



THIS TIME THEY ARE DIFFERENT:
HETEROGENEITY AND NONLINEARITY
IN THE RELATIONSHIP BETWEEN
DEBT AND GROWTH

Markus Eberhardt Andrea F. Presbitero

Working paper no. 92

December 2013

This Time They're Different: Heterogeneity and Nonlinearity in the Relationship between Debt and Growth*

Markus Eberhardt^{a,b}

Andrea F. Presbitero^{c,d}

^a School of Economics, University of Nottingham, UK

^b Centre for the Study of African Economies, Department of Economics, University of Oxford, UK

^c International Monetary Fund, Washington DC, USA

^e Money and Finance Research group (MoFiR), Ancona, Italy

2nd December 2013

Abstract: We study the long-run relationship between public debt and growth in a large panel of countries. Our analysis takes particular note of theoretical arguments and data considerations in modelling the debt-growth relationship as heterogeneous across countries. We investigate the issue of nonlinearities ('debt thresholds') in both the cross-country and within-country dimensions, employing novel methods and diagnostics from the time-series literature adapted for use in the panel. We find some support for a nonlinear relationship between debt and long-run growth across countries, but no evidence for a common debt threshold within countries over time.

Keywords: economic growth; public debt; common factor model; nonlinearity

JEL classification: E62, F34, C23, O11

*We are grateful to Thorsten Beck, Vanessa Berenguer-Rico, Andrew Berg, Gianluca Cafiso, Tito Cordella, Panicos Demetriades, Jerry Hausman, George Kapetanios, M. Hashem Pesaran, Kenneth Rogoff, Nicolas van de Sijpe and Francesco Venturini for helpful comments and suggestions, and seminar/session participants at Leicester, Strathclyde, the Annual CSAE Conference in Oxford, Nottingham, the 3rd Development Economics Workshop in Tilburg, the ICEEE in Genova, the RES Conference at Royal Holloway, the 4th RCEA Workshop on Developments in Macro and Growth, the Conference on Cross Section Dependence in Panel Data Models in Cambridge and the International Panel Data Conference at Cass Business School for feedback on earlier drafts of the paper. All remaining errors are our own. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

Correspondence: Markus Eberhardt, School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG7 2RD (UK). Tel: +44 115 846 8416. Email: markus.eberhardt@nottingham.ac.uk. Andrea Presbitero (corresponding author), International Monetary Fund, 700 19th Street NW, Washington DC 20431 (USA). Tel: +1 202 623 9899. Email: apresbitero@imf.org

1 Introduction

The presence of a common threshold or turning point beyond which the detrimental impact of debt on growth is significant or significantly increases is currently taken as a given within many policy circles.¹ This conviction has been strongly influenced by the work of Carmen Reinhart and Kenneth Rogoff, who were among the first to suggest a debt-to-GDP threshold of around 90% beyond which economic growth is seriously affected by the debt burden (Reinhart & Rogoff, 2009, 2010a,b, 2011; Reinhart, Reinhart & Rogoff, 2012), and of a large empirical literature that followed (e.g. Kumar & Woo, 2010; Cecchetti, Mohanty & Zampolli, 2011; Checherita-Westphal & Rother, 2012).²

This study is not about the analysis in Reinhart & Rogoff (2010b) – which is primarily descriptive – but about the substantial empirical literature these authors point to as a means of support for their findings. We build on this literature and approach the issue of nonlinearity in the debt-growth relationship with a number of alternative empirical strategies which enable us to distinguish a nonlinearity *across* countries from a *within-country* nonlinearity, a key distinction which has so far been entirely absent from the empirical literature. Identifying a within-country threshold effect would indeed inform policy makers of the presence of a country-specific tipping point, which may guide macroeconomic policies and fiscal adjustments. If debt-growth dynamics differ across countries, then the assumption of a common threshold across countries, as is the practice in the existing literature, leads to *one-size-fits-all* policies which are misleading at best and growth-retarding at worst.

¹In the United States, although many political battles impinge on the Congressional debate over the debt ceiling and the resulting government shutdown of October 2013, this state of affair at least in parts reflects a widespread belief that ‘debt is dangerous’ and that fiscal austerity represents the only way towards restoring sustainable growth for the world’s largest economy. In the United Kingdom Chancellor George Osborne displays a similar sentiment when telling his annual party conference that dealing with the repercussions of the financial crisis is not over “[u]ntil we’ve fixed the addiction to debt that got this country into this mess in the first place” (official Conservative Party Conference speech, Manchester, September 30, 2013).

²Note that while recently some details of their research and therefore the conclusions drawn have been seriously questioned (for a recent review see Panizza & Presbitero, 2013), the message of a common 90% debt threshold is still widely perceived in policy circles, among academics and in the media. This is the case notwithstanding that Reinhart & Rogoff (2010a) explicitly stated that they “do not pretend to argue that growth will be normal at 89% and subpar (about 1% lower) at 91% debt/GDP any more than a car crash is unlikely at 54mph and near certain at 56mph” and that causality is difficult to assess, as “for low-to-moderate levels of debt there may or may not be one; the issue is an empirical one, which merits study. For high levels of debt the evidence points to bi-directional causality.”

We analyse the debt-growth nexus within a standard neoclassical growth model for aggregate economy data, employing an empirical framework which allows for different long-run equilibrium relationships between debt and growth across countries, while simultaneously accounting for short-run effects and the impact of unobserved global shocks and local spillover effects.³ Using total public debt data from 105 developing, emerging and advanced economies over the 1972 to 2009 time horizon we find that long-run debt coefficients differ across countries and provide tentative evidence that countries with higher average debt-to-GDP ratios are more likely to see a negative effect on their long-run growth performance. We can however not find any evidence that a specific debt threshold *common to all countries* triggers a systematic parameter shift for individual countries as is widely suggested in the existing literature.

Four features of our empirical approach distinguish this study from the literature on debt and growth. First, we employ a flexible dynamic empirical framework which allows us to distinguish the long-run from the short-run relationship between debt and growth. We estimate the long-run and short-run parameters in a standard error correction model (ECM), test for the existence of a long-run equilibrium relationship (cointegration) and investigate concerns over endogeneity in the panel using recent panel time series methods.

Second, we put particular emphasis on modelling the debt-growth relationship as potentially differing across economies in an *a priori* unspecified way. Given theoretical arguments for structural (parameter) differences across countries, model and specification uncertainty, as well as serious shortcomings in the available data on public debt, we argue that flexibility in the cross-section dimensions of our panel econometric framework represents a crucial requirement and considerable strength when investigating complex entities such as national economies.⁴

Third, we identify the structural parameters of the relationship between debt

³Our analysis, like that in most of this literature, emphasises the long-run; we therefore employ a levels specification explaining income (per capita GDP), rather than the growth rate of income. However, for consistency with the literature, mainly based on the estimation of growth regressions, we use ‘growth’ instead of ‘income’ or ‘development’ in the following.

⁴There are a number of reasons to assume the equilibrium relationship between debt and growth differs across countries. First, in line with the ‘new growth’ literature (see Temple, 1999) production technology may differ across countries, and in the same vein the relationship between debt and growth.⁵ Second, vulnerability to public debt depends not only on debt levels, but also on debt composition (Dell’Erba, Hausmann & Panizza, 2013). Third, the capacity to tolerate high debts depends on a number of country-specific characteristics, related to past crises and the macro and institutional framework (Reinhart, Rogoff & Savastano, 2003; Kraay & Nehru, 2006; Manasse & Roubini, 2009).

and growth by accounting for the distorting impact of cross-section dependence in the form of unobserved global shocks and local spillover effects,⁶ both of which are likely to affect different economies in the sample to a different extent.

Fourth, we investigate the issue of nonlinearities in both the cross- and within-country dimensions, employing novel approaches and diagnostics from the time-series literature adapted for use in the panel. The presence of a nonlinearity across countries is studied by estimating a heterogeneous dynamic ECM and subsequently analysing the cross-country patterns of short-run and long-run debt coefficients. Our analysis of the within-country type of nonlinearity includes two approaches: (i) we investigate an asymmetric dynamic model, where we pick a range of threshold values, including the 90% debt-to-GDP ratio, as potential ‘tipping point’ for our debt-growth analysis; (ii) we present results from static regression models with squared and cubed debt terms.

The theoretical foundations for a negative and/or possibly non-linear relationship between debt and growth are rather tenuous (see Panizza & Presbitero, 2013, for a recent survey). While some models arrive at a negative long-run relationship (Elmendorf & Mankiw, 1999) which may be more pronounced if higher debt stocks lead to uncertainty or expectations of future financial repression (Cochrane, 2011), there are alternatives which suggest that in the presence of wage rigidities and unemployment this negative relationship disappears (Greiner, 2011). A nonlinearity or debt threshold can be motivated in developing countries by the presence of debt overhang (Krugman, 1988; Sachs, 1989) but it is difficult to extend this argument to advanced economies. Nonlinearities may also arise if there is a tipping point of fiscal sustainability (Ghosh et al., 2013; Greenlaw et al., 2013). However we are not aware of any theoretical models incorporating such debt tipping points in a growth framework.

Given the recent interest in this topic, our paper is naturally far from alone in studying the effect of the fiscal stance on growth in a cross-country regression framework (recent studies include Cordella, Ricci & Ruiz-Arranz, 2010; Kumar & Woo, 2010; Cecchetti, Mohanty & Zampolli, 2011; Checherita-Westphal & Rother, 2012; Panizza & Presbitero, 2012) – we provide a synthetic review of this literature in a Technical Appendix.⁷ Although individually quite rich in

⁶Examples for the former include the oil crises in the 1970s or the recent financial crisis, while knowledge spillovers from R&D investments in developed economies (Eberhardt, Helmers & Strauss, 2013) can be seen as illustrations for the latter.

⁷Panizza & Presbitero (2013) also provide a detailed discussion of the recent empirical literature on the relationship between debt and growth in advanced economies, focusing on causality

empirical results and proposed robustness checks, four features can broadly distinguish the analysis in these existing studies: (a) the data used (external or total debt) and country coverage (Euro area, OECD economies, developing countries, or emerging and developed countries); (b) the modelling of the hypothesised debt-growth nonlinearity/threshold (linear and squared debt terms in the regression, spline regression using preconceived thresholds, endogenous threshold regression); (c) the proposed time horizon of the results (short-run, long-run debt-growth relationship) depending on static or dynamic empirical specifications or, supposedly, the use of time-averaged or annual data; and (d) the identification strategy (standard IV/2SLS estimators, Arellano & Bond (1991)-type estimators). None of these studies however address more than one or arguably two of the four features we highlighted above (long-run versus short-run, cross-section dependence, cross-country heterogeneity, nonlinearity and asymmetry in integrated and cross-sectionally dependent macro panels), which we argue are of great importance for identification, analysis and interpretation.

The motivation of our analysis draws from earlier work by Reinhart, Rogoff & Savastano (2003), Kraay & Nehru (2006) and Reinhart & Rogoff (2010c) who, discussing debt sustainability, suggested that debt thresholds are likely to be country-specific, depending on history, policies and institutions. Along these lines, recent work emphasises the heterogeneity of the debt-growth nexus across countries, moving away from full sample analysis in homogeneous parameter regression models and investigating sub-samples along geographic, institutional or income lines (International Monetary Fund, 2012; Kourtellos, Stengos & Tan, 2014).

The remainder of this article organised as follows: Section 2 considers how the complexities of the economic theory and data realities should inform our empirical analysis. Section 3 describes our data and provides an overview of the econometric methods we apply. In Section 4 we present our empirical results and detailed analysis of heterogeneity and nonlinearity in the debt-growth relationship across and within countries. Section 5 concludes.

and non-linearities.

2 Linking Theory and Empirics

Two aspects of our approach are related to the modelling of empirical processes as common or different across countries. However, our interpretation of ‘common’ is somewhat different from what one may expect, in that we are concerned about common shocks (examples include the 1970s oil crises or the recent global financial crisis) and their distorting impact on identifying the debt-growth nexus in the data. Econometrically, we know that ignoring the impact of cross-section correlation, arising from global shocks or local spillover effects, yields seriously biased estimates for our parameters of interest (Phillips & Sul, 2003; Andrews, 2005), while nonlinearities may appear spuriously if heterogeneous relationships are erroneously modeled as common across countries (Haque, Pesaran & Sharma, 1999). In the following we provide some simple descriptive analysis highlighting the cross-sectional correlation of debt accumulation across countries, but also the cross-country heterogeneity in the relationship between debt and growth — data and sources are described in detail in Section 3.4 and the Data Appendix.

We begin with the issue of correlation across countries: Figure 1 provides a histogram for the years in which countries in our sample reach their debt-to-GDP ratio peak: although there is some heterogeneity as to the sample coverage for this period, it is notable that in over one-third of countries these peaks occurred in only three years, namely 1985, 1994 and 2009. Given that the data stretches over forty years, it is a remarkable indication of common effects across countries that the debt-to-GDP ratio peaks are clustered around a much smaller number of dates.

[Figure 1 about here]

A second illustration in Figure 2 links countries’ debt-to-GDP ratio peaks to the deviation of per capita GDP growth rate in the ‘peak years’ (defined *ad hoc* as running from two years prior to two years after the debt-to-GDP maximum) from that of the full time horizon (excluding the five peak years).⁸ We again highlight observations for the three years 1985, 1994 and 2009, as well as a small number of outliers. We can make a number of observations regarding this crude depiction of our empirical relationship of interest: first, there seems to be a negative correlation between the maximum debt level and relative growth perfor-

⁸For peaks at the start (end) of our sample we limit these averages to the peak year and the two years after (before).

mance between peak debt and other years. However, this negative relationship is not statistically significant (linear regression result reported in the figure footnote). Second, the figure highlights considerable heterogeneity across countries: for instance, among the countries for which debt-to-GDP peaked in 1994 (blue squares), one country experienced growth at around 2% *above* its growth rate in all other years, while another country experienced a ‘peak years’ average growth rate which was 4% lower.⁹ Third, and perhaps with view to the present debate in the literature most important, we note the dashed vertical line marking a debt-to-GDP ratio of 90%: a considerable number of countries had *better* growth performance in their peak debt years than at any other point since 1972, even at what some commentators refer to as ‘dangerous’ levels of debt.

[Figure 2 about here]

Figure 3 illustrates the potential for heterogeneity misspecification in the debt-growth relationship. In the first panel we plot a fractional polynomial regression line (as well as a 95% confidence interval) for per capita GDP against the debt-to-GDP ratio (both variables in logs) – the former is taken in deviation from the country-specific means (‘within’ transformation) to take account of different income levels across countries and thus focus on changes relative to the country mean.¹⁰ As can be seen there is clearly a nonlinear relationship between these two variables, in line with the standard arguments advanced by Reinhart and Rogoff as well as many others discussed above, with a ‘threshold’ of 4.5 log points (equivalent to 90% debt-to-GDP) a distinct possibility: higher debt burden is associated with lower per capita GDP. In a second plot in the same figure we add the actual observations for this regression in form of a scatter graph — the intention here is to cast some doubt over the ‘very obvious’ nonlinear relationship just discussed. In a third plot we provide country-specific fractional polynomial regression lines for all countries in our sample, while a fourth plot randomly selects thirty countries from the previous plot. In our view this highlights that

⁹Interestingly the green diamonds indicating 2009 show that all countries in which debt peaked in that year (all High-Income countries bar GRD and LCA) had *worse* growth performance in 2007-2009 than in all other years (average growth rates, respectively).

¹⁰The same pattern emerges when we use untransformed per capita GDP. In order to aid presentation in Figure 3 we exclude ‘extreme’ values (10% of observations) from this descriptive graph: the ‘full’ sample fractional polynomial regressions exclude all observations for debt-to-GDP ratio below 2 log points (<7.3% debt-to-GDP ratio), which amounts to 85 observations (primarily ARE, CHN and LUX). The scatter plots and country-specific fractional polynomial regressions further exclude all observations for which the within-transformed income change year-on-year exceeds 40%, which amounts to 281 observations (primarily BWA, IRL, KOR, THA and other fast-growing Middle-Income Countries).

the seeming nonlinearity assuming a pooled empirical model (black regression line and shaded confidence intervals) is far from obvious when we assume an empirical model which allows the relationship to differ across countries.

[Figure 3 about here]

Our descriptive analysis thus suggests that the raw data (adopting levels variables to elicit the long-run relationship) shows a clear non-linearity or threshold between the debt-to-GDP ratio and income at around 90% debt burden, provided we assume that all countries in the sample follow the same equilibrium path. However, relaxing this assumption seriously challenges this conclusion.

Of course this form of descriptive analysis is highly stylised, not to mention that there are other determinants of economic development and that such plots cannot provide any insights into any potentially causal relationship, be it from debt to growth or vice versa. Although our discussion is by no means conclusive, we feel that the illustrations provided above cast some doubt over the stringent *implicit* assumptions adopted in most of the existing literature: first, that we can carry out empirical analysis assuming that correlation across countries does not matter when running standard panel regression analysis which assumes cross-section independence. Second, the assumption that all countries, regardless of their level of economic development, their industrial structure or institutional environment, follow the same equilibrium relationship between debt and growth. Third, the notion that all countries are subject to the same debt threshold, beyond which growth is affected detrimentally, which is econometrically implemented by use of exogenous or endogenous debt thresholds or by adopting a polynomial specification for debt *within a pooled empirical model*, thus providing no insights whether the nonlinearity hides heterogeneity *across countries* or heterogeneity *within countries over time*.

Our empirical analysis of nonlinearity in the debt-growth nexus begins by considering a nonlinearity *across countries*. We adopt standard linear regression models, albeit of a fashion which accounts for both observed and unobserved heterogeneity. Identification of the long-run and short-run coefficients on debt is achieved by use of the Pesaran (2006) common correlated effects (CCE) estimator, which accounts for the presence of unobserved heterogeneity through a simple augmentation of the regression equation. Due to the dynamic setup and thus the presence of a lagged dependent variable it is necessary to adjust this augmentation following the suggestions in Chudik & Pesaran (2013). We

then analyse the relationship between the estimated long-run coefficients and country-specific averages of debt levels, of debt-to-GDP ratios as well as peak debt-to-GDP ratios.

Next, we consider nonlinearity in the debt-growth nexus *at the country-level*. We are not the first to consider such an empirical setup: Caner, Grennes & Koehler-Geib (2010) argue that provided a debt-threshold exists, this would arguably differ across countries given the heterogeneity in financial market development, openness, institutional development amongst other causes. Kourtellos, Stengos & Tan (2014) argue that if there exist heterogeneities in the debt-growth relationship (thresholds) then there may be other nonlinearities inherent in the empirical model employed to investigate this phenomenon. They find that while there does not exist a generic threshold or tipping point beyond which debt has a detrimental effect on growth, there does exist such a threshold determined by countries' level of democracy.

We adopt two approaches to investigating a nonlinearity at the country level: first we employ the nonlinear dynamic model by Shin, Yu & Greenwood-Nimmo (2013), where following selection of an exogenously given threshold (we focus on 52%, 75% and 90% in the debt-to-GDP ratio) we are able to investigate heterogeneous growth regimes (below and above the threshold) whilst accounting for cross-section dependence. As a robustness exercise, our second approach will employ the familiar microeconomic practice of including polynomial terms of the debt stock variable in a static regression model whilst accounting for cross-section dependence.

