



## Eberhardt, Markus (2017) Nonlinearities in the relationship between debt and growth: (no) evidence from over two centuries. *Macroeconomic Dynamics* . ISSN 1469-8056 (In Press)

### Access from the University of Nottingham repository:

<http://eprints.nottingham.ac.uk/45965/1/2017-04-05%20Debt-Growth-CoSummability%20journal.pdf>

### Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see: [http://eprints.nottingham.ac.uk/end\\_user\\_agreement.pdf](http://eprints.nottingham.ac.uk/end_user_agreement.pdf)

### A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

# **Nonlinearities in the relationship between debt and growth: (no) evidence from over two centuries**

Markus Eberhardt

Affiliation:

*School of Economics, Sir Clive Granger Building, University of Nottingham,*

*University Park, Nottingham NG7 2RD, United Kingdom*

Correspondence:

*email: markus.eberhardt@nottingham.ac.uk;*

*fax: +44 115 95 14159; phone: +44 115 846 8416.*

Running title:

Nonlinearities in the Debt-Growth Nexus

Corresponding author (for proofs):

Markus Eberhardt School of Economics, Sir Clive Granger Building, University of Nottingham,

University Park, Nottingham NG7 2RD, United Kingdom

email: [markus.eberhardt@nottingham.ac.uk](mailto:markus.eberhardt@nottingham.ac.uk); fax: +44 115 95 14159; phone: +44 115 846 8416.

**Abstract:** I revisit the popular concern over a nonlinearity or threshold in the relationship between public debt and growth employing long time series data from up to 27 countries. My empirical approach recognises that standard time series arguments for long-run equilibrium relations between integrated variables (cointegration) break down in nonlinear specifications such as those predominantly applied in the existing debt-growth literature. Adopting the novel co-summability approach my analysis overcomes these difficulties to find no evidence for a systematic long-run relationship between debt and growth in the bivariate and economic theory-based multivariate specifications popular in this literature.

**Keywords:** public debt; economic growth; nonlinearity; summability and co-summability

**JEL classification:** H63, C22, E62, O40

# 1 Introduction

*“The latest research [by Reinhart and Rogoff] suggests that once debt reaches more than about 90% of GDP the risks of a large negative impact on long term growth become highly significant.”*

George Osborne, Mais Lecture, February 24, 2010

*“The study [Reinhart and Rogoff (2010b)] found conclusive empirical evidence that total debt exceeding 90 percent of the economy has a significant negative effect on economic growth.”*

‘The Path to Prosperity,’ House Committee on the Budget, April 5, 2011

Despite the rhetoric adopted by a number of governments and opposition parties over recent years, determining a causal link from public debt to long-run growth as well as the potential nonlinearity of this relationship are widely regarded as unresolved empirical issues (IMF, 2012; Panizza and Presbitero, 2014). As above quotes indicate the most influential research on the debt-growth nexus in recent years is unarguably the work by Reinhart and Rogoff (2010b) which has been adopted as justification for fiscal austerity measures by politicians on both sides of the Atlantic. Although recent revelations challenged the descriptive analysis carried out in their paper, Reinhart and Rogoff

maintain that “the weight of the evidence to date — including this latest comment [by Herndon, Ash and Pollin, 2014] — seems entirely consistent with our original interpretation of the data” (*Wall Street Journal* ‘Real Time Economics’ blog, April 16, 2013), namely that “high debt/GDP levels (90 percent and above) are associated with notably lower growth outcomes” (Reinhart and Rogoff, 2010b: 577; see also Rogoff, 2013). Perhaps aware of the tension between the causal interpretation typically read into this type of statement and the descriptive nature of their analysis, some of their earlier work (Reinhart, Reinhart and Rogoff, 2012) already pointed to a set of empirical studies which are argued to address both concerns regarding causality and identification of a nonlinearity in the *long-run* debt-growth relationship (e.g. Kumar and Woo, 2010; Balasone, Francese and Pace, 2011; Cecchetti, Mohanty and Zampolli, 2011; Checherita-Westphal and Rother, 2012)<sup>1</sup> in support of their findings.

This paper investigates the debt-growth nexus from a new angle and with a somewhat more modest aim, focusing on the persistence of the long time series data used in the original Reinhart and Rogoff (2010b) study. I adopt annual data for over two centuries (1800–2010) to investigate whether linear or various nonlinear specifications of the debt-growth nexus constitute ‘long-run equilibrium relations’ in four OECD countries: the United States, Great Britain, Japan and Sweden;<sup>2</sup> additional work presented in an Online Appendix extends the analysis to 27 advanced and developing economies. The analysis employs the most popular specifications in this empirical literature – polynomial functions and piecewise linear (threshold) specifications – to model the hypothesised nonlinearity. The basic premise of my analysis is that if variable series are integrated (nonstationary), then the popular implementations of nonlinearity in the debt-growth literature (squared debt terms or debt terms interacted with threshold dummies) are invalid, since these transformations of the variable are not defined within the (co-)integration framework. Therefore any empirical results building on these polynomial or threshold specifications may be spurious.

My empirical strategy addresses this problem by adopting novel methods for summability and co-summability testing (Berenguer-Rico and Gonzalo, 2014a,b). These concepts provide a framework encompassing integration and cointegration which however extends to non-linear relationships. The analysis in this study is thus (narrowly) focused on the question of potential nonlinearities

in the long-run debt-growth relationship, bypassing any concerns over the direction of causation *which does not impact the statistical validity of the results*. If there is no evidence for (nonlinear) long-run relations, then standard empirical specifications in the literature adopting thresholds or polynomial functions are misspecified and the causal interpretation assigned in these studies is questionable: the presence of a long-run equilibrium is a pre-requisite for the existence of any long-run causal relationship in the data. Results have important policy implications given that the most vocal supporters of fiscal austerity have pointed to the above-cited studies as providing empirically sound evidence for (this type of) nonlinearities in the debt-growth relationship.

The primary empirical focus is to investigate data for debt and GDP in Great Britain, Japan, Sweden and the United States over the 1800–2010 time horizon. In additional analysis I investigate sub-periods of 60 years using rolling window analysis to allow for structural breaks in the debt-growth relationship and also to reduce the impact of global shocks such as World War II on the presence or absence of a long-run relationship. A host of further robustness checks are confined to an Online Appendix.

My core analysis finds no evidence for any long-run relationship between debt and growth in the linear or nonlinear specifications for the four countries investigated. Subsample analysis does not



fundamentally challenge this finding although it provides an indication that there may have been long-run relationships between debt and growth at different points in time, although not in the post-WWII period typically studied in the existing literature. The general patterns revealed by the subsample analysis supports the notion that the debt-growth relationship differs across countries and “with economic circumstances” (Larry Summers, Witness Statement to the US Senate Budget Committee, June 4, 2013). Additional empirical analysis goes to great lengths to determine whether the choice of countries, time periods, and/or atheoretical specifications drive this finding but arrives at a fairly consistent picture across all different modes of investigation. These findings help challenge the apparent consensus in parts of the empirical literature of both the *existence* and the *common* nature of a debt threshold across countries.

The remainder of the paper is organised as follows: the next section discusses the existing literature on debt and growth, with Section 3 providing the theoretical background for my econometric approach. Section 4 introduces the data and describes the debt-growth nexus in each of the four OECD countries which are at the core of my analysis. Results for these and a larger set of countries are presented and discussed in Section 5, before Section 6 concludes.

## 2 Existing Literature

The existing empirical literature on the debt-growth nexus builds on somewhat ambiguous theoretical foundations (for a recent survey see Panizza and Presbitero, 2014). Some theoretical models argue that higher stocks of public debt may create increased uncertainty or even fear of future financial repression among investors and thus lead to a negative long-run relationship (Elmendorf and Mankiw, 1999; Teles and Mussolini, 2014) between debt and growth. Other work maintains that this negative relationship disappears once sticky wages and unemployment are taken into account in the modelling process (Greiner, 2011). The nonlinearity or debt threshold hypothesised and investigated in most empirical work can be motivated for developing countries by pointing to the issue of debt overhang (Krugman, 1998; Sachs, 1989), although it may be difficult to extend this argument to advanced economies such as those investigated in this paper. Nonlinearities may further arise if there is a tipping point of fiscal sustainability as is developed in Ghosh, Kim, Mendoza, Ostry and Qureshi (2013), however I am not aware of any theoretical models incorporating such debt tipping points into a framework for economic growth over the long-run.

As was suggested above, the work by Reinhart and Rogoff (2009; 2010a,b; 2011) is largely descriptive in nature, although this should not distract from the significant contribution these authors

have made to the literature in the construction of long data series for empirical analysis. Regression analysis of the debt-growth nexus conducted using panel data typically shares the unease about misspecification and endogeneity with the wider cross-country growth literature (for a discussion of the latter see Durlauf, Johnson and Temple, 2005; Eberhardt and Teal, 2011). Empirical specifications in this literature are across the board partial adjustment models in the mould of Barro (1991) and Mankiw, Romer and Weil (1992) – regressing growth on a lagged level of per capita GDP and a measure for debt stock as well as typically a large number of control variables – in a pooled model specification, thus assuming away the possibility of parameter heterogeneity across countries.<sup>3</sup> The standard practice in the cross-country literature to average data over three- or five-year intervals in the panel is also adopted in all but the most recent papers (Checherita-Westphal and Rother, 2012; Baum, et al, 2013; Panizza and Presbitero, 2014). Samples differ significantly across existing studies, with the work by Kumar and Woo (2010), Cecchetti et al (2011), Checherita-Westphal and Rother(2012), Baum, et al (2013), and Panizza and Presbitero (2014) primarily focused on OECD and other high-income economies and thus most relevant to this study. Among these OECD country studies the only one to adopt a polynomial specification is the paper by Checherita-Westphal and Rother (2012), although this practice is popular in the study of developing economies (e.g. Cordella, et al, 2010; Calderon and Fuentes, 2013; Pres-

bitero, 2012). With the exception of Cecchetti, et al (2011), who apply the within (fixed effects) estimator and thus cannot address concerns over reverse causality, all of the above empirical studies implement their panel analysis adopting the Blundell and Bond (1998) System GMM estimator originally developed for firm-level panel data analysis.<sup>4</sup>

Despite different sample periods, country coverage, control variables, modelling of the nonlinearity and choice of moment conditions for identification, these studies come to remarkably similar conclusions, namely that beyond a threshold at around 90% debt-to-GDP the relationship between debt and growth is negative significant. However, as demonstrated by Panizza and Presbitero (2013), these findings are either not robust to small changes in the sample, suggesting the results are driven by outliers, or fail to formally test the coefficients on the pairwise linear terms, which on closer inspection typically cannot support the notion of a statistically significant change in the debt coefficient above the threshold.

