of experiments to investigate whether our results are robust to various extensions of the model. We first re-estimate the model reported in 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 population growth as an additional explanatory variable. For brevity we focus on results for the model containing the backward sum of investment shares, estimated in levels rather than in first-differences. As we will see, our general conclusions are robust to these variations.

More specifically, the results are robust to changes in the functional form. When the level of the investment share is used as a regressor (see Table 6), we obtain qualitatively very similar results to those reported previously using its logarithm (see Table 3). In particular, both the level and the growth effects of investment remain highly significant. All of the 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. The quantitative effect of investment on growth is now estimated to be somewhat larger. For instance, going from the median to the third quartile of the investment share is associated with a positive growth rate differential in the steady state of 1.14 percentage points per year (using the Within results in Table 6), compared with 0.92 percentage points (using the Within results in Table 3).

In Table 7 we present the results obtained when the logarithm of fixed investment as a share of GDP is used, instead of total investment. Because of data availability, the total number of observations drops considerably (from 3430 to 2262, when we use the Within estimator, and from 2456 to 1768, when we require instruments). However, we still obtain positive and significant growth and level effects for the Within and IV estimates of our basic specification, although not in this case for the model that excludes contemporaneous investment information. The size of the estimated growth effects tends to be larger here than when total investment is used.

We now add to the specification reported in Table 3 four lags of the log of the population growth rate (plus a common depreciation rate of 3% and a common fixed growth rate of 2%, as in Mankiw, Romer and Weil (1992)). This term would appear in standard tests of conditional convergence implied by the Solow growth model. The long-run parameters estimated for this specification are reported in Table 8. The sum of the coefficients on the population growth rate terms, denoted by $\widetilde{z}_{it}$, is negative, as expected (see the coefficient on $\widetilde{z}_{i,t-4}$), although this is estimated imprecisely in the Instrumental Variables results. The long-run effects of investment on both the level and the growth rate of output per worker remain very similar to those reported in Table 3, both in terms of size and statistical significance.

6 Heterogeneous Coefficients: Individual Country Results and Mean Group Estimates

In this section we will re-estimate the various models that we have developed, using annual time series data for individual countries. The main objective of this exercise is to see whether our conclusion that investment has both a level and a growth effect is also supported by the evidence when we allow all the parameters to differ across countries. The trade off here is obvious: on the one hand, by imposing equality of coefficients across countries, as we have done in the pooled models, one gains 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 steady state growth. In this case, relying on time series regressions for individual countries may provide a more accurate picture of the average effect of investment on steady state growth, and of its dispersion across countries.[24] The usefulness of this exercise is enhanced by the fact that formal tests of the validity of pooling tend to reject the restrictions of identical coefficients across countries, even at the 1% level.

As we have done so far, we will continue to measure each variable as deviations from year-specific means. If the slope coefficients are identical across countries, this exactly controls for the presence of common factors that affect growth in all countries. If the slope coefficients differ across countries, this procedure is only an approximate, although still useful, way to deal with such common influences.

The results for the model in which steady state growth depends on the country's mean share of investment are reported in Table 9; those for the model in which steady state growth depends on the mean of current and past investment shares are reported in Table 10; and those for the first-differenced version of this model are reported in Table 11. In all three tables, the first three columns report the results models that include four lags of both the log of GDP per worker and the log of the investment share for all countries. In the last column we report the results obtained for the preferred estimation method when the optimal lag length for each country is chosen according to the Akaike Criterion. We report summary count information on

---

[24] 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.

the sign and significance of individual coefficients, in addition to their median and mean values across countries. 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.[25] This is a robust variant of what Pesaran and Smith (1995) call the Mean Group estimator.

In the model using the country's overall mean share of investment, the coefficient on the time trend is $\gamma_{1i}\overline{\overline{x}}_{i.}$ (see (13)). We now include in each individual country regression a time trend. If we assume that variation across countries in average investment shares is the only source of heterogeneity in long-run growth rates, we can obtain an estimate of γ_{1i} by dividing the coefficient on the time trend by the country's value of $\overline{\overline{x}}_{i.}$. These estimates are reported in Table 9. The growth effect is estimated to be positive in around two thirds of the sample countries, and significant at the 5% level in about one third of the sample countries. The robust estimate of the mean of these growth effects is positive and significantly different from zero, and very similar to the median of the sample distribution. These results do not depend on whether we omit or instrument the current value of the investment share. When we use the optimal lag lengths for each country in the instrumental variable estimation, the size of the long run growth effect become slightly smaller, but it remains significantly different from zero. The level effect, $-\phi_{4i}/\pi_{4i}$, is also estimated to be positive in around two thirds of the countries, and the robust mean of these estimates is positive and significantly different from zero. The size of both the average level and growth effects is smaller than those estimated when the data are pooled (cf. Table 2). The average growth effect is around half the size of that obtained in the panel estimates of this model, although it is still substantial.

