What do we really know about the determinants of public spending on education?

A robustness check of three empirical models

Lisa Spantig

August, 2013

Master's Thesis in Economics, Lund University

Supervisor: Kaveh Majlesi

Abstract

A variety of theories concerning the determinants of public spending on education exist, but an encompassing one has yet to be formulated. Frequently examined determinants are regime type and globalization, but the empirical results for these variables are not consistent across studies. This paper contributes to the literature by anecdotally demonstrating that insufficient theory can lead to a lack of empirical robustness. Thereto, three different empirical models are replicated and altered. It is shown that changes in the sample, the inclusion of an additional variable or a different measurement method can lead to differing estimates. This instability is sometimes exacerbated by statistical shortcomings such as autocorrelation. In this study, both the results for regime type, measured by democracy, and globalization, captured by log trade openness, vary across samples and models, showing positive or insignificant effects. Regarding log trade openness, cross-country and within-country effects seem to differ systematically, which is contrary to previous findings and inexplicable by current theory. Given the lack of robustness, interpretation of results has to be careful and a consistent theory is needed as guidance for empirical analysis with external validity.

Keywords: Replication, Panel Data Analysis, Robustness, Public Spending on Education, Trade Openness.

Contents

1 Introduction **1**

2 Theoretical and empirical context **3**

2.1 The supply side of government spending 3

2.1.1 Theoretical framework 3

2.1.2 Empirical findings 4

2.2 The demand side of government spending 6

2.2.1 Theoretical framework 6

2.2.2 Empirical findings 8

2.3 Remarks on previous research 8

3 Empirical analysis **10**

3.1 Strategy 10

3.2 Data and variables 11

3.3 Empirical results 12

3.3.1 Cross-sectional analysis of a global sample 12

3.3.2 Time-series cross-sectional analysis of an African sample 16

3.3.3 Time-series cross-sectional analysis of a global sample 21

3.4 Discussion 25

3.4.1 Data and measurement issues 26

3.4.2 Econometric issues 27

3.4.3 Interpretation 29

4 Conclusion **33**

5 References **35**

Appendices **39**

A Additional output 39

B Data and variables 42

C Literature overview 47

Education is the most powerful weapon which you can use to change the world.
Nelson Mandela

1 Introduction

Education is a universal human right and has also been shown to have numerous positive socio-economic effects, of which economic growth is a very prominent example (Lucas, 1988; Mankiw et al., 1992; Barro, 2001). This importance has led to a lively debate on education and education policy. In case of formal public education, governments ultimately determine the supply, e.g. by school construction policies or, more generally, by spending on education. So, what induces governments to target their expenditures on education? Why do some governments spend more on education than others? What are the determinants of public education spending?

Research on these questions has explored several explanations; the most popular ones being regime type and globalization. More democratic countries are often assumed to spend more on education due to higher pressures from the constituency (Brown and Hunter, 2004; Avelino et al., 2005; Stasavage, 2005; Ansell, 2008) and more open countries are often said to invest in education to remain competitive (Kaufman and Segura-Ubiergo, 2001; Rudra and Haggard, 2001; Avelino et al., 2005; Ansell, 2008). Although a variety of theories coexists, a consistent and encompassing one is still lacking.

In addition, empirical results are inconclusive, as both the effect of regime type and globalization are not consistent across studies. This may be due to the fact that only partial effects are investigated, with an often arbitrary choice of controls. The lack of consistent theory, that ideally guides empirical analysis, might thus be one reason for the variability of empirical results.

In contrast to the few authors (e.g. Rudra and Haggard, 2001; Plümper and Martin, 2003) who briefly take note of the problem outlined above, this paper clearly demonstrates and addresses some of the consequences of lacking theory. This is done by investigating the robustness of three different empirical models for government spending on education (Mulligan et al., 2004; Stasavage, 2005; Ansell, 2008).

First, it is shown that close replication without the original dataset is extremely difficult and that the empirical results are not robust to an alteration of the original samples.

Second, the robustness of the replicated results is further examined by including an additional variable (for Mulligan et al., 2004 and Stasavage, 2005) and by using a different measurement method for a variable already included (for Ansell, 2008). For all three models, this variable is log trade openness, a commonly employed measure that captures an important

effect of globalization. As suggested by Avelino et al. (2005) and in contrast to Ansell (2008), the measure of trade openness is based on purchasing power parities (PPP).

Regarding log trade openness, a robust positive and significant effect is found in different cross-country analyses, but the within-country effects turn out to be insignificant. These results cannot be explained by current theory and the statistical explanations offered here are only tentative. Concerning the effect of regime type on education spending, the results of this paper are also inconclusive, varying by definition of the variable and specification of the model. Moreover, the statistical robustness checks show that many models might also be questioned in terms of econometric validity.

The discussion section points out several problems of the empirical analysis that arise due to the lack of theory, ranging from the choice and measurement of variables to econometric issues and interpretation problems. As robustness and thus external validity are often questionable, inference is likely to be unreliable.

Since a consistent theory cannot be offered here, it is not possible to reveal the true determinants of public education spending. However, it is shown that exactly this is also the problem of some published and peer-reviewed articles. Thereby, this paper reinforces the call for increased data sharing (e.g. Eichengreen, 2013), more general equilibrium theory (e.g. Acemoglu, 2010) and a careful interpretation of empirical results (e.g. Rodrik, 2013).

The remainder of the paper is organized as follows: first, the relevant literature is reviewed with respect to both theoretical and empirical findings and the main problems are briefly addressed. Second, the empirical strategy is outlined, the data are described and the three models are replicated, extended and subjected to statistical robustness checks. In the last section of the empirical part, the results as well as their implications are discussed, and the last section concludes.

2 Theoretical and empirical context

A variety of theories concerning the determinants of government spending exists. Here, government spending is dealt with in general, as education can be seen as a subcategory of public goods. This has the advantage of providing a more complete picture because the general literature on government spending is more extensive. Where necessary, particularities of public education spending are explained. Two mutually non-exclusive ways of categorizing this research have been proposed:

1. Two different schools of determinants of public spending are identified by Mulligan et al. (2004); the first concentrating on the role of the voting mechanism and the second emphasizing the importance of economic and demographic variables.
2. A classical demand and supply categorization of influences on government spending has been proposed by Hausken et al. (2004).

Both views can actually be combined by arguing that the voting literature investigates how the regime type constraints the way governments maximize their support, which in turn determines the supply side of government spending (Lake and Baum, 2001). Economic and demographic variables, on the other hand, describe the society and the international framework in which the country is embedded. This determines the demand for public policies.

The supply and demand framework clearly offers the advantage of general equilibrium analysis once a coherent theory will have been formulated. In what follows, the relevant literature is systematically presented in this supply and demand framework.

2.1 The supply side of government spending

2.1.1 Theoretical framework

Theories that (partially) explain "the conditions under which governments are willing to satisfy the constituents' demands" (Hausken et al., 2004: 241) usually start with the assumption that governments are utility or support maximizing (e.g. Lott, 1999; Lake and Baum, 2001). Rational and opportunistic political leaders are assumed to face trade-offs and have to mitigate distributional conflicts, but their main interest is to stay in office (Plümper and Martin, 2003; Stasavage, 2005). Usually, governments are either modeled as monopolistic firms providing public goods (e.g. Lake and Baum, 2001) or they are assumed to choose an optimal

combination of rents and public goods (e.g. Plümper and Martin, 2003). Although heavily simplifying, these models are sufficient to trace the effects of different regime types.[1]

The type of the political system can be interpreted as constraints on how politicians pursue and reach their goal of support maximization (Lake and Baum, 2001): democracies are characterized by lower costs of political participation, lower barriers to exit, and thus higher political competition. In addition, a more competitive political system is argued to produce better institutions, such as rule of law, which increases accountability (Stasavage, 2005). Accountability in combination with high political participation leads to a removal of the incumbent if a certain limit of rent extraction is exceeded (Baum and Lake, 2003). This credible threat works as an effective constraint on rent-seeking. On the contrary, the costs of political participation are higher and accountability is lower in non-democracies. Therefore, a higher level of rent extraction is tolerated (ibid.). Elections, as the means to remove politicians, can thus be interpreted as an constraint that lowers rents and thus increases provision and spending on public services (Lake and Baum, 2001.) More specifically, Bueno de Mesquita et al. (2003: 289-292) argue that the coalition feature of democracy helps explaining public expenditures: in their selectorate theory, the size of the winning coalition is positively linked to higher expenses, as a larger constituency that has to be satisfied.

This can be seen as a rough and general supply side explanation[2] of government spending (on education), that is supported by the majority of the literature[3] and which leads to the first hypothesis:

Hypothesis 1: Democracies spend more on education than non-democracies.

2.1.2 Empirical findings

A positive effect of democracy is found by a variety of articles that use various variables measuring overall government spending, the size of government or welfare effort (e.g. Hicks and Swank, 1992; Rudra and Haggard, 2001; Adsera and Boix, 2002; Rudra, 2002; Avelino et al. 2005). Table C.1 in the appendix provides a detailed overview of the empirical models that

[1] In what follows, regime type refers to the political organization of the state. In accordance with the majority of the literature, only a rough distinction between democracies and non-democracies is drawn with the voting process being the distinctive feature of a democracy. It should be kept in mind that the real distinction is much more subtly nuanced.

[2] While regime type is the most investigated supply side effect, some articles propose other or additional effects. One example is corruption, that is investigated by Mauro (1998) and Gupta et al. (2002) and that can be interpreted as a special case of rent-seeking. Higher corruption therefore is expected and shown to result in lower spending on public goods or a shift of government spending towards more targetable goods.

[3] Alternative explanations exist but they often concerns a subcategory of non-democracies and do not directly address the differences between democracies and non-democracies (e.g. Lott, 1999). Yet, they can be seen as an argument why Hypothesis 1 will not necessarily hold in empirical analysis. Therefore, they are presented in the empirical section.

are briefly presented in the following.[4] This table also contains the empirical demand side models described in section 2.2.2.

Plümper and Martin (2003) show that the level of democracy and spending is correlated in a U-shaped fashion: governments in non-democracies tend to mainly spend on rents in order to gain support. With a rising level of democracy, this practice becomes too expensive and expenditures are shifted towards public goods. With further increasing political participation, governments have the incentive to spend even more which leads to overinvestment in public goods and less efficient spending. This relationship is confirmed by Hausken et al. (2004) in a follow-up piece.

Kaufman and Segura-Ubiergo (2001) argue that the effect of regime type depends on the category of spending. While they find no difference in aggregate government spending between democracies and non-democracies in Latin America, more democratic governments in this region seem to spend more on education.

These results are confirmed by a variety of studies with different samples that examine public education spending in particular. Bueno de Mesquita et al. (2003) confirm the theoretical derivation from the above framework: if accountability is high and leaders can be easily exchanged, governments have greater incentives to spend on education than on rents. In a global sample, they show that the sizes of the winning coalition and the selectorate induce higher spending on education (ibid: 289-292). Brown and Hunter (2004) examine Latin America and confirm that democracies spend more on primary education, because the government caters to a larger constituency. Avelino et al. (2005) confirm the strong positive impact of democracy on education spending in the same region and show how demand side effects reinforce this finding.

Stasavage (2005) examines education spending in Africa and finds a positive effect of democracy that mainly operates through increased spending on primary education. He explains this finding with the fact that democracy also lends a voice to the rural population that otherwise faces high costs of collective action. Ansell (2008) arrives at the same conclusion for a global sample and provides an additional rationale why non-democracies may spend less: once the optimal level of education is reached, further investment in education leads to unemployment or a reduction in skilled labor wages. This, in turn, can lead to social unrest or even revolutions (Campante and Chor, 2012).

[4] Note that not all articles uniquely focus on the empirical model discussed below. Several papers also propose models for other social spending categories (e.g. Mulligan et al., 2004) or for a particular education sector (e.g. primary education as in Stasavage, 2005).

While the majority of the literature provides theoretical arguments and empirical results showing that democracies do spend more on public goods and education in particular, some counter arguments have been made: Lott (1999: S130) suggests and demonstrates that totalitarian regimes, a subcategory of non-democracies, use education as a means to indoctrinate the citizens since "higher levels of totalitarianism produce diminishing returns to controlling the citizenry through force and increase the marginal return to indoctrination". The same holds true for rising levels of opposition. Therefore, high levels of opposition and high levels of totalitarianism induce increased spending on education. Brown and Hunter (2004) argue that autocracies might want to increase education for developmental reasons or to maintain legitimacy. Using Cuba as an example, they claim that if non-democracies have an interest in investing in education, they might even be able to mobilize more resources than democracies.

Lastly, Mulligan et al. (2004) find no significant difference between democracies and non-democracies in terms of spending on public education. Although this is contrary to their initial hypothesis, they explain this finding by arguing that policy decisions are trade-offs that "are basic to human nature and not specific to particular political institutions" (ibid: 72).