All of the above studies are focused on pooled panel data modelling, implicitly assuming that the long-run equilibrium relationship between debt and growth is the same for all countries in the sample. Existing research has found very different results when moving away from full sample analysis in homogeneous parameter regression models and toward sub-sample analysis along

geographic, institutional or income lines (IMF, 2012; Kourtellos, et al, 2012; Eberhardt and Presbitero, 2015). There are a number of reasons to assume the equilibrium relationship between debt and growth could differ across countries. Vulnerability to public debt depends not only on debt levels, but also on debt composition IDB (2007). Unfortunately, existing data for the analysis of debt and development often represent a mix of information relating to general and central government debt, debt in different currency denominations and with different terms attached (be they explicit or implicit). All of this implies that comparability of the debt data across countries may be compromised Panizza and Presbitero (2013). In addition, even assuming that debt stocks are comparable across countries and over time, the possible effect of public debt on GDP may depend on the reason *why* debt has been accumulated and on whether it has been consumed or invested (and in the latter case in which economic activities). Furthermore, different stocks of debt may impinge differently on economic growth: debt can clearly hinder GDP growth when it becomes unsustainable, affecting interest rates and triggering a financial crisis, thus affecting the level of GDP. However, 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 and Savastano, 2003; Kraay and Nehru, 2006; Manasse and Roubini, 2009). For these reasons the focus of analysis in this paper is on country-by-country investigation of the long-run relationship

between debt and growth.

A recent study which empirically investigates the debt-growth nexus with a time series econometric approach is the paper by Balassone, et al (2011) on Italy (1861-2009). Adopting unit root and cointegration testing prior to estimation they establish a long-run relationship between per capita GDP, per capita capital stock and debt-to-GDP ratio (all in logarithms). They then go on to estimate (among other models) a piecewise linear specification for the debt-to-GDP ratio where values beyond a threshold of 100% are found to create a significantly stronger negative effect on growth – it is precisely this form of interaction between a threshold dummy and the debt-to-GDP ratio which is not defined under (linear) cointegration and which necessitates the present analysis.<sup>5</sup> It should also be noted that cointegration does not imply causation from debt to growth.

### **3 Nonlinear Relations between Integrated Processes**

#### **3.1 Methodology**

In this section I highlight the difficulties arising for conventional time series analysis when assuming a non-linear model in the presence of integrated variables and discuss a novel approach to

tackle these issues.

Suppose a time series relationship  $y_t = f(x_t, \theta) + u_t$  for a nonstationary regressor  $x_t \sim I(1)$ ,

stationary  $u_t$  and some non-linear function  $f(\cdot)$ . Assuming for illustration  $f(x_t) = \theta_1 x_t + \theta_2 x_t^2$ ,

let  $x_t = x_{t-1} + x_0 + \varepsilon_t$  and  $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$ , then we know that

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

In words, it can be shown that the (Engle and Granger, 1987, henceforth EG) characterisation of a

stationary process holds for  $\Delta x_t$  – finite variance is one of five EG characteristics. Now investigate

the same property for  $\Delta x_t^2$ :

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

Here the finite variance characteristic is clearly violated, given that the variance is a function of

time. Since this problem cannot be solved by further differencing it is not possible to determine

the order of integration of  $x_t^2$ . This in turn creates fundamental problems if the empirical analysis

of  $y_t = \theta_1 x_t + \theta_2 x_t^2 + u_t$  is to be based on arguments of cointegration. The difficulty arises from the

requirement of the EG characterisation to investigate the *differences* of a process, with the intrinsic linearity of the difference operator creating obvious problems for nonlinear processes.

The following briefly introduces a novel set of methods for nonlinear processes which closely resemble the standard toolkit in linear time series analysis (tests for unit root behaviour and cointegration). The motivation for these new methods is to create “a summary measure of the stochastic properties – such as persistence – of the time series without relying on linear structures” (Berenguer-Rico and Gonzalo, 2014b). The implementation of these tests is straightforward, involving OLS regressions of transformed variable series, where transformations avoid the first differencing so central to Dickey-Fuller-type unit root analysis and instead build on running sums. Like in the case of unit root analysis the distributions of these test statistics are non-standard, but estimates for sub-samples can be used to create confidence intervals for inference.

Berenguer-Rico and Gonzalo (2014b) build on earlier work by Gonzalo and Pitarakis (2006) to develop a non-linear alternative to linear integration, based on the ‘order of summability.’<sup>6</sup> The empirical procedure to determine the order of summability analyses the rate of convergence of a rescaled sum  $Y_k^*$  of the variable of interest  $y_t$ . Using least squares we can estimate for  $k =$



$1, \dots, T$

$$Y_k^* = \beta^* \log k + U_k^* \quad (2)$$

where  $Y_k^* = Y_k - Y_1$ ,  $U_k^* = U_k - U_1$  and  $Y_k = \log \left( \sum_{t=1}^k (y_t - m_t) \right)^2$ .<sup>7</sup> This regression yields

$$\hat{\beta}^* = \frac{\sum_{k=1}^T Y_k^* \log k}{\sum_{k=1}^T \log^2 k} \quad (3)$$

from which the estimate of the order of summability  $\hat{\delta}^* = (\hat{\beta}^* - 1)/2$  is obtained. Inference can be established using confidence intervals constructed from subsample estimation (Politis, Romano and Wolf, 1999), whereby the above procedure is applied to  $T - b + 1$  subsamples of length  $b = \text{int}(\sqrt{T}) + 1$ .<sup>8</sup>

Summability is a more general concept than integration, but they are closely related: if a series  $x_t$  is integrated of order  $d$ ,  $I(d)$  for  $d \geq 0$ , then it is also summable of order  $d$ ,  $S(d)$ ; however, not all  $S(d)$  processes are also  $I(d)$ . Summability analysis thus provides important insights into the time series properties of a variable but in contrast to unit root analysis is not limited to linear processes. In the case of the debt-growth application I pursue here this allows me to investigate the time series properties of squared and cubed debt-to-GDP ratios as well as piecewise linear debt-to-GDP series.

In a second step, Berenguer-Rico and Gonzalo (2014a) offer a test to investigate the ‘balance’ of the empirical relationship, namely the condition that the two sides of the empirical equation have the same order of summability:  $S(\delta_y) = S(\delta_z)$  for  $z = f(x_t, \theta) = \theta f(x_t)$ . Again there is a close analogy with the linear unit root and cointegration case: before cointegration between two or more variables can be tested, it is necessary to establish that these variables possess the same order of integration. Regressing stationary on nonstationary variables – as would be the case if we regressed the per capita GDP *growth rate* on the debt-to-GDP ratio in *levels* – is referred to as an inconsistent regression which leads to invalid inference. However, in the present study I do not test for balance due to an unresolved problem with the testing procedure which invalidates the results.<sup>9</sup>

It should be noted that the main arguments put forth in this paper are based on the co-summability tests, which do not suffer the same problem.

Finally, the concept of co-summability is tested by investigating the error terms of a candidate specification. In empirical practice, let  $\hat{e}_t$  be the least squares residuals from a balanced regression  $y_t = \hat{\theta}g(x_t) + \hat{e}_t$ , then ‘strong co-summability’ will imply the order of summability of  $\hat{e}_t$  is statistically close to zero,  $S(0)$  (Berenguer-Rico and Gonzalo, 2014a). Note the analogy to a linear cointegrating relationship where the residuals from a linear regression between I(1) variables will be I(0). The order of summability for  $\hat{e}_t$  can be estimated to determine whether a candidate

model is co-integrable.<sup>10</sup> Inference follows the subsampling approach as in the previous testing procedures and under the null of co-integrability the confidence interval includes zero.

## 3.2 Specifications

I adopt two specifications for nonlinearity in the debt-to-GDP ratio in line with standard approaches in the literature: *first*, in addition to a standard linear model (Model 1) I use polynomial specifications including linear and squared (Model 2) or linear, squared and cubed (Model 3) debt-to-GDP terms (in logarithms) – examples for this specification include Calderon and Fuentes (2012) and Checherita-Westphal and Rother (2012). *Second*, I adopt piecewise linear specifications where the debt-to-GDP ratio (in levels, not logs) is divided into two variables made up of values below and above a specified threshold, which is treated as exogenous (examples for this specification include Kumar and Woo, 2010; Baum, et al, 2013; Panizza and Presbitero, 2014).<sup>11</sup> For Great Britain I adopt three threshold values: 90, 70 and 50 percent. For the United States and Japan I can only adopt the 50 percent threshold since even over the full time horizon too few observations are above the other two thresholds: only 12 (Japan: 22) for 70 percent and 6 (Japan: 17) for 90 percent. In Sweden the debt-to-GDP ratio only surpasses the 50 percent threshold in 15 sample years (7% of observations) so that I cannot investigate even a 50% threshold for this country. Note that all of the

empirical approaches in the debt-growth literature discussed above are based on models which are linear in parameters but non-linear in the variables – my implementation follows this assumption.

Although there are of course alternative transformations (e.g. ‘integrable functions’ proposed by Park and Phillips, 2001) to model the potential nonlinearity in the debt-growth relationship I restrict myself to the above polynomial and threshold models since these feature in the vast majority of empirical applications – see Panizza and Presbitero (2014) for a recent survey.

