CDDRL

WORKING PAPERS

*A Monte Carlo Study of Growth Regressions**

William R. Hauk, Jr.
Romain Wacziarg
NBER

Center on Democracy, Development, and the Rule of Law
Stanford Institute for International Studies

Number 6
12 August 2004

Center on Democracy, Development,
and the Rule of Law
Stanford Institute for International Studies
Stanford University
Encina Hall
Stanford, CA 94305
Phone: 650-724-7197
Fax: 650-724-2996
http://cddrl.stanford.edu/

About the Center on Democracy, Development and the Rule of Law (CDDRL)

CDDRL was founded by a generous grant from the Bill and Flora Hewlett Foundation in October in 2002 as part of the Stanford Institute for International Studies at Stanford University. The Center supports analytic studies, policy relevant research, training and outreach activities to assist developing countries in the design and implementation of policies to foster growth, democracy, and the rule of law.

About the Author

William R. Hauk, Jr. research interests include International Political Economy, International Trade, Formal Political Theory, And Economic Growth and Development. His current research *"Projects include Protection with Mary Sellers: An Application to Legislatures with Malapportionment"*, *"The Effects of Electoral Regime Upon Trade Policy"* and *"The Estimation of Home State Effects in U.S. Presidential Elections"*.

Romaim Wacziarg is Associate Professor of Economics at Stanford University Graduate School of Business. His research focuses on economic integration, trade and economic growth, political economy of trade, economic aspects of political separatism. His publications include *Trade Liberalization and Intersectoral Labor Movements* (forthcoming, *Journal of International Economics*, 2004); *Fractionalization, Journal of Economic Growth*, June 2003; *Stages of Diversification*, *American Economic Review*, March 2003; *Measuring the Dynamic Gains from Trade, World Bank Economic Review*, October 2001

NBER

A Monte Carlo Study of Growth Regressions

William R. Hauk, Jr.
Stanford University

Romain Wacziarg*
Stanford University
and NBER

May 2004

Abstract

Using Monte Carlo simulations, this paper evaluates the bias properties of common estimators used in growth regressions derived from the Solow model. We explicitly allow for measurement error in the right-hand side variables, as well as country-specific effects that are correlated with the regressors. Our results suggest that using an OLS estimator applied to a single cross-section of variables averaged over time (the between estimator) performs best in terms of the extent of bias on each of the estimated coefficients. The fixed-effects estimator and the Arellano-Bond estimator greatly overstate the speed of convergence under a wide variety of assumptions concerning the type and extent of measurement error, while between understates it somewhat. Finally, fixed effects and Arellano-Bond bias towards zero the slope estimates on the human and physical capital accumulation variables.

---

*Hauk: Stanford Graduate School of Business, 518 Memorial Way, Stanford CA 94305, bill.hauk@stanford.edu, tel: (650) 725 4002. Wacziarg (corresponding author): Stanford Graduate School of Business, 518 Memorial Way, Stanford CA 94305, wacziarg@gsb.stanford.edu, tel: (650) 723 6069. We thank Nazrul Islam, Norman Loayza, David McKenzie and seminar participants at Stanford University, UC Davis and Duke University for useful comments. The data and programs used in this paper are available upon request, and all errors are ours.

1 Introduction

In the last decade, spurred by the early work of Baumol [1986] and Barro [1991], growth regressions have become an industry. There exists no good alternative for addressing the fundamental question of what accounts for vast observed differences in per capita income across countries. Detailed case studies, while they help identify hypotheses for further testing and theorizing, lack the generality of large sample studies. On the other hand, the detractors of growth regressions have stressed their numerous drawbacks. These include an often excessive distance between measured variables and the theoretical concepts they are meant to capture; poor grounding of estimated functional forms in economic theory, and in particular the prevalence of reduced form relationships from which structural parameters cannot be identified; unjustified claims of causality in explanations of growth; a small number of available observations; and the prevalence of prior-driven data-mining. These are but a few in a growing list, resulting in numerous methodological debates on the proper way to run growth regressions. Many of these debates are as yet unresolved, so research evaluating the effectiveness of current methodologies and suggesting improvements to cross-country growth empirics appears necessary.

This paper is such a study. We use simulation methods to evaluate the bias properties of several estimators commonly used in the empirical growth literature.

The main contribution of our approach is to consider explicitly the impact of measurement error on estimates of the determinants of growth. Measurement error is likely to be a central problem in cross-country growth empirics. Nonetheless, this issue has received little attention in the literature.[1] In the absence of measurement error and other sources of endogeneity, a fixed-effects estimator unambiguously dominates estimators that use any between-country variation, when omitted variables such as the initial level of technology are correlated with included right-hand side variables. In the presence of measurement error, however, fixed-effects estimators will tend to exacerbate measurement error bias when the right-hand side variables are more time persistent than the errors in measurement. The issue then is whether the gains from reducing omitted variables bias are offset by an increase in measurement error bias under fixed-effects. We lack any guidance

[1] A notable exception is Barro [1997], chapter 1, page 36, who briefly discusses the possible consequences of measurement error for fixed-effects estimates of the rate of convergence. Consistent with our simulation results, he argues that these estimates of the speed of convergence are likely to be too high. See Griliches and Hausman [1986] for a related point, though not in the specific context of growth regressions.

from econometric theory to evaluate the resulting net bias in the multivariate context of growth regressions, making Monte Carlo simulations necessary to address this issue.

A related contribution of this paper is to help resolve a long-running methodological debate in growth empirics: whether the appropriate way to control for time invariant cross-country heterogeneity in the level of technology is to use a fixed-effects estimator, thereby identifying parameters solely on the basis of within-country variation, or to retain at least some between-country variation in the data and directly include as right-hand side variables available proxies explaining technological differences.[2] Advancing this debate has profound implications for our understanding of the growth process. Results obtained using fixed-effects differ markedly in two main ways from those obtained by attempting to include additional correlates of growth.

First, fixed-effects lead to estimates of the speed of conditional convergence that are much higher than the conventional 2% obtained in cross sectional studies.[3] For a given history of income shocks, a fast speed of convergence indicates that most countries at any point in time are relatively close to their steady-states, so that incomes can only rise by improving the determinants of steady-state income. In contrast, a slow speed of convergence suggests there is still a lot of catching up to do, so that the force of neoclassical convergence alone can be expected to raise per capita incomes of less developed countries over time. Second, the fixed-effects estimator tends to reduce the magnitudes of the estimated coefficients on right-hand variables compared to cross-sectional alternatives, so that it becomes harder to obtain statistically significant estimates on the determinants of steady-state income: given improvements in steady-state determinants will yield smaller steady-state income gains.[4]

To evaluate methodologies for growth empirics, we start with the Solow [1956] growth model,

[2] The first approach is associated with the work of Knight, Loayza and Villanueva [1993], Islam [1995] and Caselli, Esquivel and Lefort [1996] among others. Since the mid-1990s, the use of dynamic panel data estimators in growth empirics has become prevalent. The second approach can be likened to a "kitchen-sink" method, in which the unaccounted variation in economic growth was attributed to additionnal right-hand side variables, to capture institutions, policies and economic structures. See Wacziarg [2002] for a broad discussion of fixed-effects versus the kitchen-sink approach.

[3] Barro and Sala-i-Martin [1995], chapters 11 and 12.

[4] For example, human capital variables that are highly significant in cross-sectional estimates become insignificantly different from zero in panel-fixed effects applications, and sometimes reverse signs. See Islam [1995] and Benhabib and Spiegel [1994].

in its human capital-augmented version proposed by Mankiw, Romer and Weil [1992].[5] The Solow model is arguably the only solid theoretical foundation for the specific functional form generally estimated by practitioners, which involves regressing growth rates on the log initial income and a set of steady-state income determinants. Common specifications of growth regressions can be directly derived from this model, and reasonable values for the exogenous parameters of the model can be postulated. Using such values and modelling explicitly the dynamic nature of the Solow growth specification, we generate simulated data with moments resembling those of the empirical data, and perform Monte Carlo simulations to evaluate the performance of several commonly used estimators: fixed-effects, random effects, between (OLS on country means), and the Arellano-Bond estimator first introduced to growth empirics by Caselli, Esquivel and Lefort [1996].[6]

Our results suggest that using a least-squares estimator applied to a single cross-section of variables averaged over time (the between estimator) performs best to estimate the speed of conditional income convergence, though it tends to underestimate it. The fixed-effects estimator, as well as the Arellano-Bond estimator, greatly overstate the speed of convergence under a variety of assumptions concerning the type and extent of measurement error. The random effects estimator also tends to overstate the speed of convergence, though much less drastically than fixed-effects. Finally, fixed-effects seriously biases toward zero the slope estimates on the determinants of the steady-state level of income (the accumulation and depreciation variables of the Solow model), in particular on the human capital accumulation rate. In contrast, random effects and between tend to overstate them (bias them away from zero).

Thus, our simulations are able to replicate the basic pattern of coefficients found in the literature using alternative estimators applied to real data, and suggest measurement error has a lot to do

[5] A recently proposed alternative to growth regressions has been "levels" regressions, aimed at accounting for variation in the level of income rather than in the growth rate of income. Salient examples include, in chronological order, Hall and Jones [1999], Frankel and Romer [1999] and Acemoglu, Johnson and Robinson [2001]. Arguably, these level specifications are more devoid of theoretical foundations than their growth counterparts, which in most cases can be traced back to some version of the neoclassical growth model. Because this lack of theoretical foundation makes simulation difficult for levels regressions, we focus on growth regression in this paper. These also remain vastly more prevalent in the literature on the determinants of economic development.

[6] In Section 4, we also evaluate the properties of three other estimators: 1) The Seemingly Unrelated Regression (SUR) estimator used for example in Barro and Sala-i-Martin [1995]. 2) The specific cross-sectional estimator used in Mankiw, Romer and Weil [1992], a variant of the between estimator. 3) The recently developed dynamic panel "system" estimator of Arellano and Bover (1995) and Blundell and Bond (1998).

with these differences. The punchline of our results is that the use of dynamic panel data methods leads to unreliable estimates when measurement error is present: it leads to misleading inferences on the speed of convergence, and to findings that common determinants of the steady-state income level are insignificantly different from zero when this is not the case. In this particular application, old-fashioned OLS on cross-sectional averages performs best.

Perhaps the main contribution of this paper is methodological, and carries broader implications: the type of Monte Carlo exercise we present here should be a systematic rite of passage for studies presenting new estimation methodologies in any field of empirical economics. Estimators that may seem attractive to address a specific econometric problem need to be evaluated in a setting where several sources of bias may coexist. When a potential for omitted variables bias coexists with measurement error, a cure for the first problem can be worse than the disease, as it may exacerbate the second.

This paper is structured as follows: Section 2 briefly discusses theoretical considerations related to the methodology of growth regressions. Section 3 presents our basic simulation methodology and results, contrasting OLS, fixed-effects, random effects and Arellano-Bond GMM estimators. Section 4 discusses extensions of our simulations to country-specific measurement error, regressor-specific measurement error, autocorrelated measurement error and additional estimators. Section 5 concludes by presenting new estimates of the speed of income convergence and of the effect of steady-state determinants using real data, and discusses them in light of our simulation results.

2 Theoretical Framework

2.1 Growth Regressions and the Solow Model

Mankiw, Romer and Weil [1992, henceforth MRW] and Islam [1995] have shown that the Solow growth model can be transformed in a way that allows its estimation through a simple application of linear regression techniques. This section reviews this well-known derivation.

The Solow growth model augmented to include human capital accumulation starts with a simple neoclassical production function:

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} (A(t)L(t))^{1-\alpha-\beta} \tag{1}$$

where Y is output, K is physical capital, H is human capital, L is labor and A is a labor-augmenting technology parameter. L and A are assumed to grow at exogenously determined rates n and g such that $L(t) = L(0)e^{nt}$ and $A(t) = A(0)e^{gt}$. MRW and Islam also assume that n and g do not vary between countries.

Assume that in a given period a constant fraction of output is saved and devoted to investment in physical and human capital. If we define $\hat{y} = Y/AL$, $k = K/AL$ and $h = H/AL$ to be the units per effective unit of labor, changes in physical and human capital can be represented as:

$$\dot{k}(t) = s_k \hat{y}(t) - (n+g+\delta)k(t) \tag{2}$$

$$\dot{h}(t) = s_h \hat{y}(t) - (n+g+\delta)h(t) \tag{3}$$

where s_h and s_k are the proportions of output devoted to investment in human and physical capital, respectively, and δ is the depreciation rate of both human and physical capital (which is also assumed in the literature not to vary between countries).

The dynamics in equations (2) and (3) imply that the economy converges to steady-state levels of physical and human capital k^* and h^*, derived by setting $\dot{k} = 0$ and $\dot{h} = 0$. Substituting these values back into equation (1) and taking logs, we get:

$$\begin{aligned} \log y(t) \equiv \log \frac{Y(t)}{L(t)} &= \log A(0) + gt - \frac{\alpha+\beta}{1-\alpha-\beta}\log(n+g+\delta) \\ &+ \frac{\alpha}{1-\alpha-\beta}\log s_k + \frac{\beta}{1-\alpha-\beta}\log s_h \end{aligned} \tag{4}$$

Equation (4) describes an economy in its steady-state. If one is willing to assume that countries are at their steady-states, this equation can be turned into an econometric specification for a "levels" regression - but the assumption is unlikely to hold.[7]

To derive a growth regression explicitly, we can approximate the model around the steady-state y^*:

$$\frac{d\log \hat{y}(t)}{dt} = \lambda \left[\log \hat{y}^* - \log \hat{y}(t)\right] \tag{5}$$

where $\lambda = (n+g+\delta)(1-\alpha-\beta)$ is the rate of convergence. That is, λ is the percentage of the gap between a country's steady-state and its current income that will be closed in one period, in the absence of any other shocks.[8] A convergence rate of λ would imply that, given two points in time t_1 and t_2, we can measure end-of-period output as:

$$\log \hat{y}(t_2) = \left(1 - e^{-\lambda\tau}\right)\log \hat{y}^* + e^{-\lambda\tau}\log \hat{y}(t_1) \tag{6}$$

[7] However, for a recent paper taking such an assumption seriously, see Bernanke and Gürkaynak [2001]. For a critique, see Caselli [2001].

[8] For example, if $\lambda = 0.10$, the half-life of convergence to the steady state is $\log(2)/0.10 = 0.69/10 = 6.9$ years.

where $\tau = t_2 - t_1$. The higher the convergence rate, the closer we should expect the economy to be to its steady-state at a given point in time, all else equal.

Since $\log \hat{y}(t) = \log y(t) - \log A(0) - gt$, we can substitute equation (4) into $\log \hat{y}^*$ in equation (6):

$$\begin{aligned}\log y(t_2) &= \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha-\beta}\log s_k + \left(1-e^{-\lambda\tau}\right)\frac{\beta}{1-\alpha-\beta}\log s_h \\ &-\left(1-e^{-\lambda\tau}\right)\frac{\alpha+\beta}{1-\alpha-\beta}\log(n+g+\delta) \\ &+e^{-\lambda\tau}\log y(t_1) + \left(1-e^{-\lambda\tau}\right)\log A(0) + g(t_2 - e^{-\lambda\tau}t_1) \end{aligned} \tag{7}$$

This equation is the basis for estimating growth regressions in discrete time, as derived from a continuous time Solow growth model. Adding an error term ν_{it} with mean zero conditional on all the right-hand side variables, capturing inherent randomness in $\log y_{it}$, we can rewrite equation (7) as a fixed-effects panel data regression of the form:

$$\begin{aligned}\log y_{it} &= \gamma_0 + \gamma_1 \log s_{k,it-\tau} + \gamma_2 \log s_{h,it-\tau} + \gamma_3 \log(n+g+\delta)_{it-\tau} \\ &+\gamma_4 \log y_{it-\tau} + \mu_i + \eta_t + \nu_{it}\end{aligned} \tag{8}$$

where t denotes the end of a time period of duration τ and $t-\tau$ denotes the beginning of that period.[9] The reduced form parameters and error terms are defined as:

$\gamma_1 = \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha-\beta}$

$\gamma_2 = \left(1-e^{-\lambda\tau}\right)\frac{\beta}{1-\alpha-\beta}$

$\gamma_3 = -\left(1-e^{-\lambda\tau}\right)\frac{\alpha+\beta}{1-\alpha-\beta}$

$\gamma_4 = e^{-\lambda\tau}$

$\gamma_0 + \mu_i = \left(1-e^{-\lambda\tau}\right)\log A_i(0)$ (an intercept plus a country effect)

$\eta_t = g(t - e^{-\lambda\tau}(t-\tau))$ (a time specific effect)

ν_{it} (a zero-mean error term, orthogonal to the regressors)

Equation (8) is the functional form used as the data-generating process for the remainder of this paper. In what follows we will sometimes find it useful to rewrite equation (8) as:

$$\log y_{it} = \gamma' x_{it} + \mu_i + \eta_t + \nu_{it} \tag{9}$$

[9] In our actual empirical application of equation (8) the determinants of the steady-state level of income $\log s_k$, $\log s_h$, and $\log(n+g+\delta)$ are entered as averages over the period $t-\tau$ to t, rather than their beginning of period values. This is consistent with the common practice of growth regressions, as in MRW and Islam, where introducing right-hand side variables as period averages is thought to limit the extent of classical measurement error. Note that theory gives us no guidance on this choice, as it considers these regressors to be exogenous and time invariant.

where we define $x'_{it} = [1, \log s_{k,it-\tau}, \log s_{h,it-\tau}, \log(n+g+\delta)_{it-\tau}, \log y_{it-\tau}]$ and $\gamma' = [\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4]$, with the dimension of these vectors denoted $Q = 5$.

