TESTING FOR MEAN DIFFERENCES IN EDUCATION SPENDING

Mark Millin*
11 June, 2018

Abstract

This paper does not seek to test Wagner's hypothesis *per se*, but rather uses it to justify the inclusion of the state of economic development of countries (proxied by income per capita) as part of an exploratory inquiry with respect to two level-form ratio measures of total public spending on education, namely the national effort (spending as a percentage of GDP) and budget share (spending as a percentage of total government spending) measures. Using a large sample of countries from 1989 to 2015, a sensibly simple approach to testing – best described as a (conditional) generalised-form t-test using the general linear model – is applied. The method entails a variant of fixed-effects estimation. Instead of estimating panel (country) fixed effects, more aggregated user-defined group fixed effects are estimated, where the different groups effectively comprise the time-invariant sub-samples of interest. Controlling for the state of political democracy and several other important education spending variables, what the empirical findings reveal is a novel way to validate the existence of Wagner's 'law', which can be summarised in the form of three theoretical inequality propositions. These propositions could be applied to other components of public spending to evaluate whether or not the Wagnerian hypothesis can also be validated with respect to other parts of the government's budget allocation.

JEL Codes: H52, I22, I25
Keywords: Education finance, education spending, public spending on education, Wagner's law, fixed effects, LSDV, ANOVA, ANCOVA, GLM, richer countries, poorer countries

* PhD Candidate, Department of Economics, University of Otago, Dunedin.
A paper prepared for the Annual New Zealand Association of Economists (NZAE) Conference, Auckland University of Technology, 27 to 29 June, 2018. I thank my Doctoral Supervisors, David Fielding and Dorian Owen, for their ongoing advice, support and guidance, as well as participants of the Department of Economics Brown Bag Seminar Series at the University of Otago for helpful comments. I gratefully acknowledge the University of Otago's Graduate Research School for the Doctoral Scholarship and the National Research Foundation (NRF) of South Africa for their part funding of my Doctoral research. The views expressed in this paper are my own, and do not in any way represent the views of the NRF. Any errors or omissions are my own.

1. Introduction

The modern-day human-capital revolution and associated empirical research can be traced back to the pioneering work of Mincer (1958), Schultz (1961) and Becker (1962; 1964). Education spending (or education finance) might be considered a subset of this vast field of the economic literature on human capital. Since the largest majority of spending on education in most countries comes from the public sector, a theory linking the growth of the public sector to the growth of the national economy would be a promising framework to explain changes in public spending on education.[1] The now well-known Wagnerian (after Adolph Wagner, 1835-1917) 'law of increasing state activity' (Castles, 1999, p. 2) advocates a clear role for including an economic dimension in the study of education spending patterns. By virtue of their close association, regional differences are an alternative way to distinguish economically richer versus poorer parts of the world. One of the earliest pieces of empirical work that considered both economic and regional factors explaining education spending patterns for various sub-samples of countries was that undertaken by Zymelman (1976). However, since then, there has been a whole branch of literature showing that political forces matter for education spending, too. The study conducted by Verner (1979) is a good example of one of the earliest studies incorporating political factors for a large sample of countries.

However, there is a particular problem that does not seem to be adequately addressed in the contemporary literature. An enduring question that comes to mind can be stated as follows: what do we really know about how two important level-form ratio measures of national-level total public spending on education – the national effort and budget share measures[2] – vary by economic or regional and political categorisation of countries over a relatively long period of time? Therefore, the key research question of interest in this paper asks whether or not there are significant (mean) differences in education spending between richer or poorer *and* politically distinct groups of countries, and, if so, are the empirical patterns of differences generally robust to use of alternative estimators and/or reasonable changes in specification?

Although inquiring about an interesting problem, a large part of this paper is also about using a relatively simple and less conventional way of conducting the inquiry. In this sense, the empirical analysis is somewhat exploratory in nature. Although there have been numerous studies that have tackled the subject of national-level education spending for a global sample of countries (see, for example, those listed in Table 1), none has, to the best of my knowledge, done so using the particular methodology I employ here, namely an analysis of

[1] I acknowledge the important role of private spending on education, especially in progressively richer countries, but this paper focuses on public spending as the predominant source of education finance in most countries. Note that I use the terms "public spending on education" and "education spending" interchangeably.

[2] When making reference to a level-form ratio measure, I am speaking about the 'raw' levels of the ratios, not lagged, differenced or any other functional form of the variable. Since the level-form ratios are monotonically related to the logged levels of the ratios, one could have use logged variables. For instance, using logged measures on both sides of the regression equation, results in a constant elasticities model, in which case, taking the antilog of the common intercept (or, common intercept plus respective mean difference) computes the respective raw 'fixed effect' or that level of spending not dependent on the covariates (controls) – in other words, the time-invariant portion. Bear in mind that, if no controls are included, the respective differential intercept would itself capture all of the observed mean level of education spending, otherwise referred to as an unconditional mean. However, the key point is that, regardless of whether or not the raw levels or logged levels are used, the *patterns* of mean differences should remain unchanged. As is sometimes the case in the relevant literature the national effort measure refers to (total) public spending on education (as a share or percentage of GDP), whereas, the budget share measure refers to (total) public spending on education (as a share or percentage of total government spending).

variance (ANOVA) or analysis of covariance (ANCOVA) method using a regression framework. Using a 'family' of dummies for each categorical measure of interest in a regression setup, amounts to a more efficient way of performing multiple t-tests at once instead of performing multiple two-sample t-tests. Said differently, the method might best be described as a generalised-form t-test using the general linear model (GLM).

Incorporating categorical descriptions of countries into an analysis that 'pools' time-series cross-section data for a large sample of countries from 1989 to 2015, means I am most interested in a particular state of being (heterogeneity). Heterogeneity here can be thought of as wanting to know more about spending patterns based on 'group membership', where the groups comprise user-defined group 'fixed effects'. A major advantage of using a pooled approach is that it circumvents problems related to gaps in the time-series dimension of the panel data. Pooling the data would seem to be an appropriate way of investigating differences between groups (heterogeneity) to know more about the key regression relationships (Greene, 2008, p. 193) and not differences within groups (dynamics). A criticism might be that merely using panel (country) fixed effects would have essentially captured what I am interested in – all country-specific unobserved time-invariant behaviour, including the economic, regional and political effects of interest. However, there is a conceptual problem with doing so.

For one, it is often difficult to theorise what exactly country fixed effects comprise. Because they capture all unobserved country-specific time-invariant factors, they are most often used to control for a particular type of biasing effect. For instance, instead of only being able to say there are unobserved time-invariant effects for country X, using theoretically justifiable components of heterogeneity, based on (observable) group characteristics (e.g., economically and politically similar countries), would seem like a more empirically useful interpretation. These groups, that are time-invariant, by definition, now also include all the applicable country-specific fixed effects, so one would then not need to also include country fixed effects in the estimation procedure – this would be tantamount to the double-counting of time-invariant fixed effects. Said another way, using group effects means one is in essence pooling a number of country-year observations, thereby transcending the lower-order space (or country) dimension to investigate (user-defined) grouped patterns of behaviour at a higher-order aggregation. The researcher is now able to more clearly define education spending patterns by different observable types or descriptions of countries. This would seem like a logically worthwhile exercise.

This paper makes some useful contributions to the literature on education spending. Firstly, it provides a useful way to think about fixed effects estimation or consider fixed effects that have a more meaningful (empirical) interpretation insofar as patterns of behaviour are concerned. Second, the inquiry makes some headway towards expounding how robust the *patterns* of mean differences in education spending are to the use of different estimators and changes in specification. Thirdly, using a factor-variable interaction approach allows one to evaluate to what extent political forces also impinge on what is considered to be an 'economic' process, closely tied to the Wagnerian perspective, which postulates a relationship between the process of economic development and bigger governments. Lastly, using a 'richer' (developed) versus 'poorer' (lesser-developed or developing) country perspective means a bi-modal explanation of education spending patterns can be postulated, giving rise to three testable theoretical propositions, which can be thought of as a novel way to 'reverse engineer' the Wagnerian hypothesis and augment our understanding of the developmental process with respect to education spending. These propositions could also be applied to other areas of the government's budget allocation.

The structure of this paper proceeds as follows. Section 2 conducts a review of some important theoretical and empirical literature with the view to formulating some suitable hypotheses. Section 3 outlines the research methodology, comprising the basic method to be applied, assumptions, empirical model specifications, controls to be used and various robustness checks. Section 4 presents the data collected and description of the data. Section 5 comprises the substantive analysis by reporting the results and commenting on key findings. A general discussion of these findings is then given. Section 6 concludes the paper.

2. Literature Review

2.1 Overview

The vast array of empirical literature attempting to explain national-level education spending patterns can be distilled down to using five broad theoretical dimensions, namely economic, political, demographic, social and globalisation. Although globalisation might be considered an economic dimension, it is often treated separately in the literature (see, for example, Avelino et al., 2005; Baskaran & Hessami, 2012). In most of the empirical studies, the key explanatory variables of interest relate to the economic and/or political dimensions. A list of applicable contemporary empirical studies and their general findings is given later in this section (see Table 1). What this paper then does is provide a synthesis, of sorts, for the various empirical analyses, not in a meta-analysis sense, but by bringing together the economic and political dimensions as key explanatory dimensions of interest, and use the remaining demographic, social and globalisation dimensions as the three main controls.

The demographic (size of the school-going-age population), social (urbanisation ratio) and globalisation (trade as a share of GDP) factors are, more often than not, used as control variables in the various studies – they are spoken about briefly in the Research Methodology section. However, the economic (income per capita) and political (regime type) factors, operationalised in the form of categorical variables, are of special interest in this paper. Another useful way to operationalise the economic dimension is to consider regional differences too because we know that, in general terms, there are richer regions (e.g., Western Europe) and poorer regions (e.g., Sub-Saharan Africa) of the world. In what follows, I justify the inclusion of the economic and political dimensions for my inquiry, prior to presenting a summary of the empirical literature and formulating suitable hypotheses.

2.2 Economic Development

As one of the first major contributions to the field of public finance, Wagner's 'law of increasing state activity' (Castles, 1999, p. 2) has now become a well-established theory of public sector activity, and is widely recognised as a stylised fact (Kuckuck, 2014, p. 129). Adolph Wagner's hypothesis (see, for example, Wagner 1883a, 1883b, 1958, 1892, 1911; Musgrave & Peacock, 1958) concerning state activity can be said to postulate a positive association between economic development and total public spending (Cockx & Francken, 2016, p. 397; Afonso & Alves, 2017, p. 347). However, although the hypothesis generally reveals an empirical uniformity (regularity), Wagner's law was not necessarily meant for purposes of making bold predictive claims (Peacock & Scott, 2000, p. 3).

By far the most widely used economic measure in the education spending literature is that of national income per capita (GNI per capita or GDP per capita) as a proxy for economic development. I focus on this variable (as a categorical, not a continuous, measure) to

operationalise the economic dimension of the inquiry in this paper. A substantial list of studies using income per capita is given in Table 1. However, a host of other 'economic' variables, have also been used in the literature, whether explicitly (as a key explanatory variable of interest) or implicitly (as a control variable), in the modelling of education spending or, more generally, social spending.[3]

Wagner (1892) referred to the relationship between the level of economic development and public sector activity as being attributable to continued progress in the state of the cultural and economic environment, where improved social conditions and income growth are associated with a larger public sector (Kuckuck, 2014, p. 129). In other words, from a Wagnerian perspective, with growth and development comes a growing demand for various social (e.g., education and healthcare) and welfare services that the public sector becomes instrumental in providing on a more massified (large-scale) basis. However, there is conjecture about the 'causes' of bigger governments. Whereas Wagner (1911, p. 734) points out that growth of the public sector comes at the expense of (substitutes for) growth in the private sector (see also Peacock and Scott, 2000, p. 2), a different conclusion has been argued when referring to the growth of government from a historical perspective:

> The main conclusion of Part One is that the growth of public spending was not caused by inevitable forces that made it imperative. It was thus not inevitable as assumed by some theories about the growth of spending such as Wagner's Law or Baumol's disease. On the contrary, it is argued that growth was nothing but a response to changing perceptions about what the government should do. In a way the growth of government reflected a lack of confidence in the private sector's ability to deal with some problems and a belief that public spending was the best way to deal with several risks faced by individuals. The action of government was always assumed to be additive to or complementary of the action of individuals. It was almost never assumed to be substitutive.
>
> Tanzi and Schuknecht (2000, pp. 1-2)

Three points can be advanced here. Firstly, in defence of Wagner, he was formulating his ideas at a time in the late-19th and early-20th Centuries, without the benefit of well over one hundred years of empirical experience and little in the way of mathematical tools with which to 'prove' his theory. Secondly, whether one speaks about a growing demand for various publicly provided services or changing perceptions in society about the role government should play, as Tanzi and Schuknecht suggest, these two 'causal' processes behind bigger governments might be viewed as being synonymous (and not in conflict) with one another, emanating from demand-side socio-political pressures. Thirdly, it is more likely the case that a complementary (as opposed to substitutive) relationship exists between public and private sector expansion, as Tanzi and Schuknecht also seem to suggest. From a purely logical perspective, if wealthier countries tend to have larger public sectors (as ascribed to the Wagnerian hypothesis), these wealthier countries have not become richer by osmosis (some passive process of growth and development), but have done so by virtue of the private sector of the economy having also expanded – this is at the very heart of what growth and associated development encompasses.

[3] Various measures of public debt or debt service (Tilak, 1989; Tilak, 1990; Brown & Hunter, 1999, 2004; Castles, 2007; Busemeyer, 2009; Fosu, 2010), fiscal balance (Tilak, 1989; Huber et al., 2008), inflation (Brown & Hunter, 1999, 2004; Avelino et al., 2005; Iversen & Stephens, 2008), unemployment (Avelino et al., 2005; Iversen & Stephens, 2008; Jensen, 2011), adult literacy (Tilak, 1984; Baqir, 2002), foreign direct investment (Huber et al., 2008; Iversen & Stephens, 2008), development assistance or aid (Stasavage, 2005; Fosu, 2010), labour force participation (Jensen, 2011) and income inequality (Manzano, 2013) have been used.

At this juncture it would be prudent to point out that, from a 'causal' point of view, the Wagnerian linkage flows from national income per capita to public sector activity, whereas the Keynesian linkage flows from public sector activity to national income per capita. In a sense, the Wagnerian and Keynesian formulations represent two contrasting theories of the role the public sector plays in national economic growth and development. This being said, I focus on the former (Wagnerian) linkage for my purposes in this paper. To operationalise the economic dimension, a common approach is to use the level of economic development of a particular country as measured by national income per capita because the expectation is that higher levels of economic development are associated with higher levels of public spending (as a share of GDP) – the assertion of Wagner's law (Busemeyer, 2008, p. 461).

Peacock and Scott (2000) offer a good synopsis of Wagner's salient ideas and critical discussion of the empirical testing thereof with respect to 15 studies which have been conducted in the past. One of the most intriguing points of discussion offered by them concerns the methodological approach applied in the studies they surveyed, bearing in mind that all of the studies ostensibly purported to investigate Wagner's law. They noted that most of these studies merely represented a 'race-to-the-top' with respect to using more sophisticated methods of inquiry, possibly moving further away from what Wagner himself may have intended because the testing of the 'law' should not necessarily require complicated econometric techniques to espouse what is by nature a very simple law (see Peacock & Scott, 2000, pp. 2-10).

In the more recent literature, there have been a number of studies that have focused attention on Wagner's law in more general terms – not necessarily with respect to education spending only as I do. These studies also make reference to other studies on the subject.[4] However, since public spending on education is a component of total government spending, investigating education spending patterns with respect to national income per capita, is an alternative way of investigating Wagner's law. There have been many studies that have considered the relationship between national income per capita and education spending or, more broadly, social spending (see Table 1).

As stated previously, another way to operationalise various economic differences between countries is to use regional categorisations because there are economically richer (developed) versus poorer (developing) regions of the world. The usefulness of using a regional variable can be traced back to a very early study of education spending patterns such as that of Zymelman (1976). Where no significant relationship was found with respect to the economic variable (income per capita) and education spending, a significant relationship emerged when combined with region or region in isolation (Zymelman, 1976, pp. 50-51). One reason for the importance of considering a regional dimension as opposed to purely considering income per capita groupings of countries is that countries from a similar region are more likely to also have some similar economic characteristics, so using a region-specific variable to estimate mean differences, means one is also 'controlling' for economic status too.

2.3 Political Democracy

Considering the state of economic development (or regional differences) in isolation from the political environment would be short-sighted. As operationalised in its many different forms,

[4] The interested reader can refer to the following studies: Afonso & Alves (2017); Magazzino et al. (2015); Kuckuck (2014); Kumar et al. (2012); Shelton (2007); Akitoby et al. (2006); Peacock & Scott (2000); Gemmell (1990, 1993) and Ram (1987).

the political dimension has been shown to be an important ingredient in the education-spending mix. Measures of the political environment can be broadly classified into two general streams: what I refer to as 'first-generation' (political ideology) and 'second-generation' (regime-type) measures. Measures of political (electoral) cycles might be considered a third type of political measure useful for applied work, and distinct from the political ideology and regime-type measures. This study focuses on the latter (regime-type or political democracy) and not the former (political ideology) measure.

That being said, it would be remiss not to acknowledge the important role various measures of political ideology have played in the empirical literature, whether from the point of view of general public spending or specific components thereof (education spending included).[5] Schmidt (1996) discusses the usefulness of first-generation measures of polity for applied work. A useful study presenting a meta-analysis of political ideology is that by Imbeau et al. (2001). A discussion of government ideology and budget composition is given by Potrafke (2011). A good account challenging the conventional wisdom that party politics matters is that by Garritzmann & Seng (2016).

The partisan political (political ideology) approach comprises a more straightforward class-based approach to partisan (Leftist versus Rightist) preferences, generally predicting a leaning towards primary and secondary education spending for Leftist governments. In contrast, a 'new politics' approach to education spending adopts the line of argument that, "political parties are not merely transmission belts for the economic interests of social classes, but use policies and spending strategically to attract and consolidate voter groups" (Busemeyer, 2009, p. 107). What this means in more simple terms is that, regardless of partisan political affiliation, democratic governments (as opposed to less democratic or autocratic ones) are inclined to provide a greater number of services (see Lake & Baum, 2001) and favour policies that appeal to a majority of the citizenry in order to consolidate political power.

However, for the abovementioned reason, I am more interested in using a suitable second-generation measure of political democracy for this inquiry. For instance, education is arguably a 'populist' component of the budget allocation meaning politically democratic governments are more likely to have higher levels of public spending on education, regardless of the type of spending measure used (national effort or budget share), *ceteris paribus*. Political democracy can be thought of as implying significant socio-political pressures being placed on government by a majority of the citizenry, to act increasingly in line with the needs and wants of society. An interesting proviso applies here though: regardless of the state of polity, distinguishing between societies with a relatively larger (and growing) cohort of school-going-age versus elderly citizens would be important because this would necessarily imply different policy choices to be made by government, which might inevitably favour one group over another when public choices must be made concerning how best to allocate scarce public resources. This explains why, first and foremost, including a demographic control for the school-going-age population would be paramount in any investigation of education spending.

