Economic growth and public debt: addressing unobserved heterogeneity[1]

João Basilio Pereima
Economic Department
Federal University of Paraná - UFPR, Brazil
joaobasilio@ufpr.br

Manuela Merki
Economic Department
Federal University of Paraná - UFPR, Brazil
manuela.merki@gmail.com

Fernando Motta Correia
Economic Department
Federal University of Paraná - UFPR, Brazil
fmottabr@yahoo.com.br

July 20, 2015

Abstract: The article addresses the non-linear and heterogeneous relationship between public debt and economic growth considering the recent combination of increasing public expenditures and debt with continuing low economic growth observed in many countries, controlling the results by an institutional quality multidimensional indicator. Economic theory generally supports a negative relationship between debt and growth in the long-run, whereas in the short-run fiscal stimulus may induce positive effects. Empirical literature provides some evidence for both, showing that the negative relationship might become more important after reaching a certain threshold, but the results are not absolutely conclusive. Problems of unobserved heterogeneity might be a major reason for that. To tackle the problem of unobserved heterogeneity we suggest a two-stage procedure applied to a large sample of 154 countries. We first use cluster analysis to group the countries by their governance quality using the Worldwide Governance Indicators (WGI) database. In a second step the relationship between public debt and growth is investigated using simple and threshold multiple panel regression models (TMR). The results confirm the coefficients to be negative in the long-run and to vary across country groups (clusters) depending on institutional quality.

JEL Classification Code: O47, E62, C24, F34, H63
Keywords:: Economic Growth; Public debt; Threshold models; Heterogeneity; Governance.

Resumo: O artigo analisa a relação não linear e heterogênea entre dívida pública e crescimento econômico considerando o recente aumento dos gastos e dívidas públicas dos países combinada com diminuição da taxa de crescimento anual do produto observado em muitos países, controlando os resultados usando um indicador multidimensional de qualidade institucional. A teoria econômica geralmente assume uma correlação negativa entre dívida e crescimento no longo prazo, enquanto que em algumas situações pode haver correlação positiva no curto-prazo. A literatura empírica em geral confirma estes resultados, mostrando que uma correlação negativa ocorre principalmente a partir de um certo nível de dívida, mas neste caso os resultados não são conclusivos. Problemas de heterogeneidade não observada podem ser uma importante explicação. Para analisar o problema da heterogeneidade não observada nós sugerimos um procedimento em dois estágios, usando uma amostra abrangente de 154 países. Primeiramente aplicamos análise de cluster para agrupar países de acordo com a qualidade de governança usando dados do Worldwide Governance Indicator (WGI). Num segundo passo a relação entre dívida pública e crescimento é investigada com modelos multiregressivos com efeito limiar para dados em painel. Os resultados confirmam a relação negativa de longo prazo e um coeficiente que varia muito entre os *clusters* dependendo da qualidade das instituições.

JEL Classification Code: O47, E62, C24, F34, H63
Keywords:: Crescimento Econômico; Dívida Pública; Modelos Threshold; Heterogeneidade; Governança.

Economic growth and public debt: addressing unobserved heterogeneity

1 Introduction

Debt ranks high on the public agenda, not only in certain European countries where the financial crisis turned into a crisis of public debt. Governments all over the world have tried to stimulate economic activity by means of tax relieves, subsidies and cheap money, creating and aggravating fiscal imbalances. In many cases, growth remained low or declined even further. It is therefore more than ever important to understand the impact of fiscal policy, public expenditure and debt on economic growth and social welfare. Economic literature still provides insufficient advice for economic policy. Against this background, we review the existing literature on the relationship between public debt[2]/public spending[3] and economic growth.

Economic theory generally supports a negative relationship between debt and growth in the long run through the traditional crowding out effects, whereas in the short run fiscal policy may temporarily induce growth through demand side stimulus, especially if households assume non-Ricardian behavior and economic output is far from production capacity. Empirical evidence is less clear. The findings vary considerably and are sensitive to the definition of the sample and temporal coverage. Problems of endogeneity, nonlinearities and causality have been continuously discussed. Proving causality is a very ambitious goal - probably the most ambitious one - and general knowledge on growth econometrics provides little advice. Nonlinear and threshold models indeed have been an intensive area of research in recent years and considerably improved knowledge and consistency of results. However, latest contributions have pointed to considerable biases through hidden factors and unobserved heterogeneity (e.g. Panizza and Presbitero (2013), Kourtellos et al. (2013)).

In order to estimate appropriately the nonlinear relationship and threshold given the heterogeneous sample of countries, we propose a two-stage approach to address this short-coming. We account for unobserved heterogeneity by grouping the countries based on the quality of governance and institutions, making it observable. We use a cluster methodology based on the six dimensions of the Worldwide Governance Indicators (WGI) database to classify the large sample of 154 countries into 5 more homogeneous groups. In a second step we apply the commonly used linear and nonlinear models to investigate the relationship between public debt and growth to each of the country groups.

The article is organized as follows: Section 2 reviews the existing theoretical and empirical literature including indications for unobserved heterogeneity. Section 3 discusses the possibilities to address heterogeneity and motivates the utilization of institutional quality indicators. Section 4 presents our approach to control heterogeneity by means of cluster analysis and presents the resulting country groups. Section 5 then evaluates several estimation techniques and analyses

---

[2]In this paper, government debt or public debt refer to gross general government debt, i.e. public liabilities at all levels of government - central, provincial and local. We do not address the discussion about the impact of federalism. The effects of the allocation of fiscal power, transfer payments and the various redistribution channels are not easy to measure, and the discussion on revenue sharing between central and local governments would go well beyond the scope of this analysis. Gross general government debt covers external as well as domestic debt. Even though there are valid arguments to distinguish external and domestic debt (domestic debt, issued under domestic legal jurisdiction, is more easy to get rid of by inflation or changing credit or tax regulation), we do not address them separately. The discussion seems less important today, than it has been before. Many emerging markets rely much more on domestic creditors today as in the 1980s and 1990s. Rational markets should take into account foreign and domestic debt alike.

[3]We basically focus on public debt. However, as public debt and public spending go hand in hand, it is beneficial to have a look at findings from research on public spending and other related areas as well.

the relationship between debt and growth for each of the sub-samples. The results confirm the existence of a negative relationship between debt and growth and disclose the existence of significant differences in debt coefficients between the sub-samples. Section 6 comments and concludes.

2 Literature review

According to the conventional economic view, debt reflecting deficit financing due to government spending or tax reduction, might stimulate aggregate demand and output in the short run. In the long-run, an increase of public debt will indeed reduce the growth rate through lower productivity[4]. No matter whether the debt influences the economic activity through crowding-out effects, higher interest rates, higher volatility due to reduced flexibility of the government to perform an anti-cyclical fiscal policy and/or uncertainties about future policies and taxation, at the end debt reduces investment, hamper the increase of the capital stock and thus finally slow down growth in productivity and economic growth. Lower productivity in turn means lower competitiveness and will therefore negatively impact the trade balance. And finally with growth being reduced, tax revenue is likely to decrease, increasing the probability of further budget deficits and recourse on credits, initiating a vicious cycle.

The effects of debt and fiscal policy on growth depend on the structure as well as on the level of taxation and expenditure. The literature distinguishes various types of taxes and expenditures - e.g. more and less distortionary taxes, public consumption versus investment, unproductive versus productive spending or different public expenditure categories, i.e. military, social security, education or health expenditures (e.g. Devarajan et al. (1996) or Kneller et al. (1999)). Productive expenditures are generally found to have more positive impact unless financed by distortionary taxes.

The extent of the effect of fiscal policy may be reduced by partial or complete Ricardian behaviour (Barro, 1974). That is, if individual actors anticipate the future impact of public expenditure and increasing debt and adapt their behaviour to neutralize the effect. Even in presence of such a situation of anticipatory market participants where the theory of debt neutrality or Ricardian equivalence would be perfectly met, the impact of additional debt will, however, most probably not be zero. Government consumption and future taxation may involve changes in relative utilities and have distortionary effects as a whole.

