Does public spending on tertiary education increase tertiary enrollment? Evidence from a large panel of countries

Bala, Patrick and Herzer, Dierk

9 July 2024

Does public spending on tertiary education increase tertiary enrollment? Evidence from a large panel of countries

Patrick Bala* and Dierk Herzer

Helmut Schmidt University Hamburg, Department of Economics, Holstenhofweg 85, 22043 Hamburg, Germany

Abstract

This study provides a systematic review of the few existing studies on the impact of public tertiary education spending on tertiary enrollment. It identifies several shortcomings in this literature and reexamines this impact while addressing the identified shortcomings, which include: (*i*) using public expenditures on tertiary education per student as a measure of overall public expenditures on tertiary education, (*ii*) omitting public costs per student when estimating the impact of public tertiary education spending on tertiary enrollment, (*iii*) ignoring potential endogeneity, (*iv*) ignoring possible spurious correlations in large *T* panels due to non-stationary data, and (*v*) not controlling for common time effects. In contrast to previous studies, this study finds, based on panel data for up to 149 countries between 1997 and 2018, a significant positive impact of public spending on tertiary education on tertiary enrollment that is robust to several sensitivity checks.

Keywords:

tertiary enrollment; public tertiary education spending; public costs per student; GMM

JEL classification:

H52, I23, I28

* Corresponding author. Tel.: + 49 40 6541 2775; fax: + 49 40 6541 2118. E-mail address: Patrick.bala01@outlook.com

1. Introduction

Governments worldwide allocate a portion of their limited resources to tertiary education, with the aim of improving educational access and thereby fostering economic development. A concern, however, is whether these investments effectively translate into increased tertiary enrollment rates or are a waste of public funds. The few studies on the impact of public tertiary education spending on tertiary enrollment report mixed results, ranging from significant positive relationships to insignificant and significant negative relationships. To date, there is not a single study that provides robust evidence that public spending on tertiary education increase tertiary enrollment. However, as we argue in this paper, existing studies suffer from several methodological shortcomings, including: (1) the use of a questionable measure of public expenditures on tertiary education, which likely captures the public costs per student rather than overall public tertiary education expenditures; (2) the omission of a measure of public costs per student when estimating the impact of public tertiary education spending on tertiary enrollment, which may lead to omitted variable bias; (3) ignoring potential endogeneity; (4) ignoring possible spurious correlations due to non-stationary data in panels with a large time dimension; and (5) not controlling for common time effects.

This study aims to address all these shortcomings of existing studies to examine whether public spending on tertiary education increases tertiary enrollment. Our study is novel in several respects. First, we examine the impact of public tertiary education spending on tertiary enrollment, incorporating a measure of per-student public costs. Second, we estimate not only fixed effects models (with both country and time fixed effects), in which we also use lagged regressors, but also employ the system GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998) to account for the potential endogeneity of public expenditures on tertiary education. We chose this estimator, which uses internal instruments, due to the well-known difficulty (especially in macro studies) of finding variables that qualify as strong and valid external instruments. Third, we eliminate the danger of spurious regressions in our unbalanced panel by using three-year averages to reduce the number of time series observations per country. Using three-year averages also allows us to remove cyclical effects, thereby capturing longer-term effects. Fourth, we use panel data for a large number of countries (149) over a long period (1997–2018). In contrast, both previous cross-sectional and panel studies cover a much smaller number of countries, and existing panel studies are also based on shorter observation periods. Fifth, we provide a systematic review of the literature.

To preview the main finding of this paper, in contrast to previous studies, we find a significant positive impact of public spending on tertiary education on tertiary enrollment that is robust to several sensitivity checks.

The rest of the paper is organized as follows: Section 2 provides a theoretical discussion of the impact of public tertiary education spending on tertiary enrollment and derives the empirical model. Section 3 reviews the empirical literature based on the considerations of Section 2. Section 4 presents the results of our empirical analysis. Finally, Section 5 concludes.

2. Theoretical discussion and empirical model

2.1. Framework

To provide a framework for our theoretical discussion of the impact of public tertiary education spending on tertiary enrollment and for our empirical model, including our review of the empirical literature (in the next section), we start with a simple identity that relates total tertiary enrollments, E, to public expenditures on tertiary education, P, and to the public costs per student, C:

$$E = \frac{P}{C} \tag{1}$$

By dividing both sides of the equation by the number of persons of tertiary education age (defined as 18-28 years) and by GDP, and considering that GDP is the product of GDP per capita and population, we obtain an equation that relates the gross enrollment rate in tertiary education (i.e., the number of people enrolled in tertiary education as a fraction of the total number of people aged 18 to 24), *GERT*, to public expenditures on tertiary education as a proportion of GDP, *EDUXT*, the public costs per student as a proportion of GDP per capita, *COSTS*, and the proportion of the population of tertiary education age, *TERTPOP*:

$$GERT = \frac{EDUXT}{TERTPOP \times COSTS} \tag{2}$$

The logarithmic form of this equation is:

$$\log GERT = \log EDUXT - \log TERTPOP - \log COSTS \tag{3}$$

2.2. Theoretical discussion

Although equations (2) and (3) represent an identity, they can be given an economic interpretation. The supply-side interpretation regarding *EDUXT* is as follows: An increase in the proportion of GDP allocated to tertiary education in the form of direct expenditures, such as salaries for professors and other teachers, or spending on the construction of university buildings and the purchase of textbooks and scientific journals, results in an increase in the supply of tertiary education, provided that the proportion of the population of tertiary education age and the relative costs per student enrolled remain unchanged (or increase less rapidly).

A demand-side interpretation of the identity given by equations (2) and (3) regarding *EDUXT* is that an increase in public expenditure on tertiary education in the form of indirect expenditures, such as public subsidies to students for scholarships and other grants, as a proportion of GDP, raises the demand for tertiary education and thus increases the tertiary enrollment rate, unless *TERTPOP* and *COSTS* do not change (or increase less rapidly) and the supply of tertiary education is not fully inelastic.

Regarding *TERTPOP*, it can be argued that an increase in the proportion of the population of tertiary education age implies that scarce education resources must be spread over a larger number of students, assuming that the demand for education increases with the population of tertiary education age. This increase in per capita costs reduces the supply of tertiary education per person of tertiary education age and, consequently, the tertiary enrollment rate.

Similarly, an increase in *COSTS* implies a shift of the supply curve to the left and thus a reduction in enrollment (for constant tertiary education expenditures). The simple logic is that as more funds are directed to cover the increased costs per student, less money is available for expanding access, providing scholarships, or supporting additional students. Alternatively, higher costs can make tertiary education less accessible if these costs are passed on to students through higher tuition fees. However, an increase in costs per student (as a proportion of GDP per capita) may also reflect an increase in the quality of education (better facilities, more qualified teaching personnel, advanced materials). This quality improvement could positively affect enrollment, even if the increased costs are partly passed on to the students (through higher tuition fees), if potential students perceive the increased quality as beneficial for their studies and future careers. Thus, the elasticity of *GERT* with respect to *COSTS* does not necessarily need to be negative or equal to one.

