Capital adjustment cost and inconsistency in income based dynamic panel models with fixed effects☆

Parantap Basu[*,a], Keshab Bhattarai[b], Yoseph Getachew[c]

[a] *Durham University Business School, Mill Hill Lane, Durham, DH1 3LB, UK*
[b] *Business School, University of Hull, Cottingham Road, Hull, HU6 7RX, UK*
[c] *Department of Economics, University of Pretoria, 0028, Pretoria, South Africa*

Abstract

After the seminal work of Nickell (1981), a vast literature demonstrates the inconsistency of "conditional convergence" estimator in income based dynamic panel models with fixed effects when the time horizon (T) is short but the sample of countries (N) is large. Less attention is given to the economic root of inconsistency of the fixed effects estimator when T is also large. Using a variant of the Ramsey growth model with long-run adjustment cost of capital, we demonstrate that the fixed effects estimator of such models could be inconsistent when T is large. This inconsistency arises because of the long-run adjustment cost of capital which gives rise to a negative moving average coefficient in the error term. Income convergence will be thus overestimated. We theoretically characterize the order of this inconsistency. Our Monte Carlo simulation demonstrates that the size of the bias is substantial and it is greater in economies with higher capital adjustment costs. We show that the use of instrumental variables that take into account the presence of the negative moving average term in the error will overcome this bias.

Key words:
Dynamic panel model, adjustment cost of capital, inconsistency, Ramsey growth model
JEL Classification: C5. D2. D9. O4. O5

*Corresponding authors
Email address: parantap.basu@durham.ac.uk (Parantap Basu)

1. Introduction

In the empirical economic growth literature, a typical specification of an income based dynamic panel econometric model is (e.g., Barro and Sala-i-Martin, 2004, Ch. 11, pp.462):

$$\ln\left(y_{it}/y_{it-1}\right) = -\left(1 - e^{-\beta}\right)\ln y_{it-1} + \mathbf{x}_{it}\boldsymbol{\lambda} + \eta_i + u_{it} \tag{1}$$

where y_{it} is the income of the ith cross section unit at date t; x_{it} denotes a $1 \times J$ vector of control and interest variables; $\boldsymbol{\lambda}$ is a $J \times 1$ parameter vector; η_i represents the fixed effects (FE) and u_{it} is the random disturbance term. The parameters β and $\boldsymbol{\lambda}$ summarize the list of parameters to be estimated. In the convergence literature, parameter $\beta > 0$ is a crucial parameter of interest because it measures the coefficient of conditional convergence known as *β-convergence*. Such a concept is used to understand the convergence of countries or regions conditional on certain fundamentals (such as having similar savings rate and depreciation cost of capital).[1] A higher value of β indicates a faster rate of conditional convergence among the regions or countries studied. Dynamic panel models with fixed effects have been widely employed in the literature for studying convergence among group of countries (which include the highly influential work of Islam, 1995 and Caselli et al., 1996).[2] At the micro level, there are recent applications to measure productive efficiency (Badunenko and

[1] See, for instance, Barro and Sala-i-Martin (2004, Ch. 11).

[2] See also Ho (2006).

Kumbhakar, 2016).[3]

In his seminal work, Nickell (1981) points out that fixed effects (FE) estimator of dynamic panel models is inconsistent when T is short but N is large but consistent when $T \to \infty$.[4] This is widely known as *Nickell bias* in dynamic panel regression. Our paper revisits this Nickell bias in the context of a stochastic growth model with adjustment cost of capital. The central point of this paper is that Nickell bias in an income based dynamic panel growth regression may not go away even for large waves of data if a capital adjustment cost in the technology is present. The immediate implication is that the convergence coefficient β in an income based FE dynamic panel estimation could be inconsistent. Such an adjustment cost by causing a sluggish adjustment of the capital stock in response to an idiosyncratic productivity shock could give rise to a negative moving average term in the error process resulting in a correlation between the lagged dependent variable and the error term.[5] The empirical evidence abound that such adjustment cost of capital is present (Chirinko, 1993 and Hamermesh and Pfann, 1996).

[3]Lee (2014) uses Bayesian factor and Bayesian model averaging for consistency in estimation and model selection of dynamic panel data models with fixed effects.

[4]A sizable econometric literature investigates the nature of this FE bias and possible remedies in dynamic panel data models. For instance, following the work of Nickell (1981), Kiviet (1995), Judson and Owen (1999) and Hahn and Kuersteiner (2002) examine this bias in short and long dynamic panel FE models. Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998, 2000) and Bond (2002) provide a handy way to correct the FE bias in estimation of parameters in short-time dynamic panel models applying internal instruments. However, these are also criticized as sensitive to "instrumental proliferation" (Roodman, 2009). This will be discussed later.

[5]Pesaran and Smith (1995) point out that inconsistency in the dynamic panel estimator can arise if the lagged dependent variable is correlated with the error term. They also point out that the standard corrections for serial correlation are unlikely to work because of the complexity of the error process. We demonstrate the error process can have a negative MA term due to the existence of an economic fundamental such as capital adjustment cost.

The novelty of our paper lies in identifying an economic fundamental such as capital adjustment cost as a factor contributing to inconsistency of a well known estimator. Although there is a proliferation of econometric literature suggesting possible remedies for inconsistency, little attention is given in the literature in understanding an economic root behind the inconsistency in a dynamic FE regression. Our paper is mainly an attempt in that direction whereas it also suggests solution to the problem.

The issue of inconsistency due to the presence of capital adjustment cost is not entirely uncommon in the literature. Caballero (1994) establishes that the ordinary least squares (OLS) estimate of wealth elasticity to adjustment cost generally tends to be inconsistent in a small sample when capital adjustment cost is present in the model. He does not, however, explore the implication of capital adjustment cost for FE bias. To the best of our knowledge, our paper is the first attempt in the literature to understand the role of capital adjustment cost as an economic fundamental driving the inconsistency of convergence estimator in dynamic panel models with long time horizon.

To demonstrate our key point, we develop a standard Ramsey growth model with a convex capital adjustment cost function. Such an adjustment cost function means a rising marginal cost of investment. While diminishing returns to capital facilitate the process of convergence, a rising marginal cost of investment schedule slows it down. The usual dynamic panel FE regression model fails to factor into this capital adjustment cost and thus it overestimates the rate of convergence. We show this formally by establishing that the income based dynamic models has a negative first order moving average error when capital adjustment cost is present. A negative correlation between the lagged dependent variable, $\ln y_{it-1}$, and the disturbance terms u_{it} in (1) makes the convergence coefficient inconsistent. In a multivariate context, we show that the efficiency of all coefficient estimators are affected by this inconsistency when the adjustment cost of capital is present.

