WP/13/162

IMF Working Paper

Can a Government Enhance Long-Run Growth by Changing the Composition of Public Expenditure?

Santiago Acosta-Ormaechea and Atsuyoshi Morozumi

INTERNATIONAL MONETARY FUND

 WP/13/162

IMF Working Paper

Fiscal Affairs Department

Can a Government Enhance Long-Run Growth by Changing the Composition of Public Expenditure?

Prepared by Santiago Acosta-Ormaechea and Atsuyoshi Morozumi[1]

Authorized for distribution by Bernardin Akitoby

July 2013

Abstract

This paper studies the effects of public expenditure reallocations on long-run growth. To do this, we assemble a new dataset based on the IMF's GFS yearbook for the period 1970-2010 and 56 countries (14 low-, 16 medium-, and 26 high-income countries). Using dynamic panel GMM estimators, we find that a reallocation involving a rise in education spending has a positive and statistically robust effect on growth, when the compensating factor remains unspecified or when this is associated with an offsetting reduction in social protection spending. We also find that public capital spending relative to current spending appears to be associated with higher growth, yet results are non-robust in this latter case.

JEL Classification Numbers:H50; O40; O57

Keywords: Expenditure composition; panel-data analysis; economic growth.

Author's E-Mail Address:SAcostaOrmaechea@imf.org; Atsuyoshi.Morozumi@nottingham.ac.uk

[1]We thank Bernardin Akitoby, Alberto Alesina, Mark De Broeck, Benedict Clements, Markus Eberhardt, Sourafel Girma, Francesco Grigoli, Sanjeev Gupta, Kentaro Katayama, Richard Kneller, Kevin Lee, Mico Loretan, Paolo Mauro and Harald Uhlig for their comments, and Andrea Gamba and Jiae Yoo for their help with the dataset. The conventional disclaimer applies.

Contents Page

Figures

I. INTRODUCTION

Can a government promote long-run growth by changing the composition of public expenditure? This question is relevant to many economies around the world for various reasons. For example, if a government faces high levels of indebtedness and decides to undertake fiscal austerity measures to reduce the debt burden, increasing public spending might be unfeasible for several years.[1] However, a government may still attempt to foster growth by changing the composition of its spending envelope. Amid current demographic trends of population aging, governments may also find it inevitable to increase health and social protection spending over the next several years.[2] Since at least part of the increasing bill may need to be covered by a reduction in spending in other components, policymakers will need to decide which type of spending to reduce while trying to preserve growth. One relevant historical example of spending reallocations is found in western countries after the end of the Cold War. Facing the fall in defense-related outlays, policymakers then needed to consider how to reallocate this so-called 'peace dividend' to other components such as economic infrastructure or social protection to cope with the economic and social challenges of that time.

Despite its apparent importance, the effects of public expenditure composition on growth have been rarely investigated, apart from a few notable exceptions. These include theoretical works such as Barro (1990), who shows that when a government increases 'utility-enhancing' public consumption while reducing 'production-enhancing' public spending, growth rates fall regardless of the level of total spending. While theoretical models shed valuable light on the way compositional changes exert their effects on growth, their implications are often not specific enough for active policymaking, since the contents of their classifications such as utility-enhancing expenditure can be debatable.[3] As for the empirical work, a number of papers have specifically studied how compositional changes in public spending affects economic growth. However, because they often do not clarify which components are used as compensating factors (to keep the level of total spending unchanged), their policy implications may still not be practical enough.

[1]For example, the 2012 UK's 'autumn statement' (the annual statement made by HM Treasury on economic forecasts) indicates that the ongoing fiscal austerity program would continue through 2018.

[2]See Clements, Coady, and Gupta (2012) and Clements and others (2012) for discussions on expected trends in health and pension spending in advanced and developing countries over the medium and long term.

[3]Another theoretical work that we are aware of on the link between public expenditure composition and growth is Agénor (2010). He shows how a reallocation from 'unproductive' public spending to infrastructure spending helps a country move to a steady state of higher growth.

This paper helps fill in this gap. For this purpose, we first assemble a new dataset based on historical fiscal data reported to the IMF's government finance statistics (GFS) yearbook from 1970 to 2010. The novelty of the dataset is that it directly confronts methodological changes from mid 1990s to early 2000s with the introduction of a new GFS manual (i.e., GFSM2001). These methodological changes include differences in the way in which components are categorized in the economic and functional classifications of expenditures.[4] Although we had to undertake important assumptions to bridge the two methodologies, for instance mixing cash and accrual basis concepts in certain cases, our dataset still offers *comparable* fiscal data across periods under these methodologies. This dataset, being an unbalanced panel, covers in total 56 countries (14 low-, 16 medium- and 26 high-income countries) during the period 1970–2010 at the central government level.[5]

We then attempt to capture the effects of government spending reallocation on growth both in terms of the economic and functional classifications of expenditure. To do this, we use the Generalized Method of Moments (GMM) dynamic panel estimators developed by Holtz-Eakin, Newey, and Rosen (1990) and Arellano and Bond (1991).[6] These estimators, in addition to being flexible to accommodate unbalanced panels while also handling the bias from unobserved country-specific effects, have the advantage of dealing with potential endogeneity problems. This property is important in our context. For instance, even if we observe a positive correlation between the share of education spending in total spending and economic growth, it does not necessarily imply that a higher share devoted to education causes higher growth. The causality could be reverse, making the education share endogenous.

Our results are as follows. It is *in general* difficult to find statistically significant and robust associations of compositional changes in government expenditure with growth. For instance, although an increase in capital spending financed by a fall in current spending has a positive effect on growth, the significance of this result depends on the particular specification of the

[4]The former is based on the economic characteristics of expenditure (e.g., wages, net acquisition of non-financial assets, etc.), while the latter is based on the function to which expenditure is allocated (e.g., defense, education, health, etc.).

[5]To classify countries into those three income groups, we take the following procedure. First, for each of the 41 years (1970-2010) we sort countries according to their GDP per capita level (PPP prices) into four equal-sized groups: the highest 25th percentile, between the 25th and 50th percentile, between the 50th and 75th percentile and from the 75th percentile onwards. Next, we count the number of times each country appears in those four groups during those years. Then, we categorize countries which appear above the top 25th percentile most frequently as high-income countries (HICs). Likewise, countries which appear between the 25th and 50th percentiles most frequently are grouped as middle-income countries (MICs). The rest of the countries are categorized as low-income countries (LICs). This way, our income groups reflect development levels during the whole period available. The full list of countries is given in Appendix A.

[6]Specifically, to tackle the finite sample biases caused by the use of 'difference' GMM estimators, we use the 'system' GMM approach suggested by Arellano and Bover (1995) and Blundell and Bond (1997).

model. This lack of robustness appears to hold even when a government reallocates its spending classified by its function. Specifically, we find that none of the possible public spending reallocations between defense, economic infrastructure, health, and social protection has a robust effect on long-run growth.

However, when a compositional change involves a rise in spending on *education*, a robust association with long-run growth seems to emerge. This association is particularly robust when leaving the compensating factor unspecified or when it is associated with an offsetting reduction in social protection spending. This result particularly suggest that education could have been a relatively more efficient outlay to foster growth over the last 40 years through human capital accumulation. We show that this finding is reasonably robust to various checks such as the use of lagged fiscal variables (under the assumption that the reallocation effects on growth emerge with a lag), the addition of various widely-used control variables, and the use of a particular subset of countries. Moreover, although we obtain this result using consolidated central government level data, we show that it is likely to hold even at the consolidated general government level.

In the related literature, a number of papers have examined the role of public education expenditures on economic growth. This interest is natural because at least since Lucas (1988) the important role of human capital accumulation on growth has been widely acknowledged. Then, to the extent that public education spending promotes the accumulation of human capital, one would expect that it also enhances growth. However, empirical results so far are not necessarily consistent with this common intuition. For instance, focusing on the level effects of this spending (thus causing an increase in total spending), Easterly and Rebelo (1993) show that education spending is not always growth-enhancing, pointing out that the promoting effects become statistically insignificant in some specifications. Likewise, Barro (2004), in his comprehensive study of determinants of growth, also finds that an increase in public education spending does not have a statistically significant effect on growth.[7] Facing these inconclusive empirical results, Blankenau and Simpson (2004) theoretically show that the effects of public education spending on growth may be non-monotonic.[8] Particularly, their model suggests that while public education spending has a positive impact on growth by directly promoting human capital accumulation, there are also potentially negative general equilibrium effects depending on the form in which this spending is financed. Our work adds to this literature by empirically showing that public education expenditure is robustly associ-

[7] Bose, Haque, and Osborn (2007), however, show that this spending is robustly associated with faster growth in developing countries.

[8] Other theoretical works on public education spending and growth include Kaganovich and Zilcha (1999).

ated with higher growth when the source of its financing is given by an offsetting reduction in social protection spending.

Another functional spending component often investigated in relation to economic growth is public infrastructure spending. The literature, however, again offers mixed empirical findings. For instance, while Aschauer (1989) suggests that it has a significant positive effect on growth, Holtz-Eakin (1994) indicates otherwise, showing that the positive effects completely vanish when region-specific effects are controlled for.[9] This controversy regarding the effectiveness of public infrastructure may also appear to be counter-intuitive because this spending, causing an increase in public capital and thus enhancing private firms' productivity (given their private inputs), is expected to promote growth. However, more recent works such as Pritchett (2000) and Dabla-Norris and others (2012) emphasize that not all actual accounting cost of public investment creates economically valuable capital, which can be exemplified through the expression of public investment turning into 'incomplete roads leading to nowhere'. In line with this, Agénor (2010) theoretically shows that only when the degree of public investment efficiency is high, a spending reallocation into infrastructure from 'unproductive' spending can be growth enhancing. His result may suggest that the presence of inefficient public infrastructure spending could explain the lack of robustness in the growth-promoting effects of spending reallocations toward public infrastructure found in our paper.

As noted in the literature, defense and social spending (i.e., health and social protection) may also play an important role in growth. For instance, Barro (2004) points out that defense spending can promote investment and thereby growth by enhancing entrepreneurs' property rights. Similarly, Agénor (2010) suggests that public health can influence growth by affecting labor productivity and individuals' discount factors. However, within social spending the social protection component has often been assumed not to be productive (e.g., Kneller, Bleaney, and Gemmell (1999)), which could reflect the primarily re-distributive nature of this type of outlay. Our results contribute also to this strand of economic research by suggesting that a rise in the spending devoted to either defense, health or social protection offset by a fall in other components does not appear to have a positive and statistically significant association with growth. The rest of the paper is structured as follows. Section II describes the new dataset. Section III conducts the regression analysis, and Section IV presents some concluding remarks.

[9] Glomm and Ravikumar (1997) expertly summarizes the early literature on infrastructure expenditure and growth.

