ARTICLE IN PRESS

Journal of Econometrics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

An integrated panel data approach to modelling economic growth

Guohua Feng [a,*], Jiti Gao [b], Bin Peng [b]

[a] Department of Economics, University of North Texas, Denton, TX 76201, USA
[b] Department of Econometrics and Business Statistics, Monash University, Caulfield East, VIC 3145, Australia

ARTICLE INFO

Article history:
Received 13 March 2019
Received in revised form 18 April 2020
Accepted 25 September 2020
Available online xxxx

JEL classification:
C23
O47

Keywords:
Cross-sectional dependence
Growth regressions
Parameter heterogeneity
Variable selection

ABSTRACT

Empirical growth analysis is plagued with three problems – variable selection, parameter heterogeneity and cross-sectional dependence – which are addressed independently from each other in most studies. This study is to propose an integrated framework that allows for parameter heterogeneity and cross-sectional error dependence, while simultaneously performing variable selection. We derive the asymptotic properties of the estimator, and apply the framework to a dataset of 89 countries over the period from 1960 to 2014. Our results support the "optimistic" conclusion of Sala-I-Martin (1997), and also reveal some cross-country patterns not found previously.

1. Introduction

Following the seminal works of Kormendi and Meguire (1985) and Barro (1991), a vast amount of studies in the empirical growth literature have attempted to identify salient determinants of economic growth. A main tool used by these studies is "cross-country growth regressions" — that is, to regress observed GDP growth on a plethora of possible explanatory variables that could possibly affect growth across countries. Excellent surveys of these studies and their role in the broader context of economic growth theory are provided in Durlauf and Quah (1999), Temple (1999) and Durlauf et al. (2005).

Despite the vast amount of research, the literature has identified a number of problems with conventional growth regressions, among which three deserve particular attention. The first problem is determining what variables to be included in growth regressions. This problem arises because of the nature of growth theories: although a plethora of growth theories have been proposed to identify factors that affect growth, these theories are open-ended in the sense that the validity of one causal theory of growth does not imply the falsity of another (Brock and Durlauf, 2001). In words of Durlauf et al. (2008), "a given body of candidate growth theories defines a space of possible models rather than a single specification". From an empirical perspective, this problem stems from the fact that the number of potential explanatory variables is large (over 140 identified in Durlauf et al., 2005) relative to the number of countries with enough data availability, rendering the all-inclusive regression computationally infeasible (Sala-I-Martin et al., 2004; Durlauf et al., 2005). In dealing with this problem some studies have resorted to simply "trying" combinations of variables which could

* Corresponding author.
E-mail addresses: Guohua.Feng@unt.edu (G. Feng), Jiti.Gao@monash.edu (J. Gao), Bin.Peng@monash.edu (B. Peng).

https://doi.org/10.1016/j.jeconom.2020.09.009
0304-4076/

be potentially important determinants of growth and reported the results of their preferred specification. However, as noted by Leamer (1983) and Sala-I-Martin et al. (2004) such "data-mining" could lead to spurious inference.

The second problem with conventional growth analysis is that most empirical growth studies assume that the parameters of growth regressions are identical across countries. This assumption complies with the classical Solow model (Mankiw et al., 1992), which assumes that all countries share an identical aggregate Cobb–Douglas production function. However, an increasing number of studies (e.g., Durlauf and Johnson, 1995; Durlauf et al., 2001; Salimans, 2012) have suggested that the parameters are heterogeneous across countries. These studies, though using different econometric methods, all suggest that the assumption of a single linear growth model that applies to all countries is inappropriate. For example, Durlauf and Johnson (1995) employ a regression tree analysis to show that a cross-sectional regression using the (Summers and Heston, 1991) data appears to provide support for several distinct regimes in which aggregate production functions vary among countries according to their level of development, while (Durlauf et al., 2001), employing a varying coefficient growth model, also find strong evidence of parameter heterogeneity across countries.

The third problem is that few studies in the empirical growth literature allow for cross-sectional dependence of individual countries. Panel data econometrics has recently seen an increasing interest in models with unobserved time-varying heterogeneity caused by latent common shocks influencing all units, possibly to a different degree. This type of heterogeneity introduces cross-sectional dependence to individual countries, which, when neglected, can lead to biased estimates and spurious inference (Pesaran, 2006; Bai, 2009). In the context of cross-country growth analysis, the problem of cross-sectional dependence seems particularly salient due to the omnipresence of common global shocks (such as global financial crises and world oil price shocks) that affect all countries through trade and financial linkages (Chudik et al., 2017). Durlauf and Quah (1999) discuss the possibility of cross-sectional dependence in a (Lucas, 1993) growth model with human capital spillovers. They find that these spillovers markedly change the dynamics of convergence and the authors call for the modelling of cross-country interactions in empirical convergence analysis.

The three aforementioned problems have received more or less individual attention in the growth literature. For example, Durlauf et al. (2001) address the problem of parameter heterogeneity using a varying coefficient growth model, but do not deal with the problems of variable selection and cross-sectional dependence; both (Sala-I-Martin et al., 2004) and Moral-Benito (2012) select growth determinants using Bayesian averaging, but do not account for parameter heterogeneity and cross-sectional dependence.

The main goal of this study is to propose an integrated framework that is capable of dealing with parameter heterogeneity and cross-sectional dependence, while simultaneously performing variable selection. Specifically, parameter heterogeneity is allowed for by permitting the coefficients to vary across countries according to a country's initial conditions, while cross-sectional dependence is accounted for via a factor structure. We then propose a least absolute shrinkage and selection operator (LASSO) estimator to select growth determinants, establish the associated asymptotic results, and further verify our asymptotic results through extensive simulations, which constitutes another contribution of this paper. We apply this framework to a new dataset of 89 countries over the period 1960–2014. Our findings broadly support the more "optimistic" conclusion of Sala-I-Martin (1997), that is, some variables are important regressors for explaining cross-country growth patterns. Moreover, our empirical results also provide support to some important hypotheses in the growth literature, e.g., "middle income trap hypothesis", "natural resources curse hypothesis", "religion works via belief, not practice", etc.

The rest of the paper is organized as follows. Section 2 explains how to extend the canonical cross-country growth regression to account for the aforementioned issues. Section 3 describes a procedure for estimating the extended growth regression model, and presents the associated asymptotic properties. Section 4 describes the data. The empirical results are presented in Section 5. Section 6 presents several robustness checks. Section 7 concludes. Preliminary lemmas, proofs of the main theorems and Monte Carlo simulations, together with auxiliary tables and figures, are presented in the supplementary Appendix A. The proofs of the preliminary lemmas are presented in the supplementary Appendix B, which can be found at https://ssrn.com/abstract=3348229.

2. A varying coefficient growth regression model with factor structure and sparsity

A generic representation of the canonical cross-country growth regression is

$$y_{it} = x_{it}'\beta_0 + e_{it}, \tag{2.1}$$

where $i = 1, 2, \ldots, N$ index countries; $t = 1, 2, \ldots, T$ index time; y_{it} is the rate of economic growth; x_{it} represents a set of observable explanatory variables, including those originally suggested by Solow as well as other growth theories, and e_{it} is an error term. Eq. (2.1) represents the baseline for much of growth econometrics.

However, as discussed in the introduction, (2.1) is based on two problematic assumptions. First, it assumes that the parameters (i.e., β_0) are homogeneous across all countries. Second, it assumes that there is no cross-sectional dependence across countries. To relax the two assumptions, in what follows we extend the conventional cross-country growth regression of (2.1) in two ways. In Section 2.1 we allow for parameter heterogeneity by allowing β_0 to vary across countries according to a country's initial conditions. In Section 2.2 we introduce cross-sectional dependence into the model by means of a factor structure.

2.1. Parameter heterogeneity

Following (Durlauf et al., 2001), we allow for parameter heterogeneity by generalizing (2.1) into a varying coefficient model:

$$y_{it} = x_{it}'\beta_0(z_{it}) + e_{it}, \tag{2.2}$$

where z_{it} can be interpreted as some measure of "development" (or "initial condition") of a country, and $\beta_0(z) = (\beta_{01}(z), \ldots, \beta_{0p}(z))'$ is a vector of smooth functions that map the scalar index variable z_{it} into a set of country-specific parameters.

This generalization in (2.2) provides a framework within which one can bridge the gap between cross-country regression models and new growth theories. For instance, if one believes that initial GDP per capita causally affects a country's production technology and growth as in Durlauf et al. (2001), then initial GDP per capita can be introduced as a "development" index. As pointed out by Durlauf and Johnson (1995), (2.2) is compatible both with a model in which economies pass through distinct phases of development towards a unique steady state as well as with one in which multiple steady states exist.

2.2. Cross-sectional error dependence

Having accounted for parameter heterogeneity, we next introduce cross-sectional dependence of error terms into (2.2) using a factor structure:

$$e_{it} = \gamma_{0i}'f_{0t} + \varepsilon_{it}, \tag{2.3}$$

where f_{0t} is an $r \times 1$ vector of unobservable common factors, γ_{0i} is an $r \times 1$ vector of factor loadings that capture country-specific responses to the common shocks, and ε_{it} is the idiosyncratic error term. These factors can capture unobservable impacts like world oil price shocks, global financial crises, recessions in major advanced economies, local spillover effects along with channels determined by shared culture heritage, geographic proximity, economic/social interaction, and so forth (Chudik et al., 2011). Moreover, the components of the factor structure are allowed to drive both economic growth and explanatory variables, thus partially accounting for potential endogeneity of explanatory variables, which is neglected by the traditional approaches to causal interpretation of cross-country empirical analysis.

2.3. The varying coefficient growth regression model with factor structure and sparsity

Substituting (2.3) into (2.2) yields the following growth regression model that allows for parameter heterogeneity and cross-sectional dependence:

$$y_{it} = x_{it}'\beta_0(z_{it}) + \gamma_{0i}'f_{0t} + \varepsilon_{it}, \tag{2.4}$$

which extends the local Solow growth model investigated in Durlauf et al. (2001) into a panel data context with interactive fixed effects (or factor structure). From an econometric perspective, (2.4) extends the panel data model with interactive fixed effects in Pesaran (2006) and Bai (2009) into a varying coefficient context, which is naturally motivated by the relevant empirical growth literature.

In addition to parameter heterogeneity and cross-sectional dependence, we are also interested in another issue that is prominent in the empirical growth literature — variable selection. This issue is important because (1) the dimension of x_{it} can be very large; and (2) not all elements of x_{it} drive economic growth. In other words, for those regressors not driving economic growth, it is reasonable to assume that their associated coefficients are zero, which is called "sparsity" in the literature of high dimensional econometrics.

In order to formally introduce the sparsity to the model (2.4), we assume that there exists an unknown set $\mathcal{A}^\dagger \subseteq \{1, \ldots, p\}$ satisfying that $E|\beta_{0j}(z_{it})|^2 = 0$ if and only if $j \in \mathcal{A}^\dagger$. For notational simplicity, we assume $\mathcal{A}^\dagger = \{p^* + 1, \ldots, p\}$ for an unknown integer p^* satisfying $1 \le p^* < p$. Further, let $\mathcal{A}^* = \{1, \ldots, p^*\}$, $x_{it}^* = (x_{it,1}, \ldots, x_{it,p^*})'$, and $\beta_0^*(z) = (\beta_{01}(z), \ldots, \beta_{0p^*}(z))'$. Throughout this study, we always define the variables or functions corresponding to the sets $\mathcal{A}^*$ and $\mathcal{A}^\dagger$ with super-indices * and † respectively. Thus, identifying growth determinants is equivalent to distinguishing $\mathcal{A}^*$ and $\mathcal{A}^\dagger$, which will be achieved by a LASSO estimator presented in the following section. Finally, regarding the dimension of regressors, we consider two cases where (1) p is fixed, and (2) p diverges as the sample size increases. We refer to them as the low dimensional (LD) case and the high dimensional (HD) case, respectively. In terms of econometric methodology, both cases with the sparsity setting have not been studied in the literature to the best of our knowledge.

We emphasize that as discussed in the introduction, failure to perform variable selection may result in spurious inference, failure to allow parameters to differ across countries is inconsistent with the increasing body of research that find cross-country parameters heterogeneity, and failure to account for cross-sectional dependence can lead to biased estimates and spurious inference. These possible consequences thus necessitate an integrated approach to simultaneously addressing the three issues. In the following section, we introduce a LASSO estimator that is designed specifically for performing variable selection on the extended growth regression model in (2.4).

3. Estimation

In this section, we propose a procedure to estimate model (2.4) and derive the associated asymptotic properties. Specifically, we first use a sieve method to approximate the coefficient functions of the growth regression model in (2.4), and then propose a LASSO estimator to select the significant variables. When estimating the growth regression model, we employ the principle component analysis (PCA) technique to estimate the unobservable factor structure.

