GMM Estimation of
Empirical Growth Models[*]

Stephen Bond
Nuffield College, University of Oxford
and Institute for Fiscal Studies

Anke Hoeffler
St.Antony's College, University of Oxford
and Centre for the Study of African Economies

Jonathan Temple[†]
Department of Economics, University of Bristol,
8 Woodland Road, Bristol BS8 1TN, UK and CEPR

September 12, 2001

[†]Corresponding author. Email jon.temple@bristol.ac.uk

Abstract

This paper highlights a problem in using the first-differenced GMM panel data estimator to estimate cross-country growth regressions. When the time series are persistent, the first-differenced GMM estimator can be poorly behaved, since lagged levels of the series provide only weak instruments for subsequent first-differences. Revisiting the work of Caselli, Esquivel and Lefort (1996), we show that this problem may be serious in practice. We suggest using a more efficient GMM estimator that exploits stationarity restrictions, and this approach is shown to give more reasonable results than first-differenced GMM in our estimation of an empirical growth model.

Keywords: convergence, growth, generalized method of moments, weak instruments

JEL classification: O41, O47.

1 Introduction

The last few years have seen several important advances in the empirical litera-
ture on growth and convergence. There is increasing use of relatively sophisticated
panel data and time series methods, in step with greater awareness of the econo-
metric difficulties facing growth researchers. However, the panel data method that
currently appears to be perceived as the best available, first-differenced general-
ized method of moments (GMM), has its own traps for the unwary. In this paper
we discuss a potentially serious problem with first-differenced GMM in the con-
text of empirical growth models. We also draw attention to an alternative GMM
estimator for dynamic panel data models which appears to give more reasonable
results in this context.

Before we expand on these points at greater length, we discuss the role of es-
timation within empirical growth research, and some of the associated problems.
Mankiw, Romer and Weil (1992) demonstrated that estimation could potentially
cast light on a number of issues. Unfortunately, there are well known problems
with estimating growth regressions. The right-hand-side variables are typically en-
dogenous and measured with error.[1] Another difficulty is that of omitted variables.
One variable that should be included in a conditional convergence regression, the
initial level of efficiency, is not observed. This will imply that least squares param-
eter estimates are biased, since the omitted variable is correlated with one of the
regressors, the initial level of income.

An alternative approach, associated primarily with Klenow and Rodriguez-
Claré (1997) and Hall and Jones (1999), is to carry out accounting decompositions
of differences in output levels. Yet this has problems of its own. By imposing tech-
nology parameters based upon microeconomic evidence, the approach assumes

---

[1]Both problems are well known from the microeconomic literature on estimating production
functions, and are not easily solved. See, for example, Griliches and Mairesse (1998).

away the externalities which have been emphasized in the endogenous growth literature ever since Romer (1986) and Lucas (1988). Furthermore, such exercises may sometimes take the microeconomic evidence too much at face value. For example, the microeconomic estimates of the returns to schooling may be driven by signalling effects, so that accounting decompositions overstate the contribution of education.

It is also worth noting that some of the most interesting research questions cannot be answered by the accounting approach. Accounting decompositions are silent on the growth effects of political stability, the quality of macroeconomic policy, income inequality, financial depth, and so on. Since these questions are likely to be of lasting interest, there is a clear need to develop and apply more rigorous estimation methods. Ideally, these methods should allow, where possible, for endogeneity, measurement error and omitted variables.

One prominent way to address these problems has been through first-differenced generalized method of moments estimators applied to dynamic panel data models. The relevant estimator was originally developed by Holtz-Eakin, Newey and Rosen (1988) and Arellano and Bond (1991).[2] The approach was introduced into the growth literature in the important contribution of Caselli, Esquivel and Lefort (1996), henceforth CEL. Since then, similar techniques have been applied in growth research by Benhabib and Spiegel (1997, 2000), Easterly, Loayza and Montiel (1997), Forbes (2000) and Levine et al. (2000) among others.[3]

We now describe the general form of this approach. The basic idea is to write the regression equation as a dynamic panel data model, take first-differences to remove unobserved time-invariant country-specific effects, and then instrument the right-hand-side variables in the first-differenced equations using levels of the se-

---

[2]Arellano and Bond (1991) also derived associated specification tests.

[3]We should note that the paper by Levine et al. (2000) uses not only first-differenced GMM, but also the system GMM estimator that we evaluate in this paper.

ries lagged two periods or more, under the assumption that the time-varying disturbances in the original levels equations are not serially correlated.

In studying economic growth, this procedure has important advantages over simple cross-section regressions and other estimation methods for dynamic panel data models. First, estimates will no longer be biased by any omitted variables that are constant over time (unobserved country-specific or 'fixed' effects). In conditional convergence regressions, this avoids the problem raised by the omission of initial efficiency. Secondly, as we discuss below, the use of instrumental variables allows parameters to be estimated consistently in models which include endogenous right-hand-side variables, such as investment rates in the context of a growth equation. Finally, again as we discuss below, the use of instruments potentially allows consistent estimation even in the presence of measurement error.

However, there may be a serious drawback with the method adopted by CEL and later researchers. It is now well known that large finite sample biases can occur when instrumental variables are weak, and this difficulty carries over into the GMM estimation of dynamic panel data models.[4] When the time series are persistent and the number of time series observations is small, the first-differenced GMM estimator is poorly behaved. The reason is that, under these conditions, lagged levels of the variables are only weak instruments for subsequent first-differences.

These features are typically present in empirical growth models. Output is a highly persistent series, and to avoid modelling cyclical dynamics, most growth applications consider only a small number of time periods, based on (say) five-year averages. These characteristics might lead us to predict difficulties, and this paper will show that the first-differenced GMM estimator does indeed appear to be problematic in the growth context.

We also demonstrate that more plausible results can be achieved using a system

---

[4]On weak instrument biases, see Nelson and Startz (1990a, 1990b) and Staiger and Stock (1997), among others. For a discussion in the context of panel data, see Blundell and Bond (1998).

GMM estimator suggested by Arellano and Bover (1995) and Blundell and Bond (1998). The system estimator exploits an assumption about the initial conditions to obtain moment conditions that remain informative even for persistent series, and it has been shown to perform well in simulations. The necessary restrictions on the initial conditions are potentially consistent with standard growth frameworks, and appear to be both valid and highly informative in our empirical application. Hence we recommend this system GMM estimator for consideration in subsequent empirical growth research.

The remainder of the paper is organized as follows. In section 2, we describe the first-differenced and system GMM estimators, and explain in more detail why the first-differenced estimator may not be well suited to the study of growth. We also consider the use of GMM in the presence of temporary measurement error and endogenous explanatory variables. In section 3, we set out the growth model to be estimated, and discuss whether the assumptions specific to system GMM are likely to be valid in this context. In section 4, we show that the basic first-differenced GMM estimates appear to be seriously biased, and that the system GMM estimates are more plausible. Finally, section 5 provides a brief summary of our findings and discusses their wider implications.