Empirical findings on the relationship between public debt and growth are not clear and sometimes even inconsistent. Many studies provide evidence for a negative relationship between public expenditure, government consumption, budget deficits or public debt and growth (e.g. Aiginger and Falk (2005) or Kumar and Woo (2010)). Nonetheless, evidence has been generally inconclusive, sensitive to the selection of the sample and time-period investigated and smaller than expected. Also endogeneity and causality issues have been disputed. In principle, this is little surprising. The literature on the relationship between public debt and growth builds on the vast research on the determinants of medium and long-run growth. Therefore, the same challenges and concerns about robustness, specification of the explanatory variables and intermediating factors apply (see Sala-i Martin (1997), Sala-i Martin et al. (2004) or also the comprehensive summary on growth econometrics by Durlauf et al. (2005)).

Attempts to unravel the puzzling relationship between debt and growth have intensified recently. Especially gained momentum the idea of a nonlinear relationship and the existence of a "debt threshold". Of special importance has been the analysis of Reinhart and Rogoff (2012) who - based on a comprehensive panel consisting of 44 countries and a two-centuries time span -

---

[4]Elmendorf and Mankiw (1999) give a comprehensive overview of the theoretical literature on the macroeconomic effects of government debt.

find evidence for a significant negative relationship between debt and growth for countries with "high" debt levels (higher than 90 percent of GDP), whereas the link between debt and growth seemed to be weak at "normal" debt levels. Reinhart and Rogoff (2012) have been criticized and challenged by Herndon et al. (2013) who show that the threshold effect vanishes after correcting for a coding error and using a different weighting of the data. Nevertheless the work has inspired a vast and growing empirical literature devoted to confirm the existence of causal relationships as well as nonlinearity and thresholds effects. Several contributions confirmed the existence of nonlinearity, whilst others did not find such evidence.

Besides this recent revitalized empirical research, Tanzi and Zee (1996) and Devarajan et al. (1996) already suggested a couple of years ago that the relationship between growth and fiscal variables may not be monotonic and that not only the type of use and sourcing (distortionary vs. undistortionary) is important, but also the level of the expenditure. Devarajan et al. (1996) showed that even productive expenditures could become unproductive if there is an excessive amount of them.

Searching for nonlinearity Checherita and Rother (2010) used a sample of 12 euro area countries and data between 1970-2008. They found evidence for an inversely u-shaped relationship between government debt and growth with a turning point between 90%-100%. Baum et al. (2013) also analysed 12 Euro area countries, but limited to a more narrow period between 1990-2010, and confirmed these results. For high debt ratios above 95%, additional debt has a clearly negative impact on economic activity. Cecchetti et al. (2011), for a sample of 18 OECD countries and a period of 1980-2009, find a negative impact of public debt on the subsequent average growth. A 10 percentage points increase in the debt ratio is associated with a 17-18 basis points reduction in growth. Taking into account threshold effects, they state indeed, that government debt only starts to reduce economic growth once it crosses a threshold of about 85 percent. Below this level the relationship is insignificant. On the other hand, Pescatori et al. (2014) do not find any particular debt threshold. Notwithstanding, they find evidence that the debt trajectory might have an important impact on future growth. A country's debt trajectory can even be as important as the level of debt. Countries with high but declining debt, can experience the same growth rate as a low indebted country. However, Pescatori et al. (2014) also find evidence that higher levels of debt are associated with a higher degree of output volatility.

The majority of the empirical research focuses on advanced economies for data availability reasons (OECD data) or for other considerations. But evidence exist that growth effects differ between country groups. For a more comprehensive sample of 155 countries, Afonso and Jalles (2013) find a much lower debt ratio threshold of 59%. The endogenous thresholds is slightly higher at 79% of GDP for emerging countries whereas for the Euro area it is 58%. Kumar and Woo (2010) use a series of panel data estimators for a sample of 38 emerging and developing economies over the period of 1970-2009. They find a negative relationship between debt and growth - on average a 10 percentage points increase in the debt-to-GDP ratio is associated with a slowdown in annual real per capita GDP growth of 0.2 percentage points per year - as well as evidence for nonlinearities with high debt levels having a proportionately larger negative effect. In addition they detect the impact of debt to be smaller in advanced economies. Advanced economies experience a lower slowdown of around 0.15%–0.2%, whereas the effect in emerging economies is twice as large (0.3%–0.4%).

Emerging market countries experience difficulties in raising additional funds at higher debt levels that would be manageable by advanced country criteria without any problem. At the same time they experience high incentives to increase spending - meeting need to invest in education, social security, and public goods, what favours a policy of expanding public expenditures. Reinhart et al. (2003) state that in the case of emerging markets the risk of a credit event starts to increase significantly at levels of external debt above 30-35% of GNP already. Cases of external debt default have occurred at a ratio of debt to GNP as low as 47% in the case of Mexico in

1982, at a ratio of slightly above 50% in the case of Argentina in 2001 or 20% in Ecuador in 2008 (Reinhart and Rogoff (2009)). All this evidence suggests that cross-country heterogeneity can dramatically affect the empirical analysis.

The language utilized in the debt-and-growth literature implicitly suggests a link going from debt to growth - our contribution is no exception - most probably because such a link seems desirable from an economic policy perspective. But causality remains undetermined. Panizza and Presbitero (2012) using a sample of 17 advanced economies conclude that a causal link going from public debt to economic growth may not exist. And Lof and Malinen (2014), using a panel vector auto-regression (PVAR) model for a panel of 20 developed countries, find no statistically significant effect of debt on growth for any level of debt, while GDP growth is found to have a statistically significant effect on debt. While causality will remain difficult to establish, the robustness of results can be improved by controlling for unobserved heterogeneity.

Data limitations, varying samples of countries across studies, differing time periods or problems with endogeneity or multicollinearity can result in lack of robustness and consequently inconsistent results. Researchers address these problems by delimiting to a core set of growth determinants on the basis of general growth econometrics. Many contributions also work with instrumental variables and use TSLS or GMM models. Some did robustness test as well, estimating their models for various sub-samples - using segmentations on income or development levels, geographical affiliation, membership in international organizations etc. GMM models however are only partially suitable for macroeconomic cross-country datasets as the number of observations is generally relatively small. Kumar and Woo (2010) show that the results obtained with GMM estimations are similar to the results with standard OLS, with OLS coefficients being somewhat smaller than the GMM ones. Reinhart et al. (2012) and Panizza and Presbitero (2013) provide good summaries of the empirical literature and the different approaches used.

Even though recent analysis has produced more robust evidence, there has been an increase in the literature and improvements in econometric methods that allowed for a better understanding of the nature of the relationship, the possibility to derive concrete policy measures and their acceptance is still limited. The empirical evidence has been weaker than generally assumed.

Some recent contributions have claimed attention for misleading estimates of nonlinearity and thresholds by using pooled and panel data without controlling appropriately the heterogeneity between countries. Panizza and Presbitero (2013) demonstrate that aggregate data produces an inverted U-shaped relationship between growth and public debt even though the country-specific regressions generally indicates a U-shaped or even positive relationship between debt and growth. They state: "the relationship is characterized by very large cross country heterogeneity [...]" and "[...] may depend on institutional quality, on the dimension of the public sector, on how and why debt has been accumulated, and on the structure and composition of public debt". Kourtellos et al. (2013) - who also emphasize that not accounting for parameter heterogeneity may be an important limitation of many existing work and lead to spurious results - address parameter heterogeneity by including democracy as a control variable for institutional quality. They specifically find that the relationship between debt and growth depends on the extent of democracy.

Overall however, the problem of parameter heterogeneity has received little attention so far and variables on the institutional environment have widely been ignored, even though several papers have provided evidence that countries with different institutional structures - e.g. economic and systemic arrangements - likely have different debt levels and start facing negative effects from additional debt at different levels. Controlling for institutional quality variables or for macroeconomic structures is therefore an important area for further research.

3 Controlling heterogeneity using institutional quality indicators

Heterogeneity is in the center of any cross-country study. Observed and not observed differences are combined to explain differences in other depending variables. As the independent variables rarely capture the entire cross-country differences, additional control for unobserved heterogeneity is necessary. In case of unobserved heterogeneity, the use of pooled techniques can be misleading and some authors therefore suggest to not utilize it (Panizza and Presbitero, 2013). On the other hand, switching to an individual country-specific analysis is usually not feasible, due to insufficient observations in the individual country sample. For this reason, the use panel data became the most common method to investigate empirically the correlation between public debt and growth at macroeconomic level. Fixed or random effects models have been commonly used to isolate the effect of heterogeneity in order to compute the debt coefficient in a consistent and efficient way. These traditional fixed or random effects help to isolate the cross-country heterogeneity and to estimate more significant and efficient debt regressors. But they do not help to explain the origins of the heterogeneity. Although the heterogeneity is isolated in different country's intercepts with fixed or random effects, the problem of parameter heterogeneity remains.