Another way to assess the importance of investment for steady state growth is to regress

---

[25] 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).

the country-specific coefficients on the trend terms on the average investment shares. The results of this (robust) cross-section regression are reported at the bottom of Table 9, under the heading Second Stage Regression. The coefficient on the average investment share is positive and highly significant, and very similar to the robust estimate of the mean of γ_{1i} obtained under the restriction discussed in the preceding paragraph.

This approach allows us to control for other *observed* factors that may also influence the long-run growth rate. For this reason we have also added to this second stage regression the average values of the country's inflation rate, the logs of trade and government spending (both expressed as a percentage of GDP), the percentage of the working age population with secondary schooling, and the country's initial log of GDP per worker. This is equivalent to allowing the change in $\tilde{c}_{it}$ in (9) to depend on all these variables, in addition to average investment. As a result the coefficient on the country-specific trend becomes a linear function of all of them. Even controlling for all these factors, however, the coefficient estimated on the average investment share remains positive, significant and of similar size. Note, in passing, that the estimated coefficient on public spending in this second stage regression is negative and significant at the 5% level, while that for trade is positive and significant at the 5% level. The coefficient on inflation is negative and significant at the 10% level, while the coefficients on initial income and the percentage of the working age population in secondary schooling are not significant.

In Table 10 we summarize the individual country results for the model containing the backward sum of (log) investment shares, rather than a linear trend (see (17)). Here the effect of investment on the steady state growth rate is captured by $-\theta_{1i}/\pi_{4i}$. In this case the growth effect is again found to be positive in around two thirds of the countries, and is significant at the 5% level in more than one third of them. The robust estimates of the mean value are positive and highly significant, and again similar in magnitude to the medians. The long-run effect of investment on the level of output per worker along these steady state growth paths is also positive in around two thirds of the sample countries, with a robust mean that is significantly different from zero. The size of these long-run effects of investment are somewhat smaller than those obtained for the same specification when the data are pooled (cf. Table 3). For example, a country that has an investment share equal to the third quartile has a steady state annual growth rate of output per worker that is 0.78 percentage points higher than a country with the median investment share, compared to 0.92 percentage points suggested by the corresponding panel results (using the Within estimates).

The validity of the estimates in Table 10 depends on the absence of other factors affecting the long-run growth rate, that may be correlated with levels of investment. However, the presence of any time-invariant observed or unobserved additional factors can easily be controlled for here by taking first-differences. In Table 11 we report the results for this model estimated in first-differenced form (see (18) or (21)). The inclusion of separate intercepts for each country then controls for unobserved country-specific influences on long-run growth rates. Note also that this first-differenced model can further allow for the possibility of country-specific permanent shocks (v_{it}), as in (20).

In this case we obtain average estimates of the long-run effects of investment, on both the level and the growth rate of output per worker, that are similar to those suggested by the corresponding levels model in Table 10. In particular, the preferred IV estimates of the first-differenced model suggest a robust mean growth effect that is positive and significant, and becomes somewhat higher in the specification where we select an optimal lag length for each country using the Akaike criterion. In this case the difference in long-run growth rates between countries with investment shares at the median and at the third quartile is estimated to be 0.87 percentage points. Although the estimates of the growth effects are individually significant here for relatively few countries, we still find that the significance of the average effect is robust to allowing for both country-specific drifts and permanent shocks. We again find that allowing for these forms of unobserved heterogeneity in long-run growth rates has more effect on the estimates of the autoregressive parameters, and on average these again suggest a faster 'speed of convergence' than that obtained from the corresponding specification in levels.

7 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 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 consistent with the main implication of certain endogenous growth models, such as the AK model. However it should be stressed that they do 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. 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.

We note that our results also do not rule out models in which long-run growth rates are exogenous. Investment is a forward-looking decision and investment levels are likely to rise when investors anticipate faster future growth. In subsequent work we will explore the extent to which this mechanism can be distinguished from a causal effect of current investment on future growth.

We also plan further work to explore the heterogeneity in the effects of investment suggested by our analysis of models for individual countries. Are there systematic differences by region or by level of development? 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?[26] 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.