Before proceeding further, we introduce some notations that will be used throughout this section. Let $Y_i = (y_{i1}, \dots, y_{iT})'$, $X_i = (x_{i1}, \dots, x_{iT})'$, $Z_i = (z_{i1}, \dots, z_{iT})'$, $\mathcal{E}_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$, $F_0 = (f_{01}, \dots, f_{0T})'$, and $\Gamma_0 = (\gamma_{01}, \dots, \gamma_{0N})'$. $\|\cdot\|$ denotes the Euclidean norm of a vector or the Frobenius norm of a matrix; for a square matrix W, let $\eta_{\min}(W)$ and $\eta_{\max}(W)$ stand for the minimum and maximum eigenvalues of W respectively; $M_W = I_T - P_W$ denotes the orthogonal projection matrix generated by matrix W, where $P_W = W(W'W)^{-1}W'$, and W is a matrix with full column rank.

As mentioned above, we first approximate the unknown coefficient functions of the growth regression model in (2.4), using the sieve method. Let all elements of $\beta_0(z)$ belong to a Hilbert space $L^2(\mathcal{R}, \pi(w)) = \{g \mid \int_{\mathcal{R}} g^2(w)\pi(w)dw < \infty\}$, where $\pi(\cdot)$ is a known weight function. The use of the density $\pi(w)$ makes the Hilbert space sufficiently large to include all kinds of functions that are of interest to economists and econometricians such as linear, power, and polynomial functions.

We first introduce some important properties of the Hilbert space $L^2(\mathcal{R}, \pi(w))$ and its associated orthonormal systems. Define an inner product $\langle f, g\rangle = \int_{\mathcal{R}} f(w)g(w)\pi(w)dw$ for $f, g \in L^2(\mathcal{R}, \pi(w))$ that induces norm $\|f\|_{L^2} = \sqrt{\langle f, f\rangle}$. Suppose that there exists a complete orthogonal function sequence $\{h_j(w) \mid j \geq 0\}$ in the space $L^2(\mathcal{R}, \pi(w))$ satisfying that $\langle h_i, h_j\rangle = \delta_{ij}$ (the Kronecker delta), and that $\{h_j(w) \mid j \geq 0\}$ is uniformly bounded in the sense that $\sup_{w\in\mathcal{R}} \sup_{j\geq 0} |h_j(w)\pi^{1/2}(w)| < \infty$. The existence of the complete orthogonal sequence is guaranteed by Theorem 5.4.7 in Dudley (2003, p. 169), while its uniform boundedness is fulfilled in two situations that are mostly used in the literature (Newey, 1997 and Dong and Linton, 2018). With regard to the support $\mathcal{R}$, it can be bounded or unbounded. When $\mathcal{R}$ is a compact interval on the real line, conventional orthonormal sequences such as Fourier series or polynomial sequence can be used. When $\mathcal{R}$ is unbounded such as $\mathcal{R} = (-\infty, +\infty)$ with $\pi(w) \equiv \exp(-w^2/2)$ (or $\pi(w) \equiv \exp(-w^2)$), the sequence of probabilists' (or physicists') Hermite polynomials is an orthogonal basis in $L^2(\mathcal{R}, \pi(w))$ that satisfies the uniform boundedness. As for the choice of $L^2(\mathcal{R}, \pi(w))$, one needs to take into account several factors, such as how large the function may be and what type of regression function one may employ in the space. The baseline is that $\mathcal{R}$ should cover the range of $[\min\{z_{it}\}, \max\{z_{it}\}]$ in practice. We refer interested readers to Chen (2007) and Dong and Linton (2018) for more relevant discussions.

For $\forall g \in L^2(\mathcal{R}, \pi(w))$, we can then expand $g(w)$ as follows: $g(w) = \sum_{j=0}^{\infty} c_j h_j(w)$, where $c_j = \langle g, h_j\rangle$. By the Parseval equality, $\|g\|_{L^2}^2 = \sum_{j=0}^{\infty} c_j^2$, implying the attenuation of the coefficients. Define the partial sum $g_m(w) := \sum_{j=0}^{m-1} c_j h_j(w)$ and the residue $\delta_g(w) := \sum_{j=m}^{\infty} c_j h_j(w)$ for $m \geq 1$. It is easy to see that the convergence of $g_m(w)$ to $g(w)$ as $m \to \infty$ holds in the norm sense and even in the point-wise sense when the function is smooth. Throughout this study, for any vector of functions $G(\cdot) = (g_1(\cdot), \dots, g_{d_G}(\cdot))'$, its norm is defined as $\|G\|_{L^2} = \{\sum_{\ell=1}^{d_G} \|g_\ell\|_{L^2}^2\}^{1/2}$.

Applying the above expansion to the coefficient functions of our growth regression model (i.e., (2.4)) yields

$$\beta_{0\ell}(z) = C'_{\beta_{0\ell}} H_{m_\ell}(z) + \delta_{\beta_{0\ell}}(z)$$

for any functional component $\beta_{0\ell}(z)$ with $\ell = 1, \dots, p$, where $C_{\beta_{0\ell}} = (c_{\ell,0}, \dots, c_{\ell,m_\ell-1})'$, $H_{m_\ell}(z) = (h_0(z), \dots, h_{m_\ell-1}(z))'$, $\delta_{\beta_{0\ell}}(z) = \sum_{j=m_\ell}^{\infty} c_{\ell,j} h_j(w)$, and $c_{\ell,j} = \langle \beta_{0\ell}, h_j\rangle$ for $j \geq 0$. Letting m_ℓ for $\ell = 1, \dots, p$ be the same value[1] m allows us to write

$$\beta_0(z) = C_{\beta_0} H_m(z) + \Delta_m(z) := \beta_{0,m}(z) + \Delta_m(z), \tag{3.1}$$

where

$$C_{\beta_0} = \begin{pmatrix} c_{1,0}, \dots, c_{1,m-1} \\ \vdots \\ c_{p^*,0}, \dots, c_{p^*,m-1} \\ 0_{(p-p^*)\times m} \end{pmatrix} := \begin{pmatrix} C^*_{\beta_0} \\ 0_{(p-p^*)\times m} \end{pmatrix},$$

$$\Delta_m(z) = \begin{pmatrix} \delta_{\beta_{01}}(z) \\ \vdots \\ \delta_{\beta_{0p^*}}(z) \\ 0_{(p-p^*)\times 1} \end{pmatrix} := \begin{pmatrix} \Delta^*_m(z) \\ 0_{(p-p^*)\times 1} \end{pmatrix}.$$

[1] The truncation parameter is set to be the same value for all components of $\beta_0(z)$ due to the following reasons. (1) The truncation parameters are chosen by the researcher. As long as they are within a reasonable range, the theoretical results presented in this section would carry through. (2) Using the same truncation parameter can substantially simplify the notations as shown in (3.1). (3) Using the same truncation parameter can speed up the computational process. See Section A.2 of the supplementary Appendix A for details. We will further explain how m is chosen in practice in Section 5.

Thus, the first p^* elements of $\beta_0(z)$ can be expressed as $\beta_0^*(z) = \beta_{0,m}^*(z) + \Delta_m^*(z)$, where $\beta_{0,m}^*(z) = C_{\beta_0}^* H_m(z)$.

Having approximated the coefficient functions, we then move on to estimate the growth regression model (2.4). Substituting (3.1) into (2.4) yields the following regression model in matrix form:

$$Y_i \approx \phi_i[\beta_{0,m}] + F_0\gamma_{0i} + \mathcal{E}_i = \mathcal{Z}_i \mathrm{vec}(C_{\beta_0}) + F_0\gamma_{0i} + \mathcal{E}_i$$

where $\phi_i[\beta] = (x_{i1}'\beta(z_{i1}), \ldots, x_{iT}'\beta(z_{iT}))'$ for any $p \times 1$ vector of functions $\beta(z)$, $\mathcal{Z}_i = (\mathcal{Z}_{i1}, \ldots, \mathcal{Z}_{iT})'$ and $\mathcal{Z}_{it} = H_m(z_{it}) \otimes x_{it}$. The approximation sign ($\approx$) in the above equation is due to the omission of the truncation residual.

If $F_0\gamma_{0i}$ were known, we could simply perform the group LASSO technique of Yuan and Lin (2006) on

$$Y_i - F_0\gamma_{0i} \approx \mathcal{Z}_i \mathrm{vec}(C_{\beta_0}) + \mathcal{E}_i$$

to estimate C_{β_0} and investigate the sparsity. Because both F_0 and γ_{0i} are unknown, we project out the factor structure to obtain

$$M_{F_0} Y_i \approx M_{F_0}\mathcal{Z}_i \mathrm{vec}(C_{\beta_0}) + M_{F_0}\mathcal{E}_i,$$

where M_{F_0} is a projection matrix defined above.

Our objective function can then be defined as

$$Q_\lambda(C_\beta, F) = \sum_{i=1}^N (Y_i - \phi_i[\beta_m])' M_F (Y_i - \phi_i[\beta_m]) + \sum_{j=1}^p \lambda_j \|C_{\beta,j}\|, \tag{3.2}$$

where $\beta_m(w) = C_\beta H_m(w)$, $C_{\beta,j}$ stands for the jth row of C_β, and $\lambda = (\lambda_1, \ldots, \lambda_p)'$ is the vector including the weight parameters of the coefficient functions, and is to be determined by data. The first term on the right hand side of (3.2) is commonly used when estimating a panel data model with a factor structure (e.g., $S_{NT}(\beta, F)$ of Bai, 2009). The second term is in the spirit of the group LASSO technique initially proposed by Yuan and Lin (2006) and subsequently extended by Wang and Xia (2009) to the nonparametric setting. Moreover, the penalty term in (3.2) is equivalent to performing the hard-thresholding rule of the lasso literature on the terms $\|C_{\beta,j}\|$ for $j = 1, \ldots, p$. Thus, the resulting estimator in (3.3) is a thresholding rule, which automatically sets small estimated coefficient functions to zero (i.e., penalizes small estimated coefficients) to reduce model complexity (see the *Sparsity* property mentioned in Fan and Li (2001) for details).

The estimators of C_{β_0} and F_0 that work for both the LD and HD cases can be readily obtained as

$$(\widehat{C}_\beta, \widehat{F}) = \arg\min_{C_\beta, F \in \mathrm{D}_F} Q_\lambda(C_\beta, F), \tag{3.3}$$

where $\mathrm{D}_F = \{F \mid \frac{F'F}{T} = I_r\}$. The numerical implementation of (3.3) is presented in Section A.2 of the supplementary Appendix A to conserve space of the main text. In what follows, we always partition $\widehat{C}_\beta$, according to $\mathcal{A}^*$ and $\mathcal{A}^\dagger$, as $\widehat{C}_\beta = (\widehat{C}_\beta^{*\prime}, \widehat{C}_\beta^{\dagger\prime})'$ wherever necessary.

At this point, it is convenient to state the necessary assumptions that are required for the establishment of the asymptotic results for both the LD and HD cases.

Assumption 1.

1. Let $\mathcal{F}_{-\infty}^0$ and $\mathcal{F}_\tau^\infty$ denote the σ-algebras generated by $\{(x_t, z_t, \varepsilon_t, f_{0t}) \mid t \le 0\}$ and $\{(x_t, z_t, \varepsilon_t, f_{0t}) \mid t \ge \tau\}$ respectively, where $x_t = (x_{1t}, \ldots, x_{Nt})'$, $z_t = (z_{1t}, \ldots, z_{Nt})'$, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$. Let $\alpha(\tau) = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_\tau^\infty} |\Pr(A)\Pr(B) - \Pr(AB)|$ be the mixing coefficient.
   (a) $\{X_i, Z_i, \mathcal{E}_i, \gamma_{0i}\}$ is identically distributed over i. $\{x_t, z_t, \varepsilon_t, f_{0t}\}$ is strictly stationary and α-mixing such that for some $\nu_1 > 0$, $E[\|\varepsilon_{11}\| + \|x_{11}\|]^{4+\nu_1} < \infty$, and the mixing coefficient satisfies $\sum_{t=1}^\infty [\alpha(t)]^{\nu_1/(2+\nu_1)} < \infty$.
   (b) $E[\varepsilon_{11}] = 0$, $E[\varepsilon_{11}^2] = \sigma_\varepsilon^2$, and $\{\varepsilon_{it}\}$ is independent of the other variables. Let $E[\varepsilon_{i1}\varepsilon_{j1}] = \sigma_{ij}$ for $i \ne j$, $\sum_{i \ne j} |\sigma_{ij}| = O(N)$, and $\sum_{i,j=1}^N \sum_{t,s=1}^T |E[\varepsilon_{it}\varepsilon_{js}]| = O(NT)$.
2. Let $\|\frac{1}{T}F_0'F_0 - \Sigma_f\| = O_P\left(\frac{1}{\sqrt{T}}\right)$ and $\|\frac{1}{N}\Gamma_0'\Gamma_0 - \Sigma_\gamma\| = O_P\left(\frac{1}{\sqrt{N}}\right)$, where Σ_f and Σ_γ are deterministic and positive definite. Moreover, $E\|f_{01}\|^4 < \infty$ and $E\|\gamma_{01}\|^4 < \infty$.

Assumption 2. Let $\Omega(F) = \frac{1}{NT}\left\{\Omega_1(F) - \Omega_2'(F)\Omega_3^{-1}\Omega_2(F)\right\}$, where $\Omega_1(F) = \sum_{i=1}^N \mathcal{Z}_i' M_F \mathcal{Z}_i$, $\Omega_2(F) = \sum_{i=1}^N \gamma_{0i} \otimes (M_F \mathcal{Z}_i)$, and $\Omega_3 = \Gamma_0'\Gamma_0 \otimes I_T$.