II. The Dataset

A. Construction of the Dataset

1. Merging GFSM1986 with GFSM2001

To study the compositional effects of changes in public expenditure on growth, we first assemble a new dataset using the IMF's GFS yearbook. To explain the novelty of the dataset, notice that in principle this database contains all detailed fiscal data covering a wide set of countries from 1970 onwards needed for our empirical analysis. However, a major methodological change with the introduction of GFSM2001 (from mid 1990s to early 2000s) makes the series prior to the change somewhat incomparable with those after the change. Facing this issue, yet hoping to have a long dataset covering the whole 40 years, we bridge these methodological changes to construct *comparable* data series. In what follows, we briefly describe two of the major changes that took place with the introduction of GFSM2001 and how we handled them.

First, expenditures are classified differently in GFSM2001 relative to GFSM1986 (see Wickens (2002) for details). For instance, in terms of the economic classification, the important change is that although both GFSM2001 and GFSM1986 can be roughly divided into 'current' and 'capital' expenditures, the exact definition of current and capital spending differs. That is, the capital expenditure concept under GFSM2001 (denoted as 'net acquisition of non-financial assets') adopts a net concept in the sense that the government revenue from the sales of fixed capital assets are taken into account. In contrast, capital expenditure under GFSM1986 adopts a gross concept, in which case the revenue from capital sales is not deducted. Besides, capital transfers, which were part of capital expenditure under GFSM1986, are part of the current expenditure concept under GFSM2001 (denoted as 'expense'). Regarding the functional classification, while GFSM2001 divides expenditures into 10 functional categories, GFSM1986 divides them instead into 14 categories. Second, the form in which governments report statistics have also changed. Under GFSM1986 reporting is only on a cash basis, whereas under GFSM2001 this is mainly on accrual basis. To explain the difference, under accrual basis flows are recorded at the time when a transaction accrues, independently of the flow of cash. Instead, under cash basis, transactions are recorded when cash effectively flows. Thus, the two recording bases inevitably coexist in our series. What is more, for some countries data for the different subcategories are reported in different accounting bases even within a given year under GFSM2001.

Facing these challenges, we first retrieved all historical expenditure data available for all countries that have reported data to the IMF's GFS yearbook from 1970 to 2010. Regarding the different categorizations, we followed Wickens (2002) and converted all expenditure items under GFSM1986 into the concepts defined by GFSM2001, so that in the unified series capital expenditure is defined as a net concept while the functional components are divided into 10 categories. As for the accounting issue, given that our focus is on the composition of expenditures (i.e., the expenditure shares among subcomponents), the difference in the timing of recording between cash and accrual bases appears to be less of a problem as long as all expenditure subcomponents are reported on the same accounting basis within a given year. We thus ensure that we take data under the GFSM2001 classification only when all the expenditure components of our interest (clarified below) are reported on the same basis within the same year. Further, whenever data are available for all expenditure components on the more economically-relevant accrual basis, we use that data.[10]

2. Subcomponents of Expenditure: Economic and Functional Classifications

While our dataset follows the categorization under GFSM2001, it does not attempt to cover all the detailed subcomponents provided in the manual. As for the economic classification, our main interest is in the distinction between expense and net acquisition of non-financial assets. Additionally, given the frequent interest in the 'wages' subcomponent of the former category in the literature (e.g., Gupta and others (2005)), we separate compensation of employees from the rest of expense.[11] The dataset thus have the three items in the economic classification: compensation of employees (as a proxy of wages), the rest of expense, and net acquisition of non-financial assets. Turning to the functional classification, while total expenditure is grouped into 10 categories we only cover 8 of them, leaving out public order and safety, and environmental protection.[12] The reason for this selection is due to the limited availability of these 2 subcomponents throughout our sample period of 1970 to 2010. Particularly, the latter category is only available under GFSM2001.

With the dataset in hand, we calculate the ratios of the different public spending components to total expenditure as the key regressors. However, deflators among different spending com-

[10]However, if all the expenditure components are available only on cash basis, we use that data instead.

[11]Rest of expense consists of: use of goods and services; consumption of fixed capital; interest; subsidies; grants; social benefits; and other expense.

[12]The 10 functional categories defined in GFSM2001 are: 1) general public services; 2) defense; 3) public order and safety; 4) economic affairs; 5) environmental protection; 6) housing and community amenities; 7) health; 8) recreation, culture, and religion; 9) education; and 10) social protection.

ponents may differ, implying that spending shares can be affected by both changes in relative prices and changes in relative quantities. Because of data availability issues, we will not decompose an observed change in the expenditure share into changes in relative prices or relative quantities. Likewise, issues related to the quality of spending will be left out. Since changes in the quality of public goods can affect the quantities required for a given provision of a public service, quality aspects can also be part of the underlying causes of variation in the expenditure shares. However, the limited availability of proxies for the quality of expenditure by function forces us to abstract from this potentially important source of relative variation among the different expenditure components.

3. Government Level

Another important element to clarify about the dataset is the institutional coverage level of the government. While under GFSM1986 countries report data at most at the consolidated central government (CG) level, under GFSM2001 they also provide data for the consolidated general government (GG) level. Although some countries provide fiscal data also at lower government levels (i.e., state and/or local governments) under the former, the availability of such data is severely limited. We thus use the consolidated CG level for our main analysis.[13]

However, when the degree of fiscal decentralization (measured by the share of spending at the CG level relative to that of the GG level) differs across subcomponents, the CG level data may not accurately capture the share of those subcomponents at a national level. This actually appears to be the case, because although some subcomponents such as defense tend to be centralized in most countries, others such as health and education tend to be more decentralized.[14] Moreover, when the trend of fiscal decentralization differs across those components over time, using the CG level data can be more problematic in a panel data analysis.[15]

[13]The consolidated CG level can be further divided based on whether the institutional unit is financed by the legislative budget or by extrabudgetary sources. The CG unit based on the legislative budget is called budgetary CG. Some works on fiscal policy such as Easterly and Rebelo (1993) use the budgetary CG data along with consolidated CG in an attempt to increase the number of observations. However, we often find non-trivial discrepancies between consolidated CG and budgetary CG data in some expenditure subcomponents such as social protection. We thus rely only on consolidated CG data in our dataset.

[14]For instance, Wyss and Lorenz (2000) discusses the vastly decentralized nature of the health sector in Switzerland.

[15]Dziobek, Mangas, and Kufa (2011) report that in some countries (e.g., Spain and Switzerland), the level of fiscal decentralization in total expenditure has been unstable over time. Although their report does not cover the different expenditure subcomponents (except ‘compensation of employees’ in the economic classification), those unstable trends may be caused by the decentralization of some particular subcomponents such as health and education.

Recognizing these limitations in the use of CG data, we will later check if our main results are robust when considering instead the GG level. Specifically, our robustness check will be based on a limited sample in which both CG-level and GG-level data coexist under GFSM2001. Using this restricted sample, we will present evidence suggesting that results obtained with the CG-level and GG-level data are statistically the same, thus providing certain justification to our analysis.

4. Additional Macro Variables

Our dataset also contains a few macroeconomic variables including GDP and exchange rates. They have been obtained mainly from either the World Economic Outlook (WEO) or the International Financial Statistics (IFS) databases of the IMF.[16] Other macro variables are also included to use them as control variables in our regressions. One key control variable used in our reference regressions is the average years of schooling between ages 25 and 64 (as a proxy for human capital accumulation) from the Barro and Lee (2010) dataset. The other controls used in our robustness checks include inflation rates, openness of a country (calculated as the value of imports and exports relative to GDP), population growth, and the terms of trade growth.

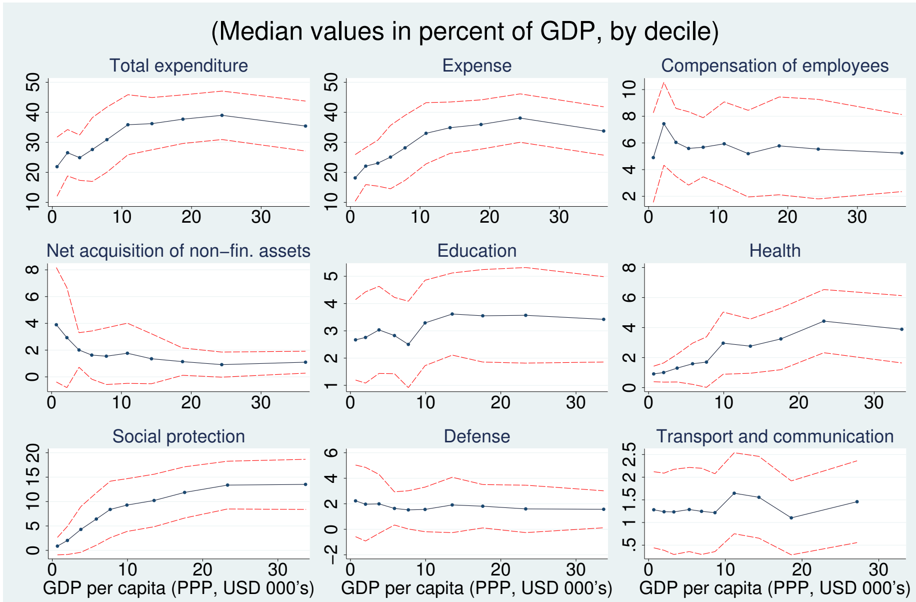
B. Graphical Description

All in all, the assembled dataset is unique and itself can reveal simple yet new facts on government expenditure over the last 40 years for a large set of countries. Therefore, before turning to our formal regression analysis on public expenditure composition and growth, we first describe the dataset from various angles. First, pooling together all countries and ordering the data according to their income level we can examine how total expenditure and the associated subcomponents vary as countries become more developed. To better understand this relation, Figure 1 divides the whole sample into deciles according to the countries' GDP per capita level (PPP prices).[17] Each point in the figure corresponds to the median value for

[16]GDP information is used to create the ratio of total expenditure over GDP, which is required to control for level effects in our regressions. Exchange rates are necessary to convert GFS data reported in national currency into US dollars, since nominal GDP in current prices is taken from WEO in US dollars.

[17]In terms of the economic classification, the figure covers 86 countries which have reported all relevant components (explained above) at least once in the period 1970 to 2010. As for the functional classification, the number of countries (again which have reported all the relevant components at least once) is 102. The transport and communication subcomponent within economic affairs is slightly limited to 101 countries, since this is not always available even when the latter category is reported.

Figure 1. Economic Development and Composition of Expenditure



Note: Central government level.
Source: Authors' calculations based on GFSY database.