In sum, different theoretical and empirical findings exist. With the majority hinting at a positive relationship between democracy and public spending on education, Hypothesis 1 can be seen as partially confirmed.

2.2 The demand side of government spending

2.2.1 Theoretical framework

Demand side determinants of government spending include characteristics of the society as well as the country's stance in the world community. In contrast to the supply side literature, research on the demand side seems to be broader but less integrated, ranging from effects of globalization, wealth and heterogeneity to individual and household characteristics. While all approaches are of interest, the most debated issue of demand side determinants in cross-country analyses is the effect of globalization. Thus, and in order to keep the overview manageable, this literature overview is restricted to globalization.[5]

[5] Future research should incorporate the most important supply- and demand-side effects into a more parsimonious model. Yet, extending the rough framework and testing the resulting hypothesis is beyond the scope of this paper. Especially individual and household characteristics play an important role not only in forming the demand for, but also the actual usage and outcome of education. But cross-country comparisons are inherently difficult and usually studies focus on one or a few countries in the same region (e.g. Mason and Rozelle, 1998).
In addition, ethnic diversity of the population is an interesting part of country characteristics that influence the demand side. Apart from differing tastes, it can lead to increased rent seeking and therefore also negatively affects public spending (Shleifer and Vishny, 1993 and Easterly and Levine, 1997). Social inclusion, in contrast, appears to have a mitigating effect (Gradstein, 2003).

Two opposing views exist on the trade-off between efficiency and welfare that comes along with globalization; the efficiency and the compensation hypothesis that are explained in detail by Garrett (2001). The efficiency hypothesis claims that governments see themselves forced to cut spending to remain competitive in the world market. Increased levels of globalization are therefore expected to decrease public spending. The compensation hypothesis, in contrast, focuses on political incentives to expand the public sector. It assumes that globalization creates an environment of uncertainty, social dislocations and unequal distributional effects for the individual, in which governments help compensate for the increased risk, e.g. by paying higher unemployment benefits (Kaufman and Segura-Ubiergo, 2001). Therefore, more globalization is expected to increase public spending. In sum, the effect of globalization on overall government spending depends on whether the efficiency or the compensation theory prevails.

Concerning education, governments' expenditure decisions might differ, as it is a special category of government spending (e.g. Avelino et al., 2005; Ansell, 2008). Spending on education can improve labor supply and enhance productivity, in which case increased expenditures leads to greater efficiency. These positive effects are likely to be larger for less developed countries that still have lower productivity, which leaves higher potential for improvement (Ansell, 2008). With respect to education spending, the efficiency and the compensation hypothesis are thus not mutually exclusive (Avelino et al., 2005).[6]

Furthermore, demand for education is subject to labor market effects that determine the returns to education. These effects depend on the degree of a country's openness since in an open economy, skilled labor can move abroad, whereas in closed countries, an increase in education can lead to reductions in relative wages for skilled labor (ibid.). Thus, in closed economies, the elite will prefer private education, profiting from higher relative wages since private education is not subject to externalities and leaves the rest of the population unskilled. In more open (globalized) economies, public spending on education is therefore expected to be higher.

In contrast to overall government spending, where the effects of globalization are ambiguous, public education spending is expected to increase with higher levels of globalization, which leads to the second hypothesis:

Hypothesis 2: Higher levels of globalization increase government spending on education.

[6] Note that this implies a steady relationship of globalization and spending on education, that is in contrast to the U-shaped relationship of democracy and general government spending found in some studies. This difference highlights the already mentioned fact that the effects explaining education spending might differ significantly from the effects that influence overall public spending.

2.2.2 Empirical findings

For both general government spending and education expenditures, the results are mixed. In empirical analyses of the effects of globalization on general government spending, a strong focus on OECD countries exists.[7] In this particular sample, the evidence is mixed, but it rather favors the compensation than the efficiency hypothesis (Garrett, 2001). In a broader sample, Rodrik (1998) and Adsera and Boix (2002) show a positive relationship of globalization and government expenditures, whereas Quinn (1997) finds the opposite.

In a sample of less developed countries, Rudra (2002) finds a negative effect which she attributes to a decline in labor's power to demand compensation. This result is consistent with Kaufman and Segura-Ubiergo (2001), who examine Latin America and find a negative effect, also emphasizing the fact that factories have greater exit options than workers. In the same region, Avelino et al. (2005) find a positive effect of globalization. Other authors find no effect of globalization (Iversen and Cusack, 2000; Iversen, 2001; Rudra and Haggard, 2001).

Focusing on education spending, Kaufman and Segura-Ubiergo (2001) do not find any significant effect of globalization in Latin America. In contrast, Avelino et al. (2005) find a positive effect of globalization on education spending in the same region. In a broader sample of developing countries, Rudra and Haggard (2001) show a negative relationship of globalization and public education spending which they relate to the influence of non-democracies on public spending. According to Avelino et al. (2005), this negative effect could also be due to non-governmental influences such as labor unions or left-wing parties that are often less present in developing countries than in OECD nations. Structural deficiencies such as weak tax collection systems might also influence the relationship between globalization and public spending (ibid.). In a global sample of 113 countries, Ansell (2008) finds a positive effect of globalization.

In sum, the effect of globalization on overall public spending varies with the sample and the same holds true for education spending. Previous research is thus inconclusive concerning Hypothesis 2.

2.3 Remarks on previous research

The issue with the empirical findings presented above has already been noted in the context of general public spending and also holds true for public spending on education: "The main problem of an econometric estimate of government spending arises from the almost complete lack of a theoretically founded baseline model" (Plümper and Martin, 2003: 39). Nonetheless, little has changed during the last ten years. In general, it appears that

[7] This is mainly due to a lack of consistent data in other samples (Garrett, 2001).

reconciliation of previous and differing research is often neglected in favor of contributing "something new", even if this contribution is the investigation of another partial effect, e.g. on the demand side, that might turn out to be very different in general equilibrium analysis or in a different sample and therefore lacks external validity.

The lack of a consistent theory implies that the control variables are chosen rather arbitrarily, which is why the articles differ significantly in this regard. While most articles that focus on demand side effects control for democracy, only a some supply side authors (e.g. Plümper and Martin, 2003; Avelino et al., 2005; Ansell, 2008) consider the effect of globalization when trying to explain education spending or other social spending categories. This has already been criticized by Rudra and Haggard (2001).

Moreover, different definitions and assumptions lead to differing variables. For instance, dependent variables are, inter alia, total public spending on education as a percentage of GDP (e.g. Mulligan et al., 2004), total spending on education as a percentage of total government spending (e.g. Stasavage, 2005) or log education spending per capita (e.g. Brown and Hunter, 2004). Furthermore, some authors focus on levels of the dependent and independent variables while others examine changes in those variables. Globalization is also captured in various ways, e.g. by trade openness (Rudra, 2002), institutional openness to trade (Plümper and Martin, 2003), capital mobility (Kaufman and Segura-Ubiergo, 2001) or financial openness (Avelino et al., 2005). Table C.1 clearly shows the variability of variables and measurement methods.

It is therefore not surprising that previous research has produced results that differ significantly across samples and models. This paper aims at enriching the above presented literature by making the robustness of empirical models its central theme. Instead of briefly mentioning the problem, some empirical problems caused by the lack of theory are examined, clearly highlighted and straightforwardly discussed. This will hopefully encourage future research to make an attempt to build an encompassing model as a theoretical guidance to empirical estimation.

3 Empirical analysis

The empirical analysis tests the robustness of three different empirical models proposed by Mulligan et al. (2004), Stasavage (2005) and Ansell (2008). These models have been chosen for several reasons with the most important ones being data availability and accessibility. Moreover, all three models use the same dependent variable which was expected to assure some comparability of the results. Lastly, the authors use different data and samples, which allows to assess whether the results are consistent across samples. Mulligan et al. (2004) conduct a cross-sectional analysis of a global sample, whereas Stasavage (2005) examines time-series cross-sectional (TSCS) data of African countries and Ansell (2008) uses a dynamic model for TSCS analysis of a global sample.

This section is organized as follows: first, the strategy is outlined; second, the data and their sources are presented. Third, the models are replicated, extended and subjected to robustness checks. In the last part, the results as well as their implications are discussed.

3.1 Strategy

The three models are examined one by one, but the strategy is the same for all three of them. First, the relevant hypothesis of the respective article is stated, the empirical model is explained and the most important findings are briefly presented.

Second, the original model is replicated as closely as possible. For comparability reasons, the data used in the present paper correspond as precisely as possible to the data of the replicated models; the same applies to the estimation technique. Replication results and problems are briefly discussed and it is shown that minor alterations, e.g. due to data availability, may yield rather different results.

Third, the replicated models are tentatively tested for omitted variables by including an additional variable. As the literature shows some support for adding globalization to the empirical analysis of public education spending, the replications of Mulligan et al. (2004) and Stasavage (2005), that do not include such a measure, are extended by one additional regressor. One of the most frequently used measures of globalization is trade openness, the sum of exports and imports over GDP.[8] As explained below and as suggested by Avelino et al. (2005), the trade openness measure used in this paper has the denominator adjusted for PPP. This implies that Ansell's (2008) model is altered insofar as the measurement of the trade openness variable differs from replicated model, which employs an unadjusted measure.

[8] Different measures of financial openness, e.g. private capital flows, are also commonly employed. Yet, in this analysis, one additional variable suffices to test the robustness of other models. As theory is lacking, the choice is rather arbitrary anyways.

Lastly, both the replicated and the extended models are subjected to specific statistical robustness checks, mainly focusing on the fulfillment of the ordinary least squares (OLS) assumptions. General concerns that are common to all models, such as omitted variables and endogeneity, are addressed in the subsequent discussion section.

3.2 Data and variables

The dependent variable is in all cases total public spending on education as a percentage of GDP from UNESCO, replicated in the World Bank Development Indicators (WDI). All models include a measure of regime type, a variable capturing the structure of the population as well as a measure of GDP. Other controls vary by model – for the rationale of including them, please refer to the respective articles. The main data sources are the WDI and other frequently employed measures such as the POLITY IV score.[9] All variables, their respective sources and transformations are listed in Tables B.5-B.7 in the appendix; Tables B.2-B.4 show the summary statistics. It has to be noted that data availability and quality varies significantly across countries.[10]

The relationship of democracy and education spending is investigated more often and a measure of democracy is already included in all three replicated models. Therefore, the variable of main interest is trade openness that is less frequently examined and has yielded mixed results in previous research. Here, trade openness is adjusted for PPP. In contrast to the commonly used measure, that uses GDP figures based on real exchange rates (in the denominator), the PPP-based one offers a more precise picture of trade relative to the size of the economy.[11] Since exchange rates are determined on the basis of tradables ignoring the non-tradable sector, using them as conversion factor may distort GDP figures: given arbitrage, export and import prices should have a lower cross-county variance than prices of non-tradable goods. In developing countries, where labor is comparatively cheap and thus non-tradable goods inexpensive, exchange rate based measures of GDP are likely to underestimate the true value of non-tradables in those countries. This, in turn, leads to an overestimation of trade relative to the economy, a shortcoming of the conventional trade openness variable. Therefore, the denominator of the trade openness variable used here is PPP-adjusted.[12] Table B.1 in the

[9] Since the Polity score measures democracy on a scale of -10 to 10 but codes e.g. cases of transition with -88, all observations that score below -10 are omitted. Otherwise, they would result in a negative bias.

[10] Potential problems are briefly discussed in Section 3.4.1. For a more in-depth discussion see e.g. Stasavage (2005). He also shows how data imputation techniques can be used for robustness checks. Unfortunately, this kind of robustness check is beyond the scope of this paper.

[11] For a more detailed explication see Avelino et al. (2005).

[12] Note that the numerator is not subject to such distortions as both imports and exports are subject to the law of one price. Moreover, it would be difficult to adjust these figures for PPP as the conversion factors are usually calculated especially for GDP adjustments.

appendix contrasts the commonly used trade openness measure and the PPP-based one. It can be seen that the latter has a lower within-country variance which Avelino et al. (2005) claim to be more realistic. Finally, the trade openness measure is used in logarithmic form which is consistent with Ansell's (2008) transformation of unadjusted trade openness and the logarithmic form of other variables, e.g. GDP and population, in all replicated models.[13]

Regarding other variables, only Mulligan et al. (2004) provide a complete dataset which makes a perfect replication possible.[14] The replication of the two remaining articles is more complicated since the authors do not provide their original datasets. Moreover, they sometimes lack explicit explication of their variables and sources. In these cases, common measures that arguably fit the idea and estimation of the respective model are used.

3.3 Empirical results

3.3.1 Cross-sectional analysis of a global sample: Mulligan et al. (2004)

The authors hypothesize that under non-democratic regimes, groups are prevented from "express[ing] their intensity of preference for economic and social policy" (p. 54), which would result in a positive and significant coefficient on the democracy variable – consistent with Hypothesis 1.