[5] There is a rich history of studies using ideological dimensions that can be referred to here, for example: Hibbs (1977); Cameron (1978); Castles & McKinlay (1979); Wilensky (1981); Castles (1982, 1986); Hibbs (1987); Iversen (2001); Adserà & Boix (2002); Kittel & Obinger (2003); Busemeyer (2009); Careja & Emmenegger (2009) and Herwartz & Theilen (2017).

There are a number of studies in the contemporary literature that have proceeded by using various kinds of second-generation measures insofar as education spending is concerned. For example, Avelino et al. (2005) used the democracy score and Fosu (2010) used the degree of political constraint on the executive in government as suitable political democracy (regime-type) measures of polity. Although using democracy score or political constraint measures are useful, they arguably include too many categorisations to be sensibly useful for my purposes. Therefore, I turn my attention to focusing on the state of political democracy (a binary categorisation – Yes or No) as my favoured second-generation measure of polity.

2.4 Political Economy

The interplay between economic and political forces, and their effect on education spending patterns (the political economy of education spending), can be thought of as the Wagnerian view contingent on political processes. Growth and development relies on particular institutional mechanisms or processes, providing a conduit through which economic prosperity can be stimulated (see, for example, North, 1991). In other words, the inter-relationship between economics and politics plays a crucial role in the process of growth and development (North et al., 2006, p. 2). Suffice to say, economic and political processes are more likely to occur simultaneously or together in the broader process of growth and development. As polar extreme explanations, institutional mechanisms might be seen as political constructs to promote better governance and market access in democratic, arguably less corrupt, societies, but chiefly as a way to maintain political power and limit market access in autocratic, arguably more corrupt, societies, in which political rent-seeking behaviour is the norm rather than the exception. Depending on the ideological motive behind the government's financing of education in non-democratic (autocratic) versus democratic countries, plausibly contrasting spending outcomes might apply to these two different types of polity. Democratic governments would reasonably be subjected to greater majoritarian political pressures to increase spending on education, thereby expanding market access opportunities to the general population over the longer term. However, non-democratic (autocratic) governments would have greater, unfettered power to substitute away from spending on education and reallocate public spending towards the state 'machinery' (military spending) used to maintain political power and control over the citizenry, thereby limiting market access opportunities over the longer term.[6]

This is a somewhat jaundiced view, and clearly not always the case, but it does, nonetheless, represent a realistic explanation of how public funding might be prioritised in one type of polity versus the other. What I try to highlight here is that economic and political forces are

[6] It would be interesting to point out that a whole other branch of Political Science research concentrates on the role played by Leviathan governments in state activity, attributable to the pioneering work of Brennan and Buchanan (1980), and associated earlier work of Wicksell (1896). The Leviathan approach might be viewed as an alternative way of thinking about how autocratic governments behave with respect to taxation and spending, when there is very little constraint on these types of government – what Buchanan (1977, p. 13) might refer to as "politics for profit". However, fiscal decentralisation is seen as a mechanism to restrict the Leviathan government's 'monopoly' power on state activity, hence, the empirical observation that, "total government intrusion into the economy should be smaller, ceteris paribus, the greater the extent to which taxes and expenditures are decentralized" (Brennan & Buchanan, 1980, p. 15). Thus, hypothetically speaking, autocratic, Leviathan-type governments, who adopt a more fiscally *decentralised* approach to taxation and spending, might not have as much of a pejorative effect on education spending, than, say, an autocratic, Leviathan-type government adopting a strongly *centralised* approach to taxation and spending. Although not directly related to the study at hand, a good example of an empirical study of fiscal decentralisation and education spending (but for OECD countries) is that of Busemeyer (2008).

not mutually exclusive – they are fundamentally connected. Thus, speaking about political democracy in isolation from the state of economic development is useful insofar as understanding political forces are concerned, but presents only part of the story. This is why using two factor-variable (economic *and* political or regional *and* political) interactions becomes very useful for my purposes in this paper, but with a caveat: since this paper focuses on an economic interpretation of education spending patterns, more or less in line with the Wagnerian hypothesis, any propositions arrived at by way of the empirical analyses must be tempered by saying that political processes play an important 'mediating' role in shaping these patterns, something not necessarily explicitly evidenced in Wagner's own writings. This caveat is well put by Peacock & Scott (2000, pp. 3-4), who speak of the causative relationship between national income and public sector growth being coordinated by choices signalled in the political system, but reference to this political machinery is not necessarily found in Wagner's own work.

2.5 Hypotheses

Having outlined some of the important theoretical arguments and examples of empirical studies with respect to general public sector expansion and public spending on education, I now present a more comprehensive list of studies on the latter (education spending), with the view to formulating appropriate hypotheses. Table 1 presents a relatively detailed list of contemporary empirical studies that have operationalised measures of economic activity (national income per capita) and regime type (political democracy). The list is not exhaustive, but comprises a substantial corpus of the contemporary literature for the purposes of formulating hypotheses with respect to the economic and political dimensions of interest.

Table 1 refers to the general *sign* of the applicable relationship estimated for each study, irrespective of whether correlation or regression methods were used, whether or not the applicable economic or political variable was the key variable of interest or merely used as a control and where the relationship was generally found to be significant (at the 5% level or less). I am most interested in the estimated sign of the relationship, owing to the variety of different functional forms of variables these studies used. Note that I have not included studies from a whole other branch of literature relating education spending to fiscal cyclicality (for example, Afonso & Jalles, 2013; Arze del Granado et al., 2013). I consider this to be a separate, albeit related, literature concerned more with cyclical variation.

What should become immediately noticeable is the dearth of empirical evidence with respect to the budget share measure, making it an interesting exercise (in itself) to test for mean differences for a large global sample of countries over an extended period of time using both the national effort and budget share measures. The general 'theme' from the 16 studies listed in Table 1 is that the national effort measure is generally *positively* related to national income per capita and political democracy. Given the lack of empirical evidence with respect to the budget share measure, it becomes difficult to proffer a meaningful expectation about the associated sign with regards to both the economic and political dimensions.

However, one can use theoretical reasoning to build an expectation about the budget share measure. Regardless of developmental status, education is likely to be an important component of fiscal spending in most countries, but its share of total public spending in different types of economies may vary depending on the number of other fiscal components to be financed by government. For instance, if richer, more developed (poorer, lesser-developed or developing) countries, do indeed, have larger (smaller) public sectors as the

Wagnerian hypothesis suggests, it would stand to reason that education's share of total government spending would tend to be smaller (larger) because richer (poorer) countries would also tend to have a greater (lesser) variety of fiscal components to fund from the public purse – what I refer to as a fiscal varieties perspective – implying education comprises a smaller (larger) share of total government spending. Therefore, the budget share measure and national income per capita are expected to be *negatively* related. Insofar as political democracy is concerned, all things being equal, one might still expect the budget share measure to be positively related to political democracy because democracy still implies the same majoritarian socio-political pressures. In other words, in isolation from the state of economic development, the budget share measure is expected to be *positively* related to the state of polity.

Figure 1 summarises the various hypotheses. From these hypotheses, we can postulate one of two outcomes. First, controlling for the state of economic development (poorer countries), democratic countries should spend more in terms of both measures of spending than do their non-democratic counterparts. Second, controlling for the state of polity (political democracy), richer countries should spend more in national effort terms, but less in budget share terms than do their poorer counterparts.

Figure 1: Summary of the Hypotheses

Dependent Variables	**Independent Variables**	
	GNI per capita	Political Democracy
National Effort	+	+
Budget Share	–	+

Table 1: Summary Findings of 16 Empirical Studies Related to Public Spending on Education

Authors (Year)	Method(s)	Sample Size	Number of Countries	Time Period	National Effort		Budget Share	
					Economic	Political	Economic	Political
Zymelman (1976)	Cross-sectional regression	8 to 69	69 developing countries	Circa 1973	+	n/a	n/a	n/a
Verner (1979)	Cross-sectional correlation	102	102 countries	1964-1965	+	+	ns	+
Castles (1989)	Cross-sectional regression	18	18 OECD countries	1960; 1981	ns	n/a	n/a	n/a
Tilak (1989)	Cross-sectional regression	16-20	20 Latin American countries	1965; 1970; 1980; 1985	+	n/a	n/a	n/a
Ram (1995)	Cross-sectional regression	18	18 OECD countries	1985	+	n/a	n/a	n/a
Baqir (2002)	Cross-sectional regression & panel-data methods	Various	59 to 106 countries	1985-1998	+	+	+	+
Avelino et al. (2005)	Panel-data methods	312; 314	19 Latin American countries	1980-1999	ns	+	n/a	n/a
Stasavage (2005)	Pooled regression & panel-data methods	365; 247; 191	44 African countries	1980-1996	+	+	+	+
Busemeyer (2007)	Panel-data methods	421	21 OECD countries	1980-2001	+ or –	n/a	n/a	n/a
Huber et al. (2008)	Panel-data methods	446	18 Latin American countries	1970-2000	+	+	n/a	n/a
Iversen & Stephens (2008)	Panel-data methods	336; 138	18 OECD countries	1960-2003	–	ns	n/a	n/a
Akanbi & Schoeman (2010)	Panel-data methods	135	15 African countries	1995-2004	+	+	n/a	n/a
Fosu (2010)	Panel-data methods	79	35 Sub-Saharan countries	1975-1994	n/a	n/a	ns or +	ns
Potrafke (2011)	Panel-data methods	552; 247	23; 20 OECD countries	1970-1997; 1990-2006	–	n/a	n/a	n/a
Cockx & Francken (2016)	Panel-data methods	320 to 349	129 to 140 countries	1995-2009	+	n/a	n/a	n/a
Garritzmann & Seng (2016)	Panel-data methods	245	21 OECD countries	1995-2010	ns or –	+	n/a	n/a

Notes: The method(s) used in each respective empirical study refer to the primary method(s) employed. The national effort measure refers to total public spending on education as a share of GDP (or similar) and the budget share measure refers to total public spending on education as a share of total government spending. Generally speaking, the economic factor refers to national income per capita and the political factor refers to some measure of political democracy. Where not applicable (n/a) appears, this means the explanatory variable of interest (economic or political) was not used/considered. Where not significant (ns) appears, this means no significant relationship was generally estimated at the 5% level or less. The study by Huber et al. (2008) uses public spending on health and education as a share of GDP as the dependent variable. The study by Akanbi & Schoeman (2010) uses the World Bank's corruption control index score and the study by Fosu (2010) uses Polity IV Project's political constraint score.

3. Research Methodology

3.1 Explanation of the Method

In a day and age increasingly preoccupied with methodological sophistication, there is merit in applying a sensibly simple approach to understanding the patterns of education spending. In its most basic form, the method applied can be said to comprise a generalised-form t-test, in which, only dummy (qualitative) variables are used. In much the same way as the conventional Student's (Gosset's) t-test is used to test for differences in the means of two sample groups, using a least-squares dummy-variable (LSDV) regression framework (a type of GLM) is an alternative way to test for differences in means across more than two sample groups at once (Gujarati & Porter, 2009, p. 278), otherwise referred to as analysis of variance (ANOVA). However, the GLM also has the added advantage of being able to condition out these differences in means by allowing for the inclusion of additional (quantitative) covariates, otherwise referred to as analysis of covariance (ANCOVA), which is very useful from a modelling point of view. In general terms then, the method of analysis is more descriptive of empirical patterns rather than anything else.

By imposing a particular, user-defined group structure on the data, one is in effect 'splitting' (partitioning) the education spending data into particular groups that arguably afford more empirically meaningful judgements about patterns of behaviour. In other words, in a panel-data context, using dummy variables on the right-hand side over the time period in question, creates time-invariant groups, by definition. In much the same way as a country fixed effect would be referred to as that part of heterogeneity that is *country-specific* and time-invariant the group fixed effect might be referred to as that part of heterogeneity that is *group-specific* and time-invariant.[7] Suffice to say, the associated categorical groupings on the right-hand side now implicitly become the sub-samples or sub-groups of interest.

For the purposes of this paper, since I am most interested in interacting two types of categorical measure (economic and political or regional and political) the method can best be descried as a factor-variable *interaction* generalised-form t-test (GLM factorial or two-way ANOVA). Alternatively, a factor-variable *interaction conditional* generalised-form t-test (GLM factorial or two-way ANCOVA) method could be used, where the word "conditional" refers to the mean differences of interest now being conditional on the inclusion of the mean moments of the respective covariates in each case. The LSDV regression model can be easily setup in Stata using factor or indicator variables, without having to explicitly generate separate dummy variables. For the interested reader, a useful practical guide to interaction effects and group comparisons in Stata is that by Williams (2015).

3.2 Methodological Assumptions

This paper proceeds under the assumption of cross-section independence, heterogeneous intercepts and homogeneous slopes. However, as part of the various robustness checks, one can make two modifying assumptions about the presence of cross-section dependence. Either,

[7] Another way to think of this is to say one is interested in 'stripping out' the time-varying nature of education spending by virtue of modelling spending (a continuous measure) as a function of the time-invariant groups of interest (the categorical measures comprising either economic and political or regional and political groupings of countries), that are assumed, by design, to remain the same over the time period in question. This means one is most interested in pooling the education spending panel data to compute the mean (or mean differences) in spending across countries with particular group characteristics over the sample time period.

under the additional assumption that any unobserved common effects are not correlated with the regressors, a suitable estimator (the Driscoll-Kraay estimator) can be used to estimate consistent standard errors. Or, assuming any endogenous unobserved common effects are homogeneous in nature – an implicit assumption often made in many papers ostensibly trying to control for 'common shocks' – time (year) fixed effects can be incorporated into the specification. Additionally, tests for regression-related problems such as heteroskedasticity, serial correlation or autocorrelation and cross-section dependence are not conducted, but rather different assumptions about the data-generating process are made that pre-suppose the presence thereof, and remedy accordingly. In making the necessary modifying assumptions, using different estimators, changing the specification and considering different sub-samples, one is able to make a judgement about the robustness of any estimated patterns of mean differences in spending when user-defined group structures are imposed on the data.

The two assumptions concerning heterogeneous intercepts and homogeneous slopes are important ones. What heterogeneous intercepts mean is I am most interested in the patterns of spending behaviour exhibited by the time-invariant user-defined group (as distinct from panel or country) fixed effects, *ceteris paribus*. What homogeneous slopes mean is I am also interested in knowing what patterns in the group fixed effects are still evidenced once other important education spending covariates – that are assumed to have a homogeneous effect across the groups of interest – have been controlled for. Using various continuous control measures forms part of checking how robust the patterns exhibited by the group fixed effects are to changes in specification. The interested reader might note that heterogeneous slopes could have also been employed. But, there is a conceptual problem in doing so because, in effect, one would need to interact the time-invariant groups with each of the time-varying controls. This would not only unnecessarily complicate the models, but would also most probably confound any time-invariant effects of interest because the heterogeneous slope parameters would most likely absorb most (if not all) of the 'fixed effects' of interest. Put simply, the objective of this paper is to perform fixed-effects estimation with a difference. Instead of estimating country fixed effects, I rather estimate more aggregated, user-defined group fixed effects.

3.3 Empirical Model Specifications

The empirical model specifications make use of pooled panel data to estimate the applicable mean differences. For this reason, the error term is always shown as having both space and time identifiers (e_{it}). Below, I show the various specifications without controls (ANOVA model specification) representative of Models (1) in the various tables of results in the text, and various specifications with controls (ANCOVA model specification) representative of Models (2) in the various tables of results in the text. However, for the models appearing as part of the robustness test results in the appendices, these models are named sequentially (i.e., Model 1A, 1B, 2A, 2B, 3A, 3B, …, 6A and 6B), to denote progressively more complex specifications, and do not refer to the naming convention used for the models in the text.

Two-way (interaction) models without controls specified:

$$Y_{it} = \sum_{j=1}^{5} \sum_{m=1}^{2} \alpha_{jm} \text{Economic}_{jit} * \text{Political}_{mit} + e_{it} \qquad \text{(Model 1a)}$$

$$Y_{it} = \sum_{l=1}^{2} \sum_{m=1}^{2} \alpha_{lm} \text{Region3}_{lit} * \text{Political}_{mit} + e_{it} \qquad \text{(Model 1b)}$$

Two-way (interaction) models with controls specified:

$$Y_{it} = \sum_{j=1}^{5} \sum_{m=1}^{2} \alpha_{jm} \text{Economic}_{jit} * \text{Political}_{mit} + \sum_{n=1}^{N} \beta_n X_{nit} + e_{it} \qquad \text{(Model 2a)}$$

$$Y_{it} = \sum_{l=1}^{2} \sum_{m=1}^{2} \alpha_{lm} \text{Region3}_{lit} * \text{Political}_{mit} + \sum_{n=1}^{N} \beta_n X_{nit} + e_{it} \qquad \text{(Model 2b)}$$

where:

Y_{it} is either the national effort or budget share measure of total education spending;
α_n refers to the coefficient of the respective dummy (interaction) variable;
β_n refers to the coefficient of the respective control variable;
"Economic" refers to a family of five dummies, with each dummy being a binary (0 or 1) categorisation for each of the income per capita country groups;
"Region3" refers to a dummy, with a binary (0 or 1) categorisation for each of the two regional country groups (richer or poorer);
"Political" refers to a dummy, with a binary (0 or 1) categorisation for each of the two states of political democracy (democratic or not democratic);
X_{nit} is a vector of continuous control variables comprising a minimum of three or maximum of eight controls;
e_{it} is the pooled error term.

I exclude showing the 'common' intercept in all specifications because this can be a user-defined group of countries that becomes the base or reference group against which comparisons (mean differences) can be made. Therefore, the common intercept is subsumed as part of the empirical specifications outlined above. In all specifications, m-1 dummies were included to avoid the dummy-variable-trap problem of perfect multicollinearity. Two-way models refer to the interaction between dummies – although the respective dummies combine in a multiplicative way, their combination still appears in the (linear) model in an additive way. These specifications are tantamount to disaggregating the intercept (autonomous) component of spending into separate parts, either without or with various covariates (controls) appearing as additional regressors, where these controls are assumed to have homogeneous slopes. An alternative approach would be to allow the slopes of the controls to vary across groups, but this would introduce other methodological problems (Greene, 2008, p. 194; see also Cornwell and Schmidt, 1984).[8]

[8] There are generally two approaches that could be used. The first is to assume heterogeneous intercepts and homogeneous slopes, in which case, because the slope parameters and means of the continuous control measures are held constant across all countries, testing for mean differences simply becomes a test of differences in the 'fixed effects' (differential intercepts). However, the second assumes heterogeneous parameters. Because intercepts *and* slopes are now assumed to vary between groups, testing for mean differences becomes more complicated. For example, one might now be more interested in testing for differences in the linear prediction of the means for each group. For the purposes of this paper, I focus on the first approach.

3.4 Control Variables

I start with an explanation of the three main controls used (pop024, urban and trade – see Tables 5 and 6). Using control variables can be thought of as a way to change the specification in order to estimate *conditional* mean differences. There are generally five dimensions (economic, political, demographic, social and globalisation) that explain education spending patterns. Since the models already explicitly consider particular economic and political groupings of countries, it stands to reason incorporating suitable controls for the other three dimensions would be necessary. The size of the school-going-age or youth-age population (pop024), urbanisation ratio (urban) and trade as a share of GDP (trade), comprise the three main control variables used to operationalise the demographic, social and globalisation dimensions, respectively. There are a number of studies in the empirical literature that have used one or more of these controls (for example, Manzano, 2013; Baskaran & Hessami, 2012; Jensen, 2011; Akanbi & Schoeman, 2010; Fosu, 2010; Busemeyer, 2008, 2009; Huber et al., 2008; Iversen & Stephens, 2008; Castles, 2007; Avelino et al., 2005 and Brown & Hunter, 2004).