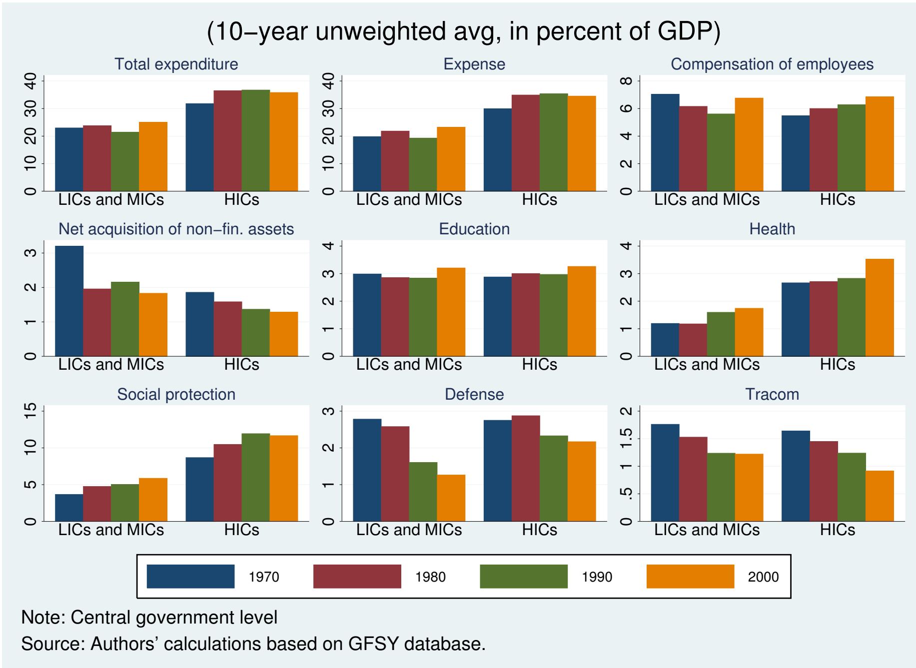
each decile's income level and the associated median value of the expenditure-related component.[18]

The first four panels show all expenditure items associated with the economic classification of expenditure, in which total expenditure is the sum of expense and net acquisition of non-financial assets. It follows that countries increase the overall expenditure envelope (as a share of GDP) as they become richer until the GDP per capita reaches around 20,000 US dollars (PPP prices), in line with the so-called Wagner's law.[19] However, after that, the size of the government flattens out and then slightly decreases, thus showing a non-monotonic relation. Importantly, this behavior of total expenditure is essentially driven by the expense subcomponent. Wages (to be precise, compensation to employees), one of the key subcomponents within the expense category, tends to show a relatively more stable pattern. In fact, the general increasing pattern of expense is rather associated with an upward trend in the social benefits subcomponent (not shown). Finally, those outlays associated with the net acquisition of non-financial assets decrease noticeably as countries become richer. Turning to the classification of expenditure by function, note that those outlays associated with health and social

[18] Easterly and Rebelo (1993) undertake a similar exercise.

[19] Wagner's law states that the size of the government rises as the associated country's income level increases. See Ram (1990) for details.

Figure 2. Long-Run Trends in Expenditure

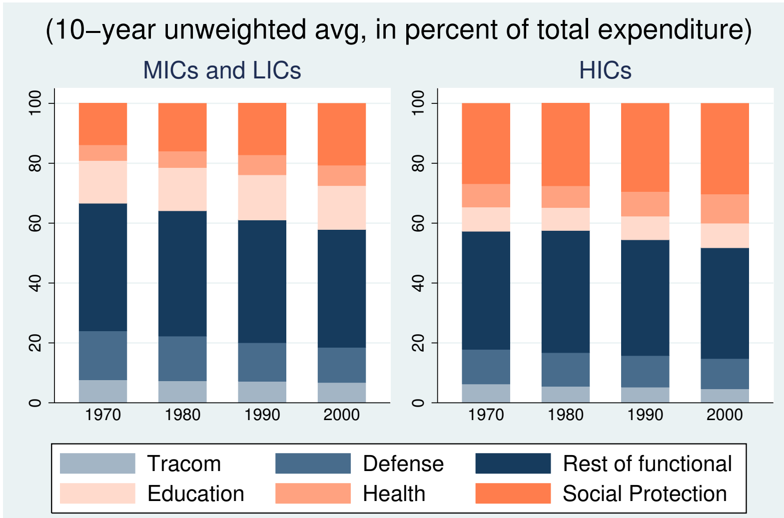


protection—the combination of which is often referred to as public social spending—increase in the level of development of the economy. Observe, however, that spending on education and transport and communication (a proxy for economic infrastructure) do not appear to increase when countries increase their income level.

Figure 2 describes the evolution of the different expenditure components considering an unweighted average of all items as percent of GDP by decades and for two broad country groups: low and medium-income countries combined (LICs and MICs) and high-income countries (HICs).[20] Consistent with the previous figure, both total expenditure and expense are higher for countries with higher income levels. There is, moreover, an upward trend over time until the 1990s and a slight fall during the 2000s in the case of HICs. In contrast, the net acquisition of non-financial assets exhibits a downward trend over the 40-years period in both groups. Particularly in the case of HICs, a similar pattern is present when observing the transport and communication subcomponent. This makes sense because a large portion of transport and communication is devoted to the acquisition of physical capital, which as noted

[20]The figures only contains countries which have reported all the economic (or functional) components for all 4 decades. Therefore, the number of countries featured in the figure is limited. For the economic classification, the figure contains unweighted averages for 19 HICs, and 7 MICs and LICs. Regarding the functional classification, there are 14 HICs, and 6 MICs and LICs (for the reason clarified above, the availability of transport and communication data is limited to 7 HICs, and 7 MICs and LICs.)

Figure 3. Composition of Expenditure: Functional Classification



Note: Central government level.
Source: Authors' calculations based on GFSY database.

above had a downward trend. Turning to education, despite certain fluctuations across time in both groups, spending on this component relative to GDP has been relatively similar and stable across groups and time. However, when focusing on those categories directly related to social spending, it follows that for both health and social protection the spending envelope has generally increased over time, though the level is significantly higher for HICs. In contrast, defense spending exhibits a clear downward pattern in both groups. Notably, the large fall during the 1990s coincides with the end of the Cold War.

We finally look at the evolution of each component by focusing on the composition of total expenditure, which is the subject of the regression analysis below (Figure 3).[21,22] Regarding the functional classification, we observe a rapid increase in the shares of spending in social protection and health, yet the shares are significantly smaller in the case of LICs and MICs. In both groups, these upward trends are accommodated by the lowering spending trends in the remaining categories including transport and communication and defense. However, in line with the previous figure, the share of education spending has not changed much over the last

[21] In terms of the economic classification (not shown), the number of countries included in the figure is the same as in Figure 2. As for the functional classification, since we only focus on the case where transport and communication data are available, the number of countries are 7 HICs, and 7 MICs and LICs.

[22] Note that 'total' expenditure in case of the functional classification is not exactly total because it does not include 2 of the 10 functional subcomponents (i.e., 'public order and safety', and 'environmental protection'). However, (as indicated below) since the shares of those expenditures are relatively small, the total of the remaining 8 components is almost equal to the actual total expenditure.

40 years in the HICs group. In the case of the economic classification, a notable downward trend in the share of net acquisition of non-financial assets is observed in both groups (not shown).

C. Descriptive Statistics

Table 9 in Appendix B summarizes the descriptive statistics for the fiscal variables relevant for the subsequent regression analysis. As in similar growth regressions, we use 5-year non-overlapping averages to abstract away from the effects of the business cycle, thus leading to a maximum of 8 observations per country (i.e., 1971–75, 1976–1980, . . . , 2006–2010). However, since our dataset is an unbalanced panel, we need to choose how many observations we require to calculate each 5-year average observation. For instance, although allowing just one observation to form a 5-year average maximizes the number of observations, this choice clearly does not handle possible business cycle fluctuations. Meanwhile, requiring full 5 observations severely reduces the sample and thereby potentially useful information could be lost. Therefore, in Table 9 and in the regression analysis presented below, we take the 5-year average if the number of observations is at least 3 within each 5-year period.[23]

III. Regression Analysis

A. Empirical Specification and Methodology

Our empirical specification is motivated by neoclassical growth models such as that of Solow-Swan.[24] The model relates real GDP per capita growth to two kinds of variables: state and control/environmental variables. The former variables give the initial position of the economy, whereas the latter determine the steady-state. As is well known, the first important implication of the model is that when the steady state is controlled for, an equiproportional increase in the state variables reduces growth, thus implying the presence of ‘conditional’ convergence. The second is that an increase in the steady state output level leads to higher growth rates during the (seemingly) long adjustment period towards the steady-state growth rate.[25]

[23]This choice is somewhat ad-hoc, but it turns out that choosing 4 as a threshold critically reduces the number of countries available in the regressions. On the other hand, choosing 2 would leave too much room for the observations to be affected by the business cycles.

[24]Barro (2004) (chapter 12) also studies the empirical determinants of growth based on these models.

[25]The steady state growth rate is determined exogenously in neoclassical growth models.

Based on this second prediction, we examine how changes in the different shares of expenditure components affect growth. Formally, our empirical specification is given by:

$$y_{i,t} - y_{i,t-1} = (\alpha - 1)y_{i,t-1} + \beta x_{i,t-1} + \bar{f}'_{i,t}\phi + \nu_i + \varepsilon_{i,t}. \tag{1}$$

The left hand side (LHS) is the growth rate of output per capita, where $y_{i,t}$ is log of output in country i at time t. Consistent with the descriptive statistics above, t designates one of the 5-year averages. Explanatory variables in the right hand side (RHS) include $y_{i,t-1}$, the initial real GDP per capita and $x_{i,t-1}$, the initial years of schooling as state variables. The former variables are meant to be a proxy for initial physical capital, while the latter variables are used as proxies for human capital accumulation. The RHS also contains a vector of control/environmental variables, $\bar{f}'_{i,t}$. Given that these variables affect the steady state of the economy during the period spanning $t-1$ and t, $\bar{f}'_{i,t}$ is obtained as an average of those variables between these two periods, i.e., $(f'_{i,t} + f'_{i,t-1})/2$. ν_i represents fixed effects (i.e., unobserved country-specific effects). Finally, the RHS also contains time dummies (though not explicitly shown in Eq. (1)).

Highlighting the fiscal variables among $\bar{f}'_{i,t}\phi$, we have

$$\bar{f}'_{i,t}\phi = \delta\bar{e}_{i,t} + \sum_{j=1}^{m}\gamma_j\bar{s}_{i,j,t} + \sum_{j=1}^{k}\eta_j\bar{z}_{i,j,t}. \tag{2}$$

In the RHS, $\bar{e}_{i,t}$, the share of total public expenditure to GDP, is included to control for the level effect of total expenditure. Next, $\bar{s}_{i,j,t}$, represents the share of the different expenditure components in total expenditure. Finally, $\bar{z}_{i,j,t}$ represents the rest of the control/environmental variables. They include the inflation rate, a proxy for trade openness, population growth, and terms of trade growth. These control variables are selected considering their availability in order to preserve the coverage of our dataset to the largest possible extent.