Mulligan et al. (2004) use 1960-1990 averages of all their independent variables and 1980-1990 averages of the dependent variable. Hence, they have comparatively few observations (114), which provides cross-country regression evidence. Using normal standard errors, they estimate the following model with OLS:

$$eduspend_i = c + \beta_1 democ_i + \gamma X_i + \varepsilon_i \qquad \text{(eq. 1)}$$

where c is a constant, $democ_i$ is the average of the democracy index from the POLITY IV data from 1960-1990, X_i is a vector of averaged control variables and ε_i is the error term. The included control variables are a dummy variable for communist countries and one for British legal origin, the share of elderly in the population, a measure of total population, real GDP per capita as well as the share of value added from agriculture. A measure of trade openness is not included.

3.3.1.1 Replication

Since the authors provide their dataset and as the estimation method is clear, the original results and the replicated ones are the identical (see Table 1, regressions (1)-(4)). Since the

[13] Tables A.1, A.2 and A.3 also show estimations using an untransformed trade openness measure.
[14] Replication without the use of their dataset yields very different results due to data ambiguities. These estimations are not reported here.

overall picture of the variables with respect to sign and significance is more important here than the actual size of the coefficient, the latter is not the main focus in the discussion of the results. Until a robust empirical model is presented, the discussion concerning the size of the estimated effects seems rather unessential.

The coefficient on the democracy score is positive but not significant at any conventional level. Thus, Hypothesis 1 is not confirmed. Public spending on education seems to be higher in communist countries, in smaller states and maybe in countries of British legal origin.[15] In regressions (1) and (3), that do not include the share of value added from agriculture, higher GDP per capita is associated with higher public spending on education. Considering the importance of the share of value added by agriculture, it is negatively and significantly related to education expenditures and renders the coefficient on GDP insignificant (regressions (2) and (4)). Both specifications, (1) and (2), explain approximately the same amount of the variation in public spending on education, namely 25%.

3.3.1.2 Extension

The extended model, also estimated by OLS, is the following:

$$eduspend_i = c + \beta_1 democ_i + \beta_2 lnTradePPP_i + \gamma X_i + \varepsilon_i \qquad \text{(eq. 2)}$$

where the inclusion of $lnTradePPP_i$, logged trade openness based on PPP, is the only modification of equation 1. For consistency reasons, now a PPP-based measure of GDP is used in the controls.[16] The main results of regressions (5) and (6) remain unchanged in comparison to (3) and (4): the coefficient on democracy is still positive, very similar in size, but insignificant in both specifications. Moreover, the coefficient on the communist dummy is slightly reduced in both size and significance and the coefficient on GDP is not significant anymore. More importantly, the size of the population now appears to have a positive effect of about the same size as the original (negative) one. The coefficient on the share of value added from agriculture in regression (6) is no longer significant and its inclusion actually lowers the adjusted R^2-value.

The estimated effect of trade openness on public education spending is significant, positive and comparatively large.[17] This confirms Hypothesis 2. Moreover, the inclusion of this variable increases the R^2-value from 24% to 43%.[18]

[15] The authors relate the significant communist dummy to Lott's (1999) findings that totalitarian regimes tend to use education for indoctrination and therefore spend more on it.
[16] The use of unadjusted GDP does not alter the results significantly (see Table A.1, (3) and (4)).
[17] Note that this is partly due to the transformation: if trade openness is not logarithmized, it remains significant but the coefficient decreases considerably: 0.021 instead of 1.476 (see Table 1, (5) and Table A.1, (5)).

3.3.1.3 Robustness checks

A general drawback of models that use data averaged over time is the loss of dynamic information, in this particular case 30 years. With respect to robustness of the results and the fulfillment of the OLS assumptions, all models show similar characteristics: Ramsey's RESET test (1969) indicates that the linear functional form is correctly specified. Multicollinearity is not a problem either. Also, since there is no time dimension in the data, serial correlation is not an issue. Yet, the Breusch-Pagan LM test (1980) shows heteroskedasticity in the residuals. Since the reason for this heteroskedasticity cannot be modeled here, robust standard errors should be applied.[19] In addition, the Jarque-Bera test (1980) indicates that the errors are not normally distributed in all cases. This does not invalidate OLS, but implies that the estimator is not asymptotically efficient. In sum, the use of OLS can be justified when heteroskedasticity robust standard errors are employed. Concerning outliers, the jackknife test indicates robust results.

While the main conclusion of Mulligan et al. (2004), democracies and non-democracies do not differ significantly in public education spending, is not altered by the inclusion of trade openness, the explanation of cross-country differences in education spending changes: the negative effect of population size has been put into question, as well as the importance of British legal origin, GDP and the share of value added from agriculture. Trade openness, in both specifications, seems to contribute significantly to explaining cross-country public spending patterns. Therefore, it can be argued that the original model suffered from omitted variable bias. Note that this might still be the case for the extend model as well.

[18] When trade openness is not adjusted for PPP, it also appears to be significant in both specifications, but the increase in the adjusted R^2-value is substantially smaller: 33% instead of 43% with the PPP-adjusted measure (see Table A.1 (1) and (2)).

[19] This is not done in the original paper but it does not alter the main findings and trade openness remains significant at 1%.

Table 1: Replication and extension of Mulligan et al. (2004)

	ORIGINAL RESULTS		REPLICATION		EXTENSION	
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy index	0.46	0.42	0.46	0.42	0.478	0.446
	(0.53)	(0.52)	(0.53)	(0.52)	(0.452)	(0.455)
Communist dummy	1.21***	1.09**	1.21***	1.09**	1.018**	0.984**
	(0.45)	(0.45)	(0.46)	(0.45)	(0.399)	(0.409)
British legal origin	0.52*	0.53*	0.52*	0.53*	0.152	0.043
	(0.30)	(0.31)	(0.30)	(0.31)	(0.270)	(0.279)
% population >65	0.04	0.07	0.04	0.07	-0.028	-0.023
	(0.06)	(0.06)	(0.06)	(0.06)	(0.053)	(0.055)
Log(population)/10	-2.34***	-2.28***	-2.34***	-2.28***	2.325**	2.211**
	(0.84)	(0.86)	(0.85)	(0.86)	(0.993)	(1.069)
Log real GDP per capita	0.46**	-0.06	0.46**	-0.06	0.203	0.161
	(0.19)	(0.29)	(0.19)	(0.29)	(0.140)	(0.220)
Share of value added from agriculture		-3.38**		-3.38**		-0.074
		(1.60)		(1.60)		(1.543)
Log trade openness (PPP)					1.476***	1.499***
					(0.236)	(0.265)
N/Countries	114	110	114	110	112	109
Adjusted R^2	0.24	0.25	0.24	0.25	0.43	0.42

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP.
All regressions include a constant (not presented), OLS estimates with normal SE.

3.3.2 Time-series cross-sectional analysis of an African sample: Stasavage (2005)

Consistent with Hypothesis 1, Stasavage (2005) assumes that if politicians need to respond to the majority of citizens in order to stay in office, governments spend more on education. He tests this in a sample of 44 African countries from 1980 to 1996. In his TSCS analysis, Stasavage (2005) uses pooled OLS with panel corrected standard errors (PCSE, eq. 3) and a fixed effects (FE) model that eliminates unobserved time-invariant country specific influences (eq. 4):

$$eduspend_{it} = c + \beta_1 multi_{it} + \gamma X_{it} + \varepsilon_{it} \quad \text{(eq. 3)}$$

$$eduspend_{it} = c + \beta_1 multi_{it} + \gamma X_{it} + \eta_i + \varepsilon_{it} \quad \text{(eq. 4)}$$

where c is a constant, $multi_{it}$ is a dummy variable that takes the value 1 if candidates from more than one party competed for the executive elections, X_{it} is a vector of control variables, η_i are country specific effects and ε_{it} is the error term. The controls include a dummy for election years, GDP per capita, total aid as a percentage of GDP, the share of rural population and the share of the population younger than 15.

The results are presented in Table 2. The coefficient on the multiparty competition dummy is positive and statistically significant at 5% in both his specifications (1) and (2), but the FE estimate is smaller in magnitude. Both estimates confirm Stasavage's (2005) hypothesis as well as Hypothesis 1.[20] Furthermore, higher GDP per capita is also associated with higher public spending on education. The effect (sign and significance level) of the other variables depends on the specification; only the election year dummy does not appear to be significant at all. This hints at the possibility that cross-country and within-country determinants of public education spending might differ or that the estimation techniques pick up different effects.

3.3.2.1 Replication

For both the election year dummy and the GDP per capita measure, the sources are unclear which makes close replication more complicated. Moreover, despite using the same sources for the remaining variables, the replication dataset is smaller which causes the number of observations in the replication to be only half of the original ones. In addition, some data points had to be removed from the dataset since these outliers significantly influenced the results. The number of countries included in the regressions thus decreases from 44 to 29. As this reduces the accuracy of the replication, the time period is tentatively expanded by 10 years: in addition

[20] As a robustness check, he runs the same regression using public spending on education as a percentage of total government spending (instead of % GDP). This also yields positive and significant coefficients on multiparty competition that are even larger in magnitude. In order to be able to compare the results of the replication and extension with the other papers, this additional measure is not used here.

to 1980-1995, the models are also replicated for the period of 1970-1995. The inclusion of more recent data was not possible due to availability of the multiparty variable, but the extended period increases the number of observations by nearly 50%. The additional estimation results for the larger time period are presented in Table A.2 in the appendix.

Regressions (3) and (4) in Table 2 are the replication of the original time period with OLS and FE respectively. As already noted, the number of observations is significantly lower which might affect the results.

In the OLS estimation (3), the coefficient on the multiparty dummy is also positive and significant, but only at 10% instead of 1%. Nonetheless, it confirms Hypothesis 1. GDP per capita is significant as well, but has the opposite sign which is hard to explain on theoretical grounds. All other variables are not significant in the OLS estimation and the R^2-value is noticeably lower in the replication (19% vs. 37%).

The FE replication (4) is even farther from the original results, with a negative and insignificant coefficient on the multiparty dummy. The only significant variable is the share of rural population, but this coefficient has the opposite sign of the one in (1). Moreover, this specification has a very low within R^2-value.[21] Given that the replication mainly differs in the sample examined, one might suspect that Stasavage's (2005) model is not entirely robust although he confirms his findings by enlarging the dataset with multiple data imputation.[22]

3.3.2.2 Extension

The original OLS and FE models (eq. 3 and 4, respectively) are extended as follows:

$$eduspend_{it} = c + \beta_1 multi_{it} + \beta_2 lnTradePPP_{it} + \gamma X_{it} + \varepsilon_{it} \qquad \text{(eq. 5)}$$

$$eduspend_{it} = c + \beta_1 multi_{it} + \beta_2 lnTradePPP_{it} + \gamma X_{it} + \eta_i + \varepsilon_{it}. \qquad \text{(eq. 6)}$$

As in the replication of Mulligan et al. (2004), only $lnTradePPP_{it}$ is added to the models and the GDP variable is now PPP-adjusted.[23] The replicated OLS results are not robust to the inclusion of trade openness (see Table 2, (6)). Multiparty competition is still significant, but its coefficient is smaller. Interestingly, various controls now turn out to be significant at the 1%-

[21] The extension of the time period in Table A.2 does not yield closer replication results and the R^2-value diminishes further. That is why the extension of Stasavage's (2005) model only includes the years 1980-1995. Nonetheless, the extension of the time period underlines the fragility of the original results.

[22] More specifically, he not only enlarges the dataset but also uses a slightly different specification by including a one-year lag of the dependent variable. Signs and significance of the variables included remain basically the same, but the size of the coefficients varies.

[23] While sign, size and significance of the trade openness measure remain unchanged, GDP interestingly switches its sign when not adjusted for PPP (compare Table 2 (6) and Table A.2 (6)). The effects of other controls are not robust to the different accounting methods either. While this could be a particularity of the African sample, further investigation might yield more insights.

level (GDP, share of rural population, share of population below the age of 15) and the GDP coefficient has switched its sign. The coefficient on trade openness is significant at the 1%-level, positive and comparatively large in size, thus confirming Hypothesis 2. Again, the inclusion of trade openness adds noticeably to the fit of the model since the R^2-value more than doubles.

The extended FE results (7) remain rather similar to the replicated results (4). Multiparty competition is not significant and the share of the population younger than 15 is negatively related to education spending. Trade openness is not significant and its inclusion does not improve the statistical fit of the model.