1. Assume all elements of $\beta_0(z)$ belong to $L^2(\mathcal{R}, \pi(w)) = \{g \mid \int_{\mathcal{R}} g^2(w)\pi(w)dw < \infty\}$, and there exists a constant μ such that $\|\Delta_m(z)\|_{L^2} = O(m^{-\mu/2})$, where $\pi(\cdot)$ is a known probability weight function. Suppose that $\sup_{w \in \mathcal{R}} f_z(w)/\pi(w) < \infty$, where $f_z(w)$ is the density function of z_{it}. Let $\eta_{\max}(\frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i'\mathcal{Z}_i') < \infty$ with a probability approaching one.
2. Suppose $\inf_{F \in \mathrm{D}_F} \eta_{\min}(\Omega(F)) > 0$.

Assumption 1 is standard in the literature. The mixing conditions are similar to Assumption C of Bai (2009) and Assumption 3.4 of Fan et al. (2016).

The order of $\|\Delta_m(z)\|_{L^2} = O(m^{-\mu/2})$ of Assumption 2.1 requires smoothness of the coefficient functions, and is the same as Assumption 3 of Newey (1997), wherein this requirement is discussed in details. The condition on $\eta_{\max}(\frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i'\mathcal{Z}_i') < \infty$ is similar to Assumption 3.1 of Fan et al. (2016). To give an example, consider a special case where $f_z(w) = \pi(w)$, $\{z_{it}\}$ and $\{x_{it}\}$ are mutually independent, $E[x_{it}] = 0$ and $E[x_{it}x_{it}'] = I_p$. In this case, it is easy to see that $\eta_{\max}(\frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i'\mathcal{Z}_i') < \infty$ holds true for both of the LD and HD cases by some standard analysis. The restriction on the density of z_{it} of Assumption 2.1 gives

$$E[h_j^2(z_{it})] = \int h_j^2(w)\pi(w)\frac{f_z(w)}{\pi(w)}dw \le O(1)\int h_j^2(w)\pi(w)dw < \infty,$$

which holds true uniformly in j and is equivalent to Assumption 3.3.ii of Fan et al. (2016).

Assumption 2.2 ensures that the estimators of (3.3) are well defined, and is equivalent to Assumption A of Bai (2009). To justify this assumption, we follow (Lam and Yao, 2012) and assume that $\frac{\Gamma_0'\Gamma_0}{N} = I_r$. With this latter assumption, $\Omega(F)$ reduces to

$$\Omega(F) = \frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i' M_F \mathcal{Z}_i - \frac{1}{N^2T}\sum_{i,j=1}^N \gamma_{0i}'\gamma_{0j}\mathcal{Z}_i' M_F \mathcal{Z}_j.$$

If we further assume $\{\gamma_{0i}\}$ is independent of $\{\mathcal{Z}_i\}$ for simplicity, then the second term in the above equation becomes negligible due to $\frac{\Gamma_0'\Gamma_0}{N} = I_r$. Thus, $\Omega(F)$ can be simplified as $\frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i' M_F \mathcal{Z}_i$, which is achievable provided that there are enough variations in $\mathcal{Z}_i$ (i.e., variations in z_{it} and x_{it}) as explained in Bai (2009, p. 1241). Moreover, if we further let $E[\mathcal{Z}_i] = 0$ (e.g., $E[x_{it}] = 0$, and assume that $\{x_{it}\}$ and $\{z_{it}\}$ are independent of each other), $\Omega(F)$ further reduces to $\frac{1}{NT}\sum_{i=1}^N \mathcal{Z}_i'\mathcal{Z}_i$ by a proof similar to (2) of Lemma A.2. In this case, it becomes even more straightforward to justify Assumption 2.2.

3.1. Low dimensional case

Given the above setting, we first investigate the asymptotic properties of the estimator in (3.3) for the LD case (i.e., p is fixed). In order to identify $\mathcal{A}^*$ and $\mathcal{A}^\dagger$ and then to establish the asymptotic properties, we impose the following assumptions.

Assumption 3.

1. $\frac{m^2}{T} \to 0$ and $\frac{\lambda_{\max}^*}{N^{\frac{3}{4}}T} \to 0$, where $\lambda_{\max}^* = \max\{\lambda_1, \ldots, \lambda_{p^*}\}$.
2. $\frac{N}{T} \to \kappa_0$ and $\frac{\lambda_{\min}^\dagger}{m^{\frac{1}{2}}N^{\frac{7}{8}}T} \to \kappa_1$, where $\lambda_{\min}^\dagger = \min\{\lambda_{p^*+1}, \ldots, \lambda_p\}$, $0 \le \kappa_0 < \infty$ and $\kappa_1 > 0$.

The conditions of Assumption 3, though seemingly complicated, can be justified. For example, let $N = \lfloor T^{b_0}\rfloor$, $m = \lfloor T^{b_1}\rfloor$, $\lambda_{\max}^* = T^{b_2}$ and $\lambda_{\min}^\dagger = T^{b_3}$, where $\lfloor a\rfloor \le a$ means the largest integer part of a real number a. Then Assumption 3 essentially requires that $0 < b_0 \le 1$, $0 < b_1 < \frac{1}{2}$, $b_2 < \frac{3}{4}b_0 + 1$ and $b_3 \ge \frac{b_1}{2} + \frac{7b_0}{8} + 1$. As for the choice of the weight parameters, λ_j's, we propose a data driven method in Section A.2 of the Supplementary Appendix A.

With Assumption 3, the following theorem holds when $\frac{N}{T} \to \kappa_0$ with $0 \le \kappa_0 < \infty$.

Theorem 3.1. *Let Assumptions 1–3.1 hold. As $(N, T) \to (\infty, \infty)$,*

1. $\|\widehat{\beta}_m - \beta_0\|_{L^2} = o_P(1)$.
   Additionally, let Assumption 3.2 hold.
2. $\Pr(\|\widehat{C}_\beta^\dagger\| = 0) \to 1$;
3. $\|\widehat{\beta}_m^* - \beta_0^*\|_{L^2} = O_P\left(\sqrt{\frac{m}{NT}} + m^{-\frac{\mu}{2}} + \frac{m\lambda_{\max}^*}{NT}\right)$, *where* $\widehat{\beta}_m^*(z) = \widehat{C}_\beta^* H_m(z)$.

The first two results of Theorem 3.1 mean that we can distinguish between $\mathcal{A}^*$ and $\mathcal{A}^\dagger$; while the third result of Theorem 3.1 provides the rate of convergence for the coefficient functions of the variables which truly drive economic growth. Moreover, provided that $\lambda_{\max}^*$ is chosen properly, the third result of Theorem 3.1 can be further simplified as $\|\widehat{\beta}_m^* - \beta_0^*\|_{L^2} = O_P\left(\sqrt{\frac{m}{NT}} + m^{-\frac{\mu}{2}}\right)$, which is the usual rate of convergence in sieve method based regressions.

To select the optimal weight parameters, we propose the following BIC type criterion:

$$\mathrm{BIC}_\lambda = \ln \mathrm{RSS}_\lambda + \mathrm{df}_\lambda \Upsilon_{NT}, \tag{3.4}$$

where $\mathrm{RSS}_\lambda = \frac{1}{NT}\sum_{i=1}^N \left(Y_i - \phi_i[\widehat{\beta}_m^\lambda]\right)' M_{\widehat{F}^\lambda}\left(Y_i - \phi_i[\widehat{\beta}_m^\lambda]\right)$, $\widehat{\beta}_m^\lambda(z) = \widehat{C}_\beta^\lambda H_m(z)$, $(\widehat{C}_\beta^\lambda, \widehat{F}^\lambda)$ are obtained by implementing (3.3) using λ as the weight vector, and df_λ is the number of nonzero coefficient functions identified using $\widehat{C}_\beta^\lambda$.

To apply the BIC in (3.4), we still need to choose Υ_{NT}. After examining the rates associated with (3.4), we require Υ_{NT} to satisfy the following condition:

$$\Upsilon_{NT} \to 0 \text{ and } \Upsilon_{NT}\sqrt[4]{N} \to \kappa_2 > 0, \text{ where } \kappa_2 \text{ is a large constant or } \infty. \tag{3.5}$$

Two values for Υ_{NT} that are commonly used in the literature and that also satisfy (3.5) are $\frac{\ln N}{\sqrt[4]{N}}$ and $\frac{N^a}{\sqrt[4]{N}}$ for some given $a \in (0, \frac{1}{4})$. In Section A.2 of the supplementary Appendix A, we consider several forms for Υ_{NT}, and investigate the finite sample performances for each of them through Monte Carlo simulations.

We select λ by

$$\widehat{\lambda} = \underset{\lambda}{\text{argmin}}\, \text{BIC}_\lambda. \tag{3.6}$$

Further letting $S_{\widehat{\lambda}} = \{j \mid \|\widehat{C}_{\beta,j}^{\widehat{\lambda}}\| > 0, 1 \le j \le p\}$ represent the set of relevant variables identified using $\widehat{C}_\beta^{\widehat{\lambda}}$, the following result then follows.

Theorem 3.2. *Let (3.5) and Assumptions 1–3 hold. Then* $\Pr(S_{\widehat{\lambda}} = \mathcal{A}^*) \to 1$ *as* $(N, T) \to (\infty, \infty)$.

Theorem 3.2 means all zero coefficient functions can be identified. In other words, all variables not driving economic growth can be identified and thus removed from the growth regression. With those variables removed, we can proceed to perform the post-selection estimation using the remaining model.

To complete the discussion on the LD case, we propose the following assumption and state the asymptotic normality associated with (3.3).

Assumption 4.

1. Suppose that for $t \ge s$, $E[f_{0t}'f_{0s} \mid \mathcal{X}_{Nt}] = a_{ts}$, and $\sum_{t=1}^T\sum_{s=1}^t |a_{ts}| = O(T)$, where $\mathcal{X}_{Nt} := \{(x_{1t}, z_{1t}), \ldots, (x_{Nt}, z_{Nt})\}$. Moreover, $\frac{NT}{m^{\mu+1}} \to 0$, $\frac{mN}{T} \to 0$, $\frac{T}{N^2} \to 0$ and $\frac{m\lambda_{\max}^*}{\sqrt{NT}} \to 0$.
2. Let $\Sigma_{\mathcal{Z}}^* = E[\mathcal{Z}_{11}^*\mathcal{Z}_{11}^{*'}]$ and $\Omega_\star = \lim_{(N,T)\to(\infty,\infty)} E[\Psi_1\Psi_1']$ both be positive definite for $\forall z \in V_z$, where

$$\Psi_1 = \sqrt{\frac{NT}{m}}\left[H_m'(z)\otimes I_{p^*}\right]\Psi_2^{-1}\Sigma_{\mathcal{Z}}^{*-1}\cdot\frac{1}{NT}\sum_{i=1}^N\left\{\mathcal{Z}_{i'} - \frac{1}{N}\sum_{j=1}^N \mathcal{Z}_j^{*'}\gamma_{0j}'\Sigma_\gamma^{-1}\gamma_{0i}\right\}\mathcal{E}_i,$$

$$\Psi_2 = I_{mp^*} + \Sigma_{\mathcal{Z}}^{*-1}E[\mathcal{Z}_{11}^*\gamma_{01}']\Sigma_\gamma^{-1}E[\gamma_{01}\mathcal{Z}_{11}^{*'}], \quad \text{and} \quad \mathcal{Z}_{it}^* = H_m(z_{it})\otimes x_{it}^*.$$

Suppose that for $\forall z \in V_z$, as $(N, T) \to (\infty, \infty)$, $\Psi_1 \to_D N(0, \Omega_\star)$.

Note that Assumption 4.1 is in the spirit of Connor et al. (2012, Eq. 3 and Eq. 20). Assumption 4.2 is equivalent to Assumption E of Bai (2009). To conserve space, we provide a detailed justification of the last part of this assumption in the supplementary Appendix A.

Theorem 3.3. *Let Assumptions 1–4 hold. For* $\forall z \in V_z$, *as* $(N, T) \to (\infty, \infty)$, $\sqrt{\frac{NT}{m}}(\widehat{\beta}_m^*(z) - \beta_0^*(z)) \to_D N(0, \Omega_\star)$, *where* $\widehat{\beta}_m^*(z) = \widehat{C}_\beta^* H_m(z)$.

It is worth emphasizing that (i). deriving the rates of convergence in Theorem 3.1 does not require Assumption 4; (ii). the asymptotic distribution in Theorem 3.3 also applies to the post-selection estimation conditional on that no regressors associated with $\mathcal{A}^*$ are removed. The associated proof is self-evident and thus is omitted here. Recently, more robust inferences on post-selection procedure have been developed, but most (if not all) of the studies (e.g., Dezeure et al., 2015; Hyun et al., 2018) focus on parametric models with i.i.d. data. It remains an open question as to how to extend these recent studies to a nonparametric panel data setting that allows for cross-sectional dependence and time series autocorrelation.