---

[26] 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] 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.

[2] 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.

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

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

[5] 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.

[6] 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.

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

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

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

[10] 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.

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

[12] 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.

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

[14] 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 Rev,* 2001, 15,177-219.

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

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

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

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

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

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

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

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

[23] 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.

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

[25] 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.

[26] Pedroni, P., “Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the Purchasing Power Parity Hypothesis”, Indiana University Working Paper no. 95-013, 1995, forthcoming in *Econometric Theory.*

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

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

[29] 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.

[30] Windmeijer, F., “A Finite Sample Correction for the Variance of Linear Two-Step GMM Estimators”, Institute for Fiscal Studies Working Paper no. W00/19, 2000, forthcoming in *Journal of Econometrics.*

Table 1A: Augmented Dickey Fuller Test

		Without trend	With trend
$\tilde{y}$	5%	4	4
	10%	3	4
$\Delta\tilde{y}$	5%	33	21
	10%	14	9
$\tilde{x}$	5%	8	5
	10%	5	6
$\Delta\tilde{x}$	5%	46	26
	10%	18	19

- The figures indicate the number of countries for which the null hypothesis of a unit root is rejected at the corresponding marginal significance level.
- The total number of countries is 98, and the sample period covers 1960-98.
- The number of lags used in the test is 4.

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

	Without trend	With trend
$\tilde{y}$	5.360	0.416
$\Delta\tilde{y}$	-12.509	-9.071
$\tilde{x}$	-2.794	-0.456
$\Delta\tilde{x}$	-14.976	-10.906

- Reported figures are the W statistic in Im, Pesaran Shin (2003) (Psi statistic in Stata outputs). Under the null hypothesis of a unit root (where the alternative is one sided), the test statistic is asymptotically distributed as a standard normal.

Table 1C: Pedroni Cointegration Test Between $\tilde{y}$ and $bs\tilde{x}$.

	Without trend	With trend
Panel ADF	2.440	-2.128
Group ADF	1.039	-4.459

- 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.

Table 2: Annual Panel Regressions: Using Overall Mean of Investment

	Within	Within (no cont. info)	IV
$\Delta\tilde{y}_{it-1}$	0.0275 (0.98)	0.0442 (1.59)	-0.0028 (-0.06)
$\Delta\tilde{y}_{it-2}$	-0.0698 (-2.69)	-0.0714 (-2.73)	-0.0327 (-0.87)
$\Delta\tilde{y}_{it-3}$	-0.0601 (-2.44)	-0.0568 (-2.27)	-0.0547 (-1.76)
$\tilde{y}_{it-4}$	-0.0545 (-7.87)	-0.0534 (-7.67)	-0.0603 (-5.96)
$\Delta\tilde{x}_{it}$	0.0449 (4.35)		0.2415 (3.68)
$\Delta\tilde{x}_{it-1}$	0.0428 (4.99)	0.0285 (2.98)	0.1154 (3.81)
$\Delta\tilde{x}_{it-2}$	0.0251 (3.81)	0.0139 (1.95)	0.0792 (3.57)
$\Delta\tilde{x}_{it-3}$	0.0342 (4.48)	0.0244 (3.20)	0.0707 (3.76)
$\tilde{x}_{it-4}$	0.0341 (6.45)	0.0243 (4.44)	0.0784 (4.62)
$\bar{\bar{x}}_i t$	0.0011 (5.43)	0.0010 (4.72)	0.0016 (4.71)
Growth effect: $-\frac{\gamma_1}{\pi_4}$	0.0206 [0.0000]	0.0184 [0.0000]	0.0263 [0.0000]
Level effect: $-\frac{\phi_4}{\pi_4}$	0.6252 [0.0000]	0.4551 [0.0000]	1.3002 [0.0000]
R^2	0.1763	0.1561	
Number of observations	3430	3430	2456
Number of countries	98	98	76
Test of first order serial correlation	-0.00 [0.9978]	0.08 [0.9394]	1.39 [0.1651]
Test of second order serial correlation	-0.68 [0.4974]	-0.79 [0.4285]	0.22 [0.8243]
Test of over-identification			[0.2499]

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.
- In the IV specification, $\Delta\tilde{x}_{it}$ is instrumented. The instrument set is lags 5 and 6 of $\tilde{x}_{it}$ and $\tilde{y}_{it}$, lags 1 and 2 of log of annual GDP 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.
- In the last row the p-value of the test of over-identifying restrictions is reported in square brackets.