3.3.2.3 Robustness Checks

The general impression obtained from the replicated and extended results is that the original models are neither robust to variations in the sample nor to a change in the estimation technique. In addition to these issues, a statistical concern can be raised: apart from autocorrelation (Avelino et al., 2005), TSCS data have an increased chance of unobserved heterogeneity as well as spatial cross-sectional dependence (Breitung and Pesaran, 2005),

Beck and Katz (1995) discuss the shortcomings of OLS in TSCS panel data in depth and show that the estimates of the coefficients might remain consistent but are inefficient. Therefore, they advocate the use of PCSE that take into account the contemporaneous correlation in the errors as well as heteroskedasticity. These PCSE have been used by Stasavage as well as in both the replication and the extension. Yet, it should be noted that OLS with PCSE does not take care of autocorrelation. Many authors thus try to address autocorrelation by including a lagged dependent variable. Yet, as e.g. Achen (2000) shows, this method can result in underestimation of the importance of other variables, especially if they vary little over time. Therefore, following Rudra and Haggard (2001) and Avelino et al. (2005), the Prais-Winsten (PW) estimator is used instead of including a lag, in order to take care of autocorrelation. This estimator assumes a first order autocorrelation process with the coefficient of this process being the same across all panels and it transforms all observations to generate homoscedastic non-autocorrelated errors (Verbeek, 2008: 107-108).

First, equation 3 is estimated using PW (regression (5), Table 2). The results differ from the OLS regression (3) insofar as the multiparty dummy is not significant and some signs of the controls are reversed. The coefficient on GDP remains negative and significant.

With respect to the inclusion of trade openness, the PW estimation of equation 5, regression (8), is slightly different from the OLS extension (6): the coefficient on trade openness is still significant and relatively large in size, yet multiparty competition and the share

of people younger than 15 are not significant. As in the OLS regressions (3) and (6), the increase in the R^2-value due to the inclusion of trade openness can also be found in the PW estimations (5) and (8).[24]

Overall, the results of the PW estimations can be interpreted as evidence that the results might suffer from autocorrelation and that more testing is needed in order to determine the appropriate model and estimation technique. This stands in contrasts to the robustness checks conducted by Stasavage (2005) that do not indicate problems of autocorrelation in his sample. The change in the controls due to the inclusion of trade openness might also be caused by omitted variables, which is discussed in section 3.4.2.

In contrast to OLS and PW, the FE models have the advantage that they take care of all time-invariant heterogeneity and thus reduce the risk of omitted variables. The Hausman test (1978) clearly confirms the choice of FE over a random effects model. Wooldridge's test (2002) shows autocorrelation in the errors for all models except the one including trade openness adjusted for PPP. Since heteroskedasticity is also present within the panels and cannot be modeled here, robust and clustered standard errors should be used. With these standard errors, all estimated coefficients are insignificant, which questions the validity of the model.[25] Moreover, the errors are not normally distributed with the implications already discussed above.

The extension of Stasavage (2005) yields the same conclusion as the one of Mulligan et al. (2004): Hypothesis 1 is not confirmed since the sign and significance of democracy vary with the specification, whereas Hypothesis 2, a positive effect of globalization, is sustained in cross-country analyses. This holds true for a variety of additional specifications presented in Table A.2. With respect to within-county variations, trade openness does not seem to have any effect, except for the case in which trade openness is not logarithmized (Table A.2 (10)). A tentative explanation for this is given in section 3.4.3.

Overall and in contrast to Mulligan et al. (2004), Stasavage (2005) appears less robust to variations in both sample and models, which might also be due to statistical shortcomings.

[24] Note that the R^2-values of the PW and OLS estimation cannot be compared since the former is derived from the final regression with transformed dependent and independent variables whereas the OLS R^2 is based on an estimation using untransformed variables (Wooldridge, 2009: 422).
[25] Note that Stasavage (2005) does not specify what kind of standard errors he uses in the FE regressions.

Table 2: Replication and extension of Stasavage (2005)

	ORIGINAL RESULTS		REPLICATION			EXTENSION		
	OLS (1)	FE (2)	OLS (3)	FE (4)	PW (5)	OLS (6)	FE (7)	PW (8)
Multiparty competition	1.10*** (0.21)	0.358** (0.168)	0.802** (0.401)	-0.420 (0.332)	0.217 (0.355)	0.486* (0.291)	-0.440 (0.350)	0.203 (0.270)
Election year	-0.085 (0.388)	0.065 (0.206)	0.117 (0.507)	0.193 (0.339)	-0.094 (0.232)	0.170 (0.451)	0.203 (0.349)	-0.084 (0.221)
Log GDP per capita	1.49*** (0.12)	0.591*** (0.214)	-0.368*** (0.109)	-0.618 (1.053)	-0.358*** (0.123)	0.736*** (0.268)	-0.398 (1.077)	0.772*** (0.279)
Aid in % GDP	-0.0004 (0.007)	-0.021** (0.009)	-0.023 (0.019)	-0.023 (0.019)	-0.017 (0.014)	-0.011 (0.019)	-0.026 (0.019)	-0.006 (0.015)
% rural population	0.035*** (0.010)	0.012 (0.015)	0.006 (0.012)	-0.083* (0.044)	-0.007 (0.014)	0.063*** (0.012)	-0.078* (0.045)	0.054** (0.021)
% population <15	0.049 (0.039)	-0.272*** (0.077)	0.049 (0.033)	-0.202 (0.160)	0.027 (0.052)	0.096*** (0.034)	-0.223 (0.166)	0.066 (0.052)
Log trade openness (PPP)						1.969*** (0.273)	-0.500 (0.518)	1.751*** (0.291)
Constant	-10.32*** (1.84)	11.84*** (3.70)	5.432** (2.138)	26.44* (15.49)	7.507*** (2.740)	-16.30*** (2.030)	24.98* (12.83)	-13.77*** (3.059)
N	365	365	176	176	176	173	173	173
Countries	44	44	29	29	29	29	29	29
R^2	0.37		0.19	0.67	0.47	0.40	0.67	0.56
Within R^2		0.26		0.08			0.08	

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP. PCSE for OLS and PW estimates, normal SE for FE. Within country R^2-values for FE models from Stata's xtreg command, adjusted R^2 for FE from areg. Standard R^2 for OLS and PW (this is the norm for OLS estimates with PCSE and PW).

3.3.3 Time-series cross-sectional analysis of a global sample: Ansell (2008)

Ansell's (2008) is the only one of the three replicated models that includes a measure of trade openness.[26] His theory leads to several hypotheses, inter alia, "the expansion of democracy increases public education spending" and "increased integration with the global economy will lead to increased education spending in both democracies and autocracies" (pp. 249, 296), consistent with Hypothesis 1 and 2. His data encompass the years 1960-2000 and 113 countries. Like Stasavage (2005), Ansell (2008) also employs OLS with PCSE (eq. 7) and FE estimation (eq. 8), but uses a dynamic specification:

$$eduspend_{it} = c + \beta_1 eduspend_{it-1} + \beta_2 democ_{it} + \beta_3 lnTrade_{it} + \gamma X_{it} + \delta region_i + \varepsilon_{it} \quad \text{(eq. 7)}$$

$$eduspend_{it} = c + \beta_1 eduspend_{it-1} + \beta_2 democ_{it} + \beta_3 lnTrade_{it} + \gamma X_{it} + \eta_i + \varepsilon_{it} \quad \text{(eq. 8)}$$

where both models include a one-year lag of the dependent variable, $democ_{it}$ is the POLITY IV score, $lnTrade_{it}$ is the logged sum of exports and imports over GDP, X_{it} is a vector of controls, $region_i$ are regional dummies, η_i represents country specific effects and ε_{it} is the error term. The controls include the fraction of young people, GDP and its square, a measure of total population and total government expenditures net of education spending as well as a linear time trend.

The original regressions (1) and (2) in Table 3 indicate a significant and positive effect of both democracy and trade openness, which is consistent with Ansell's (2008) hypotheses and the ones derived from the literature. Sign and significance of the coefficients on GDP and its square vary by specification, the other controls are consistently insignificant, except for the time trend which appears to have a negative effect.

3.3.3.1 Replication

Again, replication problems arise due to the data. Despite the use of the same source, data for the dependent variable covers 178 states with an average of 10.7 observations for each state, instead of 113 countries with an average of 15.4 observations. Nonetheless, mean and standard deviation are quite similar (4.3 vs. 4.2 and 2.1 vs. 1.9, respectively). Moreover, doubts remain concerning the exact construction of government consumption excluding education as well as the source and measurement method of GDP. Since outliers do not significantly

[26] Apart from (exports+imports)/GDP, Ansell (2008) also uses a measure of openness constructed by Hiscox and Kastner (2006) that shows the country's deviance from its optimal amount of trade (predicted by a gravity model). Since higher values of the index imply higher levels of protectionism, estimation results are expected to have the opposite sign. Findings from this index are consistent with the conventional variable and are not replicated here.

influence the results, all observations are kept to leave the sample as large as possible. In the end, the replication dataset includes more countries but fewer observations (126 vs. 113 and 1161 vs. 1501) which might be a reason for differing replication results.

In the replication, the coefficient on democracy is only significant at the 10% level in the OLS estimation (3) and is smaller in magnitude in both regressions (3) and (4). Evidence for Hypothesis 1 is therefore rather weak. The coefficient on trade openness (not adjusted for PPP) is not significant in either estimation which differs from the original results and contradicts Hypothesis 2.[27] Interestingly, the R^2-values are all slightly higher in the replication which could be due to the relatively large and significant effect of the lagged dependent variable.

With regard to replication, the original time span yields the closest results and is therefore used in the extension. The results for an increased time span are presented in Table A.3. Overall, the replication results suggest that the conventional trade openness measure as well as the polity score and thus Ansell's (2008) models are not robust in a different sample.

3.3.3.2 Extension

In the extension, the trade openness variables are substituted by the PPP-based measure which yields the following models:

$$eduspend_{it} = c + \beta_1 eduspend_{it-1} + \beta_2 democ_{it} + \beta_3 lnTradePPP_{it} + \gamma X_{it} + \delta region + \varepsilon_{it} \quad \text{(eq. 9)}$$

$$eduspend_{it} = c + \beta_1 eduspend_{it-1} + \beta_2 democ_{it} + \beta_3 lnTradePPP_{it} + \gamma X_{it} + \eta_i + \varepsilon_{it}. \quad \text{(eq. 10)}$$

Using the PPP-based measures of trade openness and GDP,[28] the sample is reduced from 1161 to 889 observations and from 126 to 125 countries. Trade openness is now significant in the OLS regression (6) in Table 3. This is consistent with Hypothesis 2 as well as the original findings. The other coefficients, including the one on democracy, are insignificant with the exception of government expenditures in the FE model. As in the extension of Stasavage (2005), trade openness is not significant in the FE model. Moreover, while the R^2-value remains unchanged in the OLS estimation, it decreases in the FE model from 66% to 62% for the within explanatory power and stays the same for the overall adjusted R^2.

[27] Since the regional dummies do not turn out to be significant, they are dropped in Table A.3, (1). Now, neither democracy nor trade is significant when estimated by OLS. Again, the sample is enlarged to include 12 more years (1960-2012) to check whether the replication results are robust A.3, (3) and (4). The sample now includes 141 countries and 2016 observations. Neither of the two main independent variables turns out to be significant but the size of the coefficients on the lag is even larger which hints at a unit root process in the dependent variable.

[28] The use of unadjusted GDP does not significantly alter the results (see Table A.3, (6-8)).

3.3.3.3 Robustness checks

As already discussed in the robustness checks of Stasavage (2005), autocorrelation might be a problem, thus justifying the inclusion of the lag. On the other hand, this lag might lead to the underestimation of other variables (Achen, 2000).[29] Therefore, the PW estimation technique is used again, without a lag (see Table 3, regression (5) and (8)). Using the conventional measure of trade openness, both coefficients on democracy and trade openness are positive, larger in size than in the original results and also statistically significant (see regression (5)). This is in line with the original results, both hypotheses and Achen's (2000) concern. With the PPP-based measure, both coefficients increase in size and remain significant at the 1%-level (see regression (8)). Again, the R^2-value increases when the PPP-based measure is used (from 51% to 56%).

Regarding the robustness of the FE model, the findings are similar to the ones in Stasavage (2005): the Hausman test (1978) favors the FE model, heteroskedasticity is present within the panels and the errors are not normally distributed. Moreover, all FE models suffer from serial correlation, as indicated by Wooldridge's test (2002). Therefore, robust and clustered standard errors have been used, both by Ansell (2008) and in this paper. Nonetheless, the FE estimator is inconsistent for finite time periods in this specification due to the correlation of the lagged dependent variable and the error term (Verbeek, 2008: 378).

These results suggest that the PPP-based measure of trade openness more robust than the unadjusted one in cross-sectional analyses. Furthermore, the use of the PW technique might be a good alternative to a lagged dependent variable when accounting for autocorrelation. Yet, trade openness, measured in either way, does not appear to have a significant within-county effect on public education spending except for when trade openness is not logarithmized. This is consistent with the replication and extension results of Stasavage (2005).

[29] Yet, it is hard to explain why Ansell's (2008) original estimations do not seem to suffer from this problem.