While equation (3) assumes that the elasticities of *GERT* with respect to *EDUXT*, *TERTPOP*, and *COSTS* are equal to one (in absolute value), it is useful for further discussion to consider the following more general version of equation (3):

$$\log GERT = \alpha \log EDUXT + \beta \log TERTPOP + \gamma \log COSTS$$

(4)

Our primary interest is in the elasticity of *GERT* with respect to *EDUXT*, α. This elasticity may be smaller than one or even zero for several reasons.

First, in addition to public education expenditures, there are private education expenditures, such as tuition fees, that finance higher education to some extent. Thus, individuals might reduce their education expenditures when higher education is financed by the government. In other words, public spending could merely act as a substitute for private spending. *Second*, a large part of government expenditures on tertiary education consists of salaries for professors and other academic staff. If the supply of academics is relatively constant, an increase in public funds for tertiary education may merely lead to higher salaries rather than an increase in the number of academics employed. *Third*, public funds for tertiary education might not be used effectively, for example, if a significant portion is consumed by administrative overheads rather than directly benefiting students, or if funding priorities are imbalanced, favoring research over teaching, scholarships, grants, and student support services. Additionally, funding decisions may be driven by political considerations rather than actual educational needs. *Fourth*, if there is already a high level of tertiary education enrollment, additional funding may yield diminishing returns. Once the majority of the population willing and able to pursue higher education is enrolled, further expenditures may not significantly increase enrollment rates. *Fifth*, many prospective students face challenges such as inadequate academic preparedness, cultural and social constraints, and family responsibilities, which financial aid alone cannot overcome. These barriers can deter students from pursuing higher education even when financial support is available. *Sixth*, if the public perceives that a tertiary education degree does not significantly enhance job prospects or earning potential, they may be less motivated to enroll, irrespective of increased funding. This perception can be influenced by labor market conditions and the visibility of successful career paths that do not require a tertiary degree.

Finally, we note that the third point might explain why the impact of public tertiary education spending on tertiary enrollment might be larger in country groups with higher income than in country groups with lower income, where public resources are often used less efficiently. In contrast, the fourth point could explain why public tertiary education spending might have a larger effect on the tertiary enrollment rate in less developed countries, where fewer people are enrolled than in more developed countries.

2.3. Baseline empirical model, data, and estimation issues

To estimate the elasticity of *GERT* with respect to *EDUXT*, we employ the following model:

$$\log GERT_{it} = \alpha \log EDUXT_{it} + \sigma X_{it} + \mu_i + f_t + \varepsilon_{it} \tag{5}$$

where i and t are country and time indices; and X_{it} is a vector of control variables. In our baseline specification, we not only control for $\log TERTPOP_{it}$ and $\log COSTS_{it}$, but we also include the log of real GDP per capita, $\log GDPPC_{it}$, the log of the urbanization rate, $\log URBAN_{it}$, and the log of the gross enrollment rate in secondary education, $\log GERS_{it}$ as control variables, based on the previous literature. In addition, we control for unobserved time-invariant country characteristics (such as geography and culture), μ_i, by including country dummies, as well as time-varying unobserved common factors (such as global business cycles or global crises), f_t, by including time dummies.

Regarding $\log COSTS_{it}$ as a control variable, we explicitly note that an increase in public education cost per student may prompt governments to allocate a higher proportion of their GDP to public education expenditures to support accessibility and affordability. Thus, $\log COSTS_{it}$ may have a positive effect on $\log EDUXT_{it}$. Given that an increase in the public costs per student as a proportion of GDP per capita may also affect the tertiary school enrollment rate, the omission of $\log COSTS_{it}$ as a covariate in an empirical model of the impact of public tertiary education spending on tertiary enrollment may result in a classical omitted variable bias. In this case, the estimated elasticity of public tertiary education spending on tertiary enrollment would capture not only the true effect of log$EDUXT_{it}$ on log$GERT_{it}$ but also the effect of education costs per student on the tertiary enrollment rate. Thus, the omission of log$COSTS_{it}$ may result in a downwardly biased estimate of α if these costs have a negative effect on the tertiary enrollment rate and a positive effect on public expenditures on tertiary education as a proportion of GDP. To avoid this bias, it is important to control for log$COSTS_{it}$.

However, if there are factors that limit the positive impact of public spending on tertiary education enrollment, as discussed above, an increase in log$EDUXT_{it}$ results in higher per-student public tertiary education expenditures or in log$COSTS_{it}$. In other words, to the extent that public tertiary education spending (as a proportion of GDP) is not fully effective, it is accompanied by an increase in public tertiary education expenditures per student (as a proportion of GDP per capita) and hence in public costs per student (as a proportion of GDP per capita). Therefore, one might be concerned that including log$COSTS_{it}$ as a covariate might introduce an 'included variable bias'. In this scenario, the estimated elasticity α does not fully capture the effect of log$EDUXT_{it}$ on log$GERT_{it}$ if public costs or expenditures per student increase due to the limited effectiveness of public expenditures. To address this concern, we also present estimation results from several specifications that exclude log$COSTS_{it}$.

As our measure of $COSTS_{it}$, we use public tertiary education expenditures per student as a percentage of GDP per capita, consistent with several other studies that also utilize public education expenditure per student as a percentage of GDP or GNP per capita to measure the unit costs of education relative to GDP or GNP per capita (see, e.g., Colclough and Al-Samarrai, 2000; Appiah and McMahon, 2002; Keller, 2006).

The definition of the variables and the sources used are presented in Table A1 in the appendix. We include all available data from these sources, resulting in an unbalanced panel of 149 countries with data between 1997 and 2018. The countries included in our sample are listed in Table A2 in the appendix. Summary statistics are presented in Table A3.

We estimate equation (5) using the OLS fixed effects (FE) estimator as the baseline estimator. Given our relatively long sample period of 22 years, there is a concern that this estimator might provide invalid inference that can be spurious when the underlying variables are non-stationary. Monte Carlo simulations by Entorf (1997) and Kao (1999) demonstrate that there is a risk of spurious regressions in fixed effects regressions with non-stationary variables, even with ten time series observations per country. They also demonstrate that an increase in the time-series dimension (T) increases this risk, and that an increase in the number of cross-sectional units (N) does not decrease the risk of spuriously indicating a relationship but may even increase it. Since our panel is highly unbalanced, characterized by numerous data gaps and only a few countries having more than ten consecutive time series observations, with none having a complete time series over the entire observation period, the risk of spurious regressions due to non-stationary data in panels with a large time series dimension should be small in our study. However, there is still a risk. To eliminate the danger of spurious regressions, we artificially reduce the time series dimension of our panel by using three-year averages, except for the period 2015–2018, which is represented by a four-year average.[1] Consequently, our panel consists of seven periods: 1997–1999, 2000–2002, 2003–2005, 2006–2008, 2009–2011, 2012–2014, and 2015–2018. Using n-year averages is a common practice in panel studies to capture long-run effects and to smooth business-cycle fluctuations. In the empirical analysis, we utilize 5-year averages as well as annual data as robustness checks.[2]