Table 3: Annual Panel Regressions: Using the Backward-sum of Investment

	Within	Within (no cont. info.)	IV
$\Delta\tilde{y}_{it-1}$	0.0280 (1.00)	0.0466 (1.67)	-0.0105 (-0.23)
$\Delta\tilde{y}_{it-2}$	-0.0694 (-2.68)	-0.0696 (-2.67)	-0.0408 (-1.10)
$\Delta\tilde{y}_{it-3}$	-0.0589 (-2.40)	-0.0541 (-2.16)	-0.0627 (-1.99)
$\tilde{y}_{it-4}$	-0.0549 (-7.56)	-0.0511 (-6.94)	-0.0757 (-6.58)
$\Delta\tilde{x}_{it}$	0.0457 (4.43)		0.2425 (3.86)
$\Delta\tilde{x}_{it-1}$	0.0431 (5.03)	0.0288 (2.97)	0.1172 (3.91)
$\Delta\tilde{x}_{it-2}$	0.0247 (3.37)	0.0136 (1.87)	0.0807 (3.66)
$\Delta\tilde{x}_{it-3}$	0.0331 (4.37)	0.0233 (3.05)	0.0717 (3.80)
$\tilde{x}_{it-4}$	0.0310 (6.12)	0.0211 (3.92)	0.0781 (4.57)
$bs\,\tilde{x}_{it}$	0.0011 (4.96)		0.0023 (5.19)
$bs\,\tilde{x}_{it-1}$		0.0008 (3.62)	
Growth effect: $-\frac{\theta_1}{\pi_4}$	0.0195 [0.0000]	0.0157 [0.0000]	0.0300 [0.0000]
Level effect: $-\frac{\phi_4}{\pi_4}$	0.5651 [0.0000]	0.4141 [0.0000]	1.0313 [0.0000]
R^2	0.1754	0.1538	
Number of observations	3430	3430	2456
Number of countries	98	98	76
Test of first order serial correlation	-0.05 [0.9598]	0.03 [0.9765]	1.60 [0.1086]
Test of second order serial correlation	-0.69 [0.4924]	-0.81 [0.4180]	0.31 [0.7593]
Test of over-identification			[0.2811]

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.
- In the IV specification, $bs\,\tilde{x}_{it}$ and $\Delta\tilde{x}_{it}$ are instrumented. The instrument set is lags 5 and 6 of $\tilde{x}_{it}$ and $\tilde{y}_{it}$, lag 1 of $bs\,\tilde{x}_{it}$, lags 1 and 2 of log of annual GDP 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.
- In the last row the p-value of the test of over-identifying restrictions is reported in square brackets.

Table 4: Annual Panel Regressions: Model in Differences

	OLS	OLS (no cont.info)	IV	IV (with country effects)	Restricted IV (with country effects)
$\Delta\tilde{y}_{it-1}$	0.0980 (3.44)	0.1165 (4.08)	1.1893 (3.04)	0.4986 (1.20)	0.6624 (3.45)
$\Delta\tilde{y}_{it-2}$	0.0004 (0.02)	-0.0033 (-0.13)	-0.1764 (-1.73)	-0.0374 (-0.43)	-0.0631 (-1.34)
$\Delta\tilde{y}_{it-3}$	0.0151 (0.64)	0.0162 (0.68)	0.0352 (0.44)	-0.0300 (-0.48)	
$\Delta\tilde{y}_{it-4}$	-0.0033 (-0.15)	-0.0059 (-0.27)	-0.0698 (-1.56)	-0.0372 (-0.89)	
$\tilde{x}_{it}$ ($\tilde{x}_{it-1}$ in the second column)	0.0128 (7.51)	0.0106 (6.24)	0.0040 (0.83)	0.0559 (1.21)	0.0376 (1.93)
$\Delta\tilde{x}_{it}$	0.0368 (3.31)		-0.0118 (-0.05)	0.2165 (1.00)	0.1312 (1.50)
$\Delta\tilde{x}_{it-1}$	0.0295 (3.66)	0.0189 (2.11)	0.1360 (1.01)	0.0493 (0.54)	0.0461 (0.98)
$\Delta\tilde{x}_{it-2}$	0.0075 (1.34)	0.0016 (0.27)	0.0108 (0.38)	0.0078 (0.33)	
$\Delta\tilde{x}_{it-3}$	0.0130 (1.97)	0.0099 (1.53)	0.0251 (1.32)	0.0082 (0.57)	
$\Delta\tilde{x}_{it-4}$	0.0136 (2.36)	0.0112 (1.96)	0.0084 (0.54)	0.0175 (1.33)	
$\alpha_1+\alpha_2+\alpha_3+\alpha_4-1$	-0.8898 [0.0000]	-0.8764 [0.0000]	-0.0218 [0.9489]	-0.6059 [0.1245]	-0.4007 [0.0244]
Growth effect: $\frac{\theta_1}{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}$	0.0144 [0.0000]	0.0121 [0.0000]	0.1818 [0.9449]	0.0922 [0.0065]	0.0939 [0.0080]
Level effect: $\frac{\beta_0+\beta_1+\beta_2+\beta_3}{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}$	0.1127 [0.0000]	0.0476 [0.0140]	7.7137 [0.9489]	0.4940 [0.0066]	0.4424 [0.0118]
R^2	0.1656	0.1403			
Number of observations	3332	3332	2447	2447	2447
Number of countries	98	98	76	76	76
Test of first order serial correlation	1.21 [0.2254]	1.42 [0.1559]	-2.84 [0.0045]	-0.79 [0.4272]	-2.39 [0.0166]
Test of second order serial correlation	-1.32 [0.1884]	-1.29 [0.1972]	-0.58 [0.5624]	-0.74 [0.4620]	-0.78 [0.4358]
Test of over-identification			[0.786]	[0.491]	[0.200]