The size of the youth population – defined as 0 to 24 years-of-age for my purposes because total spending on education comprises pre-primary, primary, secondary and tertiary levels of education – is one of the most important control factors to consider. All things being equal, a larger school-going-age population (as a share of the total population) imposes greater demand-side pressures on the education system. Therefore, a positive relationship between education spending and the size of the youth population is expected. To capture this aspect of demographic change, some studies have used various measures of student enrolment (for example, Tilak, 1984; Castles, 1989, 1999; Baskaran & Hessami, 2012; Speciale, 2012). However, the more common approach has been to use country-level population estimates of this cohort (for example, Schultz, 1988; Castles, 1989, Baqir, 2002; Brown & Hunter, 2004; Stasavage, 2005; Busemeyer, 2007, 2008; Huber et al., 2008; Akanbi & Schoeman, 2010; Jensen, 2011), and for good reason. Since enrolment is most likely endogenous – including enrolment on the right-hand side of a model explaining education spending, would most likely give rise to problems related to simultaneity bias, so a more truly exogenously-determined measure, such as the youth-age population share, would be more appropriate in a model of total education spending.

The urbanisation ratio or population density also imposes particular socio-political pressures on government and the education system. A greater concentration of people living in urban areas means majoritarian socio-political concerns can be mobilised more easily, and in conjunction with a growing youth-aged population, exacerbate the demand-side pressures for the government to provide an appropriate level of education to the broader society. Again, a positive relationship is expected to exist. A number of studies have controlled for a greater concentration of people in urban areas (for example, Verner, 1979; Schultz, 1988; Castles, 1999, Baqir, 2002; Avelino et al., 2005; Huber et al., 2008; Akanbi & Schoeman, 2010).

Trade is another important control. To incorporate an aspect of economic activity associated with the foreign sector such as total trade or trade openness (imports plus exports as a share of GDP) would be important on the grounds of two alternative hypotheses about the effect thereof. The compensation hypothesis postulates that government needs to ‘compensate’ society for the losses incurred through the process of globalisation – supposedly from greater foreign competition eroding opportunities domestically – by increasing social and other welfare spending (a positive relationship). In contrast, the efficiency hypothesis postulates

that a process of globalisation results in a more competitive, thriving domestic economy, meaning government is more easily able to curtail social and other welfare spending (a negative relationship). Which of these two effects dominates is usually a matter of empirical inquiry. A number of studies have controlled for globalisation using total trade or trade openness as a generic measure (for example, Castles, 1999; Kaufman & Segura-Ubiergo, 2001; Baqir, 2002; Brown & Hunter, 2004; Avelino et al., 2005; Castles, 2007; Schmidt, 2007; Huber et al., 2008; Iversen & Stephens, 2008; Busemeyer, 2009; Jensen, 2011; Baskaran & Hessami, 2012; Manzano, 2013; Herwartz & Theilen, 2017).

3.5 Robustness Checks

This paper uses various ways to test for robustness of the *patterns* of mean differences. The most basic way of doing so was to check for differences in the unconditional (no controls) and conditional (with three main controls) patterns of mean differences. This was extended to use different estimators of the standard errors, which potentially alter the significance of any differences, and different estimators of the parameters, which potentially alter the sign, size and significance of any differences. These aforementioned methods were all used in compiling the results presented in Tables 5a, 5b, 6a and 6b in the text. However, a couple of other robustness checks were also performed, the results of which are presented in the appendices (Appendix D, E, F and G). These checks comprised changing the specification in one of two ways, by either using additional control variables or using the additional control variables and time (year) fixed effects. This was done to see if any meaningful changes to the patterns of differences would be evidenced. I start with a discussion of the different estimators, and then continue with a discussion of the changes in specification.

Different Estimators

Eight different estimators were employed, that each take account of different problems or aspects related to the data-generating process. For example, although I do not test for regression-related problems such as heteroskedasticity, autocorrelation and cross-section dependence, these problems can be assumed to exist and remedied accordingly to see how the empirical patterns change under different assumptions about the data-generating process. Using pooled panel data in a regression framework is important for the analysis in this paper, but there are several possible problems related to the errors (or estimated residuals) that might typically occur in various dimensions of the macro-panel data, and would, therefore, need to be remedied. A brief explanation of each estimator is given.

Under the assumption that any problems with the errors are not necessarily specification-related, the coefficient estimates are not inconsistent, but are likely to have incorrectly (under-inflated) associated standard errors in most circumstances. Six estimators were used in this instance. The LSDV estimator with heteroskedasticity-robust standard errors and no controls produces the baseline estimates of the mean differences, and provides a point of comparison for other estimators of the standard errors (and coefficients). Huber/White standard errors correct for non-constant variance of the residuals (heteroskedasticity) typically arising in the space or cross-section (country) dimension of the data. Clustered standard errors (clustering by country or year or both) comprise three different ways to estimate the standard errors. However, clustering requires the number of clusters to be sufficiently large (see Cameron & Miller, 2015), and may actually impose other problems, especially when the data is more strongly unbalanced (a different number of time gaps or missing data by country in the time dimension of the panel data).

The Newey-West and Driscoll-Kraay estimators provide alternative ways to correct the estimated standard errors. The Newey-West procedure produces heteroskedasticity and autocorrelation consistent (HAC) standard errors. Since serial autocorrelation of the residuals (correlation between observations of a country ordered over time) entails a problem related to the time-series dimension of the data, under the assumption that heteroskedasticity and autocorrelation are more arbitrary in nature and not a result of misspecification, using an alternative estimator that corrects the standard errors in the presence of both of these problems seems appropriate. In addition, in a globally integrated world, where countries have stronger economic and regional ties, the likelihood of cross-sectional correlation (the problem of spatial dependence being present in the data-generating process) is heightened. As mentioned before, where any unobserved common effects ('common shocks') are exhibited across countries, and these effects are assumed to not be correlated with the regressors, the parameters are still consistent, but no longer efficient, meaning the solution becomes one of correcting the (biased) estimated standard errors (De Hoyos & Sarafidis, 2006, pp. 482-483). The Driscoll-Kraay procedure adjusts the standard errors accordingly by estimating spatial correlation-consistent standard errors.

The six estimators outlined so far comprise post-estimation approaches to adjusting the associated standard errors of the covariance matrix. They make the implicit assumption the coefficient estimates are still consistent, hence, the reason why the parameters for all of these models in Tables 5 and 6 do not change, except when the model specification is changed. Not making an adjustment to the estimated standard errors means one might be overly optimistic concerning the significance of the estimated mean differences – a greater likelihood of rejecting a true null or committing a type I error. The methods used to correct the standard errors generally produce estimates thereof that are more conservative (larger). Simply put, all of the aforementioned estimators relate to a class of covariance-matrix estimators used to correct the standard errors accordingly.[9]

Quantile regression and robust regression comprise two different ways to estimate the parameters. As said before, the LSDV results represent a baseline, of sorts, against which the coefficient estimates of the quantile (or median) and robust estimators can be compared. In its simplest form, quantile regression provides an alternative estimator of central tendency (the median), particularly useful when the underlying samples of data might not necessarily be normally distributed, and one wants to compare the mean and median differences on a purely qualitative basis. One might suggest estimating different quantiles, but this would seem almost counter-productive in an application that seeks to test for differences in the centrally-located values as opposed to any other quantiles. Instead of trying to minimise the sum of the squared residuals like ordinary least squares (OLS), quantile regression tries to minimise the sum of the absolute residuals (least absolute residual regression). Robust estimation is another useful method to deal with 'outliers' (extreme observations), often noticeable when

[9] Huber/White heteroskedasticity-robust standard errors are routinely performed in Stata using "vce(robust)" as an option after the regress command. One-way clustering of standard errors is routinely performed in Stata using "cluster(cid)" or "cluster(year)" as an option after the regress command. Two-way clustering is performed in Stata with the user-written programme "vce2way" by Hong Il Yoo (sourced from: http://fmwww.bc.edu/RePEc/bocode/v). For a more practical discussion of cluster-robust inference, one can refer to Baum et al. (2010) or Cameron and Miller (2015). The Newey-West procedure (see Newey & West, 1987; 1994) is automated in Stata using the "newey" command. The Driscoll-Kraay procedure (see Driscoll & Kraay, 1998) is performed in Stata with the user-written programme (xtscc) by Hoechle (2007). There was an update to Hoechle's programme on 3 April 2018, which meant the estimation of more 'optimistic' asymptotic standard errors was replaced with the estimation of more 'conservative' standard errors that take account of a small sample adjustment. The latter are the ones reported in this paper.

working with national-level education spending data reported for many different countries. The robust estimator uses a weighted least-squares (WLS) approach to estimation, with progressively more extreme observations being assigned a lower 'weighting' in the estimation sample based on the observation's associated deviation from the sample mean value. Both the quantile and robust estimators effectively comprise different ways to deal with outliers with respect to the dependent or response variable, and provide different estimates of the coefficients (and associated standard errors).

Changes in Specification

There are two aspects to the changes in specification that comprise various robustness checks. The first comprises the additional control variables to be used in the various models. The second comprises the use of year fixed effects. With respect to the additional controls, the expected sign of the relationship is not of particular interest, hence, the reason why I do not necessarily proffer an expectation with respect to the sign of the respective relationship. Rather, for my purposes, their inclusion is justified here on both theoretical grounds and because of their use in various parts of the empirical literature. Suffice to say, I'm more interested in simply knowing if their inclusion confounds the estimated mean differences.

Five additional control variables were tested. First, a measure of overall educational 'quality' might be a useful control measure. The human capital index (hci) for each-and-every country by year serves as one way to measure whether educational outcomes are being achieved at a macroeconomic level. Another argument can be advanced for including the aforementioned measure. If the ultimate goal of education spending (at a macroeconomic level anyway) is to improve a nation's human capital, and improvements in human capital are an important source of economic growth and development, then including a suitable measure of human capital development would be useful, because this not only foreseeably impinges on policy decisions about how much to spend on education, but might actually be correlated with the economic and political categorisations of interest. Said differently, including a suitable measure thereof controls for a possible endogeneity problem (omitted variable bias).

Theorising a cause-and-effect relationship between education spending and nationwide human capital development, might be more tenuous to show, and is beyond the scope of this paper. For one, such a relationship would be further complicated by deciding on the appropriate time lag needed to specify such a relationship, irrespective of theorising the direction of the causal linkage. Nonetheless, it would be useful to include a measure of overall human capital development, because richer (more developed) countries do tend to have a higher human capital index relative to poorer (lesser-developed) countries, which might partly explain comparative education spending differences.

Second, the share of the elderly population (65 years-of-age and older) out of the total population (pop65) might serve as a useful alternative demographic control variable. In some sense, it captures how changes in the size of the elderly population, and associated welfare and pension spending, supposedly competes with the youth-age population, and associated spending on education. This variable might be especially important in countries where transfer payments comprise a substantial portion of total government spending, implying an effect on education's share of the fiscal pie. However, the education sector and elderly population might not necessarily compete for public spending – they are not necessarily negatively related. For instance, in richer countries, a positive relationship between education spending and the size of the elderly population might actually be indicative of the fact that

these countries are able to spend more on both components – they are, in a sense, complements, not substitutes – not to mention it may be politically expedient to do so because both demographic cohorts would be appeased by this policy outcome.

Third, the share of military spending out of GDP (military) could be useful to include. Military spending might be thought to compete with education spending, especially insofar as the budget share of each is concerned. Fourth, the fiscal balance (fiscbal2) might also have a stronger effect on the budget share measure of education spending. For instance, the fiscal balance might affect real spending on education, but not necessarily the share of spending out of GDP. Also, governments that tend to run budget surpluses might more likely prioritise education spending during these times because education is arguably a more 'productive' component of public spending. Lastly, given the link between the fiscal balance and gross public debt stock, including a control measure for debt might also be useful. Debt accumulation might be viewed as a constraint on various forms of public spending.

The other sensible way to change the specification was to use time (year) fixed effects. Year fixed effects – year-specific, country-invariant effects – might also play a useful role in controlling for endogenous common effects ('common shocks') under the assumption that such effects are relatively homogeneous across panels (countries) for the particular user-defined groups tested in each case. There are a number of studies of education spending that have controlled for various kinds of common shocks or country-invariant effects, by using time dummies (see, for example, Afonso & Jalles, 2013; Arze del Granado et al., 2013; Baskaran & Hessami, 2012; Speciale, 2012; Akanbi & Schoeman, 2010; Busemeyer, 2007 and Avelino et al., 2005).

For instance, suppose a global shock has a more or less homogeneous effect on a particular group of countries, this might arguably imply similar common effects for education spending (cross-sectional correlation) – not necessarily in mean terms, but possibly in terms of how these means change over time. Many of these countries may also be in close regional proximity to one another, implying a process of spatial correlation in spending patterns. In both cases, a problem of cross-section dependence in the data-generating process may be exhibited. Therefore, it might be remiss to estimate models without the use of year fixed effects. To this end and for the sake of completeness, year fixed effects were used in all the model specifications as part of the robustness tests. Although the estimates for the various years are not reported in the applicable tables of results (see Appendix D, E, F and G) to save space, they are discussed as part of the Analysis and Discussion section.

4. Data Collection and Description of the Data

4.1 Data Collection

Table 2 presents the data that were collected. Bear in mind the World Bank acts as a repository of sorts for data from various other sources, so the sources quoted are the publicly accessible sources from which the data was obtained (see the respective sources for the original source data). For example, although one can source education spending data from the World Bank's Education Statistics (EdStats) database, these data come from UNESCO's Institute for Statistics (UIS) database. Details pertaining to the Stata variable name used; description of the applicable variable; type of measure; total sample size; effective years of data coverage and the source from which the data was obtained, are given.

Appendix A and B list the names of countries comprising the variables ypc201521 and region3, respectively. Note the use of the term 'variable' here for the categorical explanatory variables of interest, actually means each is really a single dummy (binary) variable or a 'family' (set) of dummy variables. The primary focus for the OECD group of richest countries is the 21 'core' countries, countries that have been classified as such for the entire time period under investigation (1989 to 2015). These 21 countries are also the same as those used by Busemeyer (2007) in his study of education spending for OECD democracies (from roughly 1980 to 2001), and excludes Czech Republic, Estonia, Hungary, Iceland, Israel, Korea, Luxembourg, Poland, Slovak Republic and Slovenia, which would comprise 31 OECD countries. One could have also considered using 35 OECD countries (the 31 OECD countries plus Chile, Latvia, Mexico and Turkey). However, the empirical patterns using 31 or 35 OECD countries were very similar to those when using 21 OECD countries only. Since the 21 OECD countries comprise a group that has remained unchanged since 1989, it was decided to focus on this group of OECD countries as the richest cohort with respect to GNI per capita country grouping. A list of countries is not compiled for the poldemoc dummy variable because these observations do sometimes vary (albeit very slowly) by year for a specific country, whereas for ypc201521 and region3, they remain unchanged over time.

The use of the regional dummy (region3) is very purposeful. It offers an alternative 'economic' interpretation based on the regional categorisation of countries because we know there are both richer and poorer parts of the world. In other words, one way to 'triangulate'[10] the results for the economic (GNI per capita grouping) and political (state of polity) interaction model is to test a regional (a sub-sample of richer versus poorer country regions) and political interaction model. The richer group comprises the countries of North America, Nordic Countries and Western Europe. There are no countries classified as not politically democratic in this sub-sample of countries, and predominantly comprises OECD countries as well as some high-income (non-OECD) countries located in North America and Western Europe (i.e., Bermuda, Andorra, Channel Islands, Faroe Islands, Gibraltar, Greenland, Isle of Man, Liechtenstein, Monaco and San Marino). The poorer group comprises the countries of, Central America, South America (excluding Chile and Uruguay), West Africa, Central Africa (excluding Equatorial Guinea), East Africa, Southern Africa (excluding Seychelles), South Asia and Southeast Asia (excluding Brunei Darussalam and Singapore). The poorer group comprises a sub-sample of low-income, lower middle-income and upper middle-income countries. The abovementioned exclusions are either because they are classified as being high-income countries (Chile, Uruguay, Seychelles, Brunei Darussalam and Singapore) or their economic behaviour is generally not in keeping with other 'poorer' countries over most of the time period in question (Equatorial Guinea). Irrespective of the fact that the richer cohort includes a mixture of OECD and high-income (non-OECD) countries – the latter (non-OECD) which might reasonably be expected to behave somewhat differently to the former (OECD) group of countries in education spending terms – and that the poorer cohort includes a mixture of low-income and middle-income countries, the regional dummy variable is still a useful way to check for the robustness of patterns by using different implicit sub-samples of countries to that of the economic (GNI per capita) grouping of countries.

[10] Triangulation here refers to a term I borrow from the realm of qualitative research which, for my purposes in the context of this quantitative research, loosely means to check for the validity of an empirical finding by conducting an analysis using a different, but related, economic dimension.

Table 2: Data Collected, Descriptions and Sources

Stata Variable Name	Description of the Variable	Type of Measure	Sample Size	Years	Source
	Dependent Variables				
psegdptot	Public spending on education, total (% of GDP)	Continuous (Interval or Ratio)	2551	1989-2015	World Bank EdStats
psegovtot	Public spending on education, total (% of total government expenditure.)	Continuous (Interval or Ratio)	2255	1989-2015	World Bank EdStats
	Explanatory Variables of Interest				
ypc201521	GNI per capita country grouping 2015, 21 OECD countries	Categorical (Ordinal)	5859	1989-2015	World Bank (Atlas Method)
region3	Richer versus poorer (bi-modal) country region sub-samples	Categorical (Nominal)	3024	1989-2015	Author's compilation
poldemoc	Political democracy classification: yes; no	Categorical (Nominal)	5105	1989-2015	Freedom House
	Control Variables				
pop024	Population ages 0-24 (% of total)	Continuous (Interval or Ratio)	4714	1990-2015	World Bank EdStats
urban	Urban population (% of total)	Continuous (Interval or Ratio)	5799	1989-2015	World Bank WDI
trade	Trade or exports plus imports of goods & services (% of GDP)	Continuous (Interval or Ratio)	4785	1989-2015	World Bank WDI
hci	Human capital index (PWT 9.0)	Continuous (Interval or Ratio)	3703	1989-2014	Penn World Table 9.0
pop65	Population ages 65 and above (% of total)	Continuous (Interval or Ratio)	5234	1989-2015	World Bank WDI
military	Military expenditure (% of GDP)	Continuous (Interval or Ratio)	3870	1989-2015	World Bank WDI
fiscbal2	Fiscal balance (% of GDP)	Continuous (Interval or Ratio)	4184	1990-2015	World Bank DPG
debt2	General government gross debt (IMF, % of GDP)	Continuous (Interval or Ratio)	3796	1989-2015	World Bank TCdata360

Notes: EdStats refers to the World Bank's Education Statistics database (see World Bank, 2017a). TCdata360 refers to the World Bank's TCdata360 database (see World Bank, 2017b). WDI refers to the World Bank's World Development Indicators database (see World Bank, 2017c). DPG refers to the World Bank's Development Prospects Group: A Cross-Country Database of Fiscal Space (see World Bank, 2017d). Freedom House refers to the Freedom in the World survey data (see Freedom House, 2016). The pop024 variable is the sum of pop014 and pop1524 variables available from the World Bank EdStats database. See Feenstra et al. (2015) for the Penn World Table 9.0 source.