2.2 Country-level heterogeneity

The $A(0)$ term constitutes a stumbling block for growth regressions. This term captures the initial level of technology, which can be proxied for using variables such as resource endowments, climate, institutions, government type, and so on. These variables vary widely across countries, so that we can index $A(0)$ by i. Hence, we define $\gamma_0 + \mu_i \equiv \left(1 - e^{-\lambda\tau}\right) \log A_i(0)$, where γ_0 is a constant capturing the average level of the initial technology term across countries and μ_i is a zero-mean country-specific effect. There have been three basic ways of dealing with country-level heterogeneity (i.e. the $\mu_i = \left(1 - e^{-\lambda\tau}\right) \log A_i(0) - \gamma_0$ term) to estimate growth regressions. These methods are associated with the contributions of MRW [1992], Islam [1995] and Caselli, Esquivel and Lefort [1996], respectively.

MRW [1992] and Islam [1995] MRW assumed the μ_i term had mean zero conditional on other right-hand side variables. As a result, they ran simple OLS regressions of growth on the log of initial income and time-averaged steady-state determinants (i.e. a single cross-section), including an intercept in the regression to account for γ_0. A major drawback of this approach is that it causes the estimated coefficients to be biased if the orthogonality assumption is untrue - which is likely the case in practice.

Assume that we want to estimate the parameters of the panel data regression model of equation (9):

$$\log y_{it} = \gamma' x_{it} + \mu_i + \varepsilon_{it} \tag{10}$$

where μ_i is *not* assumed to be independent from x_{it} and $\varepsilon_{it} = \eta_t + v_{it}$ is a well-behaved random-noise term.

Define X as the $NT \times Q$ matrix that stacks x_{it} over time periods $t = 1...T$ and countries $i = 1...N$, and μ as the $NT \times 1$ vector that similarly stacks μ_i. Let Σ_{xx} be the $Q \times Q$ covariance matrix of columns of X, and $\Sigma_{\mu x}$ the $Q \times 1$ vector of the covariances of μ with the columns of X. If a pooled OLS (POLS) regression is run on the stacked data, standard omitted variables bias will result:[10]

$$\text{plim}\,\hat{\gamma}^{POLS} = \gamma + \Sigma_{xx}^{-1}\Sigma_{\mu x} \tag{11}$$

Equation (11) implies that slope estimates will be biased if the country-specific effects are correlated with the regressors. In our particular application, the Solow model states that the omitted term captures some positive multiple of the initial level of technology $\log A_i(0)$. The observed data in X are initial income, rates of human and physical capital accumulation and population growth. While the Solow model strictly speaking is silent about the correlation between the $\log A_i(0)$ term and the right-hand side variables, there is a strong presumption that these four variables will be potentially highly correlated with $\log A_i(0)$. Hence, estimated coefficients could be significantly biased when the correlation between $\log A_i(0)$ and the steady-state determinants is ignored.

This is the point originally made by Islam [1995] in advocating the use of fixed-effects estimation instead of OLS on country means. Islam averaged annual data from the available sample of countries across time, into 5-year periods. μ_i is a time-invariant effect if λ is treated as a constant and τ does not vary with time (i.e. the panel involves equally spaced periods). Hence, it can be represented as a country-fixed effect in a panel regression, while the term $g(t - e^{-\lambda\tau}(t-\tau))$ is a time effect. Using a fixed-effects estimator, Islam found the estimated rate of convergence to be much higher than had been estimated by MRW, and the effect of some right-hand side variables smaller (particularly human capital).

Caselli, Esquivel and Lefort [1996]. Going one step further, Caselli, Esquivel and Lefort [1996, henceforth CEL] pointed out the *necessary* correlation between the country-specific effect μ_i and the log of initial income resulting from the dynamic nature of the specification. We can rewrite equation (9) as:

$$\log y_{it} = \gamma_0 + \gamma_s' w_{it} + \gamma_4 \log y_{it-\tau} + \mu_i + \eta_t + v_{it} \tag{12}$$

where $\gamma_s' = [\gamma_1, \gamma_2, \gamma_3]$ and $w_{it} = \left[\log s_{k,it-\tau}, \log s_{h,it-\tau}, \log{(n+g+\delta)}_{it-\tau}\right]$. Lagging equation (12) by one period, it is evident that $\log y_{it-\tau}$ contains μ_i. Thus, $\log y_{it-\tau}$ *must* be correlated with the error term unless μ_i is appropriately accounted for.

[10] This bias is also known as heterogeneity bias. We will use the terms heterogeneity bias and omitted variables bias interchangeably.

CEL transformed all variables used in the regressions into deviations from period means (thereby removing the need for a time-specific intercept η_t) and then eliminated the country-specific effects μ_i by taking first-differences. Their transformed regression is:

$$\begin{aligned} \widetilde{\log y}_{i,t} - \widetilde{\log y}_{i,t-\tau} &= \gamma_s' \left(\widetilde{w}_{i,t} - \widetilde{w}_{i,t-\tau}\right) \\ &\quad + \gamma_4 \left(\widetilde{\log y}_{i,t-\tau} - \widetilde{\log y}_{i,t-2\tau}\right) + \left(\widetilde{\nu}_{i,t} - \widetilde{\nu}_{i,t-\tau}\right) \end{aligned} \tag{13}$$

where "˜" denotes deviations of variables from period means. The problem with this specification is that, while μ_i and η_t have been differenced away, the term $\widetilde{\log y}_{i,t-\tau}$ is clearly not independent from $\widetilde{\nu}_{i,t-\tau}$.[11] Hence, some sort of instrumental variables approach is required. CEL proposed a GMM estimator similar to the Arellano and Bond [1991] estimator (henceforth, AB) to deal with the problems of heterogeneity bias and endogeneity of the differenced lagged income term in equation (13).[12] Their estimator results in a consistent estimator for the unknown parameters under the moment condition $E\left[\widetilde{\nu}_{i,t}\widetilde{\nu}_{i,t-\tau}\right] = 0$. They instrument for the differenced independent variables using all predetermined independent variables (in levels). For example, their panel consists of four time-periods, and their variables $\left(\widetilde{\log y}_{i,1} - \widetilde{\log y}_{i,0}\right)$ and $\left(\widetilde{w}_{i,1} - \widetilde{w}_{i,0}\right)$ for period 1 are instrumented using $y_{i,0}$ and $w_{i,0}$. Then, $\left(\widetilde{\log y}_{i,2} - \widetilde{\log y}_{i,1}\right)$ and $\left(\widetilde{w}_{i,2} - \widetilde{w}_{i,1}\right)$ are instrumented using $y_{i,0}$, $y_{i,1}$, $w_{i,0}$ and $w_{i,1}$ and so on. The exclusion of the current period $\widetilde{w}_{i,t}$ term from the list of instruments is meant to deal with the possible endogeneity of the variables in w_{it}, a valid procedure under the assumption that all of the instrumental variables are predetermined.[13] Consistent estimates will result even in the presence of measurement error on the right-hand side variables, as long as the instruments are not correlated with the errors in measurement, for example if these are white noise (as in the classical case).

Estimators in the class of the AB estimator may have an advantage since they address several problems with the cross-sectional approach to growth regressions. However, they require losing at least two periods of data in order to implement the IV procedure, which could affect estimates in an unknown direction when T is small. Another recently identified drawback of AB is the problem

[11] We will refer to this source of bias as endogeneity bias, to differentiate it from heterogeneity bias.

[12] Estimators in this class have been widely used in the empirical growth literature. See for instance Easterly, Loayza and Montiel [1997], among many others.

[13] The Solow model, however, treats the w_{it} variables as exogenous, so the endogeneity of w_{it} should not be a problem within the strict confines of the model. Moreover, whether these variables are in fact predetermined is subject to debate.

of weak instruments, arising in small samples: the first stage relationship between differenced independent variables and lagged level variables may be weak, biasing GMM estimates towards their fixed-effects counterparts. This is likely to be a problem especially when the convergence parameter γ_4 is large and when the variance of the country-specific effect μ_i is large relative to the variance of $\log y_{it}$. There is now a sizable literature on weak instruments. For example, Stock, Wright and Yogo [2002] show that in a two-stage least squares (2SLS) context, if the instruments in the first stage do not help *at all* in predicting the endogenous regressors, 2SLS reduces exactly to OLS.[14] Staiger and Stock [1997] provide a rule of thumb for determining whether instruments are weak in the linear IV case with one endogenous regressor: if the first stage F-test for the joint significance of the instruments is smaller than 10, then the instruments are declared to be weak. For the case of multiple endogenous regressors, Stock and Yogo [2003] propose using the Cragg-Donald [1993] test statistic for underidentification, but using appropriately corrected critical values in order to use the statistic for a test of the null hypothesis of weak instruments.

We apply this test in Section 3 in order to assess whether the AB estimates of the Solow model are likely to be subject to the weak instruments problem, finding that they are.[15] In section 4, we also examine the properties of the AB estimator when N is increased beyond the number of countries usually available in growth regressions, and we simulate the behavior of more recently developed estimators from Arellano and Bover [1995] and Blundell and Bond [1998], developed specifically to address the small sample drawbacks of the AB estimator.

The three estimators discussed here have their own strengths and weaknesses. However, the previous literature has given little attention to what happens in the presence of measurement error.

[14] Stock, Wright and Yogo [2002] define the concentration parameter μ^2 as a measure of goodness of fit of the first stage regression, or equivalently a measure of the strength of the instruments. They state that: "When $\mu^2 = 0$ (...), the instruments are not just weak, but irrelevant. In this case, the mean of the $2SLS$ estimator is the probability limit of the ordinary least squares (OLS) estimator, $\text{plim}(\hat{\beta}^{OLS})$. (...) When the instruments are relevant but weak, the $2SLS$ estimator is biased toward $\text{plim}(\hat{\beta}^{OLS})$." (p. 519).

[15] To evaluate the extent of small sample bias in dynamic panel estimators used to estimate growth regressions, Islam [2000] conducted a Monte Carlo study. He concluded that panel-IV estimators such as the AB estimator suffer from serious small sample/weak instruments problem, and that the fixed-effects (least-squares dummy variables) and the minimum distance estimators had the best small sample performance. He did not, however, consider the issue of measurement error like we do. As we argue below, measurement error is a first order problem, greatly exacerbating small sample bias in estimators of the AB class.

2.3 Measurement Error and Heterogeneity Bias

2.3.1 Measurement error without heterogeneity bias

In this subsection, we examine what happens once we allow for measurement error in the independent variables. In order to deal with various sources of bias one by one, we ignore for the moment the omitted variables and endogeneity problems identified above, and will return to them later. In other words, we assume $\log y_{it} = \gamma' x_{it} + \varepsilon_{it}$ and $E(\varepsilon_{it}|x_{it}) = 0$ where $\varepsilon_{it} = \eta_t + v_{it}$. In a univariate setting, when measurement error is white noise, pooled OLS estimates exhibit attenuation bias. However, in the more general multivariate case, it is impossible to sign the effect of measurement error on the slope estimates.

Klepper and Leamer [1984] point out that in the presence of classical measurement error in a multivariate context, few substantive restrictions can be placed on the sign and magnitude of the resulting bias unless stringent assumptions are made. In our notation, they consider a regression model where a dependent variable, $\log y_{it}$, is drawn from a normal distribution with mean $\gamma' x_{it}$ (where x_{it} is a $Q \times 1$ column vector) and variance σ_ε^2 conditional on x_{it}. Thus:

$$\log y_{it} = \gamma' x_{it} + \varepsilon_{it} \tag{14}$$

with $E[\varepsilon_{it}|x_{it}] = 0$. They assume that x_{it} cannot be observed, but that we can observe $x_{it}^* = x_{it} + d_{it}$, where $E[d_{it}|x_{it}] = 0$.[16] Define:

$$\text{var}[d_{it}|x_{it}] = D = \text{diag}\left\{\sigma_{d_0}^2, \sigma_{d_1}^2, ..., \sigma_{d_{Q-1}}^2\right\} \tag{15}$$

In our application, we will set $\sigma_{d_0}^2 = 0$, i.e. we do not shock the intercept column in X. Defining $\bar{x} \equiv E(x_{it})$, $(\log y_{it}, x_{it}^{*\prime})$ has a multivariate normal distribution with moments:

$$E\left(\log y_{it}, x_{it}^{*\prime}\right) = (\gamma' \bar{x}, \bar{x}) \tag{16}$$

$$V\left(\log y_{it}, x_{it}^{*\prime}\right) = \begin{bmatrix} \sigma_\varepsilon^2 + \gamma' \Sigma_{xx} \gamma & \gamma' \Sigma_{xx} \\ \Sigma_{xx} \gamma & \Sigma_{xx} + D \end{bmatrix} \tag{17}$$

[16] In our simulations, we consider alternative specifications for the form of measurement error. Also, in this section we are abstracting from measurement error in the dependent variable $\log y_{it}$. Obviously, in the classical measurement error case, this would be equivalent to raising the variance of the error term ν_{it}, reducing the efficiency of the estimates without introducing bias. In our simulations, we will need to consider measurement error in $\log y_{it}$ explicitly because of the dynamic nature of the Solow specification. We are ignoring this dynamic characteristic for now.

Performing a pooled OLS regression of $\log y_{it}$ on x^*_{it}, we can show:

$$\text{plim}\left(\hat{\gamma}^{POLS}\right) = \left(\Sigma_{xx} + D\right)^{-1} \Sigma_{xx}\gamma \tag{18}$$

where X^* stacks the $x^{*\prime}_{it}$ vectors over time and countries and y does the same for $\log y_{it}$. Obviously, if $D = 0$, pooled OLS produces a consistent estimator. In the generic case, however, the estimator is inconsistent. We cannot say anything more about the sign and magnitude of the bias unless we can make assumptions about the correlation structure among the various independent variables and the covariance matrix of measurement error, i.e. if we can place restrictions on $(\Sigma_{xx} + D)$ and Σ_{xx}, based on some knowledge of the nature of measurement error. This is in general a tall order.

2.3.2 Measurement error with heterogeneity bias

Assume now that $E\left(\log y_{it}|x_{it}\right) = \gamma' x_{it} + \mu_i$ where μ_i is unobserved and not necessarily orthogonal to the variables in x_{it}.[17] Thus, the true model is $\log y_{it} = \gamma' x_{it} + \mu_i + \varepsilon_{it}$, where all the variables are still defined as above. In this case, estimating the model using OLS will involve two separate problems: 1) An omitted variables bias problem due to μ_i being potentially correlated with the right hand side variables in x_{it}. 2) A measurement error problem due to x_{it} being imperfectly observed. As argued above, it is already difficult to make statements about the sign and magnitude of the bias when only measurement error is present. Such statements become even more difficult when correlated country effects are also considered. To illustrate this formally, we can derive the probability limit of the pooled OLS estimator in the presence of both measurement error and heterogeneity bias.