3.2. High dimensional case

In this subsection, we allow the dimension of x_{it} to diverge as the sample size increases (i.e., $p^* \to \infty$ and $p \to \infty$). The following assumptions are crucial for establishing asymptotic properties for the HD case.

Assumption 5.

1. $\|\mathcal{E}\|_{\text{sp}} = O_P(\max\{\sqrt{N}, \sqrt{T}\})$, where $\|\cdot\|_{\text{sp}}$ denotes the spectral norm of a matrix and $\mathcal{E} = (\mathcal{E}_1, \ldots, \mathcal{E}_N)'$;
2. $\frac{p^*\lambda_{\max}^*\sqrt[4]{\xi_{NT}}}{NT} \to 0$, $(\frac{\xi_{NT}+mp}{NT} + p^*m^{-\mu})\sqrt{\xi_{NT}} \to \kappa_2$, $\frac{\sqrt[8]{\xi_{NT}}\lambda_{\min}^\dagger}{NT} \to \kappa_3$, where $\xi_{NT} = \min\{N, T\}$, $0 \le \kappa_2 < \infty$ and $\kappa_3 > 0$.

Assumption 5.1 is identical to Assumption A.1.iii of Su et al. (2015). Assumption 5.2 can be verified in exactly the same way as shown in the paragraph following Assumption 3. Again, the data driven algorithm of Section A.2 of the supplementary Appendix A can help us bypass the complexity of Assumption 5.2 in practice.

With regard to the selection of weight parameters, we still use the BIC criterion defined in (3.4) and select λ by $\widehat{\lambda} = \text{argmin}_{\lambda}\, \text{BIC}_{\lambda}$ but with a different restriction on the penalty term:

$$\Upsilon_{NT} \to 0 \text{ and } \Upsilon_{NT}\xi_{NT}^{1/8} \to \kappa_5 > 0, \text{ where } \kappa_5 \text{ is a large constant or } \infty. \tag{3.7}$$

A natural choice of Υ_{NT} is $\frac{\ln \xi_{NT}}{\xi_{NT}^{1/8}}$, where ξ_{NT} is defined in Assumption 5. Again, in Section A.2 of the supplementary Appendix A, we consider three forms of Υ_{NT}, and implement numerical simulations to examine their finite sample performances.

Given the above setting, the following theorem holds.

Theorem 3.4. *Let Assumptions 1, 2 and 5 hold. As $(N, T) \to (\infty, \infty)$,*

1. $\Pr(\|\widehat{C}_{\beta}^{\dagger}\| = 0) \to 1$.
2. *Suppose that* $\frac{(p^*m)^2}{T} \to 0$, $p^*m^{-\mu} \to 0$, *and* $\frac{N}{T} \to \kappa_0 < \infty$. *Then* $\|\widehat{\beta}_m^* - \beta_0^*\|_{L^2} = O_P\left(\sqrt{\frac{p^*m}{NT}} + \sqrt{p^*m^{-\mu}} + \lambda_{\max}^* \frac{p^*m}{NT}\right)$.
3. *Suppose that* (3.7) *holds. Then* $\Pr(S_{\widehat{\lambda}} = \mathcal{A}^*) \to 1$.

It is worth emphasizing that if one is interested in consistent estimation with the purpose of distinguishing between $\mathcal{A}^*$ and $\mathcal{A}^{\dagger}$ only, the conditions imposed in the second result of Theorem 3.4 are not needed. Moreover, provided that $\lambda_{\max}^*$ is chosen properly, the second condition of Theorem 3.4 reduces to: $\|\widehat{\beta}_m^* - \beta_0^*\|_{L^2} = O_P\left(\sqrt{\frac{p^*m}{NT}} + \sqrt{p^*m^{-\mu}}\right)$, where the leading term $\sqrt{\frac{p^*m}{NT}}$ is as expected, and the rate $\sqrt{p^*m^{-\mu}}$ is due to the fact that the truncation residual $x_{it}'\Delta_m(z)$ is increasing with p^* in the HD setting.

Note that all central limit theorems established so far only apply to the finite dimensional case. For the HD case, one cannot derive any asymptotic normality for the estimates of interest unless some transformation is further employed (see Huang et al., 2008 for examples using i.i.d. data). However, performing a similar transformation in the presence of weak cross-sectional dependence and serial correlation is mathematically more involved, and thus is beyond the scope of this study.

In summary, in either case (LD or HD), when $\Pr(S_{\widehat{\lambda}} = \mathcal{A}^*) \to 1$, all zero coefficient functions can be identified. In the growth regression context, this is equivalent to saying that all variables not driving economic growth can be identified and thus removed from the growth regression. Moreover, the varying coefficients can be recovered using the sieve method, and the factor structure can be estimated by the PCA technique. Thus, all the three aforementioned issues that are prominent in the empirical growth literature (i.e., variable selection, parameter heterogeneity, and cross-sectional dependence) can be addressed simultaneously within a unified integrated framework. Before moving on to the empirical analysis, we next describe the data employed in this study.

4. Data

Of the many variables that have been found to be significantly correlated with growth in the literature, we select 61 variables using the following criteria. The first derives from our aim of obtaining comparable results with the existing literature, and the second comes from the fact that we need to work with a balanced panel. With these restrictions, the total size of our data set becomes 62 variables (including the dependent variable, the growth rate of per capita GDP) for 88 countries covering the period 1960 to 2014. The data set contains countries in different stages of development and with a wide geographic dispersion. The explanatory variables cover a wide range of different factors, including data on economic development, social issues, health, geography, politics, education and more. The variable names, their means, and standard deviations are presented in Table 1. Table 2 provides a list of the included countries.

A common practice in the literature is to take a five-year simple moving average of both dependent and independent variables[2]. This technique has the advantages of reducing the potential effects of short-term fluctuations and maintaining a high number of time series observations. Despite these advantages, this technique may still suffer from reverse causality or simultaneity, because causality between regressors and growth could go the other way as well or some regressors and growth may be simultaneously determined (e.g., Bils and Klenow, 2000). To mitigate this problem, we deviate from the common practice by measuring dependent and independent variables differently. Specifically, while the dependent variable is measured as a five-year moving average of economic growth, all explanatory variables are measured at the beginning of each five year period, with the exception of the variables related to war, geography, and terms of trade[3]

2 Another popular method of looking at annual data in empirical growth literature is to use averaged five-year period data. But, as is stressed by Soto (2003) and Attanasio et al. (2000), the use of n-year averages is not suitable because of the lost of information that it implies. In addition, Soto (2003) and Attanasio et al. (2000) point out that attempting to use data on averaged five-year periods severely limits the number of observations to draw from in the data.

3 Specifically, these variables include: fraction spent in war (each five-year period); number of war participation (each five-year period); number of revolutions (each five-year period); coups d'état and coup attempts within (each five-year period); time of independence; East Asian dummy; African dummy; European dummy; Latin American dummy; British colony dummy; Spanish colony dummy; landlocked country dummy; percentage of land area in Koeppen–Geiger tropics; percentage of land area within 100 km of ice-free coast; terms of trade; and terms of trade growth.

Table 1
Definitions of all variables in the regression.

Variables	Description	Formula	Mean	Std
EG	Economic growth rate	$\ln(\text{rgdpo}_t/\text{rgdpo}_{t-1})$	0.0363	0.0649
log(GPC)	log GDP per capita		6.0642	0.9861
csh_g	Government consumption share		0.2074	0.1163
Openness	Openness measure	csh_x + csh_m	−0.0322	0.1274
IP	Investment price, i.e., price level of capital formation		0.4213	0.3220
PGR	Population growth rates	$\ln(\text{pop}_t/\text{pop}_{t-1})$	0.0197	0.0127
Sch_P	Primary school enrolment		0.7396	0.2170
Sch_S	Secondary school enrolment		0.4672	0.3204
Sch_H	Higher education School Enrolment		0.1336	0.1581
LE	Life expectancy		0.5932	0.1102
PESS	Public education spending share in GDP		0.0399	0.0296
PIS	Public investment share	GFCF - GFCF_PS	0.0713	0.0501
Land	Land area (sq. km/1,000,000)		0.8719	2.1518
Exports	Percentage of primary export	Exports_OM + Exports_ARM	0.1978	0.2085
Mining	Fraction GDP in mining		0.0733	0.0874
Fertility	Fertility rate, total (births per woman)		4.7412	1.9724
Military	Military expenditure share in GDP		0.0295	0.0302
PCS	Public consumption share	GGFCE - PESS - Military	0.0809	0.0501
Malaria	Malaria prevalence: Incidence of malaria (per 1000 population at risk)		155.1712	245.5145
Inflation	Inflation rate		1.3854	7.0448
Political	Political rights		4.2100	1.9429
Civil	Civil liberties		4.1496	1.6628
Cap	Degree of capitalism		3.2697	1.7275
Trade	Terms of trade		1.3475	2.2807
Tra_Gro	Terms of trade growth		0.0073	0.0974
Locked	Landlocked country dummy (1, yes; 0, no)		0.2697	0.4438
Ind_Year	Time of independence $<= 1914 = 0$; $1915–1945 = 1$; $1946–1989 = 2$; $>= 1990 = 3$		1.4382	1.0382
kgatr	Percentage of land area in Koeppen–Geiger tropics		0.4017	0.4205
kgptr	Percentage of population in Koeppen–Geiger tropics		0.3915	0.4217
lcr100km	Percentage of land area within 100 km of ice-free coast		0.3788	0.3640
pop100cr	Ratio of population within 100 km of ice-free coast/navigable river to total population		0.4520	0.3728
cen_lat	latitude of country centroid		0.1522	0.2197
Bri_Col	British colony dummy (1, yes; 0, no)		0.2584	0.4378
Spa_Col	Spanish colony dummy (1, yes; 0, no)		0.1910	0.3931
Oil_OPEC	Oil-producing country dummy (1, yes; 0, no)		0.0674	0.2508
Gas	Proved reserves (cubic metres/10^12)		1.3789	6.1997
Oil	Proved reserves (bbl/10^9)		4.5521	19.1378
Chris	Percentage of Christian		0.5369	0.3807
Mus	Percentage of Muslim		0.3046	0.3825
Hin	Percentage of Hindu		0.0263	0.1211
Bud	Percentage of Buddhist		0.0410	0.1541
Fol	Percentage of Folk religion		0.0284	0.0628
Oth	Percentage of other religion		0.0037	0.0064
Jew	Percentage of Jewish		0.0019	0.0027
GS	Government spending share of GDP		0.1501	0.0686
Distortion	Real exchange rate distortions		129.6824	35.8479
OO	Outward orientation		−2.7398	0.7542
SIL	Ethnolinguistic fractionalization		0.4886	0.3127
ESP	English-speaking population in percentage		0.1762	0.2692
EA	East Asian dummy		0.0225	0.1482
AF	African dummy		0.4270	0.4947
EU	European dummy		0.1124	0.3158
LA	Latin American dummy		0.1573	0.3641
WarFrac	Fraction spent in war (1960–2014)		0.3265	0.4343

(*continued on next page*)

(Salimans, 2012). These latter explanatory variables are expected to be truly independent of contemporaneous economic growth, and thus also are measured as five-year moving averages (as with the dependent variable). This treatment further alleviates endogeneity, which is already mitigated by the use of multi-factor error structure as discussed in Section 2.

Given that recent literature on economic growth emphasizes the importance of institutional quality (Acemoglu et al., 2008; Acemoglu et al., 2019), we briefly elaborate on the data on institutional quality used in this paper. Three sets of variables are widely used in the literature to measure institutional quality. The first set is the survey indicators of institutional quality from the International Country Risk Guide (Hall and Jones, 1999; Acemoglu et al., 2001). The second set is the "polity2" variable from the Polity IV dataset (Acemoglu et al., 2008, 2019). It measures the degree of constraints

Table 1 (*continued*).

Variables	Description	Formula	Mean	Std
NoWars	No. of war participation (1960–2014)		0.8028	1.2609
Coup	coups d'etat and coup attempts within (1960–2014)		0.1870	0.4964
Revolution	Number of revolutions (1960–2014)		0.1941	0.5038
Pop_Dens	Population density/1000		0.0812	0.1135
WorkIR	Growth rate of work force	$\ln(WP_t/WP_{t-1})$	0.0210	0.0132
Polity2	Combined polity score		0.1132	6.5051

rgdpo — Size of economy (GDP in million)
pop — Population (in million)
csh_x — Share of merchandise exports
csh_m — Share of merchandise imports
WP — Fraction population of work force (1-A65-U15)
A65 — Fraction population over 65 years old
U15 — Fraction population under 15 years old
GFCF — Gross fixed capital formation
GFCF_PS — Gross fixed capital formation, private sector
Exports_OM — Percentage of ores and metals exports
Exports_ARM — Percentage of agricultural raw materials exports
GGFCE — General government final consumption expenditure share in GDP.

Table 2
Sample countries and their associated ISO 3166-1 alpha-3 codes.