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.
- In the first two IV specifications $\tilde{x}_{it}$, $\Delta\tilde{x}_{it}$, $\Delta\tilde{x}_{it-1}$ and $\Delta\tilde{y}_{it-1}$ are instrumented. The instrument set is lags 5 and 6 of $\tilde{x}_{it}$ and $\tilde{y}_{it}$, 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. In the restricted IV specification with country effects, $\Delta\tilde{y}_{it-3}$, $\Delta\tilde{y}_{it-4}$, $\Delta\tilde{x}_{it-2}$, $\Delta\tilde{x}_{it-3}$ and $\Delta\tilde{x}_{it-4}$ are added to the instrument set.
- The covariance matrix in all estimations allows for heteroskedasticity and MA(1) errors.
- In the last row the p-value of the test of over-identifying restrictions is reported.

Table 5: Five-Year Averages Panel: Model in Differences

	Small	Instrument Set	Augmented	Instrument Set
	IV1	GMM	IV1	GMM
$\Delta_5 \tilde{y}_{it-5}$	0.5791 (4.62)	0.5932 (3.16)	0.5636 (7.34)	0.5544 (6.47)
$\tilde{x}_{it}^{A}$	0.0275 (2.28)	0.0352 (2.53)	0.0455 (4.85)	0.0431 (4.41)
$\Delta_5 \ \tilde{x}_{it}^{A}$	0.0506 (0.64)	0.1381 (1.59)	0.1776 (3.55)	0.1728 (3.07)
Growth Effect	0.0653	0.0865	0.1043	0.0967
Level Effect	0.1202	0.3395	0.4070	0.3878
Number of observations	490	490	323	323
Number of countries	98	98	74	74
Test of first order serial correlation	-2.925 [0.003]	-2.603 [0.009]	-3.944 [0.000]	-3.569 [0.000]
Test of second order serial correlation	-0.470 [0.639]	-0.697 [0.486]	0.564 [0.573]	0.572 [0.567]
Test of over-identification		[0.093]		[0.490]

Notes:

- Small Instrument Set: lags 2, 3 and 4 of investment and GDP per worker.
- Augmented Instrument Set: small instrument set plus lags 2, 3 and 4 of trade, log of inflation and governments spending.
- IV1 corresponds to the one-step estimator with robust standard errors in DPD, and GMM corresponds to the two-step estimator with corrected standard errors (Windmeijer (2000) correction) in DPD.