The co-summability analysis thus investigates a number of specifications for the debt-growth relationship, inspired by the simple Reinhart and Rogoff (2010b) setup. The polynomial specifications are:

$$y_t = \alpha_0 + \varphi t + \phi_1 x_t + \varepsilon_t \quad (4)$$

$$y_t = \alpha_0 + \varphi t + \phi_1 x_t + \phi_2 x_t^2 + \varepsilon_t \quad (5)$$

$$y_t = \alpha_0 + \varphi t + \phi_1 x_t + \phi_2 x_t^2 + \phi_3 x_t^3 + \varepsilon_t \quad (6)$$

where  $y$  is per capita GDP and  $x$  is the debt-to-GDP ratio (both in logarithms),  $\alpha_0$  is an intercept,

$t$  a linear trend term with parameter  $\varphi$  and  $\varepsilon_t$  is white noise.

The threshold model specifications are based on

$$y_t = \alpha_0 + \varphi t + \theta_1 X_t \times \mathbb{1}(X_t < \text{threshold}) \quad (7)$$
$$+ \theta_2 X_t \times \mathbb{1}(X_t \geq \text{threshold}) + \varepsilon_t$$

where  $\mathbb{1}(X_t < \text{threshold})$  is an indicator function which is 1 for the debt-to-GDP ratio  $X_t$  below the threshold and 0 otherwise – similarly for  $\mathbb{1}(X_t \geq \text{threshold})$  at and above the threshold.

I investigate the evidence for long-run equilibrium relationships between debt burden and per capita GDP *levels* – since the focus of the applied literature is on the *long-run* relationship I adopt the levels variable for income, rather than its growth rate. The popularity of the ‘growth’ specification in the cross-country empirical literature is justified by the presence of the lagged level of per capita GDP as additional regressor (as is the case for the ‘debt-growth’ analysis of Kumar and Woo, 2010; Cecchetti, et al, 2011; Checherita-Westphal and Rother, 2012; Baum, et al, 2013; among others). This quasi-error correction specification provides estimates for a long-run *levels* relationship although researchers frequently refer to this type of specification as a ‘growth’ equation (see Eberhardt and Teal, 2011).

In addition to the analysis for the full time horizon I investigate co-summability in the four OECD

countries using a window of sixty years, which is moved along the time horizon from the 1800s to 2010. The purpose of this exercise is to provide both an indication of possible changes in the long-run debt-growth relationship over time as well as to safeguard the analysis from undue impact of severe shocks such as the two world wars or changes in the definition or the debt variable.<sup>12</sup> Due to the nature of the data this approach is only feasible for the polynomial specifications: as highlighted by Chinn (2012) in his review of Reinhart and Rogoff (2011) there are comparatively few episodes in developed economies where the debt-to-GDP ratio exceeds 90% and I can therefore not implement the moving window for the piecewise linear specification. Since this rolling window analysis represents a form of data mining I adjust the confidence intervals (CI) for all estimates following a standard Bonferroni correction, whereby  $CI^* = (1 - \alpha/m)$  for the conventional confidence level  $1 - \alpha$  (I adopt  $\alpha = .05$ ) and the number of sub-samples tested  $m$  (varies from 80 for Japan to 152 for the US, Great Britain and Sweden). In practice this makes the confidence intervals *much wider*, thus representing a more conservative approach to rejecting the null hypothesis of co-summability. A number of additional robustness checks are carried out, for which the motivation, approach, and results are presented in an Online Appendix. The focus of these robustness checks is on (i) a diverse sample of 23 additional economies (including some developing countries); (ii) a reverse specification with debt-to-GDP ratio (in logs) as the dependent variable for models includ-

ing the (log of) per capita GDP, and its squared and cubed polynomial terms as regressors; (iii) economic theory-based specifications which add a number of determinants of growth as favoured by the cross-country growth literature to the model.

## 4 Data

I use annual per capita GDP (in 1990 Geary-Khamis \$) from an updated version (Bold and Van Zenden, 2013) of the series compiled by Maddison (2010). I match these data to information on the gross government debt-to-GDP ratio (in percent) from Reinhart and Rogoff (2009). The debt figures refer to total gross central government debt, comprising domestic and external debt (see Online Appendix for exceptions). Data coverage differs across countries: for the US, Britain and Sweden data series start in 1800, for Japan in 1872 – all series end in 2010.

Descriptive statistics for these four countries are presented in the Online Appendix, where I also plot the levels and first differences of the per capita GDP and debt-to-GDP ratio variables (in logs).

Although my summability analysis provides insights into the time series properties of these data I also carry out a number of unit root tests to illustrate the difference in order of integration between the per capita GDP growth rate and the debt-to-GDP ratio in levels which rules out the existence of

any long-run relationship (cointegration, co-summability) between these two variable series.

In the Online Appendix data from a further 23 countries using the same sources are employed to carry out summability and co-summability tests. Here countries were included in the sample provided their per capita GDP and debt-to-GDP ratio series extended back to 1900 or earlier.

Extended empirical models analysed in an Online Appendix incorporate inflation and schooling data primarily taken from the Clio Infra project at the International Institute of Social History, population data from the original Maddison (2010) dataset, investment and additional debt data from Maddison (1992), Mitchell (2007a,b), and the World Bank *World Development Indicators* as well as a number of other sources (for details see Online Appendix).

**[Figure 1 about here]**

Figure 1 charts the evolution of the debt-to-GDP ratio for the four economies, where in the spirit of Reinhart and Rogoff (2010) I highlight periods with debt burden in excess of 90% of GDP. While the four time series all display idiosyncracies, it is nevertheless notable how similar in particular the patterns for British and American debt-to-GDP ratios are over much of the 20th century, albeit with substantially higher debt in the former. Britain is also the only economy studied which experienced sustained periods of debt-to-GDP above 90%.



**[Figure 2 about here]**

In Figure 2 I plot the debt-income relationship in each of the four countries, taking variables in deviation from the country-specific time-series mean. In all four economies the most significant turning points for the debt-growth nexus were marked by the Great War of 1914-18, the Great Recession of the late 1920s and World War II.

## **5 Empirical Results**

### **5.1 Main results: Order of Summability**

Table 1 provides estimates of the order of summability for all model variables, including polynomial as well as threshold terms for debt. None of the confidence intervals for tests on per capita GDP *levels* or any of the debt variables include zero, thus rejecting the null of summability of order zero. The estimated order of summability for the per capita GDP *growth rates* in contrast is always very close to zero. For the linear terms of per capita GDP and the debt-to-GDP ratio (in logs or levels) and their growth rates these results are perfectly in line with unit root and stationarity test results presented in the Online Appendix, where I establish stationary growth rates and

nonstationary levels series (whether in logarithms or not).

These findings highlight the significant persistence in the data and provide a strong motivation for the concerns over time series properties I argue are of primary importance when analysing the long-run debt-growth nexus. In analogy to integrated data, we run the risk of spurious results in any regressions containing these variables unless we can confirm our empirical models as balanced and co-summable. Note that with the exception of the study by Balassone, et al (2011) on Italy none of the papers in this literature show concern for time series properties of the data.

**[Table 1 about here]**

## **5.2 Main results: Co-Summability**

Table 2 provides results from co-summability tests using per capita GDP *levels* as dependent variable. Co-summability is rejected in *all* countries and specifications – residuals from these models were not found to be summable of order zero,  $S(0)$ . Note that the rejection of co-summability is by no means marginal, with all confidence intervals some distance away from zero. The fact that subsampling confidence intervals are at times very wide is a further strong signal for misspecification.

These findings imply that from a long-run perspective per capita income and the debt-to-GDP ratio

do not move together, *precluding any causal relationship between these variables*.<sup>13</sup>

[Table 2 about here]

### **5.3 Results from Subsample Analysis**

Subsample analysis yields three sets of results: (i) country-specific time-varying co-summability statistics for the entire 152 subsamples (80 for Japan) of sixty years, which I present in graphical form; (ii) comparison of the co-summability subsample results for the United States, Great Britain and Sweden, again in graphical form – this is intended to uncover patterns of commonality and difference in the equilibrium relationship across countries; (iii) co-summability statistics for the post-WWII period as well as results omitting the most recent years covering the global financial crisis (2008-2010).

Graphical results for the sub-sample analysis of co-summability, including Bonferroni-adjusted confidence intervals, are presented in Figure 3.<sup>14</sup> In each plot the end-year of the sixty-year window of analysis is marked on the  $x$ -axis and shading indicates the Bonferroni-adjusted 95% confidence intervals – due to different data availability this time dimension of the plots differ for Japan. Note

first that across all models and countries the confidence intervals are fairly large, typically from 0 to 2 or larger. Second, while (Japan aside) in each country the share of samples which satisfy co-summability is typically above 50%,<sup>15</sup> this data property does not appear to be satisfied consistently over longer stretches of time, but instead appears sporadically. Both of these findings provide a strong signal of misspecification and thus echo the full sample results presented above.

In Table 3 I compare the subsample periods for which the sixty-year data series constituted co-summable specifications in the data for the US, Great Britain and Sweden: Panel A refers to the linear model (Model 1), Panels B and C to the polynomial specifications with (additional) squared terms and squared and cubed debt terms, respectively (Models 2 and 3). For each country a shaded cell indicates the sixty-year subsample ending in the year specified constitutes a co-summable specification, while the intensity of the shading indicates whether this property occurs in one (lightest), two (intermediate) or all three (darkest) countries. Japan is excluded in this graphical analysis since the difference in available time series data would necessitate different shading between earlier (excluding Japan) and later periods (including Japan) which would make a mockery of my attempts to use graphs to illustrate commonality. I begin by focusing on those ‘episodes’ of long-run co-movement when the tests for *all three countries* find co-summability: in all models clusters of such episodes can be found in the 1860s (thus for the series starting

in the early 1800s), the 1890s-1910s (1830s–1850s), and the 1950s and 1960s (1890s–1900s, incorporating both World Wars). Thereafter isolated episodes pop up in the 1970s. The most recent episodes occurred in the early 2000s, which incorporate sample years during WWII and its immediate aftermath. Taken together these various episodes account for 28% of all subsamples across the three specifications.<sup>16</sup> Note that the years of the global financial crisis (2008-10) do not form part of this cluster of co-summable episodes in all three countries.