We derive an analytical expression for the inconsistency employing the Ramsey growth model with a capital adjustment cost technology. A well known parametric form for the adjustment cost function is borrowed from Lucas and Prescott (1971) that was subsequently used by Basu (1987), Hercowitz and Sampson (1991) and Basu et al. (2012). Such a capital adjustment cost differs from the investment adjustment cost (e.g. Christiano et al., 2005) in the sense that the adjustment cost persists even in the long-run. This explains why inconsistency arises even for infinite time horizon ($T \rightarrow \infty$). Our Monte Carlo experiment shows that inconsistency increases with the degree of the adjustment cost of capital and it is quantitatively substantial. We suggest the use of appropriate internal instrumental variables (IV) that take into account the presence of the negative moving average term in the error to remedy this bias.

In the next section, we develop a Ramsey-type growth model with heterogeneous countries in terms of initial wealth, tastes and productivity to characterize the inconsistency due to the adjustment cost. Section 3 reports a Monte Carlo simulation to demonstrate the sensitivity of the FE bias to the adjustment cost. Section 4 discusses the possible ways to remove this bias. Section 5 concludes.

2. The model

2.1. Preference and technology

Consider a sequence of infinitely-lived heterogenous representative citizens for each country, $i = 1, 2, ..., N$ and $t = 1, 2, ..., \infty$ where i stands for the country representative[6] and t stands for time. Countries are heterogenous in terms of (i) initial capital stock (k_{i0}), (ii) preference (discount factor, ρ_i) and (iii) idiosyncratic productivity shock (ξ_{it}).[7] We let the preference parameter ρ_i vary across countries which could give rise to country specific fixed effects. Also, assume that cross country productivity shocks are i.i.d. Households are further assumed to be both consumers and entrepreneurs.[8]

The production function facing the ith country representative resident is Cobb-Douglas with constant returns to scale as follows,

$$q_{it} = \xi_{it} \left(\prod_{j=1}^{J} (g_{ijt})^{\chi_j} \right) (k_{it})^{\omega} (m_{it})^{\varphi} \tag{2}$$

$$\varphi + \omega + \sum_{j=1}^{J} \chi_j = 1 \tag{3}$$

where q_{it} is the gross output of the ith country, k_{it} is the country's capital stock at period t and k_{i0} is given. g_{ijt} represents the jth exogenous input in the production function (such as infrastructure or a learning-by-doing knowledge spillover that could potentially give rise to technological externality, in the spirit of Arrow, 1962)

[6] Alternatively, i could represent a country in the world economy whereas each country is represented with a single representative consumer, as in Acemoglu and Ventura (2002).

[7] Variables with(out) subscript i represent individual (economy-wide) values. Variables without subscripts t and i represent economy-wide steady-state values.

[8] See Angeletos and Calvet (2006) for a similar type of entrepreneurship.

that is specific to the ith country condition. In addition, m_{it} is a flow of imported intermediate inputs that the country finances by borrowing from the international credit market at a fixed interest rate, r^*. The ith country agent treats input g_{ijt} as given while choosing consumption and investment. The production technology thus exhibits private diminishing returns to reproducible input k_{it}, imported intermediate input m_{it} and the exogenous inputs g_{ijt} but aggregate constant returns to scale, (similar to Romer, 1986 and Barro, 1990).[9]

The ith country borrows m_{it} at the start of each period and fully pays off the loan with interest rate at the end of each period. The optimal purchase of imported intermediate input thus satisfies the condition:

$$\partial q_{it}/\partial m_{it} = 1 + r^* \tag{4}$$

which gives rise to the following demand function for intermediate inputs,

$$m_{it} = [\varphi/1 + r^*]^{1/(1-\varphi)} \left(\prod_{j=1}^{J} (g_{ijt})^{\chi_j/(1-\varphi)}\right) (k_{it})^{\omega/(1-\varphi)} (\xi_{it})^{1/(1-\varphi)} \tag{5}$$

which upon plugging into (2) and after netting out the loan retirement cost, $(1 + r^*)\, m_{it}$, gives the net value added (y_{it}),

$$y_{it} = \epsilon_{it} \left(\prod_{j=1}^{J} (g_{ijt})^{\lambda_j}\right) (k_{it})^{\alpha} \tag{6}$$

[9] Barro (1990) models the production function at the individual firm level as a function of private and public capital.

where

$$\begin{aligned}
\epsilon_{it} &\equiv (1-\varphi)\left(\varphi/\left(1+r^{*}\right)\right)^{\varphi/(1-\varphi)}\left(\xi_{it}\right)^{1/(1-\varphi)} \\
\lambda_j &\equiv \chi_j/\left(1-\varphi\right) \\
\alpha &\equiv \omega/\left(1-\varphi\right)
\end{aligned}$$

The ith country agent maximizes her utility in accordance to the utility function, with a subjective discount factor ρ_i:

$$\mathrm{E}_0\left[\textstyle\sum_{t=0}^{\infty}\left(\rho_i\right)^t \ln c_{it}\right];\ \rho_i < 1 \tag{7}$$

subject to the budget constraint,

$$c_{it} + s_{it} = y_{it} \tag{8}$$

where c_{it} and s_{it} represent consumption and saving, respectively.

Following Lucas and Prescott (1971), Basu (1987) and Basu et al. (2012), the investment technology is given by the following specification:

$$k_{it+1} = k_{it}\left(1-\delta+s_{it}/k_{it}\right)^{\theta} \tag{9}$$

where $\delta \in (0,1)$ and $\theta \in (0,1)$ are rate of depreciation and degree of adjustment cost of capital (k_{it}), respectively. If $\theta = 0$, adjustment cost of capital is prohibitively high to change the capital stock. However, if $\theta = 1$, adjustment cost of capital is zero and we obtain a standard linear depreciation rule. We focus on such capital adjustment costs between these two extremes costs because it has important implications for the reduced form process for the per capita income.

Supposing capital depreciates fully each period, we may rewrite (9) as:[10]

$$k_{it+1} = k_{it} \left(s_{it}/k_{it} \right)^{\theta} \tag{10}$$

Applying standard methods of undetermined coefficient, the optimal policy functions for the ith agent are simplified as follows, (see Appendix A for details of the derivation),[11]

$$c_{it} = \left(1 - \psi_i \right) y_{it} \tag{11a}$$

$$s_{it} = \psi_i y_{it} \tag{11b}$$

where

$$\psi_i \equiv \theta \alpha \rho_i / \left(1 - \rho_i \left(1 - \theta \right) \right) \tag{12}$$

After substituting (6) and (11b) into (10), the optimal dynamic equation of capital stock of the ith country resident is given by,

$$k_{it+1} = \left(\psi_i \right)^{\theta} \left(k_{it} \right)^{\gamma} \left(\prod_{j=1}^{J} \left(g_{ijt} \right)^{\lambda_j} \right)^{\theta} \left(\epsilon_{it} \right)^{\theta} \tag{13}$$

[10] We assume complete depreciation of capital for analytical tractability, without loss of generality. Basu and Getachew (2015) show that depreciation cost has a trivial effect on convergence property.