Table 3: Replication and extension of Ansell (2008)

	ORIGINAL RESULTS		REPLICATION			EXTENSION (PPP MEASURES)		
	OLS	FE	OLS	FE	PW	OLS	FE	PW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged DV	0.792***	0.608***	0.910***	0.773***		0.898***	0.750***	
	(0.034)	(0.018)	(0.022)	(0.044)		(0.026)	(0.056)	
Polity IV score	0.016***	0.012**	0.007*	0.007	0.030***	0.008	-0.002	0.043***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.009)	(0.006)	(0.008)	(0.013)
Log trade openness	0.282***	0.232***	0.039	0.110	0.607***	0.155***	0.159	1.078***
	(0.061)	(0.089)	(0.045)	(0.110)	(0.144)	(0.006)	(0.132)	(0.250)
% population <15	0.005	-0.015	0.003	-0.002	-0.008	0.000	-0.012	-0.007
	(0.006)	(0.011)	(0.004)	(0.011)	(0.015)	(0.004)	(0.015)	(0.013)
Log GDP	-0.390	2.180***	0.191	0.733*	-0.099	0.073	0.381	-0.459
	(0.251)	(0.833)	(0.215)	(0.389)	(0.660)	(0.168)	(0.452)	(0.416)
(Log GDP)²	0.010*	-0.041**	-0.003	-0.012	0.008	-0.004	-0.023	0.017
	(0.005)	(0.017)	(0.004)	(0.008)	(0.013)	(0.007)	(0.017)	(0.017)
Log population	-0.058	-0.012	-0.043*	-0.151	-0.313***	0.044	0.344	0.124
	(0.045)	(0.225)	(0.024)	(0.262)	(0.071)	(0.039)	(0.389)	(0.196)
Government exp.	0.003	0.003	0.005	0.018**	0.034***	0.001	0.025**	0.025*
	(0.006)	(0.006)	(0.006)	(0.008)	(0.013)	(0.007)	(0.012)	(0.013)
Year	-0.019***	-0.016***	-0.004	-0.011	-0.012	-0.003	-0.002	-0.030**
	(0.004)	(0.006)	(0.002)	(0.007)	(0.007)	(0.004)	(0.013)	(0.01)
Constant	41.43***	5.829	5.145	14.22	29.01	4.365	-2.506	61.52***
	(7.494)	(14.297)	(5.916)	(13.50)	(17.04)	(7.492)	(22.97)	(23.33)
N	1501	1501	1161	1161	1578	889	889	1192
Countries	113	113	126	126	140	125	125	137
R^2	0.83		0.90	0.91	0.51	0.90	0.91	0.56
Within R^2		0.50		0.66			0.62	

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP. OLS and PW with PCSE, FE with robust SE. OLS and PW with regional dummies, omitted region is North America. Within country R^2-values for FE models from Stata's xtreg command, adjusted R^2 from areg. Standard R^2 for OLS and PW.

3.4 Discussion

Replication of (time-series) cross-sectional estimations is often difficult due to the diversity of data sources and the variation in the definition of variables. Without the original dataset, close replication is nearly impossible. Nonetheless, some specifications in this paper yield relatively close results with respect to the main variables. Therefore, it is valid to compare the extended results to the original articles.

The three replicated models use very different samples, independent variables, controls and estimation techniques: Mulligan et al. (2004) use averages of all variables for 112 countries and therefore only capture cross-country differences which they analyze using OLS. Stasavage (2005) restricts his sample to African countries, estimating both cross- and within-country variations in the OLS and FE model respectively. Ansell (2008) examines the largest sample and uses the same estimation techniques as Stasavage (2005), but includes a lagged dependent variable.

Despite these differences and in contrast to other variables, trade openness, when adjusted for PPP, appears to have a consistent, positive and relatively large effect on government spending on education in cross-country comparisons and adds to the statistical fit of the respective model. This effect appears to be more robust than the one from the conventional trade openness measure (see Table 3), that does not take into account PPP and therefore overstates the importance of trade relative to GDP in developing countries. Nonetheless, the robustness of log trade openness has to be questioned, since it is not significant in any FE model. Potential reasons for these systematic differences are discussed below in section 3.4.3. Overall, the evidence for Hypothesis 2 found is inconclusive.

The second important variable that is very popular in the literature and captures the main supply side mechanism, regime type, does not seem to be robust, either: in both the replication and extension of Mulligan et al. (2004), the average POLITY IV score is not significant. Moreover, when dropping regime type from the models, all other estimates remain roughly the same and the adjusted R^2-value does not change. Empirically, this implies that democracy could be omitted, but this would still need a theoretical foundation. In Stasavage (2005), sign and robustness of democracy, defined as multiparty competition, depend on the specification. Only the OLS results show a significant and positive effect. In addition, in the cross-country analyses, the omission of multiparty does not cause a decrease in the R^2-value, but it does so in the FE model, which is counter-intuitive. In the replication and extension of Ansell (2008), democracy, measured by the Polity score, is mostly positive and significant in the OLS and PW specifications, while it is insignificant in the FE models. The omission of democracy does not noticeably change the R^2-value of any model. In contrast to many previous studies that find a

positive effect, the results presented here suggest that there is no clear effect of democracy on education spending that holds true for different samples and specifications. The evidence for Hypothesis 1 in this paper is therefore mixed.

In sum, these inconclusive results clearly show that the robustness of empirical models, that try to reveal the determinants of public education spending, is often questionable. Do the effects of trade openness and/or democracy matter across countries but not within? If so, why? More research is needed to answer these questions. Ideally, this research would start by building a more comprehensive supply and demand model of government spending on education which will be able to guide empirical estimations.

Since this paper only replicates and extends previous models, it is necessarily subject to the same (potential) problems as the original articles. In the following, concerns with regard to data and measurement problems, econometric issues and the interpretation of empirical results are addressed.

3.4.1 Data and measurement issues

Two important concerns are briefly discussed below: data availability and measurement problems.

One frequent issue of cross-sectional analyses including developing countries is data availability. The tables B.2-B.4 in the appendix show the summary statistics for the replicated models. It can be seen that, while country characteristics such as the structure of the population are well covered by the data, variables concerning government spending in general, on education or on aid, have fewer observations. If countries that trade less or are less democratic tend to have less complete statistics of their government expenditures, this can lead to an upward bias of the estimated effects of trade openness and democracy on education spending.

Moreover, replication and extension results point toward the problem that many empirical models are not robust to changes in the sample. Since close replication combined with statistical robustness checks would increase transparency of empirical research, more accessible data would be desirable:

> "Big data promises big progress. But large datasets also make replication impossible without the author's cooperation. And the incentive for authors to cooperate is, at best, mixed. It is therefore the responsibility of editorial boards and the directors of organizations like the NBER to make open access obligatory." (Eichengreen, 2013)

When data is available, variables might be measured with error. If these errors are random, they only increase the noise, but if the errors occur systematically, e.g. because non-

democracies want to make a good impression on the international stage and therefore overstate their education expenditures, the regression results will be biased.

Moreover, different ways of accounting for the dependent variable (and most of the regressors) exist. Many studies use education spending as a percentage of GDP, which shows the global allocation of the resources of a particular country. Yet, as Rudra and Haggard (2001) argue, this conventional measure does not show how governments assign priorities and how they distribute the resources they actually control. This would be more accurately reflected in government spending as a percentage of total government expenditures. Therefore, the research question has to be precisely formulated, as it ultimately determines which measure is best to use.

Furthermore, all three models examine education spending, democracy (and trade openness) in levels. While this renders results comparable to the majority of the literature, it can reasonably be argued that both changes in the level of trade openness and democracy can lead to a shift in preferences for education spending. For a more encompassing understanding of public education spending, the effect of the respective changes should also be investigated, as for instance in Avelino et al. (2005).[30] Unfortunately, this is beyond the scope of this paper.

3.4.2 Econometric issues

In addition to data and measurement issues, the replicated models might suffer from econometric shortcomings, most importantly specification problems and violations of OLS assumptions. The robustness checks in the respective sections have shown that although many issues have been taken care of, some concerns remain. In what follows, problems common to most models, namely omitted variables, unit root processes in the variables as well as endogeneity, are briefly outlined.

The literature suggests a large variation of potentially important variables that could be included in the analysis of education spending, such that the somewhat arbitrary inclusion of independent variables and controls might lead to omitted variable bias, the direction of which is impossible to predict. In fact, the importance of the regional dummies in Ansell (2008) suggests that region-specific effects on education spending are not entirely caught by the independent variables.[31]

In many studies, the inclusion of control variables "[…] with an emphasis on those used in prior studies of policy determination […]" (Mulligan et al., 2004: 55) might offer a certain

[30] They employ an error correction model and find that both trade openness (PPP) and democracy are positive and significant but changes in those variables do not appear to be significant.

[31] Even though the dummies do not turn out significant in the replication, an omission of the dummies leads to altered results (see Table A.3 (1)). In all other specifications, at least some of coefficients on the dummies are significant.

degree of comparability to previous research but lacks theoretical reasoning. In general, some controls are included because theories of other authors argue to do so, yet the choice of which specific set of control variables to include often appears to be arbitrary. Usually, only the inclusion of controls, if anything, is justified, but never the exclusion of particular variables. For instance, Ansell (2008) explains why he uses the share of persons younger than 15, but he does not mention the "omission" of the share of rural population, although both measures are not explicitly mentioned in his theoretical framework and are common control variables.

Furthermore, Ansell's (2008) model hints at another potential problem: the fragility of the controls and the large coefficients on the lag hint at the possibility that some variables might contain a unit root, at least in the case of government education spending. The Fisher type unit root rest shows inconclusive results for the regressand, depending on the choice of the test statistic. Since the alternative hypothesis is that at least one panel is stationary, even a rejection of the null hypothesis would imply that several panels may contain a unit root which further complicates the analysis. Theoretically, all variables should be tested for a unit root to make sure that they are all integrated of the same order. Otherwise, the regressions would be spurious. In addition, cointegration tests should help determining the correct model for TSCS data, thus also for Stasavage (2005).

Lastly, a potential issue common to all models, and thus not yet discussed in the individual robustness checks, is endogeneity. For many independent variables, it is hard to argue that there is no feedback effect. With respect to regime type, for instance, increased education spending can result in better education of democratic citizens and thus positively influences democratization (Barro, 1999) or it might help maintain totalitarian regimes (Lott, 1999). Moreover, if increased education results in better education of women, birth rates may be lowered (Drèze and Murthi, 2001), such that in the long run, both total population and the share of young people are going to decrease whereas the share of older people is going to rise. Furthermore, if increased education spending also advances technology and innovation (Varsakelis, 2006), it is likely to decrease the share of value added from agriculture in the long run and might also contribute to specialization and increased trading. Therefore, education spending is likely to have a variety of long run effects on the independent variables. Since those are mostly medium and long run effects, endogeneity does not seem to be the most important concern.[32] This conclusion is consistent with Ansell's (2008) estimations that use

[32] It could be argued that Mulligan et al. (2004) might capture some endogeneity in their estimates because they use averages of their variables. Yet, their independent variables are averaged across 1960-1990 and the dependent variable only across 1980-1990 which reduces the risk of endogeneity.

lags as instrumental variables and yield the same results with respect to democracy and trade openness.[33]

In the last decade, the research on panel data robustness checks and suitable estimators has greatly enhanced.[34] As the above discussion has shown, the employed cross-sectional and within-country estimators are suboptimal. If more extensive and complete data were available, more sophisticated estimators could be used in the analysis of education spending. That is why imputations techniques might be interesting to enlarge datasets. Yet in this paper, missing data is dealt with by listwise deletion, as in the baseline regressions of the original models.

3.4.3 Interpretation

The main results, inconclusive evidence for both hypotheses, have already been discussed above. Concerning trade openness, it is noteworthy that its effect, whether measured in PPP or not, appears to be consistently significant in cross-country comparisons only: countries that trade more tend to spend more on public education. But what happens if a country opens up to international trade? According to the estimates, not much. Coefficients on log trade openness in FE models have different (mainly positive) signs in replications and extensions, but remain insignificant in all replications and extensions.[35]

Why does trade openness seem to matter in cross-country comparisons but not within countries? While the theoretical literature does not offer an explanation, statistical issues might play a role. Two tentative rationales, omitted variables and low within-country variance, are presented in the following.

In light of manifold potential determinants of public education spending, the trade openness variable might capture important omitted influences that vary across countries but not over time, while trade openness is not a significant determinant itself. In this case, the omitted influences captured by trade openness are absorbed in the FE models, rendering trade openness insignificant. The suspiciously large increase in the R^2-values when trade openness is included might also hint at the possibility that several effects are captured instead of just the one of trade openness. In Mulligan et al. (2004), for instance, the inclusion of trade openness raises the R^2-value from 24% to 43%, whereas the inclusion of the share of value added from agriculture

[33] The variations in the controls could be due to, inter alia, omitted variables or endogeneity. A final conclusion with respect to the results of the IV regressions can thus not be reached.