3 Empirical Strategy and Data

3.1 Empirical Specification: Linear Dynamic Model

The basic equation of interest in our analysis of the debt-development nexus is a static neoclassical production function augmented with a debt stock term:

$$y_{it} = \beta_i^K \text{cap}_{it} + \beta_i^D \text{debt}_{it} + u_{it} \quad u_{it} = \alpha_i + \lambda_i' f_t + \varepsilon_{it} \quad (1)$$

where y is aggregate GDP, cap is capital stock and debt is the total debt stock – all variables are in logarithms of *per capita* terms, imposing constant returns to

scale. Our specification of endogenous TFP in the form of common factors does however allow for externalities at the local and/or global level. These variables constitute the observable processes captured in our model, with their parameter coefficients β_i^j (for $j=K, D$) allowed to differ across countries¹¹ — this heterogeneity is a central feature of our empirical setup as motivated in the previous section. In addition we include country-specific Total Factor Productivity (TFP) levels (α_i) and a set of common factors f_t with country-specific ‘factor loadings’ λ_i to account for the evolution of unobservable TFP over time. The common factors can be a combination of ‘strong’ factors, representing global shocks such as the recent financial crisis, the 1970s oil crises or the emergence of China as a major economic power; and ‘weak’ factors, capturing local spillover effects following channels determined by shared culture heritage, geographic proximity, economic and social interaction (Chudik, Pesaran & Tosetti, 2011). We assume that these unobservable factors not only drive our measure of output, but also the other covariates in the above model:¹² this provides for a standard endogeneity problem whereby the β_i^j parameters are not identified unless some means to account for the unobservable factors in the error term u is found. At the same time this endogenises TFP and allows for externalities of production. We will return to the identification strategy in our discussion of the empirical implementation below. Suffice to highlight that standard instrumentation in a pooled empirical framework is invalid in the present setup as we cannot obtain instruments which are both informative and valid due to the omnipresence of unobserved factors, and/or the underlying equilibrium relationship differing across countries. Finally, we allow for the unobserved factors to be nonstationary, which has important implications for empirical analysis since all observable and unobservable processes in the model are now integrated and standard inference is invalid (Kao, 1999).

Given the importance of time series properties and dynamics in macro panel analysis, we employ an error correction model (ECM) representation of the above substantive equation of interest. This offers at least three advantages over a static model such as the above or restricted dynamic specifications such as those

¹¹We assume these parameter coefficients are fixed (Pesaran & Smith, 1995). Their moments and conditional distribution present a central interest in this study.

¹²A formal motivation for this setup from economic theory can be found in Mundlak, Butzer & Larson (2012) and Eberhardt & Teal (2013b). Note that covariates are not assumed to be only driven by common factors also contained in the estimation equation (f_t) but can have additional factors g_t exclusive to their evolution.

commonly investigated in the literature:¹³ (i) we can readily distinguish short-run from long-run behaviour; (ii) we can investigate the error correction term and deduce the speed of adjustment for the economy to the long-run equilibrium; and (iii) we can test for cointegration in the ECM by closer investigation of the statistical significance of the error correction term. The ECM representation of the above model is as follows:

$$\Delta y_{it} = \alpha_i + \rho_i (y_{i,t-1} - \beta_i^K \text{cap}_{i,t-1} - \beta_i^D \text{debt}_{i,t-1} - \lambda_i' f_{t-1}) + \gamma_i^K \Delta \text{cap}_{it} + \gamma_i^D \Delta \text{debt}_{it} + \gamma_i^{F'} \Delta f_t + \varepsilon_{it} \quad (2)$$

$$\Leftrightarrow \Delta y_{it} = \pi_{0i} + \pi_i^{EC} y_{i,t-1} + \pi_i^K \text{cap}_{i,t-1} + \pi_i^D \text{debt}_{i,t-1} + \pi_i^{F'} f_{t-1} + \pi_i^k \Delta \text{cap}_{it} + \pi_i^d \Delta \text{debt}_{it} + \pi_i^{f'} \Delta f_t + \varepsilon_{it} \quad (3)$$

where the β_i^j in equation (2) represent the long-run equilibrium relationship between GDP (y) and the measures for capital and debt in our model, while the γ_i^j represent the short-run relations. ρ_i indicates the speed of convergence of the economy to its long-run equilibrium. The term in round brackets represents the candidate cointegrating relationship we seek to identify in our panel time series approach. In equation (3) we have simply relaxed the ‘common factor restriction’ implicit in the nonlinear relationship between parameters in equation (2) and reparameterized the model to highlight that from the coefficients on the ‘levels’ terms (π_i^j for $j = K, D$) we can back out the long-run parameters

$$\beta_i^K = -\frac{\pi_i^K}{\pi_i^{EC}} \quad \beta_i^D = -\frac{\pi_i^D}{\pi_i^{EC}}$$

whereas from the coefficient on the terms in first difference (π_i^m for $m = k, d$, lowercase to distinguish from the long-run coefficients) we can read off the short-run parameters directly. π_i^{EC} indicates the speed at which the economy returns to the long-run equilibrium, with a half-life (in our data: in years) computable as

¹³It is somewhat of a tradition in the cross-country growth literature to follow Mankiw, Romer & Weil (1992) in the specification of ‘convergence’ regression models in the panel, regressing output growth Δy at time t (or averaged over a certain time horizon) on a lagged (or initial) level of output at time $t - 1$ as well inputs at time t (or averaged over a certain time horizon). This is despite the clear link Islam (1995) made between a cross-country growth model and a dynamic panel data model, which represents an Autoregressive Distributed Lag (ARDL) model of order (1,1). See Eberhardt & Teal (2011) for a survey of the cross-country growth empirics literature. Hendry (1995, 212) provides an insightful discussion on the encompassing nature of the ARDL model, for which “virtually every type of single equation model in empirical time-series econometrics is a special case”.

$(\log(0.5)/\log(1 + \pi_i^{EC}))$). Inference on this π_i^{EC} parameter will provide insights into the presence of a long-run equilibrium relationship: if $\pi_i^{EC} = \rho_i = 0$ we have no cointegration and the model reduces to a regression with variables in first differences (i.e. the term in brackets in equation (2) drops out). If $\pi_i^{EC} = \rho_i \neq 0$ we observe ‘error correction’, i.e. following a shock the economy returns to the long-run equilibrium path, and thus cointegration between the variables and processes in round brackets/levels. Note that we have included the unobservable common factors f in our long-run equation: this implies that we seek to investigate cointegration between output, capital, debt and TFP

In the spirit of Banerjee & Carrion-i-Silvestre (2011) we employ cross-section averages of all variables in the model to replace unobservables *as well as* omitted elements of the cointegration relationship.

$$\begin{aligned} \Delta y_{it} = & \pi_{0i} + \pi_i^{EC} y_{i,t-1} + \pi_i^K \text{cap}_{i,t-1} + \pi_i^D \text{debt}_{i,t-1} + \pi_i^k \Delta \text{cap}_{it} + \pi_i^d \Delta \text{debt}_{it} \\ & + \pi_{1i}^{CA} \overline{\Delta y}_t + \pi_{2i}^{CA} \overline{y}_{t-1} + \pi_{3i}^{CA} \overline{\text{cap}}_{t-1} + \pi_{4i}^{CA} \overline{\text{debt}}_{t-1} + \pi_{5i}^{CA} \overline{\Delta \text{cap}}_t + \pi_{6i}^{CA} \overline{\Delta \text{debt}}_t + \varepsilon_{it} \end{aligned} \quad (4)$$

Recent work by Chudik & Pesaran (2013) has highlighted that this approach is subject to small sample bias, in particular for moderate time series dimensions. Furthermore, these authors relax the assumption of strict exogeneity and thus allow for feedback between (in our application) debt, capital stock and output, which provides a more serious challenge to the original Pesaran (2006) approach:

$$\begin{pmatrix} \text{cap}_{it} \\ \text{debt}_{it} \end{pmatrix} = \alpha_{0i} + \alpha_{i1} y_{i,t-1} + \Gamma_i' f_t + \Upsilon_i' g_t + v_{it} \quad (5)$$

If $\alpha_{i1} = 0$ we maintain the assumption of strict exogeneity and we can proceed with the standard CCE augmentation, whereas if $\alpha_{i1} \neq 0$ this approach is only valid if and only if the dynamic common factor restrictions hold. Chudik & Pesaran (2013) provide the following empirical strategy employing cross-section averages in the presence of weakly exogenous regressors: *in addition* to the cross-section averages detailed in equation (4) they suggest (i) the inclusion of further lags of the cross-section averages, in our ECM setup

$$\sum_{\ell=1}^p \pi_{7i\ell}^{CA} \overline{\Delta y}_{t-p} + \sum_{\ell=1}^p \pi_{8i\ell}^{CA} \overline{\Delta \text{cap}}_{t-p} + \sum_{\ell=1}^p \pi_{9i\ell}^{CA} \overline{\Delta \text{debt}}_{t-p}$$

and (ii) the inclusion of cross-section averages of one or more further covariates

(other than cap and debt) which may help identify the unobserved common factors (in the spirit of Pesaran, Smith & Yamagata, 2013). In our ECM setup

$$\sum_{\ell=0}^p \pi_{g_{il}}^{CA} \overline{\Delta z}_{t-p}$$

for covariate z and similarly for further covariates. Chudik & Pesaran (2013) show that once augmented with a sufficient number of lagged cross-section averages ($p = T^{1/3}$ can be employed as a rule of thumb) the CCE mean group estimator performs well even in a dynamic model with weakly exogenous regressors.

Our empirical framework and implementation thus provide a great deal of flexibility to aid our attempts in capturing the long-run and short-run relationships between debt and growth across a set of diverse economies.¹⁴ An important feature of the empirical implementation adopted here is that all models are estimated by OLS: modelling features such as nonstationarity, cross-section correlation, heterogeneity in the equilibrium relationship across countries and non-linearity/asymmetry in the long-run and/or short-run relationship are captured by the empirical specification and the use of additional terms in the regression equation.

3.2 Empirical Specification: Weak Exogeneity Testing

In our factor model setup we have emphasised one type of endogeneity, whereby common factors drive both inputs and output, leading to identification issues unless the factors are accounted for. In the present context, a second form of endogeneity which implies reverse causality is deemed of particular importance for the interpretation of the empirical results: can we argue our empirical model derived from a neoclassical production function augmented with debt is correctly specified, or do we estimate a disguised demand equation for debt or invest-

¹⁴We do not claim that our empirical approach is ‘superior’ to the existing literature by pointing to asymptotic results in econometric theory or Monte Carlo simulation studies of known data generating processes. Instead, we highlight a set of assumptions which different empirical implementations make and provide diagnostic tests as to the validity of these assumptions. The comparison of results from different empirical estimators presented below does not constitute an exercise in data mining until the desired result emerges, but an attempt at testing the explicit and implicit assumptions made in each empirical model.

ment?¹⁵ For a basic common factor model setup this would imply

$$y_{it} = \beta_i x_{it} + u_{it} \quad u_{it} = \alpha_i + \lambda_i f_t + \varepsilon_{it} \quad (6)$$

$$x_{it} = \varrho_i f_t + \pi_i g_t + \psi_i \varepsilon_{it} + \phi_i + e_{it} \quad (7)$$

for a single covariate x and single factor f contained in both the y - and x -equations. Due to the presence of $\psi_i \varepsilon_{it}$ in the second equation we should be concerned over whether y ‘causes’ x or the reverse being the case or both. The standard approach in the literature has been to instrument for x using one or a set of variables z which satisfy the conditions of informativeness ($\mathbb{E}[zx] \neq 0$) and validity ($\mathbb{E}[z\varepsilon] = 0$). Having adopted a panel time series approach the issues of endogeneity and direction of causation can here take an alternative pathway: provided our variables are nonstationary and cointegrated, we can then apply a test for weak exogeneity. This test, described in detail in a Technical Appendix, can help us determine whether our empirical results can be interpreted as arising from a production function, rather than a misspecified input demand function. This identification strategy is not as clean as microeconomic alternatives, such as a controlled or natural experiment and instrumental variable estimation.¹⁶ We argue that neither of these strategies are suitable in a macroeconomic context: experiments may provide insights into a unique episode or a single country experience, but arguably lack the external validity by necessity required in answering our research question. We already argued above that instrument validity is difficult to justify in macro panel analyses of a globalised world. With the empirical questions addressed in this study in mind we believe our empirical strategy is the best we can do.

¹⁵The discussion in this section is based on Eberhardt & Teal (2013a), which follows the methodology of Canning & Pedroni (2008).

¹⁶We implicitly assume that debt accumulation is orthogonal to any *expectations* of future growth accelerations or slowdowns. It however bears reminding that our common factor framework allows us to account for global and local economic circumstances and business cycles. These global and local economic conditions and trends arguably play a dominant role in forming expectations about an individual country’s future growth trajectory and the orthogonality condition thus applies much more narrowly to any expectations formed on the basis of exogenous, *country-specific* shocks.

3.3 Empirical Specification: Asymmetric Dynamic Model

We follow the discussion in Shin, Yu & Greenwood-Nimmo (2013) and define the asymmetric long-run regression model

$$y_{it} = \alpha_i + \beta_i^K \text{cap}_{it} + \beta_i^{D+} \text{debt}_{it}^+ + \beta_i^{D-} \text{debt}_{it}^- + \lambda_i' f_t + \varepsilon_{it} \quad (8)$$

where we again assume observable and unobservable processes are nonstationary and debt stock has been decomposed into $\text{debt}_{it} = \text{debt}_{i0} + \text{debt}_{it}^+ + \text{debt}_{it}^-$. The latter two terms are partial sums of values above and below a specific threshold, debt_{i0} has been subsumed into the constant term. For instance, if we assume a threshold of zero then they define positive and negative changes in debt accumulation. For each country i let

$$\begin{aligned} \text{debt}_{it}^+ &= \sum_{j=1}^t \Delta \text{debt}_{ij}^+ = \sum_{j=1}^t \max(\Delta \text{debt}_{ij}, 0) \\ \text{debt}_{it}^- &= \sum_{j=1}^t \Delta \text{debt}_{ij}^- = \sum_{j=1}^t \min(\Delta \text{debt}_{ij}, 0) \end{aligned} \quad (9)$$

This setup would suit the analysis of an asymmetric response to debt accumulation and relief, whereby the hypothesised substantial growth benefits of debt relief could be shown to be questionable given a the differential relationship between debt accumulation and growth on the one hand and debt reduction and growth on the other. In our present study we instead create partial sums for debt stock below and above a number of (exogenously determined) debt-to-GDP ratio thresholds, namely 52% (sample median), 75% and the ‘canonical’ 90%. Thus the partial sums are constructed from the per capita *debt stock variable*, while the assignment to one or the other regime is determined by the *debt-to-GDP ratio* – we adopt this practice in order to be able to compare our results with those in the literature adopting the debt-to-GDP ratio as the primary variable of interest.

The ECM version of our asymmetric dynamic model is thus

$$\begin{aligned} \Delta y_{it} &= \pi_{0i} + \pi_i^{EC} y_{i,t-1} + \pi_i^K \text{cap}_{i,t-1} + \pi_i^{D+} \text{debt}_{i,t-1}^+ + \pi_i^{D-} \text{debt}_{i,t-1}^- \\ &+ \pi_i^{F'} f_{t-1} + \pi_i^k \Delta \text{cap}_{it} + \pi_i^{d+} \Delta \text{debt}_{it}^+ + \pi_i^{d-} \Delta \text{debt}_{it}^- + \pi_i^f \Delta f_t + \varepsilon_{it} \end{aligned} \quad (10)$$

The dynamic asymmetry can be included in the long-run relationship (lagged lev-

els terms), in the short-run behaviour (first difference terms) or both. As before we allow for cross-country heterogeneity in *all* long-run and short-run parameters and account for the presence of unobserved time-varying heterogeneity by augmenting the country regressions with cross-section averages of the dependent and independent variables. While in the original Shin, Yu & Greenwood-Nimmo (2013) time series approach the parameter estimates are identified by augmentation of the empirical equation with additional lagged differences, our panel approach relies on the common factor framework as developed in Section 2 above for identification. The same issues as highlighted in Section 3.1 apply and we shall augment the estimation equation with further lags of the cross-section averages (Chudik & Pesaran, 2013). Note that the implementation raises a number of problems in the case where the debt threshold is relatively high: if only a very small number/share of observations for a specific country are above the threshold, then the estimated coefficient may be very imprecise. In order to guard against this we present results of the estimated long-run debt parameters in the low and high debt regimes only for those countries where at least 20% of all time series observations are in one regime. For the 90% debt/GDP threshold this amounts to a total of 30 countries, 45 countries in case of the 75% threshold and 55 countries for the 52% threshold.¹⁷

3.4 Data

Our main variables are GDP, capital stock, constructed from gross fixed capital formation using the standard perpetual inventory method and assuming a common and constant 5% depreciation rate, and total public debt stock (all in logarithms of real US\$). Data are taken from the World Bank World Development Indicators (WDI) database and, in the case of the debt stock, from an update to Panizza (2008). The debt variable is total (external and domestic) general government debt in nominal terms (face value), in the raw data expressed as a share of GDP. We express all variables in per capita terms, including the debt stock.

¹⁷In principle the implementation of the asymmetric dynamic model by Shin, Yu & Greenwood-Nimmo (2013) is subject to the same criticism we level against the polynomial models in the following sub-section. However, given the limited number of countries which have substantial observations for debt-to-GDP in both regimes (above, below threshold), it is infeasible to adopt the subsampling strategy (working with random sets of \sqrt{N} countries) that we propose below to obtain inferential statistics.

For a small number of empirical results we further make use of cross-section averages for data on trade openness and financial development. Trade openness (imports plus exports as a share of GDP) is taken from the NYU Global Development Network Growth Database – in turn based on the World Bank’s WDI and Global Development Finance databases – while financial development (ratio of bank credit to bank deposits) is taken from Thorsten Beck and Asli Demirguc-Kunt’s Financial Structure Database, updated in 2010.¹⁸ A Data Appendix provides more details on the construction of our variables and descriptive statistics. Detailed information about the sample make-up is confined to a Technical Appendix.

4 Empirical Results

4.1 Initial Analysis

We carried out panel unit root tests following Pesaran (2007) and investigated the cross-section correlation properties of the raw data including formal CD tests following Pesaran (2004). Results are provided in a technical appendix and indicate that the levels variable series are integrated of order 1 and subject to considerable cross-section dependence.

[Figure 4 about here]

We conduct a similar descriptive analysis to that pursued in Reinhart & Rogoff (2010b) for our sample of countries, with results presented in Figure 4: within each income group (High, Upper- and Lower-Middle, Low Income) all observations are divided into four bins based on the debt-to-GDP ratio. The means (dark grey bars) and medians (light grey) for different income groups by level of indebtedness may be taken as evidence for a *differential growth performance beyond a 90% debt-to-GDP threshold*, at least for the high-, lower middle- and low-income samples. We now provide empirical evidence that this descriptive result is misleading.

¹⁸If these two variables are included in the empirical analysis the resulting estimates are for the data ending in 2008.