4.2 Descriptive Statistics

The descriptive statistics presented comprise several parts. First, the basic pooled (overall), between countries and within countries summary descriptive statistics for the continuous variables are presented (Table 3). Second, the frequency or count data for the categorical explanatory variables of interest are outlined (Table 4). Third, box plots of the two education spending variables by the separate explanatory variables of interest are shown (Appendix C).

Tables 3 and 4 report the necessary descriptive statistics. For the explanatory variables of interest (which are all categorical measures), I merely present the number of countries for each group of the respective categorical measure (see Table 4). Table 3 presents the respective descriptive statistics for the various continuous measures. Apart from detailing average behaviour and variation with respect to each variable, they also give a good idea of the overall sample size, number of countries and average number of years for each variable of the panel data. A notable comment concerning the education spending measures as compared to all the other measures is the relatively smaller average number of time-series observations – roughly half in most cases. This shows the difficulty of modelling education spending because of data availability issues, not specific to this study only, but a more pervasive problem presented in the general literature, which is not always explicitly mentioned though. Thus, more strongly unbalanced panel data for national-level education spending are usually the norm, rather than the exception, irrespective of the time period under investigation.

Table 3: Descriptive Statistics for the Continuous Variables

Variable	Observations	Countries	Years	Mean	Standard Deviation			Minimum	Maximum
					Overall	Between	Within		
psegdptot	2551	193	13.2	4.505	2.007	1.988	1.109	0.781	44.334
psegovtot	2255	181	12.5	14.849	5.036	4.566	2.742	2.563	47.279
pop024	4714	184	25.6	49.977	13.687	13.230	3.451	20.160	73.288
urban	5799	215	27.0	55.788	24.901	24.736	3.363	5.342	100.000
trade	4785	193	24.8	86.996	52.290	46.836	21.717	0.021	531.737
hci	3703	143	25.9	2.342	0.694	0.677	0.165	1.028	3.734
pop65	5234	195	26.8	7.073	4.814	4.685	1.113	0.697	26.342
military	3870	166	23.3	2.433	3.210	2.577	2.235	0	117.388
fiscbal2	4184	191	21.9	-2.299	13.715	5.706	12.383	-505.442	122.188
debt2	3796	186	20.4	57.015	49.714	38.888	33.032	0	789.833

Notes: Years refers to the average number of years (time-series observations) for each country. The only two substantive changes made to the original data (the psegdptot variable only) were the deletion of observations for Turkey in 1998 (0 was deleted because there were no other 0% values in the dataset; nil or negligible appeared in the original UNESCO source data for this observation) and Tuvalu in 1997 (3730833.5% was deleted as an obvious mistake, which also appeared as such in the original UNESCO source data for this observation).

Table 4: Descriptive Statistics for the Categorical Variables

ypc201521	Freq.	region3	Freq.	poldemoc	Freq.
Low income	837	Poorer country regions	2214	No	2063
Lower middle income	1404	Richer country regions	810	Yes	3042
Upper middle income	1512	**Observations**	3024	**Observations**	5105
High income (non-OECD)	1539				
High income (OECD)	567				
Observations	5859				

Notes: Freq. refers to frequency or count.

Appendix C shows the box plots of the two education spending measures for each of the three separate categorical explanatory variables of interest (economic, regional and political). The box plots are a useful nonparametric way of visualising the relationship between the dependent variable (education spending) and the various groups of the applicable categorical (independent) variable. Box plots are also a useful way to conceptualise the regression framework of performing multiple t-tests. A graphical rule of thumb can be applied to ascertain whether or not the medians of two groups are likely to be different.[11] The positions of the medians (and their associated absolute positions in the visual or plot space) would be referred to as unconditional medians, and if shown to be significantly different to one another, would comprise unconditional median differences. A similar logic would apply if one assumed the medians were mean values. However, when other covariates are added, as would be the case when testing for median (or mean) differences using a multiple linear regression framework, the absolute positions (and possibly even relative positions) of the medians (or means) and respective boxes in the visual space are likely to change, which might also change the significance of any differences with respect to a base or reference group in the regression model.

For the sake of generating the appropriate box plot in each case, all data was included, but for 'scaling' purposes, I excluded Zimbabwe when plotting the relationships between the national effort measure and respective categorical variable because of several extreme values. The various box plots are meant to be illustrative of possible patterns and not definitive of empirical patterns, *per se*. A general observation from the various box plots is that richer countries spend more in national effort terms, but less in budget share terms, in accordance with theoretical (and empirical) expectations. However, although politically democratic countries should be shown to spend more in terms of both spending measures, the budget share measure might seem to suggest otherwise. A simple explanation for this visual 'anomaly' is that the (negative) 'economic effect' of having many richer countries represented in the politically democratic group is likely be working against the (positive) 'political effect' in this respect. For instance, selecting 'poorer' countries only, shows a somewhat higher median value for poorer countries that are politically democratic, which would be in keeping with expectations.

5. Analysis and Discussion

The analysis and discussion comprises two parts. The first part opens with some notes and a number of general findings, including the common 'themes' running through all of the empirical results. A more specific discussion of the results and findings related to Tables 5 and 6 is then given with respect to answering the research question. The second part then presents a discussion, and contextualises the general findings by highlighting particular differences for economic or regional (and politically distinct) groupings of countries. To conclude this section a synthesis of the empirical results is given by proffering three theoretical inequality propositions (see Figure 2).

[11] A sample size of at least 30 is generally needed to make inferences about population behaviour when comparing two groups in a box plot. A graphical method to compare the medians from two box-and-whisker plots is to compute a ratio equal to the vertical distance between medians divided by the total vertical distance covered by the two boxes. For sample sizes between 30 and 100, if the ratio is greater than 0.33, then there is likely to be a difference. For sample sizes between 100 and 1000, if the ratio is greater than 0.2, then there is likely to be a difference. And, for sample sizes greater than 1000, if the ratio is greater than 0.1, then there is likely to be a difference. This explanation has been partly adapted from that currently appearing on the Nayland College, Department of Mathematics website: Retrieved from
http://maths.nayland.school.nz/Year_11/AS1.10_Multivar_data/11_Comparing_Boxplots.htm

5.1 Results and Findings

There are some notes and several general findings, which apply to the tables presented (see Tables 5 to 7). The base or reference group is the richest/richer and politically democratic group of countries in each case. For the results reported in Tables 5 and 6, I do also present Wald Test results (Table 7) to evaluate whether or not significant differences exist between various pairwise groupings of countries from the same economic or regional group. For example, it would be very useful to know whether (mean) education spending in low-income politically democratic countries differs significantly from that of low-income countries that are not politically democratic. Hence, the Wald test results control for the economic or regional grouping to show if there are mean differences by different states of polity.

Insofar as the general findings are concerned, there are a number of noteworthy comments to make prior to briefly discussing the results presented in each table. Firstly, in broad terms, the control variables in each case report the 'correct' (positive) hypothesised sign for pop024 and urban. The trade variable generally seems to come out positive in most cases, which would be more supportive of the compensation hypothesis. Thus, the applicable controls used in each case generally report significant sensible effects. Secondly, for the various estimators of the standard errors, clustering by country (as opposed to clustering by year or by country and year) seems to have the most pejorative effect on the significance of the estimated mean differences. This would not seem too surprising if one considers there are many countries for which very few observations of the dependent variable exist, meaning it becomes particularly problematic to estimate precise standard errors when clustering by country. However, this is arguably less problematic for clustering by year, or clustering by country and year (two-way clustering), for that matter. Thirdly, although the explanatory power of each model (shown by the estimated R-squared) is not of any real importance, what should be readily noticeable is how accounting for outliers using the two methods of weighting observations (quantile and robust estimators) qualitatively improves the goodness of fit. Once again, this is not a surprising result, in itself, but does highlight the importance of using an estimator that takes extreme values (with respect to the outcome variable, in particular) into consideration.

Fourthly, although the high-income (non-OECD) group of countries generally seem to spend less on education relative to their high-income (OECD) counterparts, it is the latter (OECD) group of countries I am most interested in making comparisons against because we know that high-income (non-OECD) countries are 'contaminated', in part, by oil-producing countries (e.g., Kuwait; Qatar; Saudi Arabia and United Arab Emirates), politically different types of countries (e.g., Brunei Darussalam and Singapore versus San Marino and Chile), and tax havens (e.g., Isle of Man and Monaco). In other words, although there is bound to be a greater or lesser degree of heterogeneity within income (or political) groupings of countries, the degree of heterogeneity within the high-income (non-OECD) country grouping is likely to be more pronounced, hence, why the results for the low-income and middle-income country groups is of greater interest. However, from a political perspective and for interest sake, I do show results for politically different types of high-income (non-OECD) countries too. Lastly, and most importantly, the general finding of a 'reversal' in the pattern of mean differences for the national effort versus the budget share measure of education spending between different economic (or regional) ***and*** political groupings of countries is most noteworthy. These estimated patterns lend credence to a particular empirical regularity, which is elaborated upon later in this section.

Tables 5a and 5b, and Tables 6a and 6b present the results for the two factor-variable interaction models. To reiterate, since both models essentially combine an economic and a political categorical variable, only three controls (pop024, urban and trade) were used. The two different (interaction) specifications reveal qualitatively similar results – the empirical 'theme' with respect to richer and poorer countries are very similar. Interacting the economic (or regional) and political dimensions, generally reveals a pattern of significant *negative* mean differences with respect to the national effort measure, and significant *positive* mean differences with respect to the budget share measure. These patterns are generally robust to using different estimators and changing the specification (no controls versus three controls used). An interesting, more subtle, pattern of spending behaviour emerges in these economic and political or regional and political interaction models that gives a more nuanced political-economy explanation to education spending patterns.

The Wald test results reported in Table 7 generally show that the state of polity does play an important mediating role because there are many significant mean differences within the same economic or regional grouping of countries. Put differently, within income groups (controlling for economic status) or within the poorer country regions (controlling for regional country grouping), there is evidence to suggest that different political sub-groups of countries behave differently. Regardless of the spending measure considered (national effort or budget share), one might expect that, controlling for economic (income) or regional grouping of countries, politically democratic countries should spend more than their non-democratic counterparts because of the majoritarian or populist socio-political pressures exerted upon governments in these countries. Bear in mind that, for high-income (OECD) countries or richer country regions, neither of these have countries categorised as not politically democratic.

Focusing on the models with controls (Model 2), except for the statistical comparison of the estimates for the low-income group of countries using the national effort measure, all other group comparison tests conducted (see Table 7), reported significant differences within economic or regional groups based on different states of polity. These results substantiate the abovementioned expectation that, with respect to both spending measures, politically democratic countries spend more than their non-democratic counterparts, *ceteris paribus*. Therefore, generally speaking, although relatively poorer countries spend significantly less (more) on education in national effort (budget share) terms relative to richer countries, significant patterns of mean differences also tend to exist within (poorer) economic groupings of countries, showing that the state of polity – or a political-economy interpretation – matters. This finding accords with the theorised hypotheses given in the Literature Review section concerning the effect of political democracy.

Lastly, in addition to considering the robustness of the patterns of spending behaviour using three control variables and different estimators, I also tried a couple of other robustness tests that comprised changes in specification (i.e., additional control variables and inclusion of year fixed effects). These results and concise summary conclusions are reported in Appendix D, E, F and G. Only the LSDV and robust estimators were used in conjunction with the regional and political factor-variable interaction model specification – richer versus poorer country regions interacted with the state of political democracy – because this model specification comprises a synthesis and is at the heart of what this paper attempts to show. The estimates for the coefficients of the various year fixed effects and applicable additional controls are excluded from the tables to save space. A detailed description of each model specification is given in the notes to Appendix D.

Three major points can be gleaned from Appendix D, E, F and G. First, regardless of which pair of models is considered (Model 1A and Model 1B; Model 2A and Model 2B etc.), using the Akaike or Bayesian information criterion (AIC or BIC) as the yardstick, the addition of year fixed effects did not seem to add value – in a purely informational sense – to the model specifications in all cases. However, from a purely 'raw' fit perspective (using R-squared as the yardstick), obviously the addition of year fixed effects adds value from an explained variation point of view. The difference being the information criterion takes into account the trade-off between a higher R-squared and lower F-statistic from the inclusion of the year fixed effects. On this basis, adding year fixed effects lowered the informational value of the models. The general finding was that including year fixed effects weakened the results somewhat, as expected, but left the substantive patterns of mean differences unchanged.

Second, the use of a robust estimator was purposeful. Taking cognisance of extreme observations or outliers might be of particular importance when working with education spending data that often exhibit extreme observations in one direction or the other. There is little doubt that outlier observations in certain sub-groups or sub-samples of the data might be driving particular results. In general terms, using a robust estimator meant the results and associated empirical patterns were exhibited more clearly regardless of what specification was employed.

Third, the coefficient estimates for the various additional controls (hci, pop65, military, fiscbal2 and debt2) generally reported sensible effects in most cases. Under the assumption of homogeneous slopes, introducing an additional control each time revealed more complex specifications that either did not confound or only partially confounded the empirical patterns evidenced.[12] In actual fact, the most comprehensive specification (using eight controls) generally seemed to improve or substantiate the empirical patterns. Therefore, on this basis and for the purposes of this paper, more parsimonious specifications comprising three controls only might be defensible on the grounds of the patterns of spending behaviour (in particular, those exhibited in Tables 6a and 6b) generally being robust to certain sensible changes in the specification.

[12] Partial confounding refers to the case where only poorer countries that are not politically democratic were shown to be significantly different from the base group (richer and politically democratic countries) and with the correct expected sign. No confounding refers to the case where, either, both poorer country groups (irrespective of state of polity), or, poorer and politically democratic countries were shown to be significantly different from the base group and with the correct expected sign.

Table 5a: Mean Differences in the National Effort Measure by GNI per capita & Political Democracy (Using 21 OECD Countries)

Dependent Variable: psegdptot	LSDV		LSDV (One-way; Country)		LSDV (One-way; Year)		LSDV (Two-way; Both)		Newey-West		Driscoll-Kraay		Quantile		Robust	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1#0. Low income & not politically	-1.551***	-1.635***	-1.551***	-1.635*	-1.551***	-1.635***	-1.551***	-1.635*	-1.551***	-1.635***	-1.551***	-1.635***	-1.787***	-1.885***	-1.979***	-2.013***
democratic	(0.254)	(0.359)	(0.544)	(0.911)	(0.265)	(0.394)	(0.549)	(0.926)	(0.311)	(0.486)	(0.294)	(0.531)	(0.133)	(0.260)	(0.130)	(0.207)
1#1. Low income & politically	-1.684***	-1.756***	-1.684***	-1.756**	-1.684***	-1.756***	-1.684***	-1.756**	-1.684***	-1.756***	-1.684***	-1.756***	-1.757***	-1.877***	-1.724***	-1.759***
democratic	(0.123)	(0.257)	(0.330)	(0.774)	(0.129)	(0.255)	(0.332)	(0.773)	(0.178)	(0.389)	(0.185)	(0.371)	(0.127)	(0.257)	(0.153)	(0.216)
2#0. Lower middle income & not	-1.152***	-1.412***	-1.152***	-1.412**	-1.152***	-1.412***	-1.152***	-1.412**	-1.152***	-1.412***	-1.152***	-1.412***	-1.170***	-1.578***	-1.362***	-1.611***
politically democratic	(0.130)	(0.220)	(0.406)	(0.688)	(0.108)	(0.197)	(0.399)	(0.681)	(0.196)	(0.338)	(0.152)	(0.290)	(0.209)	(0.259)	(0.116)	(0.171)
2#1. Lower middle income &	-0.454***	-0.663***	-0.454	-0.663	-0.454***	-0.663***	-0.454	-0.663	-0.454*	-0.663**	-0.454***	-0.663***	-0.739**	-1.126***	-0.979***	-1.123***
politically democratic	(0.158)	(0.199)	(0.484)	(0.627)	(0.108)	(0.131)	(0.470)	(0.609)	(0.235)	(0.305)	(0.162)	(0.169)	(0.318)	(0.287)	(0.117)	(0.159)
3#0. Upper middle income & not	-1.090***	-1.389***	-1.090**	-1.389*	-1.090***	-1.389***	-1.090**	-1.389*	-1.090***	-1.389***	-1.090***	-1.389***	-1.357***	-1.669***	-1.432***	-1.764***
politically democratic	(0.155)	(0.230)	(0.534)	(0.801)	(0.131)	(0.140)	(0.527)	(0.780)	(0.244)	(0.369)	(0.195)	(0.193)	(0.166)	(0.177)	(0.129)	(0.146)
3#1. Upper middle income &	-0.819***	-1.097***	-0.819***	-1.097**	-0.819***	-1.097***	-0.819***	-1.097***	-0.819***	-1.097***	-0.819***	-1.097***	-0.797***	-1.131***	-0.935***	-1.133***
politically democratic	(0.097)	(0.132)	(0.295)	(0.423)	(0.085)	(0.119)	(0.292)	(0.419)	(0.141)	(0.203)	(0.092)	(0.159)	(0.106)	(0.156)	(0.106)	(0.127)
4#0. High income (non-OECD) &	-1.123***	-1.987***	-1.123**	-1.987***	-1.123***	-1.987***	-1.123**	-1.987***	-1.123***	-1.987***	-1.123***	-1.987***	-1.478***	-1.933***	-1.329***	-2.047***
not politically democratic	(0.183)	(0.217)	(0.514)	(0.702)	(0.197)	(0.192)	(0.519)	(0.695)	(0.249)	(0.332)	(0.273)	(0.259)	(0.152)	(0.195)	(0.171)	(0.189)
4#1. High income (non-OECD) &	-0.814***	-0.768***	-0.814**	-0.768**	-0.814***	-0.768***	-0.814**	-0.768**	-0.814***	-0.768***	-0.814***	-0.768***	-0.644***	-0.654***	-0.777***	-0.721***
politically democratic	(0.091)	(0.101)	(0.338)	(0.357)	(0.049)	(0.064)	(0.329)	(0.348)	(0.144)	(0.156)	(0.059)	(0.085)	(0.123)	(0.130)	(0.106)	(0.111)
5#0. High income (OECD) & not	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
5#1. High income (OECD) &	5.352***	3.668***	5.352***	3.668***	5.352***	3.668***	5.352***	3.668***	5.352***	3.668***	5.352***	3.668***	5.226***	3.107***	5.325***	3.449***
politically democratic BASE	(0.055)	(0.312)	(0.225)	(0.945)	(0.058)	(0.171)	(0.226)	(0.908)	(0.089)	(0.467)	(0.083)	(0.200)	(0.056)	(0.334)	(0.073)	(0.260)
Youth population		0.016***		0.016		0.016***		0.016		0.016*		0.016**		0.022***		0.018***
		(0.006)		(0.018)		(0.005)		(0.018)		(0.009)		(0.007)		(0.006)		(0.004)
Urban population		0.009***		0.009		0.009***		0.009		0.009**		0.009***		0.011***		0.011***
		(0.002)		(0.008)		(0.001)		(0.007)		(0.004)		(0.002)		(0.003)		(0.002)
Trade		0.008***		0.008**		0.008***		0.008**		0.008***		0.008***		0.008***		0.007***
		(0.001)		(0.004)		(0.001)		(0.003)		(0.002)		(0.001)		(0.001)		(0.001)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.060	0.100	0.060	0.100	0.060	0.100	0.060	0.100	n/a	n/a	0.060	0.100	0.074	0.107	0.123	0.190
F-value	35.86***	34.89***	3.78***	3.79***	78.63***	120.42***	n/a	n/a	15.83***	15.72***	60.00***	387.40***	n/a	n/a	43.22***	48.51***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000
Countries	183	169	183	169	183	169	183	169	183	169	183	169	183	169	183	169
Years	27	26	27	26	27	26	27	26	27	26	27	26	27	26	27	26
Observations	2468	2288	2468	2288	2468	2288	2468	2288	2468	2288	2468	2288	2468	2288	2468	2288