Referring back to Figure 2 it can be seen that the first of these clusters, covering subsamples ending in the 1860s, occurred when all three countries substantially reduced their country-specific debt-burden (movement to the left in Figure 2) albeit with comparatively modest increase in growth in the US and Sweden (relatively flat line plots). No such pattern is revealed for the second cluster for subsamples ending in the 1890s and 1900s, while the third cluster with end years in the 1950s and 60s occurred when all three countries shifted from a relative debt build-up in years prior to and during WWII to significant debt reduction thereafter, whereby the latter period also represented a return to steady economic growth. The final cluster in the early 2000s again does not reveal any systematic patterns in the evolution of debt burden and growth across these three economies.

Inbetween these episodes there are stretches where two countries have co-summable specifica-

tions (around 31-38% of subsamples in each model), although these are often clustered around the episodes just described. The remainder of subsamples is made up of single country episodes (20-34% in each model) and subsamples with no co-summability in any country (7-16% in each model).

Table 4 then zooms in on the post-WWII period which forms the focal point for virtually all existing empirical studies on the debt-growth nexus. Here we find some evidence for co-summability in nonlinear specifications (these sub-samples are shaded in grey), especially for the model including a cubed term. Note however that the confidence intervals for the overwhelming majority of these results are very large (indicated with the darker shading), such that they include 1 and at times even 2: a large confidence interval is indicative of serious misspecification and these findings of co-summability should thus be treated with caution.

**[Figure 3 and Table 3 about here]**

These robustness checks provide a number of important insights: *first*, there is no overwhelming evidence that these full sample results are severely distorted by global shocks or structural breaks in the long-run debt-growth relationship, given that a very considerable share of subsamples were found to not to be co-summable across all countries and specifications. *Second*, having said that

my results point to a distinct possibility that certain countries experienced linear or nonlinear co-movement between debt and income during certain periods of time over the past two centuries, although seemingly much less so during the 20th century.

[Table 4 about here]

## 6 Concluding Remarks

This study took an alternative approach to investigating the presence of nonlinearities in the long-run equilibrium relation between public debt and growth. Empirical results for four OECD countries using data from 1800 to 2010 and the various robustness checks carried out provide limited evidence for nonlinear, or indeed linear, long-run relationships between these variables. There are however certain subperiods over this long time horizon for which tests confirm co-movement between debt and income. The timing of these subperiods of co-movement frequently appears to differ across countries. These findings are not narrowly confined to the four OECD economies studied in detail but seem to have much wider validity, and further are not an artefact of the simple model specification adopted: I investigated summability and co-summability in a sample of 23 additional countries (including some developing countries), and furthermore studied a num-

ber of theory-based extended specifications for the four OECD economies; results in the Online Appendix provide strong support for the findings presented above.

It is important to emphasise that this study *does not and cannot* address causality from high(er) debt to low(er) growth as has been the focus in most of the empirical work on this topic. This is by no means a shortcoming of the approach taken. Instead, it highlights a central inconsistency in the empirical analysis of nonlinearities in the debt-growth relationship in the existing literature: in order to establish a long-run causal relationship from debt to growth, it is necessary to *first* establish a long-run equilibrium relationship. This study documents the difficulties for establishing the latter using standard empirical specifications adopted in the literature when variable series are integrated.

Once these difficulties are addressed, I find no evidence for a long-run equilibrium relation in the data for four OECD countries. Various robustness checks provide assurance that this finding is not an artefact of sample selection. Since a long-run equilibrium relationship represents a pre-requisite for any long-run causality between variables my analysis by necessity stops at this point.

The results presented in this paper undermine some of the popular conclusions for this politically-charged issue which represent fiscal adjustment as a necessity for *long-run* economic stability and sustainability. I do not claim that a high debt burden is a matter of no concern for policymakers



or that in the short-run debt may not be detrimental to growth. Instead, I highlight the absence of evidence for nonlinearities such as the popular 90% debt-to-GDP threshold or polynomial specifications in the *long-run* relationship with growth and development, which has been the explicit focus of the empirical literature I cite and review.

## **Acknowledgements**

I am indebted to Vanessa Berenguer-Rico, Matei Demetrescu, Gerdie Everaert and Andrea Presbitero for helpful comments and suggestions. Vanessa also kindly provided the Gauss code which forms the basis for the time series analysis carried out in Gauss and the panel analysis carried out in Stata. Comments and suggestions from two anonymous referees which helped to significantly improve the paper are also gratefully acknowledged. The usual disclaimers apply.

## **References**

- [1] Balassone, Fabrizio, Maura Francese and Angelo Pace. 2011. Public Debt and Economic Growth in Italy. Bank of Italy Economic History Working Papers #11.

- [2] Barro, Robert J. 1991. "Economic growth in a cross-section of countries." *Quarterly Journal of Economics* 106(2):407-443.
- [3] Baum, Anja, Cristina Checherita-Westphal and Philipp Rother. 2013. "Debt and growth: New evidence for the euro area." *Journal of International Money and Finance* 32:809-821.
- [4] Berenguer-Rico, Vanessa and Jesus Gonzalo. 2014a. "Co-summability: from linear to non-linear co-integration." unpublished mimeo.
- [5] Berenguer-Rico, Vanessa and Jesus Gonzalo. 2014b. "Summability of stochastic processes: a generalization of integration and co-integration valid for non-linear processes." *Journal of Econometrics* 178(2):331-341.
- [6] Blundell, Richard and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87(1):115-143.
- [7] Bolt, Jutta and Jan Luiten van Zanden. 2013. "The First Update of the Maddison Project – Re-Estimating Growth Before 1820." Maddison-Project Working Paper WP-4, University of Groningen, January .
- [8] Bun, Maurice J.G. and Vasilis Sarafidis. 2013. Dynamic panel data models. In *Oxford Handbook on Panel Data*, ed. Badi H. Baltagi. Vol. forthcoming Oxford University Press.

- [9] Calderon, Cesar and J. Rodrigo Fuentes. 2013. "Government Debt and Economic Growth." Inter-American Development Bank Working Paper Series No. IDB-WP-424.
- [10] Cecchetti, Stephen, Madhusudan Mohanty and 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.
- [11] Checherita-Westphal, Cristina and 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.
- [12] Chinn, Menzie. 2012. "Book review: 'A Decade of Debt.' By Carmen M. Reinhart and Kenneth S. Rogoff." Journal of Economic Literature 50(2):525-27.
- [13] Cordella, Tito, Luca Antonio Ricci and Marta Ruiz-Arranz. 2010. "Debt Overhang or Debt Irrelevance?" IMF Staff Papers 57(1):1-24.
- [14] Dippelsman, Robert, Claudia Helene Dziobek and Carlos Gutierrez Mangas. 2012. What lies beneath: The statistical definition of public sector debt. International Monetary Fund.
- [15] Durlauf, Steven N., Paul A. Johnson and Jonathan R.W. Temple. 2005. Growth Econometrics. In Handbook of Economic Growth, ed. Philippe Aghion and Steven Durlauf. Elsevier chapter

8, pp. 555-677.

- [16] Eberhardt, Markus and Andrea F Presbitero. 2015. "Public debt and growth: Heterogeneity and non-linearity." *Journal of International Economics* 97(1):45-58.
- [17] Eberhardt, Markus and 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.
- [18] Elliott, Graham, Thomas J Rothenberg and James H Stock. 1996. "Efficient tests for an autoregressive unit root." *Econometrica* 64(4):813-836.
- [19] Elmendorf, Douglas W. and Gregory N. Mankiw. 1999. Government debt. In *Handbook of Macroeconomics*, ed. J. B. Taylor and M. Woodford. Elsevier chapter 25, pp. 1615-1669.
- [20] Engle, Robert F. and Clive W. J. Granger. 1987. "Cointegration and Error correction: representations, estimation and testing." *Econometrica* 55(2):252-276.
- [21] Ghosh, Atish R., Jun I. Kim, Enrique G. Mendoza, Jonathan D. Ostry and Mahvash S. Qureshi. 2013. "Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies." *The Economic Journal* 123(566):F4-F30.
- [22] Gonzalo, Jesus and Jean-Yves Pitarakis. 2006. "Threshold effects in co-integrating regressions." *Oxford Bulletin of Economics and Statistics* 68:813-833.

- [23] Greiner, Alfred. 2011. "Economic Growth, Public Debt and Welfare: Comparing Three Budgetary Rules." *German Economic Review* 12(2):205-222.
- [24] Henderson, Daniel and Christopher Parmeter. 2013. "Threshold Estimation in Nonparametric Regression: Government Debt and Economic Growth." Paper presented at the 4th CREA Rimini Workshop on Developments in Macro and Growth, May 23-24.
- [25] Herndon, Thomas, Michael Ash and Robert Pollin. 2014. "Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff." *Cambridge Journal of Economics* 38(2):257-279.
- [26] Inter-American Development Bank. 2006. *Living with Debt: How to Limit the Risks of Sovereign Finance*. Cambridge, MA: Harvard University Press.
- [27] International Monetary Fund. 2012. *World Economic Outlook – Coping with High Debt and Sluggish Growth*. International Monetary Fund.
- [28] Jones, Matthew T and Maurice Obstfeld. 1997. *Saving, investment, and gold: A reassessment of historical current account data*. Technical report, National Bureau of Economic Research.
- [29] Kourtellos, Andros, Thanasis Stengos and Chih Ming Tan. 2013. "The Effect of Public Debt on Growth in Multiple Regimes." *Journal of Macroeconomics* 38:35-43.