The unconditional expectations of $\log y_{it}$, x^*_{it} and μ_i are:

$$E\left(\log y_{it}, x^*_{it}, \mu_i\right) = \left(\bar{x}'\gamma + \bar{\mu}, \bar{x}, \bar{\mu}\right) \tag{19}$$

where $\bar{\mu} = E(\mu_i)$ and $\bar{x} = E(x_{it}) = E(x^*_{it})$, and the variance of μ_i is denoted σ^2_μ. Then:

$$V\left(\log y_{it}, x^*_{it}, \mu_i\right) = \begin{bmatrix} \sigma^2_\varepsilon + \gamma'\Sigma_{xx}\gamma + 2\gamma'\Sigma_{\mu x} + \sigma^2_\mu & \gamma'\Sigma_{xx} + \Sigma'_{\mu x} & \gamma'\Sigma_{\mu x} + \sigma^2_\mu \\ \Sigma_{xx}\gamma + \Sigma_{\mu x} & \Sigma_{xx} + D & \Sigma_{\mu x} \\ \gamma'\Sigma_{\mu x} + \sigma^2_\mu & \Sigma'_{\mu x} & \sigma^2_\mu \end{bmatrix} \tag{20}$$

[17] In the rest of this section we continue to ignore the problem arising from using a fixed-effects estimator in the presence of a lagged dependent variable in x_{it} - i.e. the problem identified in the discussion following equation (13). We will take this issue into account in our simulations, which explictly model the dynamic nature of the empirical Solow model.

Suppose that we estimate γ using pooled OLS, with x^*_{it} as our observed regressor. Then the limiting value of the pooled OLS estimator for γ is:

$$\operatorname{plim} \hat{\gamma}^{POLS} = (\Sigma_{xx} + D)^{-1} \Sigma_{xx}\gamma + (\Sigma_{xx} + D)^{-1} \Sigma_{\mu x} \tag{21}$$

Obviously, this estimator is inconsistent for two reasons: the first is measurement error bias, and the second is heterogeneity bias. If $D = 0$, we would recover equation (11), showing that pooled OLS would not involve measurement error bias. On the other hand, if $\Sigma_{\mu x} = 0$, we would recover equation (18), showing that pooled OLS regression would not involve any heterogeneity bias. If neither of these two issues were a problem, pooled OLS would be a consistent estimator for γ.

In a context where both problems coexist, there may be a trade-off between reducing the extent of bias due to measurement error and reducing the bias attributable to heterogeneity. The common way to deal with heterogeneity, as explained above, is to estimate γ using the fixed-effects (FE) estimator $\hat{\gamma}^{FE}$. Appendix 1 derives the limiting value of $\hat{\gamma}^{FE}$ in the presence of measurement error, showing that it gets rid entirely of heterogeneity bias. It also derives the limiting value of the between (BE) estimator $\hat{\gamma}^{BE}$obtained by computing country means of the data over time and running OLS regressions on these country means:[18]

$$\operatorname{plim} \hat{\gamma}^{BE} = \left(\Sigma^B_{xx} + \frac{1}{T}D\right)^{-1} \Sigma^B_{xx}\gamma + \left(\Sigma^B_{xx} + \frac{1}{T}D\right)^{-1} \Sigma_{\mu x} \tag{22}$$

$$\operatorname{plim} \hat{\gamma}^{FE} = \left(\Sigma^W_{xx} + \frac{T-1}{T}D\right)^{-1} \Sigma^W_{xx}\gamma \tag{23}$$

where Σ^B_{xx} denotes the between-country variation in X, Σ^W_{xx} denotes the within-country variation (as defined in the Appendix 1), and $\Sigma_{xx} = \Sigma^W_{xx} + \Sigma^B_{xx}$.

It is difficult to make general statements about what happens to bias from measurement error under FE and BE estimation in the multivariate case. Measurement error bias on specific slope coefficients under FE may or may not be exacerbated. This is the main justification for a Monte Carlo approach to evaluating the properties of these estimators.[19] However, a few statements can

[18] Appendix 1 also derives the plim of the random effects (RE) estimator $\hat{\gamma}^{RE}$, which is a matrix-weighted average of the BE and FE estimators.

[19] In subsection 2.3.3 below, and in the appendix, we discuss conditions under which measurement error is exacerbated under FE in the specific case of univariate regression. We argue these conditions are likely to hold in our particular application, as in most. But only a Monte Carlo simulation can provide definitive answers.

be made to compare the properties of pooled OLS, BE and FE estimators using equations (21), (22) and (23):

1). As noted above, FE gets rid entirely of the heterogeneity bias while there is in general both measurement error and heterogeneity bias when using the BE and pooled OLS estimators.

2). The BE estimator tends to reduce the extent of measurement error bias compared to the other estimators due to averaging the imperfectly measured variables over time, which reduces the variance of the measurement error relative to the true signal. Moreover, the greater is T, the smaller the bias from measurement error.

3). Both pooled OLS and BE will involve smaller heterogeneity bias, the greater the extent of measurement error. In other words, there is an interaction between these two sources of bias. This is because all other things equal, measurement error reduces the correlation between the regressors and the country effects, and hence alleviates the omitted variables problem. For this reason, BE will on average involve larger heterogeneity bias compared to pooled OLS, holding constant D.

4) Comparing FE and BE, if Σ_{xx}^{W} relative to $\frac{T-1}{T}D$ is "smaller" in a matrix sense than Σ_{xx}^{B} is relative to $\frac{1}{T}D$, then the bias arising from measurement error will tend to be smaller under BE compared to FE. This is likely to hold if the within variation is small compared to the between variation (most of the variation in the panel arises from the cross-section rather than the time dimension - which is the case in growth applications), or if T is large.

To conclude, despite these general lessons, we can say little about the net biases to individual parameter estimates as they would result from each estimation method. Given the multivariate nature of growth regressions, only simulations can determine which estimator dominates in terms of bias under alternative assumptions about the covariance structure of the true data Σ_{xx}, the covariances between the true variables and the country-specific effects $\Sigma_{\mu x}$ and the covariance matrix of the measurement error D.

2.3.3 Autocorrelated Measurement Error: Univariate Example

Appendix 2 analyzes in detail a simple case illustrating the trade-offs identified above in a case where net biases can be signed: the case of $Q = 2$ (a single regressor x_{it} plus an intercept). This example is also useful to illustrate what happens when measurement error is autocorrelated.

The example shows clearly that, under FE estimation, eliminating heterogeneity bias may come at the cost of exacerbating measurement error bias. The greater the time persistence in x_{it}, the greater the extent to which measurement error bias is exacerbated, as the variance of the true signal gets differenced away relative to the variance of the error in measurement. In the context of growth regressions, where right-hand side variables are highly time persistent, this point is particularly central. We cannot say analytically whether this increase in measurement error bias is worth the elimination of heterogeneity bias unless we know the moments of the true underlying data and of the measurement error.

As argued in Section 2.2, a GMM procedure such as AB could in principle deal with both sources of bias if measurement error is white noise, since measurement error in the instruments is uncorrelated with measurement error in the regressors. However, introducing measurement error weakens the first stage relationship between predetermined regressors and the instruments, potentially making the weak instruments problem worse. Moreover, the validity of this procedure relies heavily on the assumption of non-autocorrelated errors in measurement.

Appendix 2 shows that when measurement error is autocorrelated, where we define $\rho_d = \text{corr}(d_{it}, d_{it-\tau})$, FE exacerbates measurement error bias compared to pooled OLS whenever $\rho_d < \rho_x$, where ρ_x is the autocorrelation coefficient in x_{it}. In this case, instrumenting for differenced x_{it} using its lagged levels values (as in the AB procedure) no longer gets rid of measurement error bias. In words, as long as $\rho_d > 0$, we cannot produce a consistent estimator of the desired parameters using the AB estimator. We consider the case of autocorrelated measurement error in Section 4.

2.4 Summary

Five factors can cause inconsistent estimates of γ in panel growth regressions. The first is an omitted-variables bias resulting from the possible correlation between country-specific effects and the regressors, affecting the consistency of pooled OLS, BE and RE estimates. The second is the endogeneity problem specific to dynamic panels, identified after equation (13), which will make FE and RE estimates inconsistent.[20] The third is classical measurement error on the independent variables, which affects the consistency of pooled OLS, BE, RE and FE estimator, though the bias tends to be exacerbated in the latter case and partly averaged away under BE. The fourth is possible autocorrelation in measurement errors, which results in inconsistency for all estimators we consider here including the AB estimator. The fifth is the weak instruments problem that can

[20] For the sake of space and because this source of bias is well-known, we have abstracted from it in the last subsection, but we will take it into account explicitly in our simulations.

cause bias in the AB estimator in small samples.

Each of the estimators under consideration involves a trade-off: pooled OLS suffers from heterogeneity bias but limits the incidence of measurement error bias relative to FE; the BE estimator reduces measurement error through time averaging of the regressors, but does not deal with heterogeneity bias; FE addresses the problem of heterogeneity bias, but tends to exacerbate the problem of measurement error. The AB estimator is inconsistent when instruments are weak or when measurement error is autocorrelated. As a result, we cannot say a priori which estimator produces the smaller total bias. Simulations are necessary to evaluate the properties of these estimators.

3 Monte Carlo Simulations

3.1 Simulation Methodology

Since it is impossible to derive analytical results about the extent and sign of omitted variables and measurement error biases in a multivariate context, we use Monte Carlo simulations to evaluate the bias properties of of FE, BE, RE and AB estimators.[21]

The starting point for our simulations is equation (8). The data-generating process for the true data (the data not measured with error) is:

$$\begin{aligned}
\log y_{it} &= \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha-\beta}\log s_{k,it-\tau}+\left(1-e^{-\lambda\tau}\right)\frac{\beta}{1-\alpha-\beta}\log s_{h,it-\tau} \\
&\quad -\left(1-e^{-\lambda t}\right)\frac{\alpha+\beta}{1-\alpha-\beta}\log\left(n+g+\delta\right)_{it-\tau}+e^{-\lambda t}\log y_{it-\tau} \\
&\quad +\left(1-e^{-\lambda\tau}\right)\log A_i(0)+g(t-e^{-\lambda\tau}(t-\tau))+\nu_{it} \qquad (24)
\end{aligned}$$

3.1.1 Simulated Data

Underlying data. We define a period by a five year interval of time ($\tau = 5$). Our underlying data spans 40 years, from 1960 to 2000, and our 8 five year periods are defined as 1960-1965, 1965-1970, ..., 1995-2000. In equation (24), the variables $\log s_{k,it-\tau}$, $\log s_{h,it-\tau}$, $\log(n+g+\delta)_{i,t-\tau}$ and $\log y_{it-\tau}$ are simulated data with moments resembling those of the corresponding observed variables. To obtain these moments, we captured $\log s_k$ using the log of investment rates as a share of real GDP from the Penn World Tables, version 6.1 [Heston, Summers and Aten, 2002 -

[21] Results for pooled OLS estimates are available upon request. The pooled OLS estimator is rarely used in cross-sectional growth regressions since it is less efficient than random effects.

henceforth, PWT6.1]. $\log s_h$ is the log of the secondary school gross enrollment ratio from Barro and Lee [2000] and n is the rate of population growth calculated from the PWT6.1 population series. In calculating $\log(n+g+\delta)$, we postulated (as above) that $g+\delta = 0.07$. Finally, $\log y_{it-\tau}$ is the log of per capita income in purchasing power parity from PWT6.1, measured at the beginning of the first time period (1960).[22]

We averaged the variables over relevant time periods and arrayed them in a $N \times (T(Q-1)+1)$ matrix. Specifying each variable at separate time periods instead of stacking them over time allows us to simulate explicitly their time persistence characteristics. Finally, since our underlying data was available in all periods for 69 countries, and we are seeking a balanced panel, we set $N = 69$.

We assumed these variables are measured without error. In other words, we take their resulting first and second moments to be those of the "true" variables, which we will later shock by adding a white noise measurement error. As the observed data surely incorporates measurement error, we will be understating the magnitude of the covariances among the underlying "true" variables and overstating their variances.

Simulating the fixed effects term μ_i. One difficulty we face is simulating the country fixed effects $\mu_i = \left(1 - e^{-\lambda\tau}\right) \log A_i(0) - \gamma_0$.[23] Theory provides no guidance as to the values of $A_i(0)$ for different countries, since it is taken as exogenous. The problem is important because the covariance structure linking the country-specific effects to the observed regressors determines how much heterogeneity bias will be present. To obtain simulated fixed effects and their corresponding covariance structure with the right-hand side variables, we used our observed panel data set to run an FE regression on the specification in equation (24). We computed the fitted fixed effects from this regression. We then used this series and treated it as an additional variable, as if it were observed, to generate the moments of the simulated data.

Obviously, given that the underlying data must incorporate measurement error, this procedure will lead to biased estimates of the country-specific effects, as discussed in Section 2. If this is

[22] Since the model is dynamic, subsequent values of the initial income term $\log y_{it-\tau}$ will be generated by iterating on income using the Solow specification, starting from a drawn value for the first period. As explained below, we calibrate the parameters so that subsequent generated values of the income variables bear characteristics resembling those of the corresponding real data.

[23] The time fixed effect $g(t - e^{-\lambda\tau}(t-\tau))$, which is identical for all countries at each date, was generated for each period t simply by setting the parameters g and λ to their assumed values, and $\tau = 5$.

the case, the moments of μ_i and especially its covariance structure with the other right-hand side variables will also be flawed. It is therefore critical that we also present results with alternative assumptions about the covariance structure between the simulated fixed effects and the simulated regressors in x_{it}. We do so below.

Moments of the underlying data. Table 1, panel A presents the matrix of correlations among our Q +1 variables of interest in the pooled data used to generate our simulated datasets. For example, once stacked over time and countries, $\log s_{k,it}$ bears a correlation of 0.60 with $\log s_{h,it}$. The estimated country-specific effect bears high correlations with the right hand side variables, suggesting a big scope for heterogeneity bias when using estimators that do not correct for it. For instance, the correlation between initial income $\log y_{i,t-\tau}$ and μ_i is 0.93.[24] Panel B isolates the between correlations among variables, by taking time means $(\bar{x}_i, \bar{\mu}_i)$ (where obviously $\bar{\mu}_i = \mu_i$) and computing their correlation matrix. The between correlations are quite close to the pooled data correlations, suggesting that cross-sectional variation dominates in our data. For example, the between correlation between $\log s_{k,it}$ and $\log s_{h,it}$ is 0.72. Finally, Panel C displays the within correlations, obtained by computing $(x_{it} - \bar{x}_i, \mu_i - \bar{\mu}_i)$. These correlations are always much lower than either the pooled or between correlations, again suggesting that the cross-country variation dominates in the pooled data. For example, the within correlation between $\log s_{k,it}$ and $\log s_{h,it}$ is 0.21.

Draws of simulated data. We are now in the presence of N observations for $T(Q-1)+2$ variables.[25] We computed the $(T(Q-1)+2) \times 1$ vector of means for these variables, denoted $\hat{m}_{x,\mu}$ and their variance covariance matrix, denoted $\hat{\Omega}_{x,\mu}$. Stacking the data in this way (in wide format) allows us to provide a realistic simulation of the relative weights of between and within variations - by specifying explicitly the autocorrelation structure of the right-hand side variables in addition to their cross-correlations. For each run of our simulation, we then drew N observations for the $T(Q-1)+2$ variables from a multivariate normal distribution with mean $\hat{m}_{x,\mu}$ and variance $\hat{\Omega}_{x,\mu}$.

[24] For the sake of illustration, in Table 1 we used every 5-year time-interval observation between 1960 and 1995 for the *real* data on $\log y_{it-\tau}$. In contrast, in our simulations, we are generating $\log y_{it-\tau}$ from the model, for all but the first period - due to the dynamic nature of equation (24). The simulated data on $\log y_{it-\tau}$ and their observed counterparts are very highly correlated (correlations are available upon request).