4.2 Results: Linear Dynamic Models

Table 1 presents results derived from an ECM specification, with results for a standard two-way fixed effects and pooled CCE in columns [1] and [2] assuming parameter homogeneity across countries and all other models in columns [3]-[10] allowing for differential relationships. The models in columns [4] and [5] represent the standard CCE estimator in the Mean Group version, while models in columns [6], [7] and [9] add further lags of the cross-section averages as suggested in Chudik & Pesaran (2013). Models in columns [8]-[10] experiment with the cross-section averages and lags of additional covariates outside the model: we adopt proxies for trade openness ('open') and financial development ('findev'), both in logs. These variables *only* enter the empirical model in form of their cross-section averages. The aim here is to help identify the unobserved common factors f_t , which represent global shocks and local spillover effects,¹⁹ so that adopting variables which are directly linked to globalisation was deemed a suitable choice here.

[Table 1 about here]

In each model we focus on the long-run estimates as well as the coefficient on the lagged level of GDP to investigate error correction – full ECM results are available on request. LRA refers to the 'long-run average' coefficient, which is calculated directly from the pooled model ECM results in [1] and [2] and the weighted averages – we follow standard practice in this literature and employ robust regression (see Hamilton, 1992) to weigh down outliers in the computation of the averages – of the heterogeneous model ECM results in [3]-[10]. LRA standard errors are computed via the Delta method. ALR refers to the 'average long-run' coefficient in the heterogeneous models, whereby the long-run coefficients are computed from the ECM results in each country and then averaged across the panel.²⁰ In the ALR case standard errors are constructed following Pesaran & Smith (1995). For all heterogeneous models which address concerns over cross-section dependence there is evidence of error correction – the lagged GDP pc levels variable is highly statistically significant – and the average long-run coefficients appear statistically significant and positive throughout, whereas short-run coefficients are insignificant. The latter does not imply the absence of any significant effects, but rather highlights the heterogeneity across coun-

¹⁹The *country-series* for 'open' and 'findev' are not additional covariates in our regression model.

²⁰For more details on these concepts see Smith (2001).

tries with dynamics on average cancelling out. Coefficients on lagged per capita GDP levels imply reasonable estimates for the speed of convergence.²¹ Diagnostic tests highlight that the use of cross-section averages considerably reduces residual cross-section dependence – the CD statistic drops from 18 in the MG to between 2 and 3 in the CMG models. Based on work by Bailey, Kapetanios & Pesaran (2012) it is suggested that the implicit null hypothesis of this test is weak (rather than strong) cross-section dependence (Pesaran, 2013) – recall that dependence of the weak type only affects inference, whereas strong dependence can lead to an identification problem.²²

Once we move from a pooled to a heterogenous parameter specification, statistically significant positive *average* long-run coefficients as we find in our sample only provide insights regarding the *central tendency* of the panel. This result may indicate that, on average, the countries in our sample are on the ‘right’ side of an hypothetical Debt Laffer curve. This hardly surprising as the median debt-to-GDP ratio is around 50% (Table A1), a value well below the ‘tipping points’ identified by the literature on developing and advanced economies (see Table TA1). In Figure 5 we therefore provide a number of plots indicating the cross-section dispersion of the long-run debt coefficients, primarily focusing on the estimates in the dynamic CCE model with one additional lag (column [6] of Table 1) given its favourable diagnostic results. With the exception of panel (b) all plots capture the country-specific average debt-to-GDP ratio over the entire sample period (in logs) on the *x*-axis and estimated debt-coefficients on the *y*-axis (all long-run except for panel (f), which plots short-run coefficients). Panel (a) suggests that there is a nonlinear relationship between the debt-to-GDP ratio and the long-run impact of debt, which around 90% debt-to-GDP turns negative. Panel (c) makes the same point grouping countries into quintiles based on average debt/GDP ratio and providing distributional plots for each of them (group #5 represents debt burden over 90% of GDP).

Panel (b) however cautions against this conclusion: instead of average debt-to-GDP ratio we plot here the debt-to-GDP ratio *peak* for each country. It is notable that many countries still have positive coefficients despite peak debt-to-GDP ratios in excess of 90%. Panel (d) splits the data into the 25% richest

²¹The half-life indicates “the length of time after a shock before the deviation in output shrinks to half of its impact value” (Chari, Kehoe & McGrattan, 2000, 1161).

²²For models [8] and [9] we also investigated the use of ‘findev’ instead of ‘open’ and found qualitatively identical results. This includes the cross-country plots in Panels (e) and (f) of Figure 5.

countries and the rest – the nonlinearity between debt burden and the long-run debt coefficient across countries seems to primarily be driven by the poorer countries in the sample. Panels (e) and (f) provide fitted fractional polynomial regression lines for the CMG models in Table 1 for which the residual CD test is below 3: [4]-[7] and [10]. With regard to long-run results in panel (e), the average relationship emerging seems to be fairly robust to the choice of empirical specification. There is no evidence for any systematic heterogeneity in the short-run coefficients presented in panel (f).

[Figure 5 about here]

We thus find some tentative evidence for a nonlinearity in the long-run relationship between debt and growth *across countries*. We can be reasonably certain that these empirical models represent cointegrating relationships between debt and income, but this does not rule out the possibility of feedback from income to debt, which would question the validity of our empirical results. As a next step we therefore turn to weak exogeneity testings for all of our heterogeneous parameter models.

[Table 2 about here]

In Table 2 we present the results for the MG and various CMG models – models refer to the column numbering in Table 1. For each estimator we provide weak exogeneity tests using specifications with one or two lags, in each case providing three sets of results: for an output equation, a capital stock equation and a debt stock equation. If our suggestion that the empirical models analysed represent augmented production functions, rather than investment demand or debt demand equations, thus (informally) allowing us to argue for a causal relation from capital and debt stock to output and not vice-versa, we would expect a pattern whereby the various test statistics for the output equation reject the null of no causal relation from ‘inputs’ to output, whereas those in the two ‘input’ equations cannot reject their respective nulls. Taking in the results as a whole, there appears to be fairly strong evidence for the setup described: p -values for the statistic constructed from averaged t -statistics are typically below 10 percent in the output and close to unity in the input equations; the t -statistics on the averaged λ_i coefficients are typically very large in the former and typically below 1.96 in the latter.

The purpose of the analysis up to this point was to investigate the possibility of a nonlinear relationship between the debt burden and the long-run debt coefficient

in the cross-country dimension. A number of empirical models including dynamic CCE which allows for cross-section dependence in a dynamic model were evaluated and while the empirical results are somewhat fragile in a moderate- T sample, one might conclude that on balance there is some evidence for heterogeneity in the long-run coefficients across countries. We now turn to empirical models which allow for heterogeneous long-run relations across countries while at the same time allowing for thresholds in the relationship within countries over time, which represents a direct test of the consensus of a common threshold effect as propagated in the existing empirical literature.

4.3 Results: Asymmetric Dynamic Models

In Figure 6 we present results from the asymmetric (heterogeneous) dynamic regressions where we account for unobserved common factors by inclusion of cross-section averages of all covariates as well as one further lag of the cross-section averages. The three plots correspond to subsamples for an adopted threshold of 52% (top), 75% (middle) and 90% (bottom) for the debt-to-GDP ratio – in each case we only include countries which have at least 20% of their observations in one of the two regimes (below/above threshold), amounting to 55, 45 and 30 for the three thresholds, respectively. Empirical results on which these graphs are based can be found in a Technical Appendix (Table TA5, model [5] with one additional lag and asymmetry in the long- and short-run specification).

The x -axis in each plot represents the average debt burden over the entire time horizon, expressed as the average debt-to-GDP ratio (in logs) in the left column and, like in our regressions, as the total debt stock per worker (in logs) in the right column – in both sets of plots the left tip of each arrow represents the average value for the ‘low debt’ regime where debt is below 52%, 75% or 90% of GDP, while the right arrow tip marks the average value for the ‘high debt’ regime above these thresholds. The y -axis in each plot captures the estimated long-run debt coefficient which by construction is allowed to differ across regimes (and countries). Under the working hypothesis that a shift to the ‘high debt’ regime would have a negative, step-change type impact on long-run growth, we would expect most arrows to indicate a negative relationship. As can be seen, this hypothesis is not borne out by the empirical results: there is no evidence for any systematic change in the relationship between debt and growth when countries shift from a ‘low’ to ‘high’ debt regime, with only around one in two countries experiencing

an increase in the debt coefficient.²³ Average coefficient changes in each of the three cases are statistically insignificant (standard or robust means).

[Figure 6 about here]

Thus our test of within-country threshold effects in the debt-growth relationship suggests that the consensus in the empirical literature of a *common* debt threshold does not hold up for the cutoffs tested if we allow for observed and unobserved heterogeneity across countries. In the following section we investigate a popular alternative empirical representation for debt thresholds by adopting various polynomial specifications.

4.4 Robustness: Nonlinear Static Model

The previous sections we have modelled the within-country non-linearity as an asymmetry. As a robustness exercise, we estimated different static models with polynomial approximations to unknown nonlinear functions:

$$y_{it} = \alpha_i + \beta_i^K \text{inv}_{it} + \beta_i^{DL} \text{debt}_{it} + \beta_i^{DS} \text{debt}_{it}^2 + \beta_i^{DC} \text{debt}_{it}^3 + \lambda_i' \mathbf{f}_t + \varepsilon_{it} \quad (11)$$

We begin estimating a static linear model, then we introduce the squared debt term, and finally the cubed debt term. We only focus on static models given the serious difficulties arising in reconciling nonlinearities with cross-section dependence, parameter heterogeneity and a dynamic specification within a panel of moderate time series dimension.

The main concern for our analysis here is the most appropriate specification with regards to the time-series properties of the data: reliable inference on a relationship between variable series which are nonstationary involves establishing that these variables are cointegrated, and within both time series and panel time series econometrics a number of alternative approaches are available to test for this property. Crucially, however, cointegration defines a *linear* combination of variables integrated of order one (in our case) which is stationary (i.e. integrated of order zero). Difficulties for the analysis of potentially nonlinear relationships such as that between debt and growth arise given that the order of integration of the square or cube of an integrated variable is not defined within the linear integration and cointegration framework.

²³This simple count does not take statistical significance into account.

We apply novel methods on the order of summability and the concept of co-summability from the time series econometric literature (Berenguer-Rico & Gonzalo, 2013a,b) to provide pre-estimation testing as to the validity of our empirical equation incorporating country-specific nonlinearities. To the best of our knowledge our study is the first to adopt these methods in the panel context, further addressing the concerns over cross-section dependence. The details for these tests of long-run co-movement as well as the results can be found in a Technical Appendix. Estimates from the nonlinear models are presented in Table 3 – we indicate that our co-summability analysis suggested that only one of our specifications, model [7], offers strong evidence for a long-run equilibrium relationship.²⁴

Consistent with our previous results, our estimates highlight the heterogeneity in the country-specific results, with no linear or nonlinear relationship for the debt-growth nexus emerging as clearly dominant: for instance, in the linear models we find similar numbers of positive and negative slope coefficients on debt once we account for the distorting impact of cross-section correlation. In the models with linear and squared debt there is more evidence for concave relations – in line with the Reinhart & Rogoff (2010b) debt threshold story – but it would be difficult to claim that that this result is *uniform across all countries*. In the models with three debt terms the averaged coefficients hide a great deal of heterogeneity across countries. Thus on the whole we cannot provide any support for the notion that countries possess similar or even identical nonlinearities in the debt-growth relationship *over time* once we relax the assumption of common parameters *across countries*.

5 Concluding Remarks

This article empirically investigates the relationship between public debt and long-run growth and provides important insights for the current debate on threshold effects in the debt-growth nexus sparked by the work of Carmen Reinhart and Kenneth Rogoff (Reinhart & Rogoff, 2009, 2010a,b, 2011). Our paper makes three contributions to this empirical literature: first, we investigated the long-

²⁴All of the models presented are heterogeneous parameter specifications, but we also investigated various pooled model specifications (Pooled OLS, Fixed Effects, CCE Pooled) and found strong evidence of nonstationary residuals in these models, thus highlighting the potentially spurious nature of estimates from pooled empirical models (results available on request).

run relationship by means of a dynamic empirical model and adopted time series arguments to establish the presence of a long-run equilibrium, taking into account possible endogeneity issues. Since estimation results are likely to be spurious and seriously biased if these well-known data properties are not recognised and addressed in the empirical analysis our approach signals a significant departure from the standard empirical modelling in this literature and arguably provides more reliable estimates. Second, we adopted empirical specifications which allowed for heterogeneity in the long-run relationship across countries, thus reflecting a host of theoretical and empirical arguments. This heterogeneity in the specification extends to the relevant unobservable determinants of growth and debt burden, which we have addressed by means of a flexible common factor model framework. Ours is the first panel study on debt and growth to address parameter heterogeneity and cross-section dependence, thus allowing for a closer match between economic theory and data restrictions on the one hand and empirical modelling on the other. Third, we used a number of empirical estimators and testing procedures to shed light on the potential nonlinearity in the debt-growth relationship, focusing on both the possibility of a debt-growth nonlinearity *across* and *within* countries, a distinction previously entirely absent from the empirical literature. It bears emphasising that no empirical study modelling the debt-growth relationship in a pooled panel model can claim to be able to distinguish these two types of nonlinearity.

Our empirical analysis provided some evidence for systematic differences in the debt-growth relationship *across countries*, but no evidence for systematic *within-country* nonlinearities in the debt-growth relationship for all countries in our sample. With regard to the first result we observed that long-run debt coefficients appeared to be lower in countries with higher average debt burden, although the average long-run debt coefficient across countries was positive. Regarding the second result, empirical tests seemed to support a linear specification rather than the polynomial specifications popular in the empirical literature. When we employed piecewise linear specifications adopting various pre-specified thresholds, the change in the debt coefficient at the threshold was just as likely to be positive as negative. These findings imply that whatever the shape and form of the debt-growth relationship, it differs across countries, so that appropriate policies for *one* country may be seriously misguided in *another*. The commonly found 90% debt threshold is likely to be the outcome of empirical misspecification – a pooled instead of heterogeneous model – and subsequently a misinterpretation

of the results, whereby it is assumed that pooled model estimates – obtained from polynomial or piecewise linear specifications for debt – imply that a common nonlinearity detected applies within all countries over time.

References

- Afonso, Antonio & Joao Tovar Jalles. 2013. “Growth and productivity: The role of government debt.” *International Review of Economics & Finance* 25:384–407.
- Andrews, Donald W. K. 2005. “Cross-Section Regression with Common Shocks.” *Econometrica* 73(5):1551–1585.
- Arellano, Manuel & Stephen Bond. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *Review of Economic Studies* 58(2):277–97.
- Bailey, Natalia, George Kapetanios & M. Hashem Pesaran. 2012. “Exponent of Cross-sectional Dependence: Estimation and Inference.” Cambridge University, mimeo, January 2012.
- Banerjee, Anindya & Josep Carrion-i-Silvestre. 2011. Testing for Panel Cointegration using Common Correlated Effects Estimators. Paper presented at the 14th Applied Economics Meeting, Huelva, June.
- Barro, Robert J. & Jong Wha Lee. 2013. “A new data set of educational attainment in the world, 1950-2010.” *Journal of Development Economics* 104:184 – 198.
- Baum, Anja, Cristina Checherita-Westphal & Philipp Rother. 2013. “Debt and growth: New evidence for the euro area.” *Journal of International Money and Finance* 32:809–821.
- Berenguer-Rico, Vanessa & Jesus Gonzalo. 2013a. “Co-summability: from linear to non-linear co-integration.” Universidad Carlos III de Madrid Working Paper 13-12.
- Berenguer-Rico, Vanessa & Jesus Gonzalo. 2013b. “Summability of stochastic processes: a generalization of integration and co-integration valid for non-linear processes.” *Journal of Econometrics* forthcoming.
- Blundell, Richard & Stephen Bond. 1998. “Initial conditions and moment restrictions in dynamic panel data models.” *Journal of Econometrics* 87(1):115–143.

- Caner, Mehmet, Thomas Grennes & Fritzi Koehler-Geib. 2010. Finding the tipping point – when sovereign debt turns bad. Policy Research Working Paper Series 5391 The World Bank.
- Canning, David & Peter Pedroni. 2008. “Infrastructure, Long-Run Economic Growth And Causality Tests For Cointegrated Panels.” *Manchester School* 76(5):504–527.
- Cecchetti, Stephen, Madhusudan Mohanty & Fabrizio Zampolli. 2011. Achieving Growth Amid Fiscal Imbalances: The Real Effects of Debt. In *Economic Symposium Conference Proceedings*. Federal Reserve Bank of Kansas City pp. 145–196.
- Chari, Varadarajan V, Patrick J Kehoe & Ellen R McGrattan. 2000. “Sticky price models of the business cycle: can the contract multiplier solve the persistence problem?” *Econometrica* 68(5):1151–1179.
- Checherita-Westphal, Cristina & Philipp Rother. 2012. “The Impact of High Government Debt on Economic Growth and its Channels: An Empirical Investigation for the Euro Area.” *European Economic Review* 56(7):1392–1405.
- Chudik, Alexander & M. Hashem Pesaran. 2013. “Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models with Weakly Exogenous Regressors.” Federal Reserve Bank of Dallas Globalization and Monetary Policy Institute Working Paper No. 146.
- Chudik, Alexander, M. Hashem Pesaran & Elisa Tosetti. 2011. “Weak and Strong Cross Section Dependence and Estimation of Large Panels.” *Econometrics Journal* 14(1):C45–C90.
- Clements, Benedict J., Rina Bhattacharya & Toan Quoc Nguyen. 2003. External Debt, Public Investment, and Growth in Low-Income Countries. IMF Working Papers 03/249 International Monetary Fund.
- Cochrane, John H. 2011. “Understanding policy in the great recession: Some unpleasant fiscal arithmetic.” *European Economic Review* 55(1):2–30.
- Cordella, Tito, Luca Antonio Ricci & Marta Ruiz-Arranz. 2010. “Debt Overhang or Debt Irrelevance?” *IMF Staff Papers* 57(1):1–24.
- Dell’Erba, Salvatore, Ricardo Hausmann & Ugo Panizza. 2013. Debt Levels, Debt Composition, and Sovereign Spreads in Emerging and Advanced Economies. IHEID Working Papers 12-2013 The Graduate Institute of International Studies.