Notes: BASE group is high income (OECD) & politically democratic countries. The model uses 3 controls: youth population (pop024); urban population (urban) and trade (trade). Because the controls use homogeneous slopes, the parameter estimates for each control are computed at the mean of all country-year observations included in the estimation sample for the respective control variable. A pseudo R-squared is reported for the Quantile regression. Not applicable (n/a) means the respective statistic was not available or not reported. The number of countries and years were taken from the results reported for the various estimators of the standard errors. The LSDV and Quantile estimators use Huber/White heteroskedasticity-robust standard errors. The various LSDV estimators use one-way (country or year) and two-way (country and year or both) cluster-robust standard errors. The Newey-West and Driscoll-Kraay estimators use their own covariance matrix corrections to compute heteroskedasticity and autocorrelation consistent (HAC), and cross-sectional or "spatial" correlation consistent standard errors under different data-generating assumptions, respectively. The various standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5b: Mean Differences in the Budget Share Measure by GNI per capita & Political Democracy (Using 21 OECD Countries)

Dependent Variable: psegovtot	LSDV		LSDV (One-way; Country)		LSDV (One-way; Year)		LSDV (Two-way; Both)		Newey-West		Driscoll-Kraay		Quantile		Robust	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1#0. Low income & not politically democratic	3.633***	-0.012	3.633***	-0.012	3.633***	-0.012	3.633***	-0.012	3.633***	-0.012	3.633***	-0.012	3.948***	-0.162	3.387***	-0.210
	(0.436)	(0.726)	(1.208)	(2.003)	(0.292)	(0.682)	(1.163)	(1.988)	(0.644)	(1.069)	(0.314)	(0.984)	(0.608)	(0.696)	(0.392)	(0.614)
1#1. Low income & politically democratic	4.796***	1.339*	4.796***	1.339	4.796***	1.339*	4.796***	1.339	4.796***	1.339	4.796***	1.339	5.135***	1.290**	4.821***	1.467**
	(0.381)	(0.686)	(0.832)	(1.748)	(0.310)	(0.754)	(0.802)	(1.776)	(0.500)	(0.966)	(0.451)	(1.092)	(0.366)	(0.581)	(0.467)	(0.643)
2#0. Lower middle income & not politically democratic	4.304***	1.081*	4.304***	1.081	4.304***	1.081*	4.304***	1.081	4.304***	1.081	4.304***	1.081	4.231***	0.614	4.196***	0.850*
	(0.368)	(0.600)	(1.148)	(1.791)	(0.348)	(0.563)	(1.142)	(1.779)	(0.555)	(0.899)	(0.468)	(0.803)	(0.595)	(0.531)	(0.353)	(0.511)
2#1. Lower middle income & politically democratic	4.852***	2.421***	4.852***	2.421*	4.852***	2.421***	4.852***	2.421**	4.852***	2.421***	4.852***	2.421***	4.137***	1.661***	4.224***	1.830***
	(0.369)	(0.446)	(1.009)	(1.236)	(0.214)	(0.390)	(0.963)	(1.217)	(0.548)	(0.639)	(0.219)	(0.440)	(0.470)	(0.419)	(0.348)	(0.468)
3#0. Upper middle income & not politically democratic	2.520***	-0.022	2.520*	-0.022	2.520***	-0.022	2.520*	-0.022	2.520***	-0.022	2.520***	-0.022	2.058***	-0.131	2.317***	-0.128
	(0.455)	(0.503)	(1.380)	(1.502)	(0.461)	(0.378)	(1.382)	(1.465)	(0.663)	(0.749)	(0.668)	(0.547)	(0.748)	(0.371)	(0.405)	(0.450)
3#1. Upper middle income & politically democratic	3.643***	1.708***	3.643***	1.708	3.643***	1.708***	3.643***	1.708	3.643***	1.708***	3.643***	1.708**	3.556***	1.205***	3.411***	1.347***
	(0.279)	(0.354)	(0.862)	(1.053)	(0.230)	(0.415)	(0.847)	(1.076)	(0.409)	(0.524)	(0.280)	(0.617)	(0.331)	(0.179)	(0.318)	(0.374)
4#0. High income (non-OECD) & not politically democratic	0.987**	-2.057***	0.987	-2.057	0.987**	-2.057***	0.987	-2.057	0.987	-2.057**	0.987*	-2.057***	0.034	-3.107***	0.683	-2.852***
	(0.499)	(0.566)	(1.699)	(1.648)	(0.378)	(0.407)	(1.668)	(1.600)	(0.770)	(0.830)	(0.516)	(0.339)	(0.751)	(0.356)	(0.502)	(0.543)
4#1. High income (non-OECD) & politically democratic	0.568**	-0.038	0.568	-0.038	0.568***	-0.038	0.568	-0.038	0.568	-0.038	0.568**	-0.038	0.923***	0.308	0.471	-0.099
	(0.249)	(0.254)	(0.834)	(0.815)	(0.188)	(0.211)	(0.818)	(0.803)	(0.372)	(0.379)	(0.260)	(0.284)	(0.308)	(0.342)	(0.333)	(0.338)
5#0. High income (OECD) & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
5#1. High income (OECD) & politically democratic BASE	12.245***	4.332***	12.245***	4.332**	12.245***	4.332***	12.245***	4.332**	12.245***	4.332***	12.245***	4.332***	11.973***	0.885	12.217***	2.576***
	(0.122)	(0.831)	(0.500)	(2.156)	(0.088)	(0.602)	(0.493)	(2.079)	(0.196)	(1.172)	(0.128)	(0.621)	(0.166)	(0.643)	(0.214)	(0.768)
Youth population		0.146***		0.146***		0.146***		0.146***		0.146***		0.146***		0.191***		0.164***
		(0.015)		(0.038)		(0.013)		(0.038)		(0.021)		(0.017)		(0.010)		(0.013)
Urban population		0.029***		0.029		0.029***		0.029		0.029***		0.029***		0.046***		0.040***
		(0.007)		(0.020)		(0.006)		(0.020)		(0.010)		(0.009)		(0.006)		(0.006)
Trade		0.017***		0.017***		0.017***		0.017***		0.017***		0.017***		0.022***		0.021***
		(0.003)		(0.006)		(0.003)		(0.006)		(0.004)		(0.002)		(0.003)		(0.002)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.141	0.198	0.141	0.198	0.141	0.198	0.141	0.198	n/a	n/a	0.141	0.198	0.092	0.146	0.136	0.226
F-value	64.55***	56.40***	7.16***	8.08***	234.33***	201.25***	n/a	n/a	29.30***	26.94***	315.69***	476.80***	n/a	n/a	42.96***	54.58***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000
Countries	175	165	175	165	175	165	175	165	175	165	175	165	175	165	175	165
Years	27	26	27	26	27	26	27	26	27	26	27	26	27	26	27	26
Observations	2194	2069	2194	2069	2194	2069	2194	2069	2194	2069	2194	2069	2194	2069	2194	2069

Notes: BASE group is high income (OECD) & politically democratic countries. The model uses 3 controls: youth population (pop024); urban population (urban) and trade (trade). See the notes for Table 5a. The various standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6a: Mean Differences in the National Effort Measure by Richer versus Poorer Country Regions & Political Democracy

Dependent Variable: psegdptot	LSDV		LSDV (One-way; Country)		LSDV (One-way; Year)		LSDV (Two-way; Both)		Newey-West		Driscoll-Kraay		Quantile		Robust	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
0#0. Poorer country regions & not politically democratic	-1.298***	-2.031***	-1.298***	-2.031***	-1.298***	-2.031***	-1.298***	-2.031***	-1.298***	-2.031***	-1.298***	-2.031***	-1.633***	-1.756***	-1.681***	-1.760***
	(0.147)	(0.216)	(0.396)	(0.582)	(0.128)	(0.168)	(0.389)	(0.566)	(0.199)	(0.315)	(0.152)	(0.218)	(0.117)	(0.234)	(0.104)	(0.199)
0#1. Poorer country regions & politically democratic	-0.804***	-1.412***	-0.804**	-1.412***	-0.804***	-1.412***	-0.804**	-1.412***	-0.804***	-1.412***	-0.804***	-1.412***	-1.163***	-1.235***	-1.110***	-1.178***
	(0.104)	(0.166)	(0.343)	(0.467)	(0.081)	(0.135)	(0.336)	(0.457)	(0.158)	(0.239)	(0.110)	(0.162)	(0.121)	(0.180)	(0.098)	(0.166)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	5.169***	2.095***	5.169***	2.095**	5.169***	2.095***	5.169***	2.095**	5.169***	2.095***	5.169***	2.095***	5.114***	2.047***	5.213***	2.520***
	(0.069)	(0.375)	(0.264)	(0.989)	(0.052)	(0.255)	(0.260)	(0.950)	(0.110)	(0.534)	(0.064)	(0.319)	(0.062)	(0.404)	(0.074)	(0.348)
Youth population		0.034***		0.034**		0.034***		0.034**		0.034***		0.034***		0.025***		0.019***
		(0.007)		(0.017)		(0.005)		(0.016)		(0.010)		(0.005)		(0.007)		(0.006)
Urban population		0.016***		0.016**		0.016***		0.016**		0.016***		0.016***		0.022***		0.020***
		(0.003)		(0.008)		(0.002)		(0.008)		(0.004)		(0.002)		(0.003)		(0.002)
Trade		0.013***		0.013***		0.013***		0.013***		0.013***		0.013***		0.011***		0.011***
		(0.002)		(0.004)		(0.002)		(0.004)		(0.002)		(0.002)		(0.002)		(0.001)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.057	0.167	0.057	0.167	0.057	0.167	0.057	0.167	n/a	n/a	0.057	0.167	0.090	0.170	0.154	0.313
F-value	52.61***	63.60***	5.45***	6.80***	148.23***	109.08***	n/a	n/a	24.89***	31.33***	161.84***	103.56***	n/a	n/a	135.10***	125.25***
p-value	0.000	0.000	0.006	0.000	0.000	0.000	n/a	n/a	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000
Countries	102	97	102	97	102	97	102	97	102	97	102	97	102	97	102	97
Years	27	26	27	26	27	26	27	26	27	26	27	26	27	26	27	26
Observations	1486	1382	1486	1382	1486	1382	1486	1382	1486	1382	1486	1382	1486	1382	1486	1382

Notes: BASE group is richer country regions that are politically democratic. The model uses 3 controls: youth population (pop024); urban population (urban) and trade (trade). See the notes for Table 5a. The various standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6b: Mean Differences in the Budget Share Measure by Richer versus Poorer Country Regions & Political Democracy

Dependent Variable: psegovtot	LSDV		LSDV (One-way; Country)		LSDV (One-way; Year)		LSDV (Two-way; Both)		Newey-West		Driscoll-Kraay		Quantile		Robust	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
0#0. Poorer country regions & not politically democratic	3.880***	1.492**	3.880***	1.492	3.880***	1.492***	3.880***	1.492	3.880***	1.492*	3.880***	1.492***	4.046***	0.726	3.757***	1.108*
	(0.305)	(0.612)	(0.909)	(1.609)	(0.229)	(0.360)	(0.887)	(1.531)	(0.452)	(0.892)	(0.210)	(0.466)	(0.499)	(0.764)	(0.303)	(0.597)
0#1. Poorer country regions & politically democratic	5.077***	3.321***	5.077***	3.321**	5.077***	3.321***	5.077***	3.321**	5.077***	3.321***	5.077***	3.321***	4.833***	2.284***	4.923***	2.945***
	(0.233)	(0.458)	(0.711)	(1.397)	(0.141)	(0.236)	(0.686)	(1.341)	(0.345)	(0.684)	(0.157)	(0.301)	(0.287)	(0.496)	(0.283)	(0.491)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	11.944***	5.333***	11.944***	5.333	11.944***	5.333***	11.944***	5.333*	11.944***	5.333***	11.944***	5.333***	11.849***	2.093**	11.943***	3.705***
	(0.127)	(1.193)	(0.506)	(3.265)	(0.116)	(0.952)	(0.504)	(3.185)	(0.203)	(1.719)	(0.176)	(1.039)	(0.160)	(0.960)	(0.212)	(1.036)
Youth population		0.104***		0.104*		0.104***		0.104*		0.104***		0.104***		0.163***		0.127***
		(0.021)		(0.059)		(0.013)		(0.056)		(0.030)		(0.015)		(0.019)		(0.019)
Urban population		0.029***		0.029		0.029***		0.029		0.029**		0.029***		0.044***		0.037***
		(0.008)		(0.023)		(0.008)		(0.023)		(0.012)		(0.009)		(0.008)		(0.007)
Trade		0.018***		0.018**		0.018***		0.018**		0.018***		0.018***		0.017***		0.021***
		(0.004)		(0.009)		(0.004)		(0.009)		(0.005)		(0.004)		(0.004)		(0.003)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.195	0.222	0.195	0.222	0.195	0.222	0.195	0.222	n/a	n/a	0.195	0.222	0.140	0.171	0.189	0.247
F-value	270.00***	113.21***	25.86***	13.06***	723.40***	323.35***	n/a	n/a	118.36***	51.83***	611.34***	290.39***	n/a	n/a	158.19***	84.82***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000	0.000	0.000	n/a	n/a	0.000	0.000
Countries	99	96	99	96	99	96	99	96	99	96	99	96	99	96	99	96
Years	27	26	27	26	27	26	27	26	27	26	27	26	27	26	27	26
Observations	1360	1299	1360	1299	1360	1299	1360	1299	1360	1299	1360	1299	1360	1299	1360	1299

Notes: BASE group is richer country regions that are politically democratic. The model uses 3 controls: youth population (pop024); urban population (urban) and trade (trade). See the notes for Table 5a. The various standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Wald Test Results for Parameter Equality of the Factor-Variable Interactions

Wald Tests for Parameter Equality from Tables 5a and 5b

	Robust Estimator					
	(1)			(2)		
Dependent Variable: psegdptot	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.0125)**	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.0125)**
Test parameter 1#0 = 1#1	F (1, 2459) = 2.19 p = 0.1388	No	No	F (1, 2276) = 2.33 p = 0.1273	No	No
Test parameter 2#0 = 2#1	F (1, 2459) = 8.90 p = 0.0029	Yes	Yes	F (1, 2276) = 14.79 p = 0.0001	Yes	Yes
Test parameter 3#0 = 3#1	F (1, 2459) = 14.38 p = 0.0002	Yes	Yes	F (1, 2276) = 24.17 p = 0.0000	Yes	Yes
Test parameter 4#0 = 4#1	F (1, 2459) = 10.20 p = 0.0014	Yes	Yes	F (1, 2276) = 51.44 p = 0.0000	Yes	Yes
	Robust Estimator					
	(1)			(2)		
Dependent Variable: psegovtot	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.0125)**	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.0125)**
Test parameter 1#0 = 1#1	F (1, 2185) = 7.33 p = 0.0068	Yes	Yes	F (1, 2057) = 11.37 p = 0.0008	Yes	Yes
Test parameter 2#0 = 2#1	F (1, 2185) = 0.00 p = 0.9438	No	No	F (1, 2057) = 6.75 p = 0.0094	Yes	Yes
Test parameter 3#0 = 3#1	F (1, 2185) = 6.88 p = 0.0088	Yes	Yes	F (1, 2057) = 13.88 p = 0.0002	Yes	Yes
Test parameter 4#0 = 4#1	F (1, 2185) = 0.16 p = 0.6852	No	No	F (1, 2057) = 25.58 p = 0.0000	Yes	Yes

Wald Tests for Parameter Equality from Tables 6a and 6b

	Robust Estimator					
	(1)			(2)		
Dependent Variable: psegdptot	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.05)**	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.05)**
Test parameter 0#0 = 0#1	F (1, 1483) = 33.66 p = 0.0000	Yes	Yes	F (1, 1376) = 35.80 p = 0.0000	Yes	Yes
	Robust Estimator					
	(1)			(2)		
Dependent Variable: psegovtot	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.05)**	**Wald Statistic**	**Uncorrected Significance (α = 0.05)**	**Corrected Significance (α = 0.05)**
Test parameter 0#0 = 0#1	F (1, 1357) = 16.59 p = 0.0000	Yes	Yes	F (1, 1293) = 37.70 p = 0.0000	Yes	Yes

Notes: The "#" naming convention accords with that in the respective table of results. Using interaction models with applicable controls, "Yes" means the applicable interaction parameters are significantly different ("No" means not significantly different) from one another for the respective pairwise comparison at the conventional (uncorrected) 5% critical level of significance or Bonferroni (corrected) critical level of significance. As before, "(1)" refers to the model with no controls (unconditional mean differences) and "(2)" refers to the model with controls (conditional mean differences). Because there is only one pairwise test of parameter equality performed on the estimates from Tables 6a and 6b, the alpha is the same (α = 0.05) for both the uncorrected and corrected critical level of significance.

5.2 Discussion

So, what can be inferred from myriad empirical results presented? An obvious place to start would seem to emanate from the following question, which presents a 'bi-modal' (two different, commonly-occurring, states of being) perspective on education spending: all things being equal, why do richer countries seem to make a greater national effort towards education (spend more on education as a share of GDP), but have a lower budget share (spend less on

education as a share of total government spending) relative to poorer countries?[13] The intuitive answer to this question, which partly seems to reconcile these two findings, is that richer countries are more likely to have larger public sectors (in terms of expenditure as a share of GDP) as compared to poorer countries – a point consistent with the Wagnerian hypothesis. But, there is also a political aspect to this argument, which augments the Wagnerian view, because education spending patterns seem to also be different among poorer countries with contrasting political environments.

A most noticeable pattern that applies to both measures of education spending is that politics seems to matter. Moreover, there generally appears to be significant differences between politically democratic and not democratic countries from the same income group (controlling for income grouping of country), implying that: either, political pressures compel governments in more politically democratic countries to spend more; or, it could be the case that growth in these types of countries makes it more possible to leverage taxes; or, democratic governments tend to be more educationally benevolent, anyway. The empirical evidence for poorer countries that are not politically democratic, having relatively low budget share measures not necessarily significantly different from rich (democratic) countries is possibly indicative of these poorer countries not only having smaller public sectors, but a smaller relative share of spending being allocated to education, which, from a human capital point of view, might partly explain why they remain poor and under-developed.