Another concern is that the effect of public tertiary education spending on tertiary enrollment might differ across countries with different income levels, as noted above and suggested by some

[1] The structure of the available data does not allow meaningful panel cointegration analysis to alternatively address the issue of spurious regressions due to non-stationary data in panels with a large time dimension.
[2] When we use five-year averages, except for the period 2012–2018, our panel includes four periods: 1997–2001, 2002–2006, 2007–2011, and 2012–2018.

studies (discussed in the next section). To address this concern, we also present results for different income groups: low-income countries (LIC), lower-middle-income countries (LMIC), upper-middle-income countries (UMIC), and high-income countries (HIC).[3]

A further concern is the potential endogeneity of public expenditures on tertiary education as a proportion of GDP: An increase in the tertiary enrollment rate may reflect a higher demand for higher education. When governments respond to the increased enrollment rate by allocating more resources to tertiary education, the estimate of α may overstate the positive causal effect of public tertiary education spending on tertiary enrollment. Alternatively, it could be that an increase in the tertiary enrollment rate leads to an increase in human capital, and thereby an increase in GDP and thus a decrease in public expenditures on tertiary education as a proportion of GDP. In this case, the estimate of α may understate the elasticity of $GERT_{it}$ with respect to $EDUXT_{it}$.

To address the potential endogeneity of log$EDUXT_{it}$, we also estimate specifications with lagged values of all regressors. Additionally, we employ the system GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). As is well known, this estimator (designed for small-T large-N panels such as the one used here) is a dynamic panel estimator that accounts for endogeneity using internal instruments while avoiding the well-known 'Nickell bias', which arises from applying a fixed effects estimator to a lagged dependent variable model in a panel with small T.[4] To demonstrate the robustness of our results, we estimate two GMM models: one that treats only public tertiary expenditures as endogenous, and another that treats all explanatory variables (including ln$COSTS_{it}$) as endogenous.

[3] We classified the countries as low-income, lower-middle-income, upper-middle-income, or high-income countries if they were categorized as such in the World Bank's 'historical classification by income' (accessible at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519) for more than half of the years in our sample period.
[4] We use the forward orthogonal deviations transformation, an alternative to first differencing proposed by Arellano and Bover (1995), which subtracts the average of all remaining future periods from the level in period t. This transformation, unlike first differencing, preserves the sample size in unbalanced panels, such as the one used here.

3. Review of existing studies

Based on the theoretical considerations and the discussion of our empirical baseline model in the previous section, we now provide a review of studies on the impact of public tertiary education spending on tertiary enrollment. After a careful literature search, we identified only eight studies on this impact. Among them, two studies, Buckner and Khoramshahi (2021) and Wang et al. (2023), are not primarily concerned with this impact. Instead, they focus on the impact of other factors, such as tertiary school enrollment in the private sector (Buckner and Khoramshahi, 2021) and economic policy uncertainty (Wang et al., 2023), on the tertiary enrollment rate. Thus, public spending for tertiary education is not the variable of primary interest in these two studies but serves only as a control variable. Table 1 lists all identified studies, including those by Buckner and Khoramshahi (2021) and Wang et al. (2023), and also presents information on each study regarding their focus (public tertiary expenditures or other potential determinants of tertiary enrollment), the main variables of interest for our review, the type of analysis (cross-sectional analysis, panel analysis), the sample composition (number of countries and observation period), the econometric methods used, and the main results.

[Table 1]

3.1. Preliminary remarks

Before discussing the existing evidence in more detail, it is noteworthy that two studies—those by Bergh and Fink (2006, 2008)—use cross-sectional data and thus employ cross-sectional analysis, which is known to inherently suffer from omitted variable bias due to unobserved country-specific time-invariant factors. The remaining six studies—those by Winter-Ebmer and Wirz (2002), Yang and McCall (2014), Buckner and Khoramshahi (2021), Yang and St. John (2023), Hajebi et al. (2023), and Wang et al. (2023)—employ panel data models, which allow control for country-specific time-invariant factors.

In addition, it is worth mentioning that the samples used in these studies vary widely in terms of the number of countries included. Three studies use relatively small samples, with between 14 and 18 countries (Winter-Ebmer and Wirz, 2002; Hajebi et al., 2023; Wang et al., 2023). Three are based on samples of medium size, with a maximum number of countries between 64 and 86 (Bergh and Fink, 2008; Yang and McCall, 2014; Yang and St. John, 2023), and two on relatively large samples, with a maximum number of countries between 122 and 127 (Bergh and Fink, 2006; Buckner and Khoramshahi, 2021).

Moreover, the observation periods vary across the studies. One study, by Bergh and Fink (2006), considers only one year. The observation period of another study, by Hajebi et al. (2023), includes only 10 years. The majority of studies are based on periods of medium length, between 12 and 22 years (Winter-Ebmer and Wirz, 2002; Yang and McCall, 2014; Buckner and Khoramshahi, 2021; Yang and St. John, 2023; Wang et al., 2023). One study, by Bergh and Fink (2008), considers a relatively long sample period of 31 years. However, this study is purely cross-sectional.

3.2. Main results

The individual results of existing studies on the impact of public tertiary education spending on tertiary enrollment can be summarized as follows.

Only three studies—those by Buckner and Khoramshahi (2021), Hajebi et al. (2023), and Wang et al. (2023)—report statistically significant positive coefficients on their tertiary expenditure variable. However, Hajebi et al. (2023) present the results of only one regression, while Buckner and Khoramshahi (2021) and Wang et al. (2023) present the results of several regressions, where the coefficients are sometimes insignificant. Surprisingly, Wang et al. (2023) find a positive coefficient in their total sample consisting of both developed and developing countries. However, when the sample is subdivided into developed and developing countries, the coefficient from the GMM estimator becomes negative and insignificant in both subsamples. In one study—that by Winter-

Ebmer and Wirz (2002)—the coefficients are always insignificant. Two studies—those by Bergh and Fink (2006, 2008)—find both statistically significant negative coefficients and statistically insignificant coefficients. Finally, the coefficients on the public tertiary expenditure variable are consistently negative and statistically significant in two studies—Yang and McCall (2014) and Yang and St. John (2023). It is worth mentioning that Yang and McCall (2014) find a larger negative coefficient (in absolute value) for developed countries than for developing ones. Overall, the results are mixed and sometimes conflicting, with no robust evidence supporting a positive effect of public tertiary education spending on tertiary enrollment.[5]

3.3. Dependent variables and the tertiary expenditure variables

All studies use the gross tertiary enrollment rate as the dependent variable or the change in the gross tertiary enrollment rate. The only exception is Yang and St. John (2023), who use absolute enrollments in short-cycle tertiary education. To provide one explanation for the mixed results, we now discuss the tertiary expenditure variables used in these earlier studies.