[25] i.e. N observations per period for $\log s_{k,it-\tau}$, $\log s_{h,it-\tau}$, $\log(n+g+\delta)_{i,t-\tau}$, N observations for $\log y_{it-\tau}$ in 1960 and N observations on the time invariant country effects.

The next part of the data generation procedure is to simulate the residuals ν_{it}. We opted to let the variance of the residual differ across time periods, and the residuals covary across time periods. To do this, we generated the fitted residuals from the fixed-effects regression using observed data for each period, and arrayed them in a $N \times T$ matrix. We computed their $T \times T$ covariance matrix $\hat{\Omega}_\nu$. Finally, we generated N sets of T normally distributed residuals with mean zero and covariance matrix $\hat{\Omega}_\nu$. An interesting aspect of this exercise is that the variance of the fixed-effects estimated residual term $\hat{\nu}_{it}$ was found to be a small fraction of the variance of $\log y_{it}$, on the order of 1%.

Using all the parameters and simulated data, we computed the simulated dependent variable $\log y_{i,1965}$ for period 2, using equation (24) and the simulated data on $\log y_{i,1960}$. We used this generated value of $\log y_{i,1965}$ to similarly generate $\log y_{i,1970}$, and so on iteratively until we obtained $\log y_{i,2000}$.[26] Formally, log income in period t for country i was simulated as:

$$\log y_{it} = \gamma_s \sum_{j=0}^{t-1} \gamma_4^j w_{i,t-j-1} + \gamma_4^t \log y_{i0} + \mu_i \sum_{j=0}^{t-1} \gamma_4^j + \sum_{j=0}^{t-1} \gamma_4^j \nu_{it-j} \tag{25}$$

3.1.2 Parameter Values

There is no guarantee that the generated income data resembles in any way the underlying real world data. Equation (25) shows that simulated income is a function of past values of the steady-state determinants in w_{it}, the log of income at the beginning of the first period $\log y_{i0}$, the fixed effects μ_i and a weighted sum of the current and past residuals ν_{it}, as well as the model's reduced form parameters in γ. As t increases, the moments of the generated values of income might diverge more and more from those observed in the true income data.

To address this issue we calibrated the model's parameters α and β so that the generated income variables in a typical draw of the data have moments resembling those of the observed variables. We found that we did not need to diverge greatly from commonly assumed values of α and β to obtain a good calibration: in a typical draw of the data, setting $\alpha = \beta = 0.27$ delivers moments of

[26] There are several reasons for implementing a dynamic method for simulating income rather than treating $\log y_{it-\tau}$ on the right hand side of equation (24) in the same way as we treat steady-state determinants. First, internal consistency requires that income be modelled in conformity with the dynamics of the Solow model. Second, this dynamic method will allow us to isolate and quantify the extent of the endogeneity bias arising under fixed effects, as identified by CEL and discussed above in section 2.2. Third, this is computationally required for the implementation of the AB estimator.

generated income variables that look similar to those seen in the PWT6.1 data.[27] These variables are conventionally both set to 1/3 in the context of the Solow model (as discussed for instance in Barro and Sala-i-Martin, 1995].

The other parameters of the structural model, g, δ and n are set to their conventional values as in Barro and Sala-i-Martin [1995]:

$$g = 0.02;\ n = 0.01;\ \delta = 0.05$$

These parameters imply a convergence parameter $\lambda = (n + g + \delta)(1 - \alpha - \beta) = 3.68\%$, slightly higher than the value of 2.67% implied by conventional values of α and β.[28] With these assumed structural parameters and $\tau = 5$, the implied reduced form parameters are as follows:

$$\gamma_1 \approx 0.099;\ \gamma_2 \approx 0.099;\ \gamma_3 \approx -0.197;\ \gamma_4 \approx 0.832$$

Note that in empirical applications of the Solow growth model, a small contradiction exists between the theoretically derived estimating equation and the linear specification actually estimated: insofar as the rate of population growth n enters the equation as a variable (in the term $\log(n + g + \delta)$), then terms such as $\left(1 - e^{-\lambda t}\right)$, where λ depends in part on n, should not be treated as constant.[29]

3.1.3 Measurement Error

The dataset generated above is free from measurement error. If we were to run fixed-effects regressions of $\log y_{it}$ on $\log s_{k,it-\tau}$, $\log s_{h,it-\tau}$, $\log(n + g + \delta)_{i,t-\tau}$ and $\log y_{it-\tau}$ using repeated draws of the simulated data, the only source of bias in the fixed-effects regression would be the endogeneity problem that stems from the dynamic nature of the model.[30] If we were to run between regressions

[27] Details of our calibration exercise, including a detailed comparison of the moments of the generated data with those of the observed data, are available upon request.

[28] We have rerun our simulations assuming $\alpha = \beta = 1/3$ and all the results were qualitatively unchanged. These results are available upon request.

[29] However, doing so is an acceptable approximation since variation in n is likely to have a small impact on variation in $e^{-(1-\alpha-\beta)(n+g+\delta)\tau}$. On the other hand, variation in n will have a larger impact on variation in $\log(n + g + \delta)$, which justifies not treating this variable as a constant. At any rate, we follow common practice in treating $e^{-\lambda\tau}$ as a constant and $\log(n + g + \delta)$ as a variable.

[30] Additionally, if we simulated the income data without the error term ν_{it}, fixed effects would lead us to recover *exactly* the reduced form theoretical parameters γ of the model. We did this to check our simulation program for programming errors, and the corresponding results are available upon request.

on simulated data, we would obtain estimated coefficients tainted only by heterogeneity bias, i.e. the term $\left(\Sigma_x^B\right)^{-1}\Sigma_{\mu x}$ in equation (22), where D is set to 0.

To evaluate the merits of various estimators used to estimate growth regressions in the presence of measurement error, we shocked our simulated variables by adding white-noise.[31] This can be done in several ways. In our baseline simulations, we simply added a normally distributed, zero mean shock with variance equal to some fraction $F < 1$ of the variance of the underlying variable (we will refer to F as the error-to-truth ratio). This was applied to simulated variables period-by period. Formally, consider first the determinants of the steady-state level of income, $\log s_{k,it-\tau}$, $\log s_{h,it-\tau}$ and $\log(n+g+\delta)_{i,t-\tau}$. For independent variable x_q in period t, we computed $x^*_{q,it}$ as:

$$x^*_{q,it} = x_{q,it} + d_{q,it} \tag{26}$$

with $d_{q,it} \sim N\left(0, F\hat{\sigma}^2_{qt}\right)$ for all i and $q = 2...4$, where $\hat{\sigma}^2_{qt}$ is the sample variance of $x_{q,it}$ in period t. We proceeded in exactly the same way for the income variable $\log y_{it}$:

$$\log y^*_{it} = \log y_{it} + d_{y,it} \tag{27}$$

where $d_{y,it} \sim N\left(0, F\hat{\sigma}^2_{yt}\right)$ for all i, where $\hat{\sigma}^2_{yt}$ is the sample variance of $\log y_{it}$ in period t.

In the specifications above, the variance of the measurement error can vary period-by-period insofar as the variance of the underlying true data does. In other words, the $T(Q-1)\times T(Q-1)$ variance covariance matrix of the errors-in-variables, $\hat{\Omega}_d$, is diagonal, with the diagonal elements allowed to differ across regressors and time. The fraction F, however, is common to all variables in all periods. We relax some of these assumptions on measurement error in the robustness tests presented in Section 4.

The parameter F is set to four values: 0%, 5%, 10%, and 15%.[32] It is, of course, difficult to know what the appropriate extent of measurement error is in reality. Hence, it is essential to vary F to assess the robustness of our results. However, clues about whether our chosen range of values for F is reasonable can be obtained. Notice first that all variables are entered in logs, so even a value of $F = 5\%$ may imply rather large shocks, especially on the underlying income variable y_{it}. Our additive term $d_{y,it}$ translates into some multiplicative term $e^{d_{y,it}}$ applied to y_{it} (income in 1996 PPP dollars).

[31] Obviously, we did not shock the fixed-effects μ_i nor the intercept.

[32] As an additional check, we also did our simulations with extreme values of F: 25% and 50%. The qualitative properties of our results were unchanged, but the extent of bias quickly became unreasonably large.

We can display draws of the mismeasured variables pooled across time periods and compute the average absolute value deviation from their true (unshocked) values, as a summary measure of the extent of measurement error. Table 2 displays these values as well as the pooled sample averages of the underlying "true" (unshocked) variables, for comparison. To construct Table 2, we drew simulated data for $2,000$ countries in the 8-period panel, i.e. $16,000$ observations.

Consider first measurement error on income y_{it}. For $F = 5\%$, the average absolute value of measurement error was $\$1,332$, for $F = 10\%$ it was $\$2,134$ and for $F = 15\%$ it was $\$2,703$. These are to be compared to the pooled sample mean of simulated income, which is $\$4,997$.[33] The magnitudes we obtain on the other variables seem more moderate, due to the fact that their values are between 0 and 1. Consider for instance the rate of physical capital accumulation s_k: for $F = 5\%$, the average absolute value of measurement error was 2.03 percentage points, for $F = 10\%$ it was 2.95 percentage points and for $F = 15\%$ it was 3.67 percentage points. The pooled sample mean of s_k was roughly 17%. Similar relative orders of magnitudes hold for s_h and $(n + \delta + g)$, as shown in Table 2. While it is hard to know what the appropriate level of measurement error would be, the range of values displayed in Table 2 does not seem unreasonable.

3.1.4 Regressions on Simulated Data

Having generated our simulated true data and shocked it with classical measurement error, we can now evaluate the bias properties of alternative estimators in the presence of correlated country-specific effects and measurement error. We estimated equation (8) on our draw of simulated data using four estimators: fixed-effects (FE), between (BE), random effects (RE) and Arellano-Bond (AB). We stored the estimated slope coefficients from each run, and repeated this procedure 1000 times. We then computed the means of the resulting estimates, and compared those to the known true parameters. The difference between the mean estimates and the corresponding true parameters gives a measure of bias for each estimates of the slope parameters in γ. The average absolute value of these biases across parameters is used as summary measures of bias across the slope elements of γ. Although our discussion of the results focuses on bias, the standard errors of the estimates from simulated data are also available to examine the efficiency properties of the estimators (we briefly discuss this in Section 4).

[33] Since the average absolute value of the shock on income may seem too large, we examine in Section 4.1 what happens when we reduce the extent of measurement error on income without changing it for the other variables.

3.2 Baseline Simulation Results

3.2.1 10% Measurement Error

In our baseline case, we set the error-to-truth ratio F equal to 10% for all the right-hand side variables in the model, and the extent of the correlations between the fixed effects and the regressors is as described in Table 1. Table 3, column 3 presents the resulting estimates based on averages over 1000 runs. In terms of the average absolute value of bias on the slope parameters, our results reveal that the BE estimator dominates by a wide margin: average absolute bias is 33%, versus a value in the neighborhood of 200% for the other three estimators. As suggested by econometric theory, estimators that use the within variation exacerbate measurement error bias, and between averages it out.

Turning to individual coefficient estimates, BE tends to bias the estimate of the convergence parameter γ_4 upward by 19% - the average simulated coefficient is 0.990 versus a true coefficient of 0.832 (the implied speeds of convergence, i.e. λ parameter, are respectively 3.68% and 0.2%).[34] In contrast, both the FE and AB severely bias this coefficient downwards, with average biases of -78% and -89% respectively, implying very high speeds of convergence (respectively 33.99% and 47.10%). In terms of the pattern of coefficients, this broadly replicates the finding of the literature - where the FE or AB estimates of the convergence speed are an order of magnitude higher than the between estimate. CEL, for example, report a speed of convergence of 10% per year based on the AB estimator - 5 times larger than the 2% cross-sectional estimate in MRW.[35] Our results suggest that the finding of fast convergence in the literature employing fixed-effects estimators may be traceable to the incidence of exacerbated measurement error bias.

Turning to the other slope parameters of the Solow model, interesting results also emerge. While all the estimators involve some bias, the direction and magnitudes of the biases differs sharply. BE

[34] To calculate λ from the estimate of γ_4, simply compute $\lambda = -\log(\gamma_4)/\tau$ where $\tau = 5$.

[35] The precise extent to which γ_4 is biased downwards when using FE and AB in our simulations obviously depends on the postulated extent of measurement error and the postulated correlations between μ_i and the elements of x_{it}. If the error-to-truth ratio is brought down to 5%, the estimated speed of convergence is brought down to 22.05% for FE and to 32.50% for AB. These values remain higher than those reported in the literature. In Section 4.1 we discuss how to obtain more reasonable values of FE and AB estimates of the convergence speed by reducing the extent of measurement error on the income term $\log y_{it}$ - Section 3.2.3 suggested that the average absolute value of the shock to y_{it} might be too high, on the order of $\$2,000$ for the pooled sample in the baseline simulation with $F = 10\%$. This is to be compared for a pooled sample mean of simulated income equal to roughly \$5000.

tends to bias the human capital parameter slightly away from zero: the coefficient on $\log s_{h,it}$ is biased upward by 6%. The corresponding bias when using FE is a downward bias of -209% - the coefficient switches signs. Again, our simulation account for differences between estimators found in the literature - where FE typically lead to a coefficient estimate on the human capital accumulation variable that is closer to zero than BE. For example, Islam [1995] shows that the estimated BE coefficient on $\log s_{h,it}$ is roughly 0.182, and equals -0.071 when using FE. Our corresponding BE estimate is 0.105, and our FE estimate is -0.108. Similar comparisons would hold when we turn to AB rather than FE - the accumulation parameters are both severely biased towards zero, and the depreciation parameter γ_3 is biased away from zero. Again, our results suggest that the finding of smaller effects of the accumulation variables in the fixed-effects literature compared to the cross-sectional literature may be largely attributable to measurement error bias.

AB estimates are very close to the FE estimates, suggesting, as detailed in Section 2, that the weak instruments problem may be prevalent here. To evaluate this more formally, we implemented the test of the null hypothesis of weak instruments suggested by Stock and Yogo [2003], using the real world data that serves as a basis for our simulations. This test is based on computing the Cragg-Donald [1993] statistic, a generalization of the first-stage F-test for the case of multiple endogenous regressors.[36] The statistic is then compared to the critical values in Stock and Yogo [2003]. The critical values depend on parameter b, the maximum amount of squared bias that the researcher is willing to accept relative to squared OLS bias (in our case, FE bias). For instance, a value $b = 0.1$ indicates that the maximal allowable bias of the IV estimates is 10% of the maximum OLS bias. In our case, the value of the Cragg-Donald statistic was 1.513, which is smaller than all critical values whatever the value of b presented in Stock and Yogo [2003] (these range from 5% to 30%). Thus, we fail to reject the null of weak instruments at the 5% significance level even when we are willing to accept a high level of AB bias relative to FE bias.[37]

[36] In our application we allow all the right hand side variables to be predetermined - so that all four regressors are instrumented for. In doing this we follow the practice in CEL. Strictly speaking however, in our simulation only lagged income is endogenous. See the discussion after equation (13). Formally, the Cragg-Donald statistic is the smallest eigenvalue of the matrix analog of the F-statistic from the first stage regression. See Stock and Yogo [2003].

[37] We also implemented the Staiger and Stock [1997] rule of thumb based on the first stage F-statistics. Strictly speaking, the rule of thumb is only valid for the case of one endogenous regressor. But the values of our first stage F-statistics were sufficiently below 10 to reinforce our confidence that the weak instruments problem is important here. The F-statistics for the first-stage regression of lagged first-differenced initial income on its instruments was 4.73, and the corresponding values for savings on physical capital, savings on human capital and the depreciation

As long as the first stage relationship between the instruments (in levels) and the regressors (in first differences) is not exactly zero, the weak instruments bias should vanish in large samples. To further examine whether weak instruments are responsible for the AB bias we observe in our simulations even without measurement error, we reran our simulations setting $N = 1000$ instead of $N = 69$. The results (available upon request) decisively indicate that the weak instruments problem is almost entirely responsible for the AB bias. The bias almost goes away when the cross sectional sample size is raised to 1000 (the average absolute bias is 3% instead of 16%). As soon as measurement error is introduced, however, large AB biases reappear even when $N = 1000$. The reason is that the introduction of measurement error weakens the first stage relationship between the differenced regressors and the levels instruments, considerably slowing down convergence to the true parameters as the sample size is raised.