- Eberhardt, Markus. 2013. "Nonlinearities in the Relationship between Debt and Growth: Evidence from Co-Summability Testing." mimeo, October 2013.
- Eberhardt, Markus, Christian Helmers & Hubert Strauss. 2013. "Do spillovers matter when estimating private returns to R&D?" *The Review of Economics and Statistics* 95(2):436–448.
- Eberhardt, Markus & Francis Teal. 2011. "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics." *Journal of Economic Surveys* 25(1):109–155.
- Eberhardt, Markus & Francis Teal. 2013a. "No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis." *Oxford Bulletin of Economics and Statistics* 75(6):914–939.
- Eberhardt, Markus & Francis Teal. 2013b. "Structural Change and Cross-Country Growth Empirics." *World Bank Economic Review* 27(2):229–271.
- Egert, Balazs. 2013. The 90% Public Debt Threshold: The Rise and Fall of a Stylised Fact. CESifo Working Paper 4242 CESifo.
- Elmendorf, Douglas W. & Gregory N. Mankiw. 1999. Government debt. In *Handbook of Macroeconomics*, ed. J. B. Taylor & M. Woodford. Vol. 1 of *Handbook of Macroeconomics* Elsevier chapter 25, pp. 1615–1669.
- Engle, Robert F. & Clive W. J. Granger. 1987. "Cointegration and Error correction: representations, estimation and testing." *Econometrica* 55(2):252–276.
- Ghosh, Atish R., Jun I. Kim, Enrique G. Mendoza, Jonathan D. Ostry & Mahvash S. Qureshi. 2013. "Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies." *The Economic Journal* 123(566):F4–F30.
- Granger, Clive W. J. & Paul Newbold. 1974. "Spurious regressions in econometrics." *Journal of Econometrics* 2(2):111–120.
- Greenlaw, David, James Hamilton, Peter Hooper & Frederic S. Mishkin. 2013. *Crunch Time: Fiscal Crises and the Role of Monetary Policy*. NBER Working Papers 19297 National Bureau of Economic Research.
- Greiner, Alfred. 2011. "Economic Growth, Public Debt and Welfare: Comparing Three Budgetary Rules." *German Economic Review* 12(2):205–222.
- Hamilton, Lawrence C. 1992. "How Robust is Robust Regression?" *Stata Technical Bulletin* 1(2).

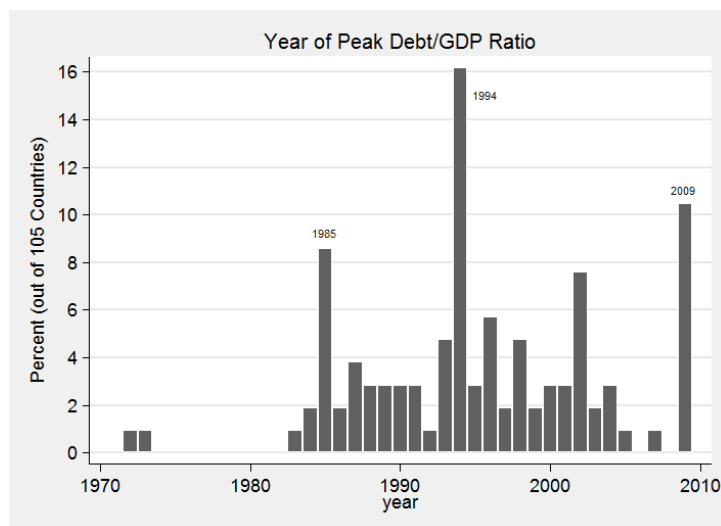
- Haque, Nadeem U., M. Hashem Pesaran & Sunil Sharma. 1999. "Neglected Heterogeneity and Dynamics in Cross-Country Savings Regressions." mimeo.
- Hendry, David. 1995. *Dynamic Econometrics*. Advanced Texts in Econometrics Oxford University Press.
- Imbs, Jean & Romain Ranciere. 2005. The overhang hangover. Policy Research Working Paper Series 3673 The World Bank.
- International Monetary Fund. 2012. *World Economic Outlook – Coping with High Debt and Sluggish Growth*. International Monetary Fund.
- Islam, Nazrul. 1995. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics* 110(4):1127–70.
- Kao, Chihwa. 1999. "Spurious regression and residual-based tests for cointegration in panel data." *Journal of Econometrics* 65(1):9–15.
- Kourtellos, Andros, Thanasis Stengos & Chih Ming Tan. 2014. "The Effect of Public Debt on Growth in Multiple Regimes." *Journal of Macroeconomics* forthcoming.
- Kraay, Aart & Vikram Nehru. 2006. "When Is External Debt Sustainable?" *World Bank Economic Review* 20(3):341–365.
- Krugman, Paul. 1988. "Financing vs. forgiving a debt overhang." *Journal of Development Economics* 29(3):253–268.
- Kumar, Manmohan S. & Jaejoon Woo. 2010. Public Debt and Growth. IMF Working Papers 10/174 International Monetary Fund.
- Manasse, Paolo & Nouriel Roubini. 2009. "Rules of thumb for sovereign debt crises." *Journal of International Economics* 78(2):192–205.
- Mankiw, N Gregory, David Romer & David N Weil. 1992. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* 107(2):407–37.
- Minea, Alexandru & Antoine Parent. 2012. Is high public debt always harmful to economic growth? Reinhart and Rogoff and some complex nonlinearities. Working Papers 8 Association Francaise de Cliometrie.
- Mundlak, Yair, Rita Butzer & Donald F. Larson. 2012. "Heterogeneous technology and panel data: The case of the agricultural production function." *Journal of Development Economics* 99(1):139–149.

- Panizza, Ugo. 2008. Domestic And External Public Debt In Developing Countries. UNCTAD Discussion Papers 188 United Nations Conference on Trade and Development.
- Panizza, Ugo & Andrea Filippo Presbitero. 2012. Public Debt and Economic Growth: Is There a Causal Effect? Mo.Fi.R. Working Papers 65.
- Panizza, Ugo & Andrea Filippo Presbitero. 2013. "Public Debt and Economic Growth in Advanced Economies: A Survey." *Swiss Journal of Economics and Statistics* 149(2):175–204.
- Pattillo, Catherine A., Helene Poirson & Luca Antonio Ricci. 2004. What Are the Channels Through Which External Debt Affects Growth? IMF Working Papers 04/15 International Monetary Fund.
- Pattillo, Catherine A., Helene Poirson & Luca Antonio Ricci. 2011. "External Debt and Growth." *Review of Economics and Institutions* 2(3). Retrieved from <http://www.rei.unipg.it/rei/article/view/45>.
- Pesaran, M. Hashem. 2004. "General Diagnostic Tests for Cross Section Dependence in Panels." University of Cambridge, mimeo.
- Pesaran, M. Hashem. 2006. "Estimation and inference in large heterogeneous panels with a multifactor error structure." *Econometrica* 74(4):967–1012.
- Pesaran, M. Hashem. 2007. "A simple panel unit root test in the presence of cross-section dependence." *Journal of Applied Econometrics* 22(2):265–312.
- Pesaran, M. Hashem. 2013. "Testing Weak Cross-Sectional Dependence in Large Panels." *Econometrics Reviews* (forthcoming).
- Pesaran, M. Hashem & Ron P. Smith. 1995. "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of Econometrics* 68(1):79–113.
- Pesaran, M. Hashem, Vanessa Smith & Takashi Yamagata. 2013. "Panel Unit Root Tests in the Presence of a Multifactor Error Structure." *Journal of Econometrics* 175(2):94–115.
- Phillips, Peter C. B. & Donggyu Sul. 2003. "Dynamic panel estimation and homogeneity testing under cross section dependence." *Econometrics Journal* 6(1):217–259.
- Presbitero, Andrea F. 2008. "The Debt-Growth Nexus in Poor Countries: A Re-assessment." *Economics: The Open-Access, Open-Assessment E-Journal* 2(30).

- Presbitero, Andrea F. 2012. "Total Public Debt and Growth in Developing Countries." *European Journal of Development Research* 24(4):606–626.
- Reinhart, Carmen M. & Kenneth S. Rogoff. 2009. *This Time is Different - Eight Centuries of Financial Folly*. Princeton, New Jersey: Princeton University Press.
- Reinhart, Carmen M. & Kenneth S. Rogoff. 2010a. Debt and Growth Revisited. MPRA Paper 24376 University Library of Munich, Germany.
- Reinhart, Carmen M. & Kenneth S. Rogoff. 2010b. "Growth in a Time of Debt." *American Economic Review Papers and Proceedings* 100(2):573–78.
- Reinhart, Carmen M. & Kenneth S. Rogoff. 2010c. Growth in a Time of Debt. Working Paper 15639 NBER.
- Reinhart, Carmen M. & Kenneth S. Rogoff. 2011. A Decade of Debt. NBER Working Papers 16827 National Bureau of Economic Research, Inc.
- Reinhart, Carmen M., Kenneth S. Rogoff & Miguel A. Savastano. 2003. "Debt Intolerance." *Brookings Papers on Economic Activity* 34(1):1–74.
- Reinhart, Carmen M., Vincent R. Reinhart & Kenneth S. Rogoff. 2012. "Public Debt Overhangs: Advanced-Economy Episodes Since 1800." *Journal of Economic Perspectives* 26(3):69–86.
- Sachs, Jeffrey D. 1989. The Debt Overhang of Developing Countries. In *Debt, Stabilization and Development*, ed. Guillermo A. Calvo, Ronald Findlay, Pentti Kouri & Jorge Braga de Macedo. Oxford: Blackwell.
- Shin, Yongcheol, Byungchul Yu & Matthew Greenwood-Nimmo. 2013. Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework. In *Festschrift in Honor of Peter Schmidt*, ed. William C. Horrace & Robin C. Sickles. New York (NY): Springer.
- Smith, Ron P. 2001. "Estimation and Inference with Non-Stationary Panel Time-Series Data." Paper prepared for the RC33 Conference.
- Temple, Jonathan. 1999. "The New Growth Evidence." *Journal of Economic Literature* 37(1):112–156.

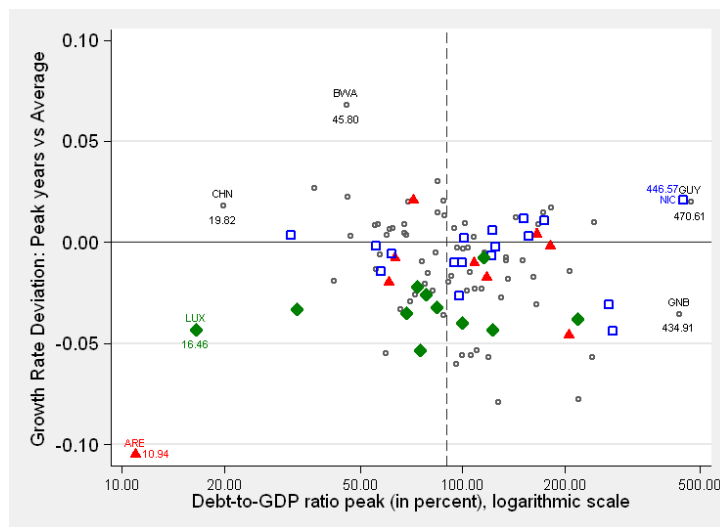
FIGURES AND TABLES

Figure 1: Peak Debt/GDP Ratio Distribution



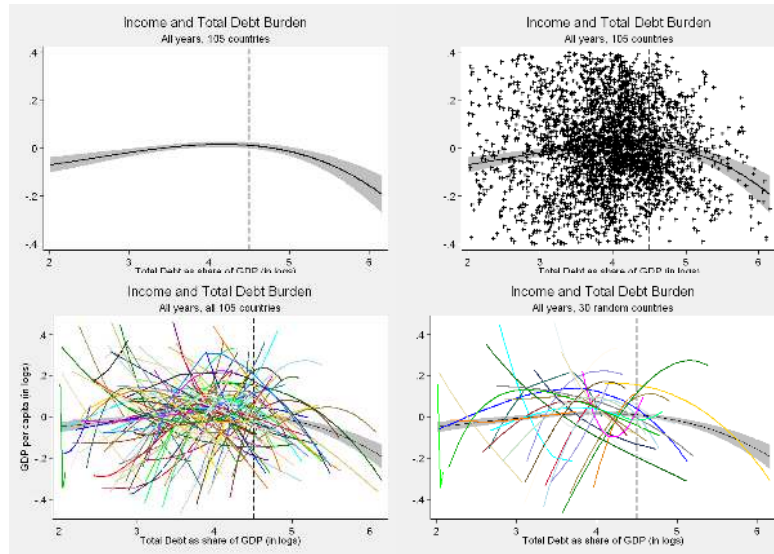
Notes: The histogram indicates the distribution of peak years for the debt-to-GDP ratio in our sample of 105 countries. Three years, 1985, 1994 and 2009 account for over one third of all debt/GDP peaks.

Figure 2: Peak Debt/GDP Ratio and Relative Growth



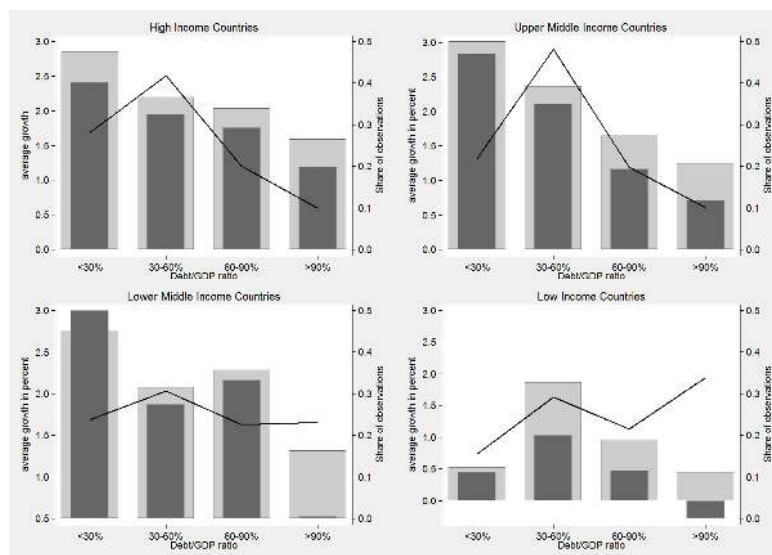
Notes: Along the x -axis we arrange countries by the value of the maximum debt-to-GDP ratio (in logarithms), highlighting three years in particular: 1985 (Triangles), 1994 (Squares), and 2009 (Diamonds). Along the y -axis we plot the deviation of countries' (i) average per capita growth rate in the five years around their peak debt year (i.e. peak debt occurs in year 3) from (ii) their average per capita growth rate over the entire time horizon 1972-2009 excluding the five 'peak debt years.' A simple (outlier-robust) linear regression of average per capita growth rates on debt-to-GDP peaks (in logarithms) yields (absolute t -ratios in brackets): $.011 [0.54] - .005 [1.12] \log(\text{debt/GDP})_i^{\max}$

Figure 3: Nonlinearities in the Country-Specific Debt-Income Nexus



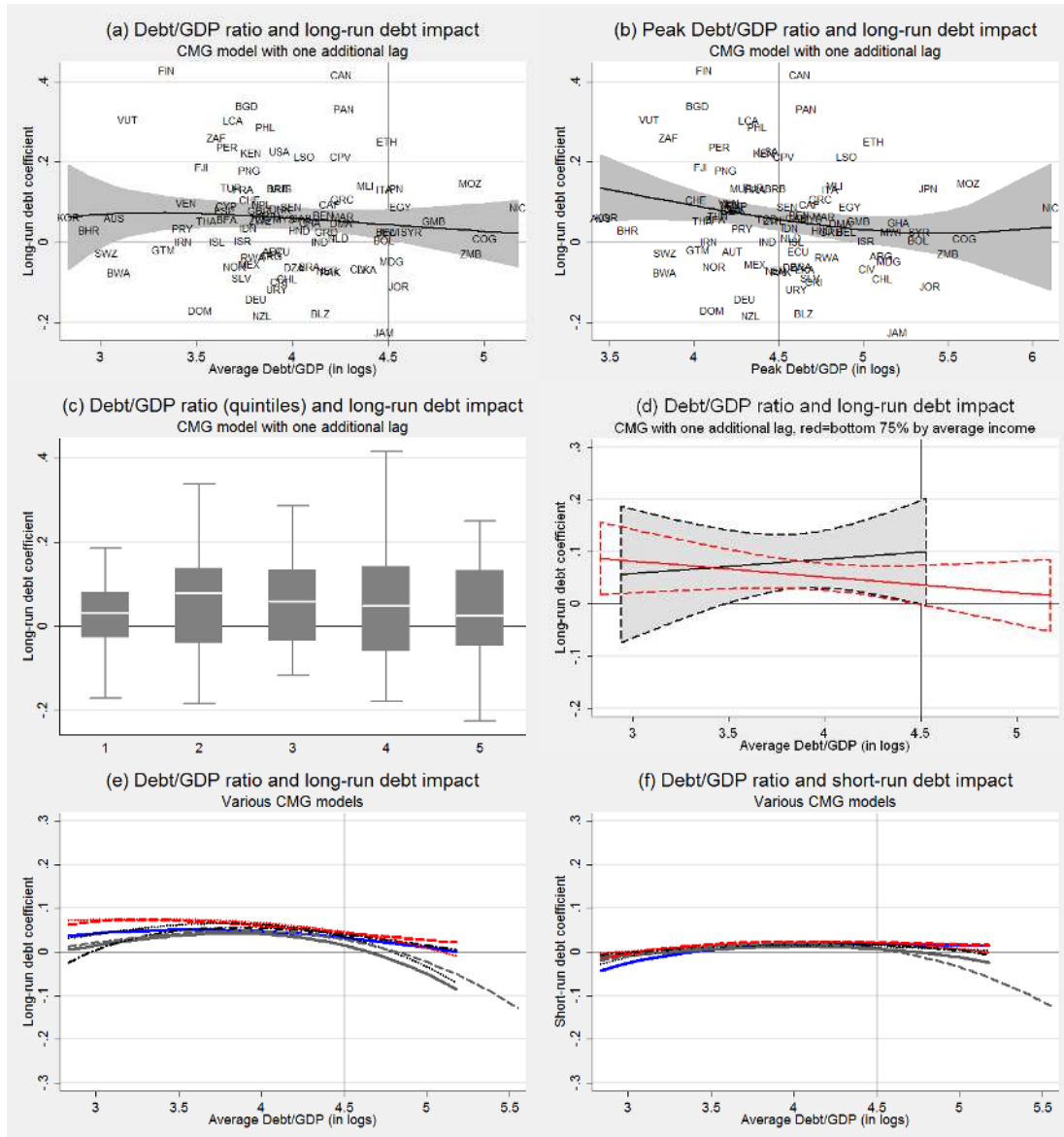
Notes: This figure plots the unconditional relation between debt/GDP ratio and within-transformed per capita GDP (both in logs). We employ fractional polynomial regression (solid regression line; shaded 95% confidence intervals) for all observations – see Footnote 10 for details on sample restriction. In the right plot of the first row we further provide a scatter of all observations; in the left plot of the second row we instead add fractional polynomial regression lines estimated for each country separately, while in the right plot we pick these for 30 countries at random.