While one study (Winter-Ebmer and Wirz, 2002) uses public tertiary education expenditures in absolute terms as the explanatory variable, three studies (Buckner and Khoramshahi, 2021; Hajebi et al., 2023; Wang et al., 2023) use public tertiary education expenditures as a percentage of GDP, similar to our study. Four studies employ public tertiary education expenditures per student, either in absolute terms (Yang and McCall, 2014) or as a percentage of GDP per capita (Bergh and Fink, 2006,

[5] For completeness, two points should be mentioned. First, Yang and St. John (2023) also consider government expenditures on tertiary education as a percentage of total government expenditures on education as an explanatory variable. The coefficient on this variable is positive and significant in the two regressions the authors conduct. Given that this result suggests a reallocation of public education expenditures towards tertiary education, and not necessarily an increase in public tertiary education spending, results in an increase in tertiary enrollment, we have decided not to include this variable and the corresponding result in Table 1. Second, while Wang et al. (2023) employ the system GMM estimator as their main estimator, they also use pooled OLS, as well as random effects and fixed effects (without time effects). The coefficient on their tertiary expenditure variable is also significant and positive for the total sample in these regressions. They also present pooled OLS results (without country and time fixed effects) for the subsamples of developed and developing countries. The coefficient on public tertiary education expenditures as a percentage of GDP from these regressions remains positive and significant only for the subsample of developing countries.

2008; Yang and St. John, 2023). All these latter four studies report negative, although not necessarily significant, coefficients. However, as discussed in Section 2, other studies use public education expenditures per student (as a percentage of GDP per capita) as a measure of public costs per student. Thus, the results of these four studies are very likely to reflect the effect of public costs per student on tertiary enrollment, rather than the effect of public education expenditures on tertiary enrollment.

Another, complementary explanation for the significant negative coefficients on public education expenditures per student (as a percentage of GDP per capita) is, of course, their likely endogeneity, given that the number of students implicitly appears in both the tertiary school enrollment rate and public education expenditures per student (as a percentage of GDP per capita).

Interestingly, all studies that report positive (but not necessarily significant) coefficients use public education expenditures (as a percentage of GDP) as a regressor. Admittedly, in two of these studies (Winter-Ebmer and Wirz, 2002; Wang et al., 2023), the coefficient on this variable is also negative (but insignificant) in some regressions.

3.4. Weaknesses of previous studies

We now briefly discuss the weaknesses of previous studies. As noted above, four out of the eight existing studies on the impact of public tertiary education spending on tertiary enrollment use public tertiary education expenditures per student as their measure of public tertiary expenditures. Therefore, they are very likely to capture, at least to some extent, the impact of public costs per student on tertiary enrollment rather than the effect of public education expenditures on tertiary enrollment.

The remaining four studies use public education expenditures (as a percentage of GDP) as a regressor. However, these studies do not control for public costs per student as a proportion of GDP per capita. It is clear, therefore, that their estimates may suffer from an omitted variable bias, as discussed in Section 2.3.

In addition, six out of the eight studies (Bergh and Fink, 2008; Yang and McCall, 2014; Buckner and Khoramshahi, 2021; Yang and St. John, 2023; Hajebi et al., 2023; Wang et al., 2023) do not address the potential endogeneity of their measure of public tertiary education spending.[6] Thus, it cannot be ruled out that the results of these studies are biased by endogeneity.

Two studies—those by Winter-Ebmer and Wirz (2002) and Bergh and Fink (2006)—use an instrumental variable (IV) approach, in addition to OLS, to address potential endogeneity problems. However, even the results of these studies can be biased if the instruments are weak (i.e., not sufficiently correlated with the potentially endogenous variable) and/or invalid (i.e., correlated with the error term and hence with the dependent variable).[7] As is well known, it is difficult, and sometimes impossible, to find external variables that qualify as strong and valid instruments in macro studies.

Moreover, the existing panel studies use annual observations and include countries with at least 10 time-series observations. Therefore, it cannot be ruled out that some of their significant results reflect spurious correlations caused by non-stationary data, as discussed in Section 2.3.

Finally, three studies—those by Buckner and Khoramshahi (2021), Hajebi et al. (2023), and Wang et al. (2023)—that employ panel data techniques do not control for time-varying unobserved common factors (such as global business cycles or global crises) that may influence both the tertiary school enrollment rate and public tertiary education expenditures. Thus, the estimates from these studies may be additionally biased.

[6] We found no information in Wang et al.'s (2023) study on whether they addressed the potential endogeneity of tertiary education expenditures (as a percentage of GDP) in their GMM estimations. Additionally, Wang et al. (2023) do not mention the potential endogeneity of tertiary education expenditures.

[7] The instruments used by Winter-Ebmer and Wirz (2002) include the sizes of the male and female populations aged 18-24 years, as well as indicators of government form and ideology. The instruments used by Bergh and Fink (2006) include public per capita expenditures on secondary education, in addition to GDP per capita and the percentage of the population younger than 15 years. Both Winter-Ebmer and Wirz (2002) and Bergh and Fink (2006) provide evidence that their instruments are strong. However, to the extent that their instruments are correlated with the tertiary enrollment rate, their results are biased.

In summary, all existing studies suffer from several methodological shortcomings. In the following analysis, we address these shortcomings and use a larger sample of countries than any previous study, as well as a longer sample period than the existing panel studies.

4. Results

4.1. Baseline fixed effects results

Column 1 in Table 2 presents the results from our baseline model, based on three-year averages. The coefficients of the control variables are as theoretically expected. However, while the coefficients on $\log COSTS_{it}$, $\log TERTPOP_{it}$, and $\log GDPPC_{it}$ are significant at least at the 10% level, those on $\log URBAN_{it}$ and $\log GERS_{it}$ are not significantly different from zero.

The insignificance of the gross secondary school enrollment rate is consistent with the findings of Bergh and Fink (2008) and Yang and St. John (2023). Other studies, however, find significant positive coefficients for the secondary school enrollment rate (see, e.g., Yang and McCall, 2014; Wang et al., 2023), as we do in some of the regressions below. Moreover, the insignificance of the urbanization rate is consistent with the findings of Bergh and Fink (2008) and Buckner and Khoramshahi (2021). The significant positive coefficient on GDP per capita accords with the findings of Bergh and Fink (2006), Yang and McCall (2014), Buckner and Khoramshahi (2021), and Wang et al. (2023), but contrasts with the results of Bergh and Fink (2008) and Yang and St. John (2023), who report insignificant coefficients. Furthermore, the significant negative coefficient on $\log TERTPOP_{it}$ is in line with the results of Bergh and Fink (2006) and Yang and McCall (2014), whose findings suggest that changes in the population structure toward a higher (lower) share of the population of tertiary education age are negatively (positively) correlated with the tertiary enrollment rate. Finally, consistent with most findings of those studies that examine public tertiary education expenditures per student (discussed in the previous section), the coefficient on $\log COSTS_{it}$ is negative and statistically significant.