[34] For an overview of recent developments in tests of both the first and the second generation (assumption of cross-sectional independence and dependence, respectively) see Hurlin and Mignon (2006).

[35] Note that in the original FE results, the coefficient on trade openness is positive and significant in Ansell (2008). All other authors cited in the literature review do not use FE and therefore do not allow further comparisons.

only adds one percentage point, although it appears to be highly significant.[36] If this explanation was correct, the omitted variable(s) would be country-specific since the omission of regional dummies in Table A.3 (1) and (2) does not significantly alter the coefficient on trade openness.[37]

Finally, this explanation leads to the important question of what omitted effects could be captured by trade openness in the cross-country comparisons. One potential factor could be institutional effects that are usually hard to observe and to account for, but as they are practically time-invariant, they can be controlled for in FE models. It can be argued that such effects would influence both trade openness and education spending: it is plausible that, (cet. par.), weak institutions lower trade since they imply higher uncertainty and costs for the trading partner. Weak institutions might also lower spending on education, e.g. in favor of corruption. As Mauro (1998) and Gupta et al. (2002) have shown, corruption decreases public spending for goods that are hard to target such as education. Therefore, institutional effects and their outcomes might be captured by trade openness which thus turns out to be significant in cross-sectional studies. In addition to institutions, cultural effects might also play a role. Tentatively, a measure for institutional quality could be included in future empirical models. Yet, for more certainty, this issue should ideally be solved theoretically.

Another rationale for the insignificance of trade openness for public education spending in the FE models might be its low within variance, since the FE estimator is inefficient when variables that vary little within the fixed units: „[... this] does not only imply low levels of significance, [but] point estimates are also unreliable“ (Plümper and Troeger, 2007: 127).

When looking at Table B.3, it can be seen that both trade variables, regardless of PPP adjustment, have a comparatively low within-country variance in Stasavage's (2005) sample; only the dummy variables and GDP have a lower variance. All these variable turn out to be insignificant in the FE estimations in Table 2 and Table A.2, whereas e.g. GDP is consistently significant in the cross-country specifications. The same picture emerges from Ansell's (2008) models: together with the log of population, the two trade variables have the lowest within-country variance (Table B.4). These three variables are never significant in the FE estimations. Yet, as the non-logarithmized trade variable shows, this problem does not seem to be due to the fact that trade openness does not carry enough information for a significant within-country estimate. Rather, the logarithmic transformation significantly reduces the within-country

[36] The other replicated authors do not vary their control variables which impedes further comparisons of increases in the R^2-value. Since panel data models tend to have lower R^2-values, the increase due to trade openness can indeed be evaluated as very large.

[37] Nor do the estimates for the coefficients change much when the dummies are omitted in the other models (results are not reported here).

variance. Indeed, when trade openness is not log transformed, it is significant in FE models as well (Table A.2 (10) and A.3 (10)). In both cases, the estimated effect of trade openness is now negative, which is consistent with the findings of Rudra and Haggard (2001). Nonetheless, this leaves new questions for future research, since it contradicts Hypothesis 2. Also, it is hard to explain why in Stasavage's (2005) sample, this FE model has a comparatively high R^2-value, whereas the fit of the FE model does not improve in Ansell's (2008) sample. Moreover, it is difficult to tell why the original results of Ansell (2008) do not suffer from this problem, although (unadjusted) log trade openness is used. In general, it is not clear, why some variables, e.g. GDP and trade openness, are log transformed but, for instance, aid in Stasavage (2005) and government expenditures in Ansell (2008) are not. In sum, the insignificant log trade openness coefficient in FE models might be a result of the log-transformation, but this is only a tentative explanation.

Both the omitted variable and the low variance explanation seem to be supported by the R^2-values that remain the same or even decrease after the inclusion of log trade openness in FE models (see Table 2, (4) and (7); Table 3, (4) and (7)). If trade openness captured time-invariant country specific effects, it should not add to explaining the within country variance of education spending and if it had too few information, its inclusion should not matter either. Nonetheless, these two approaches to an explanation are only tentative and it should be kept in mind that previous empirical research has produced many different results. Therefore, it should be investigated whether this systematic difference between cross- and within-country estimates also holds for different models or whether it is a mere coincidence and a particularity of this paper.[38] In case of confirmation, a more sophisticated and ideally theoretical explanation will be needed.

Finally, with respect to a broader interpretation of the findings, it has to be kept in mind that the level of public education spending is only a formal part of education policy and might not be tightly associated with neither quality nor effectiveness of a country's educational system. On the one hand, it has been shown that expenditure levels are a determinant of literacy, which in turn can be seen as a rough measure of the educational system's effectiveness (Bueno de Mesquita et al., 2003: 293). On the other one, not all studies find a positive relationship between government spending on education and educational attainment (e.g. Gupta et al., 2002).

[38] Since the statistical robustness checks show several problems, particularly for the FE models, it could be the case that the FE models are flawed and thus not reliable.

Moreover, different ways of accounting might lead to a misleading cross-country picture of spending data (Brown and Hunter, 2004), which cannot simply be interpreted as an indicator of a government's commitment to education:

> "Expenditure levels can be quite distorted by patterns of 'hidden' rents, disguised as higher salaries or cushy jobs for regime supporters. A politically corrupt state that is capturing larger rents and distributing them to its supporters through inflated or unnecessary expenditures, for instance, may appear to be spending more on education than a politically efficient regime, but the level of actual services delivered to citizens will be much lower in the first than in the second case." (Baum and Lake, 2003: 336)

Therefore, additional measures (ideally without measurement error) are needed to examine quality and effectiveness of educational systems across countries. Yet, comparability of these aspects is even more difficult to achieve and suitable measures do not yet encompass equally large samples.[39] Public education spending therefore remains a widely used proxy for education policy as a subset of social policy, which would be important to understand more in-depth.

Given potential omitted variables and variations in measurement (e.g. GDP vs. GDP per capita) in addition to econometric issues, it is not surprising that signs and significance of independent variables vary, making a coherent interpretation impossible. This has already been noticed elsewhere, e.g. by Plümper and Martin (2003: 40): "The interpretation of the […] regression results is, however, restricted by lack of a consistent theoretical model. As a consequence, we cannot claim to have found the 'right' model". Until a consistent theory allows empirical estimation with external validity and thus offers reliable inference, the empirical results concerning public spending on education should be interpreted very carefully.

[39] For further discussion of measurement issues of quality see e.g. Hanushek and Wößmann (2007).

Da steh ich nun, ich armer Tor!/ Und bin so klug als wie zuvor;/ Heiße Magister, heiße Doktor gar/ Und ziehe schon an die zehen Jahr/ Herauf, herab und quer und krumm/ Meine Schüler an der Nase herum-/ Und sehe, daß wir nichts wissen können![40] Faust

4 Conclusion

So, what do we really know about the determinants of public spending on education? We know that much remains to be done, if the ultimate goal of research in the field of political economy is to understand and inform politics, e.g. how and under which circumstances an efficient educational system can be built and maintained.

This paper provides anecdotal evidence that the lack of consistent theory can lead to rather arbitrary empirical models, often yielding results that are not robust to changes in the sample, the inclusion of additional variables or different measurement methods. Moreover, many models can be questioned on statistical grounds.

Although trade openness, as the additional variable investigated here, adds significantly to the goodness of fit of most models, its positive and significant effect on public education spending appears to be consistent only across countries. Current theory is not able to explain this discrepancy between estimated cross-sectional and within-country effects of trade openness and the statistical explanations offered here are only tentative. The estimated effects of regime type are also mixed, despite of more consistent previous findings. These outcomes underline the need for a coherent theory of education spending that can guide empirical analysis. In addition, more transparency and data sharing is needed which will facilitate replications and robustness checks, which in turn can lead to empirical advancements.

Moreover, the informative value of public spending on education is limited and only the consideration of different dependent variables, e.g. public education spending, enrollment rates and educational attainment can lead to a sophisticated picture of education policy. Here, it will be important to try synthesizing different branches of research and their results.

Finally, interpretation of the results ought to be careful and "economists should match honesty about what their research says with honesty about the inherently provisional nature of what passes as evidence in their profession" (Rodrik, 2013). The same criticism obviously applies to all social scientists, especially because politicians and journalists often tend to overstate the importance of a particular finding: "One thing that experts know, and that non-experts do not, is that they know less than non-experts think they do" (Basu, 2013).

[40] And here, poor fool! with all my lore/ I stand, no wiser than before:/ I'm Magister - yea, Doctor - hight,/ And straight or cross-wise, wrong or right,/ These ten years long, with many woes,/ I've led my scholars by the nose,-/ And see, that nothing can be known!

As has been demonstrated, results are seldom unambiguous, entirely robust and easy to explain. Therefore, caution concerning interpretation should prevail until a consistent theory permits empirical analysis with external validity – that is what we really know about the determinants of public spending on education so far.

5 References

Acemoglu, D. (2010): "Theory, General Equilibrium and Political Economy in Development Economics", *NBER Working Paper* 15944.

Achen, C. (2000): "Why Lagged Dependent Variables Can Suppress the Explanatory Power of the Independent Variables", presented at the Annual Meeting of the Political Methodology Section of the American Political Science Association at UCLA.

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

Ansell, B. W. (2008): "Traders, Teachers, and Tyrants: Democracy, Globalization, and Public Investment in Education", *International Organization*, 62(2): 289-322.

Avelino, G., D. S. Brown and W. Hunter (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.

Barro, R. J. (1999): "Determinants of Democracy", *Journal of Political Economy*, 107(6): S158-S183.

Barro, R. J. (2001): "Human Capital and Growth", *The American Economic Review*, 91(2): 12-17.

Basu, Kaushik (2013): "Two Policy Descriptions for the Global Crisis", *Project Syndicate*, 23.04.2013, available at: <http://www.project-syndicate.org/commentary/lessons-from-the-world-bank-imf-spring-meetings-by-kaushik-basu> [accessed 13.07.2013].

Baum, M. A. and D. A. Lake (2003): "The Political Economy of Growth: Democracy and Human Capital", *American Journal of Political Science*, 47(2): 333-347.

Beck, N. and J. N. Katz (1995): "What to do (and not to do) with Time-Series Cross-Section Data", *The American Political Science Review*, 89(3): 634-647.

Breitung, J. and M. H. Pesaran (2005): "Unit roots and cointegration in panels", *Deutsche Bundesbank Discussion Paper*, Series 1, 42/2005.

Breusch, T. S. and Pagan, A. R. (1980): "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics", *Review of Economic Studies*, 47: 239-253.

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

Bueno de Mesquita, B., A. Smith and R. M. Siverson, J. D. Morrow (2003): "*The Logic of Political Survival"*, Cambridge: The MIT Press.

Campante, F. R. and D. Chor (2012): "Why was the Arab World Poised for Revolution? Schooling, Economic Opportunities, and the Arab Spring", *Journal of Economic Perspectives*, 26(2): 167-188.

Drèze, J. and M. Murthi (2001): "Fertility, Education and Development: Further Evidence from India", *Population and Development Review*, 27(1): 33-63.

Easterly, W. and R. Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions", *The Quarterly Journal of Economics*, 112(4): 1203-1250.

Eichengreen, B. (2013): "Open-Access Economics", *Project Syndicate*, 17.05.2013, available at: <http://www.project-syndicate.org/commentary/the-real-problem-with-reinhart-and-rogoff-by-barry-eichengreen> [accessed 13.07.2013].

Garrett, G. (2001): "Globalization and Government Spending around the World", *Studies in Comparative International Development*, 35(4): 3-29.

Gradstein, M. (2003): "The Political Economy of Public Spending on Education, Inequality, and Growth", *World Bank Policy Research Working Paper* 3162.

Gupta, S., M. Verhoeven and E. R. Tiongson (2002): "The effectiveness of government spending on education and health care in developing and transition economies", *European Journal of Political Economy*, 18(4): 717-737.

Hanushek, E. A. and L. Wößmann (2007): "The Role of Education Quality in Economic Growth", *World Bank Policy Research Working Paper* 4122.

Hausken, K, C. W. Martin and T. Plümper (2004): "Government Spending and Taxation in Democracies and Autocracies", *Constitutional Political Economy*, 15(3): 239-259.

Hausman, J. A. (1978): "Specification Tests in Econometrics", *Econometrica*, 46: 1251-1271.

Hicks, A. and D. Swank (1992): "Politics, Institutions, and Welfare Spending in Industrialized Democracies", *American Political Science Review*, 86(3): 658-674.

Hiscox, M. and S. Kastner (2006): "A General Measure of Trade Policy Orientations: Gravity-Model-Based Estimates for 82 Nations, 1960-1992", unpublished manuscript, Harvard University, Cambridge, Mass.

Hurlin, C. and V. Mignon (2006): "Second Generation Panel Unit Root Tests", available at: <http://halshs.archives-ouvertes.fr/halshs-00159842> [accessed 20.07.2013].