2   GMM estimators for dynamic panel data models

In this section we briefly review the first-differenced GMM estimator for autoregressive linear regression models estimated from short panels in the presence of unobserved individual-specific time-invariant ('fixed') effects. We explain why large finite sample biases can be expected when the individual series are highly persistent, and suggest how these biases may be detected in practice. We then describe the 'system' GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The basic idea is to estimate a system of equations in

4

both first-differences and levels, where the instruments used in the levels equations are lagged first-differences of the series. These instruments are valid under restrictions on the initial conditions, and later in the paper we will discuss whether or not these restrictions are sensible in the growth context. Finally, the section also considers the extension of the estimators to the cases of temporary measurement error and endogenous regressors.

2.1   First-differenced GMM

We first set out the first-differenced GMM approach. For simplicity, consider an $AR(1)$ model with unobserved individual-specific effects

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \qquad |\alpha| < 1 \qquad (1)$$

for $i = 1, ..., N$ and $t = 2, ..., T$, where $\eta_i + v_{it} = u_{it}$ has the standard error components structure

$$E\left[\eta_i\right] = 0, \quad E\left[v_{it}\right] = 0, \quad E\left[v_{it}\eta_i\right] = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T. \quad (2)$$

We assume that the transient errors are serially uncorrelated

$$E\left[v_{it}v_{is}\right] = 0 \text{ for } i = 1, ..., N \text{ and } s \neq t \qquad (3)$$

and that the initial conditions y_{i1} are predetermined

$$E\left[y_{i1}v_{it}\right] = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T. \qquad (4)$$

Together, these assumptions imply the following $m = 0.5(T-1)(T-2)$ moment restrictions

$$E\left[y_{i,t-s}\Delta v_{it}\right] = 0 \text{ for } t = 3, ..., T \text{ and } s \geq 2 \qquad (5)$$

which can be written more compactly as

$$E\left(Z_i'\Delta v_i\right) = 0 \qquad (6)$$

5

where Z_i is the $(T-2) \times m$ matrix given by

$$
Z_i = \begin{bmatrix}
y_{i1} & 0 & 0 & ... & 0 & ... & 0 \\
0 & y_{i1} & y_{i2} & ... & 0 & ... & 0 \\
. & . & . & ... & . & . & . \\
0 & 0 & 0 & ... & y_{i1} & ... & y_{i,T-2}
\end{bmatrix}
\tag{7}
$$

and Δv_i is the $(T-2)$ vector $(\Delta v_{i3}, \Delta v_{i4}, ..., \Delta v_{iT})'$. These are the moment restrictions exploited by the standard linear first-differenced GMM estimator, implying the use of lagged levels dated $t-2$ and earlier as instruments for the equations in first-differences (cf. Arellano and Bond, 1991). This yields a consistent estimator of α as $N \to \infty$ with T fixed.

However, this first-differenced GMM estimator has been found to have poor finite sample properties, in terms of bias and imprecision, in one important case. This occurs when the lagged levels of the series are only weakly correlated with subsequent first-differences, so that the instruments available for the first-differenced equations are weak (Blundell and Bond 1998). In the $AR(1)$ model of equation (1), this occurs either as the autoregressive parameter (α) approaches unity, or as the variance of the individual effects (η_i) increases relative to the variance of the transient shocks (v_{it}).

Simulation results reported in Blundell and Bond (1998) show that the first-differenced GMM estimator may be subject to a large downward finite-sample bias in these cases, particularly when the number of time periods available is small.[5] This suggests that some caution may be warranted before relying on this method to estimate autoregressive models for a series like per capita GDP from samples containing five or six time periods of five-year averages. It may be that the presence of explanatory variables other than the lagged dependent variable, and more

---

[5]For example, with $T=4$ and $N=100$ and a true value of $\alpha = 0.9$, the distribution of the first-differenced GMM estimator has a mean of 0.23 (with a standard deviation of 0.83) in Table 2(a) of Blundell and Bond (1998).

particularly the inclusion of current or lagged values of these regressors in the instrument set, will improve the behaviour of the first-differenced GMM estimator in particular applications. But some investigation of this in the context of empirical growth models would seem to be in order.

How can we detect whether serious finite sample biases are present? One simple indication can be obtained by comparing the first-differenced GMM results to alternative estimates of the autoregressive parameter α. In the $AR(1)$ model of equation (1), it is well known that OLS levels will give an estimate of α that is biased upwards in the presence of individual-specific effects (see Hsiao, 1986, for example), and that Within Groups will give an estimate of α that is seriously biased downwards in short panels (see Nickell, 1981). Thus a consistent estimate of α can be expected to lie in between the OLS levels and Within Groups estimates.[6] If we observe that the first-differenced GMM estimate is close to or below the Within Groups estimate, it seems likely that the GMM estimate is also biased downwards in our application, perhaps due to weak instruments.

These simple bias results have been extended to models with other regressors only in the special case when all the regressors except the lagged dependent variable are uncorrelated with η_i and strictly exogenous with respect to v_{it}.[7] Nevertheless it may still be useful to compare first-differenced GMM results to those obtained by OLS levels and Within Groups. A finding that the first-differenced GMM estimate of the coefficient on the lagged dependent variable lies close to the corresponding Within Groups parameter estimate can be regarded as a signal that biases due to weak instruments may be important. In these cases, it may be appropriate to investigate the quality of the instruments by studying the reduced form equations for $\Delta y_{i,t-1}$ directly, or to consider alternative estimators that are likely to have better finite sample properties in the context of persistent series.

---

[6]Nerlove (1999a, 2000) has also made this observation in the context of empirical growth models.
[7]See, for example, Sevestre and Trognon (1996).

2.2 System GMM

We now consider one estimator that may have superior finite sample properties. To obtain a linear GMM estimator better suited to estimating autoregressive models with persistent panel data, Blundell and Bond (1998) consider the additional assumption that

$$E\left(\eta_i \Delta y_{i2}\right) = 0 \text{ for } i = 1, ..., N. \tag{8}$$

This assumption requires a stationarity restriction on the initial conditions y_{i1} which is discussed further in the Appendix. Condition (8) holds if the means of the y_{it} series, whilst differing across individuals, are constant through time for periods $1, 2, ..., T$ for each individual. Combined with the $AR(1)$ model set out in equations (1) to (4), this assumption yields $T - 2$ further linear moment conditions

$$E\left(u_{it} \Delta y_{i,t-1}\right) = 0 \text{ for } i = 1, ..., N \text{ and } t = 3, 4, ..., T. \tag{9}$$

These allow the use of lagged first-differences of the series as instruments for equations in levels, as suggested by Arellano and Bover (1995).