There are some plausible reasons for why the observed empirical patterns spoken about so far come about. Richer countries spend more as a share of national income (national effort), not because they necessarily value education more highly than do poorer countries, but because they have a greater capacity to leverage income from taxes. In other words, richer countries are less likely to be fiscally supply-side constrained relative to their poorer country counterparts, insofar as raising the necessary tax income is concerned to publicly finance various educational demands. However, poorer countries' inability to extract tax income from a relatively smaller fiscal (tax) base, constrains not only the growth of these types of countries' public sectors – a point noted by Holcombe (2005), albeit in more general terms about countries in the 21st Century – but also their 'national effort' towards education. For one, progressively poorer countries tend to have a greater preponderance of informal-sector, cash-based economic activity relative to the size of the formal private-sector economy, which implies it becomes increasingly problematic for governments in these types of economies to extract the tax income necessary to finance a larger variety of publicly provided goods and services – the fiscal varieties perspective I mentioned earlier – which is possibly one explanation for why public sector growth, as a whole, is suppressed in poorer countries. This is plausibly another contribution to the Wagnerian hypothesis along with the caveat mentioned earlier that 'politics matters' as expounded by Peacock & Scott (2000, pp. 3-4).

On the other hand, poorer countries tend to spend more as a share of total government spending (budget share) because they tend to have smaller public sectors as a whole, which means education (as a key budgetary component) would tend to comprise a larger share of the total public sector budget. However, since richer countries tend to have larger public sectors (a stylised result postulated by the Wagnerian hypothesis), they are more likely to

[13] For my purposes here, "richer" is generally synonymous with developed countries (OECD countries or regions predominantly occupied by OECD countries and some other high-income non-OECD countries) and "poorer" is synonymous with developing or lesser-developed countries (low- and middle-income countries or regions predominantly occupied by these types of countries). In this sense, I am able to proffer a bi-modal (richer versus poorer countries) explanation of education spending patterns.

have a greater variety of fiscal components to be financed through taxes and other income-generating mechanisms, which would not be the case in poorer countries. For instance, the growing role of the state in richer (developed) countries would foreseeably tend to crowd out other forms of public spending (e.g., education spending).[14] Consequently, in richer countries with larger public sectors and a greater variety of fiscal components to be serviced via the public purse (the fiscal varieties perspective), education's share of the total 'fiscal pie' would tend to be smaller, explaining then why public spending on education is lower (higher) as a share of total government spending in richer (poorer) countries.

Figure 2: Three Theoretical Inequality Propositions

Description	Richer Countries		Poorer Countries
Proposition 1 (national effort)	$\left(\frac{E}{Y}\right)_R$	>	$\left(\frac{E}{Y}\right)_P$
Proposition 2 (budget share)	$\left(\frac{E}{G}\right)_R$	<	$\left(\frac{E}{G}\right)_P$
Proposition 3 (public sector)	$\left(\frac{G}{Y}\right)_R$	>	$\left(\frac{G}{Y}\right)_P$

Notes: *E* refers to public spending on education, *Y* refers to national income (GDP) and *G* refers to total public spending. If Propositions 1 and 2 are shown to generally hold true, then, by implication, Proposition 3 will necessarily result.[15]

Figure 2 summarises the salient empirical findings by proffering three theoretical inequality propositions with respect to richer versus poorer countries, bearing in mind that different states of polity do also reveal differences in spending patterns within the poorer cohort of countries. To the best of my knowledge, this bi-modal perspective of education spending (Propositions 1 and 2), and, by implication, the relative size of the public sector (Proposition 3) in the two different types of countries, has not been presented like this before in the literature. This affords an interesting way to think about education spending patterns in richer (developed) versus poorer (developing) countries. Propositions 1 and 2 imply the existence of larger (smaller) public sectors in richer (poorer) countries, by using education spending patterns as a way to 'reverse engineer' the Wagnerian hypothesis. In essence, what the empirical work has revealed is a novel way to empirically justify Adolph Wagner's now well-established theoretical proposition.

[14] I do acknowledge here the rise or fall of the welfare state in developed countries is a matter of great contemporary debate, as covered somewhat exhaustively in Castles (2007), for example.

[15] Two interesting points to note here. First, it does not matter whether *E*, *Y* and *G* are measured in real or nominal terms, provided both the numerator and denominator of the respective ratio are both measured in the same nominal or real terms. For instance, if wanting to make a 'real' interpretation, assuming the same deflator is applied to both the numerator and denominator in each case, the ratio of the nominal measures is mathematically equivalent to the ratio of the real measures. Second, bear in mind the same estimated size of the public sector (as a share of GDP) in any one country, as given by sources such as the International Monetary Fund (IMF), cannot simply be gotten by taking the quotient of the national effort and budget share measures for that country because, for one, these education spending measures are estimates, in themselves. The quotient thereof will only give a rough approximation of the size of government, and most likely an even less accurate approximation for countries that have less accurate education spending data.

6. Conclusion

This paper sought to inquire whether or not mean differences exist in education spending among different economic (or regional) *and* political groupings of countries (two-way effects). Various empirical results were presented with respect to two different measures of education spending: the national effort and budget share measures. In general terms, it was found that there were significant mean differences between economically and politically distinct groups of countries, either without controls or including several or more important controls, patterns of differences that accord with theoretical (and other empirical) expectations. More specifically, two important empirical patterns were exhibited with respect to the (time-invariant) behaviour of education spending. Firstly, controlling for the state of economic development (same income per capita or regional grouping of countries), there was evidence to suggest that 'politics matters' – countries with different political make-ups from the same economic or regional grouping behave differently with respect to education spending. Secondly, and possibly more importantly for this inquiry, controlling for the state of polity (political democracy), it was generally found that richer (poorer) countries tend to spend a larger (smaller) share of GDP on education, but a smaller (larger) share of total government spending on education.

In other words, richer countries tend to make a greater national effort towards education and poorer countries tend to make a greater budget share prioritisation towards education. This conclusion lends itself well to proffering a 'bi-modal' perspective of education spending patterns. The bi-modal patterns of education spending evidenced are consistent with supply-side fiscal constraints in raising tax income faced by poorer countries, and richer countries being more able to finance a greater variety of societal needs and wants through the public-sector mechanism. Bigger governments or larger public sectors are a distinguishing feature of richer countries relative to their poorer country counterparts.

The general findings were summarised in the form of three theoretical inequality propositions. One might think of the empirical patterns exhibited as being an alternative way to justify the Wagnerian hypothesis, which postulates that a larger public sector is associated with the process of growth and development. Put another way, the inquiry was tantamount to using a simple methodological approach to enunciate particular patterns of education spending, but, as it turns out, also reveals a novel way to empirically validate the existence of Wagner's law.

The logical policy conclusion is that decision-makers and researchers now have a testable set of propositions with which to better understand a specific component of policy outcomes (national-level education spending) as countries move through the evolutionary process of growth and development – from poorer to richer country status. Part-and-parcel of understanding why these propositions come about concerns the way in which various countries are able to expand their respective public sectors. Supply-side fiscal constraints mean poorer country governments are less able to leverage income from taxes in a sufficient and sustainable way to meet national effort education spending targets more closely aligned with their richer country counterparts. Debt accumulation might seem like a reasonable way to raise the capacity of poorer countries to meet higher education financing targets, but this would only be a short-term measure (to finance, say, a specific short-term objective for the education sector, like building a number of new schools) and certainly not sustainable over the longer term. A larger public sector debt will only serve to further hamper progress towards growth and development in poorer countries.

A more sustainable mechanism would seem quite rhetorical: poorer countries need to grow their respective (formal) private-sector economies (using myriad public policy prescriptions), which is easier said than done! Failure to grow the formal private-sector economy or find innovative ways to leverage income from the burgeoning informal-sector economy, means the development reality for the poorest countries will only continue to be bleak. No doubt, development assistance and aid are important mechanisms to kick-start the process of growth and development in the poorest countries, but these mechanisms are arguably not sustainable over the longer term. The fact that richer countries have larger public sectors is not a matter of serendipity – they have larger formal private sectors too, meaning growth of the public and private sectors are strongly associated. The two sectors are best described as complements, as Tanzi and Schuknecht (2000, p. 2) seem to suggest, and not substitutes. In other words, with growth of the private and public sectors of the economy, public provision (and private provision too, for that matter) of education is less likely to be supply-side constrained.

There are a couple of useful avenues for future research that stem from the empirical findings and theoretical propositions advanced in this paper. For one, it would be of interest to know whether or not the spending-related theoretical propositions (Propositions 1 and 2) hold for other areas of the fiscal allocation (e.g., health spending or military spending). Another avenue concerns the political dimension considered. Ostensibly, a measure of political democracy was used. However, since all OECD countries are categorised as democratic over the time period in question (1989-2015), a study of how economic and partisan political forces (Left versus Right orientation) or more nuanced political categorisations interact with the economic measures to shape the ratio measures of total spending might be interesting.

REFERENCES

Adserà, A. & Boix, C. (2002). Trade, Democracy, and the Size of the Public Sector: The Political Underpinnings of Openness. *International Organization*, *56*(2), 229-262.

Afonso, A. & Alves, J. (2017). Reconsidering Wagner's Law: Evidence from the Functions of the Government. *Applied Economics Letters*, *24*(5), 346-350.

Afonso, A. & Jalles, J.T. (2013). The Cyclicality of Education, Health, and Social Security Government Spending. *Applied Economics Letters*, *20*(7), 669-672.

Akanbi, O.A. & Schoeman, N.J. (2010). The Determinants of Public Expenditure and Expenditure on Education in Particular in a Selection of African Countries. *South African Journal of Economic and Management Sciences*, *13*(1), 50-61.

Akitoby, B., Clements, B., Gupta, S. & Inchauste, G. (2006). Public Spending, Voracity, and Wagner's Law in Developing Countries. *European Journal of Political Economy*, *22*(4), 908-924.

Arze del Granado, J., Gupta, S. & Hajdenberg, A. (2013). Is Social Spending Procyclical? Evidence for Developing Countries. *World Development*, *42*(1), 16-27.

Avelino, G., Brown, D.S. & Hunter, W. (2005). The Effects of Capital Mobility, Trade Openness, and Democracy on Social Spending in Latin America, 1980-1999. *American Journal of Political Science*, *49*(3), 625-641.

Baqir, R. (2002). *Social Sector Spending in a Panel of Countries*. IMF Working Paper No. 02/35, International Monetary Fund, Washington, D.C. Retrieved from https://www.imf.org/external/pubs/ft/wp/2002/wp0235.pdf

Baskaran, T. & Hessami, Z. (2012). Public Education Spending in a Globalized World: Is There a Shift in Priorities Across Educational Stages? *International Tax and Public Finance, 19*(5), 677-707.

Baum, C.F., Nichols, A. & Schaffer, M.E. (2010). *Evaluating One-Way and Two-Way Cluster-Robust Covariance Matrix Estimates*. BOS'10 Stata Conference, July 2010. Retrieved from https://www.stata.com/meeting/boston10/boston10_baum.pdf

Becker, G.S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy, 70*(5), Part 2, 9-49.

Becker, G.S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Princeton, NJ: Princeton University Press.

Brennan, G. & Buchanan, J.M. (1980). *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.

Brown, D.S. & Hunter, W. (1999). Democracy and Social Spending in Latin America, 1980-1992. *American Political Science Review, 93*(4), 779-790.

Brown, D.S. & Hunter, W. (2004). Democracy and Human Capital Formation: Education Spending in Latin America, 1980 to 1997. *Comparative Political Studies, 37*(7), 842-864.

Buchanan, J.M. (1977). Why Does Government Grow? In T. Borcherding (Ed.), *Budgets and Bureaucrats: The Sources of Government Growth* (pp. 3-18). Durham, NC: Duke University Press.

Busemeyer, M.R. (2007). Determinants of Public Education Spending in 21 OECD Democracies, 1980-2001. *Journal of European Public Policy*, *14*(4), 582-610.

Busemeyer, M.R. (2008). The Impact of Fiscal Decentralisation on Education and Other Types of Spending. *Swiss Political Science Review, 14*(3), 451-481.

Busemeyer, M.R. (2009). Social Democrats and the New Partisan Politics of Public Investment in Education. *Journal of European Public Policy*, *16*(1), 107-126.

Cameron, A.C. & Miller, D.L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources, 50*(2), 317-372.

Cameron, D.R. (1978). The Expansion of the Public Economy: A Comparative Analysis. *American Political Science Review, 72*(4), 1243-1261.

Careja, R. & Emmenegger, P. (2009). The Politics of Public Spending in Post-Communist Countries. *East European Politics and Societies, 23*(2), 165-184.

Castles, F.G. (1982). The Impact of Parties on Public Expenditure. In F.G. Castles (Ed.), *The Impact of Parties: Politics and Policies in Democratic Capitalist States*. London: Sage.

Castles, F.G. (1986). Social Expenditure and the Political Right: A Methodological Note. *European Journal of Political Research, 14*(5-6), 669-676.

Castles, F.G. (1989). Explaining Public Education Expenditure in OECD Nations. *European Journal of Political Research, 17*(4), 431-448.

Castles, F.G. (1999). *Comparative Public Policy: Patterns of Post-war Transformation*. Cheltenham: Edward Elgar.

Castles, F.G. (2007). Testing the Retrenchment Hypothesis: An Aggregate Overview. In F.G. Castles (Ed.), *The Disappearing State? Retrenchment Realities in an Age of Globalisation* (pp. 19-43). Cheltenham: Edward Elgar.

Castles, F.G. & McKinlay, R.D. (1979). Does Politics Matter: An Analysis of the Public Welfare Commitment in Advanced Democratic States. *European Journal of Political Research, 7*(2), 169-186.

Cockx, L. & Francken, N. (2016). Natural Resources: A Curse on Education Spending? *Energy Policy, 92*, 394-408.

Cornwell, C. & Schmidt, P. (1984). *Panel Data with Cross-Sectional Variation in Slopes as Well as in Intercept*. Econometrics Workshop Paper No. 8404, Department of Economics, Michigan State University, Michigan.

De Hoyos, R.E. & Sarafidis, V. (2006). Testing for Cross-Sectional Dependence in Panel-Data Models. *Stata Journal, 6*(4), 482-496.

Driscoll, J.C. & Kraay, A.C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics, 80*(4), 549-560.

Feenstra, R.C., Inklaar, R. & Timmer, M.P. (2015). The Next Generation of the Penn World Table. *American Economic Review, 105*(10), 3150-3182. Data Retrieved from https://www.rug.nl/ggdc/productivity/pwt/

Fosu, A.K. (2010). The External Debt-Servicing Constraint and Public-Expenditure Composition in Sub-Saharan Africa. *African Development Review, 22*(3), 378-393.

Freedom House (2016). *Freedom in the World Survey Data*. Retrieved from https://freedomhouse.org/report/freedom-world/freedom-world-2016

Garritzmann, J.L. & Seng, K. (2016). Party Politics and Education Spending: Challenging Some Common Wisdom. *Journal of European Public Policy, 23*(4), 510-530.

Gemmell, N. (1990). Wagner's Law, Relative Prices and the Size of the Public Sector. *The Manchester School, 58*(4), 361-377.

Gemmell, N. (1993). Wagner's Law and Musgrave's Hypothesis. In N. Gemmell (Ed.), *The*

Growth of the Public Sector (pp. 103-120). Cheltenham: Edward Elgar.

Greene, W.H. (2008). *Econometric Analysis*. 6th edn. Upper Saddle River, NJ: Pearson.

Gujarati, D.N. & Porter, D.C. (2009). *Basic Econometrics*. 5th edn. Boston, MA: McGraw-Hill Irwin.

Herwartz, H. & Theilen, B. (2017). Ideology and Redistribution Through Public Spending. *European Journal of Political Economy, 46*, 74-90.

Hibbs, D.A. (1977). Political Parties and Macroeconomic Policy. *American Political Science Review*, *71*(4), 1467-1487.

Hibbs, D.A. (1987). *The Political Economy of Industrial Democracies*. Cambridge, MA: Harvard University Press.

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *Stata Journal*, *7*(3), 281-312.

Holcombe, R.G. (2005), Government Growth in the Twenty-First Century. *Public Choice*, *124*(1-2), 95-114.

Huber, E., Mustillo, T. & Stephens, J.D. (2008). Politics and Social Spending in Latin America. *Journal of Politics*, *70*(2), 420-436.

Imbeau, L.M. & Pétry, F. & Lamari, M. (2001). Left-Right Party Ideology and Government Policies: A Meta-Analysis. *European Journal of Political Research*, *40*(1), 1-29.

Iversen, T. (2001). The Dynamics of the Welfare State Expansion: Trade Openness, De-Industrialization, and Partisan Politics. In P. Pierson (Ed.), *The New Politics of the Welfare State* (pp. 45-79). Oxford: Oxford University Press.

Iversen, T. & Stephens, J.D. (2008). Partisan Politics, the Welfare State, and Three Worlds of Human Capital Formation. *Comparative Political Studies*, *41*(4/5), 600-637.

Jensen, C. (2011). Capitalist Systems, Deindustrialization, and the Politics of Public Education. *Comparative Political Studies*, *44*(4), 412-435.

Kaufman, R.R. & Segura-Ubiergo, A. (2001). Globalization, Domestic Politics, and Social Spending in Latin America: A Time-Series Cross-Section Analysis, 1973-97. *World Politics*, *53*(4), 553-587.

Kittel, B. & Obinger, H. (2003). Political Parties, Institutions, and the Dynamics of Social Expenditure in Times of Austerity. *Journal of European Public Policy*, *10*(1), 20-45.

Kuckuck, J. (2014). Testing Wagner's Law at Different Stages of Economic Development. *FinanzArchiv/Public Finance Analysis*, *70*(1), 128-168.

Kumar, S., Webber, D.J. & Fargher, S. (2012). Wagner's Law Revisited: Cointegration and Causality Tests for New Zealand. *Applied Economics*, *44*(5), 607-616.

Lake, D.A. & Baum, M.A. (2001). The Invisible Hand of Democracy: Political Control and the Provision of Public Services. *Comparative Political Studies, 34*(6), 587-621.

Magazzino, C., Giolli, L. & Mele, M. (2015). Wagner's Law and Peacock and Wiseman's Displacement Effect in European Union Countries: A Panel Data Study. *International Journal of Economics and Financial Issues, 5*(3), 812-819.

Manzano, D. (2013). Partisanship, Inequality and the Composition of Public Spending on Education. *Political Studies, 61*(2), 422-441.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy, 66*(4), 281-302.

Musgrave, R.A. & Peacock, A.T. (Eds.). (1958). *Classics in the Theory of Public Finance*. International Economic Association: Macmillan.

Newey, W.K. & West, K.D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica, 55*(3), 703-708.

Newey, W.K. & West, K.D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies, 61*(4), 631-653.

North, D.C. (1991). Institutions. *Journal of Economic Perspectives, 5*(1), 97-112.

North, D.C., Wallis, J.J. & Weingast, B.R. (2006). *A Conceptual Framework for Interpreting Recorded Human History*. National Bureau of Economic Research, Working Paper No. 12795, Cambridge, Massachusetts. Retrieved from http://www.nber.org/papers/w12795

Peacock, A. & Scott, A. (2000). The Curious Attraction of Wagner's Law. *Public Choice, 102*(1-2), 1-17.