- [30] Kraay, Aart and Vikram Nehru. 2006. "When Is External Debt Sustainable?" World Bank Economic Review 20(3):341-365.
- [31] Krugman, Paul. 1988. "Financing vs. forgiving a debt overhang." Journal of Development Economics 29(3):253-268.
- [32] Kumar, Manmohan S. and Jaejoon Woo. 2010. Public Debt and Growth. IMF Working Papers 10/174 International Monetary Fund.
- [33] Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt and Yongcheol Shin. 1992. "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" Journal of Econometrics 54(1):159-178.
- [34] Maddison, Angus. 1992. "A long-run perspective on saving." The Scandinavian Journal of Economics 94(2):181-196.
- [35] Maddison, Angus. 2010. "Historical Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD." Dataset and Documentation provided at the University of Groningen.
- [36] Manasse, Paolo and Nouriel Roubini. 2009. "Rules of thumb for sovereign debt crises." Journal of International Economics 78(2):192-205.
- [37] Mankiw, N Gregory, David Romer and David N Weil. 1992. "A Contribution to the Empirics

of Economic Growth.” *The Quarterly Journal of Economics* 107(2):407-37.

[38] Mitchell, Brian. 1971. *Abstract of British Historical Statistics*. CUP Archive.

[39] Mitchell, Brian. 2007a. *International Historical Statistics 1750-2005: Africa, Asia and Oceania*. 5th ed. Palgrave Macmillan.

[40] Mitchell, Brian. 2007b. *International Historical Statistics 1750-2005: Europe*. 6th ed. Palgrave Macmillan.

[41] Ng, Serena and Pierre Perron. 1995. “Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag.” *Journal of the American Statistical Association* 90(429):268-281.

[42] Panizza, Ugo and Andrea F. Presbitero. 2013. “Public Debt and Economic Growth in Advanced Economies: A Survey.” *Swiss Journal of Economics and Statistics* 149(2):175-204.

[43] Panizza, Ugo and Andrea F. Presbitero. 2014. “Public debt and economic growth: is there a causal effect?” *Journal of Macroeconomics* 41:21-41.

[44] Park, J. Y. and P. C. B. Phillips. 2001. “Nonlinear regressions with integrated time series.” *Econometrica* 69:117-161. Pesaran, M. Hashem and Ron P. Smith. 1995. “Estimating long-

- run relationships from dynamic heterogeneous panels.” *Journal of Econometrics* 68(1):79-113.
- [45] Phillips, Peter CB and Pierre Perron. 1988. “Testing for a unit root in time series regression.” *Biometrika* 75(2):335-346. Politis, Dimitris, Joseph Romano and Michael Wolf. 1999. *Subsampling*. New York: Springer.
- [46] Presbitero, Andrea F. 2012. “Total Public Debt and Growth in Developing Countries.” *European Journal of Development Research* 24(4):606-626.
- [47] Reinhart, Carmen M. and Kenneth S. Rogoff. 2009. *This Time is Different - Eight Centuries of Financial Folly*. Princeton, New Jersey: Princeton University Press.
- [48] Reinhart, Carmen M. and Kenneth S. Rogoff. 2010a. *Debt and Growth Revisited*. MPRA Paper 24376 University Library of Munich, Germany.
- [49] Reinhart, Carmen M. and Kenneth S. Rogoff. 2010b. “Growth in a Time of Debt.” *American Economic Review Papers and Proceedings* 100(2):573-78.
- [50] Reinhart, Carmen M. and Kenneth S. Rogoff. 2011. *A Decade of Debt*. Vol. 95 of *Policy Analyses in International Economics*, Washington D.C.: Peterson Institute for International Economics.



- [51] Reinhart, Carmen M., Kenneth S. Rogoff and Miguel A. Savastano. 2003. "Debt Intolerance." *Brookings Papers on Economic Activity* 34(1):1-74.
- [52] Reinhart, Carmen M., Vincent R. Reinhart and Kenneth S. Rogoff. 2012. "Public Debt Overhangs: Advanced-Economy Episodes Since 1800." *Journal of Economic Perspectives* 26(3):69-86.
- [53] Rogoff, Kenneth. 2013. FAQ on Herndon, Ash and Pollin's Critique of 'Growth in a Time of Debt.' Technical report. available on Rogoff's website at <http://tinyurl.com/ot8h53e>, October.
- [54] Sachs, Jeffrey D. 1989. The Debt Overhang of Developing Countries. In *Debt, Stabilization and Development*, ed. Guillermo A. Calvo, Ronald Findlay, Pentti Kouri and Jorge Braga de Macedo. Oxford: Blackwell.
- [55] Schon, Lennart and Olle Krantz. 2012. "Swedish Historical National Accounts 1560-2010." *Lund Papers in Economic History* 123.
- [56] Teles, Vladimir K. and Caio Cesar Mussolini. 2014. "Public debt and the limits of fiscal policy to increase economic growth." *European Economic Review* 66:1-15.
- [57] Van Zanden, J.L. 2001. Early modern economic growth: a survey of the European economy 1500-1800. In *Early Modern Capitalism*, ed. M. Prak. London: Routledge.

## Notes

<sup>1</sup>A further empirical study by Baum, Checherita-Westphal and Rother (2013) is cited in Reinhart, Reinhart and Rogoff (2012) but (erroneously) argues to focus on the *short-run* relationship. They note that their sample selection is driven by the finding that data for 1990–2007 appears stationary, whereas the longer 1980–2007 data appears nonstationary.

<sup>2</sup>The former two economies are presently at the centre of a policy debate relating sustainable growth to fiscal austerity (e.g. US Senate Budget Committee, June 4, 2013), Japan is at times taken as an example for sustained growth at comparatively high levels of debt, while Sweden (alongside the US and Britain) represents the country with the longest time series in my matched dataset.

<sup>3</sup>Notable exceptions include studies by Henderson and Parmeter (2013) and Kourtellos, Stengos and Tan (2013) which emphasise the heterogeneity of the debt-growth nexus across countries and adopt nonparametric methods to identify a threshold in the cross-section dimension.

<sup>4</sup>A thorough critique of this implementation in the macro panel context is beyond the scope of this paper. Eberhardt and Teal (2010) highlight the problems arising, Bun and Sarafidis (2013) provide an analysis of the impact of nonstationary initial conditions on this set of estimators while

Pesaran and Smith (1995) discuss the bias arising from heterogeneity misspecification.

<sup>5</sup>Adopting the threshold specification I find that in my data series for Italy none of the various thresholds adopted pass the co-summability test (100% threshold CI low 0.313,  $\hat{\delta}_{\hat{e}_t}=1.134$ , CI up 1.954; 90% CI low 0.486,  $\hat{\delta}_{\hat{e}_t} = 1.052$ , CI up 1.619; 70% CI low 0.910,  $\hat{\delta}_{\hat{e}_t} = 1.695$ , CI up 2.480; 50% CI low 0.811,  $\hat{\delta}_{\hat{e}_t} = 1.471$ , CI up 2.130) – see results section for notation.

<sup>6</sup>For a formal definition of summability see Definition 2 in Berenguer-Rico and Gonzalo (2014b).

<sup>7</sup>The deterministic component  $m_t$  can be accounted for by the partial mean of  $y_t$ , namely  $m_t = (1/t) \sum_{j=1}^t y_j$  in case of a constant. Given the trending behaviour of my data I focus below on the case of constant and linear trend terms, where partial demeaning of  $y_t$  is carried out twice.

<sup>8</sup>I am grateful to a referee who emphasises that the validity of the subsampling procedure has only been shown by simulation.

<sup>9</sup>A referee kindly pointed out that the properties of the balance statistic in a simple model of  $y_t = \theta x_t + u_t$ , with  $x_t \sim S(\delta)$  and  $u_t \sim S(\delta)$ , are badly affected by  $\theta$ , especially when  $\theta \neq 1$ .

<sup>10</sup>The residual series  $\hat{e}_t$  will sum to zero by default of the least squares principle if our specification includes an intercept; in practice the estimate for the intercept term is therefore not subtracted when constructing  $\hat{e}_t$ .

<sup>11</sup>Parts of the literature, including Baum, et al (2013), employ threshold regression algorithms where the threshold value is determined endogenously. Extending the co-summability approach in a similar fashion is beyond the scope of this study.

<sup>12</sup>The data used here refer to central government debt, which excludes any debt from local government, as opposed to general government debt. As shown in the work of Dippelsman, Dziobek and Mangas (2012), the quantitative implications of this choice of variable can be stark. However, in the absence of any general government debt data over the long time horizon the analysis here is forced to employ the conceptually inferior ‘central government’ measures – this choice is however aligned with the analysis in Reinhart and Rogoff (2010b).

<sup>13</sup>In additional work presented in the Online Appendix I investigate a ‘reversed’ model for debt with linear and polynomial terms of per capita GDP as the regressors and similarly find precious little evidence for a long-run equilibrium relationship.

<sup>14</sup>Not adopting the Bonferroni adjustment would lead to significantly *narrower* confidence intervals, which in all cases would yield the same or a stronger qualitative result of limited evidence for a long-run equilibrium relationship between debt and income.

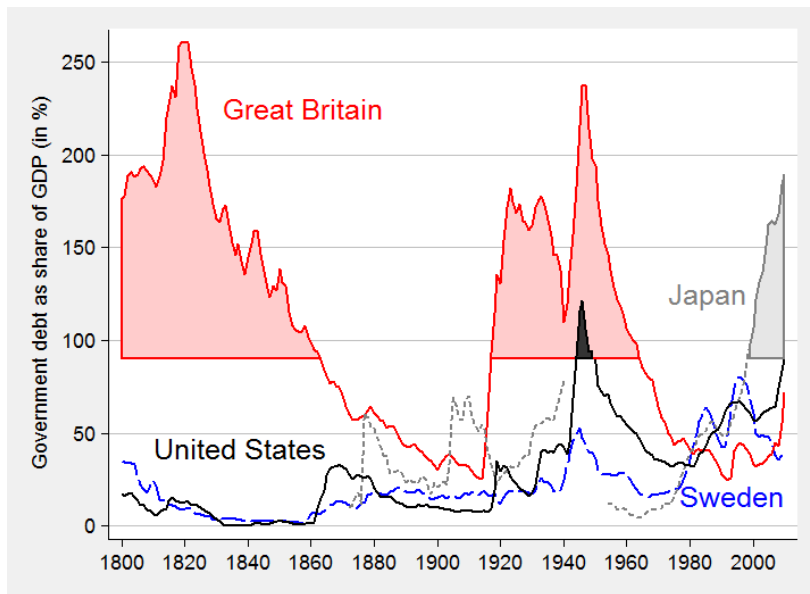
<sup>15</sup>The overall share of samples which satisfy co-summability is as follows: USA 50%, GBR

52%, SWE 49%, JPN 14% (Model 1); USA 56%, GBR 64%, SWE 64%, JPN 17% (Model 2); USA 59%, GBR 73%, SWE 64%, JPN 28% (Model 3). Results are qualitatively similar if I adopt a longer (70-year) window instead.