We can then construct a GMM estimator which exploits both sets of moment restrictions (5) and (9).[8] This uses a stacked system of $(T - 2)$ equations in first-differences and $(T - 2)$ equations in levels, corresponding to periods $3, ..., T$ for which instruments are observed.[9] The instrument matrix for this system can be written as

$$Z_i^+ = \begin{bmatrix} Z_i & 0 & 0 & ... & 0 \\ 0 & \Delta y_{i2} & 0 & ... & 0 \\ 0 & 0 & \Delta y_{i3} & ... & 0 \\ . & . & . & ... & . \\ 0 & 0 & 0 & ... & \Delta y_{i,T-1} \end{bmatrix}$$

---

[8]The use of further lags beyond $\Delta y_{i,t-1}$ as instruments in the levels equations can be shown to be redundant, given the moment conditions exploited in (5).

[9]In a balanced panel, it would suffice to use the single levels equation for period T, but this system extends less straightforwardly to the case of unbalanced panels.

where Z_i is given by equation (7). The complete set of second-order moment conditions available given assumption (8) can be expressed as

$$E\left(Z_i^{+\prime} u_i^+\right) = 0 \tag{10}$$

where $u_i^+ = (\Delta v_{i3}, ..., \Delta v_{iT}, u_{i3}, ..., u_{iT})'$.

The system GMM estimator thus combines the standard set of equations in first-differences with suitably lagged levels as instruments, with an additional set of equations in levels with suitably lagged first-differences as instruments. Although the levels of y_{it} are necessarily correlated with the individual-specific effects (η_i) given model (1), assumption (8) requires that the first-differences Δy_{it} are not correlated with η_i, permitting lagged first-differences to be used as instruments in the levels equations. As an empirical matter, the validity of these additional instruments can be tested using standard Sargan tests of over-identifying restrictions, or using Difference Sargan or Hausman comparisons between the first-differenced GMM and system GMM results (see Arellano and Bond, 1991).

The calculation of this system GMM estimator is discussed in more detail in Blundell and Bond (1998). They also report evidence from Monte Carlo simulations that compare the finite sample performance of the first-differenced and system GMM estimators. For an $AR(1)$ model, this shows that there can be dramatic reductions in finite sample bias and gains in precision from exploiting these additional moment conditions, in cases where the autoregressive parameter is only weakly identified from the first-differenced equations.[10] Blundell, Bond and Windmeijer (2000) report similar improvements for a model with a lagged dependent variable and additional right-hand-side variables, which is more typical of the equations estimated in the empirical growth literature.

It is worth noting that there are other method-of-moment-type estimators in the

---

[10]In the same experiment described in footnote 5, the distribution of the system GMM estimator has a mean of 0.94 (with a standard deviation of 0.16).

literature that may also perform better than first-differenced GMM in the growth context. Symmetrically normalized first-differenced GMM estimators proposed by Alonso-Borrego and Arellano (1999) have been shown to have smaller finite sample biases than standard first-differenced GMM estimators in situations where the instruments are weak.[11] Non-linear GMM estimators exploiting quadratic moment restrictions of the type

$$E\left(u_{it}\Delta u_{i,t-1}\right)=0 \text{ for } t=4,5,...,T \tag{11}$$

(see Ahn and Schmidt, 1995) are more efficient than first-differenced GMM in model (1) to (4), and may also have better finite sample properties. In our later empirical work, however, we focus on the system GMM estimator, which is asymptotically efficient relative to either of these alternatives provided that assumption (8) is satisfied.

2.3 Temporary measurement error

The preceding sections have explained how first-differenced and system GMM estimators can provide consistent parameter estimates in panel data models with lagged dependent variables and unobserved time-invariant individual-specific effects. We now examine how these methods can allow for transient measurement errors. Note that any permanent additive measurement errors are absorbed into the time-invariant individual effects, and hence also controlled for.

Suppose we wish to estimate the $AR(1)$ specification in (1), but instead of observing the true y_{it} series we observe

$$\widetilde{y}_{it} = y_{it} + m_{it}$$

---

[11]Continuously-updated GMM estimators proposed by Hanson, Heaton and Yaron (1996) and exponential tilting estimators proposed by Imbens, Spady and Johnson (1998) could also be considered in this context.

for $i = 1, ..., N$ and $t = 1, ..., T$, where the measurement error m_{it} is serially uncorrelated

$$E\left[m_{it}m_{is}\right] = 0 \text{ for } i = 1, ..., N \text{ and } s \neq t$$

and uncorrelated with any realizations of the disturbances except the current disturbance v_{it}

$$E\left[m_{it}v_{is}\right] = 0 \text{ for } i = 1, ..., N \text{ and } s \neq t.$$

The empirical model using the observed data is then

$$\begin{aligned}
\widetilde{y}_{it} &= \alpha\widetilde{y}_{i,t-1} + \eta_i + \varepsilon_{it} & |\alpha| < 1 & \qquad (12) \\
\varepsilon_{it} &= v_{it} + m_{it} - \alpha m_{i,t-1}
\end{aligned}$$

for $i = 1, ..., N$ and $t = 2, ..., T$, and the first-differenced equations are

$$\begin{aligned}
\Delta\widetilde{y}_{it} &= \alpha\Delta\widetilde{y}_{i,t-1} + \Delta\varepsilon_{it} & |\alpha| < 1 & \qquad (13) \\
\Delta\varepsilon_{it} &= \Delta v_{it} + \Delta m_{it} - \alpha\Delta m_{i,t-1}
\end{aligned}$$

for $i = 1, ..., N$ and $t = 3, ..., T$.

In this case it is important to notice that the error term ε_{it} in (12) is serially correlated, so that the second lag of the observed series $\widetilde{y}_{i,t-2}$ is no longer a valid instrument for the first-differenced equations in (13). Without further assumptions, this implies that no instruments are available for the first-differenced equation in period $t = 3$, and at least 4 time series observations on the mis-measured series are required to identify the parameter of interest α. Assuming that $T \geq 4$, however, the following moment conditions are available

$$E\left[\widetilde{y}_{i,t-s}\Delta\varepsilon_{it}\right] = 0 \text{ for } t = 4, ..., T \text{ and } s \geq 3,$$

implying the use of lagged levels of the observed series dated $t - 3$ and earlier as instrumental variables for the equations in first-differences.