Turning to our variable of interest, the results reveal a positive and significant coefficient on log$EDUXT_{it}$. According to the point estimate, a one percent increase in public expenditures on tertiary education as a proportion of GDP leads to a 0.829 percent increase in the gross enrollment rate in tertiary education. In terms of economic significance, this estimate implies that a one-standard-deviation increase in log$EDUXT_{it}$ is associated with a 46.78 percent increase in the standard deviation of the enrollment variable (0.829×0.6312441/1.118629),[8] indicating an economically significant effect.

[Table 2]

4.2. Robustness checks

In column 2 of Table 2, we investigate whether the significant positive coefficient on log$EDUXT_{it}$ is driven by potential outliers (or extreme values) by winsorizing our dependent variable and our public expenditure variable at the one percentile level in both tails of the distribution. While the coefficient is slightly smaller than that in Column 1, it remains highly significant, suggesting that the significant positive coefficient on log$EDUXT_{it}$ is not driven by potential outliers. However, since winsorizing replaces extreme values with less extreme values, it can introduce inaccuracies in data where extreme values represent legitimate values. Therefore, we proceed with the unwinsorized data.

In columns 3 and 4, we check the robustness of our results to the use of annual data and five-year averaged data, respectively. The coefficient on log$EDUXT_{it}$ remains positive and significant at the 1% level.

In Table 3, we test whether our results are driven by a particular income group by estimating equation (5) for low-income countries, lower-middle-income countries, upper-middle-income countries, and high-income countries. While the coefficient on log$EDUXT_{it}$ is highest for upper-middle-income countries and lowest for low-income countries, the difference across these country

[8] The standardized coefficient of 0.4678 is calculated by multiplying the unstandardized coefficient of 0.829 by the ratio of the standard deviation of log$EDUX_{it}$ to the standard deviation of log$GERT_{it}$. The values (standard deviations) used for the calculation of the standardized coefficient can be found in Table A3.

groups is not substantial, and the coefficient remains statistically highly significant for all country groups.

[Table 3]

In Table 4, we address the potential endogeneity of log$EDUX_{it}$. Column 1 presents the results with lagged variables. The coefficient on log$EDUX_{it}$ becomes somewhat smaller but is still significant at the 1% level.

[Table 4]

In columns 2 and 3 of Table 4, the GMM results are reported. In column 2, only log$EDUX_{it}$ is treated as endogenous, whereas the other explanatory variables are predetermined. In column 3, all explanatory variables are treated as endogenous.

While it is perhaps needless to say that the lagged dependent variable is always treated as predetermined, it is important to say that we use the two-step estimator as it is more efficient than the one-step estimator. However, a well-known property of the two-step estimator is that the standard errors may be severely biased downwards in small samples. To address this problem, we adopt the finite sample correction to the standard errors proposed by Windmeijer (2005).

Moreover, it is well known that GMM can exhibit the problem of 'too many instruments' when the number of instruments exceeds the number of cross-sectional units. The proliferation of instruments can lead to unreliable inference and weaken the Hansen-J test of overidentifying restrictions. To avoid this problem, we use only the second and third lags of the regressors as instruments and collapse the instrument matrix.

Following common practice, we also report the *p*-values of the Arellano and Bond (1991) test for second-order serial correlation (AR2), the *p*-values of the Hansen-J test of overidentifying restrictions (Hansen), and the number of instruments. The AR2 test indicates that the errors exhibit no second-order serial correlation, the Hansen test does not reject the validity of the instruments, and the number of instruments is always less than the number of cross-sectional units. We thus conclude that the models presented in columns 2 and 3 are correctly specified.

As before, the coefficient on log$EDUXT_{it}$ is positive and statistically significant in columns 2 and 3. While the estimated coefficient on log$EDUXT_{it}$ represents short-run effects, the long-run effects can be calculated by dividing the estimated short-run coefficients by one minus the coefficients on the lagged dependent variable. Thus, the estimates in column 2 imply a long-run elasticity of 0.993, and those in column 3 imply a long-run elasticity of 0.957. The former implies that, in the long run, a one-standard-deviation increase in log$EDUXT_{it}$ is associated with a 56.04 percent increase in the standard deviation of log$GERT_{it}$, and the latter implies a 54.00 percent increase in the standard deviation of log$GERT_{it}$ as a result of a one-standard-deviation increase in log$EDUXT_{it}$. These values are greater than, but still close to, the estimated magnitude (46.78 percent) implied by the coefficient in column 1 of Table 2.

A concern with the above results could be that they suffer from an 'included variable bias', where the estimated elasticities do not capture the effect of public tertiary education spending on the tertiary enrollment rate when ineffective public tertiary education expenditures increase the public costs or expenditures per student (as a proportion of GDP per capita). To address this concern, and to ensure that our results suffer neither from an omitted variables bias nor from an included variable bias, we reestimate our main models excluding log$COSTS_{it}$ as a control. The results from this robustness check are reported in Table 5.

[Table 5]

The estimated coefficients on log$EDUXT_{it}$ are still positive and highly significant. As expected, the estimated elasticities and their economic magnitudes are smaller than those in Tables 1 and 3. However, while the differences between the estimated coefficients on log$EDUXT_{it}$ in columns 1 and 2 of Table 5 (0.100 and 0.158, respectively), and those in column 1 of Table 2 (0.829) and column 1 of Table 4 (0.547), are very large in relative terms, the discrepancies in the estimated long-run elasticities implied by the results from columns 1 and 2 of Table 5—0.767 (0.283/(1-0.631)) and 0.603 (0.284/(1-0.529))—and the estimated long-run elasticities implied by the results from columns 2 and 3 of Table 4—0.993 and 0.957, respectively—are moderate.

Finally, we note that the estimated elasticities for public tertiary education spending in the regressions that do not control for endogeneity (in column 1 of Table 2 and in column 1 of Table 5) are smaller than the long-run elasticities for public tertiary education spending from the GMM regressions (in Tables 3 and 4), which do control for endogeneity. This difference might suggest that not controlling for endogeneity induces a downward bias in the estimated coefficient on $\log EDUXT_{it}$.

5. Conclusions

The question of the effectiveness of public spending on tertiary education is crucial not only for individual socioeconomic advancement but also for national economic development. Knowing whether public spending on tertiary education increases enrollment is essential for policymakers to optimize public investments.

In this study, we have conducted a systematic review of existing studies on the impact of public tertiary education spending on tertiary enrollment. Our review of the literature indicates that there are few studies examining this impact, and these studies report mixed results—ranging from significant positive relationships to both insignificant and significant negative relationships. Among the studies reviewed, there is no indication that public tertiary education spending has a robust positive effect on tertiary enrollment. However, our review has identified several shortcomings in these studies, of which some (such as the use of public expenditures on tertiary education per student as a measure of public expenditures on tertiary education, omitting public costs per student, ignoring potential endogeneity, not controlling for common time effects) may explain the failure to find a robust positive effect of public tertiary education spending on tertiary enrollment.