Another predominant practice has been to group countries according to some pre-defined criteria, generally default definitions like OECD, Asian or Euro area countries for example, without motivating the selection. A common characteristic to classify countries - apart from geographical ones - is income per capita. Well known examples include the Worldbank classification distinguishing low, medium and high-income countries. The existing empirical literature on the relationship between debt and growth provides evidence for different debt-growth relationships in emerging market economies as opposed to advanced economies. Income-based classifications can possibly be useful, of course, depending on the purpose at stake, but cannot explain or disentangle the existence of heterogeneity. At best, these criteria allow to compute more significant and efficient regressors, without explaining the origins of heterogeneity and its effects over the debt-growth relationship. The heterogeneity itself and its consequences on the relationship between debt and growth continue unknown. An explicit control for heterogeneity variables, e.g. using methods for grouping such as cluster analysis, can overcome this problem and therefore improve the results. The advantage will be illustrated in the next sections.

We further argue that income-based classification criteria alone do not capture well the driving factors behind a country's fiscal policy. A country's institutional setting may be a key factor explaining the design and reliability of its fiscal policy. Kourtellos et al. (2013) provide evidence in this direction. They find higher public debt to have a negative effect on growth in low-democracy regimes, whereas in high-democracy regimes no statistically significant relationship can be found.

There are several reasons why institutional quality and macroeconomic stability may shape the relationship between debt and growth. Markets can be more sensitive to debt increases and changes in fiscal and economic policy in the cases of countries with weak institutions and/or with unstable macroeconomic environment. A variety of channels may connect the institutional and macroeconomic environment to the relationship between debt and growth. In case of weak institutional quality:

- Uncertainty may generally be higher.

- Interest rates may react more strongly as lenders request a higher risk premium.

- Crowding-out effects may be more pronounced.

- Pressure to pursue an expansionary monetary policy may be more accentuated due to lower central bank autonomy, leading to an increase in inflation. Reinhart and Rogoff (2012) show

that high public debt levels empirically coincide with higher inflation in emerging market countries, whereas in advanced economies it generally does not.

- Political and economic power may be less balanced - favoring some sectors at the detriment of others - leading to public and social costs through inefficient allocation.

- Tax revenue needed to cover debt and interest rate payments may have a more distortionary and consequently more negative impact as taxation is supposed to be more difficult in case of countries with weak institutions. Tax evasion is supposed to be high, the tax base generally small and the collection of revenues more costly and control-intensive, leading governments to develop tax systems that are inefficient from an economic perspective and exploit whatever they can (Tanzi and Zee (2000)).

- Corruption may be more prevalent and part of the money will dissipate.

Making a long story short, the quality of the prevailing institutions may influence the "return on investment" of (additional) public expenditure and public debt.

The Worldwide Governance Indicators

To capture the effect of governance and institutional quality, we refer to the Worldwide Governance Indicators, WGI. We briefly resume the main characteristics of the indicators below which are well suited to control heterogeneity and help to estimate threshold models appropriately. For a complete methodology of the WGI dataset see Kaufmann et al. (2010) and Kaufmann et al. (1999b).

The WGI summarize the multiple views on the quality of governance. It has been developed with the support of the World Bank in 1999. The database covers more than 200 countries and compiles several hundred individual variables on governance taken from 32 data sources. The WGI are based exclusively on subjective or perceptive measures of governance, an approach that is justified first by the fact that this kind of information (perceptions, impressions, views) is what drives agents' actions, second because there are no objective measures available in many areas of governance, and third because the distinction between "subjective" and "objective" measures is often less clear-cut than it might seem. The data are gathered from a number of survey institutes, think tanks, non-governmental organizations, international organizations, and private sector firms.

Six aggregate indicators are derived using a statistical tool known as unobserved components model (UCM)[5] and available for a period of 1996-2013.[6] The full WGI dataset is available at www.govindicators.org.

The six indicators refer to three aspects of governance (Kaufmann et al., 2010):

1. The process by which governments are selected, monitored and replaced;

    - *Voice and Accountability* - capturing the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, association and media;

    - *Political Stability and Absence of Violence* - capturing the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically motivated violence and terrorism;

2. The capacity of the government to effectively formulate and implement sound policies;

---

[5]See Kaufmann et al. (1999a) for a detailed discussion of the aggregation methodology used.

[6]Between 1996 and 2002 the indicators are available every second year only. As from 2002, they are updated on a yearly basis.

- *Government Effectiveness* - capturing the quality of public services, the quality of the civil service and its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies;

- *Regulatory Quality* - capturing the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development;

3. The respect of citizens and the state for the institutions.

- *Rule of Law* - capturing the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police and the courts, as well as the likelihood of crime and violence;

- *Control of Corruption* - capturing the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.

As any synthesizing indicator, also the WGI database has its strengths and weaknesses. Main critics involve scaling (the data always relies on observed characteristics, i.e. the world average is the same in each year) or the source and quality of the underlying perception data which may show correlated errors, be biased towards perceptions of the elite or influenced by the recent performance of the country. Kaufmann et al. (2007) respond to these critics, which we acknowledge but consider to be of minor importance and generally inevitable.

The alternative would be to rely on objective or *de jure* measures. The democracy variable from the "Polity IV Project" database used by Kourtellos et al. (2013) as a proxy for institutional quality is based on *de jure* criteria. The major disadvantages of *de jure* measures are that they contain limited information about peoples' real incentives and show little variance especially for advanced economies in the recent past. We consider these drawbacks to be more substantial.

Finally, the WGI data are convenient to use in cluster analysis as all of them are quantitative variables and are already standardized.

4   Cluster Analysis

We classify the countries according to their institutional quality measured by the WGI indicators and utilize cluster technique to group them into more homogeneous clusters. Thereafter the resulting sub-samples can be used in panel regressions. The two-stage procedure allows to explore at the same time the pooled data (increasing the degree of freedom of the model), estimate fixed or random effects and control for institutional variables that account for heterogeneity.

Cluster analysis is a major technique to group heterogeneous individuals by finding relatively homogeneous groups based on multivariate characteristics. The cluster methodology is frequently used in medical research, in biology, in marketing and many other social sciences, but still less common in economics.

Some recent research has applied clustering techniques to cross-country data to isolate heterogeneity. Zielenkiewicz (2014) and Unlukaplan (2011) both apply cluster methods to WGI data to analyse the socio-economic development process. But both papers restrict to EU countries. Vázquez and Sumner (2012) and Aubyn (2007) use cluster techniques to identify cross-country groups on a global scale, Aubyn (2007) uses data from the World Economic Forum, Vázquez and Sumner (2012) create their own database based on various sources.

Cluster analysis groups individuals in order to maximize within-group similarity while at the same time maximizing dissimilarity between groups, based on a set of distinct variables

$X_p = \{X_1, ..., X_P\}$ characterizing each individual $i = \{1, ...j, ..., n\}$. In principle, cluster analysis does not start from a theoretical argument and does not even make an explanation of why a cluster exists. Cluster are based on a distance measuring in $\mathbb{R}^P$ space. Nevertheless it is important to have a conceptual basis to explain why groups might exist.

There are three important decisions in any cluster analysis: the determination of (1) the clustering variables, (2) the dissimilarity or distance measure, and (3) where to stop the agglomerating of the individuals or decide how many group must exist. In contrast to other multivariate techniques, the clustering variables are specified by the researcher. The focus of a cluster analysis lies on the comparison of the individual distances, not on the variables.

We argue that the governance indicators are appropriate variables to account for institutional weakness and the quality of the prevailing institutions in turn influences the "return on investment" of public expenditure and public debt, as explained in the previous section. Clusters based on the quality of governance may be an important differential, therefore, in the relationship between debt and economic growth.

Similarity or dissimilarity between the different countries is determined on the basis of the selected variables and only refer to the P variables entering the cluster analysis, i.e. the WGI in our case. The hypothesis that the resulting groups show similar or dissimilar behaviour regarding variables not entering the grouping (e.g. government debt or GDP growth) remains purely speculative and has to be detected separately. This is convenient for our purpose, as the cluster analysis does neither anticipate the result nor interfere with the following analysis of the relationship between debt and growth. We use all six dimensions covered by WGI as cluster variables and thereby assume that countries are similar if they show similar institutional quality. Using only the WGI indicator to cluster the countries is an advantageous strategy, as we can isolate the variables used to observe heterogeneity from that macroeconomic variables used in the economic model.