11

Assume now that $E\left(\eta_i \Delta y_{i2}\right) = 0$ for $i = 1, ..., N$, so that the additional moment conditions for the levels equations discussed in section 2.2 would be available in the absence of measurement error. The first-order moving average serial correlation in ε_{it} again implies that $\Delta \widetilde{y}_{i,t-1}$ is no longer a valid instrument for the equations in levels. However provided the measurement error m_{it} induces no correlation between the observed first-differences $\Delta \widetilde{y}_{it}$ and the individual effects η_i, that is provided

$$E\left[\eta_i \Delta m_{it}\right] = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T,$$

then the following moment conditions are available

$$E\left(\Delta \widetilde{y}_{i,t-2}(\eta_i + \varepsilon_{it})\right) = 0 \text{ for } i = 1, ..., N \text{ and } t = 4, ..., T.$$

Thus suitably lagged first-differences of the observed series can still be used as instrumental variables for the levels equations in the presence of serially uncorrelated measurement error. As before, it is likely that the validity of these additional moment conditions will be crucial to the construction of GMM estimators with good finite sample properties in the context of highly persistent series.

2.4 Endogenous regressors

As a further extension, we now consider a model with an additional right-hand-side variable x_{it}

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \eta_i + v_{it} \qquad |\alpha| < 1 \qquad (14)$$

for $i = 1, ..., N$ and $t = 2, ..., T$, where x_{it} is correlated with η_i and endogenous in the sense that

$$E\left[x_{it} v_{is}\right] \neq 0 \text{ for } i = 1, ..., N \text{ and } s \leq t,$$

which allows both contemporaneous correlation between the current shock v_{it} and x_{it}, and feedbacks from past shocks $v_{i,t-s}$ onto the current value of x_{it}. The er-

12

ror components satisfy the assumptions given in section 2.1 above. Taking first-differences to eliminate the individual effects η_i, the moment conditions

$$E\left[x_{i,t-s}\Delta v_{it}\right] = 0 \text{ for } t = 3, ..., T \text{ and } s \geq 2$$

are available here, in addition to those given in (5). Lagged values of endogenous x_{it} variables dated $t-2$ and earlier can then be used as instruments for the equations in first-differences.[12]

Similarly if in addition to condition (8) from section 2.2 we are willing to assume that first-differences of x_{it} are uncorrelated with the individual-specific effects,

$$E\left[\eta_i \Delta x_{it}\right] = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T,$$

then the following moment conditions are available

$$E\left(\Delta x_{i,t-1} u_{it}\right) = 0 \text{ for } i = 1, ..., N \text{ and } t = 3, ..., T$$

in addition to those given in (9). Suitably lagged first-differences of endogenous x_{it} variables can then be used as instruments for the levels equations.[13]

Finally, we consider the presence of measurement error and endogenous right-hand-side variables combined. One observation is that temporary measurement error in the observed x_{it} series, with the properties outlined in section 2.3 above, will have no consequences for the estimation of model (14). Since we are already allowing for simultaneous correlation between x_{it} and the disturbance here, lagged values of the observed x_{it} (and y_{it}) series dated $t - 2$ and earlier continue to be valid instruments for the first-differenced equations in this case.

---

[12] Additional instruments are available for the equations in first-differences if the x_{it} variables satisfy more restrictive assumptions, for example if they are predetermined with respect to v_{is} (which rules out contemporaneous correlation but not feedbacks from past shocks) or strictly exogenous with respect to v_{is} (which rules out correlation between x_{it} and v_{is} at any dates). See Arellano and Bond (1991) for further discussion of these cases.

[13] Again there may be additional moment conditions available if the x_{it} variables are predetermined or strictly exogenous. See Blundell, Bond and Windmeijer (2000) for further discussion.

More generally, if the model contains a lagged $x_{i,t-1}$ variable that is measured with error or, as before, if the lagged dependent variable is measured with error, this will require period $t-2$ values of the variables measured with error to be omitted from the set of instruments used for the equations in first-differences, and period $t-1$ first-differences of the variables measured with error to be omitted from the set of instruments for the equations in levels. If only $y_{i,t-1}$ is measured with error, this may or may not affect the validity of some x_{is} instruments, depending on whether or not the measurement error in the y_{it} series is correlated with x_{it}. Finally the approach can in principle be extended to allow for low order moving average serial correlation in the measurement errors, which would require only longer lags of the series to be used as the instrumental variables.

The potential for obtaining consistent parameter estimates even in the presence of measurement error and endogenous right-hand-side variables is a considerable strength of the GMM approach in the context of empirical growth research. Whilst there are a number of maximum likelihood[14] and bias-corrected Within Groups[15] estimators that have been proposed for dynamic panel data models, it is far from clear how these are affected by the presence of measurement error and endogenous right-hand-side variables. It should also be noted that whilst different assumptions about the presence of measurement errors and the endogeneity of right-hand-side variables will have implications for the validity of specific instruments, these assumptions can be tested in the GMM framework, for example by the use of Sargan tests of over-identifying restrictions and related tests.

---

[14]See, for example, Bhargava and Sargan (1983), Hsiao (1986), Blundell and Smith (1991), Nerlove (1999b), and Hsiao, Pesaran and Tahmiscioglu (2001).
[15]See Kiviet (1995).

3 Estimating growth models by system GMM

In this section, we briefly set out the Solow growth model to be estimated by system GMM. We go on to discuss whether the assumptions needed to use system GMM are likely to be valid in this context.

The growth equation we wish to estimate has the following form:

$$\Delta y_{it} = \gamma_t + (\alpha - 1)y_{i,t-1} + x'_{it}\beta + \eta_i + v_{it} \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T \quad (15)$$

where Δy_{it} is the log difference in per capita GDP over a five year period, $y_{i,t-1}$ is the logarithm of per capita GDP at the start of that period, and x_{it} is a vector of characteristics measured during, or at the start of, the period. In empirical applications of the Solow model these include the logarithm of the investment rate (s_{it}), and the logarithm of the population growth rate (n_{it}) plus 0.05, where 0.05 represents the sum of a common exogenous rate of technical change (g) and a common depreciation rate (δ). In the augmented Solow model the regressors may also include measures of human capital, such as the logarithm of the secondary-school enrollment rate (enr_{it}).

Among other things, the unobserved country-specific effects (η_i) reflect differences in the initial level of efficiency, whilst the period-specific intercepts (γ_t) capture productivity changes that are common to all countries. Country effects and time effects may also reflect country-specific and period-specific components of measurement errors.

Clearly the above model can be written equivalently as:

$$y_{it} = \gamma_t + \alpha y_{i,t-1} + x'_{it}\beta + \eta_i + v_{it} \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T. \quad (16)$$

We now consider the additional assumptions that are required to estimate this equation by system GMM. Blundell and Bond (2000) consider a similar model without the time effects (γ_t). Similar to the result for the basic $AR(1)$ specification

15

considered in section 2.2, they show that in this case constant means of both the y_{it} and x_{it} series through time for each country would be sufficient for the validity of the moment conditions $E\left(\eta_i \Delta y_{it}\right) = 0$ and $E\left(\eta_i \Delta x_{it}\right) = 0$. This allows for the *levels* of the x_{it} variables (and y_{it}) to be correlated with the unobserved country-specific effects, but permits suitably lagged *first-differences* of x_{it} (and y_{it}) to be used as instruments in the levels equations.