[11] See also, Basu (1987) and Hercowitz and Sampson (1991) for a similar closed form solution.

where $\gamma \equiv 1 - (1-\alpha)\theta$. Therefore, the optimal capital stock in period $t+1$ for country i is a function of the country's current capital stock (k_{it}), the idiosyncratic shock (ϵ_{it}), time-dependent country specific exogenous factors (g_{ijt}) and a time-independent country specific factor (ψ_i), and adjustment cost of capital θ.

2.2. Role of long-run adjustment cost in determining the bias

According to (13), the adjustment cost of capital ($\theta \neq 1$) impacts not only the dynamics of capital at the individual country level but also the steady-state capital.[12] In this respect, it differs from investment adjustment cost as in Christiano et al. (2005) which do not have such long run effects.[13] In particular, we see below the long-run variance of the capital stock is given by $\upsilon^2\theta^2/(1-\gamma^2)$ where $\upsilon^2 = \text{var}[\ln \epsilon_{it}]$, which depend on the adjustment cost parameter θ.[14] Thus the adjustment cost does not disappear in the long-run when ($\theta \neq 1$).

In the present context, this long-lasting nature of the adjustment cost is particularly reflected on its effect on the idiosyncratic shock. First, this idiosyncratic shock forms the disturbance term (u_{it}) in an estimation equation (1). Second, the shock relates to country's contemporaneous income (6), which appears as a lagged variable in dynamic panel regression models. Therefore, such effects of the idiosyncratic shock will manifest as a source of inconsistency in the estimate of the lagged income in (1).

[12] Note that in the present model, the country specific fixed effect arises solely due to differences in the taste parameter ρ_i. A more general specification can allow for differences in technology which we do not pursue here.

[13] See Groth and Khan (2010) for a specification of a general adjustment cost function which nests capital and investment adjustment cost.

[14] Refer to Corollary 1 below for details of the derivation of the variance.

Based on the production function (6), the log of income of the ith country at date t is given by,

$$\begin{aligned} \ln y_{it} &= \alpha \ln k_{it} + \textstyle\sum_{j=1}^{J} (\lambda_j \ln g_{ijt}) + \ln \epsilon_{it} \\ &= \alpha \ln k_{it} + \mathbf{g}_{it}\boldsymbol{\lambda} + \ln \epsilon_{it} \end{aligned} \tag{14}$$

where $\mathbf{g}_{it}$ is $1 \times J$ vector of (exogenous) regressors, $\mathbf{g}_{it} \equiv (\ln g_{i1t}, \ln g_{i2t}, ..., \ln g_{iJt})$ and $\boldsymbol{\lambda} \equiv (\lambda_1, \lambda_2, ..., \lambda_J)'$ is a $J \times 1$.

Finally, applying (6) and (13) to (14), we obtain the following representation for the dynamic panel model:

$$\ln y_{it} = \gamma \ln y_{it-1} + \textstyle\sum_{j=1}^{J} \lambda_j x_{ijt} + \eta_i + u_{it} \tag{15}$$

where $\eta_i \equiv \alpha\theta \ln \psi_i$ and,

$$x_{ijt} \equiv (\theta - 1) \ln g_{ijt-1} + \ln g_{ijt} \tag{16a}$$

$$u_{it} = \ln \epsilon_{it} - (1 - \theta) \ln \epsilon_{it-1} \tag{16b}$$

In vector form,

$$\ln y_{it} = \gamma \ln y_{it-1} + \mathbf{x}_{it}\boldsymbol{\lambda} + \eta_i + u_{it} \tag{17}$$

where $\mathbf{x}_{it} \equiv (x_{i1t}, x_{i2t}, ..., x_{iJt})$ is a $1 \times J$, $\boldsymbol{\lambda} \equiv (\lambda_1, \lambda_2, ..., \lambda_J)'$ as x_{ijt} is defined in (16a) while u_{it} is given by (16b). By virtue of (1), one can write $-(1-\gamma) = -\left(1 - e^{-\beta}\right)$. Then, using the linear approximation that $e^{-\beta} \approx 1 - \beta$ one gets:

$$1 - \gamma = \beta \tag{18}$$

Thus, the higher the value of γ, the slower the convergence.

Eq. (17) is an autoregressive moving average (ARMA) evolution of the income of the ith country which looks similar to (1). It represents the true specification for the estimation model of conditional convergence based on the Ramsey growth model with a non-zero long-run adjustment cost of capital. Such adjustment cost can be seen as a permanent tax on capital imposed by the mother nature. The long-run impact of such idiosyncratic shock (13) is responsible for the negative moving average disturbance term (u_{it}) in (17).

The income based dynamic panel regression model thus involves an error term which is negatively correlated with the lagged dependent variable ($\ln y_{it-1}$). Therefore, the dynamic panel estimators with FE γ are inconsistent, even when $T \to \infty$.

To see the order of inconsistency involved in the lagged income term, set $\lambda_j = 0$ for all j to simplify exposition. Then, the following Proposition and Corollary can be stated for the univariate case:

Proposition 1. *The inconsistency from the dynamic panel model with FE estimator $(\widehat{\gamma})$ of the coefficient γ with respect to (17), when $\forall j\ \lambda_j = 0$, is given by:*

$$p\lim\left(\widehat{\gamma}\right) = \gamma + \left[\left(\theta - 1\right)\upsilon^2 / \,\mathrm{E}_i \operatorname{var}\left\{\ln y_{it-1}\right\}\right] \tag{19}$$

where $\upsilon^2 = var(\epsilon_{it})$.