Several clustering algorithms or methods and measures of dissimilarity exist. Since the number of countries is fixed and we do not know the number of clusters that will emerge, we use the Ward (1963) hierarchical method which directly bases upon the squared Euclidean distance measure. The Ward method takes the minimization of the variance of the sum of all squared Euclidean distances before and after fusion in a certain step k as criterion for fusion.

The Euclidean distance and squared Euclidean distance are amongst the most frequently used distance measures in cluster analysis. It is defined as the distance between two observations (i and j) which corresponds to the sum of the squares of the differences between observations i and j for all variables X_p. The squared Euclidean distance is given by equation (1):

$$d_{ij} = \sum_{p=1}^{P}(X_{ip} - X_{jp})'(X_{ip} - X_{jp}) \quad \text{where} \quad i = \{1, ..., j, ..., n\} \tag{1}$$

where i, j identify the individual (country), $p = \{1, ..., P\}$ is the number of characteristic variable used, d_{ij} is a matrix with nxn dimension and n is the number of cross-section units or countries in our case.

A Ward's Euclidean distance matrix (nxP) in equation (2) is computed in each step k in order to determine the minimum variance. It is slightly different from the Euclidean i, j distance in equation (1) as it is computed against the average of the existing groups $\bar{X}_p$.

$$d_{Ward}^2 = \sum_{p=1}^{P}\left(X_{ip} - \bar{X}_p\right)'\left(X_{ip} - \bar{X}_p\right) \tag{2}$$

where d_{Ward}^2 is a column vector with n elements containing the sum of distance between the $p = \{1, ..., P\}$ cluster variables.

The Ward method is a hierarchical agglomerative method. Starting from the bottom, where each individual (country) is a cluster by itself, a recursive procedure runs up and continuously groups individual by individual into an existing cluster. At the top, after the last step, only one big cluster remains, including all the countries.

The procedure is executed $k = 1, ..., n-1$ times. In each step, the procedure takes the minimum distance found in matrix d^2_{Ward} which we can name $d^2_{(ij)}|k$ to emphasize the subscripts. Two cluster will thus be merged according to equation (3) in step k. In sequence, the row and column d^2_{ij} is suppressed, reducing the dimension of the matrix to the next step of the analysis. In each step k a new bigger cluster C_{ij} will emerge merging the closest clusters C_i and C_j, as the follow equation indicates:

$$C_{(ij)k} = C_i \cup C_j|_k = \min\{d^2_{(ij)}|k\} \quad \forall \quad k = 1, ..., n-1 \tag{3}$$

Our clustering variables are first the three-year averages between 2010 and 2012 for all the six dimensions covered by WGI, i.e. voice and accountability, political instability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption. The three year averages avoid giving too much emphasis on one single year. Second, we also include the variation in each of the six WGI dimensions over the whole period 1996-2012. Therefore, the Ward matrix has a dimension nx12.

There is no single objective procedure to determine the correct number of clusters. For hierarchical methods, the change in total heterogeneity as the number of clusters decreases provides an indication. We base our decision on a graphical analysis (cf. dendrograms in figure (1)) and on a sample of commonly used dissimilarity criteria (table (1)).
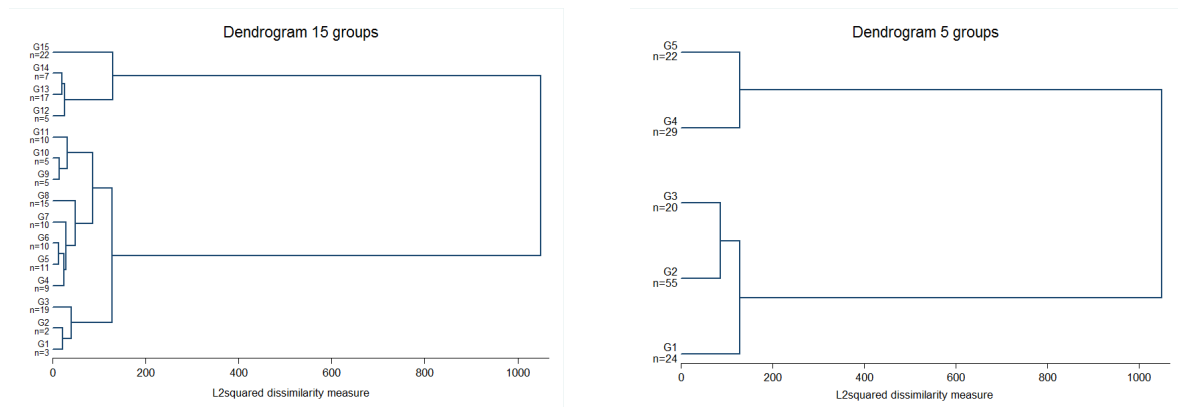
Figure 1: Dendrogram

Table 1: Stopping rules / Measures of dissimilarity

Number of groups	Cluster height	d absolute	d%	Calinski -Harabasz Pseudo-F	Duda-Hart Je(2)/Je(1)	Pseudo T-squared
15	13.019	1.151	0.1	43.17	0.734	5.43
14	14.067	1.048	0.1	44.75	0.711	7.72
13	14.869	0.801	0.1	46.61	0.669	3.96
12	20.377	5.508	0.4	48.14	0.751	7.31
11	22.708	2.331	0.1	49.91	0.316	6.50
10	25.623	2.916	0.1	51.97	0.751	9.29
9	26.506	0.882	0.0	54.81	0.791	7.13
8	29.971	3.466	0.1	58.35	0.811	8.88
7	32.958	2.987	0.1	63.11	0.714	7.21
6	41.165	8.208	0.2	69.08	0.665	11.09
5	49.620	8.454	0.2	77.56	0.802	13.10
4	86.163	36.543	0.7	86.27	0.809	17.20
3	128.651	42.488	0.5	100.01	0.817	21.72
2	129.429	0.778	0.0	155.95	0.557	38.99
1	1048.797	919.367	7.1	0.00	0.487	155.95

We decide to work with five country clusters. The Calinski-Harabasz test does not provide a decision criterion in our case, as it does not show a maximum[7]. In case of the Duda-Hard stopping rule, the strategy is to identify one of the largest Je(2)/Je(1) values that corresponds to a low pseudo-T-squared value. Larger values for Je(2)/Je(1) statistics indicate more distinct clustering, whereas at the same time smaller pseudo-T-squared values indicate more distinct clustering. Based on the Duda-Hard index alone, we would probably select to work with 8 groups then. However, as the values for 5 clusters are fairly good and Calinski-Harabasz as well as practical reasons suggest to work with a smaller number of clusters, we decide in favour of 5 groups. These groups are used subsequently to split up the data on debt and growth into different sub-samples.

The analysis of the variance of the clustering variables between the detected clusters demonstrates that the model is well-shaped as is displayed in table (2). All clusters are significantly different across the WGI variables. Furthermore, the variance with regard to GDP growth and government debt also provides first evidence for our hypothesis on the impact of institutional quality on GDP growth and government debt. The different country clusters show significantly different growth and government debt levels.

Table 2: Variance across clustering variables, GDP growth and government debt

| Variable[8]	F	Prob $>F$	$R2$
WGI Corruption	212.55	0.0000	0.8543
WGI Government Effectiveness	206.39	0.0000	0.8506
WGI Rule of Law	271.06	0.0000	0.8820
WGI Political Stability	56.95	0.0000	0.6111
WGI Regulatory Quality	187.43	0.0000	0.8379
WGI Voice and Accountability	84.95	0.0000	0.7009
GDP Growth	8.60	0.0000	0.1917
Government Debt	5.78	0.0002	0.1376

Table (3) and figure (2) show the within-group mean values for the WGI variables. Based on this descriptive analysis of the cluster characteristics we denominate the groups as poor, low, medium, good and top institutional quality.