Finally, Table 3 also reveals that RE performs more poorly than the other three estimators when the summary measure of bias is the average absolute value of the bias (here 235% compared to 33% for the between estimator). This is significant because this estimator is frequently used by growth regression practitioners who wish to retain the panel dimension but are unwilling to discard all the between variation in the data. However, RE does quite well in estimating the convergence parameter, displaying a bias of only -16%. The other slope parameters are all biased upwards. For instance the coefficient on $\log s_{k,it}$ is biased upward by 107%. One possible reason is that, contrary to between, RE does not average measurement error over time, nor does it address the problem of heterogeneity bias. The interaction of these two biases, which is hard if not impossible to characterize analytically, turns out to result in large net biases in this particular application.

3.2.2 Varying the Extent of Measurement Error

Increasing the error-to-truth ratio to 15% or reducing it to 5% does not generally change the conclusions reached above (Table 3, columns 2 and 4). As expected, the average absolute bias tends to (weakly) increase with the error-to-truth ratio for most estimators, though this is not necessarily true for individual parameter estimates. An interesting feature of our simulations is how increases in the extent of measurement error across columns of Table 3 seem to little affect the BE estimate of the coefficient on lagged income per capita - in fact the upward bias on γ_4 remains equal to $18 - 19\%$ whatever the value of F.

term were, respectively, 4.03, 2.71 and 2.22.

Interesting lessons can also be learnt when measurement error is shut down entirely. While unrealistic, this exercise allows us to isolate the incidence of heterogeneity bias in BE estimates, and of endogeneity bias in FE estimates. Table 3, column (1) presents simulation results when $F = 0$. As predicted, BE still tends to create an upward bias on the lagged income coefficient, an upward bias on the effect of human capital, and a small downward bias on the depreciation term. Clearly, the strong positive correlation between the country-specific effect and the lagged income term, built into our simulations, accounts for the upward bias on $\log y_{it-\tau}$ when country fixed effects are not included in the regression. The results suggest that most of the bias in the BE estimates seen when F is set to a value different from zero is attributable to heterogeneity bias, as the biases on individual coefficients change little as F is increased. In other words, BE does a good job at averaging away measurement error.

When $F = 0$, FE estimates are also biased. This is due to the endogeneity problem inherent in this type of dynamic panel. Our simulations allow us to quantify this problem. The biases are relatively small, especially on the main parameter of interest γ_4 (biased downwards by -5%). This bias quickly gets swamped by measurement error bias when F is increased. The AB estimator, which is supposed to get rid of endogeneity bias, does display biases of similar magnitudes as FE. Moreover the biases on the various slope parameters are similar in signs and relative magnitudes to the FE biases. As discussed above, this is due to the weak instruments problem, which tends to bias AB estimates towards FE.

To summarize, as predicted by theory, when measurement error is not present, BE is tainted by heterogeneity bias, the other estimators perform better and the AB estimator in particular performs best. However, this case is unrealistic since we are unable to replicate the broad findings obtained across estimators with real data when setting $F = 0$. Even in the presence of a small amount of measurement error ($F = 5\%$), large biases appear when using FE, and the BE estimator asserts itself as the dominant estimator in terms of average absolute bias. Moreover, in that case we are able to broadly replicate the pattern of estimates found in the literature across estimators.

3.2.3 Varying the Extent of Heterogeneity Bias

We now examine how our results change when we vary the extent of heterogeneity bias, holding measurement error fixed at some baseline level. As described in Section 3.1.2, the assumed correlations between the right hand side variables and the country fixed effect μ_i used to draw simulated data were based on estimated values of μ_i from an FE regression. We know from econometric theory that in the presence of measurement error, these estimated μ_is will be inconsistent estimates of the true country fixed-effects, therefore their sample correlations with the regressors will themselves be flawed. Hence, it is critical to examine what happens when we change these assumed correlations.

Table 4 displays simulation results when varying the correlations between the country-specific effects and the regressors, while setting $F = 10\%$. Table 1 showed that the correlations used for our baseline simulations were high. For example, in the pooled sample the correlation between our estimated μ_i and $\log y_{it-\tau}$ was 0.93, and the correlation with $\log s_{h,it-\tau}$ was 0.80. Here we simply multiply all these assumed correlations, variable-by-variable and period-by-period, by a single constant $C < 1$, prior to generating the simulated data.[38] We allowed C to take on the values 0%, 5% and 75%. For the sake of comparison, column (5) of Table 4 also reports the results obtained when $C = 100\%$ (i.e. column (3) of Table 3).

Table 4 demonstrates that the biases obtained under FE when varying C do not change drastically, suggesting that most of the FE bias comes from measurement error.[39] Notably, the estimate of γ_4 changes relatively little (from -56% when $C = 0\%$ to -78% when $C = 100\%$). As expected from econometric theory, the AB estimate of γ_4 exhibits an even greater degree of stability across values of C.

We now turn to the BE estimator in the extreme case where $C = 0$. Again, this is an unrealistic assumption, but it allows us to evaluate the incidence of measurement error in isolation from heterogeneity bias. We observe that the average absolute value bias is increased (to 157% compared to 33% when $C = 100\%$), but that the pattern of signs and relative magnitudes for the bias is roughly in line with the results in column (3) of Table 3. The convergence parameter γ_4 now exhibits a larger bias (34% rather than 19%). The same pattern holds for all the other slope parameters: as C rises, the extent of bias is progressively reduced, illustrating nicely a main message of this paper: as the incidence of heterogeneity bias rises, it increasingly mitigates the problem of measurement error for the BE estimator. In this case, the two sources of bias tend to *cancel each other out*.

[38] That is, we modified the relevant entries of the data covariance matrix $\hat{\Omega}_{x,\mu}$ used to generate the simulated series.

[39] We should not expect the FE estimates to remain unchanged when varying the value of C. While it is true that this estimator eliminates the country specific effect by differencing the data from country means, the extent of bias may vary as C is modified, as such a change will affect the properties of the simulated μ_i and thus those of the simulated $\log y_{it-\tau}$.

As C is increased to 50% and then to 75%, the mean estimated slope parameters from BE progressively converge to the values in column (5), and mean absolute value bias decreases steadily. To summarize, variations in the incidence of heterogeneity bias do not change the lessons of our baseline simulations regarding convergence: even when we change the incidence of heterogeneity bias, BE and RE provides the closest estimates of the speed of convergence. The signs of the biases on the parameters of the steady-state determinants are robust to changes in C for all estimators. As before FE and AB always bias downwards estimates on the accumulation variables, bias upwards the estimate on $\log(n+g+\delta)_{it-\tau}$ and result in much too fast an estimated speed of convergence.

4 Extensions

This section considers various modifications of our basic simulation method. We consider what happens when we change assumptions on the nature of measurement error. We also examine the properties of two additional estimators frequently used in the empirical literature on economic growth.

4.1 Varying measurement error on income

While our baseline simulation results replicate the broad differences in past findings on convergence and the determinants of steady-state income level across estimators, the estimated speed of convergence under the FE and AB estimators was too high relative to the BE estimate - we obtained an FE estimate of $\lambda = 33.99\%$ and a BE estimate of $\lambda = 0.21\%$, while the literature finds values in the neighborhood of 10% and 2%, respectively. Moreover, in Section 3.1.3 we showed that an error-to-truth ratio of $F = 10\%$ implies an average absolute value error in measurement of roughly \$2,000, while the mean of simulated income was \$5,000. While it is difficult to know what the appropriate extent of measurement error is, this is probably too big.[40] In this subsection we exam-

[40] In contrast, the extent of measurement error on the other variables, implied by setting $F = 10\%$, seemed more realistic. Some have argued that per capita income may be better measured than savings rates on human capital s_h, physical capital s_k and the depreciation variable $(n+g+\delta)$. In principle population growth n will be well-measured, but recall that we had to make an assumption of constancy across time and countries for δ and g, which surely introduces error. Similarly, s_h in the Solow model should be measured by dollars saved per unit of time for the purpose of financing education, but we followed the literature in proxying for this using gross enrollment rates in secondary education. However, it is well-known that different methods of computing price indices and PPP exchange rates can deliver vastly different estimates of PPP income.

ine whether reducing the extent of measurement error on log income can help match convergence speeds estimated in the literature: we reduce the extent of measurement error on the income term (F_y), while maintaining $F = 10\%$ on the other variable. We consider values of F_y equal to 0%, 1%, 2.5%, and 5%.

We might expect reduced measurement error on $\log y_{it}$ to reduce attenuation bias on its coefficient and thereby improve the performance of the FE and AB estimators (though this of course is not always true in the presence of measurement error on the other variables). Our intuition turns out to be borne out: Table 5 demonstrates that with a value of $F_y = 1\%$, we obtain FE and AB estimates of λ that are much closer to those found in the past literature: 9.19% and 10.81%, respectively. With $F_y = 1\%$, the average absolute value of the error in measurement on y_{it} is $618, which is perhaps a more realistic value than those implied by $F_y = 5\%$ and more. The BE estimate of γ_4 is unchanged compared to our baseline simulations, with a 19% upward bias, confirming our suggestion that most of this bias is attributable to the omission of the country-specific effect μ_i.

In general, the average absolute value bias becomes much lower for the AB and FE estimators, largely because the bias on the income term is now reduced. In fact, when we set $F_y = 0$, these estimators get convergence almost right, suggesting that measurement error in this variable is important to replicate the pattern of γ_4 estimates found in the literature.

To summarize, when we allow for a smaller error-to-truth ratio on income, we are able to obtain FE and AB estimated speeds of convergence that are much closer to those obtained when using real data. The extent of bias on the other parameters is not affected very much by measurement error on income. We still get BE estimates of γ_4 that are too large (and therefore BE estimates of λ that are too small) relative to the existing cross-country literature. We will see why in Section 4.4.

4.2 Country-specific measurement error

The extent of measurement error probably varies from country to country, and is correlated with country characteristics such as per capita income. Fortunately, we do have some information on the accuracy of statistics used in most empirical growth studies: PWT6.1, acknowledging that all of their PPP adjusted data is probably measured with some error, includes a data quality rating for each country. The scale of the rating runs from A to D, which we recode as running from 1 to 4, so that countries with less accurate data have a higher rating.

We exploited this information by adjusting the error-to-truth ratio to allow it to vary across countries in proportion to the PWT6.1 data quality rating. Since the extent of measurement error is probably correlated with some of the other variables used as regressors, we added the PWT data quality indicator as another variable for the purpose of computing the covariance matrix of the observed right-hand side variables, thus allowing simulated data quality ratings to covary with the other simulated variables. For example, lower income is associated with lower data quality. This expanded covariance matrix for the actual data was then used to draw the simulated data, including a simulated data quality rating. Finally, after drawing the errors in measurement, we multiplied them by the country's normalized simulated data quality rating. The simulated data quality rating is normalized by its average. So if $F = 10\%$, a country with the average level of data quality would have an error-to-truth ratio of 10%, while a country with twice the average data quality rating would have an error-to-truth ratio of 20%.

Results displayed in Table 6 show that this extension does not change our basic results. With $F = 10\%$, BE still performs best at estimating the rate of convergence with an upwards bias of 19% on γ_4. On the other hand, FE displays a downwards bias of 78% and AB leads to a downwards bias of 89%. BE still outperforms the other estimators in general, with an average absolute bias of 33% compared to 212% for the fixed-effects estimator and 213% for the AB estimator. In fact, a comparison of the entries of Table 6 with those of Table 3 reveals very little difference. Thus, allowing the variance of measurement error to depend on observables does not change our findings.

4.3 Autocorrelated measurement error

So far, we have assumed classical measurement error, i.e. the error in measurement was purely white noise. However, errors in measurement could be autocorrelated across time. For instance, if a country has over-reported the amount of savings in physical capital in one period, it may be more likely to do so in subsequent periods. Hence, measurement error can be expected to persist over time. Moreover, as discussed in Section 2, persistent measurement error invalidates the IV procedure of the AB estimator when it comes to addressing measurement error bias, since the error in measurement in lagged regressors (the instruments) is no longer independent from error in measurement in the regressors. In this subsection, we run simulations where autocorrelation is built into measurement error, as a further robustness check. Specifically, for each regressor x_q

$(q = 2...4)$ and for the income term $\log y_{it}$, we set:

$$\begin{cases} d_{q,it} = \rho_d d_{q,it-\tau} + \zeta_{it} \\ d_{y,it} = \rho_d d_{y,it-\tau} + \vartheta_{it} \end{cases} \tag{28}$$

where we still maintain $d_{q,it} \sim N\left(0, F\hat{\sigma}^2_{qt}\right)$ and $d_{y,it} \sim N\left(0, F\hat{\sigma}^2_{yt}\right)$. In other words, the $T(Q-1) \times T(Q-1)$ variance covariance matrix of the errors-in-variables, $\hat{\Omega}_d$, is now block diagonal, with the diagonal elements identical to what they were before and the off-diagonal elements of each $T \times T$ block as implied by equation (28). Errors are then drawn from a multivariate normal distribution with mean zero and covariance matrix $\hat{\Omega}_d$, as before.

In this case, we do have some theoretical priors as to how persistence in measurement error might affect the results. If measurement error is highly persistent over time, we would expect FE to perform better relative to the case of $\rho_d = 0$, for two reasons: first, with persistence in $d_{q,it}$, some of the measurement error will be differenced away when the data is differenced from its country means. In other words, the greater the autocorrelation in measurement error, the larger the between component of measurement error and the smaller the within component. Since FE will difference away the between variation, we expect greater autocorrelation in the measurement error term, all else equal, to improve the performance of FE relative to BE. In the limit, when the autocorrelation coefficient ρ_d is 1, we would expect FE to get rid of all of the measurement error, as it will get entirely differenced away.[41]

We ran simulations when the measurement error autocorrelation term was $\rho_d = 0.5$, $\rho_d = 0.75$ and $\rho_d = 0.90$, setting $F = 10\%$. The results in Table 7 show that our theory-driven priors are confirmed by the simulations. Average absolute value bias declines with ρ_d in FE and AB estimators. When the autocorrelation term is only 50%, the convergence coefficient γ_4 exhibits a 20% bias using the BE estimator and a -49% bias using the FE estimator (in the baseline case of $\rho_d = 0$, the corresponding numbers were 19% and -78%, respectively). When the autocorrelation term is increased to 75%, the bias on the BE estimate of γ_4 rises to only 21%, and the bias on the FE estimate declines to -35%. Finally, when the autocorrelation term is increased all the way to 90%, which is probably much too high, the BE coefficient has a bias of 21% and the FE bias of -25%. Therefore, while we confirm our intuition that the FE estimator improves relative to the BE estimator when we increase persistence in measurement error, BE tends to do better or as well

[41] We checked that this is the case, and the results are available upon request.

as alternative estimators even as ρ_d is raised to implausibly high levels. Even high persistence in measurement error does not invalidate the overall conclusions reached in the baseline case.

4.4 Variance Properties

Up until this point, we have focused only on the bias generated by the various estimators. However, when evaluating the properties of estimators in an empirical study, one might also look at the standard errors and t-statistics of the estimators to check their efficiency properties. In order to check to see whether or not our gains or losses in bias were being offset by gains and losses in the standard errors of the estimators, we ran simulations that checked the mean squared errors and t-statistics of the estimators for our baseline case of $F = 10\%$. For each estimator the mean squared error reported is the mean of each coefficient's squared error across 1000 runs, and the t-statistic reported is the mean of 1000 individual t-statistics from each run. The results are reported in Table 8. As we can see from these results, there is no trade-off between bias and efficiency. In the baseline case, BE, which has the lowest average absolute bias, also has the lowest sum of mean squared errors and significantly high t-statistics on every coefficient except for $\log(n + g + \delta)_{it-\tau}$. Looking at individual coefficients, the BE estimator has the lowest MSE on every coefficient except the convergence coefficient. Even on this last coefficient, where RE is the better performer, the difference between BE and RE is quite small. Because MSE is a widely used device to measure the trade-off between bias and error, this observation makes a strong case that BE is superior on both bias and variance grounds.