Proof. See Appendix B. ■

Corollary 1. *The size of the bias is given by*

$$\Phi \equiv \frac{(1-\gamma^2)(1-\theta)}{\alpha^2\theta^2 + (1-\gamma^2)}. \tag{20}$$

Proof. Using (14), when $\forall j\ \lambda_j = 0$, we can rewrite the denominator in (19) as,

$$\begin{aligned} \mathrm{E}_i \operatorname{var}[\ln y_{it-1}] &= \mathrm{E}_i \operatorname{var}[\alpha \ln k_{it-1} + \ln \epsilon_{it-1}] \\ &= \alpha^2\, \mathrm{E}_i \operatorname{var}[\ln k_{it-1}] + \upsilon^2 \end{aligned} \tag{21}$$

Then, from (13),

$$\begin{aligned} \mathrm{E}_i \operatorname{var}[\ln k_{it-1}] &= \mathrm{E}_i \operatorname{var}[\ln \eta_i + \gamma \ln k_{it-2} + \theta \ln \epsilon_{it-2}] \\ &= \gamma^2\, \mathrm{E}_i \operatorname{var}[\ln k_{it-2}] + \theta^2\upsilon^2 \end{aligned} \tag{22}$$

since η_i is fixed over time and $\operatorname{cov}(\ln k_{it}, \ln \epsilon_{it}) = 0$. Next note from (13) that for a generic ith country agent, $\lim_{t\to\infty} \operatorname{var}[\ln k_{it-1}] = \operatorname{var}[\ln k_{it-2}] = \upsilon^2\theta^2/(1-\gamma^2)$ because $0 < \gamma < 1$.[15] Since α, υ^2 and θ are the same for all i, all agents converge to the same variance of the capital stock which implies

$$\mathrm{E}_i \operatorname{var}[\ln k_{it-1}] = \upsilon^2\theta^2/\left(1-\gamma^2\right) \tag{23}$$

[15] To see this, rewrite (22) as:

$$\begin{aligned} \mathrm{E}_i \operatorname{var}[\ln k_{it}] &= \gamma^{2T}\, \mathrm{E}_i \operatorname{var}[\ln k_{i0}] + \theta^2\upsilon^2 \textstyle\sum_{t=0}^{T} \gamma^{2t} \\ &= \gamma^{2T}\, \mathrm{E}_i \operatorname{var}[\ln k_{i0}] + \theta^2\upsilon^2 \left(1-\gamma^{2T+2}\right)/\left(1-\gamma^2\right) \end{aligned}$$

As $T \to \infty$, the terms in the right hand side converge to $\upsilon^2\theta^2/\left(1-\gamma^2\right)$.

Substitute (21) into (19) after substituting (23) into the former to derive the closed form solution for the degree of inconsistency. ■

Thus in the presence of capital adjustment cost in a growth model, the panel estimator with FE of γ in (1) is inconsistent regardless of the time dimension of the panel. This overestimates the conditional convergence. It is straightforward to verify that this bias is greater in economies with a lower value of θ meaning a higher adjustment cost. The inconsistency of the FE estimator is absent if there is no capital adjustment cost ($\theta = 1$).

The inconsistency of the FE estimator arises due to a negative contemporaneous correlation between $\ln y_{it-1}$ and $\ln \epsilon_{it-1}$. The size of this correlation is proportional to the degree of adjustment cost $(1 - \theta)$. To get the (economic) intuition further for such a negative correlation, let the ith country experience a positive TFP shock ($\Delta \ln \epsilon_{it-1}\%$) at date $t-1$. The optimal investment rule (11b) dictates that the ith country resident's contemporaneous investment rises by the same percent because the elasticity of s_{it-1} with respect to ϵ_{it-1} is unity. Such a blip in investment ($\Delta \ln s_{it-1}$) increases the current capital stock ($\ln k_{it}$) by only $\theta\%$ (see 10). The remaining $(1-\theta)\%$ of the investment is lost due to the presence of the long-run capital adjustment cost. This loss enters the error term in (17) with a negative coefficient ($-(1-\theta) \ln \epsilon_{it-1}$). The standard regression equation (1) ignores this negative correlation between $\ln y_{it-1}$ and $\ln \epsilon_{it-1}$. As a result, the estimate of γ will be inconsistent and, hence, the "conditional convergence" will be overestimated.

In the multivariate case where $\exists j \ \lambda_j \neq 0$, the inconsistency in FE panel estimator affects the estimators of all variables due to a correlation between lagged output and exogenous technological variables. Proposition 2 below demonstrates this.

Proposition 2. *The inconsistency from the dynamic panel model with FE estimator of $\widehat{\gamma}$ and $\widehat{\boldsymbol{\lambda}}$ of the parameters γ and λ with respect to (17) are given by, when $\exists j$ $\lambda_j \neq 0$:*

$$p\lim\left(\widehat{\widetilde{\boldsymbol{\lambda}}}\right) = \widetilde{\boldsymbol{\lambda}} + \left(\mathrm{E}_i\,\mathrm{E}\left[\mathbf{b}_{it}'\mathbf{b}_{it}\right]\right)^{-1}\mathbf{p}(\theta-1)\upsilon^2 \tag{24}$$

where $\mathbf{b}_{it} \equiv \widetilde{\mathbf{x}}_{it} - \widetilde{\mathbf{x}}_i$, $\widetilde{\boldsymbol{\lambda}} \equiv (\gamma, \boldsymbol{\lambda}')'$, $\mathbf{p} \equiv (1, \mathbf{0})'$ and $\mathbf{0}$ is $1 \times J$ zero vector.

Proof. See Appendix C. ∎

The last term in (24) is different from zero with a non-zero adjustment cost of capital ($\theta \neq 1$). Therefore, in the multivariate case the efficiency of all coefficient estimators (17) are affected by the inconsistency when the adjustment cost of capital is present. This happens because the lagged income is correlated with the exogenous variables in the production function. Imbens and Lancaster (1994) had illustrated substantial gains in efficiency in panel estimation from use of marginal information in micro and macro data but they were not explicit about the role of adjustment costs like this.

3. Simulation

In this section, we report the results of a Monte Carlo simulation based on our ARMA(1, 1) specification of the income process (15) to ascertain the quantitative magnitude of the bias resulting from the capital adjustment. As in our model income process (17), we allow the subjective discount factor ρ_i to vary across countries in the range $[0.9, 0.99]$ which is the source of the fixed effect. All other structural parameters are assumed to be the same for all countries. The capital elasticity parameter α is fixed at 0.9 which is higher than the conventional level of 0.36 with a view to target a plausible rate of convergence. The higher value of α is not unreasonable in an open economy context given the fact that α equals $\omega/(1-\varphi)$ which could be higher if the share φ of foreign intermediate input is higher. In addition, Romer (1986) alludes to a higher capital share estimate in view of the broad based nature of capital that includes knowledge. The adjustment cost parameter is first fixed at a baseline level of 0.2 which gives rise to a γ equal to 0.98 meaning a 2% conditional convergence (see 18) which is in accord with Barro and Sala-i-Martin (2004).

Table 1: Sensitivity of the bias to alternative θ values

θ	γ	$p\lim(\widehat{\gamma})$	Bias
0.2	0.98	0.5458	0.4342
0.4	0.96	0.7329	0.2271
0.6	0.94	0.8253	0.1147
0.8	0.92	0.8738	0.0462
1.00	0.90	0.90	0.0000

For a fixed i, we take 10000 draws of TFP (ϵ_{it}) from an i.i.d. lognormal distribution and pass it through the true ARMA (1, 1) process (17) for income (normalizing the initial income at the unit level) to generate draws of log y_{it}. We take 1000 draws of ρ_i from a rectangular distribution with the support [0.9, 0.99] to capture the fixed effects. The pooled estimator (B.8) as shown in the appendix is then computed for sufficiently large T and N. The bias (the difference between γ and $\widehat{\gamma}$) is about 0.44 for the baseline case. The failure to include the adjustment cost could potentially give rise to an overestimation of convergence by 44%. Table 1 illustrates the sensitivity of the bias to alternative adjustment cost parameter values. The bias decreases in economies with lower adjustment cost (higher θ). For no adjustment cost scenario, the bias nearly disappears.