---

[7]According to Calinski-Harabasz, the fact that F statistics increases monotonically signifies that there is no

Table 3: Country cluster: Within-group mean value

Governance Quality Cluster	Control of corruption	Government effectiveness	Rule of law	Political instability and absence of violence	Regulatory quality	Voice and accountability
1 Poor	-1.106766	-1.165616	-1.212473	-1.22652	-1.205016	-1.235357
2 Low	-.5868858	-.4793372	-.5641854	-.4730593	-.4104204	-.5330608
3 Medium	-.3019282	-.1638614	-.3579136	-.2857403	.0954569	-.0246447
4 Good	.6508028	.7422237	.680869	.6496806	.6689776	.5576136
5 Top	1.889143	1.71387	1.704357	.9977463	1.590965	1.319288
Total	-.0081726	.0374739	-.0441902	-.1107685	.0667941	-.0628122

Figure 2: Cluster characteristics
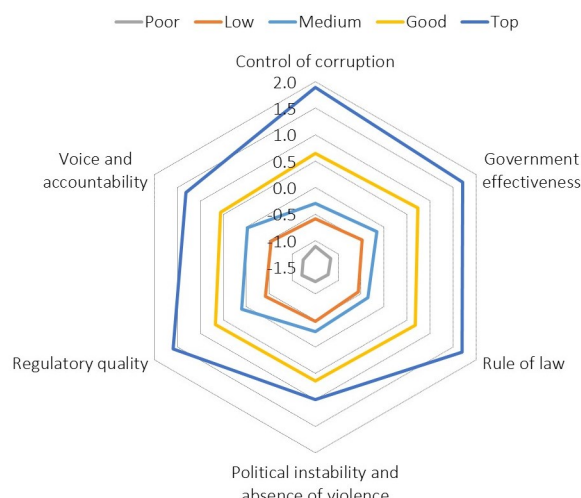
Cluster analysis has some major deficiencies of course. Adding a new individual, or a new country in our case, can significantly change the resulting groups and a country that belonged to one group may change to another. Cluster analysis is therefore unsuitable for application in sample studies, where the number of cross-section units varies. This restriction does not affect our case, as we are working with a complete and predetermined set of countries. Another critical point is that the grouping variables are determined discretionarily by the researcher and not due to a corresponding previous analysis. In addition, Nielsen (2011) criticizes the large degree of freedom in choosing among alternative distance measures and cluster algorithms, and the difficulty to periodically update the taxonomy.[9]

5 Models on the relationship between debt and growth

Given heterogeneity across countries, we investigate how the detected differences in institutional characteristics help to explain the relationship between debt and growth. We use simple and threshold multiple regression models (TMR), adapting the threshold autoregressive models developed by Tong (1978, 1983). We start with some descriptive statistics of the data. To make

---

natural partition of the data.

[8]WGI indicators three-year averages 2010-2012, GDP growth and government debt in 2012.

[9]Nielsen (2011) develops an alternative approach to overcome this deficiencies, noting lack of clarity with regard to the underlying rationale of the existing UNDP, World Bank, and IMF classification. Nielsen's approach is based on Lorenz curves. In his classification methodology a discretionary number of country groups are built in a way that all groups show the same error between the actual and an even distribution within each group.

the procedure as transparent as possible, we then perform estimations for a linear model without threshold comparing various estimation methodologies such as pooled OLS, between estimator (BE)[10], fixed effects (FE), two-stage-least-square (TSLS) and general method of moments (GMM) in a panel framework. Regressions are first computed without clustering the countries. In sequence, we proceed with the analysis for the sub-samples previously determined by cluster analysis using GMM estimator with period and country fixed effects to compute linear models, and finally go on to nonlinear sample-split models with exogenous threshold. The analysis focuses on the observed heterogeneity across clusters and controls remaining unobserved heterogeneity of macroeconomic variables across countries by fixed effect estimator.

5.1 Data and descriptive statistics

We analyse data for 154 countries covering the period 1990-2012 and refrain from using a longer time span in order to exclude structural changes and to get more meaningful results for today's economy. The key variable, gross government debt-to-GDP ratio, is compiled by the IMF for a large cross-section of countries and a long time-period.[11]. The other variables - GDP growth as well as control and instrumental variables - originate from the World Development Indicators (WDI) database.[12] Time-series on inflation as measured by consumer prices are taken from the IMF World Economic Outlook database, as these series are more complete.

The boxplot in figure (3) shows the distribution of our key variables GDP growth and government debt for the five country clusters. Considerable differences exist depending on the level of institutional quality. Better institutional quality (e.g. cluster 5 compared to cluster 1-4) seems to go hand in hand with increasing debt but also with decreasing growth. However, the overlap between the groups is relatively large. Within-group variation is considerable and displays a non-continuous development with regard to the institutional quality dimension. Within-group variation in GDP growth is typically larger than the variation of government debt except for cluster 5 whose members show very little variation with regard to economic growth.

Figure 3: Distribution of GDP growth and government debt by country cluster
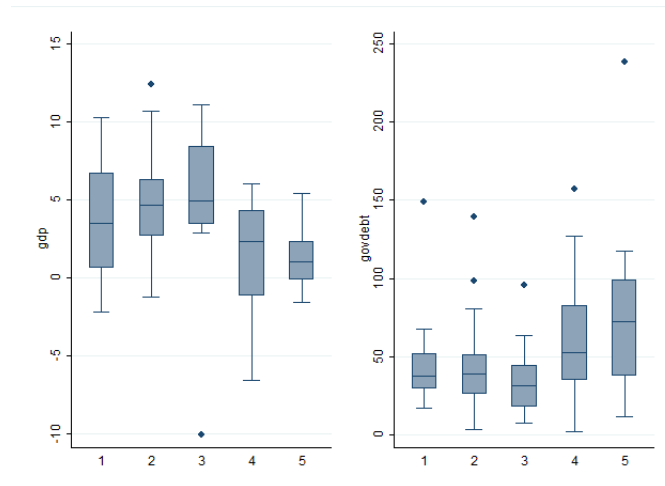
In the subsequent analysis of the relationship between debt and growth, we check for a set of control variables, using adjusted $R2$, the Sum of Squared Residuals (SSR), F statistic and

---

[10]Cross-section estimation using time-averages of the regressors.

[11]Historical Public Debt Database of the IMF Fiscal Affairs Department (version Fall 2013 from 30.01.2014, http://www.imf.org/external/pubs/ft/wp/2010/data/wp10245.zip) For details on data sources and methodology see Abbas et al. (2010).

[12]The WDI is compiled by the World Bank. We work with the version updated on December 19, 2014.

J-statistic as a criteria to select the control variables. As the panel is not balanced, we refrain from including desirable variables in a few cases, e.g. in case of education or interest rates. The inclusion of these variables would result in a massive loss of observations. The resulting list of control variables is given below:

- *Investment*, % of GDP - measured as gross capital formation (formerly gross domestic investment) consists of outlays on additions to the fixed assets of the economy plus net changes in the level of inventories;

- *Trade openness*, % of GDP - (imports+exports)/GDP;

- *Inflation*, annual % - consumer prices;

- *GDP per capita*, in current USD;

- *Government consumption*, % of GDP - to approximate the share of public sector production in contrast to private sector production;

- *Government consumption growth*, annual %;

- *Population growth*, annual %.

Table (4) gives some descriptive statistics - mean values and standard deviation - for the control variables for each of the sub-samples as well as the entire sample of 154 countries[13].

Table 4: Descriptive statistics - Within-cluster mean value (1st line) and standard deviation (2nd line)

	1 - Poor	2 - Low	3 - Medium	4 - Good	5 - Top	All
Gov. Debt	71.5162	63.8799	53.3040	52.5781	59.5276	60.5864
	44.4958	42.5059	36.3730	34.2911	34.5067	40.5187
GDP Growth	4.0348	4.3446	4.1593	3.2788	2.6704	3.7789
	5.2028	4.1062	4.4500	3.9630	2.8605	4.1946
Investment	22.2823	24.0019	21.2317	24.7666	22.7829	23.3159
	16.5222	9.2148	6.2182	6.9465	3.6010	9.1865
Trade	63.8466	81.5303	69.4427	92.6030	102.9502	84.3741
	41.7442	36.4760	35.4208	38.6663	90.6683	49.8435
Inflation	99.9461	22.7433	85.4903	6.2177	2.4781	31.9960
	1253.6870	185.3303	520.3240	9.5439	2.4510	466.9269
GDPpcUSD	1342.1548	2691.6720	3126.1883	13339.3267	35256.5514	9212.3583
	2977.5814	5052.3163	2877.7577	11581.5348	17358.8175	14348.3405
Gov. Cons.	12.4751	14.0044	15.1500	18.2414	18.5570	15.8871
	5.1199	5.8985	5.6283	5.4266	4.6364	6.0884
Gov. Cons. Growth	10.1958	5.3967	3.6533	3.1281	2.4140	4.5017
	57.0100	17.8514	9.8062	7.1212	2.4942	20.6215
Population Growth	2.3759	1.9089	1.0375	1.2610	0.8288	1.5200
	0.8438	1.2133	1.6473	2.2534	0.7036	1.5042

In the next sections we estimate a variety of models using aggregated and clustered data applying linear and nonlinear models to empirically analyse the dynamic between growth and public debt.