At first sight this condition may not look too promising for the estimation of an empirical growth model. Although stationary means of investment rates and population growth rates are quite consistent with the Solow growth model, constant means of the per capita GDP series clearly are not. Fortunately the inclusion of the time dummies allows for common long-run growth in per capita GDP, consistent with common technical progress, without violating the validity of the additional moment restrictions used by the system GMM estimator.[16] This assumption of common technical progress has been standard in empirical applications of the Solow model since the work of Mankiw, Romer and Weil (1992).

Further, whilst the assumption of constant means in the y_{it} and x_{it} series after conditioning on common time effects is sufficient for the validity of these additional moment conditions exploited by the system GMM estimator, Blundell and Bond (2000) also show that this condition is not necessary. Consider equation (16) in first-differences

$$\Delta y_{it} = \gamma_t - \gamma_{t-1} + \alpha \Delta y_{i,t-1} + \Delta x_{it}'\beta + \Delta v_{it} \text{ for } i = 1,...,N \text{ and } t = 3,...,T.$$

Given $E\left(\eta_i \Delta x_{it}\right) = 0$ for all t, if this process has been generating the per capita GDP series for long enough, prior to our sample period, for any influence of the true start-up conditions to be negligible, then $E\left(\eta_i \Delta y_{it}\right) = 0$ as required. This

---

[16]The inclusion of time dummies is equivalent to transforming the variables into deviations from time means (i.e. the mean across the N individual countries for each period). Thus any arbitrary pattern in the time means is consistent with a constant mean of the transformed series for each country.

will hold even if the means of the x_{it} variables, and hence y_{it}, are not constant, even after removing common time-specific components. The requirement for first-differences of investment rates and population growth rates to be uncorrelated with country-specific effects does not seem unreasonable in the growth context. Note that, if these first-differences were correlated with country-specific effects, this would have implausible long-run implications.

We should stress that the assumption $E(\eta_i \Delta y_{it}) = 0$ does not imply that the country-specific effects play no role in output determination. These effects will be one determinant of the steady-state level of output per efficiency unit of labour, conditional on initial output and other steady-state determinants like investment and population growth. The nature of the assumption is, loosely speaking, that there is no correlation between output growth and the country-specific effect in the absence of conditioning on other variables. Again, such a correlation would tend to have implausible long-run implications.

The brief analysis of this section suggests that it is not unreasonable to consider the system GMM estimator in the context of empirical growth models. It remains to be seen whether the additional instruments that this estimator exploits for equations in levels will prove to be valid and useful in an empirical application, and this will be investigated in the remainder of the paper.

4   Estimating the Solow growth model

We now consider the results of applying GMM to estimation of the Solow and augmented Solow growth models. We use the same data set used by CEL, and will compare our findings with theirs. As in their paper, all variables are expressed as deviations from time means, which eliminates the need for time dummies. Another point worth noting is that CEL appear to have reported the results of a two-step GMM estimator (see Arellano and Bond, 1991). For the special case of spherical

17

disturbances, the one-step and two-step GMM estimators are asymptotically equivalent for the first-differenced estimator. Otherwise the two-step estimator is more efficient, and this is always true for system GMM. Unfortunately, Monte Carlo studies have shown that the efficiency gain is typically small, and that the two-step GMM estimator has the disadvantage of converging to its asymptotic distribution relatively slowly. In finite samples, the asymptotic standard errors associated with the two-step GMM estimators can be seriously biased downwards, and thus form an unreliable guide for inference. With this in mind, we prefer to report the results for the one-step GMM estimators, with standard errors that are not only asymptotically robust to heteroskedasticity but have also been found to be more reliable for finite sample inference (see Blundell and Bond, 1998).

Our results for the basic Solow growth model are reported in Table 1.[17] In this table and those that follow, we use Y_{it} to denote GDP per capita of country i at time t. The first three columns of Table 1 report the results using OLS levels, Within Groups and first-differenced GMM estimators respectively. In the first-differenced and system GMM estimates reported here, both investment rates and population growth rates are treated as potentially endogenous variables.

Although we were not able to replicate the CEL results exactly, our results for the first-differenced GMM estimator are qualitatively similar. The differences between the corresponding coefficient estimates are small relative to their standard errors. In particular, our estimate of the coefficient on initial income in the first-differenced GMM results (-0.537) is very similar to that reported by CEL (-0.473). However, we can see that this point estimate lies *below* the corresponding Within Groups estimate, which itself is likely to be seriously biased downwards in a short panel like this one. CEL suggest that the high rate of convergence implied by first-differenced GMM favours open economy versions of the neoclassical growth

---

[17]All results are computed using the DPD98 software for GAUSS. See Arellano and Bond (1998).

18

model (CEL, p. 381). In contrast, we interpret the results in Table 1 as suggesting that the first-differenced GMM estimate of the coefficient on initial income is likely to be seriously biased, consistent with the known properties of this estimator in the presence of weak instruments.

The fourth column of Table 1 reports the results from using system GMM. Here the estimate of the coefficient on initial income lies comfortably above the corresponding Within Groups estimate, and below the corresponding OLS levels estimate. Neither the basic Sargan test of over-identifying restrictions nor the Difference Sargan test, which focuses on the additional instruments used by the system GMM estimator, detects any problem with instrument validity. These additional instruments therefore seem to be valid and highly informative in this context. Overall, the results suggest that there is indeed a serious finite sample bias problem caused by weak instruments in the first-differenced GMM results, which can be addressed by system GMM. The system GMM estimator also yields a considerable improvement in precision compared to first-differenced GMM.

By treating both the investment rate and the population growth rate as potentially endogenous variables, these estimates already allow for the possibility of a serially uncorrelated measurement error in either of these explanatory variables. In the final column of Table 1 we consider the possibility of a serially uncorrelated measurement error in the per capita GDP series, which would invalidate both the level of this series dated $t-2$ as an instrument for the first-differenced equations, and the first-difference of this series dated $t-1$ as an instrument for the levels equations. The Sargan tests of over-identifying restrictions for the GMM estimators in columns three and four do not indicate a serious problem with the validity of these instrumental variables. Nevertheless the final column reports the results for the system GMM estimator when these instruments are excluded. The estimated coefficients can be seen to be very similar to those in column four, which again

suggests no serious problem resulting from transient measurement error in the per capita GDP series.[18]

The system GMM results indicate a rate of convergence of around 2% a year, which is surprisingly similar to the standard cross-section finding. Importantly, they also indicate that the investment rate has a significant positive effect on the steady state level of per capita GDP, even after controlling for unobserved country-specific effects and allowing for the likely endogeneity of investment.