Potrafke, N. (2011). Does Government Ideology Influence Budget Composition? Empirical Evidence from OECD Countries. *Economics of Governance, 12*(2), 101-134.

Ram, R. (1987). Wagner's Hypothesis in Time-Series and Cross-Section Perspectives: Evidence from "Real" Data for 115 Countries. *Review of Economics and Statistics, 69*(2), 194-204.

Ram, R. (1995). Public Educational Expenditures in the United States: An Analytical Comparison with Other Industrialized Countries. *Economics of Education Review, 14*(1), 53-61.

Schmidt, M.G. (1996). When Parties Matter: A Review of the Possibilities and Limits of Partisan Influence on Public Policy. *European Journal of Political Research, 30*(2), 155-183.

Schmidt, M.G. (2007). Testing the Retrenchment Hypothesis: Educational Spending, 1960-2002. In F.G. Castles (Ed.), *The Disappearing State? Retrenchment Realities in an Age of Globalisation* (pp. 159-183). Cheltenham: Edward Elgar.

Schultz, T.P. (1988). Expansion of Public School Expenditures and Enrollments: Intercountry Evidence on the Effects of Income, Prices, and Population Growth. *Economics of Education Review, 7*(2), 167-183.

Schultz, T.W. (Ed.) (1961). *Investment in Human Beings*. Chicago, Il: University of Chicago Press.

Shelton, C.A. (2007). The Size and Composition of Government Expenditure. *Journal of Public Economics, 91*(11-12), 2230-2260.

Speciale, B. (2012). Does Immigration Affect Public Education Expenditures? Quasi-Experimental Evidence. *Journal of Public Economics*, *96*(9-10), 773-783.

Stasavage, D. (2005). Democracy and Education Spending in Africa. *American Journal of Political Science*, *49*(2), 343-358.

Tanzi, V. & Schuknecht, L. (2000). *Public Spending in the 20th Century: A Global Perspective*. Cambridge: Cambridge University Press.

Tilak, J.B.G. (1984). Political Economy of Investment in Education in South Asia. *International Journal of Educational Development*, *4*(2), 155-166.

Tilak, J.B.G. (1989). The Recession and Public Investment in Education in Latin America. *Journal of Interamerican Studies and World Affairs*, *31*(1-2), 125-146.

Tilak, J.B.G. (1990). External Debt and Public Investment in Education in Sub-Saharan Africa. *Journal of Education Finance*, *15*(4), 470-486.

Verner, J.G. (1979). Socioeconomic Environment, Political System, and Educational Policy Outcomes: A Comparative Analysis of 102 Countries. *Comparative Politics, 11*(2), 165-187.

Wagner, A. (1883a). *Lehr und Handbuch der Politischen Ökonomie, Vierte Hauptabeilung: Finanzwissenschaft*. 3rd edn. Leipzig: Winter.

Wagner, A. (1883b, 1958). The Nature of the Fiscal Economy. In R.A. Musgrave & A.R. Peacock (Eds.), *Classics in the Theory of Public Finance* (pp. 1-8). London: Macmillan.

Wagner, A. (1892). *Grundlegung der Politischen Ökonomie*. 3rd edn. Leipzig: Winter.

Wagner, A. (1911). Staat in Nationalökonomischer Hinsicht. *Handwörterbuch der Staatswissenschaften* (pp. 743-745). 3rd edn, Book VII. Jena: Lexis.

Wicksell, K. (1896). *Finanztheoretische Untersuchungen nebst Darstellung und Kritik des Steuerwesens Schwedens*. Jena: Fischer.

Wilensky, H. (1981). Leftism, Catholicism, and Democratic Corporatism: The Role of Political Parties in Recent Welfare State Development. In P. Flora & A.J. Heidenheimer (Eds.), *The Development of Welfare States in Europe and America* (pp. 345-382). New Brunswick: Transaction.

Williams, R. (2015). *Interaction Effects and Group Comparisons*. Working Paper, University of Notre Dame. Retrieved from https://www3.nd.edu/~rwilliam/stats2/l51.pdf

World Bank (2017a). *Education Statistics: Education Expenditure*. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=Education-Statistics:-Education-Expenditure

World Bank (2017b). *TCdata360 database*. Retrieved from http://tcdata360.worldbank.org/

World Bank (2017c). *World Development Indicators*. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators

World Bank (2017d). *World Bank Development Prospects Group: A Cross-Country Database of Fiscal Space*. Retrieved from http://www.worldbank.org/en/research/brief/fiscal-space

Zymelman, M. (1976). *Patterns of Educational Expenditures*. World Bank Staff Working Paper No. 246, International Bank for Reconstruction and Development, World Bank, Washington, D.C. Retrieved from http://documents.worldbank.org/curated/en/899691468766202748/Patterns-of-educational-expenditures

APPENDICES

Appendix A: List of Countries by GNI per capita Group in 2015 (ypc201521)

Low Income (31)	Lower Middle Income (52)	Upper Middle Income (56)	High Income (non-OECD) (57)	High Income (OECD) (21)
Afghanistan	Armenia	Albania	Andorra	Australia
Benin	Bangladesh	Algeria	Antigua and Barbuda	Austria
Burkina Faso	Bhutan	American Samoa	Aruba	Belgium
Burundi	Bolivia	Angola	Bahamas, The	Canada
Central African Republic	Cabo Verde	Argentina	Bahrain	Denmark
Chad	Cambodia	Azerbaijan	Barbados	Finland
Comoros	Cameroon	Belarus	Bermuda	France
Congo, Dem. Rep.	Congo, Rep.	Belize	British Virgin Islands	Germany
Eritrea	Cote d'Ivoire	Bosnia and Herzegovina	Brunei Darussalam	Greece
Ethiopia	Djibouti	Botswana	Cayman Islands	Ireland
Gambia, The	Egypt, Arab Rep.	Brazil	Channel Islands	Italy
Guinea	El Salvador	Bulgaria	Chile	Japan
Guinea-Bissau	Ghana	China	Croatia	Netherlands
Haiti	Guatemala	Colombia	Curacao	New Zealand
Korea, Dem. People's Rep.	Honduras	Costa Rica	Cyprus	Norway
Liberia	India	Cuba	Czech Republic	Portugal
Madagascar	Indonesia	Dominica	Estonia	Spain
Malawi	Kenya	Dominican Republic	Faroe Islands	Sweden
Mali	Kiribati	Ecuador	French Polynesia	Switzerland
Mozambique	Kosovo	Equatorial Guinea	Gibraltar	United Kingdom
Nepal	Kyrgyz Republic	Fiji	Greenland	United States
Niger	Lao PDR	Gabon	Guam	
Rwanda	Lesotho	Georgia	Hong Kong SAR, China	
Senegal	Mauritania	Grenada	Hungary	
Sierra Leone	Micronesia, Fed. Sts.	Guyana	Iceland	
Somalia	Moldova	Iran, Islamic Rep.	Isle of Man	
South Sudan	Mongolia	Iraq	Israel	
Tanzania	Morocco	Jamaica	Korea, Rep.	
Togo	Myanmar	Jordan	Kuwait	
Uganda	Nicaragua	Kazakhstan	Latvia	
Zimbabwe	Nigeria	Lebanon	Liechtenstein	
	Pakistan	Libya	Lithuania	
	Papua New Guinea	Macedonia, FYR	Luxembourg	
	Philippines	Malaysia	Macao SAR, China	
	Samoa	Maldives	Malta	
	Sao Tome and Principe	Marshall Islands	Monaco	
	Solomon Islands	Mauritius	Nauru	
	Sri Lanka	Mexico	New Caledonia	
	Sudan	Montenegro	Northern Mariana Islands	
	Swaziland	Namibia	Oman	
	Syrian Arab Republic	Palau	Poland	
	Tajikistan	Panama	Puerto Rico	
	Timor-Leste	Paraguay	Qatar	
	Tonga	Peru	San Marino	
	Tunisia	Romania	Saudi Arabia	
	Ukraine	Russian Federation	Seychelles	
	Uzbekistan	Serbia	Singapore	
	Vanuatu	South Africa	Sint Maarten (Dutch part)	
	Vietnam	St. Lucia	Slovak Republic	
	West Bank and Gaza	St. Vincent and the Grenadines	Slovenia	
	Yemen, Rep.	Suriname	St. Kitts and Nevis	
	Zambia	Thailand	St. Martin (French part)	
		Turkey	Trinidad and Tobago	
		Turkmenistan	Turks and Caicos Islands	
		Tuvalu	United Arab Emirates	
		Venezuela, RB	Uruguay	
			Virgin Islands (U.S.)	

Source: Adapted from the World Bank's historical classification (see the notes and source below).
Notes: These groups are adapted from the World Bank's Country and Lending Groups for the 2015 calendar year using the World Bank Atlas Method, except for the high-income (OECD) group of countries, which includes the 21 countries comprising the 'core' OECD nations (excludes Chile, Czech Republic, Estonia, Hungary, Iceland, Israel, Korea, Latvia, Luxembourg, Mexico, Poland, Slovak Republic, Slovenia & Turkey, which would comprise the broader 35 OECD countries). The numbers in parentheses show the total number of countries in each group. The historical classification of the 217 countries is available from:
https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

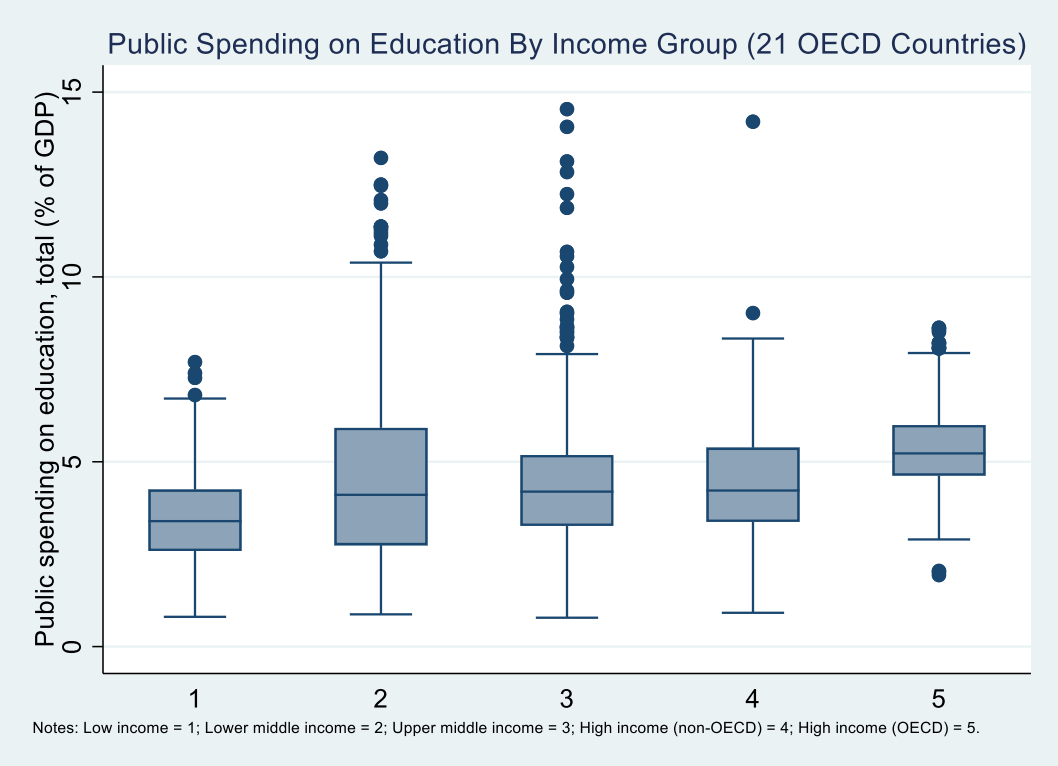
Appendix B: List of Countries by 2 Regional Country Groups (region3)

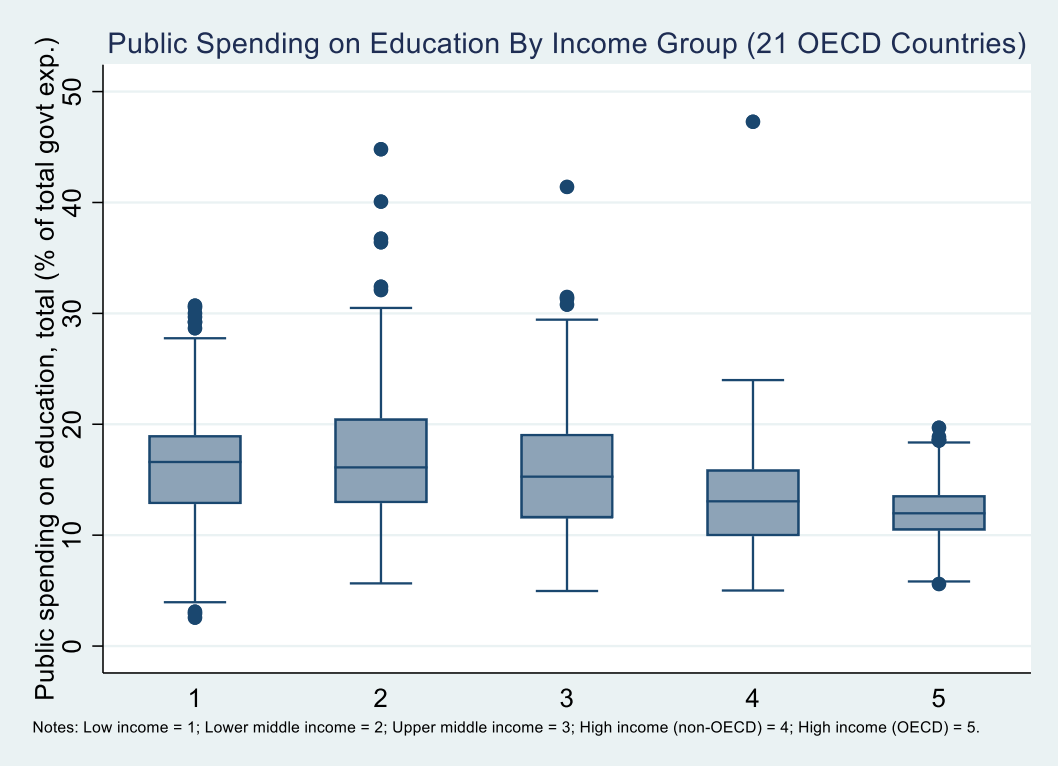
Poorer Country Regions				
Central Africa (8)	**Central America (8)**	**East Africa (12)**	**South America (12)**	**South Asia (8)**
Cameroon	Belize	Burundi	Argentina	Afghanistan
Central African Republic	Costa Rica	Comoros	Bolivia	Bangladesh
Chad	El Salvador	Djibouti	Brazil	Bhutan
Congo, Dem. Rep.	Guatemala	Eritrea	Chile	India
Congo, Rep.	Honduras	Ethiopia	Colombia	Maldives
Equatorial Guinea	Mexico	Kenya	Ecuador	Nepal
Gabon	Nicaragua	Rwanda	Guyana	Pakistan
Sao Tome and Principe	Panama	Somalia	Paraguay	Sri Lanka
		South Sudan	Peru	
		Sudan	Suriname	
		Tanzania	Uruguay	
		Uganda	Venezuela, RB	
Southeast Asia (11)	**Southern Africa (13)**	**West Africa (16)**		
Brunei Darussalam	Angola	Benin		
Cambodia	Botswana	Burkina Faso		
Indonesia	Lesotho	Cabo Verde		
Lao PDR	Madagascar	Cote d'Ivoire		
Malaysia	Malawi	Gambia, The		
Myanmar	Mauritius	Ghana		
Philippines	Mozambique	Guinea		
Singapore	Namibia	Guinea-Bissau		
Thailand	Seychelles	Liberia		
Timor-Leste	South Africa	Mali		
Vietnam	Swaziland	Mauritania		
	Zambia	Niger		
	Zimbabwe	Nigeria		
		Senegal		
		Sierra Leone		
		Togo		
Richer Country Regions				
North America (3)	Nordic Countries **(5)**	**Western Europe (22)**		
Bermuda	Denmark	Andorra		
Canada	Finland	Austria		
United States	Iceland	Belgium		
	Norway	Channel Islands		
	Sweden	Faroe Islands		
		France		
		Germany		
		Gibraltar		
		Greece		
		Greenland		
		Ireland		
		Isle of Man		
		Italy		
		Liechtenstein		
		Luxembourg		
		Monaco		
		Netherlands		
		Portugal		
		San Marino		
		Spain		
		Switzerland		
		United Kingdom		

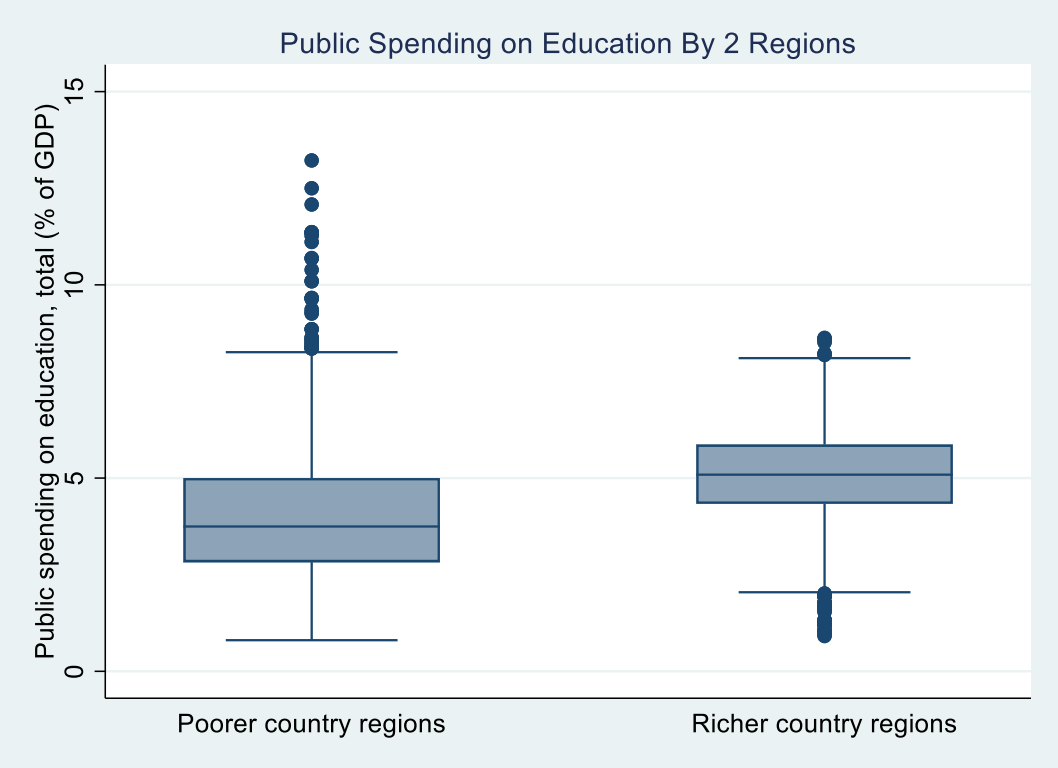
Source: Author's compilation.