5.2 Model without cluster and without threshold

The first model is the more aggregated and general one. We compare several estimation methods, pooled, between estimator (BE), fixed effect (FE) computed by OLS, TSLS estimation as well as

---

[13]Depending on the availability of data the total number of countries may vary between 154 and 176.

generalized moments (GMM), all of them applied to the entire data sample without clustering and without threshold. Estimations of this kind are frequently reported in the literature. The results given in table (5) can thus be contrasted to a multitude of existing results by various authors. An auto-regression term of GDP growth in $t-1$ is included to avoid potential correlation between the endogenous variable and the error term, thereby converting the model into a dynamic panel.

Our base model takes the following form:

$$y_{it} = \alpha + \lambda D_{it} + \beta y_{it-1} + \mathbf{X}'_{it}\boldsymbol{\gamma} + \mu_t + \nu_i + \epsilon_{it} \tag{4}$$

where $D_{i,t}$ is the variable for government debt, $\mathbf{X}_{it\mathrm{x}K}$ depicts the matrix of control variables $\{1,...,K\}$ and μ_t and ν_i are the terms for period and country fixed effects.

In the case of TSLS and GMM, instrumental variables are added to the model, which assume the following forms:

$$y_{it} = \alpha + \lambda D_{it} + \beta y_{it-1} + \mathbf{X}'_{it}\boldsymbol{\gamma} + \mathbf{Z}'_{it}\boldsymbol{\delta} + \mu_t + \nu_i + \epsilon_{it} \tag{5}$$

where $\mathbf{Z}_{it}$ is a matrix of dimension $\mathbf{Z}_{it\mathrm{x}L}$ corresponding to $\{1,...,L\}$ instrumental variables and $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are a $L\mathrm{x}1$ column vector of the coefficients.

Multiple regression models of this kind are a typical starting point in the empirical literature on public debt and growth. The list of control variables may differ somewhat. Almost always price level, interest rates and the business cycle need to be adjusted, however. Price level, because the real value of the debt is, for most purposes, more important than the nominal value - in our case, as debt data is already PPPGDP adjusted, this is not necessary; interest rate, to capture the market value of the debt - we refrain from including it due to many missing values; and business cycle, because the deficit rises automatically when economic activity slows and vice versa. The short-term business cycle fluctuation is often captured using five-year-ahead overlapping averages of GDP growth rates. As five-year averages also level out variances that are not attributable to business cycles, we prefer to stick to the annual data and alternatively decide to include an auto-regressive term for GDP growth. Additionally, where specified, period and country fixed effects are controlled.

The model formalized in equation (4) is computed by pooled and OLS panel and equation (5) by instrumental variables techniques. In the latter case two-stage-least-square (TSLS) and generalized method of moment (GMM)[14] estimators are used. Our results are consistent with the existing empirical literature, despite the differences with regard to the sample, covering a large number of countries and unbalanced data between 1990 to 2012, and the prevailing differences in the selected control variables. All the estimators in table (5) compute a negative coefficient of *GovDebt* but vary based on consistence and efficiency reasons, resulting from omitted and endogenous variables. The global government debt coefficient is negative in the range of -0.008 to -0.0518 depending on the estimation methodology used. Similar result were found by Kourtellos et al. (2013), Cecchetti et al. (2011), Kumar and Woo (2010) and many others.

Economic growth and government debt are both highly aggregated macroeconomic variables and therefore subject to many influences from a multitude of other variables, and even more than this, in a complex temporal dynamic. For this reason, estimating a complete model with exogenous variables is an almost impossible mission, and the best option is to use instrumental variables techniques. As the results given in table (5) confirm, the coefficient of *GovDebt* significantly improves in comparison to the traditional pooled, between and fixed effects models in the case of instrumental variables techniques (TSLS and GMM). Significant difference exists between the TSLS and GMM estimators. The SSR statistics is not useful to select the best estimator in this case. We decide in favour of GMM estimator because the significance of all

---

[14]Since the model is a dynamic and linear one, the Arellano Bond two step estimator or White period system covariances is applied.

Table 5: Models without cluster and without threshold

Regressors	Pooled	BE	FE	FE+TE	TSLS	GMM
Cross-section		141	154	154	154	147
Total Obs (unbalanced)	2782	141	2665	2665	2504	2193
C	2.2454	2.4035	3.2727	2.9468	3.0957	
	0.0000	0.0004	0.0000	0.0000	0.0053	
GOVDEBT	-0.0080	-0.0090	-0.0127	-0.0137	-0.0149	-0.0518
	0.0000	0.0377	0.0000	0.0000	0.0001	0.0006
GDP(-1)	0.2899		0.1889	0.2111	0.1447	-0.0892
	0.0000		0.0000	0.0000	0.2156	0.1948
INVEST	0.0684	0.0885	0.0809	0.0777	0.0413	-0.1917
	0.0000	0.0000	0.0000	0.0000	0.1848	0.0678
TRADE	0.0038	0.0044	0.0198	0.0121	0.0240	0.5259
	0.0043	0.1136	0.0000	0.0040	0.0005	0.0000
INFLATION	-0.0005	-0.0009	-0.0003	-0.0003	-0.0071	-0.0281
	0.0005	0.2709	0.0166	0.0251	0.1520	0.0068
GDPPCUSD	0.0000	0.0000	-0.0001	0.0000	-0.0001	-0.0008
	0.0000	0.0003	0.0000	0.0006	0.0000	0.0006
GOVCONS	-0.0750	-0.0505	-0.1843	-0.1272	-0.0833	-2.0021
	0.0000	0.0529	0.0000	0.0000	0.2387	0.0000
GOVCONSGR	0.0149	-0.0017	0.0140	0.0178	-0.0063	0.0130
	0.0000	0.8742	0.0000	0.0000	0.7362	0.0745
POP	0.2981	0.5069	0.3400	0.3335	0.4218	0.3386
	0.0000	0.0000	0.0000	0.0000	0.0012	0.7889
R2 Adj	0.2477	0.3816	0.3005	0.3971	0.3672	
SSR	33496.0	305.6	27307.7	23337.8	22398.8	106578.7
F	102.7491	11.8007	8.0644	10.5894	8.2852	
AIC	5.3333	3.7391	5.2872	5.1458		
DW	1.9661	0.0000	1.9825	2.0121	1.8955	
J-Statistic						13.7484
J-Prob						0.1316

coefficients is higher comparing with TSLS. This result is important for the re-estimation of the model considering clustered data. We will do this in the next sections using the GMM estimator.

As Kumar and Woo (2010) describe in detail, the pooled and BE estimations suffer from many biases - both due to omitted-variables and measurement errors. They argue however that other things equal, these measurement errors tend to reduce the correlation between the regressors and the country fixed effects, and therefore reduce the heterogeneity bias. Controlling for fixed-effects across countries and over the time (FE and FE+TE) addresses the problem of the omitted-variables bias but exacerbates the measurement error problem. However, TSLS and GMM would at least theoretically and as long as instruments are valid, i.e. uncorrelated with the errors, address the omitted-variables bias, endogeneity, and measurement errors. Strong and valid instruments are difficult to encounter. Yet a further and major limiting factor of GMM models is their infeasibility for the analysis of small samples.

For the analysis of macro data GMM models should be generally used with caution. The main motive is the difficulty to find appropriate instruments, mainly in the case of multiple regression model, as in our case. The common practice in macroeconomics with regard to instrumental variables is the use of lagged values of endogenous variables as instruments. We use two lagged observations of each potential endogenous variables as instruments.