Table 1 here

Table 2 reports our results for a version of the augmented Solow model, where the logarithm of the secondary-school enrollment rate is included as an additional explanatory variable, as in CEL.[19] It is interesting that the inclusion of the school enrollment variable in the model and the instrument set produces a somewhat more reasonable coefficient on initial income using the first-differenced GMM estimator. This now coincides with the Within Groups estimate, rather than being substantially below it. Nevertheless since Within Groups is itself likely to be seriously biased in a panel with $T = 5$, the system GMM estimates in the final column are again our preferred results. As for the basic Solow model there is no indication of instrument invalidity, and again our results indicate a rate of convergence considerably slower than found by CEL.

Table 2 here

We can illustrate the weak instruments problem with the basic first-differenced

---

[18]Results for the first-differenced GMM estimator with $\ln(Y_{i,t-2})$ omitted from the instrument set were also similar to those reported in column three of Table 1. Further omitting the period $t-2$ values of the investment rate and the population growth rate from the instrument set also made no significant difference to our basic results, although the point estimates became less precise.

[19]School enrollment is measured at the start of each five-year period, and treated as a predetermined variable in the results reported here. Our main findings were robust to alternative treatments. School enrollment was not available for 1985 for the Congo and Switzerland, and the relevant observations are dropped from the sample used in Tables 2 and 3.

20

GMM estimator in another way. Our preferred system GMM estimates in Table 2 suggest that the particular human capital measure used here can be omitted from the specification of the model.[20] This suggests that we may be able to strengthen the instrument set used to estimate the basic Solow growth model in first-differences, by including lags of school enrollment as instruments, and testing their validity.

In Table 3 we report the basic first-differenced and system GMM results, using the slightly smaller sample for which school enrollment is measured. These results are very close to those previously reported in Table 1. In the final column we report the first-differenced GMM results using an extended instrument set, which also includes the lags of school enrollment. These additional instruments do make a substantial difference to the first-differenced GMM results, and illustrate the fragility of first-differenced GMM in this context. It is striking that the extended instrument set produces results which are much closer to system GMM. The implied rate of convergence falls from 14% to 4% a year, and the coefficient on the investment rate becomes significant at conventional levels (it is not in the differenced GMM results in the first column of Table 3). There is, however, some indication that the lags of school enrollment may not be valid instruments in this specification.

These results imply that lagged school enrollment helps to predict growth in per capita GDP in the reduced form equations of the first-differenced estimator, even though current school enrollment may not have a significant effect on the steady state level of per capita GDP after controlling for unobserved country-specific effects, investment and population growth. One possible explanation for these findings is that school enrollment affects growth through the rate of investment.

Table 3 here
_____

[20]The first-differenced GMM results suggest a perverse negative effect of school enrollment, but we have stressed that these estimates are likely to be biased. In any case, it is not particularly sensible to expect school enrollment rates to affect growth almost instantaneously.

Throughout the paper, our preferred results indicate a low convergence rate, in the region of 2% to 4% a year. It is important to qualify this finding, and to acknowledge that there is a great deal of uncertainty in measuring convergence rates. As Nerlove (1999a, 1999b, 2000) has emphasized, much depends on the choice of estimator. This should not be altogether surprising. Allowing for unobserved heterogeneity in the estimation of autoregressive parameters, in a short panel based upon a series as persistent as output, is intrinsically challenging.

As we have shown, some techniques are likely to work poorly in this context, and even our preferred estimates are quite imprecise. They may also be biased. We have emphasized the importance of controlling for unobserved heterogeneity in the intercepts of our empirical growth model (country-specific effects) but there may also be heterogeneity in the slope parameters (Lee, Pesaran and Smith 1997). Heterogeneous slope coefficients would invalidate the use of lagged values of serially correlated regressors as instruments. In principle such misspecification would be detected by our tests of over-identifying restrictions, but we should acknowledge that these tests may not be very powerful in the present context.

Unfortunately it is not possible to allow for unrestricted heterogeneity in both the intercepts and the slope coefficients for all the countries in our data set, without the availability of longer time series. One potentially fruitful line of research would be to develop specifications that allow for some limited heterogeneity in slope coefficients, and to investigate the extent of such heterogeneity using sub-samples of countries where longer time series are available. Work along these lines might give very different results. Our principal aim has not been to present definitive estimates of rates of convergence but, more modestly, to highlight the problems with using first-differenced GMM estimators in estimating empirical growth models.

5 Summary and conclusions

The work of Caselli, Esquivel and Lefort (1996) has been influential in its recommendation of the first-differenced GMM estimator for empirical growth models. In this paper, we have shown that the estimator does not appear to perform well in this context. In particular, we pointed out that the first-differenced GMM estimates of the coefficient on the lagged dependent variable tend to lie below the corresponding Within Groups estimates. This suggests that the first-differenced GMM estimates are seriously biased. One plausible explanation, given the high degree of persistence in output, is that the instruments are weak.

We considered two possible solutions to this problem, which both amount to using more informative sets of instruments. The first solution is to use the system GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). This estimator uses lagged first-differences of the variables as instruments for equations in levels, in combination with the usual approach. These additional instruments are valid under a restriction on the initial conditions which is potentially consistent with the Solow growth framework. In our application, we did not reject the validity of these instruments, and they turn out to be highly informative.[21] The second solution we tried is to strengthen the instrument set used for the equations in first-differences by using other variables that are not included in the model, for example through the use of lags of school enrollment as instruments in estimating the basic Solow model.

In both cases, we found that the estimates of the coefficient on the lagged dependent variable then lie above the Within Groups estimates. We take this as a signal that the system GMM approach is probably preferable in this context, and that earlier results in the literature may be seriously biased due to the weakness of

---

[21]We have found similar results for more general growth specifications. For further applications of system GMM to empirical growth models, see Hoeffler (1998).

the instruments. Our preferred results suggest much slower speeds of convergence than those found by CEL, but confirm the importance to growth of investment rates even after allowing for simultaneity.

We round off with two messages for growth researchers. The first and most important is that, since strengthening the instrument set with outside instruments is usually not an easy task, it may be preferable to use the system GMM estimator rather than the first-differenced estimator in empirical growth work. At the very least, researchers who report the standard first-differenced GMM estimates should probably check their results against those of alternative estimators, as illustrated here.