<sup>16</sup>In Model 1 they make up 17% of all subsamples, in Models 2 and 3 30% and 36% respectively.

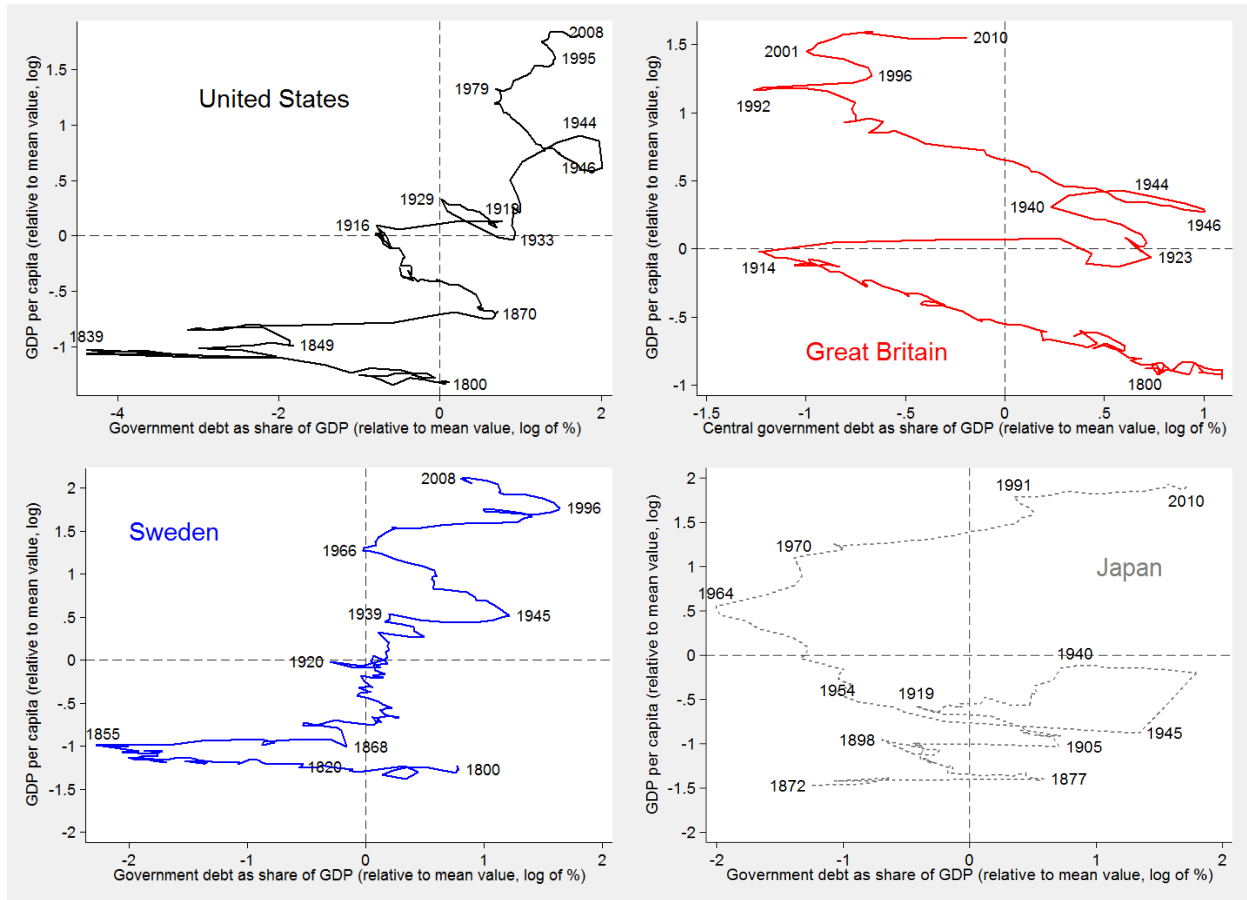
# Figures and Tables

Figure 1: Evolution of Debt/GDP ratios



**Notes:** The shaded areas represent the periods where debt/GDP exceeded 90%.

Figure 2: Debt Ratio and Income Per Capita



**Notes:** Debt ratios and per capita GDP series (both in logarithms) are presented in deviation from their country-specific time-series means (within transformation).

Figure 3: Co-Summability Testing (Sub-Samples)