In this paper, we have reexamined the impact of public spending on tertiary education enrollment, addressing the shortcomings of prior studies. Using panel data from up to 149 countries between 1997 and 2018, we found that public spending on tertiary education increases the tertiary enrollment rate. To the best of our knowledge, this is the first study to provide robust empirical support for a positive impact of public spending on tertiary education on enrollment.

References

Appiah, E. N., & McMahon, W. W. (2002). The social outcomes of education and feedbacks on growth in Africa. *Journal of Development Studies*, *38*(4). https://doi.org/10.1080/00220380412331322411

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, *58*.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Bergh, A., & Fink, G. (2006). Higher Education: Does Public Expenditure Increase Enrollment? *Unpublished Working Paper*.

Bergh, A., & Fink, G. (2008). Higher education policy, enrollment, and income inequality. *Social Science Quarterly*, *89*(1). https://doi.org/10.1111/j.1540-6237.2008.00529.x

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143. https://doi.org/10.1016/S0304-4076(98)00009-8

Buckner, E., & Khoramshahi, C. (2021). Does the private sector expand access to higher education? A cross-national analysis, 1999-2017. *International Journal of Educational Development*, *84*. https://doi.org/10.1016/j.ijedudev.2021.102410

Colclough, C., & Al-Samarrai, S. (2000). Achieving schooling for all: Budgetary expenditures on education in sub-Saharan Africa and South Asia. *World Development*, *28*(11). https://doi.org/10.1016/S0305-750X(00)00065-6

Entorf, H. (1997). Random walks with drifts: Nonsense regression and spurious fixed-effect estimation. *Journal of Econometrics*, *80*(2). https://doi.org/10.1016/S0304-4076(97)00041-9

Hajebi, E., Billing, C., & Hajebi, M. (2023). The Effect of Government Expenditure on Education on the Enrollment Rate of Different Educational Levels in Selected OECD Countries. *International Journal of Scientific Research and Management*, *11*(05). https://doi.org/10.18535/ijsrm/v11i05.el03

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, *90*(1). https://doi.org/10.1016/S0304-4076(98)00023-2

Keller, K. R. I. (2006). Education expansion, expenditures per student and the effects on growth in Asia. *Global Economic Review*, *35*(1). https://doi.org/10.1080/12265080500537243

Wang, D., Wang, Z., & Zong, X. (2023). Impact of economic policy uncertainty on higher education expansion. *Finance Research Letters*, *52*. https://doi.org/10.1016/j.frl.2022.103564

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, *126*(1). https://doi.org/10.1016/j.jeconom.2004.02.005

Winter-Ebmer, R., & Wirz, A. (2002). Public Funding and Enrolment into Higher Education in Europe. *IZA Discussion Paper Series*, *503*.

Yang, L., & McCall, B. (2014). World education finance policies and higher education access: A statistical analysis of World Development Indicators for 86 countries. *International Journal of Educational Development*, *35*. https://doi.org/10.1016/j.ijedudev.2012.11.002

Yang, L., & St. John, E. P. (2023). Public Investment in Short-Cycle Tertiary Vocational Education: Historical, Longitudinal, and Fixed-Effects Analyses of Developed and Less-Developed Countries. *Education Sciences*, *13*(6). https://doi.org/10.3390/educsci13060573

Table 1. Summary of studies examining the impact of public tertiary expenditure on tertiary school enrollment

(1) Study [Focus]	(2) Dependent variable {Variable for tertiary education expenditures}	(3) Type of analysis and number of countries [Period]	(4) Main econometric methods	(5) Main results
1. Winter-Ebmer and Wirz (2002) [Impact of public tertiary expenditures on tertiary school enrollment]	Male gross tertiary enrollment rate, female gross tertiary enrollment rate (measured in logs) {Public tertiary education expenditures (measured in logs)}	Panel data analysis covering 14 European countries [1980–1996]	Fixed effects OLS and fixed effects IV models	- Insignificant coefficients for both the male and the female tertiary enrollment rates in fixed effects regressions. In the OLS regressions, the coefficients are positive, while in the IV regressions, they are negative.
2. Bergh and Fink (2006) [Impact of public tertiary expenditures on tertiary school enrollment]	Gross tertiary enrollment rate {Public tertiary education expenditures per student as a percentage of GDP per capita}	Cross-sectional analysis covering between 56 and 127 countries [2000]	OLS, IV	- Significant negative coefficients in most OLS regressions. - Negative but insignificant coefficients in all IV regressions.
3. Bergh and Fink (2008) [Impact of public tertiary expenditures on tertiary school enrollment and income inequality]	Change in the gross tertiary school enrollment rate over different periods between 1970 and 2000 {Public tertiary education expenditures per student at the beginning of each period}	Cross-sectional analysis covering between 45 and 64 countries [1970–2000]	OLS	- Both statistically significant negative coefficients, and coefficients that are both positive and negative, but not statistically significant.
4. Yang and McCall (2014) [Impact of public tertiary expenditures on tertiary school enrollment]	Gross tertiary school enrollment rate {Public tertiary education expenditures per student (measured in logs)}	Panel data analysis covering 86 countries [1998–2009]	Random and fixed-effects OLS models	- Significant negative coefficients. - The negative coefficient becomes larger (in absolute value) for developed countries.
5. Buckner and Khoramshahi (2021) [Impact of tertiary school enrollment in the private sector on overall tertiary enrollment]	Gross tertiary enrollment rate {Public tertiary education expenditures as a percentage of GDP}	Panel data analysis covering 122 countries [1999–2017]	Fixed-effects OLS models (*without time effects*)	- Both statistically significant positive coefficients and positive coefficients that are not statistically significant.
6. Yang and St. John (2023) [Impact of public investment in tertiary education on short-cycle tertiary enrollment]	Enrollment (headcount) in short-cycle tertiary education {Public tertiary education expenditures per student as a percentage of GDP per capita}	Panel data analysis covering 67 countries [2000–2018]	Fixed-effects OLS models	- Significant negative coefficients in two regressions.
7. Hajebi et al. (2023) [Impact of public tertiary expenditures on tertiary school enrollment]	Gross tertiary school enrollment rate (measured in logs) {Public tertiary education expenditures as a percentage of GDP (measured in logs)}	Panel data analysis covering 17 OECD countries [2010–2019]	Random effects OLS model (*without time effects*)	- Significant positive coefficient.

(1) Study [Focus]	(2) Dependent variable {Variable for tertiary education expenditures}	(3) Type of analysis and number of countries [Period]	(4) Main econometric methods	(5) Main results
8. Wang et al. (2023) [Impact of economic policy uncertainty on tertiary school enrollment]	Gross tertiary school enrollment rate {Public tertiary education expenditures as a percentage of GDP}	Panel data analysis covering 18 developed and developing countries [1998–2019]	System GMM (*without time effects*)	- Significant positive coefficient in the total sample. - When the sample is subdivided into developed and developing countries, the coefficient becomes negative and insignificant in both subsamples.