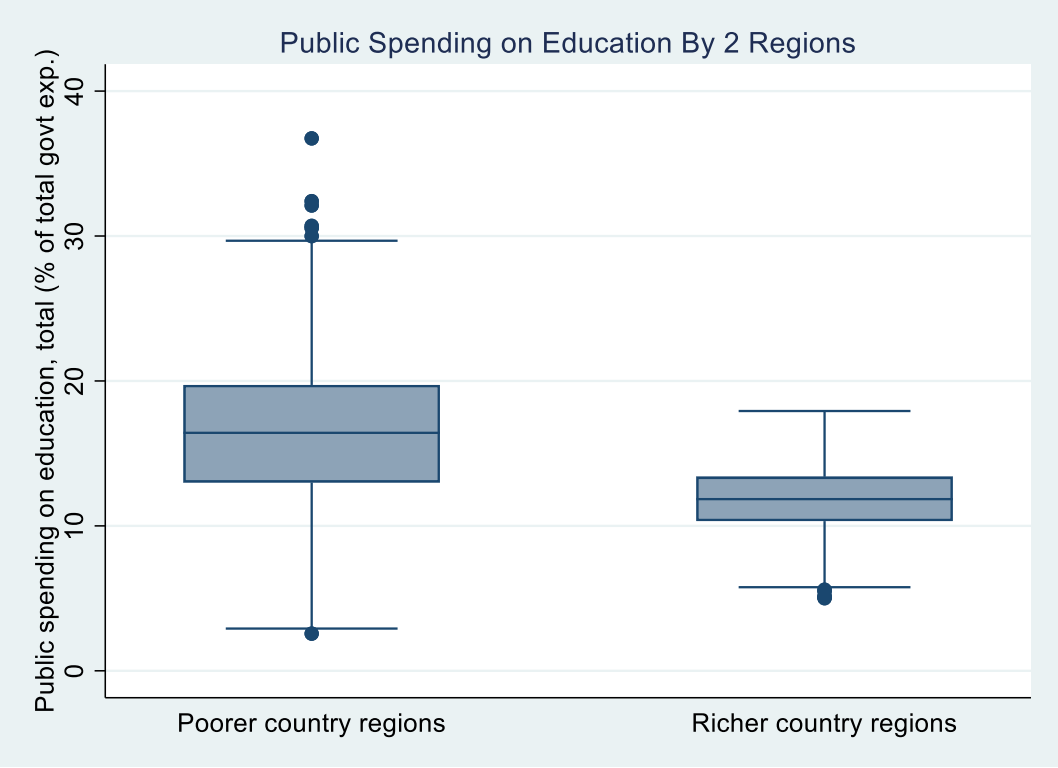
Notes: The numbers in parentheses show the total number of countries in each sub-group of the respective country region. For the poorer country regions, Equatorial Guinea, Chile, Uruguay, Brunei Darussalam, Singapore and Seychelles (the highlighted countries) are excluded for various reasons.

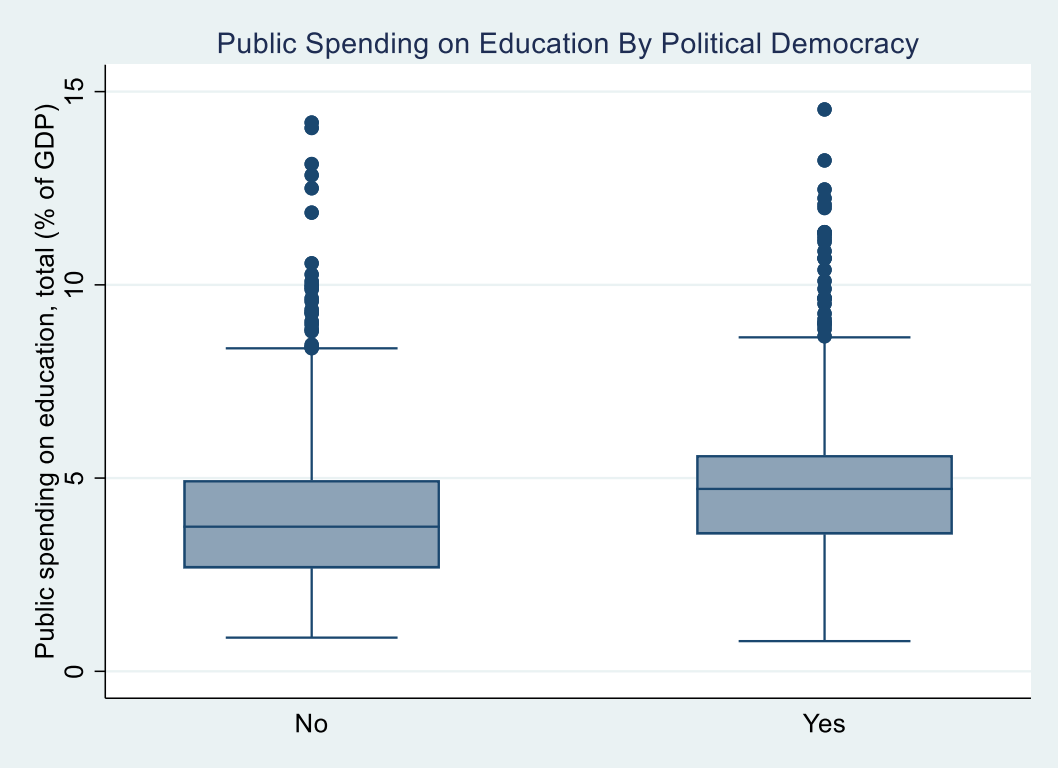
Appendix C: Box Plots of Education Spending by the Various Categorical Measures

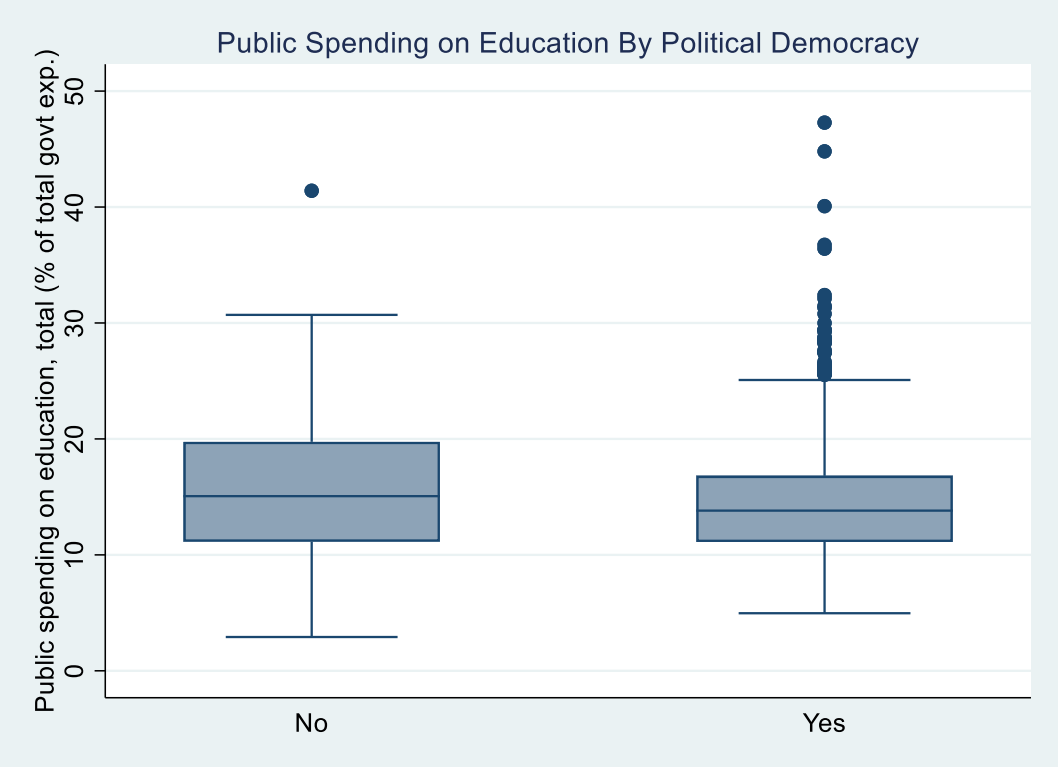





Public Spending on Education By 2 Regions
Public spending on education, total (% of GDP)
0
5
10
15
Poorer country regions
Richer country regions


Public Spending on Education By 2 Regions
Public spending on education, total (% of total govt exp.)
0
10
20
30
40
Poorer country regions
Richer country regions


Public Spending on Education By Political Democracy
Public spending on education, total (% of GDP)
0
5
10
15
No
Yes


Public Spending on Education By Political Democracy
Public spending on education, total (% of total govt exp.)
0
10
20
30
40
50
No
Yes

Appendix D: A Summary of Various Changes to the Model Specification (Using the National Effort Measure and LSDV Estimator)

Dependent Variable: psegdptot	**Model 1A**	**Model 1B**	**Model 2A**	**Model2B**	**Model 3A**	**Model 3B**	**Model 4A**	**Model 4B**	**Model 5A**	**Model 5B**	**Model 6A**	**Model 6B**
0#0. Poorer country regions & not politically democratic	-2.031***	-2.183***	-1.764***	-1.866***	-0.991***	-1.072***	-0.568*	-0.692*	-0.800***	-1.146***	-0.906***	-1.163***
	(0.216)	(0.242)	(0.224)	(0.258)	(0.275)	(0.365)	(0.296)	(0.383)	(0.268)	(0.293)	(0.279)	(0.307)
0#1. Poorer country regions & politically democratic	-1.412***	-1.543***	-1.379***	-1.448***	-0.578**	-0.631**	-0.028	-0.102	-0.021	-0.277	-0.042	-0.241
	(0.166)	(0.155)	(0.176)	(0.167)	(0.237)	(0.288)	(0.258)	(0.315)	(0.257)	(0.279)	(0.265)	(0.288)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	2.095***	1.504***	-1.863**	-2.243***	-5.403***	-5.626***	-7.503***	-7.711***	-6.685***	-6.649***	-6.539***	-6.743***
	(0.375)	(0.518)	(0.760)	(0.790)	(1.427)	(1.549)	(1.550)	(1.713)	(1.241)	(1.295)	(1.297)	(1.370)
R-squared	0.167	0.187	0.201	0.223	0.208	0.229	0.223	0.245	0.353	0.375	0.365	0.384
F-value	63.60***	15.33***	103.09***	23.20***	91.08***	22.81***	80.33***	21.89***	69.88***	21.30***	60.49***	19.72***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AIC	5782	5798	5190	5202	5181	5194	4858	4872	3866	3876	3581	3598
BIC	5813	5960	5225	5361	5222	5358	4904	5040	3916	4046	3635	3771
Countries	97	97	86	86	86	86	85	85	84	84	84	84
Years	26	26	25	25	25	25	25	25	25	25	25	25
Year fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Covariates (controls)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of control variables	3	3	4	4	5	5	6	6	7	7	8	8
Are the patterns confounded?	No	No	No	No	No	No	Partially	Partially	Partially	Partially	Partially	Partially
Observations	1382	1382	1256	1256	1256	1256	1174	1174	1101	1101	1030	1030

Notes: BASE group is richer country regions that are politically democratic. All models use a LSDV estimator and robust standard errors. Models 1A and 1B use homogeneous slopes and three controls (pop024, urban and trade). Models 2A and 2B use homogeneous slopes and four controls (pop024, urban, trade and hci). Models 3A and 3B use homogeneous slopes and five controls (pop024, urban, trade, hci and pop65). Models 4A and 4B use homogeneous slopes and six controls (pop024, urban, trade, hci, pop65 and military). Models 5A and 5B use homogeneous slopes and seven controls (pop024, urban, trade, hci, pop65, military and fiscbal2). Models 6A and 6B use homogeneous slopes and eight controls (pop024, urban, trade, hci, pop65, military, fiscbal2 and debt2). See Table 4.2 for a description of each control variable used. Time (year) fixed effects are used in each alternative model (Model *B). The estimates for the various controls and year fixed effects are excluded to save space. Huber/White heteroskedasticity-robust standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix E: A Summary of Various Changes to the Model Specification (Using the Budget Share Measure and LSDV Estimator)

Dependent Variable: psegovtot	**Model 1A**	**Model 1B**	**Model 2A**	**Model2B**	**Model 3A**	**Model 3B**	**Model 4A**	**Model 4B**	**Model 5A**	**Model 5B**	**Model 6A**	**Model 6B**
0#0. Poorer country regions & not politically democratic	1.492**	0.985	1.296**	0.737	1.358**	0.499	2.207***	1.514**	1.962***	1.269*	1.803***	1.668**
	(0.612)	(0.664)	(0.644)	(0.700)	(0.637)	(0.718)	(0.661)	(0.755)	(0.656)	(0.749)	(0.667)	(0.756)
0#1. Poorer country regions & politically democratic	3.321***	2.799***	2.970***	2.451***	3.035***	2.204***	3.784***	3.186***	3.717***	3.162***	3.561***	3.391***
	(0.458)	(0.499)	(0.481)	(0.530)	(0.518)	(0.596)	(0.548)	(0.628)	(0.539)	(0.621)	(0.552)	(0.630)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	5.333***	3.726***	-3.448**	-4.803***	-3.737	-3.794	-4.396	-4.194	-1.745	-1.109	-1.285	-1.221
	(1.193)	(1.443)	(1.686)	(1.810)	(2.738)	(2.805)	(2.840)	(2.923)	(2.867)	(2.948)	(2.897)	(2.955)
R-squared	0.222	0.237	0.296	0.311	0.296	0.311	0.331	0.343	0.347	0.360	0.415	0.423
F-value	113.21***	20.02***	110.62***	23.58***	100.86***	24.23***	91.63***	23.86***	92.91***	27.10***	96.81***	31.64***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AIC	7456	7481	6612	6636	6614	6637	6133	6162	6070	6097	5567	5599
BIC	7487	7641	6648	6793	6655	6800	6178	6327	6120	6267	5621	5772
Countries	96	96	85	85	85	85	84	84	84	84	84	84
Years	26	26	25	25	25	25	25	25	25	25	25	25
Year fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Covariates (controls)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of control variables	3	3	4	4	5	5	6	6	7	7	8	8
Are the patterns confounded?	No	No	No	No	No	No	No	No	No	No	No	No
Observations	1299	1299	1175	1175	1175	1175	1099	1099	1091	1091	1024	1024

Notes: BASE group is richer country regions that are politically democratic. All models use a LSDV estimator and robust standard errors. See the notes for Appendix D. Huber/White heteroskedasticity-robust standard errors are given in parentheses. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix F: A Summary of Various Changes to the Model Specification (Using the National Effort Measure and Robust Estimator)

Dependent Variable: psegdptot	**Model 1A**	**Model 1B**	**Model 2A**	**Model2B**	**Model 3A**	**Model 3B**	**Model 4A**	**Model 4B**	**Model 5A**	**Model 5B**	**Model 6A**	**Model 6B**
0#0. Poorer country regions & not politically democratic	-1.760***	-2.022***	-1.489***	-1.729***	-1.240***	-1.615***	-0.926***	-1.311***	-0.930***	-1.332***	-1.077***	-1.418***
	(0.199)	(0.206)	(0.199)	(0.208)	(0.260)	(0.275)	(0.274)	(0.291)	(0.279)	(0.297)	(0.274)	(0.297)
0#1. Poorer country regions & politically democratic	-1.178***	-1.403***	-1.111***	-1.310***	-0.844***	-1.188***	-0.455*	-0.785***	-0.424*	-0.734***	-0.522**	-0.797***
	(0.166)	(0.172)	(0.165)	(0.172)	(0.239)	(0.252)	(0.253)	(0.267)	(0.256)	(0.271)	(0.250)	(0.268)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	2.520***	1.764***	0.251	-0.321	-1.013	-0.848	-2.635***	-2.416**	-3.139***	-2.862***	-2.688***	-2.624**
	(0.348)	(0.425)	(0.520)	(0.566)	(0.924)	(0.951)	(0.965)	(0.998)	(1.002)	(1.041)	(0.988)	(1.053)
R-squared	0.313	0.335	0.331	0.351	0.333	0.351	0.338	0.357	0.340	0.363	0.373	0.392
F-value	125.25***	22.67***	103.17***	22.06***	88.81***	21.32***	74.25***	19.81***	62.34***	18.39***	60.71***	18.87***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AIC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
BIC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Countries	97	97	86	86	86	86	85	85	84	84	84	84
Years	26	26	25	25	25	25	25	25	25	25	25	25
Year fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Covariates (controls)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of control variables	3	3	4	4	5	5	6	6	7	7	8	8
Are the patterns confounded?	No	No	No	No	No	No	No	No	No	No	No	No
Observations	1382	1382	1256	1256	1256	1256	1174	1174	1101	1101	1030	1030

Notes: BASE group is richer country regions that are politically democratic. All models use a robust estimator. See the notes for Appendix D. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix G: A Summary of Various Changes to the Model Specification (Using the Budget Share Measure and Robust Estimator)

Dependent Variable: psegovtot	**Model 1A**	**Model 1B**	**Model 2A**	**Model2B**	**Model 3A**	**Model 3B**	**Model 4A**	**Model 4B**	**Model 5A**	**Model 5B**	**Model 6A**	**Model 6B**
0#0. Poorer country regions & not politically democratic	1.108*	0.522	1.000*	0.310	1.122	0.172	1.972**	1.255	1.798**	1.077	1.614**	1.315*
	(0.597)	(0.626)	(0.588)	(0.617)	(0.761)	(0.806)	(0.771)	(0.829)	(0.751)	(0.804)	(0.702)	(0.767)
0#1. Poorer country regions & politically democratic	2.945***	2.335***	2.626***	1.919***	2.760***	1.777**	3.292***	2.617***	3.268***	2.608***	3.037***	2.724***
	(0.491)	(0.517)	(0.479)	(0.503)	(0.697)	(0.737)	(0.709)	(0.757)	(0.690)	(0.734)	(0.642)	(0.694)
1#0. Richer country regions & not politically democratic	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
1#1. Richer country regions & politically democratic BASE	3.705***	1.652	-4.570***	-6.683***	-5.296*	-5.963**	-5.074*	-5.589**	-2.121	-2.119	-1.306	-1.495
	(1.036)	(1.400)	(1.533)	(1.748)	(2.708)	(2.810)	(2.701)	(2.838)	(2.692)	(2.821)	(2.531)	(2.719)
R-squared	0.247	0.264	0.328	0.348	0.328	0.347	0.373	0.386	0.402	0.421	0.483	0.492
F-value	84.82***	15.18***	94.92***	20.35***	81.47***	19.61***	81.11***	20.98***	80.84***	23.30***	94.69***	28.19***
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AIC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
BIC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Countries	96	96	85	85	85	85	84	84	84	84	84	84
Years	26	26	25	25	25	25	25	25	25	25	25	25
Year fixed effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Covariates (controls)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of control variables	3	3	4	4	5	5	6	6	7	7	8	8
Are the patterns confounded?	No	No	No	No	No	No	No	No	No	No	No	No
Observations	1299	1299	1175	1175	1175	1175	1099	1099	1091	1091	1024	1024

Notes: BASE group is richer country regions that are politically democratic. All models use a robust estimator. See the notes for Appendix D. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.