To proceed, however, notice in Eq. (2) that $\sum_{j=1}^{m} s_{i,j,t} = 1$ by construction. Thus, to avoid exact multicollinearity, we need to leave out at least one component, say component 'm'. Doing this yields

$$y_{i,t} - y_{i,t-1} = (\alpha - 1)y_{i,t-1} + \beta x_{i,t-1} + \delta\bar{e}_{i,t} + \gamma_m + \sum_{j=1}^{m-1}\left(\gamma_j - \gamma_m\right)\bar{s}_{i,j,t} + \sum_{j=1}^{k}\eta_j\bar{z}_{i,j,t} + \nu_i + \varepsilon_{i,t}. \tag{3}$$

We can further rewrite this expression to have a dynamic equation in which the lagged dependent variable appears in the RHS. By simply adding $y_{i,t-1}$ to both sides, we obtain:

$$y_{i,t} = \alpha y_{i,t-1} + \beta x_{i,t-1} + \delta \bar{e}_{i,t} + \gamma_m + \sum_{j=1}^{m-1} (\gamma_j - \gamma_m)\, \bar{s}_{i,j,t} + \sum_{j=1}^{k} \eta_j \bar{z}_{i,j,t} + \nu_i + \varepsilon_{i,t}. \qquad (4)$$

Observe that the coefficients on the expenditure components are now interpreted as the effects of a rise in those components on growth when they are compensated by a fall in the factor that is left out. In short, they represent reallocation effects among the different spending components.

We estimate this dynamic panel data model using a GMM approach. There are various reasons for this choice. First, the GMM framework is flexible enough to accommodate our unbalanced panel. Second, it allows us to deal with country fixed effects. Third, it enables us to handle the potential endogeneity of all explanatory variables through the use of internal instruments (i.e., instruments based on lagged values of those variables). This is important because endogeneity issues appear to be non-trivial concerns in our context. In addition to the reverse causality issue mentioned in the introduction, omitted variable problems are also likely to be present. For example, Mauro (1998) finds that corruption reduces spending on education, while Mauro (1995) also shows that corruption reduces growth by lowering investment. Further, while aging societies tend to increase social spending, this demographical change, possibly lowering overall productivity by reducing the fraction of the population in the prime age category of 15–65 (e.g., Barro (2004)), can also affect growth negatively.

However, while the GMM approach yields consistent estimators, the original ‘difference’ GMM estimators developed by Holtz-Eakin, Newey, and Rosen (1990) and Arellano and Bond (1991) may suffer from finite sample biases. These biases arise if the time series are persistent, which in turn let instruments become weak. In fact, Bond, Hoeffler, and Temple (2001) point out that these biases are likely to be large in the context of empirical growth models since output tends to be a largely persistent variable. They thus recommend the alternative ‘system’ GMM estimators developed by Arellano and Bover (1995) and Blundell and Bond (1997), which augments the difference estimator by combining the regression in differences with the regression in levels in a system in which the two equations are separately instrumented.[26] We use this system procedure in what follows.

[26]Revisiting Caselli, Esquivel, and Lefort (1996), who use difference GMM estimators in growth regressions, Bond, Hoeffler, and Temple (2001) show that the use of system GMM can improve on the finite sample biases present in that paper. Other papers using system GMM estimators in growth regressions include Levine, Loayza, and Beck (2000) and Rodrik (2008).

To be specific, the difference part of our system includes the following moment conditions:

$$E\left[y_{i,t-s}(\varepsilon_{i,t}-\varepsilon_{i,t-1})\right]=0$$

and

$$E\left[x_{i,t-s}(\varepsilon_{i,t}-\varepsilon_{i,t-1})\right]=0,$$

where $s \geq 2$ and $t = 3,\ldots,T$. The rest of this part consists of the analogous relations for the other explanatory variables $\bar{e}_{i,t-s}$, $\bar{s}_{i,j,t-s}$, and $\bar{z}_{i,j,t-s}$. Regarding the level part of the system, where the instruments used are the lagged differences of the variables, the moment conditions are:

$$E\left[(y_{i,t-1}-y_{i,t-2})(\nu_i+\varepsilon_{i,t})\right]=0$$

and

$$E\left[(\bar{x}_{i,t-1}-\bar{x}_{i,t-2})(\nu_i+\varepsilon_{i,t})\right]=0,$$

where $t \geq 3$. Again, the rest of the conditions consist of the analogous relations for $\bar{e}_{i,t-1} - \bar{e}_{i,t-2}$, $\bar{s}_{i,j,t-1} - \bar{s}_{i,j,t-2}$, and $\bar{z}_{i,j,t-1} - \bar{z}_{i,j,t-2}$. In what follows, however, to reduce the number of instruments generated in the system, we combine instruments through additions to smaller sets. This can be done by asking the estimator to minimize the magnitude of empirical moments only for each lag length rather than for each lag length and time.[27] We take this measure because as Roodman (2009b) emphasizes, having too many instruments (relative to the number of countries) makes estimation results unreliable.[28]

To ensure the validity of this system approach in our context, we conduct a number of specification tests. The first is the Arellano-Bond test. Its purpose is to examine the hypothesis that the error term is not serially correlated, which is assumed to draw all the orthogonality conditions. The second is the Hansen test, which checks the overall validity of the the various instruments of the system. The third is the difference Hansen test, which examines the validity of the different sets of instruments used in the level part of the system.

We finally comment on our choice of a (system) GMM approach relative to other dynamic panel estimators used in the literature. Particularly, with increasing availability of data covering a large number of time series observations (T) and a large number of countries (N), some recent works on fiscal policy and growth (e.g., Gemmell, Kneller, and Sanz (2011) and

[27]We do this in our estimations using the ‘collapse’ option in Roodman’s ‘xtabond2’ Stata command (Roodman (2009a)).

[28]For those regressions presented in Table 1, however, we consider an alternative measure for reducing the number of instruments, namely, using only one lag (instead of all available lags).

Arnold and others (2011)) use the Mean-Group (MG) and/or Pooled Mean-Group (PMG) estimators developed by Pesaran and Smith (1995) and Pesaran, Shin, and Smith (1999), respectively. These estimators have their own advantages. Most notably, they allow for a simultaneous investigation of long-run equilibrium relations and short-run adjustments processes, in which key parameters are allowed to be heterogeneous.[29] Since there is no particular reason to think that the effects of fiscal policy on long-run growth should be homogeneous, this could be an advantage over our GMM approach, where only the long-run relation is considered and heterogeneity is allowed only in terms of an intercept. However, one potential downside of these alternative approaches is that since 'large T' requires the use of annual data, the effect of business cycles can be more problematic than in our 5-year averaged case. Besides, from a practical viewpoint, the fact that our highly disaggregated fiscal expenditure dataset is unbalanced (and thus many years are missing for several countries, particularly for LICs) does not allow us to use either of these alternative estimators.

B. Results

1. Economic Classification

We now present the results on the reallocation effects for the economic classification. As mentioned, our focus is on the following three items: compensation of employees, the rest of expense, and net acquisition of non-financial assets (for brevity, non-financial assets). Table 1 presents the estimation results of a reallocation between expense and non-financial assets.[30] In column (1), the compensating component is expense. Thus, as Eq (3) indicates, the coefficients of non-financial assets represent the effect of a rise in this spending when compensated by an equal fall in expense. Given that each period spans 5 years, the coefficient of 1.4 means that a rise in the share of non-financial assets by 1 percentage point, offset by an equal fall in expense, increases growth by about 1.4 percentage points over the 5 years period (thus about 0.27 percentage points per annum). Next, a rise in the share of total expenditure to GDP has a negative effect (with statistical significance at the 10 percent level), probably because the corresponding rise in tax revenues (to finance the increase in expenditure) can be distortionary. The coefficient on initial GDP, a proxy for initial physical capital, is negative, being consistent with the presence of conditional convergence. Meanwhile, a rise in the average years of schooling has a positive effect on its own.

[29] In the case of PMG, the heterogeneity is assumed only in the short-run coefficients.

[30] Results based on a more detailed disaggregation are presented in Appendix D, which also divides the sample according to the different development levels of the countries.

Table 1. Expenditure Composition and Growth: Economic Classification

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)
Total exp/GDP	−0.416∗	−0.537∗∗
	(0.250)	(0.248)
Expense/Total Exp		−1.209∗∗∗
		(0.391)
Nonfin. Assets/Total Exp	1.375∗∗∗	
	(0.436)	
Initial GDP p.c.	−0.805	−0.763
	(0.855)	(0.782)
Initial Human Capital	4.600∗∗∗	4.452∗∗∗
	(0.996)	(0.998)
Constant	−20.711	106.409∗∗∗
	(16.789)	(32.142)
Compensating factor	Expense	Nonfinancial assets
Observations	190	190
No. of countries	52	52
No. of instruments	51	51
Arellano-Bond AR(1), p-value	0.01	0.01
Arellano-Bond AR(2), p-value	0.11	0.11
Hansen, p-value	0.55	0.58
Diff Hansen 1, p-value	0.80	0.95
Diff Hansen 2, p-value	1.00	1.00

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, *** $p < 0.05$, * $p < 0.1$.
Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

In terms of specification tests, the Arellano-Bond tests indicate that the error term is not serially correlated, thus supporting the use of GMM. Next, the Hansen test validates the instruments used both in the difference and level parts of the system as a whole. We further conduct two difference Hansen tests to focus on the validity of particular subsets of instruments. The first test examines the validity of the exogeneity of the extra instruments used in the level part of the system as a whole; the second difference test checks the exogeneity of the lagged output used as an instrument in the level part.[31] Overall, the corresponding p-values validate the use of system (instead of difference) GMM estimators.

Turning to Column (2), it has non-financial assets as the compensating factor. As expected, the result on the compositional effect is opposite in sign to that of column (1).[32] These findings suggest that a reallocation involving an increase in net acquisition of non-financial assets (i.e., capital spending) financed through an offsetting reduction in expense (i.e., current spending) has a positive effect on growth. This result, however, deserves an additional comment. Although the signs of coefficients remain essentially the same under different model specifications, their statistical significance varies substantially. For instance, slight variations in the samples and the associated econometric technique provide coefficients which suddenly become non-significant (not shown to preserve space). These facts suggest that results under the economic classification of expenditure tend to be non-robust.

2. Functional Classification

We now turn to the functional classification of expenditure. Among the 8 functional categories covered in the dataset, we focus on the reallocation effects among the following 5 components: defense, transport and communication, health, education, and social protection. Our interest in these components is based on the fact that their effects on growth have often been studied in the related literature, as reviewed above. This literature generally argues that spending on education, infrastructure, defense and health can promote growth while social protection spending is not directly productive. We therefore expect that growth-enhancing reallocations should involve an increase in these four expenditure outlays compensated with a fall in social protection. However, formally detecting a growth-enhancing reallocation among them (if any) requires a rigorous empirical investigation.

Table 2 summarizes the results on the reallocation effects among those 5 functional components. (Full estimation results are left to Appendix C). Each of the five columns in the table designates the expenditure component that is increased in the reallocation, whereas each row indicates the associated component that is decreased to offset the change. Although the 5 different components yield 25 cells, only 20 are relevant for the analysis. Further, as clarified in the previous case with the economic classification, the symmetric nature of the analysis

[31] This second test is recommended by Roodman (2009a), who points out that a lagged dependent variable is often problematic among the sets of instruments used in the level part.