AGO	Angola	HND	Honduras	PAK	Pakistan
ALB	Albania	HRV	Croatia	PAN	Panama
ARM	Armenia	HTI	Haiti	PER	Peru
AZE	Azerbaijan	IND	India	PHL	Philippines
BDI	Burundi	IRN	Iran, Islamic Republic of	POL	Poland
BEN	Benin	JAM	Jamaica	PRY	Paraguay
BFA	Burkina Faso	JOR	Jordan	RUS	Russian Federation
BGD	Bangladesh	JPN	Japan	RWA	Rwanda
BGR	Bulgaria	KAZ	Kazakhstan	SDN	Sudan
BLR	Belarus	KEN	Kenya	SEN	Senegal
BRA	Brazil	KGZ	Kyrgyzstan	SLE	Sierra Leone
BWA	Botswana	KHM	Cambodia	SLV	El Salvador
CAF	Central African Republic	LAO	Lao People's Democratic Republic	SWZ	Swaziland
CIV	Côte d'Ivoire	LBN	Lebanon	SYR	Syrian Arab Republic
CMR	Cameroon	LKA	Sri Lanka	TCD	Chad
COG	Congo	LSO	Lesotho	TGO	Togo
COL	Colombia	MDA	Moldova, Republic of	THA	Thailand
DOM	Dominican Republic	MDG	Madagascar	TTO	Trinidad and Tobago
DZA	Algeria	MEX	Mexico	TUN	Tunisia
ECU	Ecuador	MKD	Macedonia	TUR	Turkey
EGY	Egypt	MLI	Mali	TZA	Tanzania, United Republic of
ETH	Ethiopia	MNG	Mongolia	UGA	Uganda
GAB	Gabon	MOZ	Mozambique	UKR	Ukraine
GBR	United Kingdom	MWI	Malawi	URY	Uruguay
GEO	Georgia	MYS	Malaysia	USA	United States
GHA	Ghana	NAM	Namibia	VEN	Venezuela, Bolivarian Republic of
GIN	Guinea	NER	Niger	YEM	Yemen
GMB	Gambia	NIC	Nicaragua	ZAF	South Africa
GNB	Guinea-Bissau	NPL	Nepal	ZWE	Zimbabwe
GTM	Guatemala	OMN	Oman		

on politicians and politically connected elites through five indicators: intensity of political competition, regulation of political participation, competitiveness of executive recruitment, openness of executive recruitment, and the constraints it places on its chief executive. The third set includes the "civil liberties" and "political rights" from Freedom House (Barro, 1998; Sala-I-Martin, 1997; Acemoglu et al., 2019). The civil liberties index is made up of four subcategories: freedom of expression and belief, associational and organizational rights, rule of law, and personal autonomy and individual rights, while the political rights index is composed of three subcategories: electoral process, political pluralism and participation, and functioning of government. In this paper, we choose not to use the first set of variables, because it only dates back to 1984 (much later than the beginning year of our study period) and also is not available for many of our sample countries. With regard to the other two sets of institutional quality variables, we mainly rely on the third one[4], while using the second one as a robustness check.

[4] We also use secondary sources to resolve ambiguous cases or those without data coverage in Freedom House. The secondary sources are the dichotomous measures by Cheibub et al. (2010) and Boix et al. (2013). See Acemoglu et al. (2019).

5. Empirical results

In this section, we present results obtained from the varying coefficient growth regression model with factor structure and sparsity.

5.1. Choices of the number of factors, the development index and the truncation parameter

In Section 3, we assume that the number of factors r is known. In practice, r is unknown and has to be estimated. The main tool for estimating the number of factors of large dimensional datasets is the use of information criterion. In view of the fact that there are 59 observable explanatory variables in our case, we follow (Ando and Bai, 2017) to choose the number of factors by minimizing the following criterion function:

$$\mathrm{PIC}(r) = \widehat{\sigma}_{\varepsilon}^2 \cdot \left(1 + r \cdot \frac{N+T}{NT} \log(NT)\right), \tag{5.1}$$

where $\widehat{\sigma}_{\varepsilon}^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - x_{it}'\widehat{\beta}_m - \widehat{f}_t'\widehat{\gamma}_i)^2$, and for $\forall r$, $\widehat{\beta}_m$, $\widehat{f}_t$ and $\widehat{\gamma}_i$ are the corresponding estimates obtained using the approach presented in Section 3.

Given the range of the index variable, there are many basis functions we can use (e.g., Dong and Linton, 2018). The choice of basis functions is equivalent to the choice of kernel functions in the literature of kernel regression. In practice, the results are not sensitive to the choice of the basis functions, so long as $\mathcal{R}$ defined in the beginning of Section 3 can cover $[\min\{z_{it}\}, \max\{z_{it}\}]$. In this study, we adopt the Hermite functions studied in Dong and Linton (2018), and also provide some Monte Carlo simulations on these functions in the supplementary Appendix A of this paper.

We now turn to the choice of the development index z_{it}. In Section 2, we have specified a varying coefficient growth regression model capable of capturing parameter heterogeneity by means of a development index. Of the possible development indices, output and human capital are believed to be the most important ones in previous studies (e.g., Durlauf and Johnson, 1995; Salimans, 2012). Following those studies, we consider four alternative development indices[5]: (1) log initial GDP per capita, (2) initial primary schooling enrolment rate, (3) initial secondary schooling enrolment, and (4) initial higher education enrolment rate.

When choosing among the four alternative development indices and the truncation parameter m, we use the root mean squared error (RMSE) which is consistent with the criterion function used in estimation. Specifically, for each development index and each truncation parameter, we first choose the number of factors and select the regressors. Then we run post selection regression by letting the weight parameters be zero to calculate RMSE. The results are summarized in Table 4. For the purpose of comparison, we also consider the constant coefficient growth regression model with interactive fixed effects (where the coefficient β_0 is a vector of constant), and the varying coefficient (with log initial GDP per capita as the development index) panel data model with only fixed effects. Table 4 shows that the varying coefficient panel data model with log initial GDP per capita as the development index has the smallest RMSEs in general, and thus fits the data best.

Having selected log initial GDP per capita as the development index, we next move on to a more subtle question, i.e., the choice of the truncation parameter. As explained in the simulation study of Section A.2 of the supplementary Appendix A, choosing the optimal m is theoretically challenging when both the factor structure and variable selection procedure get involved. In practice, we would like to pick an m, which can explain the model as much as possible. Therefore, we use the cumulative variation of the residuals explained by the factors as our criteria in choosing m. Table 5 presents the cumulative variation of the residuals for $m = 3, 4, 5$ respectively. As can be seen, when $m = 3, 4, 5$, the number of factor identified by the data is 4, 6, 6, respectively. In addition, $m = 4$ with the number factors equal to 6 is favoured over $m = 3$ with the number of factors equal to 4 and $m = 5$ with the number of factors equal to 6.

In summary, the model with six factors and log initial GDP per capita as the development index (i.e., $r = 6$, $m = 4$ and $z =$ log initial GDP per capita) receives the most support from the data. Hence, in what follows we concentrate on the results obtained from this model.

5.2. Estimates of the common factors and loadings

Figs. 1 and 2 plot the estimates of the factors identified above and their corresponding loadings respectively. The former shows that all the common factors vary considerably over time with the exception of the first factor which exhibits a relatively small amount of variation during the sample period, while the latter shows that all the factor loadings vary substantially across countries.

5 More discussions on the index variable are provided in Section A.1.3 of the supplementary Appendix A.

5.3. Estimates of the coefficient functions of selected variables

5.3.1. General findings

We have identified a number of robust growth determinants (variables), thus broadly supporting the "optimistic" conclusion of Sala-I-Martin (1997), Fernández et al. (2001) and Sala-I-Martin et al. (2004), that is, some variables are important regressors for explaining cross-country growth patterns. Specifically, we have identified 31 "robust" growth determinants, including initial GDP per capita, institutional quality (civil liberties), human capital, trade, and macroeconomic policies, as well as natural resources, geographical characteristics, colonial origin, religion, and war-related variables. Since the coefficients in our specification are functions of log initial GDP per capita, we report in Table 3 the estimated coefficient function for each of the 31 robust growth determinants at ln(initial GDP per capita)= 3.98 (minimum), 5, 6, 7, 8, and 8.81 (maximum), together with their associated 95% bootstrapped confidence intervals (CI)[6]. To offer a more detailed look at these coefficient functions, we also plot each of them against log initial GDP per capita in Figure A.2 of the supplementary Appendix A.

Although our results broadly support the conclusion of previous studies, we note three differences in results between this and previous studies. First, our set of robust growth determinants differs from those identified in previous studies, in spite of many overlaps between them. Specifically, some variables appear to be robust in our study but not in previous studies (such as secondary school enrolment rate and terms of trade growth) or vice versa (such as primary school enrolment rate and fraction of mining in GDP). There are three possible reasons for this difference: (1) we use a different model (i.e., we allow for parameter heterogeneity and cross-sectional dependence); (2) we use a different variable selection method (i.e., we use a LASSO method while most previous studies use Bayesian averaging methods); and (3) we use a different dataset (i.e., our dataset spans a longer time period and covers a slightly different set of countries).

Second, our estimated coefficients vary considerably across countries according to their level of development, while those in most previous studies are identical across countries. Specifically, some of our estimated coefficients have the same sign but different values across different levels of initial GDP per capita (such as civil liberty, terms of trade growth, and percentage of land area in tropics), while others not only have different signs but also different magnitudes across different levels of development (such as consumption share of government, life expectancy, military expenditure, and OPEC dummy). These results suggest that it is inappropriate to apply growth regressions with homogeneous parameters.

Third, our estimated coefficients reveal some cross-country patterns not found in previous studies. Taking the coefficient of initial GDP per capita for example, while it is negative for all other countries, it is positive for countries with GDP per capita between \$1780 and \$2117 in 1960 U.S. dollars (between \$13,166 and \$15,665 in 2010 U.S. dollars). This finding is in line with the "middle income trap" hypothesis suggesting that some developing countries get stuck at middle-income levels and fail to advance into high-income countries (Eichengreen et al., 2014). To give another example, our estimated coefficient of oil reserve starts out negative, increases monotonically with GDP per capita, and then eventually becomes positive for economies with initial GDP per capital above \$2175 in 1960 U.S. dollars (\$16,094 in 2010 U.S. dollars). This latter finding is consistent with previous studies (e.g., Leite and Weidmann, 1999) that stresses the role of economic institutions in determining the effects of natural resources on growth. Specifically, these studies suggest that in developed economies where economic institutions are generally well-developed, natural resources tend to promote economic growth, whereas in developing economies where economic institutions are generally weak, natural resources tend to hamper economic growth. We will discuss these two examples in more details below.

5.3.2. Specific findings

Due to space limitations, we will not discuss all of the 31 robust growth determinants in details but, instead, will concentrate on some of the robust growth determinants that have received relatively more attention in the literature, such as initial GDP per capita, price for investment goods, human capital, natural resources, trade, religion, and institutional quality.

Fig. 3.1 presents our estimate of the coefficient of initial GDP per capita. This figure reveals three findings. First, this coefficient is negative for most GDP per capita levels, largely supporting the conditional convergence hypothesis. Second, the coefficient has an inverse U-shaped relationship with initial GDP per capita. This finding is consistent with (Durlauf et al., 2001) who finds that the coefficient of initial GDP per capita do not exhibit any sort of monotonicity with respect to level of development. It is also in line with (Salimans, 2012) who finds that coefficient of initial GDP per capita first increases with level of development up to a point and then declines afterwards. Third, the coefficient is positive for countries with GDP per capita between \$319 and \$1812 in 1960 U.S. dollars (or between \$2554 and \$14,510 in 2014 U.S. dollars), suggesting that these middle-income countries have failed to catch up with more developed countries. As an example, we find two countries in our sample (South Africa and Colombia) have never been able leave the "middle-income

6 Note that these confidence intervals need to be interpreted carefully. As well understood, one cannot establish the confidence intervals for the estimates under HD case unless certain transformation is further employed (e.g., Huang et al., 2008). However, if one regards 31 (the number of selected variables) as a relatively small number, then one can treat our regression as a LD case and employ the same bootstrap procedure as in Su et al. (2015). In order to ensure the validity of the bootstrap procedure, stronger assumptions on the error terms are needed. For example, one can employ the martingale difference type of assumptions (see Assumption A.4 of Su et al., 2015), or simply assume that the error terms are i.i.d. over both i and t. Generally speaking, when the error term exhibits both cross-sectional and serial correlation, the bootstrap results are not reliable or incorrect.

range'' over the entire sample period since their GDP per capita fell into this range at the beginning of the sample period (i.e., 1960). This finding is consistent with the ''middle income trap hypothesis'', which postulates a dim chance for middle income countries to advance to high-income status (e.g., Eichengreen et al., 2014).