Iversen, T. (2001): "The Dynamics of Welfare State Expansion: Trade Openness, Deindustrialization and Partisan Politics". In: P. Pierson (ed.) *The New Politics of the Welfare State*, Oxford University Press.

Iversen, T. and T. R. Cusack (2000): “The Causes of Welfare State Expansion: Deindustrialization or Globalization?”, *World Politics*, 53(3): 313-349.

Jarque, C. M. and A. K. Bera (1980): “Efficient Test for Normality, Homoscedasticity and Serial Independence of Regression Residuals”, *Economics Letters*, 6: 255-259.

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

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

Lott, J. R. (1999): “Public Schooling, Indoctrination, and Totalitarianism”, *Journal of Political Economy*, 107(S6): S127-S157.

Lucas, R. E. (1988): “On the Mechanics of Economic Development”, *Journal of Monetary Economics*, 22(1): 3-42.

Mankiw, G., D. Romer and D. Weil (1992): “A Contribution to the Empirics of Economic Growth”, *Quarterly Journal of Economics*, 107(2): 407-437.

Mason, A. D. and S. D. Rozelle (1998): “Schooling Decisions, Basic Education, and the Poor in Rural Java”, unpublished manuscript, UC Davis, available at: <agecon.ucdavis.edu/people/faculty/scott-rozelle/docs/Publications/ARTICLE4.pdf>, [accessed 13.07.2013].

Mauro, P. (1998): “Corruption and the composition of government expenditure”, *Journal of Public Economics*, 69(2): 263-279.

Mulligan, C. B., R. Gil, X. Sala-i-Martin (2004): “Do Democracies Have Different Public Policies than Nondemocracies?”, *The Journal of Economic Perspectives*, 18(1): 51-74.

Plümper, T. and C. W. Martin (2003): “Democracy, government spending, and economic growth: A political-economic explanation of the Barro-effect”, *Public Choice*, 117 (1/2): 27-50.

Plümper, T. and V. E. Troeger (2007): “Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analysis with Unit Fixed Effects”, *Political Analysis*, 15(2): 124-139.

Quinn, D. (1997): “The Correlates of Changes in International Financial Regulation”, *American Political Science Review*, 91(3): 531-552.

Ramsey, J. B. (1969): “Test for Specification error in Classical Linear Least Squares Regression Analysis”, *Journal of the Royal Statistical Society, Series B*, 31: 350-371.

Rodrik, D. (1998): "Why Do More Open Economies Have Bigger Government?", *Journal of Political Economy*, 106(5): 997-1032.

Rodrik, D. (2013): "What use are Economists?", *Project Syndicate*, 10.05.2013, available at: <http://www.project-syndicate.org/commentary/the-provisional-nature-of-economic-research-by-dani-rodrik#zs2qh3e2JIMZmdsl.01> [accessed 13.07.2013].

Rudra, N. (2002): "Globalization and the Decline of the Welfare State in Less-Developed Countries", *International Organization*, 56(2): 411-45.

Rudra, N. and S. Haggard (2001): "Globalization, Domestic Politics, and Welfare Spending in the Developing World", presented at the Annual Meeting of the American Political Science Association, San Francisco.

Shleifer, A. and R. W. Vishny (1993): "Corruption", *The Quarterly Journal of Economics*, 108(3): 599-617.

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

Varsakelis, N. C. (2006): "Educations, political institutions and innovative activity: A cross-country empirical investigation", *Research Policy*, 35: 1083-1090.

Verbeek, M. (2008): "*A Guide to Modern Econometrics*", 3rd ed., West Sussex: Wiley.

Wooldridge, J. M. (2002): "*Econometric Analysis of Cross Section and Panel Data*", Cambridge, MA: MIT Press.

Wooldridge, J. M. (2008): "*Introductory Econometrics. A Modern Approach*", 4th ed., Mason: South-Western.

Appendices

A Additional output

Table A.1: Additional output for Mulligan et al. (2004)

	EXTENSION					
	Unadjusted trade (log) and GDP		Trade (PPP and log), unadjusted GDP		Trade (PPP, no log), GDP (PPP)	
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy index	0.865*	0.816	0.525	0.481	0.472	0.367
	(0.509)	(0.512)	(0.463)	(0.462)	(0.507)	(0.506)
Communist dummy	1.078**	1.029**	0.944**	0.894**	1.024**	0.852*
	(0.428)	(0.429)	(0.400)	(0.400)	(0.453)	(0.459)
British legal origin	0.175	0.080	0.116	0.038	0.314	0.272
	(0.296)	(0.313)	(0.271)	(0.281)	(0.304)	(0.311)
% population >65	0.038	0.046	-0.006	0.004	0.021	0.045
	(0.057)	(0.057)	(0.054)	(0.054)	(0.059)	(0.060)
Log(population)/10	1.642	1.489	2.023*	1.835*	-0.379	-0.810
	(1.311)	(1.353)	(1.050)	(1.062)	(0.949)	(0.980)
Log real GDP per capita	0.171	-0.012	0.124	-0.036	0.281*	-0.032
	(0.196)	(0.279)	(0.180)	(0.261)	(0.157)	(0.241)
Share of value added from agriculture		-0.979		-0.919		-2.496
		(1.672)		(1.483)		(1.645)
Log trade openness	1.405***	1.149***	1.457***	1.438***	0.021***	0.018**
	(0.362)	(0.403)	(0.247)	(0.260)	(0.007)	(0.007)
N/Countries	112	109	111	108	112	109
Adjusted R^2	0.33	0.32	0.42	0.41	0.29	0.29

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP.
All regressions include a constant (not presented), OLS estimates with normal SE.

Table A.2: Additional output for Stasavage (2005)

	1970-1995 REPLICATION		1980-1995 EXTENSION								
			Unadjusted trade (log) and GDP			Trade (PPP and log), unadjusted GDP			Trade (PPP, no log), GDP(PPP)		
	OLS	FE	OLS	FE	PW	OLS	FE	PW	OLS	FE	PW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Multiparty competition	0.932*	-0.414	0.837**	-0.427	0.312	0.366	-0.456	0.057	0.234	-0.405	0.011
	(0.533)	(0.575)	(0.344)	(0.342)	(0.316)	(0.319)	(0.351)	(0.326)	(0.286)	(0.323)	(0.235)
Election year	-0.210	-0.089	0.333	0.205	-0.038	0.206	0.204	-0.032	0.125	0.215	-0.098
	(0.619)	(0.583)	(0.470)	(0.349)	(0.254)	(0.496)	(0.349)	(0.270)	(0.416)	(0.337)	(0.202)
Log GDP per capita	-0.387**	0.958	-0.243***	-0.695	-0.260***	-0.147***	-0.590	-0.174**	1.212***	-0.018	1.11***
	(0.157)	(1.187)	(0.085)	(1.070)	(0.100)	(0.052)	(1.069)	(0.081)	(0.221)	(1.043)	(0.288)
Aid in % GDP	-0.030	0.001	-0.052***	-0.025	-0.037***	-0.020	-0.027	-0.012	0.014	-0.030	0.000
	(0.023)	(0.032)	(0.014)	(0.020)	(0.014)	(0.017)	(0.019)	(0.015)	(0.018)	(0.019)	(0.015)
% rural population	-0.040***	-0.062	0.033***	-0.085*	0.022	0.033***	-0.080*	0.022	0.054***	-0.069	0.047**
	(0.015)	(0.051)	(0.010)	(0.046)	(0.014)	(0.011)	(0.045)	(0.017)	(0.012)	(0.044)	(0.022)
% population <15	0.109**	-0.505**	0.106***	-0.195	0.070	0.051	-0.232	0.026	0.084***	-0.256	0.044
	(0.049)	(0.202)	(0.031)	(0.164)	(0.045)	(0.039)	(0.167)	(0.049)	(0.031)	(0.156)	(0.054)
Trade openness			1.907***	0.153	1.797***	1.630***	-0.498	1.504***	0.045***	-0.038***	0.067***
			(0.247)	(0.591)	(0.306)	(0.261)	(0.518)	(0.228)	(0.009)	(0.012)	(0.010)
Constant	6.71***	22.28	-8.019***	26.57*	-4.838*	-4.047	29.05*	-1.355	-13.82***	22.65*	-10.21***
	(2.46)	(18.32)	(2.022)	(15.79)	(2.523)	(2.671)	(16.17)	(2.798)	(1.800)	(11.82)	(3.104)
N	264	264	174	174	174	173	173	173	175	175	177
countries	30	30	29	29	29	29	29	29	29	29	29
R^2	0.11	0.35	0.38	0.67	0.55	0.37	0.73	0.52	0.38	0.75	0.54
Within R^2		0.06		0.08			0.08			0.14	

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP. PCSE for OLS and PW estimates, normal SE for FE. Within country R^2-values for FE models from Stata's xtreg command, adjusted R^2 from areg. Standard R^2 for OLS and PW.

Table A.3: Additional output for Ansell (2008)

	REPLICATION					EXTENSION					
	1960-2000, No dummies		1960-2012			Trade (PPP and log), unadjusted GDP			Trade (PPP, no log), GDP (PPP)		
	OLS	PW	OLS	FE	PW	OLS	FE	PW	OLS	FE	PW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Lagged DV	0.914***		0.943***	0.811***		0.894***	0.746***		0.903***	0.746***	
	(0.020)		(0.017)	(0.024)		(0.025)	(0.057)		(0.025)	(0.022)	
Polity IV score	0.005	0.014*	0.003	0.007	0.037***	0.006	-0.001	0.043***	0.008	-0.003	0.046***
	(0.003)	(0.008)	(0.003)	(0.005)	(0.007)	(0.006)	(0.008)	(0.012)	(0.006)	(0.007)	(0.014)
Log trade openness	0.034	0.625***	-0.015	0.038	0.358***	0.138**	0.058	1.068***	0.003**	-0.005*	0.022***
	(0.044)	(0.125)	(0.041)	(0.082)	(0.124)	(0.061)	(0.136)	(0.236)	(0.001)	(0.002)	(0.007)
% population <15	0.006**	0.020**	0.002	0.000	-0.014	0.003	-0.006	-0.006	-0.001	-0.012	-0.014
	(0.003)	(0.010)	(0.003)	(0.007)	(0.012)	(0.005)	(0.014)	(0.015)	(0.004)	(0.015)	(0.013)
Log GDP	0.183	-0.605	0.162	0.851**	-0.119	0.296	1.225*	-0.668	0.070	0.528	-0.544
	(0.180)	(0.509)	(0.158)	(0.360)	(0.531)	(0.298)	(0.656)	(0.732)	(0.166)	(0.395)	(0.436)
(Log GDP)²	-0.003	0.020*	-0.003	-0.017**	0.006	-0.006	-0.023*	0.012	-0.004	-0.024	0.023
	(0.003)	(0.010)	(0.003)	(0.008)	(0.011)	(0.006)	(0.013)	(0.015)	(0.007)	(0.016)	(0.018)
Log population	-0.055**	-0.397***	-0.020	-0.259	-0.301***	-0.003	0.272	0.113	0.016	-0.238	-0.033
	(0.023)	(0.058)	(0.022)	(0.171)	(0.065)	(0.036)	(0.454)	(0.121)	(0.036)	(0.396)	(0.191)
Government exp.	0.006	0.045***	0.007*	0.015**	0.036***	0.002	0.028***	0.028**	0.002	0.025***	0.026**
	(0.006)	(0.012)	(0.004)	(0.006)	(0.010)	(0.007)	(0.011)	(0.014)	(0.007)	(0.009)	(0.013)
Year	-0.003	-0.006	0.001	0.005	-0.003	-0.002	-0.008	-0.029***	-0.002	0.011	-0.024**
	(0.002)	(0.007)	(0.002)	(0.005)	(0.006)	(0.004)	(0.011)	(0.011)	(0.004)	(0.011)	(0.011)
Constant	4.533	22.28	-3.190	-15.07	13.74	0.526	-4.707	66.75***	3.103	-19.89	56.63***
	(5.540)	(15.40)	(4.996)	(11.18)	(13.49)	(10.01)	(20.01)	(23.12)	(7.326)	(18.20)	(21.36)
N	1161	1578	2016	2016	2577	900	900	1205	889	889	1192
Countries	126	140	141	141	148	125	125	137	125	125	137
R^2	0.90	0.50	0.91	0.92	0.49	0.90	0.91	0.55	0.90	0.91	0.56
Within R^2				0.71			0.63			0.62	

Coefficients and SE in (.), significance levels: 0.1*, 0.05**, 0.01*** Dependent variable: total public spending on education as a percentage of GDP. OLS and PW with PCSE, FE with robust SE. OLS and PW with regional dummies, omitted region is North America. Within country R^2-values for FE models from Stata's xtreg command, adjusted R^2 from areg. Standard R^2 for OLS and PW.