Notes: The instruments used by Winter-Ebmer and Wirz (2002) include the sizes of the male and female populations aged 18-24 years, as well as indicators of government form and ideology. The instruments used by Bergh and Fink (2006) include public per capita expenditures on secondary education, in addition to GDP per capita and the percentage of the population younger than 15 years. It is unclear whether the study by Wang et al. (2023) accounts for the endogeneity of tertiary education expenditure (% of GDP) in the GMM estimations. Short-cycle tertiary education refers to education below the level of a bachelor's degree or its equivalent.

Table 2. Baseline results

	(1) 3-year averages	(2) Winsorized data	(3) Annual data	(4) 5-year averages
log$EDUXT_{it}$	0.829***	0.799***	0.609***	0.781***
	(0.062)	(0.063)	(0.164)	(0.064)
log$COSTS_{it}$	-0.852***	-0.805***	-0.607***	-0.773***
	(0.061)	(0.066)	(0.180)	(0.060)
log$TERTPOP_{it}$	-0.831***	-0.779***	-0.636***	-0.856***
	(0.073)	(0.080)	(0.170)	(0.095)
log$GDPPC_{it}$	0.076*	0.117**	0.222**	0.128*
	(0.044)	(0.049)	(0.102)	(0.075)
log$URBAN_{it}$	0.177	0.184	0.234	0.118
	(0.154)	(0.159)	(0.225)	(0.192)
log$GERS_{it}$	0.074	0.121	0.288*	0.165**
	(0.074)	(0.077)	(0.151)	(0.073)
R-squared	0.942	0.934	0.886	0.942
No. of obs.	577	577	1,172	388
No. of countries	149	149	144	149

Notes: The dependent variable is log$GERT_{it}$. All regressions include both country and time fixed effects. Since some countries (France, Liberia, Angola, Tonga, and Palau) have only single observations for different years within our averaged periods, these countries are missing from the sample with annual observations. Robust standard errors are in parentheses. *** (**) [*] indicates significance at the 1% (5%) [10%] level.

Table 3. Results for different groups of countries

	(1) Low-income countries	(2) Lower-middle-income countries	(3) Upper-middle-income countries	(4) High-income countries
log$EDUXT_{it}$	0.678***	0.834***	0.994***	0.854***
	(0.116)	(0.112)	(0.055)	(0.095)
log$COSTS_{it}$	-0.708***	-0.890***	-0.998***	-0.829***
	(0.090)	(0.119)	(0.054)	(0.113)
log$TERTPOP_{it}$	-0.090	-0.723***	-1.066***	-0.934***
	(0.282)	(0.140)	(0.083)	(0.105)
log$GDPPC_{it}$	0.275*	0.132	-0.065	-0.010
	(0.140)	(0.093)	(0.047)	(0.064)
log$URBAN_{it}$	0.275	0.250	-0.151	-0.395
	(0.265)	(0.396)	(0.170)	(0.347)
log$GERS_{it}$	0.219*	-0.111	0.260***	0.072
	(0.103)	(0.110)	(0.088)	(0.071)
R-squared	0.964	0.918	0.985	0.937
No. of obs.	124	168	100	185
No. of countries	37	44	28	40

Notes: The dependent variable is log$GERT_{it}$. All regressions include both country and time fixed effects. Robust standard errors are in parentheses. *** [*] indicates significance at the 1% [10%] level.

Table 4. Results with lagged variables and GMM results

	(1) FE model with one-period lags	(3) GMM	(4) GMM
log$GER_{it\text{-}1}$		0.191***	0.123*
		(0.070)	(0.069)
log$EDUX_{it}$	0.547***	0.803***	0.839***
	(0.070)	(0.086)	(0.090)
log$COSTS_{it}$	-0.471***	-0.717***	-0.732***
	(0.068)	(0.120)	(0.099)
log$TERTPO_{it}$	-0.520***	-1.016***	-1.000***
	(0.144)	(0.107)	(0.091)
log$GDPPC_{it}$	0.149	-0.028	-0.055
	(0.129)	(0.047)	(0.041)
log$URBAN_{it}$	0.331	-0.246	-0.034
	(0.378)	(0.160)	(0.116)
log$GERS_{it}$	0.340**	0.291	0.314**
	(0.135)	(0.251)	(0.148)
Variables treated as exogenous		Time dummies	Time dummies
Variables treated as predetermined		log$GER_{it\text{-}1}$, log$COSTS_{it}$, log$TERTPOP_{it}$, log$GDPPC_{it}$, log$URBAN_{it}$, log$GERS_{it}$	log$GER_{it\text{-}1}$
Variables treated as endogenous		log$EDUXT_{it}$	log$EDUXT_{it}$, log$COSTS_{it}$, log$TERTPOP_{it}$, log$GDPPC_{it}$, log$URBAN_{it}$, log$GERS_{it}$
AR2 (*p*-value)		0.727	0.832
Hansen (*p*-value)		0.361	0.828
No. of instruments		41	41
R-squared	0.818		
No. of obs.	461	478	478
No. of countries	135	138	138

Notes: The dependent variable is log$GERT_{it}$. All specifications control for both country and time fixed effects. AR2 is the Arellano-Bond test for second-order autocorrelation in differenced residuals. Hansen is the Hansen test of overidentifying restrictions. Robust standard errors are in parentheses. *** (**) [*] indicates significance at the 1% (5%) [10%] level.

Table 5. Results for regressions without public costs per student

	(1) FE baseline model	(2) FE model with one-period lags	(3) GMM	(4) GMM
log$GERT_{it-1}$			0.631***	0.529***
			(0.162)	(0.151)
log$EDUXT_{it}$	0.100**	0.158***	0.283***	0.284***
	(0.050)	(0.051)	(0.103)	(0.093)
log$TERTPOP_{it}$	-0.095	-0.108	-0.595***	-0.509***
	(0.151)	(0.166)	(0.188)	(0.168)
log$GDPPC_{it}$	0.359***	0.416***	-0.180***	-0.140***
	(0.127)	(0.136)	(0.054)	(0.052)
log$URBAN_{it}$	0.372	0.585	-0.098	-0.035
	(0.424)	(0.453)	(0.224)	(0.178)
log$GERS_{it}$	0.796***	0.728***	0.937***	0.975***
	(0.085)	(0.126)	(0.306)	(0.294)
Variables treated as exogenous			Time dummies	Time dummies
Variables treated as predetermined			logGER_{it-1}, log$TERTPOP_{it}$, log$GDPPC_{it}$, log$URBAN_{it}$, log$GERS_{it}$	logGER_{it-1}
Variables treated as endogenous			log$EDUXT_{it}$	log$EDUXT_{it}$, log$TERTPOP_{it}$, log$GDPPC_{it}$, log$URBAN_{it}$, log$GERS_{it}$
AR2 (*p*-value)			0.153	0.666
Hansen (*p*-value)			0.241	0.227
No. of instruments			36	36
R-squared	0.742	0.763		
No. of obs.	628	499	510	510
No. of countries	149	135	138	138

Notes: The dependent variable is log$GERT_{it}$. All specifications control for both country and time fixed effects. AR2 is the Arellano-Bond test for second-order autocorrelation in differenced residuals. Hansen is the Hansen test of overidentifying restrictions. Robust standard errors are in parentheses. *** (**) indicates significance at the 1% (5%) level.