Figure 4: The Rogoff and Reinhart (2010) approach in our dataset



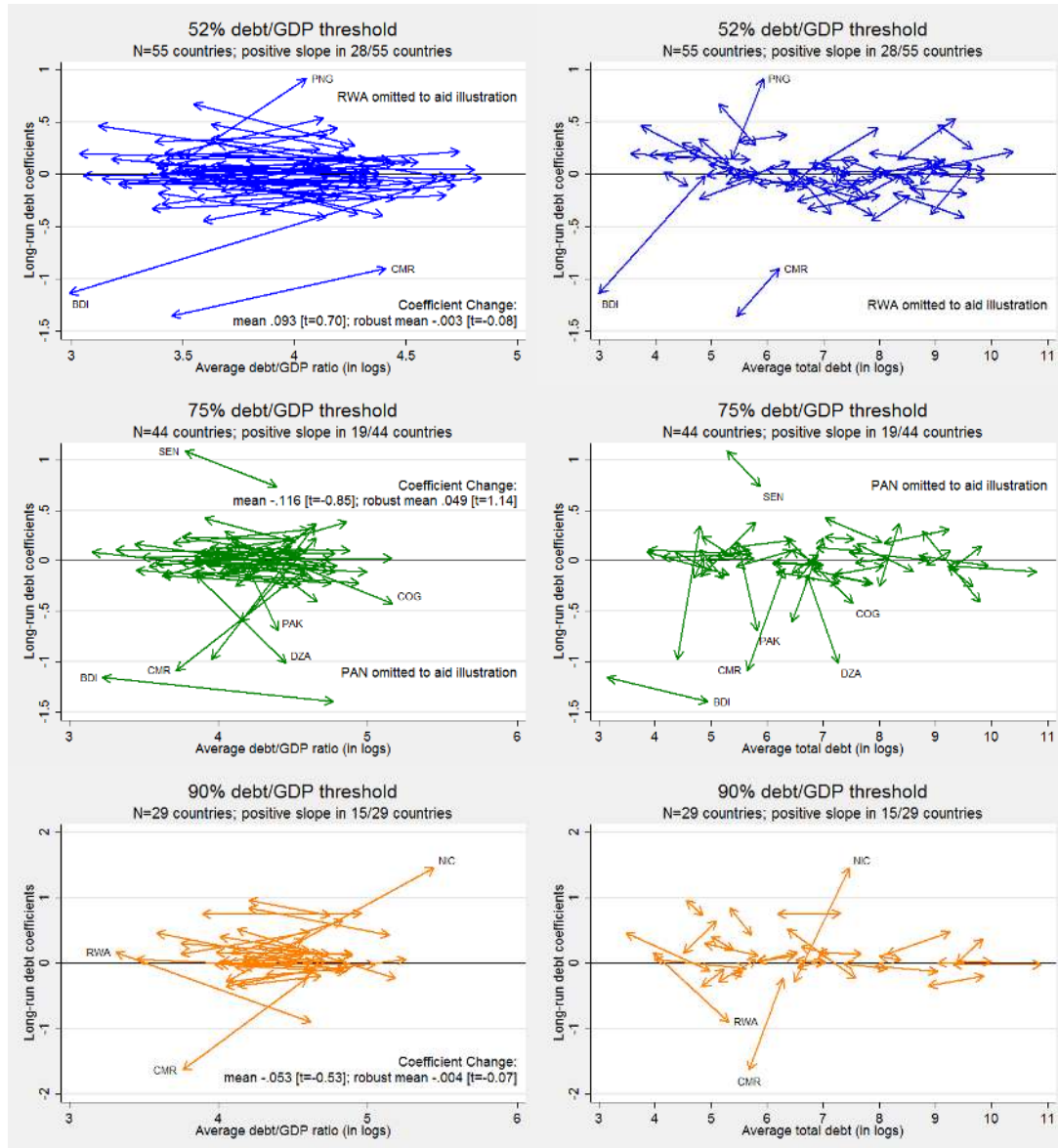
Notes: In each plot the light-grey bars represent median growth rates, the dark-grey bars the mean growth rates (both left axis), the black line the share of total observations (right axis) for each group respectively. For High Income Countries we have a total of 1,001 observations (29 countries), for the Upper Middle Income, Lower Middle Income and Low Income Countries these figures are 752 (23), 1,021 (30) and 685 (23), respectively. Income classification follows the World Bank approach.

Figure 5: Patterns for CMG debt coefficients



Notes: We plot the country specific long-run coefficients for debt in each country, taken from the dynamic CMG model with one additional lag (in column [6] of Table 1) against (a) the country-specific average debt/GDP ratio (in logs), and (b) the country-specific peak value for debt/GDP (in logs) — for both plots we reduce the number of countries as detailed below to improve illustration. In both cases we added fitted fractional polynomial regression lines along with 5% and 95% confidence bands (shaded area). We further provide (c) box plots for all 105 country-estimates divided into quintiles of the average country debt/GDP ratio distribution — outliers are omitted from these box plots and we focus on the medians and interquartile ranges (shaded). In (d) we split the sample into the top 25% and bottom 75% by average income and fit fractional polynomial regression lines alongside 5% and 95% confidence bands for each grouping (reduced sample in the plot for illustration). The final set of plots in (e) and (f) presents fitted fractional polynomial regression lines of long-run and short-run debt coefficients against average debt/GDP ratio for all CMG models (columns [4]-[10]), respectively. In each case (as in the first two scatter plots) we omit the observations with average debt/GDP ratio below 12% (ARE, CHN, LUX) as well as countries with absolute long-run debt coefficients (ALR) over 0.5 resulting in 92 [4], 95 [5], 92 [6], 91 [7], 95 [8], 94 [9] and 93 [10] out of 105 observations. This practice excludes the following country estimates in four or more of the seven models: GNB, GUY, HUN, IRL, SLE, SYC, TGO. In all plots we add a horizontal line to mark zero, in most plots we also add a vertical line at 4.5 log points ($\equiv 90\%$) of the debt/GDP ratio.

Figure 6: Debt Coefficient Comparison: three debt-to-GDP thresholds



Notes: We plot the long-run debt coefficients in the low and high debt regime for (top) 52%, (middle) 75% and (bottom) 90% debt/GDP thresholds. In each case we use the CCEMG results with one additional lag of cross-section averages (model [5] in Table TA5) for 55, 44 and 29 countries, respectively — countries are only included if they have at least 20% of their observations in one of the two regimes (below/above threshold). The values on the x-axis represents the average debt/GDP ratio and the average debt stock per capita (both in logarithms) for the lower and higher regimes (average over all years in each regime), in the left and right plot respectively. We carried out empirical tests for statistical significance of average coefficient changes at each threshold and report the mean and robust mean estimates together with respective t -ratios.

Table 1: Linear Dynamic Models

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
	2FE	CCEP	MG	CMG	CMG	CMG	CMG	CMG	CMG	CMG
additional covariate(s) †						one lag	two lags	open	open	open, findev
additional lagged CA									one lag	
country trends ‡			×		×					
Debt coefficients										
LRA	-0.034 [0.023]	0.000 [0.018]	-0.004 [0.011]	0.035 [0.013]***	0.016 [0.012]	0.050 [0.013]***	0.044 [0.013]***	0.027 [0.014]*	0.034 [0.014]**	0.031 [0.017]*
ALR			-0.011 [0.014]	0.036 [0.013]***	0.016 [0.012]	0.046 [0.013]***	0.040 [0.012]***	0.049 [0.017]***	0.029 [0.015]**	0.053 [0.017]***
SR	0.000 [0.007]	0.001 [0.009]	-0.015 [0.007]**	0.007 [0.007]	-0.001 [0.007]	0.009 [0.008]	0.009 [0.009]	0.010 [0.009]	0.003 [0.008]	0.004 [0.009]
ALR threshold			96.3 %	90.2%	73.4%	22.5%	0.3%	126.7%	28.4%	18.5%
EC coefficient										
$\gamma_{i,t-1}$	-0.108 [0.014]***	-0.339 [0.028]***	-0.487 [0.027]***	-0.559 [0.031]***	-0.655 [0.031]***	-0.587 [0.033]***	-0.656 [0.037]***	-0.608 [0.033]***	-0.674 [0.038]***	-0.634 [0.034]***
t -statistic ^b	-8.23	-12.31	-19.46	-18.98	-22.22	-18.00	-17.87	-18.98	-18.84	-18.76
\bar{t} -statistic			-3.45	-3.38	-3.94	-3.14	-2.98	-3.46	-3.25	-3.38
Implied half-life (years)	6.06	1.67	1.04	0.85	0.65	0.78	0.65	0.74	0.62	0.69
RMSE	0.040	0.036	0.030	0.025	0.023	0.022	0.019	0.02	0.018	0.020
GD test	-0.89	12.57	18.15	2.05	2.04	1.84	2.60	3.74	3.07	2.87
Observations	3,485	3,485	3,485	3,485	3,485	3,419	3,354	3,403	3,337	3,403

Notes: Results for full sample of $N = 105$ countries, based on an error correction model with the first difference of log real GDP per worker as dependent variable. We report the robust mean of coefficients across countries in the heterogeneous parameter models in [3]-[10] (Hamilton, 1992) unless indicated; standard errors in these models are constructed following Pesaran & Smith (1995). † The CMG estimator (Pesaran, 2006; Chudik & Pesaran, 2013) is implemented using further cross-section averages (CA) of (a) additional lags and/or (b) other variables (open – trade/GDP, findev – financial development, both in logs) as indicated – see main text for details. ‡ These models are augmented with country-specific linear trend terms. ‘LRA’ refers to the long-run average coefficient, which is calculated directly from the pooled model ECM results in [1] and [2] and the averages of the heterogeneous model ECM results (standard errors computed via the Delta method) in [3]-[10]. ‘ALR’ refers to the average long-run coefficient in the heterogeneous models, whereby the long-run coefficients are computed from the ECM results in each country and then averaged across the panel. ‘SR’ refers to the short-run coefficients. ‘ALR threshold’ indicates the cross-country implied threshold (i.e. where the impact of debt becomes negative) derived from a linear regression of the long-run debt coefficient on average debt-to-GDP ratio over the entire time period. ^b The first set of t -statistics are non-parametric statistics derived from the country-specific coefficients following Pesaran & Smith (1995). The second set represent averages across country-specific t -statistics. *, **, and *** indicate significance at 10%, 5% and 1% level respectively. RMSE is the root mean squared error, CD test reports the Pesaran (2004) test, which under the null of cross-section independence is distributed standard normal.

Table 2: Weak Exogeneity Testing

Model	Equation	lags	Panel A: Without CA				Panel B: With CA			
			GM- t	p	Avg $\hat{\lambda}_i$	t -stat	GM- t	p	Avg $\hat{\lambda}_i$	t -stat
MG [3]	Output	1	-2.54	0.01	-0.928	-21.98				
	Capital	1	-0.22	0.82	-0.030	<u>-2.09</u>				
	Debt stock	1	0.09	0.93	0.254	<u>1.71</u>				
	Output	2	-2.18	0.03	-0.990	-19.60				
	Capital	2	-0.12	0.90	-0.014	-0.89				
	Debt stock	2	0.12	0.90	0.239	1.11				
CMG [4]	Output	1	-2.23	0.03	-0.842	-19.32	-2.06	0.04	-0.800	-17.00
	Capital	1	-0.09	0.93	-0.012	-1.14	0.18	0.86	0.023	1.70
	Debt stock	1	0.30	0.76	0.649	<u>3.70</u>	0.17	0.86	0.490	<u>2.56</u>
	Output	2	-2.03	0.04	-0.875	-17.62	-1.64	0.10	-0.916	-12.67
	Capital	2	-0.05	0.96	-0.010	-0.78	0.14	0.89	0.031	1.65
	Debt stock	2	0.28	0.78	0.712	<u>3.33</u>	-0.01	0.99	0.405	1.54
CMG with trend [5]	Output	1	-2.56	0.01	-1.000	-24.87	-2.44	0.01	-0.956	-20.92
	Capital	1	-0.25	0.80	-0.029	-2.32	-0.08	0.94	-0.002	-0.10
	Debt stock	1	0.24	0.81	0.454	<u>2.80</u>	0.16	0.87	0.369	<u>1.92</u>
	Output	2	-2.25	0.02	-1.048	-22.12	-1.85	0.06	-1.141	-16.01
	Capital	2	-0.14	0.89	-0.024	-1.49	-0.04	0.97	-0.020	-1.01
	Debt stock	2	0.21	0.84	0.535	<u>2.62</u>	0.02	0.98	0.306	0.98
CMG with 1 add. lag [6]	Output	1	-1.88	0.06	-0.773	-18.04	-1.73	0.08	-0.742	-16.72
	Capital	1	0.02	0.99	0.003	0.22	0.22	0.83	0.025	1.48
	Debt stock	1	0.29	0.77	0.525	<u>2.63</u>	0.23	0.82	0.544	<u>2.36</u>
	Output	2	-1.76	0.08	-0.775	-14.76	-1.68	0.09	-0.860	-13.62
	Capital	2	0.05	0.96	0.010	0.66	0.20	0.84	0.026	1.34
	Debt stock	2	0.29	0.77	0.662	<u>3.08</u>	0.05	0.96	0.358	1.51
CMG with 2 add. lags [7]	Output	1	-1.73	0.08	-0.815	-17.08	-1.68	0.09	-0.813	-19.46
	Capital	1	-0.11	0.91	-0.007	-0.46	0.03	0.98	0.011	0.64
	Debt stock	1	0.16	0.87	0.325	1.52	0.16	0.88	0.386	1.58
	Output	2	-1.60	<u>0.11</u>	-0.775	-13.34	-1.61	<u>0.11</u>	-0.874	-13.19
	Capital	2	-0.06	0.95	0.006	0.32	0.01	<u>0.99</u>	0.004	0.19
	Debt stock	2	0.16	0.87	0.348	1.45	0.01	1.00	0.012	0.05
CMG with 1 add. covariate [8]	Output	1	-2.06	0.04	-0.815	-18.06	-1.93	0.05	-0.764	-16.25
	Capital	1	-0.08	0.94	-0.014	-1.42	0.11	0.91	0.010	0.86
	Debt stock	1	0.32	0.75	0.591	<u>3.15</u>	0.21	0.84	0.456	<u>2.22</u>
	Output	2	-1.95	0.05	-0.851	-16.20	-1.60	<u>0.11</u>	-0.892	-12.38
	Capital	2	-0.09	0.92	-0.017	-1.31	0.01	0.99	0.009	0.45
	Debt stock	2	0.28	0.78	0.608	<u>2.75</u>	0.03	0.97	0.308	0.99
CMG with 1 add. covariate & lag [9]	Output	1	-1.71	0.09	-0.802	-15.00	-1.61	<u>0.11</u>	-0.737	-13.93
	Capital	1	-0.01	0.99	-0.007	-0.58	0.12	0.91	0.018	1.22
	Debt stock	1	0.30	0.77	0.371	<u>1.74</u>	0.27	0.79	0.436	<u>1.91</u>
	Output	2	-1.66	0.10	-0.804	-12.77	-1.57	<u>0.12</u>	-0.847	-11.65
	Capital	2	0.02	0.98	0.007	0.46	0.12	<u>0.91</u>	0.015	0.73
	Debt stock	2	0.31	0.76	0.505	<u>2.14</u>	0.11	0.91	0.234	0.78
CMG with 2 add. covariates [10]	Output	1	-1.93	0.05	-0.826	-17.07	-1.80	0.07	-0.740	-15.31
	Capital	1	-0.06	0.95	-0.005	-0.43	0.09	0.93	0.017	1.14
	Debt stock	1	0.34	0.74	0.663	<u>3.41</u>	0.24	0.81	0.560	<u>2.77</u>
	Output	2	-1.83	0.07	-0.833	-15.91	-1.46	<u>0.14</u>	-0.773	-11.90
	Capital	2	-0.04	0.97	-0.005	-0.38	0.05	0.96	0.011	0.68
	Debt stock	2	0.31	0.76	0.587	<u>2.49</u>	0.06	0.96	0.224	0.81

Notes: Numbers in brackets correspond to the columns in Table 1. For the tests in Panel B cross-section averages of all variables are added to the estimation equation, whereas in Panel A we do not include these. All results are for $N = 105$. Equation refers to the ECM regression where the named variable is on the LHS, lags reports the number of lagged differences included in the regression. GM- t gives the group-mean average of country-specific t -ratios for the coefficient on the disequilibrium term ($\hat{\lambda}_i$) which is distributed $N(0, 1)$, p indicates the corresponding p -value. Avg $\hat{\lambda}_i$ refers to the robust mean coefficient on the ECM term, t -stat the corresponding t -statistic. Underlined p -values or 'robust' t -statistics indicate evidence against the hypothesis of a well-specified production function.

Table 3: Static Linear and Nonlinear Models

CA	No augmentation			Standard CA		
	[1]	[2]	[3]	[4]	[5]	[6]
Estimator	MG	MG	MG	CMG	CMG	CMG
Bal & Co-Sum						
$debt_{it}$	-0.059 [0.010]***	0.286 [0.156]*	1.989 [2.129]	0.000 [0.010]	0.338 [0.131]***	1.307 [1.697]
$debt_{it}^2$		-0.024 [0.011]**	-0.330 [0.331]		-0.027 [0.010]***	-0.178 [0.263]
$debt_{it}^3$			0.019 [0.017]			0.009 [0.014]
Nonlinearity †						
# of Countries	/10 \46	∪19 ∩38	~32	/18 \22	∪13 ∩27	~28
Diagnostics ‡						
Observations	3,613	3,613	3,613	3,613	3,613	3,613
RMSE	0.056	0.051	0.048	0.047	0.039	0.035
CD Test	25.69	25.85	22.97	4.48	4.06	3.48

CA	Add CA of additional covariates		
	[7]	[8]	[9]
Estimator	CMG	CMG	CMG
Bal & Co-Sum	×		
$debt_{it}$	-0.005 [0.010]	0.257 [0.121]**	3.115 [1.703]*
$debt_{it}^2$		-0.021 [0.009]**	-0.462 [0.266]*
$debt_{it}^3$			0.019 [0.013]
Nonlinearity †			
# of Countries	/18 \24	∪16 ∩28	~25
Diagnostics ‡			
Observations	3,531	3,531	3,531
RMSE	0.042	0.035	0.031
CD Test	1.87	1.69	1.98

Notes: We report the estimates and diagnostic tests for static production functions with linear and polynomial debt terms. All estimates are robust means (see Table 1). The MG models further include trend terms, we also omitted to report the averaged capital stock coefficients and constant terms in all models (available on request). ‘Bal & Co-Sum’ indicates the specification which was found to be balanced and co-summable (Tables TA7 and TA8). † We report the number of countries with statistically significant (5% level) positive and negative debt coefficient using / and \, respectively; ∪ and ∩ report the number of countries with statistically significant (5% level) convex and concave debt-growth relationships, respectively. ~ reports the number of countries for which all three debt terms are statistically significant at the 5% level. ‡ All residual series were found to be stationary.

DATA APPENDIX

A-I Data construction

The principle data sources for our empirical analysis are the World Bank World Development Indicators and an update to the dataset provided by Panizza (2008). From the former we take real GDP in year 2000 US\$ values, the per capita series of the same variable, population as well as gross fixed capital formation (investment) as a share of GDP. The Panizza data provides total debt series, comprising domestic and external debt, in face value terms as a percentage of GDP, enabling us to construct the real debt stock series.

With the investment series we can construct real capital stock by adopting the perpetual inventory method with a standard annual depreciation rate of 5%. If country series contained gaps of less than three years length we used cubic spline interpolation. This resulted in changes to a total of 53 observations in 19 country series,²⁵ thus an average of 2.8 per country. This amounts to a total of 1.36% of our full sample data. Note that this interpolation does not affect the overall sample size, since the observations in question are also missing for GDP and other variables; it does however aid the construction of the capital stock series.

In the process of constructing the capital stock series we investigated a number of basic magnitudes, including the investment-to-GDP ratio in 1970 (found to be between 10 and 40%) and the capital-output ratio in 1970 (found to be between 1.5 and 4.7), which were all within reasonable bounds and thus no adjustments were made. We did however limit our analysis to countries with at least 21 years of data, which effectively included any transition economy as well as a small number of African and Latin American countries.²⁶ The final sample contains

²⁵The affected countries are BEL, BEN, BHR, DNK, FIN, FRA, GHA, HUN, IND, KEN, LUX, MDG, MWI, NLD, NOR, NZL, TON, TTO, VEN – see Technical Appendix for country iso-codes. Note that of these Bosnia and Herzegovina (BHR), Tonga (TON) and Trinidad & Tobago (TTO) were later dropped due to insufficiently long time series.