5.3 Linear model with cluster without threshold

We apply GMM estimation controlling for period and country fixed effects to the sub-samples previously determined by cluster analysis. Table (6) lists the respective results for all of the five country groups in comparison to the entire sample. The debt coefficient is positive for cluster

1 (poor institutional quality countries) and negative in all other cases. Nonetheless, the size and significance of the impact of the debt coefficient varies considerably, ranging from -0.045 for cluster 2 to -0.159 for cluster 3, which indicates an impact that is about three times higher. Debt coefficients for clusters 1, 3 and 4, are considerably higher than for the overall estimation (column 6). These results confirm the advantage of observed heterogeneity controlled by institutions, as suggested in this paper. Controlling the unobserved heterogeneity by cross-country fixed effect in panel data is not sufficient to capture the appropriate correlation between debt and growth.

It is interesting to note that the institutional quality heterogeneity does not display a monotonically increasing or decreasing correlation regarding the relationship between growth and debt. On the contrary. The results indicate that in addition to a possible within-group quadratic or otherwise nonlinear relation between debt and growth (which we did not investigate so far), also the influence of institutional quality on the relationship between debt and growth is highly complex. Further research on the channel between institutional quality and the debt-and-growth relationship is therefore certainly indicated.

An interesting linkage to explore might be the propensity of changes in institutional quality or the reliability of the persistence of a country's institutional setting in periods of crisis, which might be higher for some groups than for others. Countries figuring in medium (3) and good (4) governance quality environments could be more difficult to classify from a political credibility point of view as countries in group (5). They might be more suspect to succumb to populist policy measures, to experience a higher risk of institutional quality changes and therefore suffer a more negative impact from additional debt. At the both extremes, poor and top, there are little doubt about the permanence of institutional environment and this, in a certain way, remains priced. In the middle, the institutional quality may be more prone to change. A possible starting point to test this hypotheses could be a detailed investigation of the variation in the institutional quality variables or the propensity of changes in credit market ratings. Risks of downgrading are typically very low for countries with top institutional governance as their credibility is scarcely called into question. On the other hand, the positive coefficient in cluster 1 can be theoretically justified by myopia or very strong preferences for present public expenditures given the uncertain future of these countries.

if we consider that in the poor institutional countries, each public cents help to burst the growth because the per capita income or productivity as a hole is equally so low that even a bad allocation of public resources has greater returns than the low productivity observed abroad, overcoming the negative effects.

The coefficients of the control variables generally have the expected sign and are in most cases statistically significant. It is interesting to note that a higher level of government consumption is associated with lower growth in all institutional conditions. This means that a higher share of public consumption is generally supposed to be negative e.g. because of lower efficiency and high risk of persistence. Surprisingly however, also government consumption growth shows a negative sign in most cases but for cluster 2 and for the insignificant coefficient in cluster 3. This suggests that successful short-term fiscal stimulus may be a very challenging policy in most situations. Of special interest is the comparison of the goodness of fit of the sub-sample model estimations compared to the estimation for the entire sample. For all sub-samples but the medium one, the models are clearly superior to the overall estimation. We can check this dividing the total SSR by the number of observations. This substantiates the relevance of institutional quality as a parameter for the debt and growth relationship and confirms the suitability of the variables to control for unobserved heterogeneity. Further analysis to investigate other existing sources of unobserved heterogeneity and sources of the nonlinearity in institutional quality are still needed.

Table 6: Linear model with cluster without threshold - estimated with GMM

Cluster	1 - Poor	2 - Low	3 - Medium	4 - Good	5 - Top	All
Cross-Section	16	51	19	25	22	147
Total Obs (Unbalanced)	175	792	270	389	401	2193
GOVDEBT	0.0317	-0.0451	-0.1596	-0.1214	-0.0692	-0.0518
	0.0043	0.0460	0.0069	0.0007	0.2159	0.0006
GDP(-1)	-0.0902	-0.2910	0.0483	-0.0820	-0.3710	-0.0892
	0.0459	0.0002	0.5940	0.2262	0.0013	0.1948
INVEST	-0.0154	0.2082	-1.0323	-0.3961	-0.0059	-0.1917
	0.5997	0.4193	0.0236	0.0004	0.9888	0.0678
TRADE	0.0319	-0.0274	-0.7285	0.2369	0.2180	0.5259
	0.1093	0.8741	0.0024	0.0000	0.0001	0.0000
INFLATION	0.0011	-0.2372	-0.0018	-0.4972	-0.0488	-0.0281
	0.5915	0.0003	0.5870	0.0000	0.8902	0.0068
GDPPCUSD	0.0011	-0.0025	-0.0071	-0.0003	0.0002	-0.0008
	0.0012	0.0318	0.0000	0.2144	0.3129	0.0006
GOVCONS	-1.3704	-3.0779	-0.5398	-2.3428	-2.9021	-2.0021
	0.0000	0.0000	0.3582	0.0000	0.0076	0.0000
GOVCONSGR	-0.0090	0.0262	-0.0026	-0.0773	-0.4647	0.0130
	0.0000	0.0412	0.9464	0.0360	0.0014	0.0745
POP	3.6266	-1.4322	-5.4399	0.3196	-7.2599	0.3386
	0.0014	0.4300	0.0052	0.5657	0.0015	0.7889
SSR	4787.5	29308.7	46416.4	9198.4	6413.7	106578.7
SSR/Obs	27.4	37.0	171.9	23.6	16.0	48.6
J-Statistic	5.1949	10.1436	10.4416	18.5878	16.4260	13.7484
J-Prob	0.6362	0.3390	0.3159	0.0289	0.0585	0.1316

5.4 Nonlinear model with cluster and exogenous threshold

As a next step we investigate the existence of a nonlinear relationship within each of the sub-samples. The common practice in debt and growth literature follows the methodology developed by Caner and Hansen (2004), that is using a sample split function of the form (6):

$$y_{it} = \alpha + \lambda_1 d1_{it} * D_{it} + \lambda_2 d2_{it} * D_{it} + \beta y_{it-1} + \mathbf{X}'_{it}\boldsymbol{\gamma} + \mathbf{Z}'_{it}\boldsymbol{\delta} + \mu_t + \nu_i + \epsilon_{it} \qquad (6)$$

where $d1$ and $d2$ are dummy variables taking on the value 1 if debt is above a given threshold d^* in case of $d1$, or if debt is below a given threshold d^* in case of $d2$.

We exogenously determine the threshold at a level of 90% as this level has been of special importance in the economic literature and political discussion and estimate the threshold model (6) using GMM estimator.[15] The number of observations above and below the threshold varies considerable and ranges from 10% above vs. 90% below in case of the medium institutional quality cluster and 31% above vs. 69% below in case of the countries with poor institutional quality. Table (7) depicts the respective results for all sub-samples.

The results are very similar to the findings from the previous linear regression without threshold and the quality of the models is practically identical. The estimated influence of government debt on growth above the determined threshold differs only slightly from the estimated coefficient below threshold.

Altogether the result suggest that the existence of a government debt threshold at a level of 90% of GDP cannot be confirmed. This outcome should not be misinterpreted in the sense that no threshold exists or the relationship is better approximated by a linear model indeed. In fact, to exogenously set a threshold at 90% is too short sighted and a mere attempt to test an existing presumption. In reality, the level of the threshold might vary between the sub-samples and

---

[15]We refrain from testing different threshold levels by two simple reason: we focus on the observed heterogeneity rather than in threshold options and by restriction of space.