[32] This is expected because if increasing item A and decreasing item B promotes growth, decreasing A and increasing B should suppress it.

prompts us to highlight only 10 cases. When an enhancing/reducing effect is statistically significant, a star superscript is attached to the coefficient. Specifically, 1, 2, and 3 stars indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table 2. Summary Results of Compositional Effects: Functional Classification

		Component increased				
		Defense	Health	Education	Soc. prot.	Tracom
	Defense	n/a	Insignificant	Insignificant	Insignificant	Insignificant
	Health		n/a	Enhancing**	Insignificant	Insignificant
Component decreased	Education			n/a	Reducing**	Insignificant
	Soc. prot.				n/a	Insignificant
	Tracom					n/a

Notice that the table indicates that only education spending has growth-enhancing effects that are statistically significant. This happens specifically when an increase in education spending is financed by a fall in health or social protection spending. None of the other reallocations, even when involving a rise in economic infrastructure, produces statistically significant effects on growth. This result highlights the particular importance of education spending as a growth-enhancing component.[33]

Given this summary result, Table 3 elaborates on the effects of a rise in education spending financed by a fall in each of the other spendings components.[34] To explain, Column (2) estimates the regression in which fiscal components include (apart from the ratio of total expenditure to GDP) the ratio of education spending (to total expenditure) and the ratio of the addition of all the remaining 6 spending items but education and defense. Defense spending is thus treated as a compensating factor in this column. Columns (3) to (5) can be seen in a similar way except that the compensating factors are health, social protection, and transport and communication, respectively.[35] Lastly in Column (1), the only fiscal component included is education spending, implying that the compensating factor consists of all the remaining 7 components. Although this exercise does not provide a precise interpretation (since it does

[33]The growth-enhancing nature of education spending relative to the other spending components is carefully examined below. In particular, it will be studied to what extent this result holds under different specifications of the model.

[34]Since the economic affairs component of the functional classification is not always available at a more disaggregated level, the composition effects involving transport and communication, a subcomponent of it, are considered based on a smaller sample.

[35]When transport and communication is a compensating factor, spending excluding education and transport and communication includes the rest of the economic affairs spending category as well.

Table 3. Effects of Public Education Spending on Growth

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)
Total exp/GDP	−0.369	−0.375	−0.532*	−0.463	−1.063*
	(0.382)	(0.357)	(0.297)	(0.347)	(0.557)
Educ/Total exp	1.098*	0.539	1.553**	1.090**	−0.816
	(0.646)	(0.903)	(0.747)	(0.411)	(2.251)
Spend. ex. Defense and Educ/Total exp		−0.357			
		(0.506)			
Spend. ex. Health and Educ/Total exp			0.262		
			(0.730)		
Spend. ex. Educ and Soc Prot/Total exp				0.191	
				(0.253)	
Spend. ex. Educ and Tracom					−0.880
					(1.861)
Initial gdp p.c.	−0.066	−0.569	−2.155	−1.545	−1.581
	(2.666)	(2.571)	(2.690)	(1.840)	(1.974)
Initial human capital	5.644***	5.216***	4.127**	4.963***	2.349*
	(1.862)	(1.888)	(1.954)	(1.407)	(1.372)
Constant	−46.684	−0.963	−29.441	−31.743	120.712
	(44.379)	(77.324)	(100.962)	(38.351)	(167.526)
Compensating factor	All the rest	Defense	Health	Soc. Prot	Tracom
Observations	175	175	175	175	151
No. of countries	56	56	56	56	55
No. of instruments	33	39	39	39	37
Arellano-Bond AR(1), p-value	0.01	0.01	0.01	0.01	0.03
Arellano-Bond AR(2), p-value	0.19	0.19	0.20	0.19	0.27
Hansen, p-value	0.26	0.45	0.31	0.53	0.56
Diff Hansen 1, p-value	0.42	0.77	0.77	0.71	0.48
Diff Hansen 2, p-value	0.68	0.59	0.59	0.48	0.23

Robust standard errors in parentheses. Time dummies are not shown.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

not clarify how exactly other spendings fall), this still gives information about the general usefulness of education spending relative to all other categories.

This table suggests that the growth-enhancing effects of education may be quantitatively important, particularly when the rise in this spending component is compensated by social protection or health. For instance, a 1 percentage point increase in education spending offset by a

1 percentage point fall in social protection spending causes a 1.09 percentage points increase in growth over the 5-year period (i.e., about 0.2 percentage points increase per annum). A slightly larger effect is obtained if the compensating component is, instead, health spending. Finally, column (1) indicates that education is in general growth enhancing relative to all the other functional components. Turning to the remaining variables, the results are all in line with those of the economic classification. First, the effect of an increase in the level of total expenditure has a negative effect on growth. Next, initial GDP per capita has a negative effect on growth, while initial human capital has a positive effect on its own. Again, all the specification tests support the use of a system GMM approach for the estimation of the dynamic model.

Finally, Table 4, being parallel to Table 3, presents detailed results on the reallocation effects involving an increase in spending on transport and communication.[36] This is done to simply highlight its effect compared to that of education. As the table shows, none of the coefficients on the share of transport and communication to total spending has a statistically significant effect. Notice that this is still the case even when a rise in this spending is offset by a fall in health and social protection outlays.

C. Robustness

The main results from the above analysis are threefold. First, a simple reallocation of public spending that involves an increase in capital expenditure financed through an offsetting reduction in current spending has a positive yet not statistically robust association with growth. Second and most importantly, a reallocation involving an increase in education spending appears to have a statistically significant association with higher growth. Third, none of the possible reallocations among the other key functional components has a statistically significant relation with growth. This includes reallocations involving a rise in spending on economic infrastructure compensated by a fall in social spending. In what follows, we mainly check the robustness of our most important result on the statistical association of education spending with higher long-run growth.

[36] In fact, this table is a subset of that presented later in Appendix C.

Table 4. Effects of Transport and Communication Spending on Growth

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)
Total exp/GDP	−0.952*	−0.478	−1.423***	−1.107**	−0.076
	(0.508)	(0.508)	(0.441)	(0.549)	(0.654)
Tracom/Total exp	1.246	−1.097	0.414	0.241	1.420
	(1.937)	(1.353)	(1.494)	(2.240)	(1.881)
Spend. ex. Defense and Tracom/Total exp		−1.155**			
		(0.436)			
Spend. ex. Health and Tracom/Total exp			1.671**		
			(0.745)		
Spend. ex. Educ and Tracom/Total exp				−0.102	
				(0.878)	
Spend. ex. Tracom and Soc Prot/Total exp					0.932
					(0.576)
Initial GDP p.c.	−1.234	−2.490	−4.398	−2.049	−1.127
	(2.063)	(1.714)	(3.066)	(2.233)	(2.904)
Initial Human capital	2.344*	2.873*	2.067	2.106	5.899*
	(1.327)	(1.535)	(1.473)	(1.305)	(3.447)
Constant	24.675	134.600***	−60.299	53.370	−95.184
	(40.083)	(46.020)	(72.558)	(89.037)	(91.400)
Compensating factor	All the rest	Defense	Health	Education	Soc Prot
Observations	151	151	151	151	151
No. of countries	55	55	55	55	55
No. of instruments	33	37	37	37	37
Arellano-Bond AR(1), p-value	0.03	0.01	0.15	0.04	0.01
Arellano-Bond AR(2), p-value	0.22	0.36	0.33	0.33	0.43
Hansen, p-value	0.69	0.53	0.59	0.63	0.86
Diff Hansen 1, p-value	0.52	0.17	0.13	0.41	0.96
Diff Hansen 2, p-value	0.44	0.84	0.35	0.27	0.38

Robust standard errors in parentheses. Time dummies are not shown.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

1. Lagged Fiscal Variables

We first check the robustness of these results by changing the timing at which fiscal policy affects growth. We assumed previously that fiscal policy, by changing the steady state of the economy simultaneously, affects growth without a delay (see Eq. (3)). However, one may instead assume that fiscal policy affects the economy only with lags. To gauge the potentially delayed effect of public education expenditure, it is useful to acknowledge that the channels through which this spending affects the economy can be diverse. For instance, Aghion and others (2009) show that increasing education spending on research universities promotes growth through technological innovation (particularly in technologically advanced areas). Regarding this channel, it may be more sensible to consider that the enhancing effect of education spending emerges only with a delay. Turning to spending on transport and communication, it may take a while for a local community (including businesses) to take full advantage of the improved economic infrastructure such as roads, bridges, and airports. In light of this, we now assume in Eq. (3) that the steady state of the economy during the period spanning $t-1$ and t is affected solely by a fiscal policy change taking place in period $t-1$ (rather than the average over periods $t-1$ and t).

Table 5 presents the results only highlighting the coefficients on the share of education spending to total spending for brevity. The table reconfirms the importance of education spending as a growth-enhancing component. In particular, the coefficients on education spending with health and social protection as compensating components are similar to the ones in Table 3, although the reallocation with health is not statistically significant in this specification. Though not shown in the table, the coefficients on the rest of variables (i.e., initial GDP, initial human capital, the share of total expenditure to GDP, and the other expenditure share variables) are also in line with those reported in Table 3.

2. Different Development Levels

We next run regressions focusing on a smaller set of countries. Given that the reallocation effects may differ depending on the development level of a country, it would be ideal to run separate regressions for country groups with different income levels (e.g., LICs, MICs, and HICs as defined above). However, having a smaller sample quickly makes estimation results unreliable, because the number of instruments become too many relative to the number of samples (countries). It is thus difficult to restrict drastically the number of countries. Below, however, we attempt to address partially this issue while preserving the robustness of the reference regression (as in Table 3) by excluding the G20-Advanced countries.[37] While not en-

[37]The G20-advanced countries group (G20-Advanced) includes: Australia, Canada, France, Germany, Italy, Japan, Korea, United Kingdom and the United States. Since we do not have sufficient data for Germany and Japan, in practice this country group includes the remaining 7 economies.

Table 5. Robustness Check with Lagged Fiscal Variables

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)
Educ/Total exp	1.278*	0.761	1.423	1.234**	0.785
	(0.683)	(0.826)	(0.959)	(0.552)	(1.235)
Compensating factor	All the rest	Defense	Health	Social Prot	Tracom
Observations	175	175	175	175	151
No. of countries	56	56	56	56	55
No. of instruments	33	39	39	39	37
Arellano-Bond AR(1), p-value	0.00	0.00	0.00	0.00	0.00
Arellano-Bond AR(2), p-value	0.24	0.26	0.22	0.21	0.38
Hansen, p-value	0.27	0.43	0.46	0.32	0.64
Diff Hansen 1, p-value	0.47	0.65	0.62	0.39	0.68
Diff Hansen 2, p-value	0.15	0.18	0.27	0.73	0.43

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

tirely satisfactory, this check is meant to examine the robustness of the results to the subset of countries with a lesser degree of development.[38]

Table 6 again only presents coefficients on the share of education to total spending. Although it is true that the coefficient on this variable with social spending as the compensating factor is somewhat smaller than that of Table 3, the importance of education spending remains in this smaller sample. Again, the significance on education when the compensating factor is health spending is eroded, highlighting its relatively less robust effect. The coefficients on the other variables, though again not shown in the table, are again in line with Table 3.