Fig. 3.2 shows the coefficient of price for investment goods. Three findings emerge from this figure. First, this coefficient is negative for countries with initial GDP per capita up to $543 in 1960 U.S. dollars (or $4348 in 2014 U.S. dollars), suggesting that for these countries a relative low price of investment goods in the first year of each five-year period is strongly and positively related to subsequent income growth. This finding is not surprising because a low investment price stimulates investment (including investment in machinery and equipment), which further spurs economic growth (De Long and Summers, 1991, 1992). Second, the coefficient falls in absolute value as initial GDP per capita increases, meaning that the marginal effect of investment price on growth is stronger for poor countries than for rich countries. This latter finding is consistent with (Temple, 1999) who finds that the growth-spurring effects of investment is greater for developing countries. Such a finding may be because of the all-encompassing nature of the equipment and machinery category in the data. It is likely that purchase of any type of machinery will be of assistance in developing countries whereas growth in developed countries may be fostered more by the purchase of innovative equipment that is subsumed within the equipment and machinery category (ab Iorwerth, 2005). Third, for countries with initial GDP per capita above $543 in 1960 U.S. dollars (or $4348 in 2014 U.S. dollars) the estimated coefficient of investment goods price is positive but insignificant, because the associated confidence intervals contain zero. This suggests that investment goods price has no growth effects for these countries. A possible reason is that the data from the Penn World Table is disaggregated enough to distinguish price of equipment investment, which has strong growth effects, and price of other forms of investment, which have little growth effects (De Long and Summers, 1991).

Fig. 3.3 shows the coefficient of secondary schooling enrolment. Two findings stand out from this figure. First, for most low income countries secondary schooling enrolment is either statistically or economically insignificant. This finding can be possibly explained by Bils and Klenow (2000)'s argument on reverse causality (i.e., from growth to education). Specifically, Bils and Klenow (2000) find that the impact of schooling on growth can explain a small fraction of the empirical relationship between education and growth, while reverse causality (i.e., from growth to education) possibly accounts for most of the relationship. In our case where reverse causality and omitted variables are partially removed by the use of lagged independent variables and multi-factor structure, the coefficient of secondary schooling is likely to mainly reflect the impact of secondary schooling on growth instead of the reverse. This may in turn suggest that the impact of secondary schooling on growth may indeed be quite small for low income countries. A possible reason for this small impact is the low quality of secondary education in these countries that fails to translate additional years of secondary schooling into better cognitive skills (Hanushek and Wößmann, 2012). The second finding that emerges from Fig. 3.3 is that for most high income countries secondary schooling enrolment has a positive, strong, and statistically significant effect on growth. One plausible reason for this positive strong effect is that high income countries in general have much better quality of secondary education (Hanushek and Wößmann, 2012). The strong impact of higher education may also be partially explained by the argument of Bils and Klenow (2000). Specifically, although reverse causality is partially mitigated in this study by using lagged secondary schooling enrolment, it cannot be completely ruled out, thereby causing the coefficient of secondary schooling enrolment for high income countries to appear stronger than it actually is.

Fig. 3.4 shows that the coefficient of higher education enrolment. Three findings emerge. First, this coefficient is lower than the coefficient of secondary schooling enrolment at any level of development. This finding is consistent with the concavity argument which suggests that labour market returns are characterized by a concave relationship with education that implies decreasing returns to additional years of school (Psacharopoulos and Sanyal, 1982; Psacharopoulos and Patrinos, 2004). Second, the coefficient of higher education enrolment is lower for low income countries than for high income countries. One likely reason is the low quality of higher education in low income countries that fails to translate additional years of higher education into an increase in human capital. For example, studies find that in Middle East and North Africa countries public higher education institutions tend to focus on the production of credentials rather than the mix of skills demanded in a competitive private-sector-led economy (Psacharopoulos and Patrinos, 2018; Assaad, 2014). Another possible reason is that a high proportion of higher education graduates in low income countries choose to work in the public sector[7] because this sector provides a fertile ground for engaging in rent-seeking activities, which yield very high private returns but negative social returns (Pritchett, 2001; Owusu, 2005, 2006). A third finding that emerges from Fig. 3.4 is that the coefficient of higher education enrolment is negative for almost all countries with the exception of middle income countries. This finding is consistent with (Salimans, 2012) who, using a Bayesian model averaging method that allows for parameter heterogeneity, finds that the effect of higher education is negative for 90% of his sample countries and positive for the rest. It is also broadly consistent with (Sala-I-Martin et al., 2004) who, using a Bayesian model averaging method that does not allow for parameter heterogeneity, finds a negative effect of higher education. While the negative effect of higher education for low income countries is understandable as explained above, that for high income countries is counter-intuitive and may reflect the crudity of the higher education enrolment measure. Specifically, ''higher education'' is not clearly defined as to what types of schools (e.g., community colleges, vocational schools, trade schools, liberal arts colleges, and universities) are included and what forms of training and skills acquisition are fostered (Becker,

[7] For example, according to Darvas et al. (2017), in Sub-Saharan countries approximately 50% of higher education graduates were employed by the public sector in 2011. This figure is even higher prior to 2011.

1992). In addition, as is well known in the empirical growth literature, years of schooling is a crude measure of schooling experience, as it ignores the influence of student–teacher ratios, teacher experience, peer grouping, curricula, and other factors that are believed to influence the amount of knowledge and skills embodied in individuals during the years of schooling (Becker, 1992).

Fig. 3.5 shows the coefficient of oil reserve. Two findings from this figure are noteworthy. First, this coefficient is negative for countries with initial GDP per capita below $2373 in 1960 U.S. dollars (or $19,002 in 2014 U.S. dollars). This finding is consistent with the "natural resources curse hypothesis" (e.g., Sachs and Warner, 2001) and can be explained by the rent-seeking behaviour of countries with large endowments of natural resources. Second, this coefficient increases monotonically with GDP per capita and eventually becomes positive for economies with initial GDP per capita above $2373 in 1960 U.S. dollars (or $19,002 in 2014 U.S. dollars). This latter finding is line with recent studies (e.g., Leite and Weidmann, 1999) that suggest that the contribution of natural resources to a country's economy does not take place in isolation, but rather in the overall context of the country's economic management and institutions. Put differently, it is the quality and competency of these policies and institutions that determines whether natural resources can promote economic growth. Specifically, in developed economies where economic institutions are generally well-developed, natural resources tend to promote economic growth; whereas in developing economies where economic institutions are generally weak, natural resources tend to hamper economic growth.

Fig. 3.6 presents the coefficient of terms of trade growth. This coefficient is positive for all countries, suggesting that growth tends to be faster in countries where terms of trade growth is higher. This finding is consistent with previous studies (Bleaney and Greenaway, 2001) that find that an improvement in terms of trade leads to higher levels of investment and hence long-run economic growth. In addition, this figure shows that this coefficient increases with GDP per capita, indicating that the marginal effect of terms of trade growth is larger in richer countries than in poorer ones. This latter finding is consistent with recent studies (e.g., Blattman et al., 2007) that suggest that higher volatility in terms of trade reduces investment and hence growth because of aversion to risk. Therefore, terms of trade instability is likely to have a smaller negative impact on rich countries because these countries have more sophisticated institutions and markets and thus are more likely to have cheaper ways to insure against price volatility than poor countries.

Here we note that as in Sala-I-Martin et al. (2004), trade openness (defined as exports plus imports as a share of GDP) is insignificant, presumably reflecting the crudity of this measure, and perhaps the distinction between opening to international trade generating a one-time step increase in income as factors are reallocated according to comparative advantage versus an ongoing growth impact associated with greater openness.

Fig. 3.7–3.9 present the coefficients of fraction of Christian, Muslim, and Jewish respectively. These coefficients are negative at all levels of development or nearly all levels of development. This finding is consistent with (Barro and McCleary, 2005) who find that religion works via belief, not practice. They argue that higher church attendance uses up time and resources and eventually runs into diminishing returns. The "religion sector", as they call it, can consume more than it yields.

We are also very interested in the two institutional quality variables (i.e., civil liberties and political rights). From Figure A.2 of the supplementary Appendix A, we see that civil liberties is a significant growth determinant while political rights is not. This finding is consistent with (Acemoglu et al., 2019) who finds that among the components of democratic institutions that matter for growth, civil liberties is the most important. For readers' convenience, we reproduce the estimated coefficient of civil liberties in Fig. 3.10. As can be seen, this coefficient is positive for nearly all level of GDP per capita, thus being consistent with (Acemoglu et al., 2019) who find that democratic institutions are not only good for developed economies but also for low income economies. To investigate the robustness of this result, we re-estimated our model using an alternative measure of institutional quality — the "polity2" variable from the Polity IV dataset. Our results indicate that "polity2" is still identified as a significant growth determinant. In addition, Figure A.3 shows that the estimated coefficient of "polity2" is still statistically non-negative for most of the sample countries, confirming that institutions are important for growth. Here we note that the magnitude of the estimated coefficient of "polity2" is much smaller than that of civil liberties. The finding is consistent with that of BenYishay and Betancourt (2010) who find that civil liberties (i.e., individual freedoms and rights) dominates constraints on politicians in every possible comparison in terms of predictive performance and statistical significance.

6. Robustness check

In order to further confirm the above empirical results, we conduct a number of robustness checks. First, we examine the sensitivity of our results by considering alternative choices of Υ_{NT}. Second, we test for autocorrelation and cross-sectional dependence in residuals in order to investigate the validity of our bootstrap procedure. Finally, we consider longer lags for initial GDP per capita to assess if the above findings carry through.

6.1. Different choices of Υ_{NT}

We consider three alternative choices for Υ_{NT}: (1) $\frac{\ln \xi_{NT}}{\sqrt[8]{\xi_{NT}}}$, (2) $\frac{\ln(N+T)}{\sqrt[8]{N+T}}$, (3) $\frac{\sqrt[10]{\xi_{NT}}}{\sqrt[8]{\xi_{NT}}}$. Note that these three choices are taken from the HD case in our simulation studies presented in supplementary Appendix A, where the three choices of Υ_{NT} have been proved to be reasonable. Our results show that the variables selected using each of the three alternative Υ_{NT}'s, their associated coefficients, and the number of factors are identical to those presented in Section 5, indicating that our empirical results are robust to different choices of Υ_{NT}.

Table 3
Estimates of coefficients at log(GPC) = 3.98, 5, 6, 7, 8 and 8.81.