4.5 Other Estimators

4.5.1 Flexible random effects (SUR)

In this subsection, we evaluate the bias properties of the Seemingly Unrelated Regressions (SUR) estimator, commonly used in the empirical growth literature (see, for example, Barro and Sala-i-Martin 1995, chap. 12). This estimator is computationally close to the RE estimator in that it also weighs the between and within variations in the data. However, in contrast to RE, the SUR estimator does not assume the within-country correlation in the error term to be the same across subsequent time periods, but instead allows it to vary. For example, the covariance between ε_{i1} and ε_{i2} is allowed to differ from the covariance between ε_{i2} and ε_{i3}. Thus, we can refer to the SUR estimator as a flexible RE estimator, as the residual covariance matrix is less restrictive. This is expected to lead to efficiency gains. Moreover, the weighing of the between and within variations will now differ from the RE weighing scheme, and be a complicated function of the variance of the fixed-effects, the variance of the error term $\mu_i + \nu_{it}$ and its autocorrelation structure. Thus, the bias properties of SUR may differ from those of random effects if the country-specific effects are correlated with the regressors, since it will weigh the within and between variations differently.

Results in Table 9 show that the estimates are not very different from RE (for the sake of comparison Table 9 also includes the RE results already presented in Table 3), but SUR does better overall than RE across values of F. For example, the convergence coefficient γ_4 displays a bias of -4% with SUR and -16% with RE when $F = 10\%$. In fact, SUR overall is the best estimator when it comes to estimating the speed of convergence λ. In terms of average absolute bias, SUR does better than RE whatever the level of measurement error. Both estimators do particularly well in estimating the speed of convergence, but both tend to greatly bias away from zero the estimates on steady-state determinants.

4.5.2 The Mankiw, Romer and Weil estimator

The BE estimator does not strictly correspond to the cross-sectional estimator often used in the cross-country growth literature. Indeed, it involves the time averaging of all variables, including the income term on the left-hand side and lagged income on the right-hand side. In contrast, cross-sectional estimators in the class of MRW's OLS estimator are based on the following regression:

$$\log y_{i,2000} = \gamma_0 + \gamma_1 \overline{\log s_{k,i}} + \gamma_2 \overline{\log s_{h,i}} + \gamma_3 \overline{\log (n+g+\delta)}_i + \gamma_4 \log y_{i,1960} + \vartheta_{it} \tag{29}$$

where "" denotes averages computed over the whole period. Contrary to the BE estimator, income enters as end and beginning of period values (our total period spans 1960-2000), so measurement error in the initial income term does not get averaged away.[42]

Table 9 displays the simulation results, with the appropriate correction. For comparability, we also reproduce the BE results from Table 3. The MRW and BE biases are very similar in terms of magnitudes and signs. However, MRW does slightly better than BE in terms of average absolute bias across values of F. This partly reflects a lower bias on the convergence parameter γ_4. Since lagged income does not get averaged, measurement error on this variable counteracts the upward

[42] Note also that in this application τ is set to 40, requiring a correction to ensure the comparability of the estimated reduced form coefficients with BE estimates.

bias from the omission of μ_i to a greater degree than BE. These results confirm that simple OLS cross-sectional estimators are best at limiting net overall bias resulting from heterogeneity and measurement error.

4.5.3 The Arellano-Bover/Blundell-Bond estimator

We argued in section 3.2 that weak instruments led to bias in the AB estimates when the sample was small, and that measurement error exacerbates this small sample bias and reduces the speed at which estimates converge to the true parameters as N increases. This problem has received increasing attention in the dynamic panel data literature, and Arellano and Bover [1995] and Blundell and Bond [1998] have developed a "system GMM" estimator to address it.[43] In addition to instrumenting for differenced variables using lagged levels (see equation 13 and the discussion that follows), both papers suggest using lagged differences to instrument for levels variables. Both the equation in first differences (with levels instruments) and the equation in levels (with first-differenced instruments) are included in the system, so as to exploit the additional moment conditions valid for the latter equation. This procedure will be valid under a stationarity assumption that covariance between $\log y_{i,t-\tau}$ and μ_i is constant across values of τ (this assumption must be made in addition to all the assumptions required for the AB estimator). Indeed, if $E(\mu_i y_{i,t-s}) = E(\mu_i y_{i,t-r})$ for all s, r, then the moment condition $E(\mu_i (y_{i,t-s} - y_{i,t-r})) = 0$ can be used in a GMM procedure that employs $y_{i,t-s} - y_{i,t-r}$ as an instrument for $y_{i,t-\tau}$ (where $s > r \geq \tau$).

Table 9 presents simulation results using the Blundell-Bond (BB) estimator. This estimator behaves in ways similar to the RE and pooled OLS estimators. Namely it does pretty well at estimating the speed of convergence, but generates large upwards biases on the estimated coefficients on the steady-state determinants. Since this estimator includes specifications in levels in addition to first differences, the similarity between the biases it exhibits and those of pooled OLS might indicate the weak instruments problem remains prevalent. However, the biases did not vanish as the sample size was raised to $N = 1000$ or even $N = 5000$. An alternative explanation for the bias is that the stationarity assumption required for the moment condition $E(\mu_i (y_{i,t-s} - y_{i,t-r})) = 0$ is violated. In the context of our simulations, this assumption can be evaluated directly.[44] We computed the covariances between $y_{i,t-s}$ and μ_i for all periods in a simulated dataset with $N = 2000$. The

[43] This estimator has been used in the growth literature by Levine, Loayza and Beck [2000].

[44] We are not aware of a formal test for stationarity in this context, but simply examining the covariances over time can help determine whether they are roughly equal.

covariances increased steadily with time, ranging from 1.63 from period 1 to 3.46 for period 8. The violation of the stationarity assumption may therefore account for the disappointing performance of the BB estimator in the specific context of growth regressions.

5 Conclusion

In this paper, we have used Monte Carlo simulations to evaluate the econometric methods commonly used to estimate growth regressions. Our results suggest that, in the presence of measurement error, fixed-effects and the Arellano-Bond GMM estimator lead researchers to overestimate the speed of convergence and to underestimate the impact of several common determinants of the steady-state level of income, such as human capital. Simple OLS on variables averaged over time provides a closer estimate of the speed of convergence, but overestimates the magnitude of the effect of steady-state determinants. These findings were shown to be robust to changes in the specification of measurement error and to varying assumptions about the incidence of heterogeneity bias. Running Monte Carlo simulations required that we take the Solow model seriously as the data generating process for our simulated series. However, many of the lessons learnt here can be applied readily to other related specifications in the empirical literature on cross-country comparisons of income, growth and other variables.

Until now, differences in speeds of convergence across estimators were interpreted as implying that heterogeneity bias was prevalent in cross-sectional growth regressions, since fixed-effects methods were thought to correct for this bias and led to a speed of convergence roughly 5 times higher than that estimated using a between estimator. This paper has overturned this interpretation, showing instead that the difference in estimated convergence speeds is in fact attributable to greater bias from measurement error when using this class of estimators. The estimated speed from traditional cross-sectional regressions is in all likelihood closer to the true speed of convergence.

Of course, our simulations can only characterize the properties of the estimators. They cannot inform us as to the actual speed of convergence or the impact of steady-state determinants, since we simulated our data by assuming values for these parameters implied by a strict application of the Solow model. This strict application led to postulated parameters that may or may not hold true in actual data. Table 10 displays the results from estimating our basic specification using the new PWT version 6.1 data and updated series for the secondary school enrollment rate from Barro and Lee. We are able to replicate the basic findings of the past literature in this data: the speed of convergence is roughly 5 times larger under FE and AB (respectively 4.6% and 5.3%) compared to BE (0.8%). Our simulation results suggest the latter number is likely to be much closer to the truth. All the other estimators, that do not isolate the within variation in the data (namely MRW, SUR and RE) lead to estimated speeds of convergence that lie between 0.8% and 1.6%, while FE and AB lead to estimates in the neighborhood of 5%.

The speed of convergence we report based on the application of the BE estimator, less than 1%, falls short of the number typically reported in the cross-sectional literature. Barro and Sala-i-Martin [1995] cite a number closer to 2%, based on the previous version of PWT.[45] This difference is not attributable to our use of new and extended data.[46] Past cross-sectional estimates rely on an OLS specification closer to the MRW estimator described above, where the current and lagged income terms do not get averaged over time. Implementing this estimator, we obtain a convergence speed of 1.62%, which is closer to existing estimates. Given that BE and MRW have been shown above to somewhat understate the speed of convergence, a number in the neighborhood of 2% for λ does not seem unreasonable.

The slope parameters on the determinants of the steady-state level of income are reduced in magnitude when using FE or the AB estimator. They are similar across estimators that use at least some between variation in the data (BE, SUR, RE and MRW). For example, the impact of the log of the enrollment rate is equal to 0.04 using the BE estimator, and is significant at the 1% level. As in Islam, this estimate switches signs (to -0.03) when using the FE estimator, and is statistically insignificant.

Our paper illustrates an econometric second best property: by addressing one source of bias (stemming from omitted variables), the application of a certain class of estimators makes another source of bias worse (measurement error). Unfortunately, simulation exercises such as ours almost always come too late in empirical economics.

[45] This is also true for the (biased) FE estimated speed, which at 5% is about half of the convergence speeds reported in CEL [1996] and Islam [1995].

[46] We removed the 1990-1999 decade to check this, and our results were unchanged. In an effort to maintain a balanced panel and cover as wide a time period as possible, our regressions also feature only 69 countries. This was not responsible either for the reduced speed of convergence, as results obtained with more countries using an unbalanced panel were essentially the same. All these results are available upon request.

Appendix 1 - Limiting Values of the Between, Fixed-Effects and Random Effects Estimators

This appendix derives the limiting values of the BE, FE and RE estimators in the multivariate case. As in the text, assume that the true model is:

$$\log y_{it} = \gamma' x_{it} + \mu_i + \varepsilon_{it} \tag{30}$$

where all these variables are as defined in Section 2. As in subsection 2.3, we also abstract from the dynamic nature of the model (i.e. assume $\operatorname{plim} X'\varepsilon = 0$).

Assume also that x_{it} is imperfectly measured. Instead of observing x_{it}, we can only see $x^*_{it} = x_{it} + d_{it}$, where $E\left[d_{it}|x_{it}\right] = 0$ for all observations and $\operatorname{var}\left[d_{it}|x_{it}\right] = D = \operatorname{diag}\left\{\sigma^2_{d_1}, \sigma^2_{d_2}, ..., \sigma^2_{d_k}\right\}$. We can derive the following unconditional expectations for three variables of interest:

$$E\left(y_{it}, x^*_{it}, \mu_i\right) = \left(\gamma'\bar{x} + \bar{\mu}, \bar{x}, \bar{\mu}\right) \tag{31}$$

Their unconditional variance is:

$$V\left(y_{it}, x^*_{it}, \mu_i\right) = \begin{bmatrix} \sigma^2_\varepsilon + \gamma'\Sigma_{xx}\gamma + 2\gamma'\Sigma_{\mu x} + \sigma^2_\mu & \gamma'\Sigma_{xx} + \Sigma'_{\mu x} & \gamma'\Sigma_{\mu x} + \sigma^2_\mu \\ \Sigma_{xx}\gamma + \Sigma_{\mu x} & \Sigma_{xx} + D & \Sigma_{\mu x} \\ \gamma'\Sigma_{\mu x} + \sigma^2_\mu & \Sigma'_{\mu x} & \sigma^2_\mu \end{bmatrix} \tag{32}$$

In order to analyze the properties of the BE and FE estimators further, it is useful to break down the variation on each variable into within-country variation and between-country variation. Define the between-country variance for x_{it} as:

$$\Sigma^B_{xx} \equiv E\left[\left(\frac{1}{T}\sum_{t=1}^{T} x_{it}\right)\left(\frac{1}{T}\sum_{t=1}^{T} x'_{it}\right)\right] - E\left(\frac{1}{T}\sum_{t=1}^{T} x_{it}\right) E\left(\frac{1}{T}\sum_{t=1}^{T} x'_{it}\right) \tag{33}$$

and the within-country variance to be:

$$\begin{aligned} \Sigma^W_{xx} \equiv\ & E\left[\left(x_{it} - \frac{1}{T}\sum_{t=1}^{T} x_{it}\right)\left(x'_{it} - \frac{1}{T}\sum_{t=1}^{T} x'_{it}\right)\right] \\ & - E\left[x_{it} - \frac{1}{T}\sum_{t=1}^{T} x_{it}\right] E\left[x'_{it} - \frac{1}{T}\sum_{t=1}^{T} x'_{it}\right] \end{aligned} \tag{34}$$

It is well-known that:

$$\Sigma_{xx} = \Sigma^W_{xx} + \Sigma^B_{xx} \tag{35}$$

It is also easy to show that, for the covariance between x_{it} and μ, $\Sigma_{\mu x} = \Sigma_{\mu x}^B$. Finally, $\Sigma_{xx^*}^B$, the between covariance matrix of the imperfectly observed data x^*, is defined as:

$$\begin{aligned}
\Sigma_{xx^*}^B &\equiv E\left[\left(\frac{1}{T}\sum_{t=1}^{T} x_{it} + d_{it}\right)\left(\frac{1}{T}\sum_{t=1}^{T} x_{it}' + d_{it}'\right) - E\left[x_{it} + d_{it}\right] E\left[x_{it}' + d_{it}'\right]\right] \\
&= \Sigma_{xx}^B + \frac{1}{T} D
\end{aligned} \tag{36}$$

It is also easy to show that:

$$\Sigma_{xx^*}^W = \Sigma_{xx}^W + \frac{T-1}{T} D \tag{37}$$

We are now ready to derive the plims of the BE and FE estimators in the presence of measurement error and in the multivariate case.

First consider the BE estimator (OLS on country means across time). Using standard OLS results, we can derive:

$$\operatorname{plim} \hat{\gamma}^{BE} = \left(\Sigma_{xx}^B + \frac{1}{T} D\right)^{-1} \Sigma_{xx}^B \gamma + \left(\Sigma_{xx}^B + \frac{1}{T} D\right)^{-1} \Sigma_{\mu x} \tag{38}$$

Now consider FE. To eliminate the heterogeneity bias arising through the correlation between the time invariant country-specific effects and the regressors, the most obvious solution is to use the FE estimator. By the Frisch-Waugh theorem, we can show that:

$$\hat{\gamma}^{FE} = \left(X^{*\prime} M_c X^*\right)^{-1} X^{*\prime} M_c y \tag{39}$$

where:

$$M_c = I - C\left(C'C\right)^{-1} C' \tag{40}$$

and C is an $(NT \times N)$ matrix that stacks dummy variables for the different countries (with sub-vectors of T ones along the diagonals, zero elsewhere). Then:

$$\operatorname{plim} \hat{\gamma}^{FE} = \left(\Sigma_{xx}^W + \frac{T-1}{T} D\right)^{-1} \Sigma_{xx}^W \gamma \tag{41}$$

Finally, as is well-known, RE is simply a matrix-weighted average of BE and FE estimates:

$$\hat{\gamma}^{RE} = \left(\hat{\Sigma}_{xx^*}^W + \hat{\theta}\hat{\Sigma}_{xx^*}^B\right)^{-1}\left(\hat{\Sigma}_{xx^*}^W \hat{\gamma}^{FE} + \hat{\theta}\hat{\Sigma}_{xx^*}^B \hat{\gamma}^{BE}\right) \tag{42}$$

where $\hat{\Sigma}_{xx^*}^W$ and $\hat{\Sigma}_{xx^*}^B$ are the sample estimates of $\Sigma_{xx^*}^W$ and $\Sigma_{xx^*}^B$, respectively, and $\hat{\theta}$ is an estimate of θ where:

$$\theta = \frac{\sigma_\varepsilon^2}{T\sigma_\mu^2 + \sigma_\varepsilon^2} \tag{43}$$

i.e. θ is the weights given to the BE estimator. Then:

$$\begin{aligned}\operatorname{plim}\hat{\gamma}^{RE} &= \left[\left(\Sigma_{xx}^{W}+\frac{T-1}{T}D\right)+\theta\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)\right]^{-1} \\ &\quad\times\left[\left(\Sigma_{xx}^{W}+\frac{T-1}{T}D\right)\gamma^{FE}+\theta\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)\gamma^{BE}\right]\end{aligned} \tag{44}$$

Note that when the variance of the error term ε_{it} is zero, RE reduces to FE.