4. Overcoming the bias

The key issue raised in this paper is that the well known Nickell bias in the income convergence regression does not disappear for large waves of data if there is a sizable capital adjustment cost. The root of this bias in the context of our growth model is that the long run capital adjustment cost gives rise to a negative moving average coefficient in the error term of the transformed dynamic panel model (B.4). This bias can be eliminated by proper use of IV. We show in Appendix D using the second distant lag of dependent variables, $a_{it-2} = \ln y_{it-2} - \ln y_{i-2}$,[16] this bias is eliminated given the moment conditions:

$$E(\ln y_{it-2} \ln \epsilon_{it}) = E(\ln y_{it-2} \ln \epsilon_{it-1}) = 0$$

Our procedure of correcting the bias by using distant lag of dependent variables as IV is similar in spirit to Anderson and Hsiao (1981) and Arellano and Bond (1991) with an important difference. This literature applies first difference while we use orthogonal deviation as in Nickell (1981). Our analysis can be extended to first differencing in which case we need to use longer distant lags of the dependent variable as internal IVs to satisfy the moment conditions. Doing so, however, one loses more degrees of freedom. Moreover, if there are missing data, first differencing

[16] $\ln y_{i-2}$ is the mean of the variable over time and formally is defined: $\ln y_{i-2} \equiv (T-2)^{-1} \sum_{t=3}^{T} \ln y_{it-2}$.

becomes more problematic.

Arellano and Bond (1991) use internal instruments which are built from past observations of IVs to deal with this type of inconsistency and other endogeneity problems related to the rest of the covariates. Under the assumption of sequential exogeneity, and no serial correlation in the error term, they instrument $\Delta \ln y_{it-1} = \ln y_{it-1} - \ln y_{it-2}$ with distant lags of the dependent variable starting from $\ln y_{it-3}$. In our case, due to the MA(1) error term, $\Delta \ln y_{it-1}$ should be instrumented with distant lags of the dependent variable starting from $\ln y_{it-4}$.

It is well known that the Difference generalized moment method (GMM) estimator might still be inconsistent if the time series are persistent. In this case, the lagged levels of the variables are weak instruments for the subsequent first differences. Thus, our procedure could still result in biased coefficients with large asymptotic variances. To deal with these problems, Arellano and Bover (1995), and Blundell and Bond (1998) developed System GMM – estimating the equation of interest simultaneously in differences as well as levels, in which case the two equations have been distinctly instrumented. While the instruments for the difference equations are the same as above, the instruments for the level equations become the lagged difference of the corresponding variables. The application of these methods in our present growth regressions should take into consideration the presence of the MA(1) error term which means that one should choose the internal instruments at least a period further back than normally recommended.

Standard assumptions for consistency of the GMM estimators are related to the validity of the instruments and whether the error terms are serially correlated.[17] One should also test for the validity of instruments using standard diagnostic tests such as the Sargan test of over-identifying restrictions or the Hanson's J-statistics and the test for second-order serial correlation can also be easily conducted during GMM estimation. The weak-instrumentation can also be tested using Cragg and Donald Wald F test and Kleibergen-Paap rk Wald F stat.[18] Another important concern, emphasized by Roodman (2009) was "instrument proliferation" which requires reducing the number of instruments, particularly, in System GMM.

5. Conclusion

The economic fundamentals that could generate inconsistency in models have rarely received any attention in macroeconometrics literature. Such inconsistency could arise due to several economic fundamentals. We identify one such fundamental in income based dynamic panel models which is the long-run capital adjustment cost. Using a parametric form for such an adjustment cost technology in a standard Ramsey growth model, we have demonstrated that the dynamic panel regression with fixed individual effects gives rise to an inconsistent estimator of convergence for an infinite time horizon akin to Nickell bias (1981). The inconsistency arises because of the presence of a negative moving average term in the error of the dynamic panel regression. Unlike Nickell (1981), this asymptotic bias does not disappear in

[17] Grossmann and Osikominu.(2019) advice a structural approach to obtain credible results of IV estimations and use of a rigorous theoretical foundation that helps to determine valid, excluded instruments.

[18] See Baum et al. (2007) for a good survey of diagnostic tests of weak instruments including rank of matrix test for identification.

an infinite horizon. The bias is larger in economies with a higher adjustment cost of capital which is verified in a Monte Carlo simulation. The implication of this inconsistency is that the FE estimator of the "conditional convergence" of countries or regions could be seriously overestimated. We have shown analytically that the use of distant lags of the dependent variable as instruments could remove this bias. We have suggested that in remedying this bias standard procedures should also take into account the presence of a negative moving average term in the error. Our analysis is illustrative as it is based on a specific functional form for the capital adjustment cost technology that admits an analytical expression of the asymptotic bias. Future extension of our work would be to take a more general adjustment cost specification which includes short run investment adjustment cost such as Christiano et al. (2005) and explore the implication for inconsistency in dynamic panel with fixed effects.

Appendix

A. Optimal capital accumulation

The proof mimics Basu (1987). Write the value function for this problem as:

$$v(k_{it}, \epsilon_{it}, g_{i1t}, ..., g_{iJt}) = \max_{k_{it+1}} \left[\begin{array}{c} \ln\left\{\epsilon_{it}\left(\prod_{j=1}^{J}(g_{ijt})^{\lambda_j}\right)(k_{it})^{\alpha} - (k_{it+1}/k_{it})^{1/\theta}\,k_{it}\right\} \\ +\rho_i\,\mathrm{E}_t\,v(k_{it+1}, \epsilon_{it+1}, g_{i1t+1}, ..., g_{iJt+1}) \end{array} \right]$$

where E_t is the conditional expectation operator.