Table 6: Annual Panel Regressions: Using Level Investment and Backward sum

	Within	Within (no cont. info)	IV
$\Delta\tilde{y}_{it-1}$	0.0221 (0.76)	0.0470 (1.67)	-0.0575 (-1.13)
$\Delta\tilde{y}_{it-2}$	-0.0705 (-2.65)	-0.0682 (-2.59)	-0.0497 (-1.40)
$\Delta\tilde{y}_{it-3}$	-0.0617 (-2.38)	-0.0518 (-2.01)	-0.0911 (-2.51)
$\tilde{y}_{it-4}$	-0.0519 (-7.10)	-0.0473 (-6.34)	-0.0752 (-5.88)
$\Delta\tilde{x}_{it}$	0.0046 (5.33)		0.0256 (4.02)
$\Delta\tilde{x}_{it-1}$	0.0033 (5.33)	0.0025 (3.70)	0.0079 (4.96)
$\Delta\tilde{x}_{it-2}$	0.0020 (3.80)	0.0011 (2.07)	0.0063 (3.73)
$\Delta\tilde{x}_{it-3}$	0.0025 (4.37)	0.0015 (2.49)	0.0070 (4.06)
$\tilde{x}_{it-4}$	0.0022 (5.64)	0.0014 (3.09)	0.0066 (4.29)
$bs\,\tilde{x}_{it}$	0.0001 (4.03)		0.0002 (4.78)
$bs\,\tilde{x}_{it-1}$		0.0001 (2.51)	
Growth effect: $-\frac{\theta_1}{\pi_4}$	0.0014 [0.0000]	0.0010 [0.0014]	0.0023 [0.0000]
Level effect: $-\frac{\phi_4}{\pi_4}$	0.0429 [0.0000]	0.0287 [0.0003]	0.0873 [0.0000]
R^2	0.1785	0.1508	
Number of observations	3430	3430	2456
Number of countries	98	98	76
Test of first order serial correlation	-0.25 [0.8037]	-0.07 [0.9454]	0.94 [0.3465]
Test of second order serial correlation	-0.84 [0.4021]	-0.95 [0.3435]	0.11 [0.9108]
Test of over-identification			[0.176]

Notes:

- See table 3 for notes.
- $\tilde{x}_{it}$ in this specification refers to the level of investment as a share of GDP, instead of its log.

Table 7: Annual Panel Regressions: Using Log of Fixed Investment and Backward sum

	Within	Within (no cont. info)	IV
$\Delta\tilde{y}_{it-1}$	0.0877 [2.56]	0.1311 [3.62]	0.0096 [0.16]
$\Delta\tilde{y}_{it-2}$	-0.0551 [-1.90]	-0.0473 [-1.66]	-0.0404 [-0.94]
$\Delta\tilde{y}_{it-3}$	-0.0682 [-2.55]	-0.0552 [-1.99]	-0.0865 [-2.17]
$\tilde{y}_{it-4}$	-0.0416 [-4.46]	-0.0336 [-3.45]	-0.0508 [-4.45]
$\Delta\tilde{x}_{it}$	0.1210 [7.59]		0.3443 [4.27]
$\Delta\tilde{x}_{it-1}$	0.0274 [2.04]	0.0159 [1.13]	0.0443 [2.15]
$\Delta\tilde{x}_{it-2}$	0.0127 [1.25]	-0.0115 [-1.07]	0.0476 [2.23]
$\Delta\tilde{x}_{it-3}$	0.0336 [3.12]	0.0109 [1.03]	0.0852 [3.58]
$\tilde{x}_{it-4}$	0.0224 [3.46]	-0.0002 [-0.04]	0.0629 [3.58]
$bs\,\tilde{x}_{it}$	0.0017 [2.91]		0.0036 [3.73]
$bs\,\tilde{x}_{it-1}$		0.0004 [0.62]	
Growth effect: $-\frac{\theta_1}{\pi_4}$	0.0403 (0.0002)	0.0110 (0.4885)	0.0705 (0.0000)
Level effect: $-\frac{\phi_4}{\pi_4}$	0.5386 (0.0011)	-0.0071 (0.9710)	1.2377 (0.0013)
R^2	0.2854	0.2142	
Number of observations	2262	2262	1768
Number of countries	95	95	73
Test of first order serial correlation	-0.25 [0.8060]	-0.63 [0.5296]	-0.40 [0.6898]
Test of second order serial correlation	-0.85 [0.3930]	-1.28 [0.1993]	-0.96 [0.3372]
Test of over-identification			0.3923

Notes:

- See table 3 for notes.
- $\tilde{x}_{it}$ in this specification refers to the log of fixed investment as a share of GDP.