²⁶The following 67 countries for which *some* debt-to-GDP information was available in the Panizza data (covering 172 countries) were thus dropped from the analysis, either due to lack of observations for *all* variable series or due to the time series restriction: Albania, Angola, Armenia, Azerbaijan, Bahamas, Belarus, Bhutan, Bosnia and Herzegovina, Brunei Darussalam, Bulgaria, Cambodia, Colombia, Croatia, Czech Republic, Djibouti, Equatorial Guinea, Eritrea, Estonia, Georgia, Guinea, Haiti, Hong Kong SAR, Iraq, Kazakhstan, Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Liberia, Lithuania, FYR Macedonia, Maldives, Mauritania, Moldova, Mongolia, Myanmar, Namibia, Nigeria, Oman, Poland, Portugal, Qatar, Romania, Russian Federation, Samoa, San Marino, Sao Tome and Principe, Saudi Arabia, Slovak Republic, Slovenia,

3,485 observations (dynamic specification) from 105 countries (23 Low-Income, 30 Lower Middle-Income, 23 Upper Middle-Income and 29 High Income countries based on World Bank classification), thus on average 33.2 years per country (range of 21 to 38 country observations) from 1972 to 2009.

Some of the empirical models make use of the cross-section averages of two proxies for trade openness and financial development: we use the data series provided by Bill Easterly for the former (exports plus imports as a share of GDP – data is from the WDI and Global Development Finance) and by Thorsten Beck and Asli Demirguc-Kunt (Financial Structure Database) for the latter (ratio of bank credit to bank deposits). Both raw variables are in logs.

We present descriptive statistics for our sample in Table A1. Detailed information about the sample make-up is confined to a Technical Appendix.

Solomon Islands, St. Kitts and Nevis, St. Vincent and the Grenadines, Sudan, Taiwan, China, Tajikistan, Tanzania, Tonga, Trinidad and Tobago, Turkmenistan, Uganda, Ukraine, Uzbekistan, Vietnam, Yemen (Rep).

A-II Descriptive Statistics

Table A1: Descriptive Statistics

PANEL A: RAW VARIABLES AND TRANSFORMATIONS						
variable	type	mean	median	sd	min	max
GDP	level	2.48E+11	1.25E+10	9.39E+11	1.18E+08	1.17E+13
GDP growth	%age growth rate	3.429	3.673	4.821	-69.812	33.280
GDP per capita	level	6,844	1,950	9,572	102	57,215
GDP pc growth	%age growth rate	1.637	2.050	4.780	-63.285	32.091
Population	level	3.96E+07	8.32E+06	1.35E+08	6.65E+04	1.32E+09
Population growth	%age growth rate	1.792	1.824	1.265	-8.271	14.734
Investment/GDP ratio	%age share of GDP	21.730	21.107	7.158	3.412	76.693
Capital Stock	level	7.01E+11	2.97E+10	2.55E+12	4.21E+08	2.91E+13
Capital Stock growth	%age growth rate	3.557	3.158	2.761	-3.629	28.813
Capital Stock per capita	level	1.99E+04	5.22E+03	2.88E+04	1.89E+02	1.56E+05
Capital Stock pc growth	%age growth rate	1.765	1.690	2.783	-9.607	25.157
Debt (total)	level	1.54E+11	6.98E+09	7.35E+11	9.34E+06	1.06E+13
Debt growth	%age growth rate	4.924	4.034	17.748	-142.053	136.954
Debt (total) per capita	level	3.52E+03	1.01E+03	6.37E+03	2.56E+00	8.33E+04
Debt pc growth	%age growth rate	3.132	2.492	17.749	-144.855	133.551
Debt/GDP ratio	%age share of GDP	62.177	51.728	49.618	0.971	470.610
PANEL B: REGRESSION VARIABLES (IN LOGS OR FIRST DIFFERENCES OF LOGS)						
variable		mean	median	sd	min	max
Δy_{it}		0.016	0.020	0.048	-0.633	0.321
$y_{i,t-1}$		7.706	7.554	1.617	4.628	10.955
$cap_{i,t-1}$		8.659	8.544	1.734	5.239	11.959
$debt_{i,t-1}$		6.949	6.890	1.628	0.939	11.272
Δcap_{it}		0.018	0.017	0.028	-0.096	0.252
$\Delta debt_{it}$		0.031	0.025	0.177	-1.449	1.336

Notes: We present descriptive statistics for the full sample of 3,485 observations from $N = 105$ countries (average $T=33.2$). In Panel A we added a number of standard transformations of the data applied, e.g. the debt/GDP ratio and the investment/GDP ratio as well as per capita GDP and its growth rate. Some of these variables are applied in the post-estimation analysis. In Panel B we present descriptives for the error correction model regression variables, namely Δy_{it} — GDP per capita growth rate, $y_{i,t-1}$ — lagged level of GDP per capita (in logs), $cap_{i,t-1}$ — lagged level of capital stock per capita (in logs), $debt_{i,t-1}$ — lagged level of debt stock per capita (in logs), Δcap_{it} — growth rate of capital stock per capita, $\Delta debt_{it}$ — growth rate of debt stock per capita.

TECHNICAL APPENDIX — not for publication

TA-I Selective Review of the Empirical Literature

In the following we provide a selective review of recent studies in the empirical literature of debt and growth. Table TA1 provides an overview of characteristics related to the sample (N) and its makeup, the period and time-series dimension of the data (T) and whether and how any data aggregation over time was carried out: until the most recent contributions which use annual data all studies investigated averaged the data over time, in line with the standard practice in the cross-country growth literature. Further details provided cover the empirical model setup, namely the dependent variable and covariates (including proxies for debt stock and debt service), as well as the empirical specification. We focus on the most general parametric and semi-parametric results in each paper. Most studies reviewed carried out a large number of regressions (robustness checks), including some adopting nonparametric methods; results for these are sometimes indicated but we abstract from a more detailed discussion for conciseness. The final column of the table indicates which specific regression results we base our discussion on. Regarding the variables entering the model there are minor differences across studies, although investment, trade openness (trade/GDP) and a measure of human capital are typically included. With regard to the latter, it is particularly notable that all papers reviewed adopt a pooled partial adjustment model (PAM) which includes some lagged level of GDP or GDP per capita as covariate and GDP growth or the per capita equivalent as the dependent variable. Pooled here indicates that it is assumed the equilibrium relationship is the same across all countries in the sample. Implementations are again typical of the standard in the cross-country growth literature, including OLS, FE and various instrumentation strategies (including Arellano and Bond (1991)-type estimators). All studies consider nonlinearities in the debt-growth relationship.

In our synthetic review we use subscript i for countries and t for time periods, *also* for averaged time periods (see column on averaging; in few cases time-periods overlap (indicated as OL), but typically periods are non-overlapping). LD refers to ‘long differences’ (e.g. $t-(t-30)$). Other abbreviations used are defined as follows:

Sample: LDCs – less developed economics; LICs – low income countries; MICs

– middle income countries; HICs – high income countries; EMEs – emerging economies; EAC – Euro Area countries. n refers to the number of observations in the regression (may be time-averaged); n refers to the number of observations; since this at times differs across specifications we provide ballpark figures.

Averaging: Time-period over which data is averaged; annual – no averaging. Note that the ‘convergence term’ $(\ln(Y/L))_{i,t-1}$, i.e. lagged level of log per capita GDP) is typically the value for the first year in a three or five-year period rather than an average value (annual data obviously excepted).

Dependent Variables: Y – real GDP; Y/L – real GDP per capita; $\Delta \ln(Y/L)$ – real GDP per capita growth.

Debt Variables: ED/Y – external debt to GDP ratio; ED/EX – external debt to exports ratio; similarly for Total Debt (TD). FV and NPV refer to face- and net present value of debt.

Other Variables: Inv/Y – investment to GDP ratio; HC – human capital proxy (typically Barro & Lee, 2013); ΔL – population growth; ΔTOT – changes in terms of trade; $[EX+IM]/Y$ – trade openness; (LL/Y) – liquid liabilities to GDP ratio; fin/Y – financial openness; FB/Y – fiscal balance over GDP; IR – interest rate; REER – real effective exchange rate; inf – (CPI) inflation; DR – dependency ratio; $FCDebt/Debt$ – foreign currency debt to total debt ratio; D – dummy for certain events (crises; EMU membership) amongst others; σ^{inf} – inflation volatility; σ^Y – GDP pc growth volatility; $Credit/GDP$ – private credit to GDP ratio; Aid/Y – foreign aid to GDP ratio, $inst$ – institutional quality/good governance. * Values are decadal averages with the exception of the ‘initial income’ variable, which is the value for the first year of each decade; ‘lags’ are 5-year averages for the 5 years immediately preceding a decade.

Model: PAM – partial adjustment model (growth regressed on lagged level of dependent variables and contemporaneous covariates); MRW – cross-section convergence regression model following Mankiw, Romer & Weil (1992) with the panel aspect unexploited in the empirics.

Estimators: OLS – ordinary least squares; FE – one-way fixed effects; BE – between groups estimator (cross-section of averages); DGMM – Arellano & Bond (1991); SGMM – Blundell & Bond (1998); — STR-GMM – structural threshold regression model combined with GMM approach; PTSR – Panel Smooth Transition Regression.

Table TA1: Continued

Authors	Sample	Period	Averaging	Dep. Var.	Debt stock	Debt service	Covariates†	Model	Estimator(s)	Nonlinearity	Reference
Cordella, Ricci & Ruiz-Arranz (2010)	79 LDCs (n ≈ 700)	1970-2002	3-year	$\Delta \ln(Y/L)_{it}$	$\ln(ED/Y)_{it}$, $\ln(ED/Y)_{it}^2$, $\ln(ED/Y)_{it} \times D$, $\ln(ED/Y)_{it}^2 \times D$ (NPV)	$(TDS/EX)_{it}$, $(TDS/EX)_{it} \times D$	$\ln(Y/L)_{it-t-1}$, $\ln(Inv/Y)$, $\ln(HC)$, $\ln(\Delta L)$, ΔTOT , $[EX+IM]/Y$, FB/Y , Aid/Y , High-Debt Dummy	Pooled PAM	OLS, SGMM	Squared debt term; spline regression (dummy), split at median	Table 2
<p>OLS: overhang → insignificant irrelevance as ED ↑ (SGMM all insignificant); differs by institutional quality</p>											
Kumar & Woo (2010)	38 HICs and EMEs (n ≈ 179)	1970-2007	5-year LD, 4-year lags	$\ln(Y/L)_{it} - \ln(TD/Y)_{it-t}$	$\ln(ED/Y)_{it-t}$ (average) and interactions (< 30%, 30 – 90%, > 90%)		$\ln(Y/L)_{it-t}$, ΔTOT , $(HC)_{it-t}$, D^{crisis} , $\ln(1+\ln d)_{it-t}$, $(C^{gov}/Y)_{it-t}$, FD , $\ln([EX+IM]/Y)_{it-t}$, $\ln(LL/Y)_{it-t}$; also: $\ln(Inv)_{it-t}$, I_{it-t} , ...	Pooled PAM	OLS, BE, FE, SGMM	Interactions (< 30%, 30 – 90%, > 90% of GDP)	Tables 1, 5
<p>negative significant (linear); some evidence of overhang >90%</p>											
Cecchetti, Mohanty & Zampolli (2011)	18 HICs (n ≈ 360)	1980-2010	5-year (OL)	$\Delta \ln(Y/L)_{it}$	$(TD/Y)_{it}$; also corporate, hh and total (all three) debt; threshold interactions		$\ln(Y/L)_{it-1}$, (Sav/Y) , HC , ΔL , DR , D^{crisis} , $\ln([EX+IM]/Y)$, (LL/Y)	Pooled PAM	FE	threshold regression	Tables 5, 6 (gov debt)
<p>negative linear effect driven by threshold</p>											
Pattillo, Poisson & Ricci (2011)	93 LDCs (n ≈ 630)	1969-98	3-year	$\Delta \ln(Y/L)_{it}$	$\ln(ED/Y)_{it}$, $\ln(ED/Y)_{it}^2$, $\ln(ED/EX)_{it}$, $\ln(ED/EX)_{it}^2$ (FV or NPV)	$(TDS/EX)_{it}$	$\ln(Y/L)_{it-1}$, $\ln(HC)$, $\ln(Inv/Y)$, $\ln(\Delta L)$, ΔTOT , $[EX+IM]/Y$, FB/Y	Pooled PAM	FE, SGMM	Squared debt term (also spline regression)	Table 4
<p>significant, concave relationship in most models</p>											
<p>debt detrimental beyond threshold at 96% debt/GDP</p>											
<p>growth-retarding threshold > 35% (ED/Y), > 160% (ED/EX)</p>											

(Continued)

Table TA1: Continued

Authors	Sample	Period	Averaging	Dep. Var.	Debt stock	Debt service	Covariates†	Model	Estimator(s)	Nonlinearity	Reference
Calderon & Fuentes (2012)	116 countries (n ≈ 740)	1970-2010	5-year	$\Delta \ln(Y/L)_{it}$	$\ln(TD/Y)_{it}$		$\ln(Y/L)_{it,t-1}$, $\ln(HC)$, $\ln([\text{Credit}/\text{GDP}]^*100)$, $\ln(\text{Inst})$, $\ln(\text{inf})_{it}$, $(\text{FB}/Y)^*100$, $\ln(100*[\text{EX}+\text{IM}]/Y)$, $\ln([\text{fin}/Y]^*100)$; σ^y , debt-interactions	Pooled PAM	SGMM	Squared debt term; debt-interactions	Tables 2, 3
					negative significant, mitigated in rich, financially-developed countries with good institutions					inverted-U; debt negative for $Y/L > \$7k$	
Checherita-Westphal & Rother (2012)	12 EAC (n ≈ 390)	1970-2008	annual (also: 5-year)	$\Delta \ln(Y)_{it}$	$(TD/Y)_{it}$, $(TD/Y)_{it}^2$		$\ln(Y/L)_{it,t-1}$, Inv/Y , ΔL , tax , FB/Y , IR , REER , $[\text{EX}+\text{IM}]/Y$, ΔTOT , inf , DR	Pooled PAM	FE, 2SLS (lags); (also GMM)	Squared debt term	Tables 1, 3
					significant, concave relationship					inverted-U with 90-100% debt/GDP threshold	
Minea & Parent (2012)	20 HICs (n ≈ 1,300)	1945-2009	annual	$\Delta \ln(Y)_{it}$	$\ln(TD/Y)_{it}$		-	Pooled	PSTR	multiple smooth thresholds nonlinear debt-growth relation, +ve above 115%	Table 1 & Figure 2
Presbitero (2012)	92 LICs & MICs (n ≈ 320)	1990-2007	3-year	$\Delta \ln(Y/L)_{it}$	$\ln(TD/Y)_{it}$; $\ln(TD/Y)_{it}^2$; $t-1$	also at $t-1$	$\ln(Y/L)_{it,t-1}$, $\ln(\text{Inv})$, $\ln(HC)$, ΔTOT , $\ln([\text{EX}+\text{IM}]/Y)$, CPIA , σ^{inf}	Pooled PAM	SGMM	Squared debt term; spline regression (dummy); CPIA interactions	Tables 3, 4
					debt irrelevance threshold >90%; overhang only in high CPIA economies					inverted-U relation driven by a few observations	
Panizza & Presbitero (2012)	17 OECD	1970-2008	annual	forward $\Delta \ln(Y/L)_{it}$ (t + 1 to t + 6)	$(TD/Y)_{it}$; also corporate, hh and total (all three) debt; threshold interactions		$\ln(Y/L)_{it}$, (Sav/Y) , HC , ΔL , DR , D^{Crisis} , inf , $([\text{EX}+\text{IM}]/Y)$, (LL/Y) , $(\text{FCDebt}/\text{Debt})_{it,t-1}$, $\text{REER}_{t,t-1}$	Pooled PAM	OLS, IV (valuation effect on debt)	threshold regression	Table 3
					negative linear effect disappears with instrumentation					no significant effect found	

(Continued)

TA-II Sample Makeup

Table TA2: Sample details

wbcode	Country	Region	Income	Obs	Coverage	Missing
ARE	United Arab Emirates	MENA	HIC non-OECD	31	1978-2007	
ARG	Argentina	LAC	Upper MIC	23	1973-2009	1980-95
AUS	Australia	EAP	High income: OECD	37	1973-2008	
AUT	Austria	EEE	High income: OECD	38	1973-2009	
BDI	Burundi	SSA	LIC	35	1973-2006	
BEL	Belgium	EEE	High income: OECD	35	1973-2009	1989-92
BEN	Benin	SSA	LIC	23	1985-2006	
BFA	Burkina Faso	SSA	LIC	26	1982-2006	
BGD	Bangladesh	SA	LIC	28	1983-2009	
BHR	Bahrain	EAP	HIC non-OECD	24	1983-2008	1990-93
BLZ	Belize	LAC	Lower MIC	27	1983-2008	
BOL	Bolivia	LAC	Lower MIC	38	1973-2009	
BRA	Brazil	LAC	Upper MIC	30	1981-2009	
BRB	Barbados	LAC	HIC non-OECD	26	1973-2002	1974-79
BWA	Botswana	SSA	Upper MIC	38	1973-2009	
CAF	Central African Republic	SSA	LIC	31	1980-2008	
CAN	Canada	NA	High income: OECD	38	1973-2009	
CHE	Switzerland	EEE	High income: OECD	38	1973-2009	
CHL	Chile	LAC	Upper MIC	38	1973-2009	
CHN	China	EAP	Lower MIC	27	1984-2009	
CIV	Cote d'Ivoire	SSA	Lower MIC	38	1973-2009	
CMR	Cameroon	SSA	Lower MIC	31	1978-2007	
COG	Congo, Rep.	SSA	Lower MIC	34	1977-2008	
CPV	Cape Verde	SSA	Lower MIC	22	1989-2008	
CRI	Costa Rica	LAC	Upper MIC	38	1973-2009	
CYP	Cyprus	EEE	HIC non-OECD	32	1978-2008	
DEU	Germany	EEE	High income: OECD	38	1973-2009	
DMA	Dominica	LAC	Upper MIC	31	1983-2006	
DNK	Denmark	EEE	High income: OECD	29	1979-2009	1997-2000
DOM	Dominican Republic	LAC	Upper MIC	38	1973-2009	
DZA	Algeria	MENA	Upper MIC	38	1973-2009	
ECU	Ecuador	LAC	Lower MIC	38	1973-2009	
EGY	Egypt	MENA	Lower MIC	38	1973-2009	
ESP	Spain	EEE	High income: OECD	38	1973-2009	
ETH	Ethiopia	SSA	LIC	27	1984-2009	
FIN	Finland	EEE	High income: OECD	34	1973-2009	1979-83
FJI	Fiji	EAP	Upper MIC	37	1973-2008	
FRA	France	EEE	High income: OECD	34	1973-2009	1978-82
GAB	Gabon	SSA	Upper MIC	38	1973-2009	
GBR	United Kingdom	EEE	High income: OECD	38	1973-2009	

Continued on the following page.