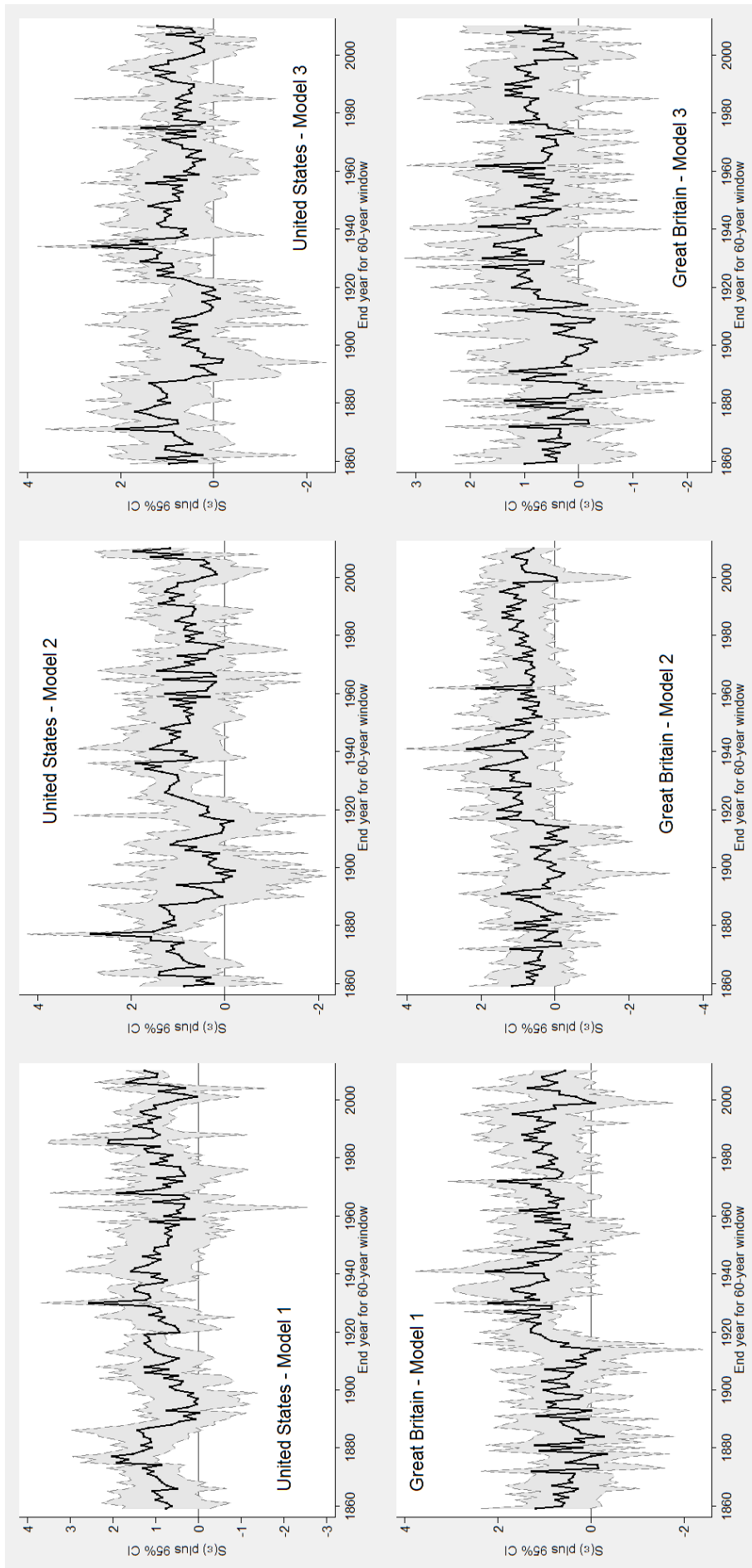


Figure continued on the following page



Figure 3: Co-Summability Testing (Sub-Samples) continued

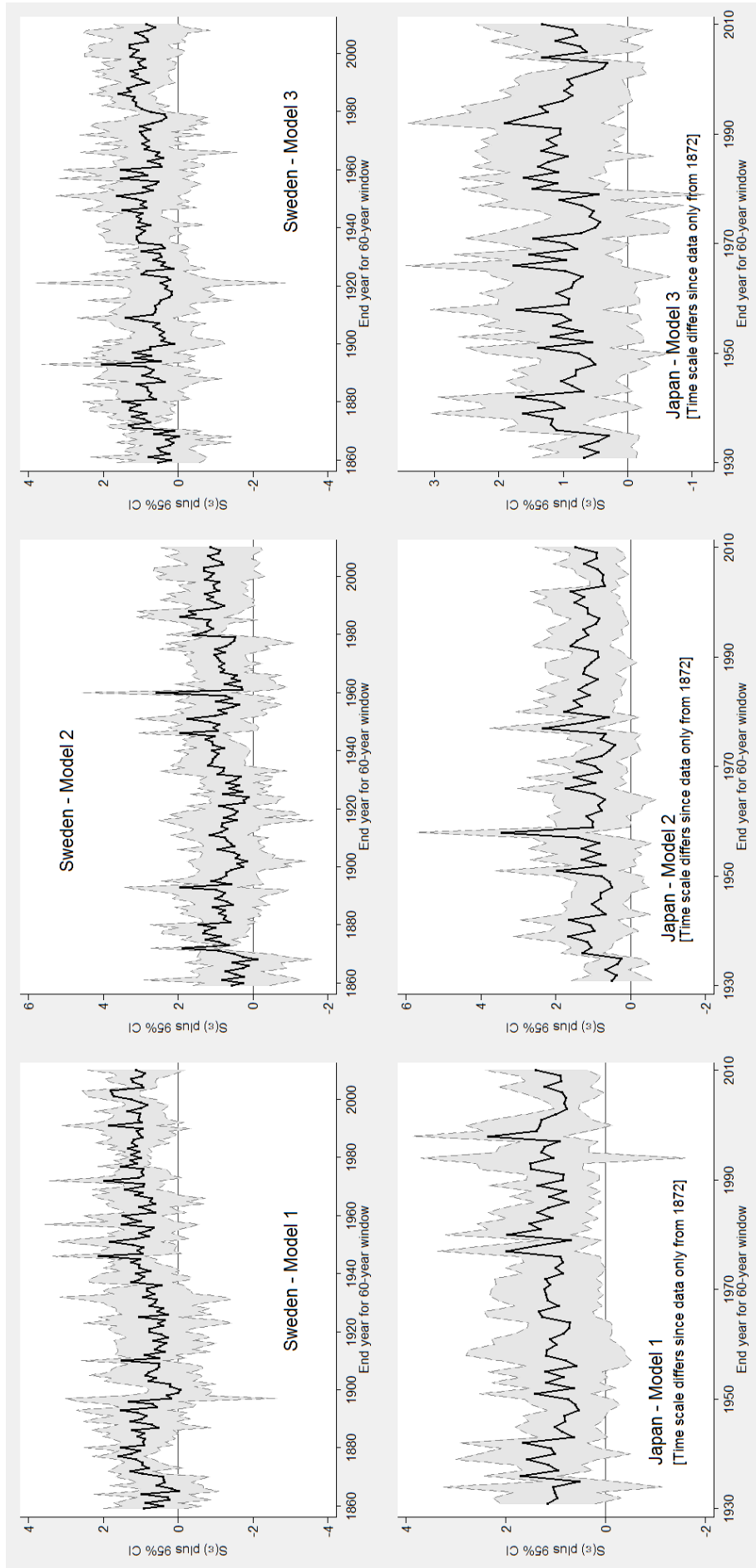


Figure Notes overleaf.

**Notes (to Figure 3):** The shaded areas represent the Bonferroni-corrected 95% Confidence

Intervals for the Co-Summability statistic computed in a moving window of 60-year time periods.

The solid black line represents the computed Co-Summability statistic. I allow for an intercept in

the co-summability analysis. The coverage of the data differs across countries: for the US, Great

Britain and Sweden I have data from 1800-2010 (152 subsamples), for Japan from 1872-2010 (80

subsamples). Model 1 refers to a specification with linear debt terms only, Model 2 to a

specification with linear and quadratic debt terms, Model 3 further includes a cubed debt term.

Table 1: Estimated Order of Summability

Country	Start & End Year	Obs	Variable	CI low	$\hat{\delta}$	CI up	
USA	1800	2010	211	ln(GDP pc)	0.652	1.490	2.329
				$\Delta$ ln(GDP pc)	-0.524	0.066	0.657
				ln(Debt/GDP)	0.551	1.082	1.613
				ln(Debt/GDP) squared	0.383	0.860	1.336
				ln(Debt/GDP) cubed	0.404	0.993	1.582
				168 Debt/GDP < 50%	0.313	0.825	1.337
				43 Debt/GDP $\geq$ 50%	0.691	1.409	2.127
GBR	1800	2010	211	ln(GDP pc)	0.802	1.839	2.877
				$\Delta$ ln(GDP pc)	-0.452	0.130	0.712
				ln(Debt/GDP)	0.540	0.967	1.393
				ln(Debt/GDP) squared	0.509	0.948	1.386
				ln(Debt/GDP) cubed	0.475	0.931	1.387
				100 Debt/GDP < 90%	0.511	1.062	1.613
				111 Debt/GDP $\geq$ 90%	0.405	0.936	1.467
				86 Debt/GDP < 70%	0.428	1.200	1.972
				125 Debt/GDP $\geq$ 70%	0.465	0.923	1.381
				64 Debt/GDP < 50%	0.447	1.068	1.689
				147 Debt/GDP $\geq$ 50%	0.459	0.898	1.336
SWE	1800	2010	211	ln(GDP pc)	0.495	0.897	1.298
				$\Delta$ ln(GDP pc)	-1.110	-0.378	0.355
				ln(Debt/GDP)	0.645	1.624	2.603
				ln(Debt/GDP) squared	0.704	1.577	2.451
				ln(Debt/GDP) cubed	0.677	1.538	2.399
JPN	1872	2010	139	ln(GDP pc)	0.987	2.390	3.792
				$\Delta$ ln(GDP pc)	-0.692	-0.004	0.683
				ln(Debt/GDP)	0.427	1.091	1.755
				ln(Debt/GDP) squared	0.433	1.101	1.769
				ln(Debt/GDP) cubed	0.410	1.114	1.819
				85 Debt/GDP < 50%	0.543	1.282	2.022
				54 Debt/GDP $\geq$ 50%	0.192	1.025	1.858

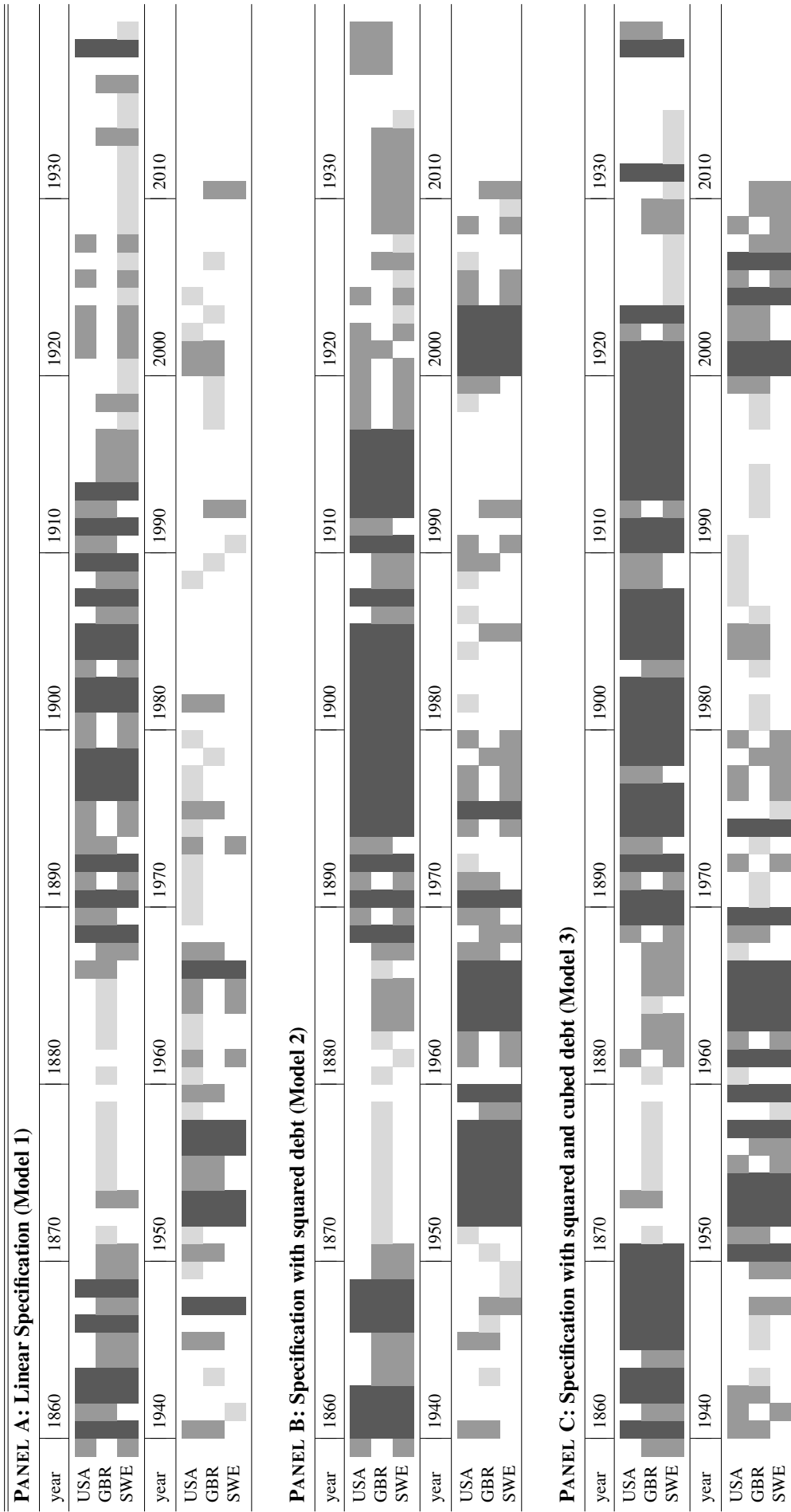
**Notes:** CI low and up indicate the 95% confidence interval for the summability estimate  $S(\delta)$  constructed from subsampling – shaded cells indicate variable series where the summability confidence interval includes zero. In all tests conducted I allow for deterministic terms (constant and trend).