To summarize, we have derived the following:

$$\begin{cases}\operatorname{plim}\widehat{\gamma}^{POLS}=\left(\Sigma_{xx}+D\right)^{-1}\Sigma_{xx}\gamma+\left(\Sigma_{xx}+D\right)^{-1}\Sigma_{\mu x} \\ \operatorname{plim}\widehat{\gamma}^{BE}=\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)^{-1}\Sigma_{xx}^{B}\gamma+\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)^{-1}\Sigma_{\mu x} \\ \operatorname{plim}\widehat{\gamma}^{FE}=\left(\Sigma_{xx}^{W}+\frac{T-1}{T}D\right)^{-1}\Sigma_{xx}^{W}\gamma \\ \operatorname{plim}\widehat{\gamma}^{RE}=\left[\left(\Sigma_{xx}^{W}+\frac{T-1}{T}D\right)+\theta\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)\right]^{-1} \\ \qquad\qquad\times\left[\left(\Sigma_{xx}^{W}+\frac{T-1}{T}D\right)\gamma^{FE}+\theta\left(\Sigma_{xx}^{B}+\frac{1}{T}D\right)\gamma^{BE}\right]\end{cases}$$

Appendix 2 - A Simple Univariate Example

To illustrate the effects at play in the presence of both heterogeneity bias and measurement error, we consider the case where x_{it} is unidimensional, and contrast estimation by pooled OLS and FE. Consider the following relationship with a single observed regressor and an intercept term:

$$y_{it}=\gamma_0+\gamma_1 x_{it}+\mu_i+\varepsilon_{it} \tag{45}$$

Suppose that the observed variable x_{it}^{*} incorporates measurement error:

$$x_{it}^{*}=x_{it}+d_{it} \tag{46}$$

where d_{it} is independent of the true x_{it}. The variance of the measurement error and of x_{it} are, respectively, σ_d^2 and σ_x^2.

By estimating (45) using pooled OLS, we get both an omitted variables bias due to the fact that μ_i is potentially correlated with x_{it}, and a measurement error bias due to the correlation between ε_{it} and x_{it}^{*}. The limiting value of the pooled OLS estimate of γ_1 is as follows:

$$\operatorname{plim}\widehat{\gamma}_1^{POLS}=\frac{\gamma_1}{1+\frac{\sigma_d^2}{\sigma_x^2}}+\frac{\operatorname{cov}(x_{it},\mu_i)}{\sigma_x^2+\sigma_d^2} \tag{47}$$

In equation (47), the two sources of bias appear clearly. The variance of measurement error contributes to lessen the extent of heterogeneity bias, as it appears in the denominator of the expression on the right hand side of (47).

Consider now FE estimation, still in the univariate case. To simplify things and without loss of generality, assume that we difference away the time invariant individual effects by taking first differences, rather than by taking differences from country means of the data. The limiting value of the FE estimate of γ_1 is then:

$$\operatorname{plim} \widehat{\gamma}_1^{FE} = \frac{\gamma_1}{1 + \frac{\sigma_{\Delta d}^2}{\sigma_{\Delta x}^2}} \tag{48}$$

where $\sigma_{\Delta d}^2$ is the variance of the first differenced measurement error, and $\sigma_{\Delta x}^2$ is the same for the "true" regressor x_{it}.

We have derived formal expressions for our estimate of interest in two cases. The second method, FE, allows us to remove the heterogeneity bias but will exacerbate measurement error bias. To see why, note that the error-to-truth ratio in the denominator of equation (48) will *always* have increased compared to that under pooled OLS:

$$\sigma_{\Delta d}^2 = \operatorname{var} d_{it} + \operatorname{var} d_{it-\tau} - 2\operatorname{cov}(d_{it}, d_{it-\tau}) = 2\sigma_d^2 \tag{49}$$

$$\sigma_{\Delta x}^2 = \operatorname{var} x_{it} + \operatorname{var} x_{it-\tau} - 2\operatorname{cov}(x_{it}, x_{it-\tau}) = 2\sigma_x^2(1-\rho_x) \tag{50}$$

where $\rho_x = \operatorname{corr}(x_{it}, x_{it-\tau})$ is the autocorrelation of x_{it}. Thus,

$$\frac{\sigma_{\Delta d}^2}{\sigma_{\Delta x}^2} = \frac{\sigma_d^2}{\sigma_x^2(1-\rho_x)} > \frac{\sigma_d^2}{\sigma_x^2} \tag{51}$$

In words, $\sigma_{\Delta x}^2$ will be smaller relative to σ_x^2 the greater the time persistence in x_{it} (i.e. the higher is ρ_x).

We have assumed until now that there was no time persistence in measurement error (i.e. we had white noise errors-in-variables). This assumption is problematic in the context of data used for growth regressions, where errors in measurement from one period are likely to carry over to the next. In the case of autocorrelated measurement error, where we define $\rho_d = \operatorname{corr}(d_{it}, d_{it-\tau})$, the error-to-truth ratio under FE is:

$$\frac{\sigma_{\Delta d}^2}{\sigma_{\Delta x}^2} = \frac{\sigma_d^2(1-\rho_d)}{\sigma_x^2(1-\rho_x)} \tag{52}$$

It is then trivial to show that FE will exacerbate measurement error bias compared to pooled OLS whenever $\rho_d < \rho_x$.

References

[1] Arellano, Manuel and Stephen Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Questions", *Review of Economic Studies*, vol. 58, no.2, pp. 277-297.

[2] Arellano, Manuel and Olympia Bover (1995), "Another Look at the Instrumental Variables Estimation of error-components models", *Journal of Econometrics,* vol. 68, pp. 29-51.

[3] Barro, Robert (1991), "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, vol. 106, no. 2, May, pp. 407-443.

[4] Barro, Robert (1997), *Determinants of Economic Growth*, Cambridge: MIT Press.

[5] Barro, Robert J. and Jong-Wha Lee (2000), "*International Data on Educational Attainment: Updates and Implications*", Center for International Development at Harvard University, *Working Paper* no. 42, April.

[6] Barro, Robert and Sala-i-Martin, Xavier (1995), *Economic Growth*, New York: McGraw Hill.

[7] Baumol, William (1986), "Productivity Growth, Convergence and Welfare: What the Long Run Data Show", *American Economic Review*, vol. 76, pp. 1072-1085.

[8] Benhabib, Jess and Mark Spiegel (1994), "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data", *Journal of Monetary Economics*, vol. 34, pp. 143-173.

[9] Bernanke, Ben S. and R. S. Gürkaynak (2001), "Is Growth Endogenous? Taking Mankiw, Romer, and Weil Seriously", *NBER Macroeconomics Annual.*

[10] Blundell, Richard and Stephen Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, vol. 87, pp. 115-143.

[11] Caselli, Francesco (2001), "Comment on Is Growth Endogenous? Taking Mankiw, Romer, and Weil Seriously", *NBER Macroeconomics Annual.*

[12] Caselli, Francesco, Gerardo Esquivel and Fernando Lefort (1996), "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, vol. 1, no. 3, pp. 363-389.

[13] Cragg, J. G. and S. G. Donald (1993), "Testing Identifiability and Specification in Instrumental Variable Models", *Econometric Theory*, vol. 9, pp. 222-240.

[14] Easterly, William, Norman Loayza and Peter Montiel (1997), "Has Latin America's Post-Reform Growth Been Disappointing?", *Journal of International Economics*, vol. 43, pp. 287-311.

[15] Frankel, Jeffrey A. and David Romer (1999), "Does Trade Cause Growth?", *American Economic Review*, vol. 89, no. 3, June, pp. 379-399.

[16] Griliches, Zvi and Jerry Hausman (1986), "Errors in Variables in Panel Data", *Journal of Econometrics*, vol. 31, no. 1, pp.93-118.

[17] Hall, Robert and Jones, Charles I. (1999), "Why Do Some Countries Produce so Much More Output Per Worker than Others?", *Quarterly Journal of Economics*, vol. 114 no. 1, pp. 83-116, February.

[18] Heston, Alan, Robert Summers and Bettina Aten (2002), *Penn World Table Version 6.1*, Center for International Comparisons at the University of Pennsylvania (CICUP), October.

[19] Islam, Nazrul (1995), "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, vol. 110, no.4, pp. 1127-1170.

[20] Islam, Nazrul (2000), "Small Sample Performance of Dynamic Panel Estimators in Estimating the Growth Convergence Equation: a Monte Carlo Study", *Advances in Econometrics*, vol. 15, pp. 317-339.

[21] Klepper, Steven and Edward Leamer (1984), "Consistent Sets of Estimates for Regressions with Errors in All Variables", *Econometrica*, vol. 52, no.1, pp. 163-184.

[22] Knight, Malcolm, Norman Loayza, Delano Villanueva (1993), "Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach", *International Monetary Fund Staff Papers,* vol. 40, no. 3, September, pp. 512-41.

[23] Levine, Ross, Norman Loayza and Thorsten Beck (2000), "Financial Intermediation and Growth: Causality and Causes", *Journal of Monetary Economics*, vol. 46, pp. 31-77.

[24] Mankiw, N. Gregory, David Romer and David N. Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, vol. 107, no. 2, pp. 407-437.

[25] Solow, Robert (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, vol. 70, no. 1, February, pp. 65-94.

[26] Staiger, D., and J. H. Stock (1997), "Instrumental Variables Regression With Weak Instruments", *Econometrica*, vol. 65, pp. 557–586.

[27] Stock, James H., Jonathan Wright and Motohiro Yogo (2002) "A Survey of Weak Instruments and Weak Identification in GMM", *Journal of Business and Economic Statistics,* vol. 20, no. 4, pp. 518-529.

[28] Stock, James H. and Motohiro Yogo (2003), "Testing for Weak Instruments in Linear IV Regression", forthcoming in D. W. K. Andrews and J. H. Stock, eds., *Festschrift in Honor of Thomas Rothenberg*, Cambridge (UK): Cambridge University Press.

[29] Wacziarg, Romain (2002), "Review of Easterly's The Elusive Quest for Growth", *Journal of Economic Literature*, vol. 40, no. 3, September, pp. 907-918.

Table 1 – Correlation Structure Among Regressors and Fixed-Effects

	$\log s_{k,it-\tau}$	$\log s_{h,it-\tau}$	$\log(n+g+\delta)_{it-\tau}$	$\log y_{it-\tau}$	μ_i
Panel A – Pooled Data					
$\log s_{k,it-\tau}$	1.0000				
$\log s_{h,it-\tau}$	0.6046	1.0000			
$\log(n+g+\delta)_{it-\tau}$	-0.3800	-0.5763	1.0000		
$\log y_{it-\tau}$	0.6220	0.8086	-0.6640	1.0000	
μ_i	0.6248	0.8031	-0.5957	0.9273	1.0000
Panel B – Between Variation					
$\log s_{k,it-\tau}$	1.0000				
$\log s_{h,it-\tau}$	0.7160	1.0000			
$\log(n+g+\delta)_{it-\tau}$	-0.5004	-0.6594	1.0000		
$\log y_{it-\tau}$	0.7107	0.8691	-0.7154	1.0000	
μ_i	0.7096	0.9070	-0.6629	0.9622	1.0000
Panel C – Within Variation					
$\log s_{k,it-\tau}$	1.0000				
$\log s_{h,it-\tau}$	0.2104	1.0000			
$\log(n+g+\delta)_{it-\tau}$	0.0763	-0.2531	1.0000		
$\log y_{it-\tau}$	0.1497	0.5400	-0.3799	1.0000	
μ_i	0.0000	0.0000	0.0000	0.0000	1.0000

Table 2 – Magnitude of Measurement Error on the Underlying Data (based on various values of F)

		Income ($)	Investment Rate (% GDP)	Secondary Enrollment Rate (%)	Population growth (%)
Average Value of Unshocked Variable		4997.15	17.17%	51.40%	1.59%
Average Absolute Value of Shock with F=	1%	617.646	0.92%	4.33%	0.06%
	2.5%	1010.391	1.46%	6.52%	0.10%
	5%	1332.400	2.03%	8.84%	0.14%
	10%	2134.202	2.95%	13.94%	0.20%
	15%	2702.867	3.67%	16.05%	0.25%

Averages computed from simulated data for 2,000 countries in the 8-period panel, i.e. 16,000 pooled observations.

Table 3 – Baseline Simulation Results – Average Estimated Coefficients and Bias (1000 runs)

		(1)		(2)		(3)		(4)	
Error-to-Truth Ratio		**F=0%**		**F=5%**		**F=10%**		**F=15%**	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Avg Coeff**	**Bias (%)**	**Avg Coeff**	**Bias (%)**	**Avg Coeff**	**Bias (%)**
Fixed Effects									
log $s_{k,it-\tau}$	0.099	0.093	-5%	0.026	-74%	0.003	-97%	-0.007	-107%
log $s_{h,it-\tau}$	0.099	0.076	-23%	-0.095	-197%	-0.108	-209%	-0.103	-204%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.287	46%	-0.991	402%	-1.107	461%	-1.069	442%
log $y_{it-\tau}$	0.832	0.787	-5%	0.332	-60%	0.183	-78%	0.110	-87%
Avg. Abs. Bias			20%		183%		211%		210%
Implied λ	3.68%	4.79%	30%	22.07%	500%	33.99%	824%	44.11%	1099%
Between									
log $s_{k,it-\tau}$	0.099	0.081	-18%	0.079	-19%	0.079	-20%	0.078	-21%
log $s_{h,it-\tau}$	0.099	0.114	16%	0.110	12%	0.105	6%	0.100	1%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.031	-84%	-0.033	-83%	-0.027	-86%	-0.031	-84%
log $y_{it-\tau}$	0.832	0.983	18%	0.986	18%	0.990	19%	0.993	19%
Avg. Abs. Bias			34%		33%		33%		31%
Implied λ	3.68%	0.35%	-91%	0.29%	-92%	0.21%	-94%	0.14%	-96%
Random Effects									
log $s_{k,it-\tau}$	0.099	0.110	12%	0.163	65%	0.204	107%	0.234	137%
log $s_{h,it-\tau}$	0.099	0.171	74%	0.268	172%	0.322	226%	0.347	252%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.182	-8%	-0.899	355%	-1.361	590%	-1.666	744%
log $y_{it-\tau}$	0.832	0.924	11%	0.789	-5%	0.699	-16%	0.639	-23%
Avg. Abs. Bias			26%		150%		235%		289%
Implied λ	3.68%	1.58%	-57%	4.75%	29%	7.17%	95%	8.96%	143%
Arellano-Bond									
log $s_{k,it-\tau}$	0.099	0.095	-4%	-0.038	-139%	-0.072	-173%	-0.083	-185%
log $s_{h,it-\tau}$	0.099	0.068	-31%	-0.280	-384%	-0.254	-357%	-0.221	-325%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.243	23%	-0.695	252%	-0.642	225%	-0.553	180%
log $y_{it-\tau}$	0.832	0.795	-4%	0.207	-75%	0.095	-89%	0.046	-94%
Avg. Abs. Bias			16%		213%		211%		196%
Implied λ	3.68%	4.58%	24%	31.52%	756%	47.10%	1180%	61.45%	1570%

Table 4 - Varying Heterogeneity Bias (alternative correlations between μ_i and the regressors) – F = 10%, 1000 runs.