[38]To give a further sense of how different results might be across country groups depending on their income levels, Appendix D provides a preliminary inspection of these differences when countries are divided roughly into HICs and MICs+LICs. Among other caveats, in the regressions presented there we do not address properly the possible endogeneity issues related to the aforementioned fact that the number of instruments tends to be larger than the number of samples.

Table 6. Robustness Check without G20-Advanced Countries

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)
Educ/Total exp	0.741	0.811	1.154	0.804*	−0.601
	(0.528)	(1.003)	(0.756)	(0.446)	(2.112)
Compensating factor	All the rest	Defense	Health	Social Prot	Tracom
Observations	142	142	142	142	125
Number of ifs	49	49	49	49	48
No. of instruments	33	39	39	39	35
Arellano-Bond AR(1), p-value	0.04	0.05	0.04	0.04	0.04
Arellano-Bond AR(2), p-value	0.12	0.11	0.12	0.11	0.17
Hansen, p-value	0.41	0.82	0.76	0.74	0.77
Diff Hansen 1, p-value	0.42	0.87	0.52	0.61	0.72
Diff Hansen 2, p-value	0.35	0.72	0.97	0.80	0.38

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

3. Additional Explanatory Variables

Third, we check the sensitivity of the results on education spending by adding an extra control/environmental variable. The additional variables considered here are the inflation rate, the degree of openness, population growth, and terms of trade growth. All of them have often been discussed as potential determinants of growth in the related literature (e.g., Barro (2004)).[39] We treat these variables as endogenous as is most likely the case in growth regressions. However, while the dynamic GMM framework allows us to deal with these additional endogeneity issues with internal instruments, adding more endogenous controls would again quickly make estimation results unreliable if the number of instruments becomes too many. We thus consider only a specification in which the compensating factors are not individually specified while adding these extra control variables one by one.

[39] Since our dataset covers a long time and contains a large number of countries including many LICs, the potential set of control/environmental variables to choose from tends to be limited. For instance, while we would like to also consider some institutional variables such as the degree of corruption as a control variable, it cannot be included in our regressions since it reduces our sample dramatically.

Table 7. Robustness Check with Additional Variables

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)
Educ/Total exp	1.098*	1.417**	0.594	1.380***	1.555***
	(0.646)	(0.580)	(0.601)	(0.515)	(0.557)
Inflation		−0.673*			
		(0.379)			
Openness			0.100		
			(0.106)		
Population growth				−1.340	
				(3.716)	
Terms of trade growth					0.516
					(0.782)
Compensating factor			All the rest		
Observations	175	150	158	165	153
No. of countries	56	50	53	52	50
No. of instruments	33	39	39	39	39
Arellano-Bond AR(1), p-value	0.01	0.06	0.01	0.01	0.07
Arellano-Bond AR(2), p-value	0.19	0.15	0.15	0.28	0.18
Hansen, p-value	0.26	0.64	0.29	0.44	0.62
Diff Hansen 1, p-value	0.42	0.97	0.24	0.55	0.89
Diff Hansen 2, p-value	0.68	0.86	0.84	0.37	0.47

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

Results are summarized in Table 7, which only presents the coefficients on the share of education and the added control variables. For comparison, Column (1) of the table replicates the result from the basic specification (without any added control variables) given in Column (1) of Table 3. We see that in all cases education spending appears to have a robust association with higher growth. The effects are in fact statistically significant in all but the case with the openness variable as an additional control. The added controls also have coefficients that are in line with those generally found in the literature (e.g., inflation having a negative impact on growth). Finally, as in the previous two checks, all the other key coefficients are consistent with the reference case without extra controls.

D. Central vs General Government Level

The final robustness check is related to the government level used in the analysis. So far this has been based on the consolidated CG level. As explained, this is due to the limited availability of GG level data. However, given that it is the GG which more accurately captures the state of public finances at a country level, it becomes important to examine whether our main results on education spending still hold when considering this more aggregate government level.

For this purpose, we now focus on a smaller sample under GFSM2001 in which expenditure composition information is available at both the CG and GG consolidated levels. Discarding the GFSM1986 yearbook dataset entirely and looking only at a limited number of countries under GFSM2001, the sample is inevitably restricted. Specifically, the data now cover three 5-year periods at most (i.e., 1996–2000, 2001–2005, and 2006–2010) and often only 2 periods (2001–2005 and 2006–2010), depending on the year at which a country migrates to GFSM2001. Using this more limited sample, we conduct again a panel data analysis (see Appendix E for details about this dataset). However, this shorter panel with a smaller coverage of countries implies that it is not feasible to use the dynamic GMM approach with internal instruments. As a compromise, we study a static fixed effects model without using those instruments, in which the dependent variable is now the average of the annual growth rate of real GDP per capita over 5 years, and the independent variables are the same as in the reference case presented before (see Eq. (1)).[40]

We now compare the results from the static panel analysis (Table 8). As mentioned, our purpose now is not to obtain accurate estimates of the reallocation effects of public expenditures but rather to compare the results obtained at the different CG and GG levels.[41] Specifically, we check if the difference between these results about reallocation effects involving education spending is statistically significant. If the difference is insignificant, this would suggest that the coefficients on the reallocation effects are statistically the same in both cases.

Column (1) in this table examines the effect of education spending on growth at the CG level when the compensating factor is not individually specified, whereas Column (2) does the

[40]That is, the independent variables contain initial real GDP per capita (as the real GDP per capita in the first year of the five years), initial human capital (as the average schooling years between ages 25 and 64 in the first year), relevant fiscal variables, fixed effects, and time dummies.

[41]Because we do not deal with the potential endogeneity issue with instruments, the result is likely to be not consistent. Moreover, the limited sample used here should make estimations less accurate than those presented previously.

Table 8. Effects of Education Spending at the Central and General Government Levels

Dependent variable: Annual GDP per capita growth, averaged over 5 years

Regressors	(1)	(2)	(3)	(4)
Total exp/GDP	−0.240	−0.350***	−0.359***	−0.364***
	(0.162)	(0.100)	(0.129)	(0.097)
Educ/Total exp	0.286	0.723*	0.602	0.712*
	(0.432)	(0.384)	(0.522)	(0.352)
Spend. ex. Social and Educ spend./Total exp			−0.327	−0.181
			(0.300)	(0.300)
Initial GDP p.c.	−8.294**	−10.981***	−10.612**	−11.640***
	(3.121)	(3.210)	(3.976)	(4.082)
Initial human capital	0.290	0.431	−0.762	0.370
	(0.998)	(1.133)	(1.556)	(1.058)
Constant	91.570**	117.404***	141.183**	132.278**
	(34.109)	(36.030)	(61.624)	(54.041)
Level of government	CG	GG	CG	GG
Compensating factor	All the rest	All the rest	Social Spend.	Social Spend.
Observations	59	59	59	59
R-squared	0.45	0.54	0.53	0.56
Number of countries	32	32	32	32

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Fixed effects model is used.

same at the GG level. Notice that the coefficients on the share of education spending are positive in both columns, implying that education spending is generally growth-enhancing relative to the other subcomponents. Though we do not claim its reliability, the value of 0.286 implies that one percentage point increase in the share of education spending is associated with 0.286 percentage points increase in growth per annum.[42] Notice that the coefficient is actually much larger at the GG level, with a statistical significance at 10 percent level. However, importantly, the paired t-test indicates that the difference between these coefficients is not statistically significant. Regarding the coefficients on the other variables, the signs are the same as the ones in Table 3. Column (3) and (4), using the CG and GG level data respectively, consider the case in which the compensating factor is specified as social spending. Again, the crucial result here is that the coefficients on education spending are not statistically different between the two cases. Overall, a general implication of these exercises is that results on the

[42] Remember that the dependent variable is now an (averaged) annual growth rate.

reallocation effects between education and social spending are likely to be robust to different government levels.[43]

IV. Concluding Remarks

The results of this paper suggest that in general, reallocations of government spending across different outlays (when taking as given the overall expenditure envelope) do not appear to be robustly associated with growth. However, a key exception stands out in the case of education spending. In particular, when an increase in this spending is offset by a fall in social protection spending, there seems to be a robust relation with higher growth. Moreover, this result stands to the alternative assumption of delayed fiscal policy effects, the use of a smaller sample excluding the richest countries in the dataset, and the addition of various extra control variables. Although our main analysis is conducted using the consolidated central government level data, our key finding appears to hold also at the consolidated general government level. Turning to the economic classification of expenditure, increases in capital spending financed through reductions in current spending (including public wages) seem to be positively associated with growth. Yet this particular result does not appear to be robust to different specifications of the model.

These results are important for various reasons. Focusing on one of the reasons mentioned in the introduction, the ongoing fiscal austerity measures in many advanced countries (often triggered by the fiscal expansion to overcome the negative effects of the 2007–8 financial crisis on economic activity) are expected to stay for several years. Under these circumstances in which a government may not be able to raise 'total' expenditure for a long period of time, it is important to highlight the potential effects of possible 'reallocation' measures. In this regard, we contribute to this strand of the policy debate by highlighting the relevance of education expenditure as a type of government spending that tends to promote growth. Also, our results tend to highlight the importance of public capital spending to boost growth in the long run, albeit with caveats due to the lack of robustness found in this particular case.

While this exercise provides more specific insights relative to previous works, we acknowledge that there is still room for further improvements. One obvious aspect is related to the efficiency/quality of spending. That is, since we simply look at the quantity (i.e., how much)

[43] We also run a set of regressions comparing results for the economic classification under both the CG and the GG levels (not shown). Results indicate that coefficients under the two different government levels do not appear to be statistically different. This provides further support to our previously-discussed findings considering the economic classification of spending under the CG level.

of spending without explicitly considering its quality, our results should still be taken with some caution. For instance, as mentioned in the introduction, the recent literature emphasizes the impact on the economy of different degrees of efficiency in public infrastructure spending across countries. Then, without controlling for the efficiency of this spending, our result obtained based only on quantities can be to some extent misleading: to the extent that the quality is high enough, an increase in this spending may be found to be robustly associated with higher growth when compensated by a fall in other types of spending. A challenge in this regard is to find adequate proxies to control for the quality of public spending among the different expenditure components considered in this paper, which is a fruitful topic for further research.

Moreover, despite its undeniable importance, economic growth is surely not the only criteria a government wants to take into account when deciding how to allocate public spending. While this paper focuses on growth, there are other crucial elements such as employment and income equality that should also be considered. In fact, even when social protection spending may not be a growth-enhancing type of spending, it may help promote income equality. Examining the effects of the public expenditure composition on these other key variables is also an important avenue for future work.

Appendix A. List of 56 Countries Used in Main Regressions

The following is the list of 56 countries used to estimate the reallocation effects of public spending on growth. Although the specific number of countries varies across regressions depending on the availability of the different spending components for a particular type of regression, we present below the list involving the analysis of a change in public education spending (Table 3: one of our main results). Note that our classification of countries by income levels reflect the entire 41 years of development between 1970 and 2010, as indicated in the main text. In alphabetical order, the classification of countries is as follows:

High-income countries (HICs): Australia, Austria, Bahrain (Kingdom of), Barbados, Belgium, Canada, Cyprus, Czech Republic, Denmark, Finland, France, Ireland, Israel, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Singapore, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States.