Table 2: Co-Summability – ln(GDP pc) specifications

	Start	End	obs	Nonlinearity	CI low	$\hat{\delta}_{\hat{\epsilon}_t}$	CI up	Verdict
USA	1870	2010	141	-	0.467	1.049	1.631	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$b = 13$	Debt/GDP squared	0.277	0.943	1.609	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$M = 129$	Debt/GDP cubed	0.351	0.900	1.449	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
				Threshold 50%	0.526	1.124	1.721	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
GBR	1830	2010	181	-	0.660	1.194	1.728	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$b = 14$	Debt/GDP squared	0.699	1.196	1.693	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$M = 168$	Debt/GDP cubed	0.702	1.196	1.689	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
				Threshold 90%	0.653	1.208	1.763	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
				Threshold 70%	0.656	1.224	1.791	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
				Threshold 50%	0.668	1.284	1.899	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
SWE	1820	2010	191	-	0.777	1.546	2.314	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$b = 15$	Debt/GDP squared	0.658	1.602	2.546	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$M = 177$	Debt/GDP cubed	0.697	1.598	2.499	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
JPN	1872	2010	139	-	0.580	1.128	1.676	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$b = 13$	Debt/GDP squared	0.236	0.873	1.511	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
			$M = 127$	Debt/GDP cubed	0.181	0.821	1.460	$S(\delta_{\hat{\epsilon}_t}) \neq 0$
				Threshold 50%	0.857	2.056	3.256	$S(\delta_{\hat{\epsilon}_t}) \neq 0$

**Notes:** In all models I take per capita GDP (in logarithms) as the dependent variable. CI low and up indicate the 95% confidence interval for the co-summability estimates. In all tests conducted I allow for deterministic terms (constant only; additional trends do not qualitatively change the results).  $\hat{\delta}_{\hat{\epsilon}_t} \neq (=)0$  implies that co-summability is (not) rejected. Obs reports the number of observations,  $b = \text{int}\sqrt{T} + 1$  refers to the time series length of each subsample,  $M = T - b + 1$  to the number of subsamples used in the analysis. Regarding the ‘Nonlinearity,’ the model with  $\ln(\text{Debt}/\text{GDP})^2$  also includes  $\ln(\text{Debt}/\text{GDP})$ , while the model with  $\ln(\text{Debt}/\text{GDP})^3$  also includes  $\ln(\text{Debt}/\text{GDP})^2$  and  $\ln(\text{Debt}/\text{GDP})$ .

Table 3: Co-Summability – Cross-Country Comparison



**Notes:** Increasing shading indicates the number (0-3) of countries for which the sixty-year subsample ending in the year indicated has a co-summable specification.

Table 4: Co-Summability – ln(GDP pc) specifications (sub-sample results for post-WWII period)

$M$	start	end	United States				United Kingdom			
			CI low	$\hat{\delta}_{\hat{e}_t}$	CI up	Verdict	CI low	$\hat{\delta}_{\hat{e}_t}$	CI up	Verdict
1	1946	2005	0.051	0.699	1.347	$S(\delta_{\hat{e}_t}) \neq 0$	0.080	0.936	1.792	$S(\delta_{\hat{e}_t}) \neq 0$
	1947	2006	0.965	1.697	2.428	$S(\delta_{\hat{e}_t}) \neq 0$	-0.126	0.725	1.577	$S(\delta_{\hat{e}_t}) = 0$
	1948	2007	0.886	1.435	1.984	$S(\delta_{\hat{e}_t}) \neq 0$	0.072	1.024	1.977	$S(\delta_{\hat{e}_t}) \neq 0$
	1949	2008	0.698	0.987	1.277	$S(\delta_{\hat{e}_t}) \neq 0$	0.360	1.051	1.742	$S(\delta_{\hat{e}_t}) \neq 0$
	1950	2009	0.704	0.964	1.225	$S(\delta_{\hat{e}_t}) \neq 0$	0.220	0.738	1.256	$S(\delta_{\hat{e}_t}) \neq 0$
	1951	2010	0.836	1.263	1.690	$S(\delta_{\hat{e}_t}) \neq 0$	-0.145	0.551	1.247	$S(\delta_{\hat{e}_t}) = 0$
2	1946	2005	-0.250	0.347	0.944	$S(\delta_{\hat{e}_t}) = 0$	0.158	0.780	1.402	$S(\delta_{\hat{e}_t}) \neq 0$
	1947	2006	-0.128	0.444	1.016	$S(\delta_{\hat{e}_t}) = 0$	0.101	0.765	1.428	$S(\delta_{\hat{e}_t}) \neq 0$
	1948	2007	0.504	1.573	2.641	$S(\delta_{\hat{e}_t}) \neq 0$	0.375	1.170	1.966	$S(\delta_{\hat{e}_t}) \neq 0$
	1949	2008	-0.741	0.903	2.548	$S(\delta_{\hat{e}_t}) = 0$	0.385	1.048	1.711	$S(\delta_{\hat{e}_t}) \neq 0$
	1950	2009	1.132	1.954	2.777	$S(\delta_{\hat{e}_t}) \neq 0$	0.130	0.709	1.289	$S(\delta_{\hat{e}_t}) \neq 0$
	1951	2010	0.777	1.169	1.561	$S(\delta_{\hat{e}_t}) \neq 0$	-0.167	0.585	1.336	$S(\delta_{\hat{e}_t}) = 0$
3	1946	2005	-0.924	0.277	1.478	$S(\delta_{\hat{e}_t}) = 0$	0.057	0.648	1.238	$S(\delta_{\hat{e}_t}) \neq 0$
	1947	2006	-0.913	0.233	1.379	$S(\delta_{\hat{e}_t}) = 0$	-0.168	0.490	1.148	$S(\delta_{\hat{e}_t}) = 0$
	1948	2007	0.337	1.017	1.697	$S(\delta_{\hat{e}_t}) \neq 0$	-0.361	0.488	1.336	$S(\delta_{\hat{e}_t}) = 0$
	1949	2008	-0.066	0.411	0.887	$S(\delta_{\hat{e}_t}) = 0$	0.433	1.319	2.205	$S(\delta_{\hat{e}_t}) \neq 0$
	1950	2009	0.008	0.482	0.956	$S(\delta_{\hat{e}_t}) \neq 0$	-0.969	0.525	2.018	$S(\delta_{\hat{e}_t}) = 0$
	1951	2010	0.790	1.217	1.644	$S(\delta_{\hat{e}_t}) \neq 0$	-0.159	0.982	2.123	$S(\delta_{\hat{e}_t}) = 0$
$M$	start	end	Sweden				Japan			
			CI low	$\hat{\delta}_{\hat{e}_t}$	CI up	Verdict	CI low	$\hat{\delta}_{\hat{e}_t}$	CI up	Verdict
1	1946	2005	0.558	0.974	1.391	$S(\delta_{\hat{e}_t}) \neq 0$	0.205	0.794	1.383	$S(\delta_{\hat{e}_t}) \neq 0$
	1947	2006	0.625	1.223	1.821	$S(\delta_{\hat{e}_t}) \neq 0$	0.039	0.879	1.720	$S(\delta_{\hat{e}_t}) \neq 0$
	1948	2007	0.575	1.265	1.955	$S(\delta_{\hat{e}_t}) \neq 0$	0.311	1.234	2.156	$S(\delta_{\hat{e}_t}) \neq 0$
	1949	2008	0.311	0.984	1.657	$S(\delta_{\hat{e}_t}) \neq 0$	0.173	0.893	1.613	$S(\delta_{\hat{e}_t}) \neq 0$
	1950	2009	0.075	0.899	1.722	$S(\delta_{\hat{e}_t}) \neq 0$	0.099	0.913	1.726	$S(\delta_{\hat{e}_t}) \neq 0$
	1951	2010	-0.174	1.127	2.428	$S(\delta_{\hat{e}_t}) = 0$	0.380	1.402	2.423	$S(\delta_{\hat{e}_t}) \neq 0$
2	1946	2005	-0.205	0.847	1.900	$S(\delta_{\hat{e}_t}) = 0$	0.041	0.733	1.425	$S(\delta_{\hat{e}_t}) \neq 0$
	1947	2006	0.039	1.071	2.103	$S(\delta_{\hat{e}_t}) \neq 0$	0.225	0.876	1.526	$S(\delta_{\hat{e}_t}) \neq 0$
	1948	2007	0.052	1.203	2.354	$S(\delta_{\hat{e}_t}) \neq 0$	0.396	1.237	2.078	$S(\delta_{\hat{e}_t}) \neq 0$
	1949	2008	-0.074	0.933	1.941	$S(\delta_{\hat{e}_t}) = 0$	0.145	0.899	1.653	$S(\delta_{\hat{e}_t}) \neq 0$
	1950	2009	-0.260	0.881	2.023	$S(\delta_{\hat{e}_t}) = 0$	0.185	0.905	1.626	$S(\delta_{\hat{e}_t}) \neq 0$
	1951	2010	-0.197	1.137	2.471	$S(\delta_{\hat{e}_t}) = 0$	0.405	1.482	2.560	$S(\delta_{\hat{e}_t}) \neq 0$
3	1946	2005	-0.092	0.823	1.738	$S(\delta_{\hat{e}_t}) = 0$	-0.095	0.643	1.381	$S(\delta_{\hat{e}_t}) = 0$
	1947	2006	-0.200	0.947	2.094	$S(\delta_{\hat{e}_t}) = 0$	-0.240	0.738	1.717	$S(\delta_{\hat{e}_t}) = 0$
	1948	2007	-0.241	1.146	2.532	$S(\delta_{\hat{e}_t}) = 0$	0.106	1.120	2.133	$S(\delta_{\hat{e}_t}) \neq 0$
	1949	2008	-0.853	0.804	2.460	$S(\delta_{\hat{e}_t}) = 0$	-0.104	0.786	1.677	$S(\delta_{\hat{e}_t}) = 0$
	1950	2009	-0.201	0.621	1.444	$S(\delta_{\hat{e}_t}) = 0$	-0.388	0.858	2.103	$S(\delta_{\hat{e}_t}) = 0$
	1951	2010	-0.020	0.853	1.727	$S(\delta_{\hat{e}_t}) = 0$	0.317	1.332	2.347	$S(\delta_{\hat{e}_t}) \neq 0$

**Notes:** This table presents results for sub-sample co-summability testing. Model ( $M$ ) 1-3 refer to the following specifications: 1 – linear, 2 – linear and squared, 3 – linear, squared and cubed debt/GDP terms (in logs). In each case I report statistics from the sixty-year samples, with start and end years as indicated in the table, focusing on the period after 1945. Shaded cells indicate sub-samples where the co-summability confidence interval includes zero – note however that these confidence intervals are at times very large, such that they further include 1 and at times even 2 (the latter options are represented by darker shading). See Table 2 for all other details.