Conjecture that the value function is loglinear in state variables as follows:

$$v(k_{it}, \epsilon_{it}) = \pi_0 + \pi_1 \ln k_{it} + \pi_2 \ln \epsilon_{it} + \pi_3 \sum_{j=1}^{J} \lambda_j \ln g_{ijt}$$

which after plugging into the value function

$$\begin{aligned} &\pi_0 + \pi_1 \ln k_{it} + \pi_2 \ln \epsilon_{it} + \pi_3 \sum_{j=1}^{J} \lambda_j \ln g_{ijt} \\ &= \max_{k_{it+1}} \left[\begin{array}{c} \ln\left\{\epsilon_{it}\left(\prod_{j=1}^{J}(g_{ijt})^{\lambda_j}\right)(k_{it})^{\alpha} - (k_{it+1}/k_{it})^{1/\theta}\,k_{it}\right\} \\ +\rho_i\,\mathrm{E}_t\left\{\pi_0 + \pi_1 \ln k_{it+1} + \pi_2 \ln \epsilon_{it+1} + \pi_3 \sum_{j=1}^{J} \lambda_j \ln g_{ijt+1}\right\} \end{array} \right] \end{aligned} \quad \text{(A.1)}$$

Differentiating with respect to k_{it+1} and rearranging terms one gets:

$$k_{it+1} = \left[(\pi_1\rho_i\theta/(1+\pi_1\rho_i\theta))\right]^{\theta}(\epsilon_{it})^{\theta}\left(\prod_{j=1}^{J}(g_{ijt})^{\lambda_j}\right)^{\theta}(k_{it})^{\alpha\theta+1-\theta} \quad \text{(A.2)}$$

which after plugging into (A.1) and comparing left hand and right side coefficients of $\ln k_{it}$ uniquely solves:

$$\pi_1 = \alpha / \left(1 - \rho_i(\alpha\theta + 1 - \theta)\right)$$

which after plugging into (A.2) we get:

$$k_{it+1} = \left\{\alpha\rho_i\theta / \left(1 - \rho_i(1-\theta)\right)\right\}^{\theta} \left(\epsilon_{it}\right)^{\theta} \left(\prod_{j=1}^{J} \left(g_{ijt}\right)^{\lambda_j}\right)^{\theta} \left(k_{it}\right)^{\alpha\theta+1-\theta} \tag{A.3}$$

Note that the decision rule for the capital stock depends only on π_1. The remaining coefficients, π_0, π_2 and π_3 can also be solved by using the same method of undetermined coefficients and one can check that they are also uniquely determined by π_1.

B. Proof of proposition 1

First rewrite (17), when $\forall j\ \lambda_j = 0$, as:

$$\ln y_{it} = \gamma \ln y_{it-1} + \eta_i + u_{it} \tag{B.4}$$

where u_{it} is given by (16b). Then, rewrite (B.4) in a deviation (from individual steady-state mean) form as follows to eliminate the unobserved individual heterogeneity (η_i):

$$a_{it} = \gamma a_{it-1} + v_{it} \tag{B.5}$$

where

$$a_{it} \equiv \ln y_{it} - \ln y_i \text{ and } a_{it-1} \equiv \ln y_{it-1} - \ln y_{i-1} \tag{B.6}$$

$$v_{it} \equiv (\theta - 1)(\ln \epsilon_{it-1} - \ln \epsilon_{i-1}) + (\ln \epsilon_{it} - \ln \epsilon_i) \tag{B.7}$$

For any z, $\ln z_i \equiv T^{-1} \sum_{t=1}^{T} \ln z_{it}$ and $\ln z_{i-1} \equiv (T-1)^{-1} \sum_{t=2}^{T} \ln z_{it-1}$.

The FE estimator of γ is the pooled OLS estimator of the model (B.5),

$$p \lim (\widehat{\gamma}) = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (a_{it} a_{it-1})}{\sum_{i=1}^{N} \sum_{t=1}^{T} a_{it-1}^2} = \gamma + \frac{\sum_{i=1}^{N} T^{-1} \sum_{t=1}^{T} (v_{it} a_{it-1})}{\sum_{i=1}^{N} T^{-1} \sum_{t=1}^{T} a_{it-1}^2} \tag{B.8}$$

When $T \to \infty$, the terms in the right side of (B.8) can be rewritten as,

$$p \lim (\widehat{\gamma}) = \gamma + \sum_{i=1}^{N} \mathrm{E}\left[v_{it} a_{it-1}\right] / \sum_{i=1}^{N} \mathrm{E}\left[a_{it-1}^2\right] \tag{B.9}$$

where $\mathrm{E}(.)$ stands for the time expectation operator.

Substituting back (B.6) into (B.9), we obtain,

$$\begin{aligned} p \lim (\widehat{\gamma}) &= \gamma + \frac{\sum_{i=1}^{N} \mathrm{E}\left[(\ln y_{it-1} - \mathrm{E}\left[\ln y_{it-1}\right])\left((\theta - 1) \ln \epsilon_{it-1} + \ln \epsilon_{it}\right)\right]}{\sum_{i=1}^{N} \mathrm{E}\left[(\ln y_{it-1} - \mathrm{E}\left[\ln y_{it-1}\right])(\ln y_{it-1} - \mathrm{E}\left[\ln y_{it-1}\right])\right]} \\ &= \gamma + \sum_{i=1}^{N} \mathrm{cov}\left((\ln y_{it-1}, (\theta - 1) \ln \epsilon_{it-1} + \ln \epsilon_{it}\right) / \sum_{i=1}^{N} \mathrm{var}\left[\ln y_{it-1}\right] \end{aligned} \tag{B.10}$$

Note that from (6), $\mathrm{cov}(\ln y_{it-1}, \ln \epsilon_{it}) = 0$.[19] Thus, (B.10) becomes

[19] This is also refereed as sequential exogeneity (see Wooldridge, 2010, Ch. 10 & 11).

$$p\lim(\widehat{\gamma}) = \gamma + (\theta - 1)\sum_{i=1}^{N}\mathrm{cov}(\ln y_{it-1}, \ln \epsilon_{it-1}) / \sum_{i=1}^{N}\mathrm{var}(\ln y_{it-1}) \tag{B.11}$$

Then, substitute (6) into (B.11) to obtain,

$$\begin{aligned} p\lim(\widehat{\gamma}) &= \gamma + (\theta - 1)\,\mathrm{cov}(\alpha \ln k_{it-1} + \ln \epsilon_{it-1}, \ln \epsilon_{it-1}) / \mathrm{var}\,(\ln y_{it-1}) \\ &= \gamma + (\theta - 1)\, N^{-1}\sum_{i=1}^{N}\mathrm{var}(\ln \epsilon_{it-1}) / N^{-1}\sum_{i=1}^{N}\mathrm{var}\,(\ln y_{it-1}) \end{aligned} \tag{B.12}$$

since k_{it-1} is predetermined and, hence, uncorrelated with ϵ_{it-1} (see (6)).