Table 8: Annual Panel Regressions with Population Growth Variables (Using Backward sum of Investment -- Selected Coefficients)

	Within	Within- no $\Delta\tilde{x}_{it}$	IV
$\tilde{y}_{it-4}$	-0.0583 (-8.10)	-0.0548 (-7.46)	-0.0750 (-6.65)
$\tilde{z}_{it-4}$	-0.0737 (-3.52)	-0.0791 (-3.80)	-0.0267 (-0.90)
$\tilde{x}_{it-4}$	0.0306 (5.94)	0.0202 (3.71)	0.0767 (4.61)
$bs\,\tilde{x}_{it}$	0.0011 (5.13)		0.0022 (4.94)
$bs\,\tilde{x}_{it-1}$		0.0008 (3.54)	
Growth effect: $-\dfrac{\theta_1}{\pi_4}$	0.0183 (0.0000)	0.0141 (0.0000)	0.0292 (0.0000)
Level effect: $-\dfrac{\phi_4}{\pi_4}$	0.5243 (0.0000)	0.3679 (0.0000)	1.0214 (0.0000)
R^2	0.1840	0.1592	
Number of observations	3332	3332	2456
Number of countries	98	98	76
Test of first order serial correlation	1.00 [0.3164]	1.03 [0.3043]	1.55 [0.1211]
Test of second order serial correlation	-2.11 [0.0349]	-2.16 [0.0310]	0.19 [0.8456]
Test of over-identification			[0.298]

- See table 3 for notes.
- $\tilde{z}$ is the log of the population growth rate (plus 0.05).

Table 9: Country Specific Regressions: Using Overall Mean of Investment

		OLS	OLS- No $\Delta\tilde{x}_{it}$	IV	IV With optimal lag
$\overline{\tilde{x}}_i t$	mean	0.0032 (3.31)	0.0036 (3.09)	0.0047 (4.39)	0.0020 (2.15)
	median	0.0023	0.0034	0.0032	0.0014
$\tilde{x}_{it-4}$	mean	0.0593 (5.85)	0.0277 (2.34)	0.0533 (3.90)	0.0581 (5.47)
	median	0.0529	0.0186	0.0535	0.0522
$\tilde{y}_{it-4}$	mean	-0.2503 (-13.43)	-0.2659 (-14.34)	-0.2668 (-13.03)	-0.2207 (-11.83)
	median	-0.2280	-0.2548	-0.2523	-0.2105
Growth effect: $-\frac{\gamma_{1i}}{\pi_{4i}}$	Positive	66	66	50	44
	Positive and signf. at 10%	39	42	29	31
	Positive and signf. at 5%	37	39	28	31
	mean	0.0120 (3.49)	0.0103 (2.57)	0.0129 (4.02)	0.0095 (2.32)
	median	0.0122	0.0132	0.0128	0.0077
Level effect: $-\frac{\phi_{4i}}{\pi_{4i}}$	Positive	74	65	51	57
	Positive and signf. at 10%	27	19	20	22
	Positive and signf. at 5%	19	17	15	20
	mean	0.2851 (6.35)	0.1001 (2.47)	0.2430 (4.00)	0.2858 (5.08)
	median	0.2104	0.0932	0.1893	0.2043
Second Stage Regression	$\overline{\tilde{x}}_i$	0.0037 (5.99)	0.0041 (5.36)	0.0051 (6.52)	0.0036 (4.37)
Second Stage Regression (with additional controls)	$\overline{\tilde{x}}_i$	0.0044 (4.02)	0.0058 (4.09)	0.0064 (5.45)	0.0035 (2.27)

Notes:

- The number of countries is 98 for the first two specifications, and 76 for IV.
- Robust means are reported. t-ratios are in parentheses.
- In the second stage regression the dependent variable is the coefficient estimated on the time trend for each country, and the reported figures are the coefficient of the regressor, $\overline{\tilde{x}}_i$, and the corresponding t-ratio.
- The additional controls in the second stage regression are the averages log of annual GDP inflation plus one, log of trade (sum of exports and imports) as a percentage of GDP, log of government spending as a percentage of GDP, secondary schooling, and initial income.
- In the last column optimal lag length for each country has been chosen by maximizing the Akaike Criterion. Therefore the row headings require an appropriate reinterpretation.