	log(GPC) = 3.98	log(GPC) = 5	log(GPC) = 6	log(GPC) = 7	log(GPC) = 8	log(GPC) = 8.81
log(GPC)	−0.0766	−0.0447	0.0117	0.0255	−0.0531	−0.2055
	(−0.1145, −0.0429)	(−0.0618, −0.0305)	(−0.0038, 0.0249)	(0.0080, 0.0469)	(−0.0886, −0.0034)	(−0.2883, −0.0898)
csh_g	0.3341	0.1483	−0.0587	−0.2143	−0.2619	−0.1964
	(0.1077, 0.5534)	(0.0706, 0.2248)	(−0.1110, −0.0107)	(−0.3141, −0.1454)	(−0.5782, 0.0082)	(−0.9416, 0.4733)
IP	−0.1284	−0.0545	−0.0093	0.0160	0.0282	0.0325
	(−0.1971, −0.0621)	(−0.0686, −0.0422)	(−0.0191, 0.0059)	(−0.0046, 0.0406)	(−0.0146, 0.0638)	(−0.0895, 0.1209)
PGR	3.2821	0.5353	0.8485	1.5689	0.6502	−2.2853
	(1.7325, 4.7072)	(0.0426, 0.9486)	(0.4054, 1.2542)	(1.1371, 1.9781)	(−0.1228, 1.3620)	(−3.5441, −0.7206)
School_S	0.1322	0.0823	−0.0420	−0.0147	0.3303	0.9119
	(−0.1505, 0.4845)	(0.0108, 0.1495)	(−0.0780, −0.0081)	(−0.0548, 0.0177)	(0.1989, 0.4201)	(0.5495, 1.1719)
School_H	−1.5869	−0.1175	0.2220	0.0326	−0.2066	−0.1808
	(−2.4454, −0.7631)	(−0.3214, 0.1119)	(0.1562, 0.3020)	(−0.0157, 0.0921)	(−0.2810, −0.1039)	(−0.3986, 0.1401)
LE	0.3148	0.2935	−0.0597	−0.2370	0.1455	1.0373
	(−0.0171, 0.6403)	(0.1811, 0.3938)	(−0.1739, 0.0389)	(−0.3918, −0.0990)	(−0.2875, 0.5411)	(0.0000, 1.9922)
PESS	−1.2399	−0.5737	−0.2476	−0.0158	0.3070	0.7260
	(−2.4950, 0.0107)	(−0.7969, −0.3513)	(−0.4210, −0.0944)	(−0.2163, 0.1716)	(−0.6249, 1.3302)	(−1.5776, 3.4530)
Military	−1.1748	−0.1436	−0.0428	−0.0020	0.6417	1.9197
	(−2.4253, −0.0639)	(−0.3588, 0.1047)	(−0.1667, 0.0719)	(−0.2609, 0.1609)	(0.0134, 1.1085)	(0.1982, 3.4369)
Inflation	−0.0055	0.0006	−0.0011	−0.0031	0.0007	0.0107
	(−0.0094, −0.0012)	(0.0000, 0.0012)	(−0.0016, −0.0005)	(−0.0039, −0.0023)	(−0.0043, 0.0048)	(−0.0020, 0.0209)
Civil	0.0183	0.0050	−0.0004	0.0033	0.0166	0.0340
	(0.0048, 0.0308)	(0.0011, 0.0085)	(−0.0027, 0.0024)	(0.0000, 0.0064)	(0.0027, 0.0282)	(−0.0036, 0.0661)
Tra_Gro	0.0009	0.0304	0.0048	0.0018	0.0796	0.2279
	(−0.1138, 0.1100)	(0.0117, 0.0458)	(−0.0162, 0.0252)	(−0.0248, 0.0351)	(−0.0062, 0.1727)	(−0.0124, 0.4799)
kgatr	−0.3099	−0.0815	−0.0763	−0.1422	−0.1597	−0.0776
	(−0.4367, −0.1693)	(−0.1281, −0.0391)	(−0.1161, −0.0447)	(−0.1887, −0.0969)	(−0.2602, −0.0333)	(−0.3270, 0.2396)
lcr100km	0.6793	0.1930	0.0511	−0.0621	−0.3850	−0.9133
	(0.5005, 0.8498)	(0.1396, 0.2461)	(0.0127, 0.0940)	(−0.1305, −0.0005)	(−0.5977, −0.2168)	(−1.4617, −0.5189)
cen_lat	0.5894	0.0293	0.0360	0.1693	0.0873	−0.3058
	(0.3011, 0.8198)	(−0.0723, 0.1294)	(−0.0315, 0.1114)	(0.0985, 0.2655)	(−0.0526, 0.2459)	(−0.6344, 0.0789)
Spa_Col	−0.5174	−0.1564	−0.0204	0.0548	0.1945	0.4167
	(−0.6748, −0.3222)	(−0.2111, −0.1080)	(−0.0503, 0.0091)	(0.0171, 0.0950)	(0.1040, 0.3131)	(0.1867, 0.6800)
Oil_OPEC	−0.1254	−0.0460	−0.0947	−0.0815	0.1358	0.5247
	(−1.0835, 0.6246)	(−0.2325, 0.1181)	(−0.1496, −0.0423)	(−0.1383, −0.0344)	(−0.0124, 0.2413)	(0.1090, 0.8785)
Oil	−0.0100	−0.0021	−0.0008	−0.0011	0.0008	0.0057
	(−0.0217, 0.0018)	(−0.0051, 0.0009)	(−0.0017, 0.0000)	(−0.0022, −0.0001)	(−0.0010, 0.0023)	(0.0004, 0.0099)
Chris	0.2689	0.1112	−0.0056	−0.1262	−0.2818	−0.4469
	(0.1164, 0.4267)	(0.0650, 0.1711)	(−0.0535, 0.0461)	(−0.2001, −0.0629)	(−0.4824, −0.1333)	(−0.8908, −0.0694)
Mus	−0.0736	0.0130	−0.0292	−0.1685	−0.3744	−0.5692
	(−0.1864, 0.0735)	(−0.0383, 0.0599)	(−0.0774, 0.0195)	(−0.2344, −0.1004)	(−0.5710, −0.2335)	(−1.0085, −0.2474)
Oth	−3.7386	0.6890	−1.1133	2.9016	21.6126	51.3537
	(−13.0771, 4.1928)	(−1.3860, 2.4081)	(−3.0584, 0.7782)	(0.2919, 6.3889)	(11.6269, 32.3911)	(27.3474, 75.2005)
Jew	−32.4452	−10.5477	−2.0163	−0.9826	−2.7502	−3.6114
	(−55.7335, −10.0679)	(−18.6022, −3.8609)	(−7.6269, 3.6243)	(−6.9105, 4.0461)	(−12.1471, 5.8912)	(−27.3730, 18.2231)
GS	−0.2516	−0.2079	−0.0476	−0.0738	−0.5101	−1.2583
	(−0.5709, 0.0529)	(−0.3047, −0.1254)	(−0.1142, 0.0215)	(−0.1676, 0.0172)	(−0.8230, −0.1881)	(−2.1023, −0.4396)
Distortion	0.0231	0.0884	0.1085	0.0482	−0.1148	−0.3280
	(−0.0449, 0.1048)	(0.0580, 0.1187)	(0.0701, 0.1435)	(−0.0064, 0.1025)	(−0.2793, 0.0061)	(−0.6718, −0.0148)
OO	1.0569	4.1811	5.1496	2.3016	−5.4384	−15.5641
	(−2.1588, 4.9579)	(2.7557, 5.6164)	(3.3263, 6.8047)	(−0.2707, 4.8851)	(−13.1848, 0.3040)	(−31.9253, −0.7429)
ESP	0.3819	0.1039	0.0581	0.1010	0.1193	0.0577
	(0.2151, 0.5483)	(0.0575, 0.1465)	(0.0214, 0.0975)	(0.0494, 0.1631)	(−0.0059, 0.2319)	(−0.2260, 0.2967)
EA	0.2924	0.0841	0.0263	−0.1256	−0.5515	−1.1753
	(−0.1574, 0.7975)	(−0.0589, 0.2081)	(−0.0668, 0.1185)	(−0.2975, 0.0293)	(−1.0307, −0.1357)	(−2.2484, −0.2191)
EU	−1.0237	−0.2206	−0.0275	0.0738	0.4780	1.2174
	(−1.7313, −0.3070)	(−0.3896, −0.0438)	(−0.0797, 0.0197)	(0.0169, 0.1225)	(0.3182, 0.6845)	(0.7527, 1.7964)
WarFrac	−0.0044	−0.0035	−0.0120	−0.0130	0.0058	0.0411
	(−0.0372, 0.0351)	(−0.0121, 0.0072)	(−0.0179, −0.0040)	(−0.0208, −0.0036)	(−0.0157, 0.0235)	(−0.0227, 0.0882)
Coup	−0.0190	−0.0004	−0.0028	−0.0070	0.0016	0.0251
	(−0.0287, −0.0089)	(−0.0031, 0.0020)	(−0.0051, −0.0005)	(−0.0102, −0.0038)	(−0.0086, 0.0113)	(−0.0005, 0.0502)
Revolution	−0.0052	0.0044	−0.0016	−0.0066	0.0021	0.0252
	(−0.0198, 0.0090)	(0.0007, 0.0075)	(−0.0042, 0.0011)	(−0.0106, −0.0020)	(−0.0054, 0.0084)	(−0.0034, 0.0464)

6.2. Testing for autocorrelation and cross-sectional dependence in residuals

As discussed in the footnote of Section 5.3.1, the validity of the bootstrap procedure used for generating the confidence intervals for the coefficient functions hinges on some assumptions concerning the error terms. For example, Su et al. (2015) employ the martingale difference type of assumptions (see Assumption A.4 of their paper). Alternatively, one can assume

Table 4
Comparison among different models.

	z of varying coefficient model				FE	CC
	ln(GPC)	School P	School S	School H		
$m = 3$	0.0199	0.0222	0.0217	0.0277	0.0346	
$m = 4$	0.0170	0.0202	0.0231	0.0231	0.0310	0.0220
$m = 5$	0.0160	0.0179	0.0184	0.0168	0.0295	

1. "FE" refers to the varying coefficient (with ln(GPC) as the development index) panel data model with fixed effects
2. "CC" refers to the constant coefficient panel data model with interactive fixed effects, where the coefficients are constant and six factors are selected.

Table 5
Cumulative variation of the residuals explained by the factors with different truncation parameters.

No. Factors	1	2	3	4	5	6
$m = 3$	67.43%	82.83%	89.65%	93.45%		
$m = 4$	**87.86%**	**94.87%**	**96.65%**	**98.10%**	**98.71%**	**99.01%**
$m = 5$	71.78%	88.84%	93.77%	95.70%	97.17%	97.99%

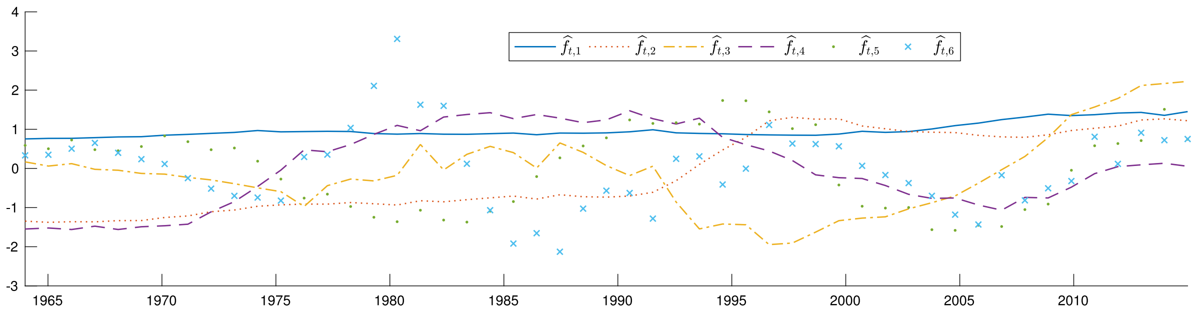

Fig. 1. Estimates of Common Factors.

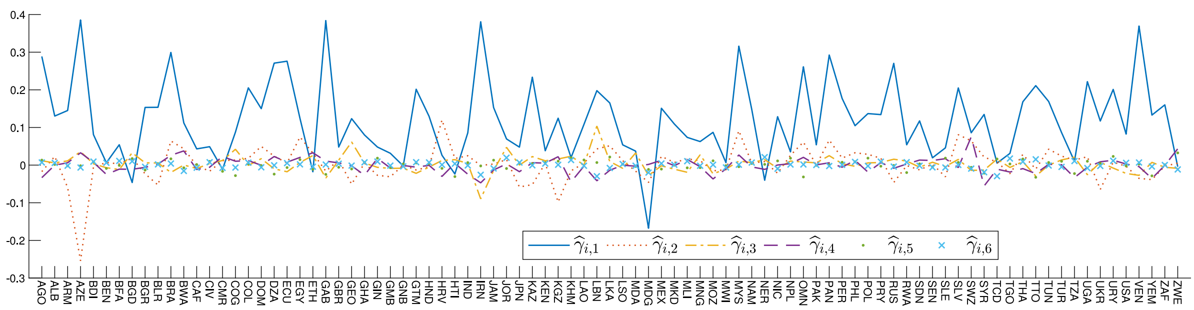

Fig. 2. Estimates of Factor Loadings.

that the error terms are cross-sectionally independent. Put differently, as long as we can show that correlation along either of the two dimensions (the time and cross-sectional dimensions) is negligible, the confidence intervals produced using the bootstrap procedure are reliable if one regards the number of selected variables (i.e., 31) as relatively small.

Specifically, we conduct two tests: (1) the Ljung–Box Q-Test to test for autocorrelation in each of the time series $\{\widehat{\varepsilon}_{i1}, \ldots, \widehat{\varepsilon}_{iT}\}$; and (2) a cross-sectional dependence (CD) test to examine cross-sectional dependence in residuals. The test statistic of this latter test is of the following form:

$$\mathrm{CD} = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{ij},$$

where ρ_{ij} measures the correlation between $\{\widehat{\varepsilon}_{i1}, \ldots, \widehat{\varepsilon}_{iT}\}$ and $\{\widehat{\varepsilon}_{j1}, \ldots, \widehat{\varepsilon}_{jT}\}$. See Juodis and Reese (2019) for more details on the CD test.

Our results from the Ljung–Box Q-Test show that 50.56% individual countries reject the null hypothesis of no autocorrelation at 1% significance level, while our statistic from the CD test is 0.2876, thus failing to reject the null that

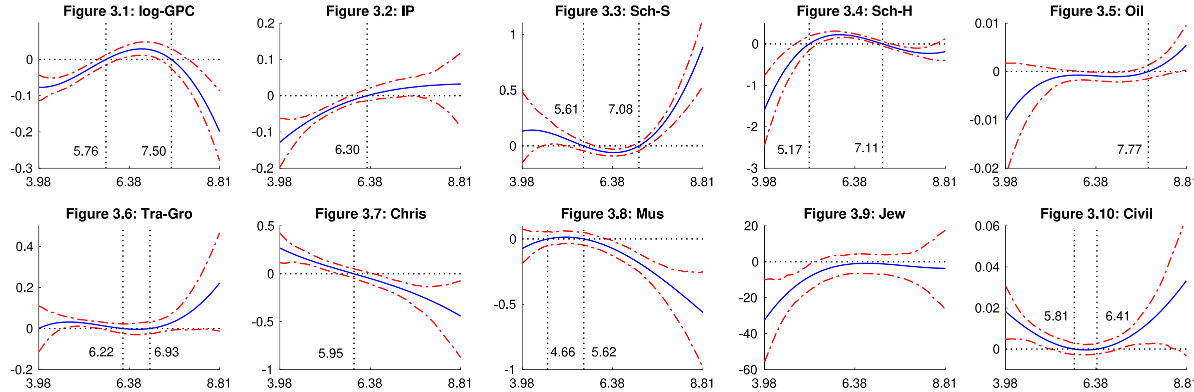

Fig. 3. Estimates of Selected Coefficient Functions.

there is no cross-sectional independence. These results suggest that correlation along the cross-sectional dimension is negligible due to the use of the factor structure, thus ensuring the validity of the bootstrap procedure used in our empirical study.