Appendix

Table A1. Variable definitions and data sources

Variable	Source
GERT is the gross enrollment rate in tertiary education, which is the number of people enrolled in tertiary education as a percentage of the total population of tertiary age (i.e., the total number of people aged 18 to 24).	World Development Indicators (available at https://databank.worldbank.org/source/world-development-indicators, last accessed in June 2024)
EDUXT denotes public tertiary education expenditures as a percentage of GDP.	UNESCO (available at http://data.uis.unesco.org/, last accessed in June 2024)
COSTS denotes the public costs per student as a proportion of GDP per capita, measured by public tertiary education expenditures per student as a percentage of GDP per capita.	World Development Indicators (available at https://databank.worldbank.org/source/world-development-indicators, last accessed in June 2024)
TERTPOP denotes the percentage of the population of tertiary age.	UNESCO (available at http://data.uis.unesco.org/, last accessed in June 2024)
GDPPC represents real GDP per capita, measured in 2017 PPP dollars.	World Development Indicators (available at https://databank.worldbank.org/source/world-development-indicators, last accessed in June 2024)
URBAN is the urbanization rate, measured by the percentage of the population living in urban areas with one million or more people.	World Development Indicators (available at https://databank.worldbank.org/source/world-development-indicators, last accessed in June 2024)
GERS is the gross enrollment rate in secondary education, which is the number of students enrolled in secondary schools as a percentage of the total population of secondary age.	World Development Indicators (available at https://databank.worldbank.org/source/world-development-indicators, last accessed in June 2024)

Table A2. Countries in the sample and their classification

Country		Country		Country		Country	
Afghanistan	LIC	Czech Republic	HIC	Latvia	HIC	San Marino	HIC
Albania	LMIC	Dem. Rep. of the Congo	LMIC	Lesotho	LMIC	Sao Tome and Principe	LIC
Andorra	HIC	Denmark	HIC	Liberia	LIC	Senegal	LIC
Angola	LMIC	Ecuador	LMIC	Lithuania	UMIC	Serbia	UMIC
Antigua and Barbuda	UMIC	El Salvador	LMIC	Luxembourg	HIC	Seychelles	UMIC
Argentina	UMIC	Eritrea	LIC	Madagascar	LIC	Singapore	HIC
Armenia	LMIC	Estonia	HIC	Malawi	LIC	Slovakia	HIC
Australia	HIC	Eswatini	LMIC	Malaysia	UMIC	Slovenia	HIC
Austria	HIC	Ethiopia	LIC	Mali	LIC	South Africa	UMIC
Azerbaijan	UMIC	Fiji	LMIC	Malta	HIC	Spain	HIC
Bahrain	HIC	Finland	HIC	Marshall Islands	LMIC	Sri Lanka	LMIC
Bangladesh	LIC	France	HIC	Mauritania	LIC	Sweden	HIC
Barbados	HIC	Gambia	LIC	Mauritius	UMIC	Switzerland	HIC
Belarus	LMIC	Georgia	LMIC	Mexico	UMIC	Syria	LMIC
Belgium	HIC	Germany	HIC	Moldova	LMIC	Tajikistan	LIC
Belize	LMIC	Ghana	LIC	Mongolia	LMIC	Tanzania	LIC
Benin	LIC	Greece	HIC	Morocco	LMIC	Thailand	LMIC
Bhutan	LIC	Grenada	UMIC	Mozambique	LIC	Timor-Leste	LMIC
Bosnia and Herzegovina	UMIC	Guatemala	LMIC	Myanmar	LIC	Togo	LIC
Botswana	UMIC	Guinea	LIC	Namibia	LMIC	Tonga	LMIC
Brazil	UMIC	Guyana	LMIC	Nepal	LIC	Tunisia	LMIC
Brunei Darussalam	HIC	Honduras	LMIC	Netherlands	HIC	Turkey	UMIC
Bulgaria	LMIC	Hong Kong, China	HIC	Niger	LIC	Uganda	LIC
Burkina Faso	LIC	Hungary	UMIC	North Macedonia	LMIC	Ukraine	LMIC
Burundi	LIC	Iceland	HIC	Norway	HIC	United Kingdom	HIC
Cabo Verde	LMIC	India	LIC	Oman	UMIC	United States	HIC
Cameroon	LMIC	Indonesia	LMIC	Pakistan	LIC	Uruguay	UMIC
Canada	HIC	Iran	LMIC	Palau	UMIC	Vanuatu	LMIC
Central African Rep.	LIC	Ireland	HIC	Panama	UMIC	Zimbabwe	LIC
Chad	LIC	Israel	HIC	Paraguay	LMIC		
Chile	UMIC	Italy	HIC	Peru	LMIC		
China	UMIC	Jamaica	LMIC	Philippines	LMIC		
Colombia	LMIC	Japan	HIC	Poland	HIC		
Comoros	LIC	Jordan	LMIC	Portugal	HIC		
Congo (Rep. of the)	LIC	Kazakhstan	UMIC	Romania	LMIC		
Costa Rica	UMIC	Kenya	LIC	Russian Federation	UMIC		
Côte d'Ivoire	LMIC	Korea (Rep. of)	HIC	Rwanda	LIC		
Croatia	UMIC	Kuwait	HIC	Saint Kitts and Nevis	UMIC		
Cuba	LMIC	Kyrgyzstan	LIC	Saint Lucia	UMIC		
Cyprus	HIC	Lao P.D.R.	LIC	Samoa	LMIC		

Table A3. Summary statistics based on three-year averages

Variable	Obs.	Mean	Std. dev.	Min	Max
GERT	577	38.093	27.042	0.294	132.592
EDUXT	577	0.935	0.536	0.070	4.318
COSTS	577	80.431	192.085	3.974	2489.753
TERTPOP	577	8.314	1.758	4.116	12.645
GDPPC	577	14190.000	18182.770	253.637	106833.100
URBAN	577	58.562	22.925	8.463	100.000
GERS	577	81.594	29.625	5.435	161.460
log*GERT*	577	3.214636	1.118629	-1.22312	4.887278
log*EDUXT*	577	-0.2434	0.631244	-2.65555	1.462772
log*COSTS*	577	3.685336	0.970928	1.379855	7.819939
log*TERTPOP*	577	2.094003	0.223184	1.414898	2.537232
log*GDPPC*	577	8.671948	1.465822	5.535902	11.57902
log*URBAN*	577	3.967141	0.497176	2.135704	4.60517
log*GERS*	577	4.29951	0.520257	1.69294	5.084254