Medium-income countries (MICs): Argentina, Chile, Croatia, El Salvador, Estonia, Hungary, Iran (I.R. of), Kazakhstan, Korea (Republic of), Latvia, Malaysia, Mexico, Romania, Slovak Republic, Turkey, Uruguay.

Low-income countries (LICs): Bolivia, Cameroon, Dominican Republic, Egypt, Guyana, Indonesia, Mongolia, Morocco, Paraguay, Thailand, Tunisia, Ukraine, Bangladesh, Myanmar.

Appendix B. Summary Statistics: CG Dataset

Table 9 presents the summary statistics of the CG dataset used in the paper. This table shows that our sample countries grew, on average, around 12 percent in per capita terms over the 5-year period. Looking at the fiscal variables in the economic classification, the share of total public expenditure in GDP is about 33 percent. While the share of expense accounts for more than 90 percent of the total, the wage share takes about 20 percent. The total expenditure share in the functional classification is obtained at the slightly lower value of 28.4 percent. The reason why this is different from the total expenditure under the economic classification is twofold. First, samples are different: under the economic classification, the sample size is 190, whereas under the functional, it is 151. Second, even when the same sample is used, total expenditure under the functional is not larger than the one under the economic classification by definition. This is because we left out two components of the former (i.e., public order and safety and environmental protection). Note, however, that the shares taken

Table 9. Summary Statistics
(in percent)

Variable	Mean	Standard Deviation	Minimum	Maximum
Growth rate of real GDP pc (over 5 years)	11.94	9.61	−24.21	39.98
Economic classification				
Total Exp/GDP	33.03	8.90	13.05	53.87
Comp. of Employees/Total Spend.	20.20	9.47	5.85	54.04
Rest of Expense/Total Spend.	72.56	14.05	33.83	92.47
Nonfin. Assets/Total Spend.	7.24	6.50	1.24	36.47
Functional classification				
Total exp/GDP	28.40	9.72	10.43	52.88
Defense/Total Spend.	9.58	7.00	1.98	33.06
Transport and communication/Total Spend.	5.69	2.80	1.26	13.72
Health/Total Spend.	8.41	4.79	0.83	20.29
Education/Total Spend.	11.39	5.53	1.82	23.34
Social Prot./Total Spend.	27.11	15.43	1.33	54.89
Rest/Total Spend.	37.83	11.19	16.64	76.76

by these components are relatively small, thus the total of the remaining 8 is almost equal to the actual total expenditure.[44] In terms of the functional classification, while the share of education spending represents about 11.4 percent of the total, that of transport and communication is notably lower, at about half of it (5.7 percent). Social protection accounts for more than a quarter of the share of total functional expenditure (27 percent). When health is combined with social protection, the total (i.e., social spending) represents more than a third of total spending (35.5 percent).

Appendix C. Estimation Results Behind Table 2

The following table provides the detailed estimation results required to produce Table 2. The way to interpret the results is analogous to the other related tables including Table 3.

[44]For instance, if we focus on the 75 observations (out of 151) in which we also observe total functional spending, the difference is rather small: the total of the 8 components (relative to GDP) is 28.35 percent while the total of the 10 components is 29.66 percent.

Table 10. Effects of Compositional Changes on Growth: Functional Classification

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total exp/GDP	−1.048**	−0.375	−0.532*	−0.706	−0.828**	−0.561	−0.478	−1.423***	−1.107**	−0.076
	(0.416)	(0.357)	(0.297)	(0.465)	(0.348)	(0.386)	(0.508)	(0.441)	(0.549)	(0.654)
Health/Total exp	−0.502									
	(0.919)									
Educ/Total exp		0.539	1.553**							
		(0.903)	(0.747)							
Soc Prot/Total exp				−0.528	0.120	−0.864**				
				(0.497)	(0.868)	(0.429)				
Tracom/Total exp							−1.097	0.414	0.241	1.420
							(1.353)	(1.494)	(2.240)	(1.881)
Rest/Total exp	−0.397	−0.357	0.262	−0.277	0.473	−0.554	−1.155**	1.671**	−0.102	0.932
	(0.613)	(0.506)	(0.730)	(0.691)	(0.882)	(0.576)	(0.436)	(0.745)	(0.878)	(0.576)
Initial gdp p.c.	−1.213	−0.569	−2.155	−0.343	−1.411	−1.061	−2.490	−4.398	−2.049	−1.127
	(3.536)	(2.571)	(2.690)	(2.473)	(2.097)	(2.026)	(1.714)	(3.066)	(2.233)	(2.904)
Initial human capital	3.927	5.216***	4.127**	5.561***	5.598***	5.552***	2.873*	2.067	2.106	5.899*
	(2.727)	(1.888)	(1.954)	(2.082)	(1.733)	(1.543)	(1.535)	(1.473)	(1.305)	(3.447)
Constant	55.619	−0.963	−29.441	14.997	−34.271	42.901	134.600***	−60.299	53.370	−95.184
	(107.214)	(77.324)	(100.962)	(84.309)	(97.019)	(53.509)	(46.020)	(72.558)	(89.037)	(91.400)
Compensating factor	Defense	Defense	Health	Defense	Health	Educ	Defense	Health	Educ	Soc Prot
Observations	175	175	175	175	175	175	151	151	151	151
No. of countries	56	56	56	56	56	56	55	55	55	55
No. of instruments	39	39	39	39	39	39	37	37	37	37
Arellano-Bond AR(1), p-value	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.15	0.04	0.01
Arellano-Bond AR(2), p-value	0.20	0.19	0.20	0.18	0.20	0.18	0.36	0.33	0.33	0.43
Hansen, p-value	0.24	0.45	0.31	0.24	0.27	0.33	0.53	0.59	0.63	0.86
Diff Hansen 1, p-value	0.24	0.77	0.26	0.16	0.06	0.07	0.17	0.13	0.41	0.96
Diff Hansen 2, p-value	0.88	0.59	0.62	1.00	0.09	0.65	0.84	0.35	0.27	0.38

Robust standard errors in parentheses. Time dummies are not shown.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

Rest/Total exp differs across columns (e.g., In column (1), it is spending excluding defense and health).

Appendix D. Preliminary Estimation Results by Country Groups

We now attempt to shed light on whether the heterogeneity present in the dataset may affect results, by estimating the same model as in equation (3) but considering instead the HICs and LICs+MICs groups separately. However, in order to increase the number of samples/observations we have made slight adjustments to the model, including the use of the first year value of GDP (per capita) among each 5-year period as an initial GDP instead of using the previous 5 year averages. Subsequently, we managed to increase the number of total samples/observations for this particular analysis. Further, in an attempt to conduct this analysis, we adjust slightly the definition of these two income groups. Particularly, when we refer to the HICs group below, it includes the MICs group as well. Thereby countries belonging to the MICs group appear in both HICs and LICs+MICs in the regressions presented below.

Even after taking these measures, however, the reliability of the analysis is still relatively weak. There are a few reasons for this. For instance, the concern over the number of instruments (relative to the number of samples) here prompts us to treat fiscal variables as exogenous. The relatively weak nature of the results is reflected in the fact that in some of the specifications of the model, the estimated coefficient on initial GDP per capita turned out to be positive, thereby violating the conditional income convergence assumed in the model. Yet with caveats, dividing the sample in this way still sheds valuable light on the differences through which public spending variables affect growth across country groups.

D.1. Economic Classification

Table 11, below, summarizes results for the economic classification when the compensating factor is the net acquisition of non-financial assets. Importantly, results indicate that coefficient estimates do not differ significantly across country groups, suggesting that unobserved heterogeneity may not be a major concern in the sample used in our estimations. In addition, the sign of coefficients are consistent with those reported in Table 1, yet they now become statistically significant. This highlights that the net acquisition of non-financial assets has a positive effect on growth under most specifications, yet the robustness of such result tends to be rather limited as mentioned in the main text.

Table 11. Regressions by Country Groups: Economic Classification

Dependent variable: GDP per capita growth over 5 years

Regressors Country group	(1) HICs	(1) LICs+MICs	(2) HICs	(2) LICs+MICs
Expense/Total Exp	−1.150***	−0.767***		
	(0.389)	(0.220)		
Comp. of Employees/Total Exp			−0.912*	−0.654*
			(0.544)	(0.362)
Expense-no-Comp. of Employees/Total Exp			−1.129***	−0.775***
			(0.371)	(0.226)
Compensating factor	Nonfin. Assets	Nonfin. Assets	Nonfin. Assets	Nonfin. Assets
Observations	228	104	228	104
No. of countries	54	37	54	37
No. of instruments	38	38	39	39
Arellano-Bond AR(1), p-value	0.00	0.05	0.00	0.05
Arellano-Bond AR(2), p-value	0.69	0.99	0.74	0.99
Hansen, p-value	0.35	0.88	0.32	0.81
Diff Hansen 1, p-value	0.75	0.96	0.74	0.88
Diff Hansen 2, p-value	0.72	1.00	0.69	0.98

Robust standard errors in parentheses. Time dummies are not shown.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

D.2. Functional Classification

Turning to the functional classification of expenditure, Table 12 and Table 13, below, summarize results by country groups for education and transport and communication. Results reported in these tables do not differ significantly across country groups. Also, signs are consistent with those reported in Table 3 and Table 4. However, it may be worth noting that the effects of education spending appears to be significant only in HICs particularly when offset by a fall in spending on health and social protection (see Column (3) and (4) of Table 12). Potentially, the former observation relative to health spending may suggest that in LICs and MICs, public health also plays a key role in fostering human capital and thereby growth.