Table TA2: Continued

wbcode	Country	Region	Income	Obs	Coverage	Missing
GHA	Ghana	SSA	LIC	35	1973-2009	1990-92
GMB	The Gambia	SSA	LIC	22	1984-2004	
GNB	Guinea-Bissau	SSA	LIC	22	1982-2002	
GRC	Greece	EEE	High income: OECD	33	1973-2009	1976-81
GRD	Grenada	LAC	Upper MIC	31	1980-2008	
GTM	Guatemala	LAC	Lower MIC	38	1973-2009	
GUY	Guyana	LAC	Lower MIC	37	1973-2008	
HND	Honduras	LAC	Lower MIC	38	1973-2009	
HUN	Hungary	LAC	Lower MIC	21	1984-2008	
IDN	Indonesia	SA	Lower MIC	29	1982-2009	1993-96
IND	India	EAP	Lower MIC	24	1973-2009	1974-78, 83-93
IRL	Ireland	EEE	High income: OECD	38	1973-2009	
IRN	Iran	MENA	Upper MIC	26	1983-2007	1991-95
ISL	Iceland	EEE	High income: OECD	36	1975-2009	
ISR	Israel	MENA	High income: OECD	38	1973-2009	
ITA	Italy	EEE	High income: OECD	38	1973-2009	
JAM	Jamaica	LAC	Upper MIC	26	1972-1997	
JOR	Jordan	MENA	Lower MIC	32	1979-2009	
JPN	Japan	EAP	High income: OECD	38	1973-2009	
KEN	Kenya	SSA	LIC	34	1973-2009	1977-81
KOR	Korea	EAP	High income: OECD	38	1973-2009	
LCA	St. Lucia	LAC	Upper MIC	27	1984-2009	
LKA	Sri Lanka	SA	Lower MIC	38	1973-2009	
LSO	Lesotho	SSA	Lower MIC	38	1973-2009	
LUX	Luxembourg	EEE	High income: OECD	31	1977-2009	1990-93
MAR	Morocco	MENA	Lower MIC	38	1973-2009	
MDG	Madagascar	SSA	LIC	35	1973-2008	1983-86
MEX	Mexico	LAC	Upper MIC	38	1973-2009	
MLI	Mali	SSA	LIC	36	1973-2007	
MOZ	Mozambique	SSA	LIC	24	1987-2009	
MUS	Mauritius	SSA	Upper MIC	32	1983-2009	
MWI	Malawi	SSA	LIC	32	1976-2009	2002-05
MYS	Malaysia	EAP	Upper MIC	38	1973-2009	
NER	Niger	SSA	LIC	24	1983-2005	
NIC	Nicaragua	LAC	Lower MIC	33	1973-2009	1988-93
NLD	Netherlands	EEE	High income: OECD	35	1973-2009	2000-03
NOR	Norway	EEE	High income: OECD	34	1973-2009	1981-85
NPL	Nepal	SA	LIC	33	1978-2008	
NZL	New Zealand	EAP	High income: OECD	35	1973-2009	1999-02
PAK	Pakistan	SA	Lower MIC	33	1973-2009	1991-96
PAN	Panama	LAC	Upper MIC	28	1983-2009	
PER	Peru	LAC	Upper MIC	38	1973-2009	
PHL	Philippines	EAP	Lower MIC	38	1973-2009	
PNG	Papua New Guinea	EAP	Lower MIC	35	1976-2009	
PRY	Paraguay	LAC	Lower MIC	38	1992-2009	
RWA	Rwanda	SSA	LIC	38	1973-2008	
SEN	Senegal	SSA	Lower MIC	38	1973-2009	
SLE	Sierra Leone	SSA	LIC	26	1983-2007	
SLV	El Salvador	LAC	Lower MIC	38	1973-2009	
SWE	Sweden	EEE	High income: OECD	38	1973-2009	

Continued on the following page.

Table TA2: Continued

wbcode	Country	Region	Income	Obs	Coverage	Missing
SWZ	Swaziland	SSA	Lower MIC	38	1973-2009	
SYC	Seychelles	SA	Upper MIC	28	1983-2009	
SYR	Syrian Arab Republic	MENA	Lower MIC	38	1973-2009	
TCD	Chad	SSA	LIC	26	1985-2008	
TGO	Togo	SSA	LIC	24	1983-2005	
THA	Thailand	EAP	Lower MIC	38	1973-2009	
TUN	Tunisia	MENA	Lower MIC	38	1973-2009	
TUR	Turkey	EEE	Upper MIC	33	1979-2009	
URY	Uruguay	LAC	Upper MIC	38	1972-2009	
USA	United States	NA	High income: OECD	38	1973-2009	
VEN	Venezuela	LAC	Upper MIC	34	1973-2009	1992-96
VUT	Vanuatu	EAP	Lower MIC	23	1986-2007	
ZAF	South Africa	SSA	Upper MIC	38	1973-2009	
ZMB	Zambia	SSA	LIC	38	1973-2009	
ZWE	Zimbabwe	SSA	LIC	28	1979-2005	

Notes: Regional codes are EAP — East Asia & Pacific; EEE — Emerging Economies in Europe; LAC — Latin America & Caribbean; MENA — Middle East & North Africa; SA — South Asia; SSA — Sub-Saharan Africa.

Economies are divided among income groups according to 2009 gross national income (GNI) per capita, calculated using the World Bank Atlas method (see <http://tinyurl.com/pc8rpn>): LIC — Low-Income Country (\$995 or less); Lower MIC — Lower Middle-Income Country (\$996-3,945); Upper MIC — Upper Middle-Income Country (\$3,946-12,195).

'Obs' indicates the time-series observations available.

TA-III Stationarity testing

Table TA3: Panel Stationarity Testing

Maddala and Wu (1999) Fisher test (deterministics: constant)						
Lags	GDP pc	(p)	Debt pc	(p)	Cap pc	(p)
0	222.59	0.26	388.92	0.00	703.69	0.00
1	175.52	0.96	294.86	0.00	212.81	0.43
2	150.92	1.00	333.46	0.00	223.75	0.25
3	155.73	1.00	437.48	0.00	241.54	0.07
4	182.75	0.91	337.02	0.00	221.66	0.28

Maddala and Wu (1999) Fisher test (deterministics: constant and trend term)						
Lags	GDP pc	(p)	Debt pc	(p)	Cap pc	(p)
0	128.12	1.00	120.86	1.00	259.54	0.01
1	219.89	0.31	157.22	1.00	447.84	0.00
2	207.42	0.54	165.42	0.99	301.60	0.00
3	207.41	0.54	233.60	0.13	306.65	0.00
4	174.95	0.96	163.06	0.99	258.33	0.01

Pesaran (2007) CIPS test (deterministics: constant)						
Lags	GDP pc	(p)	Debt pc	(p)	Cap pc	(p)
0	3.86	1.00	4.07	1.00	2.67	1.00
1	4.10	1.00	3.45	1.00	3.47	1.00
2	2.98	1.00	4.66	1.00	2.98	1.00
3	2.51	0.99	3.78	1.00	3.86	1.00
4	5.29	1.00	4.89	1.00	5.82	1.00

Pesaran (2007) CIPS test (deterministics: constant and trend)						
Lags	GDP pc	(p)	Debt pc	(p)	Cap pc	(p)
0	3.39	1.00	4.14	1.00	10.80	1.00
1	-1.12	0.13	3.45	1.00	-3.74	0.00
2	-0.03	0.49	4.56	1.00	3.60	1.00
3	0.87	0.81	5.27	1.00	3.79	1.00
4	5.94	1.00	6.50	1.00	7.75	1.00

Notes: All variables in logarithms. For the Maddala and Wu (1999) test we report the Fisher statistic and associated p -value, for the Pesaran (2007) test the standardised Z-tbar statistic and its p -value. The null hypothesis for both tests is that all series are nonstationary. Lags indicates the lag augmentation in the Dickey Fuller regression employed. Augmentation of the Dickey Fuller regressions with a constant or a constant and trend as indicated. We used the Stata routine `multipur` by Markus Eberhardt, which wraps the routines `xtfisher` and `pesca`df written by Scott Merryman and Piotr Lewandowski respectively.

TA-IV Cross-section dependence

Table TA4: Cross-Section Correlation

	PANEL A: LEVELS			PANEL B: FIRST DIFFERENCES		
	y_{it}	$debt_{it}$	cap_{it}	Δy_{it}	$\Delta debt_{it}$	Δcap_{it}
avg ρ	0.39	0.36	0.38	avg ρ	0.07	0.07
avg $ \rho $	0.66	0.52	0.75	avg $ \rho $	0.19	0.18
CD	157.62	148.36	154.54	CD	29.93	28.03
p -value	0.00	0.00	0.00	p -value	0.00	0.00

	PANEL C: HETEROG. AR(2)			PANEL D: HETEROG. AR(2) CCE		
	y_{it}	$debt_{it}$	cap_{it}	y_{it}	$debt_{it}$	cap_{it}
avg ρ	0.09	0.13	0.08	avg ρ	0.00	0.00
avg $ \rho $	0.20	0.20	0.30	avg $ \rho $	0.18	0.18
CD	35.82	50.93	31.80	CD	1.67	0.91
p -value	0.00	0.00	0.00	p -value	0.10	0.36

Notes: We present the average and average absolute correlation coefficients across the $N(N - 1)$ sets of correlations. CD reports the Pesaran (2004) cross-section dependence statistic, which is distributed $N(0, 1)$ under the null of cross-section independence. Panels A and B test the variable series in levels and first differences respectively. In Panel C each of the three variables in levels is entered into a time-series regression

$z_{it} = \pi_{0,i} + \pi_{1,i}z_{i,t-1} + \pi_{2,i}z_{i,t-2} + \pi_{3,i}t + \varepsilon_{it}$, conducted separately for each country i . In Panel D the country-regressions are augmented with cross-section averages of all variables (in the Pesaran (2006) CCE fashion) instead of a linear trend. The correlations and cross-section dependence statistic in Panels C and D are then based on the residuals from these AR(2) regressions. We used the Stata routine `xtcd` written by Markus Eberhardt. The contrast between results in Panel C and D show the power of the simple cross-section average approach in addressing residual cross-section dependence.

TA-V Weak Exogeneity Testing in the Panel

We adopt a simplified empirical setup of the common factor model with a single covariate x and single factors f and g . Let

$$y_{it} = \beta_i x_{it} + u_{it} \quad u_{it} = \alpha_i + \lambda_i f_t + \varepsilon_{it} \quad (12)$$

$$x_{it} = \varrho_i f_t + \pi_i g_t + \psi_i \varepsilon_{it} + \phi_i + e_{it} \quad (13)$$

Then provided there exists a cointegrating relationship between variables the Granger Representation Theorem (Engle & Granger, 1987) states that these series can be represented in the form of a dynamic ECM. Generically, for the pair

of cointegrated variables x and y we can write

$$\Delta y_{it} = c_{1i} + \lambda_{1i} \hat{e}_{i,t-1} + \sum_{j=1}^K \psi_{11ij} \Delta y_{i,t-j} + \sum_{j=1}^K \psi_{12ij} \Delta x_{i,t-j} + \varepsilon_{1it} \quad (14)$$

$$\Delta x_{it} = c_{2i} + \lambda_{2i} \hat{e}_{i,t-1} + \sum_{j=1}^K \psi_{21ij} \Delta y_{i,t-j} + \sum_{j=1}^K \psi_{22ij} \Delta x_{i,t-j} + \varepsilon_{2it} \quad (15)$$

where $\hat{e}_{i,t-1}$ represents the ‘disequilibrium term’ $\hat{e} = y - \hat{\beta}_i x - \hat{d}$ constructed using the estimated cointegrating relationship between these two variables (d represents deterministic terms). Equations (14) and (15) further include lagged differences of the variables in the cointegrating relationship. In the above example there are only two equations, since we have two variables in the cointegrating relationship. The Granger Representation Theorem implies that for a long-run equilibrium relationship to exist between y and x at least one of λ_{1i} and λ_{2i} must be non-zero: if (and only if) $\lambda_{1i} \neq 0$ then x has a causal impact on y , if (and only if) $\lambda_{2i} \neq 0$ then the causal impact is reversed. If both λ_{1i} and λ_{2i} are non-zero they determine each other jointly.

TA-VI Asymmetric Dynamic Regressions

Table TA5: Asymmetric Dynamic Models

	[1]	[2]	[3]	[4]	[5]	[6]
	MG	CMG	CMG	CMG	CMG	CMG
country trends	×		×			
asymmetry	LR, SR	LR, SR	LR, SR	LR	LR, SR	LR, SR
lagged CA					1 lag	2 lags
52% threshold						
ALR debt >52% GDP	-0.038 [0.024]	-0.004 [0.024]	-0.049 [0.021]**	-0.014 [0.024]	0.008 [0.026]	0.028 [0.027]
ALR debt <52% GDP	0.002 [0.023]	-0.009 [0.022]	-0.015 [0.022]	0.014 [0.020]	0.012 [0.027]	0.020 [0.027]
$y_{i,t-1}$	-0.566 [0.036]***	-0.711 [0.045]***	-0.767 [0.043]***	-0.692 [0.043]***	-0.732 [0.053]***	-0.754 [0.054]***
t -statistic ^b	-15.56	-15.64	-17.71	-16.10	-13.81	-14.00
\bar{t} -statistic	-3.71	-3.95	-4.19	-4.16	-3.71	-3.84
RMSE	0.029	0.022	0.020	0.023	0.018	0.017
CD Test	10.97	3.95	3.85	3.45	5.03	5.88
Obs (N)	1,873 (55)	1,873 (55)	1,873 (55)	1,873 (55)	1,804 (55)	1,768 (55)
75% threshold						
ALR debt >75% GDP	-0.053 [0.041]	-0.032 [0.043]	-0.040 [0.039]	-0.023 [0.036]	0.013 [0.034]	0.013 [0.032]
ALR debt <75% GDP	-0.046 [0.028]	0.002 [0.021]	-0.014 [0.019]	-0.003 [0.021]	0.018 [0.027]	-0.004 [0.029]
$y_{i,t-1}$	-0.547 [0.049]***	-0.726 [0.055]***	-0.790 [0.050]***	-0.673 [0.052]***	-0.781 [0.061]***	-0.771 [0.056]***
t -statistic ^b	-11.27	-13.33	-15.83	-12.84	-12.91	-13.87
\bar{t} -statistic	-3.56	-3.81	-4.06	-3.71	-3.77	-6.36
RMSE	0.031	0.023	0.022	0.025	0.019	0.018
CD Test	3.46	1.78	-0.32	1.72	1.36	0.97
Obs (N)	1,509 (45)	1,509 (45)	1,509 (45)	1,509 (45)	1,434 (44)	1,402 (44)
90% threshold						
ALR debt >90% GDP	-0.001 [0.054]	0.003 [0.030]	-0.021 [0.033]	-0.010 [0.033]	0.069 [0.054]	0.037 [0.055]
ALR debt <90% GDP	-0.005 [0.034]	0.054 [0.060]	0.001 [0.042]	0.084 [0.063]	0.049 [0.055]	0.120 [0.051]**
$y_{i,t-1}$	-0.549 [0.054]***	-0.611 [0.065]***	-0.703 [0.057]***	-0.602 [0.062]***	-0.698 [0.074]***	-0.774 [0.076]***
t -statistic ^b	-10.09	-9.35	-12.28	-9.66	-9.47	-10.17
\bar{t} -statistic	-3.37	-3.32	-3.65	-3.49	-3.48	-3.17
RMSE	0.034	0.025	0.023	0.027	0.021	0.020
CD Test	0.78	0.42	-0.58	-0.25	-1.35	-1.06
Obs (N)	996 (30)	996 (30)	996 (30)	996 (30)	940 (29)	921 (29)

Notes: We present average long-run coefficients (based on country-specific long-run results) for debt from models which allow for asymmetry in the debt coefficients, adapted to the panel from the time-series approach by Shin, Yu & Greenwood-Nimmo (2013). The dependent variable is the GDP per capita growth rate. Three thresholds are adopted to split the data into two (high/low debt) ‘regimes’: 52% (sample median), 75% and 90% debt/GDP ratio. Countries are only included in the analysis if they have at least 20% of their observations in one of the two regimes (below/above threshold), resulting in 55, 45 and 30 countries, respectively. Models [2]-[6] add cross-section averages to the regressions, those in [5] and [6] further add lags of the cross-section averages in the spirit of Chudik & Pesaran (2013). All models allow for long-run (LR) and short-run (SR) asymmetry, with the exception of model [4], which only allows for long-run asymmetry. ^b For the coefficient on lagged GDP per capita we report the Pesaran & Smith (1995) nonparametric t -statistics as well as the average of country-specific t -statistics (\bar{t}), the CD Test is distributed $N(0, 1)$ under the null of cross-section independence. RMSE is the root-mean squared error.

TA-VII Summability, Balance and Co-Summability

TA-VII.1 Theory

In the following we discuss the fundamental difficulties arising for conventional empirical analysis when assuming a non-linear model in the presence of integrated variables and introduce a novel time series approach to deal with these issues. Suppose a single time series relationship $y_t = f(x_t, \theta) + u_t$ for a nonstationary covariate $x_t \sim I(1)$, stationary u_t and some non-linear function $f(\cdot)$.¹ In this context, it becomes difficult to apply our standard notion of integration to $f(\cdot)$, given that integration is a linear concept: although we may be able to determine the order of integration of x_t , the order of integration of $f(x_t, \theta)$ (and thus y_t) may not be well defined for many non-linear transformations $f(\cdot)$. Assuming for illustration $f(x_t) = \theta x_t^2$ we can make this point somewhat clearer: let $x_t = x_{t-1} + \varepsilon_t$ and $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$, then we know that

$$\mathbb{V}[x_t - x_{t-1}] = \sigma_\varepsilon^2 \Rightarrow x_t \sim I(1) \quad (16)$$

In words, we can show that the Engle & Granger (1987) characterisation of a stationary process holds for Δx_t (finite variance is one of five characteristics, albeit the crucial one for our illustration), such that x_t can be concluded to follow an $I(1)$ process. Now investigate the same property for Δx_t^2 :

$$\mathbb{V}[x_t^2 - x_{t-1}^2] = \mathbb{E}[\varepsilon_t^4] + 4(t-1)\sigma_\varepsilon^4 - \sigma_\varepsilon^4 \Rightarrow x_t^2 \sim I(?)$$

We can see that the finite variance characteristic is violated, given that it is a function of time t – further differencing does not change this outcome. Although we can define x_t within the integration framework, we cannot state the order of integration of x_t^2 , which creates fundamental problems if the empirical analysis of $y_t = f(x_t, \theta) + u_t$ is to be based on arguments of cointegration.