The use of more sophisticated techniques should not become an end in itself, and our second message relates to what we have learned about growth from the use of system GMM. Previous work, notably that by CEL, found a rapid rate of conditional convergence. This would imply that externalities from capital inputs are relatively unimportant, and that most of the cross-country variation in output arises through differences in total factor productivity. Our work, by indicating a lower rate of convergence, suggests that such a conclusion would be too hasty. There is a great deal of uncertainty in measuring the convergence rate, and one consequence is that significant externalities to physical and human capital should not yet be ruled out.

6   Appendix

In this appendix we elaborate on the nature of assumption (8) in the $AR(1)$ model

$$E\left(\eta_i \Delta y_{i2}\right) = 0 \text{ for } i = 1, ..., N$$

which is a restriction on the initial conditions process generating y_{i1}. To see this, first write y_{i1} without loss of generality as

$$y_{i1} = \left(\frac{\eta_i}{1-\alpha}\right) + e_{i1}. \tag{17}$$

Now consider equation (1) for the first period observed

$$y_{i2} = \alpha y_{i1} + \eta_i + v_{i2}.$$

Subtracting y_{i1} from both sides of this equation

$$\Delta y_{i2} = (\alpha - 1)\, y_{i1} + \eta_i + v_{i2} \tag{18}$$

and using (17) we obtain

$$\begin{aligned} \Delta y_{i2} &= (\alpha - 1)\left(\frac{\eta_i}{1-\alpha}\right) + (\alpha - 1)\, e_{i1} + \eta_i + v_{i2} \\ &= (\alpha - 1)\, e_{i1} + v_{i2}. \end{aligned}$$

Hence given the error components structure in (2), assumption (8) is equivalent to the restriction $E\left(e_{i1}\eta_i\right) = 0$ for $i = 1, ..., N$. A sufficient condition is thus that the initial conditions y_{i1} satisfy the mean stationarity restriction $E\left(y_{i1}|\eta_i\right) = \eta_i/(1-\alpha)$ for each individual. Note that this requires only the first moments of the y_{it} series to be constant, and does not require constant second moments.

References

[1] Ahn, S. C. and P. Schmidt. (1995). "Efficient estimation of models for dynamic panel data." *Journal of Econometrics*, 68, 5-28.

[2] Alonso-Borrego, C. and M. Arellano. (1999). "Symmetrically normalized instrumental-variable estimation using panel data." *Journal of Business and Economic Statistics*, 17(1), 36-49.

[3] Arellano, M. and S. Bond. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58, 277-297.

[4] Arellano, M. and S. Bond. (1998). "Dynamic panel data estimation using DPD98 for GAUSS: a guide for users." Mimeo, Institute for Fiscal Studies, London.

[5] Arellano, M. and O. Bover. (1995). "Another look at the instrumental variable estimation of error-components models." *Journal of Econometrics*, 68, 29-52.

[6] Benhabib, J. and M. M. Spiegel. (1997). "Cross-country growth regressions." *C. V. Starr Center for Applied Economics, NYU, research report 97-20*.

[7] Benhabib, J. and M. M. Spiegel (2000). "The role of financial development in growth and investment." *Journal of Economic Growth*, 5(4), 341-360.

[8] Bhargava, A. and Sargan, J.D. (1983). "Estimating dynamic random effects models from panel data covering short time periods." *Econometrica*, 51, 1635-1659.

[9] Blundell, R. and S. Bond. (1998). "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, 87(1), 115-143.

[10] Blundell, R. and S. Bond. (2000). "GMM estimation with persistent panel data: an application to production functions." *Econometric Reviews*, 19(3), 321-340.

[11] Blundell, R., S. Bond and F. Windmeijer. (2000). "Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator." In B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Elsevier Science.

26

[12] Blundell, R. and R. Smith. (1991). "Initial conditions and efficient estimation in dynamic panel data models." *Annales d'Economie et de Statistique*, 20/21, 109-123.

[13] Caselli, F., G. Esquivel and F. Lefort. (1996). "Reopening the convergence debate: a new look at cross-country growth empirics." *Journal of Economic Growth*, 1, 363-389.

[14] Easterly, W., N. Loayza and P. Montiel. (1997). "Has Latin America's post-reform growth been disappointing?" *Journal of International Economics*, 43(3-4), 287-311.

[15] Forbes, K. (2000). "A reassessment of the relationship between inequality and growth." *American Economic Review*, 90(4), 869-887.

[16] Griliches, Z. and J. Mairesse. (1998). "Production functions: the search for identification". In S. Strom (ed.), *Essays in Honour of Ragnar Frisch*, Econometric Society Monograph Series, Cambridge University Press, Cambridge.

[17] Hall, R. and C. Jones. (1999). "Why do some countries produce so much more output per worker than others?" *Quarterly Journal of Economics*, 114(1), 83-116.

[18] Hansen, L. P., J. Heaton and A. Yaron. (1996). "Finite-sample properties of some alternative GMM estimators.", *Journal of Business and Economic Statistics*, 14(3), 262-280.

[19] Hoeffler, A. (1998). *Econometric studies of growth, convergence and conflicts*. D. Phil. thesis, University of Oxford.

[20] Holtz-Eakin, D., W. Newey and H. Rosen. (1988). "Estimating vector autoregressions with panel data." *Econometrica*, 56, 1371-1395.

[21] Hsiao, C. (1986). *Analysis of panel data*. Cambridge: Cambridge University Press.

[22] Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu. (2001). "Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods." mimeo, University of Cambridge.

[23] Imbens, G. W., R. H. Spady and P. Johnson. (1998). "Information theoretic approaches to inference in moment condition models." *Econometrica*, 66(2), 333-357.

[24] Kiviet, J. (1995). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models." *Journal of Econometrics*, 68(1), 53-78.

[25] Klenow, P. and A. Rodriguez-Clare. (1997). "The neoclassical revival in growth economics: has it gone too far?" *NBER Macroeconomics Annual*, 73-103.

[26] Lee, K., M.H. Pesaran, and R. Smith. (1997). "Growth and convergence in a multi-country empirical stochastic Solow model." *Journal of Applied Econometrics*, 12, 357-392.

[27] Levine, R., N. Loayza and T. Beck (2000). "Financial intermediation and growth: causality and causes." *Journal of Monetary Economics*, 46(1), 31-77.

[28] Lucas, R. E. (1988). "On the mechanics of economic development." *Journal of Monetary Economics*, 22(1), 3-42.

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

[30] Nelson, C. and R. Startz. (1990a). "Some further results on the exact small sample properties of the instrumental variable estimator", *Econometrica*, 58, 967-976.

[31] Nelson, C. and R. Startz. (1990b). "The distribution of the instrumental variable estimator and its t-ratio when the instrument is a poor one", *Journal of Business*, 63(1), S125-S140.