Table 12. Regressions by Country Groups: Education

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(1)	(2)	(2)	(3)	(3)	(4)	(4)	(5)	(5)
Country group	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs
Educ/Total exp	0.688*	0.336	0.498	0.414	1.067**	−0.016	0.787**	0.379	0.023	0.164
	(0.376)	(0.462)	(0.409)	(0.478)	(0.506)	(0.602)	(0.360)	(0.440)	(0.797)	(0.716)
Compensating factor	All the rest	All the rest	Defense	Defense	Health	Health	Social Prot	Social Prot	Tracom	Tracom
Observations	211	124	211	124	211	124	211	124	211	124
Number of ifs	55	53	55	53	55	53	55	53	51	48
No. of instruments	38	38	39	39	39	39	39	39	39	39
Arellano-Bond AR(1), p-value	0.10	0.22	0.10	0.22	0.10	0.22	0.11	0.23	0.04	0.08
Arellano-Bond AR(2), p-value	0.38	0.77	0.38	0.77	0.37	0.76	0.36	0.86	0.41	0.12
Hansen, p-value	0.37	0.32	0.33	0.33	0.41	0.36	0.20	0.27	0.18	0.69
Diff Hansen 1, p-value	0.96	0.33	0.94	0.34	0.97	0.39	0.81	0.27	0.16	0.81
Diff Hansen 2, p-value	0.75	0.18	0.75	0.16	0.75	0.22	0.65	0.19	0.06	0.99

Robust standard errors in parentheses. Time dummies are not shown. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

Table 13. Regressions by Country Groups: Transport and Communication

Dependent variable: GDP per capita growth over 5 years

Regressors	(1)	(1)	(2)	(2)	(3)	(3)	(4)	(4)	(5)	(5)
Country group	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs	HICs	LICs+MICs
Tracom/Total exp	0.603	0.328	0.452	0.213	1.039**	0.108	−0.023	−0.164	0.905*	0.488
	(0.421)	(0.491)	(0.544)	(0.593)	(0.434)	(0.688)	(0.797)	(0.716)	(0.469)	(0.502)
Compensating factor	All the rest	All the rest	Defense	Defense	Health	Health	Social Prot	Social Prot	Education	Education
Observations	170	104	170	104	170	104	170	104	170	104
No. of countries	51	48	51	48	51	48	51	48	51	48
No. of instruments	38	38	39	39	39	39	39	39	39	39
Arellano-Bond AR(1), p-value	0.04	0.09	0.04	0.08	0.03	0.09	0.04	0.08	0.04	0.06
Arellano-Bond AR(2), p-value	0.55	0.15	0.65	0.15	0.68	0.15	0.41	0.12	0.81	0.14
Hansen, p-value	0.47	0.77	0.33	0.77	0.41	0.76	0.18	0.69	0.27	0.58
Diff Hansen 1, p-value	0.52	0.81	0.27	0.81	0.36	0.79	0.16	0.81	0.33	0.74
Diff Hansen 2, p-value	0.25	0.83	0.06	0.87	0.13	0.85	0.06	0.99	0.13	0.74

Robust standard errors in parentheses. Time dummies are not shown. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Diff Hansen 1 tests the exogeneity of the instruments used in the level part (of the system) as a whole. Diff Hansen 2 tests the exogeneity of the lagged level of output used as an instrument in the level part.

Table 14. Government Expenditure Composition: Central and General Government Levels (in percent)

Variable	Mean	Std. dev.	Minimum	Maximum
Total exp/GDP (CG)	33.48	7.04	15.73	44.29
Total exp/GDP (GG)	40.24	7.45	20.52	57.66
Educ/Total exp (CG)	9.75	3.96	2.32	20.44
Educ/Total exp (GG)	13.60	2.77	6.35	19.69
Social/Total exp (CG)	49.93	9.62	8.22	63.19
Social/Total exp (GG)	48.92	8.11	26.23	59.74
Rest/Total exp (CG)	40.33	10.00	26.36	86.24
Rest/Total exp (GG)	37.48	7.37	26.39	55.32

Appendix E. Summary Statistics: CG vs GG Datasets

Table 14 presents the summary statistics for the 5-year averaged fiscal data used in the static panel regression in subsection III.D.[45] The number of countries covered is 32 (20 HICs, 9 MICs, and 3 LICs). For simplicity, we focus on the reallocation effects between education and social spending (as a combination of health and social protection). Further, to focus on this effect, we do not isolate defense spending and merge this subcomponent with the rest of spending. First, as for the mean of the share of total spending to GDP, the table indicates the obvious fact that when looking at the same sample, spending at the GG level is larger (40 percent) than the one at the CG level (33 percent).[46] Turning to the spending composition, notice that the share of education spending is actually larger at the GG level, implying that this spending is relatively more decentralized than the other subcomponents. However, when considering social spending, these shares are roughly the same between CG and GG. Also, though not shown in the table, when social spending is divided into health and social protection, we observe that spending on health is also relatively more decentralized (the share of 11.39 percent at the CG level and 13.50 percent at the GG level).

[45]To be consistent with our main regression analysis conducted above, we calculate the 5-year average if we have 3 observations or more (of all the expenditure subcomponents of our interest) out of the maximum of 5.

[46]As before, total expenditure is calculated as the addition of the 8 functional subcomponents out of 10. That is, public order and safety and environmental protection are excluded.

References

Agénor, Pierre-Richard, 2010, "A theory of infrastructure-led development," *Journal of Economic Dynamics & Control*, Vol. 34, pp. 932–950.

Aghion, Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche, 2009, "The Causal Impact of Education on Economic Growth: Evidence from U.S." *Brookings Papers on Economic Activity*, Vol. Spring, pp. 1–73.

Arellano, Manuel, and Stephen Bond, 1991, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, Vol. 58, pp. 277–297.

Arellano, Manuel, and Olympia Bover, 1995, "Another look at the instrumental-variable estimation of error-components models," *Journal of Econometrics*, Vol. 68, pp. 29–52.

Arnold, Jens Matthias, Bert Brys, Christopher Heady, Åsa Johansson, Cyrille Schwellnus, and Laura Vartia, 2011, "Tax policy for economic recovery and growth," *Economic Journal*, Vol. 121, pp. F59–F80.

Aschauer, David Alan, 1989, "Is public expenditure productive?" *Journal of Monetary Economics*, Vol. 23, pp. 177–200.

Barro, Robert, 1990, "Government spending in a simple model of endogenous growth," *Journal of Political Economy*, Vol. 98, No. 5, pp. S103–S125.

———, 2004, *Economic Growth* (Cambridge, Massachusetts: MIT Press), second ed.

Barro, Robert, and Jong-Wha Lee, 2010, "A New Data Set of Educational Attainment in the World, 1950–2010," NBER Working Paper 15902.

Blankenau, William F., and Nicole B. Simpson, 2004, "Public education expenditures and growth," *Journal of Development Economics*, Vol. 73, pp. 583–605.

Blundell, Richard, and Stephen Bond, 1997, "Initial conditions and moment restrictions in dynamic panel data models," University College London, Discussion Paper in Economics 97-07.

Bond, Stephen, Anke Hoeffler, and Jonathan Temple, 2001, "GMM Estimation of Empirical Growth Models," Discussion Paper No. 2048, Centre for Economic Policy Research.

Bose, Niloy, M. Emranul Haque, and Denise R. Osborn, 2007, "Public expenditure and economic growth: A disaggregated analysis for developing countries," *The Manchester School*, Vol. 75, No. 5, pp. 533–556.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort, 1996, "Reopening the convergence debate: A new look at cross-country growth empirics," *Journal of Economic Growth*, Vol. 1, pp. 363–389.

Clements, Ben, David Coady, Frank Eich, Sanjeev Gupta, Alvar Kangur, Baoping Shang, and Mauricio Soto, 2012, *The Challenge of Health Care Reform in Advanced and Emerging Economies*, Occasional Paper No 275 (Washington DC: International Monetary Fund).

Clements, Ben, David Coady, and Sanjeev Gupta, 2012, "The Challenge of Health Care Reform in Advanced and Emerging Economies," in Ben Clements, David Coady, and Sanjeev Gupta (eds.), *The Economics of Public Health Care Reform in Advanced and Emerging Economies*, pp. 3–21 (Washington DC: International Monetary Fund).

Dabla-Norris, Era, Jim Brumby, Annette Kyobe, Zac Mills, and Chris Papageorgiou, 2012, "Investing in public investment: An index of public investment efficiency," *Journal of Economic Growth*, Vol. 17, No. 3, pp. 235–266.

Dziobek, Claudia, Carlos Gutierrez Mangas, and Phebby Kufa, 2011, "Measuring fiscal decentralization – Exploring the IMF's database," IMF, Working Paper 11/126.

Easterly, William, and Sergio Rebelo, 1993, "Fiscal policy and economic growth," *Journal of Monetary Economics*, Vol. 32, pp. 417–458.

Gemmell, Norman, Richard Kneller, and Ismael Sanz, 2011, "The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries," *Economic Journal*, Vol. 121, pp. F33–F58.

Glomm, Gerhard, and B. Ravikumar, 1997, "Productive government expenditures and long-run growth," *Journal of Economic Dynamics and Control*, Vol. 21, pp. 183–204.

Gupta, Sanjeev, Benedict Clements, Emanuele Baldacci, and Carlos Mulas-Granados, 2005, "Fiscal policy, expenditure composition, and growth in low-income countries," *Journal of International Money and Finance*, Vol. 24, pp. 441–463.

Holtz-Eakin, Douglas, 1994, "Public-sector capital and the productivity puzzle," *Review of Economics and Statistics*, Vol. 76, No. 1, pp. 12–21.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey Rosen, 1990, "Estimating vector autoregressions with panel data," *Econometrica*, Vol. 56, pp. 1371–1395.

Kaganovich, Michael, and Itzhak Zilcha, 1999, "Education, social security and growth," *Journal of Public Economics*, Vol. 71, pp. 289–309.

Kneller, Richard, Michael F. Bleaney, and Norman Gemmell, 1999, "Fiscal policy and growth: evidence from OECD countries," *Journal of Public Economics*, Vol. 74, pp. 171–190.

Levine, Ross, Norman Loayza, and Thorsten Beck, 2000, "Financial intermediation and growth: Causality and causes," *Journal of Monetary Economics*, Vol. 46, pp. 31–77.

Lucas, Robert, 1988, "On the mechanics of economic development," *Journal of Monetary Economics*, Vol. 22, pp. 3–42.

Mauro, Paolo, 1995, "Corruption and growth," *Quarterly Journal of Economics*, Vol. 110, pp. 681–712.

———, 1998, "Corruption and the composition of government expenditure," *Journal of Public Economics*, Vol. 69, pp. 263–279.

Pesaran, M. Hashem, Y Shin, and Ron P. Smith, 1999, “Pooled mean group estimation of dynamic heterogeneous panels,” *Journal of the American Statistical Association*, Vol. 94, pp. 621–34.

Pesaran, M. Hashem, and Ron P. Smith, 1995, “Estimating long-run relationships from dynamic heterogeneous panels,” *Journal of Econometrics*, Vol. 68, pp. 79–113.

Pritchett, Lant, 2000, “The tyranny of concepts: CUDIE (cumulated, depreciated, investment effort) is not capital,” *Journal of Economic Growth*, Vol. 5, pp. 361–384.

Ram, Rati, 1990, “Wagner’s Hypothesis in Time-Series and Cross-Section Perspectives: Evidence from “ Real” Data for 115 Countries,” *Review of Economics and Statistics*, Vol. 69, No. 2, pp. 194–204.

Rodrik, Dani, 2008, “The real exchange rate and economic growth,” *Brookings Papers on Economic Activity*, Vol. Fall, pp. 365–412.

Roodman, David, 2009a, “How to do xtabond2: An instroduction to difference and system GMM in Stata,” *Stata Journal*, Vol. 9, No. 1, pp. 86–136.

———, 2009b, “A Note on the Theme of Too Many Instruments,” *Oxford Bulletin of Economics and Statistics*, Vol. 71, No. 1, pp. 135–58.

Wickens, Tobias, 2002, “Government Finance Stataistics Manual 2001 Companion Material: Classification of GFSM 1986 data to the GFSM 2001 framework,” International Monetary Fund.

Wyss, Kaspar, and Nicolaus Lorenz, 2000, “Decentralization and central and regional coordination of health services: The case of Switzerland,” *International Journal of Health Planning and Management*, Vol. 15, pp. 103–114.