Taking $N \to \infty$, (B.12) can be rewritten as:

$$p\lim(\widehat{\gamma}) = \gamma + (\theta - 1)\,\mathrm{E}_i\,\mathrm{var}(\ln \epsilon_{it-1}) / \mathrm{E}_i\,\mathrm{var}\,(\ln y_{it-1}) \tag{B.13}$$

where $\mathrm{E}_i(.)$ represents the cross sectional expectation. Since $\mathrm{var}(\ln \epsilon_{it-1}) = \upsilon^2$ is the same for all i, $\mathrm{E}_i\,\mathrm{var}(\ln \epsilon_{it-1}) = \upsilon^2$.

C. The multivariate case

For the case $\exists j \;\; \lambda_j \neq 0$, first rewrite (17) as:

$$\ln y_{it} = \widetilde{\mathbf{x}}_{it}\widetilde{\boldsymbol{\lambda}} + \eta_i + u_{it} \tag{C.14}$$

where $\widetilde{\mathbf{x}}_{it} \equiv (\ln y_{it-1}, \mathbf{x}_{it})$ is a $1 \times (J+1)$ and $\widetilde{\boldsymbol{\lambda}} \equiv (\gamma, \boldsymbol{\lambda}')'$ is a $(J+1) \times 1$ vector of parameters.

Then, transform the equation in (C.14) to eliminate the fixed effects (η_i):

$$a_{it} = \mathbf{b}_{it}\widetilde{\boldsymbol{\lambda}} + v_{it} \tag{C.15}$$

where $\mathbf{b}_{it} \equiv \widetilde{\mathbf{x}}_{it} - \widetilde{\mathbf{x}}_i$.

Recall that:

$$a_{it} \equiv \ln y_{it} - \ln y_i \tag{C.16a}$$

$$v_{it} \equiv (\theta - 1)\left(\ln \epsilon_{it-1} - \ln \epsilon_{i-1}\right) + \left(\ln \epsilon_{it} - \ln \epsilon_i\right) \tag{C.16b}$$

$$\widetilde{\mathbf{x}}_{it} \equiv \left(\ln y_{it-1}, \mathbf{x}_{it}\right) \tag{C.16c}$$

$$\begin{aligned}\mathbf{x}_{it} &\equiv \left(x_{i1t}, x_{i2t}, ..., x_{iJt}\right) \\ &= (\theta - 1)\, \mathbf{g}_{it-1} + \mathbf{g}_{it}\end{aligned} \tag{C.16d}$$

$$x_{ijt} \equiv (\theta - 1) \ln g_{ijt-1} + \ln g_{ijt} \tag{C.16e}$$

$$\mathbf{g}_{it} \equiv \left(\ln g_{i1t}, \ln g_{i2t}, ..., \ln g_{iJt}\right) \tag{C.16f}$$

$$\widetilde{\boldsymbol{\lambda}} \equiv \left(\gamma, \lambda_1, \lambda_2, ..., \lambda_J\right)' \tag{C.16g}$$

$$\begin{aligned}\mathbf{b}_{it} &= \left(\ln y_{it-1} - \ln y_{i-1}, x_{i1t} - x_{i1}, ..., x_{iJt} - x_{iJ}\right) \\ &\equiv \left(b_{i0t}, b_{i1t}, ..., b_{iJt}\right)\end{aligned} \tag{C.16h}$$

The FE estimator of $\widetilde{\boldsymbol{\lambda}}$ is the pooled OLS estimator of the model (C.15)**:**

$$p\lim\left(\widehat{\widetilde{\boldsymbol{\lambda}}}\right) = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{b}_{it}'\mathbf{b}_{it}\right)\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{b}_{it}'a_{it}\right)$$
$$= \widetilde{\boldsymbol{\lambda}} + \left(\sum_{i=1}^{N}T^{-1}\sum_{t=1}^{T}\left(\mathbf{b}_{it}'\mathbf{b}_{it}\right)\right)^{-1}\sum_{i=1}^{N}T^{-1}\sum_{t=1}^{T}\left(\mathbf{b}_{it}'\upsilon_{it}\right) \qquad \text{(C.17)}$$

When $T \to \infty$, the terms in the right hand side of (C.17) can be rewritten as,

$$p\lim\left(\widehat{\widetilde{\boldsymbol{\lambda}}}\right) = \widetilde{\boldsymbol{\lambda}} + \left(\sum_{i=1}^{N}\mathrm{E}\left[\mathbf{b}_{it}'\mathbf{b}_{it}\right]\right)^{-1}\sum_{i=1}^{N}\mathrm{E}\left[\mathbf{b}_{it}'\upsilon_{it}\right] \qquad \text{(C.18)}$$

Note that, the variance and covariance matrix is given by,

$$\mathbf{b}_{it}'\mathbf{b}_{it} = \begin{bmatrix} b_{i0t}^2 & b_{i0t}b_{i1t} & \dots & b_{i0t}b_{iJt} \\ b_{i1t}b_{i0t} & b_{i1t}^2 & \dots & b_{i1t}b_{iJt} \\ . & . & \dots & . \\ b_{iJt}b_{i0t} & b_{iJt}b_{i1t} & \dots & b_{iJt}^2 \end{bmatrix}$$

and, considering that x_{ijt} are exogenous and thus $\mathrm{E}\left[x_{ijt}\upsilon_{it}\right] = 0$, we can simplify the last term in (C.18) as,

$$\mathrm{E}\left[\mathbf{b}_{it}'\upsilon_{it}\right] = (\theta - 1)\,\mathrm{cov}\left(\ln y_{it-1},\ \ln \epsilon_{it-1}\right)\mathbf{p} \qquad \text{(C.19)}$$

where $\mathbf{p} \equiv (1, \mathbf{0})'$ and $\mathbf{0}$ is $1 \times J$ zero vector.

Substituting (C.19) into (C.18), we obtain:

$$p\lim\left(\widehat{\widetilde{\boldsymbol{\lambda}}}\right) = \widetilde{\boldsymbol{\lambda}} + \left(\mathrm{E}_i\,\mathrm{E}\left[\mathbf{b}_{it}'\mathbf{b}_{it}\right]\right)^{-1}\mathbf{p}(\theta - 1)\upsilon^2 \qquad \text{(C.20)}$$

since, from Appendix B, $\mathrm{cov}\left(\ln y_{it-1},\ \ln \epsilon_{it-1}\right) = \upsilon^2$.

D. Using lagged internal instruments

Considering the univariate case, first, instrumenting (B.5) with a_{it-2} leads to the the IV estimator:

$$\widehat{\gamma}_{IV} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(a_{it}a_{it-2}\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(a_{it-2}a_{it-1}\right)} \tag{D.21}$$

Plugging (B.5) into (D.21), we can write:

$$\widehat{\gamma}_{IV} = \gamma + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(v_{it}a_{it-2}\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(a_{it-2}a_{it-1}\right)} \tag{D.22}$$

When $T \to \infty$, the terms in the right side of (D.22) can be rewritten as,

$$p\lim\left(\widehat{\gamma}\right) = \gamma + \sum_{i=1}^{N} \mathrm{E}\left[v_{it}a_{it-2}\right] / \sum_{i=1}^{N} \mathrm{E}\left[\left(a_{it-2}a_{it-1}\right)\right] \tag{D.23}$$

where $\mathrm{E}(.)$ stands for the time expectation operator.