Table 7: Nonlinear model with cluster and exogenous threshold at 90% - estimated with GMM

Cluster	1 - Poor	2 - Low	3 - Medium	4 - Good	5 - Top	All
Cross-Section	16	51	19	25	22	147
Total Obs (Unbalanced)	175	792	270	389	401	2193
D_1DEBT	0.0325	-0.0460	-0.1481	-0.1206	-0.0583	-0.0488
	0.0087	0.0311	0.0045	0.0004	0.0609	0.0014
D_2DEBT	0.0211	-0.0415	-0.2938	-0.1062	-0.0874	-0.0636
	0.1094	0.1439	0.0003	0.0017	0.0286	0.0003
GDP(-1)	0.0565	-0.3195	0.0194	-0.0939	-0.3504	-0.0936
	0.6774	0.0000	0.8547	0.1523	0.0013	0.1770
INVEST	0.0058	0.1302	-0.9829	-0.3912	0.0361	-0.1865
	0.8671	0.5480	0.0363	0.0007	0.9334	0.0472
TRADE	0.0359	0.0568	-0.5733	0.2429	0.2257	0.5157
	0.1139	0.6899	0.0000	0.0000	0.0000	0.0000
INFLATION	-0.0026	-0.2558	-0.0031	-0.4638	-0.1150	-0.0245
	0.4887	0.0000	0.2546	0.0000	0.7443	0.0044
GDPPCUSD	0.0009	-0.0023	-0.0079	-0.0002	0.0002	-0.0008
	0.0050	0.0346	0.0000	0.2598	0.3394	0.0004
GOVCONS	-1.6285	-2.9450	-1.3369	-2.3030	-2.7939	-1.9604
	0.0000	0.0000	0.0077	0.0000	0.0079	0.0000
GOVCONSGR	-0.0096	0.0277	0.0350	-0.0735	-0.4648	0.0131
	0.0000	0.0229	0.3570	0.0071	0.0015	0.0336
POP	3.7028	-1.2097	-5.8945	0.2235	-7.2901	0.2078
	0.0006	0.4974	0.0060	0.6066	0.0009	0.8688
SSR	5817.8	28401.7	43240.0	9029.8	6276.5	102774.1
SSR/Obs	33.2	35.9	160.1	23.2	15.7	46.9
J-Statistic	4.6222	11.3410	14.4834	18.8529	16.4955	14.2082
J-Prob	0.5931	0.3316	0.1061	0.0422	0.0863	0.1637

should be assumed to be unknown. Further testing using endogenously determined thresholds are therefore the next step to consider.

6   Concluding remarks

The empirical analysis of the relationship between debt and growth is demanding. Unobserved parameter heterogeneity can considerably bias the results and might be one of the major problems in the existing literature.

Addressing heterogeneity by prepending a cluster analysis to detect similarities in the countries' institutional quality turns out to be a well-suited approach to address this short-coming. More complex estimation techniques and model specifications could potentially account for the problem of unobserved parameter heterogeneity as well. But these kind of models are increasingly difficult to understand and the results are therefore hard to interpret. This makes them difficult to translate to a broader audience (e.g. politicians) and to apply to real world situations. A two-stage procedure, as suggested in this paper, seems more appropriate and has the great advantage to explicitly disclose the heterogeneity between the countries.

Our approach delivers sound and very promising results. The estimated coefficients for government debt vary considerably between the detected country groups and the quality of the estimation is in many cases superior to a combined overall estimation. The results are valid especially for countries that have passed a fair level of institutional quality.

Nevertheless several limitations of the approach still persist and additional investigation on endogenous threshold, on the channel between institutional quality and the debt-and-growth relationship, and on further sources of unobserved heterogeneity are necessary. We are convinced that future research will have to move in this direction.

References

Abbas, S. A., Belhocine, N., ElGanainy, A., and Horton, M. (2010). A historical public debt database. *IMF Working Paper*. No. 10/245.

Afonso, A. and Jalles, J. T. (2013). Growth and productivity: The role of government debt. *International Review of Economics and Finance*, 25:384–407.

Aiginger, K. and Falk, M. (2005). Explaining differences in growth among oecd countries. *Empirica*, 32(1):19–43.

Aubyn, M. S. (2007). Modernising public administration and economic growth. *Working Paper*. Paper presented at the European Central Bank Public Finance Workshop: "Challenges for government spending in the EU", Frankfurt, 6 December 2007.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, 82(6).

Baum, A., Checherita-Westphal, C., and Rother, P. (2013). Debt and growth: New evidence for the euro area. *Journal of International Money and Finance*, 32:809–821.

Caner, M. and Hansen, B. E. (2004). Instrumental variable estimation of a threshold model. *Econometric Theory*, 20:813–843.

Cecchetti, S., Mohanty, M., and Zampolli, F. (2011). Achieving growth amid fiscal imbalances: the real effects of debt. *Proceedings - Economic Policy Symposium - Jackson Hole*, pages 145–196.

Checherita, C. and Rother, P. (2010). The impact of high and growing government debt on economic growth. an empirical investigation for the euro area. *European Central Bank Working Papers*. No. 1237.

Devarajan, S., Swaroop, V., and Zou, H.-F. (1996). The composition of public expenditures and economic growth. *Journal of Monetary Economics*, 37(2):313–344.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. (2005). *Handbook of Economic Growth*, volume 1, chapter 8 - Growth Econometrics, pages 555–677. North Holland.

Elmendorf, D. and Mankiw, N. G. (1999). Government debt. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1C. North-Holland, Amsterdam.

Herndon, T., Ash, M., and Pollin, R. (2013). Does high public debt consistently stie??? economic growth? a critique of reinhart and rogoff. *Working Paper Series*. No. 322.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2007). The worldwide governance indicators project: Answering the critics. *World Bank Policy Research Working Paper*, (4149).

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). The worldwide governance indicators: Methodology and analytical issues. *Policy Research Working Paper*. No. 5430.

Kaufmann, D., Kraay, A., and Zoido-Lobatón, P. (1999a). Aggregating governance indicators. *Policy Research Working Paper*, (2195).

Kaufmann, D., Kraay, A., and Zoido-Lobatón, P. (1999b). Governance matters. *Policy Research Working Paper*, (2196).

Kneller, R., Bleaney, M. F., and Gemmell, N. (1999). Fiscal policy and growth: evidence from oecd countries. *Journal of Public Economics*, 74:171–190.

Kourtellos, A., Stengos, T., and Tan, C. M. (2013). The effect of public debt on growth in multiple regimes. *Journal of Macroeconomics*, 38:35–43.

Kumar, M. S. and Woo, J. (2010). Public debt and growth. *IMF Working Paper*. No. 10/174.

Lof, M. and Malinen, T. (2014). Does sovereign debt weaken economic growth? a panel var analysis. *Economic Letters*, 122:403–407.

Nielsen, L. (2011). Classifications of countries based on their level of development: How it is done and how it could be done. *IMF Working Paper*, (31).

Panizza, U. and Presbitero, A. F. (2012). Public debt and economic growth: Is there a causal effect? *POLIS Working Papers*. No. 168.

Panizza, U. and Presbitero, A. F. (2013). Public debt and economic growth in advanced economies: A survey. *Swiss Society of Economics and Statistics*, 149 (2):175–204.

Pescatori, A., Sandri, D., and Simon, J. (2014). Debt and growth: Is there a magic threshold? *IMF Working Paper*. No. 14/34.

Reinhart, C. M., Reinhart, V. R., and Rogoff, K. S. (2012). Public debt overhangs: Advanced-economy episodes since 1800. *Journal of Economic Perspectives*, 26(3):69–86.

Reinhart, C. M. and Rogoff, K. S. (2009). *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton, NJ.

Reinhart, C. M. and Rogoff, K. S. (2012). Growth in a time of debt. *American Economic Review*, 100(2):573–578.

Reinhart, C. M., Rogoff, K. S., and A., S. M. (2003). Debt intolerance. *National Bureau of Economic Research - Working Paper*. No. 9908.

Sala-i Martin, X. (1997). I just ran two million regressions. *American Economic Review*, 87(2):178–183.

Sala-i Martin, X., Doppelhofer, G., and Miller, R. I. (2004). Determinants of long-term growth: A bayesian averaging of classical estimates (bace) approach. *American Economic Review*, 94(4):813–835.

Tanzi, V. and Zee, H. H. (1996). Fiscal policy and long-run growth. *IMF Working Papers*. No. 96/119.

Tanzi, V. and Zee, H. H. (2000). Tax policy for emerging markets: Developing countries. *Working Paper*. No. 00/35.

Tong, H. (1978). *On a Threshold Model*. Sijhoff & Nordhoff, AmsterdanUSA.

Tong, H. (1983). *Threshold models in non-linear time series analysis*. Springer, New York, USA.

Unlukaplan, I. (2011). Multivariate investigation of governance indicators in european union. *Journal of US-China Public Administration*, 8:66–76.

Vázquez, S. T. and Sumner, A. (2012). Beyond low and middle income countries: What if there were five clusters of developing countries? *IDS Working Paper*, (404/2012).

Ward, Joe H., J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):pp. 236–244.

Zielenkiewicz, M. (2014). Institutional environment in the context of development of sustainable society in the european union countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, (9 (1)):21–37.