[32] Nerlove, M. (1999a). "Properties of alternative estimators of dynamic panel models: an empirical analysis of cross-country data for the study of economic growth." in C. Hsaio, K. Lahiri, L.-F. Lee and M. H. Pesaran (eds.) *Analysis of panels and limited dependent variable models*. Cambridge University Press, Cambridge, pp. 136-170.

[33] Nerlove, M. (1999b). "Likelihood inference for dynamic panel models." *L'Annales d'Economie et de Statistique*, 55-56, 1-41.

[34] Nerlove, M. (2000). "Growth rate convergence, fact or artifact?" In J. Krishnakumar and E. Ronchetti (eds.) *Panel data econometrics: future directions*. North Holland, Amsterdam, pp. 3-34.

[35] Nickell, S. (1981). "Biases in dynamic models with fixed effects." *Econometrica*, 49, 1417-1426.

[36] Romer, P. (1986). "Increasing returns and long-run growth." *Journal of Political Economy*, 94(5), 1002-1037.

[37] Sevestre, P. and A. Trognon. (1996). "Linear dynamic models." In Matyas, L. and Sevestre, P. (eds.) *The econometrics of panel data: handbook of theory and applications*. Boston: Kluwer Academic Publishers.

[38] Staiger, D. and J. Stock. (1997). "Instrumental variables regression with weak instruments". *Econometrica*, 65, 557-586.

7 Tables

Table 1
Estimation of the Solow growth model
Dependent variable $\Delta \ln Y_{i,t}$

Estimation	OLS	WG	DIF-GMM	SYS-GMM	SYS-GMM
Observations	479	382	382	479	479
$\ln(Y_{i,t-1})$	-0.031	-0.319	-0.537	-0.101	-0.112
	(0.012)	(0.055)	(0.138)	(0.052)	(0.047)
$\ln(s_{it})$	0.089	0.128	0.047	0.188	0.197
	(0.015)	(0.038)	(0.074)	(0.047)	(0.044)
$\ln(n_{it} + g + \delta)$	-0.111	-0.083	-0.169	-0.309	-0.410
	(0.056)	(0.146)	(0.305)	(0.264)	(0.328)
Implied λ	0.006	0.077	0.154	0.021	0.024
	(0.003)	(0.016)	(0.060)	(0.012)	(0.011)
Sargan test	-	-	0.24	0.43	0.69
Dif Sargan test	-	-	-	0.74	0.99

Notes

Standard errors in parentheses. 'WG' is Within Groups estimation. The figures reported for the Sargan test and Difference Sargan test are the p-values for the null hypothesis, valid specification. Difference Sargan tests the additional instruments used by the SYS-GMM estimator.

Instruments used for DIF-GMM (column (iii)) are $ln(Y_{i,t-2})$, $ln(s_{i,t-2})$, $ln(n_{i,t-2} + g + \delta)$ and all further lags.

Additional instruments used for levels equations in SYS-GMM (column (iv)) are $\Delta ln(Y_{i,t-1})$, $\Delta ln(s_{i,t-1})$ and $\Delta ln(n_{i,t-1} + g + \delta)$.

SYS-GMM estimates in column (v) omit $ln(Y_{i,t-2})$ from the instruments used for the first-differenced equations, and replace $\Delta ln(Y_{i,t-1})$ by $\Delta ln(Y_{i,t-2})$ in the instruments used for the levels equations.

Data are for five-year intervals between 1960 and 1985, as used in CEL.

Table 2

Estimation of the augmented Solow growth model

Dependent variable $\Delta \ln Y_{i,t}$

Estimation	OLS	WG	DIF-GMM	SYS-GMM
Observations	477	380	380	477
$\ln(Y_{i,t-1})$	-0.052	-0.323	-0.331	-0.081
	(0.016)	(0.056)	(0.107)	(0.077)
$\ln(s_{it})$	0.080	0.136	0.131	0.187
	(0.015)	(0.038)	(0.056)	(0.044)
$\ln(enr_{it})$	0.030	-0.046	-0.149	-0.018
	(0.015)	(0.027)	(0.064)	(0.046)
$\ln(n_{it} + g + \delta)$	-0.098	-0.025	0.297	-0.295
	(0.053)	(0.143)	(0.370)	(0.293)
Implied λ	0.011	0.078	0.080	0.017
	(0.003)	(0.017)	(0.032)	(0.017)
Sargan test	-	-	0.26	0.25
Dif Sargan test	-	-	-	0.23

Notes

Standard errors in parentheses. 'WG' is Within Groups estimation. The figures reported for the Sargan test and Difference Sargan test are the p-values for the null hypothesis, valid specification. Difference Sargan tests the additional instruments used by the SYS-GMM estimator.

Instruments used for DIF-GMM are $ln(Y_{i,t-2})$, $ln(s_{i,t-2})$, $ln(n_{i,t-2} + g + \delta)$, $ln(enr_{i,t-1})$ and all further lags.

Additional instruments used for levels equations in SYS-GMM are $\Delta ln(Y_{i,t-1})$, $\Delta ln(s_{i,t-1})$, $\Delta ln(n_{i,t-1} + g + \delta)$, and $\Delta ln(enr_{it})$.

Data are for five-year intervals between 1960 and 1985, as used in CEL.

Table 3
Estimation of the Solow growth model, extended instrument set
Dependent variable $\Delta \ln Y_{i,t}$

Estimation	DIF-GMM	SYS-GMM	DIF-GMM
Observations	380	477	380
Instrument set	Basic	Basic	Extended
$\ln(Y_{i,t-1})$	-0.502	-0.096	-0.191
	(0.128)	(0.050)	(0.096)
$\ln(s_{it})$	0.054	0.186	0.135
	(0.071)	(0.046)	(0.054)
$\ln(n_{it} + g + \delta)$	-0.156	-0.281	-0.297
	(0.291)	(0.252)	(0.250)
Implied λ	0.139	0.020	0.042
	(0.051)	(0.011)	(0.024)
Sargan test	0.23	0.45	0.13
Dif Sargan test		0.83	0.16

Notes

Standard errors in parentheses. The figures reported for the Sargan test and Difference Sargan test are the p-values for the null hypothesis, valid specification. Difference Sargan tests the additional instruments used by the SYS-GMM estimator.

Instruments used for DIF-GMM (Basic) are ln($Y_{i,t-2}$), ln($s_{i,t-2}$), ln($n_{i,t-2} + g + \delta$) and all further lags.

Additional instruments used for levels equations in SYS-GMM are Δln($Y_{i,t-1}$), Δln($s_{i,t-1}$), and Δln($n_{i,t-1} + g + \delta$).

Instruments used for DIF-GMM (Extended) are ln($Y_{i,t-2}$), ln($s_{i,t-2}$), ln($n_{i,t-2} + g + \delta$), ln($enr_{i,t-1}$) and all further lags.

Data are for five-year intervals between 1960 and 1985, as used in CEL.