Substituting back (B.6) into (D.22), we obtain

$$p\lim\left(\widehat{\gamma}\right) = \gamma + \sum_{i=1}^{N} \operatorname{cov}\left(\left(\ln y_{it-2}, (\theta-1)\ln\epsilon_{it-1} + \ln\epsilon_{it}\right) / \sum_{i=1}^{N} cov(\ln y_{it-1}, \ln y_{it-2}\right) \tag{D.24}$$

But $\operatorname{cov}(\ln y_{it-2}, \ln\epsilon_{it}) = \operatorname{cov}(\ln y_{it-2}, \ln\epsilon_{it-1}) = 0$. Thus, (D.24) becomes

$$p\lim\left(\widehat{\gamma}\right) = \gamma \tag{D.25}$$

given $cov(\ln y_{it-1}, \ln y_{it-2}) \neq 0$, or the instrument variable is sufficiently correlated with the endogenous regressor. Thus the use of lagged internal instrument eliminates the Nickell bias in our model.

In the multivariate case, equation (C.19) turns to

$$\mathrm{E}\left[\mathbf{b}_{it}' v_{it}\right] = (\theta - 1)\,\mathrm{cov}\left(\ln y_{it-2},\ \ \ln \epsilon_{it-1}\right)\mathbf{p} = \mathbf{0} \tag{D.26}$$

which implies that (C.20) becomes now

$$p\lim\left(\widehat{\widetilde{\boldsymbol{\lambda}}}\right) = \widetilde{\boldsymbol{\lambda}} \tag{D.27}$$

Acknowledgement

Authors wish to acknowledge a referee for useful comments. Wiji Arulampalam, Anurag Banerjee, and Thomas Ziesemer are gratefully acknowledged for comments. The usual disclaimer applies.

References

[1] Acemoglu, D., and J. Ventura (2002): "The World Income Distribution," *The Quarterly Journal of Economics*, 117(2), 659–694.

[2] Anderson, T. W., and C. Hsiao (1981): "Estimation of dynamic models with error components," *Journal of the American statistical Association*, 76(375), 598–606.

[3] Angeletos, G.-M., and L.-E. Calvet (2006): "Idiosyncratic production risk, growth and the business cycle," *Journal of Monetary Economics*, 53(6), 1095–1115.

[4] 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(2), 277–297.

[5] Arellano, M., and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68(1), 29–51.

[6] Arrow, K. J. (1962): "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29(June), 155–173.

[7] Badunenko, O., and S. C. Kumbhakar (2016): "When, where and how to estimate persistent and transient efficiency in stochastic frontier panel data models," *European Journal of Operational Research*, 255(1), 272–287.

[8] Barro, R. J. (1990): "Government Spending in a Simple Model of Endogeneous Growth," *The Journal of Political Economy*, 98(5), S103–S125.

[9] Barro, R. J., and X. Sala-i Martin (2004): *Economic Growth*. The MIT Press, Cambridge, Massachusetts London, England, second edn.

[10] Basu, P. (1987): "An Adjustment Cost Model of Asset Pricing," *International Economic Review*, 28(3), 609–621.

[11] Basu, P., and Y. Getachew (2015): "An adjustment cost model of social mobility," *Journal of Macroeconomics*, 44, 177 – 190.

[12] Basu, P., M. Gillman, and J. Pearlman (2012): "Inflation, human capital and Tobin's q," *Journal of Economic Dynamics and Control*, 36(7), 1057–1074.

[13] Baum, C. F., M. E. Schaffer, and S. Stillman (2007): "Enhanced routines for instrumental variables/generalized method of moments estimation and testing," *The Stata Journal*, 7(4), 465–506.

[14] Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1), 115–143.

[15] Bond, S. (2002): "Dynamic panel data models: a guide to microdata methods and practice," CeMMAP working papers CWP09/02, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

[16] Caballero, R. J. (1994): "Small Sample Bias and Adjustment Costs," *The Review of Economics and Statistics*, 76(1), 52–58.

[17] 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(3), 363–389.

[18] Chirinko, R. S. (1993): "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications," *Journal of Economic Literature*, 31(4), 1875–1911.

[19] Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

[20] Grossmann, V., and A. Osikominu (2019): "Let the Data Speak? On the Importance of Theory-Based Instrumental Variable Estimations," *German Economic Review*.

[21] Groth, C., and H. Khan (2010): "Investment Adjustment Costs: An Empirical Assessment," *Journal of Money, Credit and Banking*, 42(8), 1469–1494.

[22] Hahn, J., and G. Kuersteiner (2002): "Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T Are Large," *Econometrica*, 70(4), 1639–1657.

[23] Hamermesh, D. S., and G. A. Pfann (1996): "Adjustment costs in factor demand," *Journal of Economic Literature*, 34(3), 1264.

[24] Hercowitz, Z., and M. Sampson (1991): "Output Growth, the Real Wage, and Employment Fluctuations," *American Economic Review*, 81(5), 1215–1237.

[25] Ho, T.-W. (2006): "INCOME THRESHOLDS AND GROWTH CONVERGENCE: A PANEL DATA APPROACH," *Manchester School (14636786)*, 74(2), 170–189.

[26] Imbens, G. W., and T. Lancaster (1994): "Combining micro and macro data in microeconometric models," *The Review of Economic Studies*, 61(4), 655–680.

[27] Islam, N. (1995): "Growth empirics: A panel data approach," *Quarterly Journal of Economics*, 110(4), 1127–1170.

[28] Judson, R. A., and A. L. Owen (1999): "Estimating dynamic panel data models: a guide for macroeconomists," *Economics Letters*, 65(1), 9–15.

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

[30] Lee, Y.-J. (2014): "Testing a linear dynamic panel data model against nonlinear alternatives," *Journal of Econometrics*, 178, 146–166.

[31] Lucas, Robert E., J., and E. C. Prescott (1971): "Investment Under Uncertainty," *Econometrica*, 39(5), 659–681.

[32] Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417–1426.

[33] Pesaran, M. H., and R. Smith (1995): "Estimating long-run relationships from dynamic heterogeneous panels," *Journal of econometrics*, 68(1), 79–113.

[34] Romer, P. M. (1986): "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94(5), 1002–1037.

[35] Roodman, D. (2009a): "How to do xtabond2: An introduction to difference and system GMM in Stata," *Stata Journal*, 9(1), 86–136.

[36] Roodman, D. (2009b): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.

[37] Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts, second edition edn.