		(1)		(2)		(3)		(4)	
FE Correlation:		C=0%		C=50%		C=75%		C=100%	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**
Fixed Effects									
log $s_{k,it-\tau}$	0.099	0.070	-29%	0.037	-63%	0.022	-78%	0.003	-97%
log $s_{h,it-\tau}$	0.099	0.049	-50%	-0.018	-118%	-0.061	-162%	-0.108	-209%
$\log(n+g+\delta)_{it-\tau}$	-0.197	-0.234	19%	-0.577	192%	-0.845	328%	-1.107	461%
log $y_{it-\tau}$	0.832	0.363	-56%	0.275	-67%	0.228	-73%	0.183	-78%
Avg. Abs. Bias			39%		110%		160%		211%
Implied λ	3.68%	20.27%	451%	25.85%	602%	29.60%	704%	33.99%	824%
Between									
log $s_{k,it-\tau}$	0.099	0.010	-90%	0.027	-73%	0.043	-57%	0.079	-20%
log $s_{h,it-\tau}$	0.099	-0.069	-170%	-0.026	-127%	0.019	-81%	0.105	6%
$\log(n+g+\delta)_{it-\tau}$	-0.197	0.462	-334%	0.349	-277%	0.221	-212%	-0.027	-86%
log $y_{it-\tau}$	0.832	1.118	34%	1.092	31%	1.059	27%	0.990	19%
Avg. Abs. Bias			157%		127%		94%		33%
Implied λ	3.68%	-2.23%	-161%	-1.77%	-148%	-1.14%	-131%	0.21%	-94%
Random Effects									
log $s_{k,it-\tau}$	0.099	0.090	-9%	0.134	36%	0.165	67%	0.204	107%
log $s_{h,it-\tau}$	0.099	0.078	-21%	0.172	75%	0.239	142%	0.322	226%
$\log(n+g+\delta)_{it-\tau}$	-0.197	-0.196	0%	-0.595	201%	-0.927	370%	-1.361	590%
log $y_{it-\tau}$	0.832	0.854	3%	0.802	-4%	0.758	-9%	0.699	-16%
Avg. Abs. Bias			8%		79%		147%		235%
Implied λ	3.68%	3.15%	-14%	4.41%	20%	5.54%	51%	7.17%	95%
Arellano-Bond									
log $s_{k,it-\tau}$	0.099	0.011	-89%	-0.049	-150%	-0.061	-162%	-0.072	-173%
log $s_{h,it-\tau}$	0.099	0.032	-67%	-0.120	-221%	-0.195	-298%	-0.254	-357%
$\log(n+g+\delta)_{it-\tau}$	-0.197	-0.494	150%	-0.613	211%	-0.646	228%	-0.642	225%
log $y_{it-\tau}$	0.832	0.060	-93%	0.003	-100%	0.038	-95%	0.095	-89%
Avg. Abs. Bias			100%		170%		196%		211%
Implied λ	3.68%	56.25%	1429%	114.83%	3020%	65.32%	1675%	47.10%	1180%

Table 5 - Varying Measurement Error on Initial Income, 1000 runs (F=10% on the other variables)

		F_y=0%		F_y=1%		F_y=2.5%		F_y=5%	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**
Fixed Effects									
log $s_{k,it-\tau}$	0.099	0.066	-33%	0.052	-47%	0.035	-64%	0.018	-81%
log $s_{h,it-\tau}$	0.099	0.030	-69%	-0.003	-103%	-0.038	-139%	-0.072	-173%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.140	-29%	-0.371	88%	-0.606	207%	-0.844	328%
log $y_{it-\tau}$	0.832	0.773	-7%	0.632	-24%	0.486	-42%	0.337	-59%
Avg. Abs. Bias			35%		66%		113%		160%
Implied λ	3.68%	5.15%	40%	9.19%	150%	14.42%	292%	21.74%	491%
Between									
log $s_{k,it-\tau}$	0.099	0.078	-21%	0.079	-20%	0.079	-19%	0.079	-20%
log $s_{h,it-\tau}$	0.099	0.107	8%	0.106	7%	0.106	7%	0.105	6%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.031	-84%	-0.031	-84%	-0.029	-85%	-0.034	-83%
log $y_{it-\tau}$	0.832	0.989	19%	0.989	19%	0.988	19%	0.989	19%
Avg. Abs. Bias			33%		33%		33%		32%
Implied λ	3.68%	0.23%	-94%	0.22%	-94%	0.23%	-94%	0.22%	-94%
Random Effects									
log $s_{k,it-\tau}$	0.099	0.104	5%	0.116	18%	0.134	36%	0.160	62%
log $s_{h,it-\tau}$	0.099	0.127	29%	0.151	53%	0.186	89%	0.239	142%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.155	-22%	-0.298	51%	-0.516	161%	-0.850	331%
log $y_{it-\tau}$	0.832	0.951	14%	0.922	11%	0.876	5%	0.808	-3%
Avg. Abs. Bias			17%		33%		73%		135%
Implied λ	3.68%	1.00%	-73%	1.62%	-56%	2.65%	-28%	4.27%	16%
Arellano-Bond									
log $s_{k,it-\tau}$	0.099	0.064	-36%	0.036	-63%	-0.007	-107%	-0.042	-143%
log $s_{h,it-\tau}$	0.099	0.016	-84%	-0.048	-149%	-0.128	-230%	-0.198	-301%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.123	-38%	-0.319	62%	-0.448	127%	-0.556	182%
log $y_{it-\tau}$	0.832	0.770	-7%	0.582	-30%	0.386	-54%	0.223	-73%
Avg. Abs. Bias			41%		76%		130%		175%
Implied λ	3.68%	5.24%	42%	10.81%	194%	19.04%	417%	29.97%	714%

Table 6 – Country-Specific Measurement Error, 1000 runs

Error-to-Truth Ratio:		F=5%		F=10%		F=15%	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**
Fixed Effects							
log $s_{k,it-\tau}$	0.099	0.022	-77%	0.003	-97%	-0.015	-115%
log $s_{h,it-\tau}$	0.099	-0.100	-201%	-0.110	-211%	-0.096	-198%
log$(n+g+\delta)_{it-\tau}$	-0.197	-1.019	417%	-1.106	461%	-1.035	425%
log $y_{it-\tau}$	0.832	0.328	-61%	0.183	-78%	0.110	-87%
Avg. Abs. Bias			189%		212%		206%
Implied λ	3.68%	22.28%	506%	33.93%	822%	44.21%	1101%
Between							
log $s_{k,it-\tau}$	0.099	0.080	-19%	0.079	-20%	0.077	-22%
log $s_{h,it-\tau}$	0.099	0.109	11%	0.104	6%	0.101	2%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.025	-87%	-0.025	-87%	-0.029	-85%
log $y_{it-\tau}$	0.832	0.987	19%	0.990	19%	0.993	19%
Avg. Abs. Bias			34%		33%		32%
Implied λ	3.68%	0.27%	-93%	0.20%	-95%	0.15%	-96%
Random Effects							
log $s_{k,it-\tau}$	0.099	0.163	66%	0.204	107%	0.230	134%
log $s_{h,it-\tau}$	0.099	0.268	172%	0.321	225%	0.350	255%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.893	353%	-1.357	588%	-1.666	744%
log $y_{it-\tau}$	0.832	0.789	-5%	0.700	-16%	0.638	-23%
Avg. Abs. Bias			149%		234%		289%
Implied λ	3.68%	4.73%	29%	7.14%	94%	9.00%	145%
Arellano-Bond							
log $s_{k,it-\tau}$	0.099	-0.045	-145%	-0.071	-172%	-0.091	-192%
log $s_{h,it-\tau}$	0.099	-0.286	-390%	-0.257	-361%	-0.215	-318%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.710	260%	-0.651	230%	-0.526	167%
log $y_{it-\tau}$	0.832	0.200	-76%	0.091	-89%	0.042	-95%
Avg. Abs. Bias			218%		213%		193%
Implied λ	3.68%	32.23%	776%	47.91%	1202%	63.45%	1624%

Table 7 – Autocorrelated Measurement Error (F=10%, 1000 runs)

		ρ_d=50%		ρ_d=75%		ρ_d=90%	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**
Fixed Effects							
log $s_{k,it-\tau}$	0.099	0.029	-71%	0.047	-52%	0.066	-33%
log $s_{h,it-\tau}$	0.099	-0.060	-161%	-0.035	-135%	-0.002	-102%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.818	315%	-0.627	218%	-0.537	172%
log $y_{it-\tau}$	0.832	0.425	-49%	0.538	-35%	0.627	-25%
Avg. Abs. Bias			149%		110%		83%
Implied λ	3.68%	17.13%	366%	12.40%	237%	9.33%	153%
Between							
log $s_{k,it-\tau}$	0.099	0.075	-24%	0.074	-25%	0.072	-27%
log $s_{h,it-\tau}$	0.099	0.094	-5%	0.086	-13%	0.078	-21%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.031	-84%	-0.027	-86%	-0.044	-78%
log $y_{it-\tau}$	0.832	0.998	20%	1.003	21%	1.007	21%
Avg. Abs. Bias			33%		36%		37%
Implied λ	3.68%	0.05%	-99%	-0.06%	-102%	-0.14%	-104%
Random Effects							
log $s_{k,it-\tau}$	0.099	0.157	59%	0.132	33%	0.115	16%
log $s_{h,it-\tau}$	0.099	0.231	134%	0.184	87%	0.151	53%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.826	319%	-0.493	150%	-0.300	52%
log $y_{it-\tau}$	0.832	0.817	-2%	0.880	6%	0.923	11%
Avg. Abs. Bias			129%		69%		33%
Implied λ	3.68%	4.05%	10%	2.56%	-30%	1.60%	-57%
Arellano-Bond							
log $s_{k,it-\tau}$	0.099	-0.017	-118%	0.007	-93%	0.032	-68%
log $s_{h,it-\tau}$	0.099	-0.177	-279%	-0.146	-248%	-0.107	-208%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.367	86%	-0.202	3%	-0.206	5%
log $y_{it-\tau}$	0.832	0.426	-49%	0.537	-35%	0.599	-28%
Avg. Abs. Bias			133%		95%		77%
Implied λ	3.68%	17.09%	364%	12.43%	238%	10.24%	178%

Table 8 - Efficiency Properties of the Estimators (F=10%, 1000 runs)

Variable	Mean	Abs (Bias%)	MSE	t-stat*
Fixed Effects				
log $s_{k,it-\tau}$	0.003	97%	0.0140	0.05
log $s_{h,it-\tau}$	-0.108	209%	0.0463	-1.78
log$(n+g+\delta)_{it-\tau}$	-1.107	461%	1.1530	-1.99
log $y_{it-\tau}$	0.183	78%	0.4241	4.33
Average		211%	1.6373	
Between				
log $s_{k,it-\tau}$	0.079	20%	0.0007	4.92
log $s_{h,it-\tau}$	0.105	6%	0.0003	6.76
log$(n+g+\delta)_{it-\tau}$	-0.027	86%	0.0424	-0.24
log $y_{it-\tau}$	0.990	19%	0.0250	85.10
Average		33%	0.0683	
Random Effects				
log $s_{k,it-\tau}$	0.204	107%	0.0129	4.13
log $s_{h,it-\tau}$	0.322	226%	0.0513	7.87
log$(n+g+\delta)_{it-\tau}$	-1.361	590%	1.4855	-3.67
log $y_{it-\tau}$	0.699	16%	0.0187	23.74
Average		235%	1.5684	
Arellano-Bond				
log $s_{k,it-\tau}$	-0.072	173%	0.0394	-0.77
log $s_{h,it-\tau}$	-0.254	357%	0.1323	-2.78
log$(n+g+\delta)_{it-\tau}$	-0.642	225%	0.8059	-0.86
log $y_{it-\tau}$	0.095	89%	0.5484	1.52
Average		211%	1.5260	

* t-statistics averaged over 1000 runs.

Table 9 – Simulation Results for Alternative Estimators– 1000 runs

Error-to-Truth Ratio:		F=5%		F=10%		F=15%	
Variable	**True Coeffs**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**	**Mean**	**Bias (%)**
Random Effects							
log $s_{k,it-\tau}$	0.099	0.163	65%	0.204	107%	0.234	137%
log $s_{h,it-\tau}$	0.099	0.268	172%	0.322	226%	0.347	252%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.899	355%	-1.361	590%	-1.666	744%
log $y_{it-\tau}$	0.832	0.789	-5%	0.699	-16%	0.639	-23%
Avg. Abs. Bias			150%		235%		289%
Implied λ	3.68%	4.75%	29%	7.17%	95%	8.96%	143%
SUR (Flexible Random Effects)							
log $s_{k,it-\tau}$	0.099	0.133	35%	0.160	62%	0.183	85%
log $s_{h,it-\tau}$	0.099	0.210	113%	0.244	148%	0.268	171%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.531	169%	-0.857	334%	-1.092	454%
log $y_{it-\tau}$	0.832	0.865	4%	0.802	-4%	0.751	-10%
Avg. Abs. Bias			80%		137%		180%
Implied λ	3.68%	2.91%	-21%	4.40%	20%	5.74%	56%
Between							
log $s_{k,it-\tau}$	0.099	0.079	-19%	0.079	-20%	0.078	-21%
log $s_{h,it-\tau}$	0.099	0.110	12%	0.105	6%	0.100	1%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.033	-83%	-0.027	-86%	-0.031	-84%
log $y_{it-\tau}$	0.832	0.986	18%	0.990	19%	0.993	19%
Avg. Abs. Bias			33%		33%		31%
Implied λ	3.68%	0.29%	-92%	0.21%	-94%	0.14%	-96%
Mankiw-Romer-Weil (modified Between)[a]							
log $s_{k,it-\tau}$	0.099	0.087	-12%	0.085	-14%	0.083	-16%
log $s_{h,it-\tau}$	0.099	0.134	35%	0.126	27%	0.121	22%
log$(n+g+\delta)_{it-\tau}$	-0.197	-0.109	-45%	-0.099	-50%	-0.087	-56%
log $y_{it-\tau}$	0.832	0.959	15%	0.966	16%	0.971	17%
Avg. Abs. Bias			27%		27%		28%
Implied λ	3.68%	0.84%	-77%	0.69%	-81%	0.60%	-84%
Arellano-Bover/Blundell-Bond							
log $s_{k,it-\tau}$	0.099	0.179	82%	0.234	137%	0.270	174%
log $s_{h,it-\tau}$	0.099	0.282	186%	0.330	235%	0.352	256%
log$(n+g+\delta)_{it-\tau}$	-0.197	-1.508	664%	-2.200	1015%	-2.515	1175%
log $y_{it-\tau}$	0.832	0.732	-12%	0.622	-25%	0.553	-34%
Avg. Abs. Bias			236%		353%		410%
Implied λ	3.68%	6.23%	69%	9.49%	158%	11.84%	222%

a: MRW estimates adjusted with $\tau=5$ instead of $\tau=40$ to ensure comparability of the point estimates with the other estimators.

Table 10 – Growth Regressions on Actual Data – Dependent Variable: Log Per Capita Income – 1960-1999.

	(1)	(2)	(2')	(3)	(4)	(5)	(6)	(7)
	BE	**MRW**	**MRW adj.**[b]	**RE**	**SUR**	**FE**	**AB**	**BB**
log $s_{k,it-\tau}$	0.058	0.382	0.062	0.065	0.066	0.062	0.072	0.085
	(0.016)**	(0.104)**		(0.012)**	(0.011)**	(0.017)**	(0.024)**	(0.016)**
log $s_{h,it-\tau}$	0.041	0.405	0.066	0.040	0.038	-0.028	-0.060	0.025
	(0.014)**	(0.085)**		(0.011)**	(0.010)**	(0.020)	(0.033)*	(0.016)
$\log(n+g+\delta)_{it-\tau}$	-0.214	-2.173	-0.355	-0.385	-0.284	-0.421	-0.085	-0.617
	(0.111)*	(0.734)**		(0.091)**	(0.078)**	(0.153)**	(0.217)	(0.157)**
log $y_{it-\tau}$	0.960	0.523	0.922	0.940	0.948	0.796	0.795	0.944
	(0.014)**	(0.078)**		(0.011)**	(0.010)**	(0.026)**	(0.042)**	(0.018)**
# of Observations	-	-	-	552	552	552	414	483
Number of Countries	69	69	69	69	69	69	69	69
Implied λ	0.82%	1.62%	1.62%	1.24%	1.07%	4.56%	4.59%	1.15%

Standard errors in parentheses
* significant at 10%; ** significant at 1%
All regressions include time effects (except columns 1 and 2). Regressions include a constant term (not reported), where applicable.
a: Income at the end of each 5-year period in all but columns 2 and 2', where the dependent variable is income in 1999.
b: MRW estimates adjusted with τ=5 instead of τ=40 to ensure comparability of the point estimates with the other columns.