B Data and variables

Table B.1: Comparison of the two (non-logarithmized) trade openness measures

									Avelino et al. (2005)			
	Trade Openness PPP				Trade Openness				Trade Openness PPP		Trade Openness	
Country	Mean	Std. dev.	Min	Max	Mean	Std. dev.	min	max	Min	Max	Min	Max
Argentina	9.586	4.95	3.145	17.80	16.95	3.481	11.54	23.32	4.8	16.3	11.6	23.3
Bolivia	13.42	2.123	10.76	17.42	48.07	3.828	41.89	58.47	15.4	25.3	41.9	58.5
Brazil	6.678	2.184	3.905	10.45	17.43	2.259	14.39	21.47	6.1	13.0	13.2	22.2
Chile	22.60	5.265	14.04	30.46	54.84	6.502	40.62	66.04	22.7	47.4	40.6	67.1
Colombia	8.841	3.181	5.860	14.73	31.83	4.288	23.67	36.15	6.6	13.5	23.7	37.2
Costa Rica	27.57	10.51	16.43	49.52	73.04	13.94	53.98	97.67	19.0	44.0	61.8	97.6
Dominican Republic	23.78	10.02	11.26	38.10	67.15	15.84	32.71	83.18	14.4	35.6	32.7	78.0
Guatemala	13.57	3.654	8.915	20.77	39.14	6.744	24.93	47.11	10.4	21.0	24.9	47.1
Honduras	20.19	4.596	15.34	30.56	73.48	17.61	48.79	100.5	17.4	33.5	47.9	100.5
Mexico	12.85	7.281	5.235	27.93	39.12	14.12	23.34	63.51	10.2	39.9	23.3	63.5
Nicaragua	17.31	6.201	9.302	28.85	56.61	17.70	25.53	97.73	9.6	23.5	25.5	119.2
Panama	62.23	14.23	44.14	84.29	161.44	26.56	125.21	198.8	37.8	66.4	63.0	99.1
Peru	8.855	3.159	4.397	14.03	32.34	5.311	23.69	41.82	5.8	17.6	23.7	41.8
El Salvador	14.67	5.909	8.449	25.68	53.31	7.933	36.93	67.41	9.5	25.4	36.9	67.4
Uruguay	19.17	6.108	12.23	29.84	39.92	4.900	31.62	49.29	17.8	32.6	31.6	49.3

A selection of Latin American countries and their values of trade openness (PPP) and unadjusted trade openness shows that the importance of trade is usually overstated when PPP are not considered. In addition, the standard deviation for the conventional measure is higher in most cases. The Latin American sample has been chosen for a comparison with Avelino et al. (2005), who propose the PPP-adjusted measure. In most countries, the results are comparable.

Table B.2: Summary statistics for Mulligan et al. (2004)

VARIABLE	N	MEAN	STD. DEV.	MIN	MAX
Education spending	116	3.797	1.594	1.022	10.22
Democracy index	138	0.322	0.388	0	1
Communist dummy	138	0.167	0.374	0	1
British legal origin	138	0.312	0.465	0	1
% population >65	138	5.209	3.484	1.643	14.93
Log(population)/10	138	0.882	0.161	0.537	1.372
Log real GDP p.c.	129	7.721	1.039	5.662	9.79
Log trade	133	3.952	0.575	2.433	5.735
Log trade PPP	127	3.013	0.703	1.458	4.971

Table B.3: Summary statistics for Stasavage (2005), 1980-1996

VARIABLE	N	MEAN	OVERALL STD. DEV.	WITHIN STD. DEV.	MIN	MAX
Education spending	650	4.357	2.184	1.398	0.580	16.06
Multiparty competition	2491	0.099	0.298	0.274	0	1
Election year	1661	0.105	0.307	0.300	0	1
Log GDP p.c.	661	10.85	2.080	0.184	5.654	14.41
Log GDP p.c. PPP	1273	7.202	0.981	0.256	5.016	10.35
Aid in % GDP	670	15.07	14.20	8.379	0.001	94.44
% rural population	2444	72.72	14.87	9.032	13.85	97.96
% population <15	2392	43.88	3.627	2.529	21.40	50.04
Log trade	2051	4.086	0.534	0.303	2.310	5.618
Log trade PPP	1242	3.167	0.646	0.350	1.509	5.348
Trade PPP	1246	30.05	25.95	12.17	2.567	210.2

Table B.4: Summary statistics for Ansell 1960-2000

VARIABLE	N	MEAN	OVERALL STD. DEV.	WITHIN STD. DEV.	MIN	MAX
Education spending	1870	4.318	2.112	1.274	0	44.33
Polity IV	5311	-0.338	7.653	4.011	-10	10
Log Trade	5382	4.082	0.661	0.281	0.062	6.021
Log Trade PPP	3038	3.244	0.733	0.292	0.943	5.586
Trade PPP	3038	33.66	28.79	11.18	2.567	266.7
Population >15	7461	36.82	9.427	3.498	14.32	53.03
Log GDP	5991	22.22	2.440	0.958	15.99	29.92
Log GDP PPP	6087	8.883	2.404	0.938	2.404	16.11
Log population	8407	14.67	2.385	0.286	8.361	20.99
Government Exp.	1741	12.33	5.710	2.968	1.836	64.28

Table B.5: Data sources and transformations for Mulligan et al. (2004)

DESCRIPTION	NAME IN SOURCE/TRANSFORMATION	SECONDARY SOURCE	PRIMARY SOURCE
Public education spending	Average (se_xpd_totl_gd_zs)	Mulligan et al. dataset	UNESCO
Democracy index	Average of Polity IV democracy, 1960-1990	Mulligan et al. dataset	PolityIV
Communist	Dummy	Mulligan et al. dataset	Kornai
British legal origin	Dummy	Mulligan et al. dataset	World Bank Global Development Network Growth Database
Percentage of population aged 65+	Average (sp_pop_65up_to_zs)	Mulligan et al. dataset	ILO (International Labor Organization)
Population	Log(average population)/10	Mulligan et al. dataset	ILO
GDP p.c.	Log(average GDP per capita), 1960-89	Mulligan et al. dataset	PWT (Penn World Tables)
Share of value added from agriculture	Average (nv_agr_totl_zs)	Mulligan et al. dataset	WDI (World Bank Development Indicators)
GDP p.c. PPP	Log(average of cgdp)		PWT
Trade openness PPP	Log(average[(ne_exp_gnfs_cd+ ne_imp_gnfs_cd)/ ny_gdp_pcap_pp_cd]*100)		WDI

Table B.6: Data sources and transformations for Stasavage (2005)

DESCRIPTION	NAME IN SOURCE/TRANSFORMATION	SECONDARY SOURCE	PRIMARY SOURCE
Public education spending	se_xpd_totl_gd_zs		WDI (UNESCO)
Multiparty competition	Dummy if sexec2b=6 (Pessimistic Executive Scale)		Institutions Data Set, Africa Research Program, Harvard
Election year	Presidential election dummy (gol_preel)		Quality of Government Institute
GDP p.c.	Log (ny_gdp_pcap_kn)		WDI
Aid in % GDP	Dt_oda_alld_gd_zs		WDI
% population rural	Sp_rur_totl_zs		WDI
% population <15	Sp_pop_0014_to_zs		WDI
Trade openness	Log (ne_trd_gnfs_zs)		WDI
GDP p.c. PPP	Log(wdi_gdpc)	Quality of Government Institute	WDI
Trade openness PPP	Log([(ne_exp_gnfs_cd+ ne_imp_gnfs_cd)/ ny_gdp_pcap_pp_cd]*100)		WDI

Table B.7: Data sources and transformations for Ansell (2008)

DESCRIPTION	NAME IN SOURCE/TRANSFORMATION	PRIMARY SOURCE
Public Education Spending	se_xpd_totl_gd_zs	WDI (UNESCO)
Polity IV score	Polity2	PolityIV
Trade openness	Log (ne_trd_gnfs_zs)	WDI
% population < 15	Sp_pop_0014_to_zs	WDI
GDP	Log(ny_gdp_mktp_cd)	WDI
Log population	Log(sp_pop_totl)	WDI
Government expenditures	ne_con_govt_zs - se_xpd_totl_gd_zs	WDI
Region dummies	Region1 = East Asia and Pacific, Region2 = Europe and Central Asia, Region3 = Latin America, Region4 = MENA Region5 = North America (omitted in all models) Region6 = South Asia Region7 = Sub-Saharan Africa	WDI
GDP PPP	Log(tcgdp)	PWT
Trade openness PPP	Log([(ne_exp_gnfs_cd+ ne_imp_gnfs_cd)/ ny_gdp_pcap_pp_cd]*100)	WDI

C Literature overview

Table C.1: Overview of empirical models explaining public spending (on education)

Article	Data	Estimation	Dependent variable	Lagged DV	Main independent variables	Interaction effects	Results
Lott (1999)	TSCS, 99 countries, 1985-1992	FE	Current real public school expenditures per capita	no	Totalitarianism index by Freedom House	Totalitarianism*GDP/Population	Totalitarianism (+)***, interaction (-) sometimes significant
Kaufman and Segura-Ubiergo (2001)	TSCS, 14 Latin American countries, 1973-1997	Time and country fixed effects, OLS with PCSE	Public welfare spending and health+education expenditures (in per capita 1995 dollars, % of GDP, % of central government spending)	yes	Polity, trade and capital mobility; all in levels and first differences	For welfare spending: trade*capital mobility	Welfare spending: Polity (-), trade (-)**, capital mobility mostly insignificant, interaction effect (-)***; Health+education expenditures: Polity mostly (+)**, trade (-), capital mixed results
Rudra and Haggard (2001)	TSCS, 57 less developed countries, 1972-1997	PW with PCSE	Social security and welfare spending, education spending (% of total government spending)	no	Trade, capital flows, Polity, potential labor power	no	Social security and welfare: Trade(-), capital flows (+)*, Polity (+)*** Education spending: Trade(-)**, capital flows (+), Polity (+)
Rudra (2002)	TSCS, 53 less developed countries, 1972-1995	FE	Welfare spending (% of GDP, % of total government spending, per capita)	yes	Trade, capital flows, democracy, potential labor power (PLP)	Trade*PLP, capflows*PLP	Trade mixed, capital flows (+)*, democracy (+)**, PLP (+)** Interaction effects (-)*
Bueno de Mesquita et al. (2003)	TSCS, N>3000	FE	Education Expenditures	no	Size of the winning coalition, size of the selectorate	no	Winning coalition (+)***, selectorate (+)***
Plümper and Martin (2003)	Cross-sectional, 83 countries	OLS	Government spending (% of GDP)	no	Polity, Polity squared, institutional openness	no	Polity (-)**, Polity squared (+)**, institutional openness (+)

Sign of the coefficient in (.), significance levels: 0.1*, 0.05**, 0.01***.

Article	Data	Estimation	Dependent variable	Lagged DV	Main independent variables	Interaction effects	Results
Brown and Hunter (2004)	TSCE, 17 Latin American countries, 1980-1997	PW with PCSE	Per capita education spending	no	Polity, health and social security spending, trade	no	Polity (+)**, health and social security spending (+)**, trade (-)
Hausken et al (2004)	Cross-sectional, 83 countries	OLS	Government spending (% of GDP)	no	Polity, Polity squared, institutional openness to trade	no	Polity (-)***, Polity squared (+)***, institutional openness to trade (-)**
Mulligan et al. (2004)	Cross-sectional, 110 countries, 1960-1990 averages	OLS	Education spending (% of GDP)	no	Polity, share of value added from agriculture, communist	no	Polity (+), agriculture (-)***, communist (+)***
Avelino et al. (2005)	TSCS, 19 Latin American countries, 1980-1999	OLS with PCSE, PW	Social spending (% of GDP), education spending (% of GDP)	yes	Democracy dummy based on Alvarez et al. (1996), trade openness (in PPP), financial openness (Quinn, 1997)	For social spending: democracy*trade, democracy*financial openness	Social spending: Democracy (+)***, trade (+)**, financial openness not significant, democracy*trade (-), democracy*financial openness (-)** Education spending: Democracy (+)***, trade (+)***, financial openness (-)
Stasavage (2005)	TSCS, 44 African countries, 1980-1996	OLS with PCSE, FE	Total government spending on education (% of GDP, % of total spending)	no	Multiparty competition, election year, aid	no	Education spending in % of GDP: Multiparty (+)**, election year (mixed), aid (-)* Education spending in % of total spending: Multiparty (+)***, election year (-), aid (-)**
Ansell (2008)	TSCS, 113 countries, 1960-2000	OLS with PCSE, FE	Public expenditure on education (% of GDP)	yes	Polity, trade openness, Hiscox/Kastner openness	no	Polity (+)**, trade openness (+)**, Hiscox/Kastner (-)***

Sign of the coefficient in (.), significance levels: 0.1*, 0.05**, 0.01***.