6.3. Initial GDP per capita with longer lags

In Section 5 the dependent variable is measured as a five-year moving average of economic growth, while the explanatory variables (including initial GDP per capita) are measured at the beginning of each five-year period with the exception of the variables related to war, geography, and terms of trade (Salimans, 2012). In other words, in Section 5 initial GDP per capita is lagged by three years. However, as is well known in the literature, initial GDP per capita may have a much longer effect on GDP growth. Therefore, in the subsection we check the robustness of our empirical results by using longer lags for initial GDP per capita. Specifically, we lag initial GDP per capita by 20 years whenever it is possible. For the first nineteen periods initial GDP per capita lagged by 20 years is not available, we use the initial GDP per capita in 1960. We also estimate the model using GDP per capita lagged by 15 and 25 years respectively, and the results are qualitatively similar to those reported below. Thus, in what follows we concentrate on results of the lag length of 20.

As shown in Figure A.4 of the supplementary Appendix A, our findings using lag length of 20 are, in general, consistent with those presented in Section 5. Specifically, a comparison of this figure and Figure A.2 reveals that most of the variables selected here coincide with those selected in Section 5, implying that most of the selected variables are robust to the choice of lag. More specifically, there are 21 variables that are selected in both cases. These variables include initial GDP per capita, institutional quality (civil liberties), high school enrolment, terms of trade growth, etc. In addition, we find that these variables exhibit cross-country patterns very similar to those obtained in Section 5. For example, the coefficient of initial GDP per capita is negative for all countries with the exception of middle income countries, confirming the "middle-income trap hypothesis" discussed above. To give another example, the coefficient of civil liberties is positive for nearly all countries, confirming that institutional quality is important for growth.

7. Conclusion

A rigorous cross-country growth regression analysis should simultaneously account for three major problems identified in the literature — variable selection, parameter heterogeneity, and cross-sectional dependence. Though these three problems have received individual attention, little or no research has sought to integrate them into a single, comprehensive framework. The purpose of this study is to fill this void by proposing a new, integrated framework that is capable of dealing with parameter heterogeneity and cross-sectional dependence, while simultaneously performing variable selection. Specifically, parameter heterogeneity is allowed for by means of a varying coefficient growth regression model, while cross-sectional dependence is introduced into the model via a multi-factor structure. For simplicity, we refer to the resulting growth regression model as the "varying coefficient growth regression model with factor structure and sparsity". We then propose a LASSO estimator that is capable of performing variable selection on this model. In addition, we have established the associated asymptotic results for this estimator and further investigate the performance of the estimator by conducting extensive simulations.

We apply the above framework to a new data set that covers 89 countries over the period from 1960 to 2014. We have identified 31 robust growth determinants, providing evidentiary support for the canonical neoclassical growth variables; i.e., initial income, investment, and population growth, as well as macroeconomic policies, geography, institutions, religion and ethnic fractionalization. Moreover, we find that all the coefficients of the robust growth determinants vary considerably across countries according to their level of development, which reveals some interesting cross-country patterns not found previously.

Acknowledgements

The authors would like to thank the Co-Editor: Professor Elie Tamer, the Associate Editor and the two referees for their constructive suggestions and comments on earlier versions of this paper. Gao would like to acknowledge the Australian Research Council Discovery Projects for its financial support under Grant Numbers: DP170104421 and DP200102769. Peng would like to acknowledge the Australian Research Council Discovery Projects for its financial support under Grant Number: DP210100476.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2020.09.009.

References

Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. Amer. Econ. Rev. 91 (5), 1369–1401.
Acemoglu, D., Johnson, S., Robinson, J.A., Yared, P., 2008. Income and democracy. Amer. Econ. Rev. 98 (3), 808–842.
Acemoglu, D., Naidu, S., Restrepo, P., Robinson, J.A., 2019. Democracy does cause growth. J. Polit. Econ. 127 (1), 47–100.
Ando, T., Bai, J., 2017. Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. J. Amer. Statist. Assoc. 112 (519), 1182–1198.
Assaad, R., 2014. Making sense of arab labor markets: The enduring legacy of dualism. IZA J. Labor Dev. 3 (1), 1–25.
Attanasio, O., Picci, L., Scorcu, A.E., 2000. Saving, growth, and investment: A macroeconomic analysis using a panel of countries. Rev. Econ. Stat. 82 (2), 182–211.
Bai, J., 2009. Panel data models with interactive fixed effects. Econometrica 77 (4), 1229–1279.
Barro, R.J., 1991. Economic growth in a cross section of countries. Q. J. Econ. 106 (2), 407–443.
Barro, R.J., 1998. Determinants of Economic Growth: A Cross-Country Empirical Study. The MIT Press.
Barro, R.J., McCleary, R.M., 2005. Which countries have state religions? Q. J. Econ. 120 (4), 1331–1370.
Becker, W.E., 1992. Why go to college? The value of an investment in higher education. In: Becker, W.E., Lewis, D.R. (Eds.), The Economics of American Higher Education. Springer, Dordrecht.
BenYishay, A., Betancourt, R., 2010. Civil liberties and economic development. J. Inst. Econ. 6 (3), 281–304.
Bils, M., Klenow, P.J., 2000. Does schooling cause growth? Amer. Econ. Rev. 90 (5), 1160–1183.
Blattman, C., Hwang, J., Williamson, J.G., 2007. The impact of the terms of trade on economic development in the periphery, 1870–1939: Volatility and secular change. J. Dev. Econ. 82 (1), 156–179.
Bleaney, M., Greenaway, S.D., 2001. The impact of terms of trade and real exchange rate volatility on investment and growth in sub-Saharan Africa. J. Dev. Econ. 65 (2), 491–500.
Boix, C., Miller, M., Rosato, S., 2013. A complete data set of political regimes, 1800–2007. Comparative Polit. Stud. 46 (12), 1523–1554.
Brock, W.A., Durlauf, S.N., 2001. Discrete choice with social interactions. Rev. Econom. Stud. 68 (2), 235–260.
Cheibub, J.A., Gandhi, J., Vreeland, J.R., 2010. Democracy and cictatorship revisited. Publ. Choice 143 (1/2), 67–101.
Chen, X., 2007. Large sample sieve estimation OF semi-nonparametric models. In: Handbook of Econometrics, Vol. 6. pp. 5549–5632 (chapter 76).
Chudik, A., Mohaddes, K., Pesaran, M.H., Raissi, M., 2017. Is there a debt-threshold effect on output growth? Rev. Econ. Stat. 99 (1), 135–150.
Chudik, A., Pesaran, M.H., Tosetti, E., 2011. Weak and strong cross-section dependence and estimation of large panels. Econom. J. 14 (1), C45–C90.
Connor, G., Hagmann, M., Linton, O., 2012. Efficient semiparametric estimation of the Fama–French model and extensions. Econometrica 80 (2), 713–754.
Darvas, P., Gao, S., Shen, Y., Bawany, B., 2017. Sharing higher education's promise beyond the few in sub-Saharan Africa. http://documents.worldbank.org/curated/en/862691509089826066/Sharing-higher-education-s-promise-beyond-the-few-in-Sub-Saharan-Africa.
De Long, J.B., Summers, L., 1991. Equipment investment and economic growth. Q. J. Econ. 106 (2), 445–502.
De Long, J., Summers, L., 1992. Equipment investment and economic growth: How strong is the nexus? Brook. Pap. Econ. Act. 2, 157–199.
Dezeure, R., Bühlmann, P., Meier, L., Meinshausen, N., 2015. High-dimensional inference: Confidence intervals, p-values and R-software hdi. Statist. Sci. 30 (4), 533–558.
Dong, C., Linton, O., 2018. Additive nonparametric models with time variable and both stationary and nonstationary regressors. J. Econometrics 207 (1), 212–236.
Dudley, R., 2003. Real Analysis and Probability. In: Cambridge studies in advanced mathematics 74, Cambridge University Press, Cambridge, U.K..
Durlauf, S.N., Johnson, P.A., 1995. Multiple regimes and cross country growth behaviour. J. Appl. Econometrics 10 (4), 365–384.
Durlauf, S., Johnson, P.A., Temple, J.R., 2005. Growth econometrics. In: Handbook of Macroeconomics, Vol. 1. pp. 555–677, Part A.
Durlauf, S., Kourtellos, A., Minkin, A., 2001. The local solow growth model. Eur. Econ. Rev. 45 (4–6), 928–940.
Durlauf, S.N., Kourtellos, A., Tan, C.M., 2008. Are any growth theories robust. Econom. J. 118 (527), 329–346.
Durlauf, S.N., Quah, D.T., 1999. The new empirics of economic growth. In: Handbook of Macroeconomics, Vol. 1. pp. 235–308, Part A.
Eichengreen, B., Park, D., Shin, K., 2014. Growth slowdowns redux: New evidence on the middle-income trap. Japan World Econ. 32, 65–84.
Fan, J., Li, R., 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. J. Amer. Statist. Assoc. 96 (456), 1348–1360.
Fan, J., Liao, Y., Wang, W., 2016. Projected principal component analysis in factor models. Ann. Statist. 44 (1), 219–254.
Fernández, C., Ley, E., Steel, M.F.J., 2001. Model uncertainty in cross-country growth regressions. J. Appl. Econometrics 16 (5), 563–576.
Hall, R.E., Jones, C.I., 1999. Why do some countries produce so much more output per worker than others? Q. J. Econ. 114 (1), 83–116.
Hanushek, E.A., Wößmann, 2012. Education quality and economic growth. In: The 4 Percent Solution: Unleashing The Economic Growth America Needs. pp. 227–239 (chapter 16).
Huang, J., Horowitz, J.L., Ma, S., 2008. Asymptotic properties of bridge estimators in sparse high-dimensional regression models. Ann. Statist. 36 (2), 587–613.
Hyun, S., G'Sell, M., Tibshirani, R.J., 2018. Exact post-selection inference for the generalized lasso path. Electron. J. Stat. 12 (1), 1053–1097.
ab Iorwerth, A., 2005. Machines and the economics of growth. In: Working paper available at https://EconPapers.repec.org/RePEc:fca:wpfnca:2005-05.
Juodis, A., Reese, S., 2019. The incidental parameters problem in testing for remaining cross-section correlation. arXiv:1810.03715.
Kormendi, R.C., Meguire, P.G., 1985. Macroeconomic determinants of growth: Cross-country evidence. J. Monetary Econ. 16 (2), 141–163.
Lam, C., Yao, Q., 2012. Factor modeling for high-dimensional time series: Inference for the number of factors. Ann. Statist. 40 (2), 694–726.
Leamer, E., 1983. Let's take the con out of econometrics. Amer. Econ. Rev. 73 (1), 31–43.

Leite, C., Weidmann, J., 1999. Does mother nature corrupt? natural resources, corruption, and economic growth. IMF Working paper 99/85.
Lucas, R., 1993. Making a miracle. Econometrica 61 (2), 251–272.
Mankiw, N.G., Romer, D., Weil, D.N., 1992. A contribution to the empirics of economic growth. Q. J. Econ. 107 (2), 407–437.
Moral-Benito, E., 2012. Determinants of economic growth: A Bayesian panel data approach. Rev. Econ. Stat. 94 (2), 566–579.
Newey, W.K., 1997. Convergence rates and asymptotic normality for series estimators. J. Econometrics 79 (1), 147–168.
Owusu, F., 2005. Opportunities and challenges for development in Africa: An introduction. Geojournal 62 (1–2), 115–116.
Owusu, F.Y., 2006. On public organizations in Ghana: what differentiates good performers from poor performers? Afr. Dev. Rev. 18 (3), 471–485.
Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74 (4), 967–1012.
Pritchett, L., 2001. Where has all the education gone? World Bank Econ. Rev. 15 (3), 367–391.
Psacharopoulos, G., Patrinos, H.A., 2004. Returns to investment in education: A further update. Educ. Econ. 12 (2), 111–134.
Psacharopoulos, G., Patrinos, H.A., 2018. Returns to investment in education: A decennial review of the global literature. Educ. Econ. 26 (5), 445–458.
Psacharopoulos, G., Sanyal, B., 1982. Student expectations and graduate market performance in Egypt. Higher Educ. 11, 27–49.
Sachs, J.D., Warner, A., 2001. The curse of natural resources. European Econ. Rev. 45 (4–6), 827–838.
Sala-I-Martin, X., 1997. I just ran two million regressions. Amer. Econ. Rev. 87 (2), 178–183.
Sala-I-Martin, X., Doppelhofer, G., Miller, R.I., 2004. Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. Amer. Econ. Rev. 94 (4), 813–835.
Salimans, T., 2012. Variable selection and functional form uncertainty in cross-country growth regressions. J. Econometrics 171 (2), 267–280.
Soto, M., 2003. Taxing capital flows: An empirical comparative analysis. J. Dev. Econ. 72 (1), 203–221.
Su, L., Jin, S., Zhang, Y., 2015. Specification test for panel data models with interactive fixed effects. J. Econometrics 186 (1), 222–244.
Summers, R., Heston, A., 1991. The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988. Q. J. Econ. 106 (2), 327–368.
Temple, J., 1999. The new growth evidence. J. Econ. Lit. 37 (1), 112–156.
Wang, H., Xia, Y., 2009. Shrinkage estimation of the varying coefficient. J. Amer. Statist. Assoc. 104 (486), 747–757.
Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. J. R. Stat. Soc. Ser. B Stat. Methodol. 68 (1), 49–67.