Berenguer-Rico & Gonzalo (2013b) develop an alternative approach, based on the ‘order of summability’ $S(\delta)$ of linear or non-linear processes: “[t]he order of summability, δ , gives a summary measure of the stochastic properties – such as persistence – of the time series without relying on linear structures” (p.3). Using OLS we estimate for each country i

$$Y_{ik}^* = \beta_i^* \log k + U_{ik}^* \quad (17)$$

where $k = 1, \dots, T$, $Y_{ik}^* = Y_{ik} - Y_{i1}$, $U_{ik}^* = U_{ik} - U_{i1}$ and $Y_{ik} = \log \left(\sum_{t=1}^k (y_{it} - m_t) \right)^2$, with m_t the country-specific partial mean of y_{it} , namely $m_t = (1/t) \sum_{j=1}^t y_j$. This is the definition for m_t in the ‘intercept only’ case. Given the trending nature of our data we further investigate the ‘constant and linear trend’ case, where $m_t = (1/t) \sum_{j=1}^t y_{ij} - (2/t) \sum_{j=1}^t (y_{ij} - (1/j) \sum_{\ell=1}^j y_{i\ell})$. This implies

$$\hat{\beta}_i^* = \frac{\sum_{k=1}^T Y_{ik}^* \log k}{\sum_{k=1}^T \log^2 k} \quad (18)$$

¹Our discussion in this section as well as the implementation follow Berenguer-Rico & Gonzalo (2013a) and Berenguer-Rico & Gonzalo (2013b).

from which we then obtain our estimate of the order of summability $\hat{\delta}_i^* = (\hat{\beta}_i^* - 1)/2$. This approach essentially investigates the rate of convergence of a rescaled sum constructed from the variable series y_{it} . In the single time series inference can be established using confidence intervals constructed via estimation in subsamples; here, in the panel, where there is no natural ordering of countries in the cross-section dimension we take random draws of \sqrt{N} countries (and in each country the full time series T), each time capturing the mean and median summability statistic, to create subsample estimates for inference.

It bears emphasising that summability is a more general concept than integration, but that that latter is closely related to the former in the following fashion: if a time series x_t is integrated of order d , $I(d)$ with $d \geq 0$, then it is also summable of order d , $S(d)$. It is the breakdown of the reverse of this condition in cases where x_t is a nonlinear transformation which necessitates our adoption of the concept of summability. In our empirical application we will analyse the order of summability of all variables entering the polynomial specifications.

Next, in analogy to the analysis of integrated variables, the ‘balance’ of the empirical relationship needs to be tested, namely the condition that both sides of the empirical equation of interest have the same order of summability: $S(\delta_y) = S(\delta_z)$ for $z = f(x_t, \theta) = \theta f(x_t)$ – see below for a comment on the linearity in parameters we assume here. Such a test of balance is equivalent to testing the null of $\beta_{ni} \equiv (\beta_{yi} - \beta_{zi}) = 0$ in the country-specific regression

$$Y_{yik}^* - Y_{zik}^* = \underbrace{(\beta_{yi} - \beta_{zi})}_{\beta_{ni}} \log k + (U_{yik} - U_{zik}) \quad (19)$$

where Y_{yik}^* is for the LHS variable y and defined as in the summability analysis above, and Y_{zik}^* is the partially demeaned sum of all RHS processes $Y_{zik} = \log \left(\sum_{t=1}^k (z_{it} - m_t) \right)^2$, accounting for initial conditions in the same fashion as above by taking the deviation from the first observation. In practice, all elements of z (RHS variables) are summed, appropriately partially demeaned and their estimated order of summability is subtracted from that for y and the result divided by 2.² Again inference in the single time series test is based on subsample estimation. In the panel we employ the same strategy to create subsample estimates and thus confidence bands as detailed above. Under the null of balance the resulting confidence interval includes zero and balancedness is a necessary but not sufficient condition for a valid empirical specification.³

Finally, let \hat{e}_t be the OLS residuals from a balanced country-specific regression $y_{it} = \hat{\theta} g(x_{it}) + \hat{e}_{it}$, then ‘strong co-summability’ will imply the order of summability of \hat{e}_{it} , $S(\delta_{\hat{e}_{it}})$, is statistically close to zero. We employ the above approach to estimate the order of summability for \hat{e}_{it} which enables us to determine whether our balanced model is co-summable or not. Note that the residual series \hat{e}_{it} as defined above will sum to zero by default of the least squares principle, we therefore in

²Given $\hat{\delta}_y^* = (\hat{\beta}_y^* - 1)/2$ and $\hat{\delta}_z^* = (\hat{\beta}_z^* - 1)/2$ it is easy to see that $\hat{\delta}_y^* - \hat{\delta}_z^* = (\hat{\beta}_y^* - \hat{\beta}_z^*)/2$.

³Again the parallels with the theory of integration and cointegration may help to illustrate this point: the seminal Granger & Newbold (1974) paper investigated spurious regression by regressing two independent random walks, Y_t and X_t . Since both processes are $I(1)$ the regression equation $Y_t = \beta_0 + \beta_1 X_t$ is balanced. Since they constitute independent processes, later work by Engle & Granger (1987) would suggest that the residual series from this regression are not $I(0)$, so that Y_t and X_t are not cointegrated. Similarly, balancedness of $y_t = f(x_t, \theta)$ is a necessary prerequisite for co-summability between y_t and $f(x_t, \theta)$.

practice do not subtract the estimate for the intercept term in each country regression. Inference in the original time series and in our panel application follow the same principles as the previous two testing procedures.

The above routines imply a sequence of tests (summability, balance, co-summability) which in principle bear close resemblance to the integration-cointegration concepts and testing procedures. The simplicity of the above approach is marred by the presence of deterministic components in the variable evolution. Intercept and trend terms are addressed by repeated partial demeaning of the variable series as suggested in Berenguer-Rico & Gonzalo (2013b).⁴ We assume non-linearity in variables but not in parameters:

$$y_t = g(x_t, \theta) + \varepsilon_t = \theta g(x_t) + \varepsilon_t \quad (20)$$

The econometric theory of the approach is at present being extended to nonlinearity in parameters. However, the restriction to linearity in parameters is in line with the standard implementations in the literature adopting debt thresholds (endogenous or endogenous debt/GDP threshold with subsequent analysis splitting observations into separate below/above threshold values/terms) or nonlinearities through polynomial functions (linear, squared and cubed debt terms).

We provide an extension to the above panel versions of the balance and co-summability tests, whereby in the spirit of the recent panel time series literature we include the cross-section averages (CA) of all variables in the specification of the empirical test (Pesaran, 2006; Chudik & Pesaran, 2013). The motivation for this approach is the same we provided for our panel models above: country-by-country investigation of the variable and specification properties assumes these to be cross-sectionally independent. Both theorising and empirical practice have shown that in a globalising world where countries trade and are subject to similar social, economic and/or cultural heritage this assumption is likely to be violated.

We adopt two variants of the cross-section average augmentation: (i) a standard approach such as that outlined above, (ii) an approach where in addition to the CA of all model variables we also include the CA of ‘other covariates,’ similar to the approach in the dynamic heterogeneous panel estimations (Chudik & Pesaran, 2013).

TA-VII.2 Test results

Table TA6 presents the summability results, with models assuming a constant term in the left panel and constant and trend terms in the right panel, with the latter a more natural choice given the trending nature of our data. It appears that all of the variables investigated reject summability of order 0, $S(0)$, which justifies our concern about time series properties – recall the analogy with unit root tests, whereby integrated data of order 1 or higher provides evidence for nonstationarity. In the lower panel we carry out summability testing for the growth rates of per capita GDP, debt stock and capital stock. For the former two we can broadly conclude that

⁴We do not pursue the analysis of a quadratic trend here due to the limited time-series dimension of our panel.

these first difference series are $S(0)$, while the capital stock growth rate appears to reject this null hypothesis.

Table TA7 presents the results from balance tests, with (unaugmented) ‘standard’ specifications in Panel A, specifications augmented in the common correlated effect fashion in Panel B and specifications which further add cross-section averages from two ‘openness’ variables in Panel C. Recall that for the two sides of the equation to be balanced, i.e. be made up of variables with the same order of summability, the balance statistic should be close to zero. We highlight all those specifications where this requirement is statistically rejected by underlining the estimate and 95% confidence bands. In each of the three panels we provide results for a specification with a constant and a specification with a constant and trend term, where again the latter appears *a priori* the more suitable choice. Across all three panels there is relatively strong evidence for the linear specification to represent a balanced model, with mean and median estimates for the balance statistics close to zero. There is comparatively less evidence for the two nonlinear specifications, with only the median estimates and 95% confidence intervals for the Model with linear, squared and cubed debt terms containing zero. Having said that, the rejection of the null of equal order of summability on both sides of the model equation is marginal in the specification with linear and squared debt terms of both Panels B and C.

In the co-summability results presented in Table TA8 we highlight those specifications for which balance was somewhat uncertain by printing them in grey, whereas the specifications which were confirmed as balanced are printed in black. We again have three blocks of results, for a standard panel version of co-summability (equivalent to Panel A in the balance results in Table TA7), for a version which includes cross-section averages of all model variables (Panel B) and for a version which in addition to these cross-section averages includes those from ‘other covariates’ (Panel C). None of the specifications *without* cross-section averages is co-summable, and the estimated test statistics – summability statistics for model residuals – are some distance away from zero (which would signify co-summability). Results for the specifications with cross-section averages are noticeably closer to zero, but still reject co-summability in the linear model. Results for the final set of specifications which include further cross-section averages in the empirical model then move even closer to zero, with the linear specification now co-summable if we focus on the median statistic. The nonlinear models in this case also appear to be co-summable, however it bears reminding that there was comparatively less evidence for these models to be balanced, which as a prerequisite for co-summability renders these models at best as uncertain with regard to the presence or absence of a long-run equilibrium relationship.

We draw three conclusions from this analysis: first, there is strong evidence for significant persistence in the data investigated, which as argued above may seriously impact estimation and inference. Second, it appears that results from an approach which assumes cross-section independence yields very different results from one which relaxes this assumption. In the context of the recent panel econometric literature this finding is not at all surprising, given the importance of accounting for cross-section correlation in the analysis of macro panel datasets. Further investigation of this result is beyond the scope of this article and left for future research. Third, the only empirical model tested for which we found fairly convincing evidence of it representing a balanced and co-summable specification is the linear model augmented with standard and additional cross-section averages. There is less convincing evidence for nonlinear models, even

though some only fail the balance tests marginally. It bears reminding that the purpose of this exercise was to identify linear or nonlinear specifications which represent long-run equilibrium relationships.⁵ Whatever the identification strategy of existing studies in the literature, these results suggest that the adoption of linear and squared debt terms in a flexible specification to model debt thresholds may represent a seriously misspecified empirical model which could lead to spurious regression results.

Table TA6: Estimated Order of Summability

	Deterministics: Constant					Deterministics: Constant and Trend				
	y_{it}	$debt_{it}$	$debt_{it}^2$	$debt_{it}^3$	cap_{it}	y_{it}	$debt_{it}$	$debt_{it}^2$	$debt_{it}^3$	cap_{it}
Lower CI band	0.948	1.011	1.027	1.045	1.217	0.392	0.638	0.612	0.468	0.745
Mean	1.096	1.168	1.180	1.205	1.357	0.786	0.870	0.830	0.773	1.078
Upper CI band	1.243	1.325	1.334	1.364	1.497	1.180	1.101	1.048	1.077	1.411
Lower CI band	0.960	0.999	1.026	1.034	1.084	0.414	0.626	0.614	0.468	0.719
Median	1.109	1.135	1.156	1.174	1.286	0.768	0.823	0.819	0.783	1.048
Upper CI band	1.257	1.272	1.286	1.315	1.487	1.123	1.020	1.025	1.098	1.376

	Deterministics: Constant			Deterministics: Constant and Trend		
	Δy_{it}	$\Delta debt_{it}$	Δcap_{it}	Δy_{it}	$\Delta debt_{it}$	Δcap_{it}
Lower CI band	-0.048	0.090	0.094	-0.725	-0.537	0.202
Mean	0.062	0.221	0.248	-0.323	-0.173	0.391
Upper CI band	0.171	0.351	0.402	0.079	0.190	0.580
Lower CI band	-0.146	0.017	0.066	-0.751	-0.630	0.083
Median	0.000	0.182	0.213	-0.354	-0.205	0.284
Upper CI band	0.146	0.347	0.359	0.043	0.220	0.484

Notes: The table presents the panel statistics for $N = 105$ country-specific estimates of the order of summability $\hat{\delta}^*$ (see main text for further details). All variables are in logarithms. We account for the constant term by partial demeaning, and for the additional linear trend term by double partial demeaning as detailed in Berenguer-Rico & Gonzalo (2013b). For each variable we present two sets of statistics: the upper (lower) panel presents mean (median) $\hat{\delta}^*$ across the panel as well as the mean- (median-)based subsampling results (lower and upper 95% confidence bands). Each of the $N - b + 1 = 95$ subsamples of size $b = \text{int}(\sqrt{N}) + 1 = 11$ countries is a random draw of countries from our full sample of $N = 105$.

⁵Eberhardt (2013) applies these methods to the long time series data provided by Reinhart & Rogoff (2009) and finds no evidence for a non-linear long-run relationship between debt and growth in the data for the US, Great Britain, Sweden and Japan as well as 23 other (primarily High-Income) countries.

Table TA7: Estimated Balance

Panel A – Standard Specification							
Deterministics	Constant			Constant & Trend			
	[1]	[2]	[3]	[4]	[5]	[6]	
Nonlinearity	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	
Lower CI band	-0.316	-0.274	-0.278	Lower CI band	-0.173	<u>0.039</u>	<u>0.078</u>
Mean	-0.147	-0.097	-0.106	Mean	0.082	<u>0.386</u>	<u>0.493</u>
Upper CI band	0.023	0.079	0.066	Upper CI band	0.338	<u>0.733</u>	<u>0.908</u>
Lower CI band	<u>-0.348</u>	-0.295	-0.357	Lower CI band	-0.134	<u>0.003</u>	<u>0.014</u>
Median	<u>-0.191</u>	-0.097	-0.144	Median	0.111	<u>0.345</u>	<u>0.453</u>
Upper CI band	<u>-0.033</u>	0.100	0.069	Upper CI band	0.357	<u>0.688</u>	<u>0.893</u>

Panel B – Specification with CA							
Deterministics	Constant			Constant & Trend			
	[1]	[2]	[3]	[4]	[5]	[6]	
Nonlinearity	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	
Lower CI band	<u>-0.265</u>	-0.260	<u>-0.281</u>	Lower CI band	-0.218	<u>0.002</u>	<u>0.023</u>
Mean	<u>-0.147</u>	<u>-0.136</u>	<u>-0.149</u>	Mean	0.067	<u>0.404</u>	<u>0.470</u>
Upper CI band	<u>-0.029</u>	<u>-0.012</u>	<u>-0.017</u>	Upper CI band	0.351	<u>0.805</u>	<u>0.918</u>
Lower CI band	<u>-0.301</u>	<u>-0.312</u>	<u>-0.342</u>	Lower CI band	-0.141	<u>0.030</u>	-0.009
Median	<u>-0.166</u>	<u>-0.179</u>	<u>-0.216</u>	Median	0.112	<u>0.397</u>	0.449
Upper CI band	<u>-0.031</u>	<u>-0.045</u>	<u>-0.089</u>	Upper CI band	0.364	<u>0.764</u>	0.908

Panel C – Specification with Additional CA							
Deterministics	Constant			Constant & Trend			
	[1]	[2]	[3]	[4]	[5]	[6]	
Nonlinearity	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	-	$debt_{it}^2$	$debt_{it}^2, debt_{it}^3$	
Lower CI band	<u>-0.354</u>	<u>-0.264</u>	<u>-0.281</u>	Lower CI band	-0.220	<u>0.001</u>	<u>0.022</u>
Mean	<u>-0.232</u>	-0.143	<u>-0.149</u>	Mean	0.048	<u>0.400</u>	<u>0.469</u>
Upper CI band	<u>-0.110</u>	<u>-0.021</u>	<u>-0.017</u>	Upper CI band	0.316	<u>0.799</u>	<u>0.917</u>
Lower CI band	<u>-0.317</u>	<u>-0.317</u>	<u>-0.343</u>	Lower CI band	-0.167	<u>0.004</u>	-0.008
Median	<u>-0.188</u>	<u>-0.183</u>	<u>-0.216</u>	Median	0.083	<u>0.391</u>	0.450
Upper CI band	<u>-0.060</u>	<u>-0.050</u>	<u>-0.089</u>	Upper CI band	0.333	<u>0.777</u>	0.907

Notes: The table presents distributional statistics for $N = 105$ country-specific estimates of the balance in the indicated regression models ($\hat{\delta}_y - \hat{\delta}_g$) (see main text for further details). The RHS of each model always includes cap_{it} and $debt_{it}$. All variables are in logarithms. We account for a constant term by partial demeaning, and for an additional linear trend term by repeated partial demeaning as detailed in Berenguer-Rico & Gonzalo (2013b). CA refers to the augmentation of the static country regression with cross-section averages following Pesaran (2006): in Panel B we include the model variables, in Panel C we include CA of the openness and financial development variables in addition to the model variables. Underlined mean or median balance statistics indicate evidence *against* the hypothesis of a balanced regression model, $(\hat{\delta}_y - \hat{\delta}_g) = 0$.

Table TA8: Co-Summability

	Standard Specification			Specification with CA			
	[1]	[2]	[3]	[4]	[5]	[6]	
Nonlinearity	-	debt_{it}^2	$\text{debt}_{it}^2, \text{debt}_{it}^3$	-	debt_{it}^2	$\text{debt}_{it}^2, \text{debt}_{it}^3$	
CA	-	-	-	×	×	×	
Lower CI band	0.792	0.775	0.979	Lower CI band	0.109	0.030	-0.012
Mean	<u>0.929</u>	<u>0.907</u>	<u>1.188</u>	Mean	<u>0.270</u>	<u>0.210</u>	0.222
Upper CI band	1.065	1.038	1.396	Upper CI band	0.431	0.390	0.456
Lower CI band	0.670	0.789	0.951	Lower CI band	0.093	-0.118	-0.264
Median	<u>0.873</u>	<u>0.952</u>	<u>1.130</u>	Median	<u>0.277</u>	0.095	-0.032
Upper CI band	1.075	1.114	1.309	Upper CI band	0.461	0.309	0.200
	Specification with Additional CA						
	[7]	[8]	[9]				
Nonlinearity	-	debt_{it}^2	$\text{debt}_{it}^2, \text{debt}_{it}^3$				
CA	×	×	×				
Lower CI band	0.052	-0.051	-0.089				
Mean	<u>0.206</u>	0.136	0.143				
Upper CI band	0.360	0.322	0.374				
Lower CI band	-0.016	-0.244	-0.314				
Median	0.160	-0.062	-0.127				
Upper CI band	0.336	0.120	0.060				

Notes: The table presents distributional statistics for $N = 105$ country-specific order of summability estimates for the respective model residuals. The RHS of each model always includes cap_{it} and debt_{it} . All variables are in logarithms. CA refers to the augmentation of the static country regression with cross-section averages following Pesaran (2006), 'Additional CA' refer to the CA for the openness and financial development variables (in logs) as described in the main text. We account for a constant term by partial demeaning as detailed in Berenguer-Rico & Gonzalo (2013b). Underlined mean or median co-summability statistics indicate evidence against the hypothesis of a co-summable model specification, $\hat{\delta}_\varepsilon = 0$. Since co-summability is conditional on balance, we only print those specifications in black for which we have convincing evidence from the balance testing in Table TA7, whereas all other specifications are printed in grey.