Table 10: Country Specific Regressions: Using Backward sum of Investment

		OLS	OLS- (no cont. info)	IV	OLS With optimal lag
$bs\tilde{x}_{it}$	mean	0.0041 (4.16)	0.0031 (2.65)	0.0059 (4.50)	0.0038 (3.98)
$(bs\tilde{x}_{it-1})$	median	0.0027	0.0029	0.0045	0.0028
$\tilde{x}_{it-4}$	mean	0.0552 (4.94)	0.0273 (2.33)	0.0267 (1.93)	0.0566 (5.46)
	median	0.0515	0.0243	0.0243	0.0466
$\tilde{y}_{it-4}$	mean	-0.2553 (-11.81)	-0.2565 (-11.90)	-0.2902 (-10.91)	-0.2264 (-12.21)
	median	-0.2235	-0.2466	-0.2588	-0.2094
Growth effect: $-\frac{\theta_{1i}}{\pi_{4i}}$	Positive	67	68	56	72
	Positive and signf. at 10%	46	41	41	41
	Positive and signf. at 5%	43	38	40	30
	mean	0.0165 (4.83)	0.0161 (4.04)	0.0206 (5.57)	0.0184 (5.07)
	median	0.0144	0.0134	0.0152	0.0171
Level effect: $-\frac{\phi_{4i}}{\pi_{4i}}$	Positive	72	68	47	73
	Positive and signf. at 10%	33	19	22	25
	Positive and signf. at 5%	27	16	20	17
	mean	0.2489 (5.43)	0.0935 (2.04)	0.1578 (3.13)	0.2840 (6.12)
	median	0.1982	0.1173	0.1120	0.2124

Notes:

- The number of countries is 98 for the first two specifications, and 76 for IV.
- Robust means are reported. t-ratios are in parentheses.
- In the last column optimal lag length for each country has been chosen by maximizing the Akaike Criterion. Therefore the row headings require an appropriate reinterpretation.

Table 11: Country Specific Regressions: Model in Differences

		OLS	OLS- (no cont. info)	IV	IV With optimal lag
$\tilde{x}_{it}$ ($\tilde{x}_{it-1}$ in the second column)	Mean	0.0212 (4.38)	-0.0005 (-0.09)	0.0080 (1.24)	0.0080 (1.41)
$\beta_0+\beta_1+\beta_2+\beta_3+\beta_4$ ($\beta_1+\beta_2+\beta_3+\beta_4$ in the second column)	Mean	0.1432 (5.80)	0.0347 (1.63)	0.1314 (4.76)	0.0876 (3.70)
	Median	0.1231	0.0396	0.0838	0.0660
$\alpha_1+\alpha_2+\alpha_3+\alpha_4-1$	mean	-1.0319 (-23.94)	-1.0077 (-17.68)	-0.7235 (-11.56)	-0.5329 (-10.61)
	median	-1.0656	-1.0562	-0.7519	-0.5400
Growth effect: $\dfrac{\theta_1}{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}$	Positive	68	45	47	47
	Positive and signf. at 10%	24	13	8	8
	Positive and signf. at 5%	19	12	6	8
	mean	0.0212 (4.57)	0.0033 (0.60)	0.0203 (2.27)	0.0290 (2.51)
	median	0.0207	-0.0021	0.0203	0.0269
Level effect: $\dfrac{\beta_0+\beta_1+\beta_2+\beta_3+\beta_4}{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}$	Positive	68	59	54	53
	Positive and signf. at 10%	26	13	14	13
	Positive and signf. at 5%	16	6	7	10
	mean	0.1398 (5.89)	0.0553 (2.56)	0.1785 (5.07)	0.1933 (4.67)
	median	0.1229	0.0344	0.1437	0.1913

Notes:

- The number of countries is 98 for the first two specifications, and 76 for IV.
- In the IV regression $\tilde{x}_{it}$, $\Delta\tilde{x}_{it}$, $\Delta\tilde{x}_{it-1}$ and $\Delta\tilde{y}_{it-1}$ are instrumented. The instrument set is lags 5 and 6 of $\tilde{x}_{it}$ and $\tilde{y}_{it}$, lags 2 and 3 of log of annual GDP 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.
- Robust means are reported. t-ratios are in parentheses.
- In the last column optimal lag length for each country has been chosen by maximizing the Akaike Criterion. Therefore the row headings require an appropriate reinterpretation.

DATA APPENDIX

LIST OF COUNTRIES

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

Note: Countries that are marked with stars are excluded from the 76-country panel.

DESCRIPTIVE STATISTICS

		Growth rate of output per worker	Investment share
FULL SAMPLE	Mean	0.0167	15.24
	Standard Deviation	0.0613	9.24
	First Quartile	-0.0087	7.93
	Second Quartile	0.0199	13.47
	Third Quartile	0.0478	21.61
	Min	-0.3990	0.50
	Max	0.4421	52.88
	Number of observations	3724	3822

- Data Source: Penn World Table 6.0
- The investment share is measured as total investment as a percentage of GDP.
- The sample covers the 1960-98 period.