Factors Affecting the Educational Budget in the United States

Andrew Du, Jalal Haider, Stella Liu

May 14, 2024

Spring 2024: Econometrics

Professor Sebastiano Manzan

NYU Stern School of Business

I. Introduction

Education spending includes the allocation of financial resources dedicated to the educational systems by governments. It includes investments in infrastructure like school buildings, teacher salaries, and learning materials along with administrative services aimed at improving educational quality and equity across diverse populations. The public has paid close attention to the topic of education spending because it serves as the basis for a promising future for US citizens. Sufficient resources for various educational levels from primary to higher education play a crucial role in shaping educational outcomes. It is a crucial area of public policy, influencing not only the quality of education provided but also broader economic and social outcomes. Our research aims to explore the drivers of education spending, looking at factors such as population age, urbanization, labor force participation, and GDP.

The most important factor we consider is the population age between states. The American population today is much older than previous generations and is still growing older. By 2030, the percentage of the population aged 65 or older will increase to over 20%. This aging population has consequences for the government's spending and budget. Many studies have found that as the population grows older, there will be less support for education spending. To confirm the impact, our research aims to explore this relationship more precisely by adding other controls and variables that might have an impact on education spending.

Our data comes primarily from the National Center for Education Statistics, which contains information on education spending differences between states. We supplemented the data with urbanization data from the World Bank, age data from the US Census Bureau, GDP, and Labor force data from Fred. Together, with the data collected, we wish to determine important factors that have an impact on education spending. We found that the education budget as a percentage of GDP is negatively associated with changes in urbanization and labor force participation as well as the change in over 65s 6 years ago. Additionally, it is positively associated with the percentage last year but negatively with the percentage two years ago. Following this introduction, this paper will include a literature review that can provide a basic understanding of the topic, the model and data we utilized, a review of the empirical application of our findings, and lastly a conclusion that summarizes this paper.

II. Survey of the Literature

Our primary inspiration was Harris's (2001) research, "Education Spending in an Aging America". The research utilized a national panel of public school districts to study the impact of an aging population on public education spending. The thesis revolves around how an aging population can increase competition for public funds between pensions and healthcare for the elderly, and education for the younger population. They found that education spending is less sensitive to the elderly population in a school district than total spending. But overall, in all regions, there is a negative elasticity between the aging population and education spending with relatively minor effects. They consider variables such as population age, race, wealth, and federal revenue.

Another relevant paper is Grob's (2007) "Demographic Change and Public Education Spending: A Conflict between Young and Old". The research focused on the conflict in Sweden. It included dependent variables such as unemployment rate, degree of urbanity, share of homeowners, and strength of governing parties. However, these variables were all insignificant. Only variables like the number of people, number of retired, and per-capita income were significant in their findings. Overall, this paper also found that an aging population significantly reduces education budgets, potentially forcing steeper cuts in spending.

Lastly, we looked at the research "Effect of Education Expenditure on Per Capita GDP in Developing Countries" (2017). This paper explores the impact of education expenditure on economic growth in developing countries. The study uses a dynamic panel model and the 'system' General Method of Moments (GMM) estimator to estimate the effect of increased spending on education on per capita GDP. The findings indicate that an expansion in education expenditure in developing countries affects per capita GDP positively. The study also highlights the importance of education in promoting economic growth and emphasizes the need for quality education as a means to improve labor productivity and, ultimately, per capita GDP. So there might be a simultaneous causality bias between education spending and GDP. Nevertheless, we believe that GDP is a relevant factor in our study as the causation plays out over the longer term.

III. Model

We tested the following hypotheses:

1. Is educational expenditure as a percentage of GDP affected by changes in GDP per capita, changes in urbanization, changes in the senior population, and changes in labor force participation (LFP)?
    a. H_0: the changes in these independent variables do not affect the educational expenditure budget
    b. H_a: at least one of these changes does affect the educational expenditure budget
    c. Model 1:
    $$edexp/GDP \;=\; \beta_0 + \beta_1 \Delta(urban\%)_{detr} + \beta_2 \Delta(\%\ over\ 85) \;+ \beta_3 \Delta(LFP\%)$$
    d. Model 2:

$$edexp/GDP \;=\; \beta_0 + \beta_2 \Delta(GDP/cap)_{detr} \;+\; \beta_2 \Delta(urban\%)_{detr} + \beta_3 \Delta(\%\ over\ 85) \;+ \beta_4 \Delta(LFP\%)$$

2. Do past values of educational expenditure as a percentage of GDP, past changes in GDP/capita, and past changes in urbanization predict and/or affect the current educational expenditure as a percentage of GDP?
    a. H_0: none of the lags of these independent variables or the endogenous variable predict the educational expenditure budget %
    b. H_a: at least one of the lags of these independent variables or the endogenous variable predicts the educational expenditure budget %
    c. ARDL(p, q) Model: $Y_t \;=\; \beta_0 \;+\; \sum_{i=1}^{p} \beta_{yi} Y_{t-i} \;+\; \sum_{j=1}^{q} \beta_{xj} X_{t-j}$
3. Is the ***change*** in edexp/GDP affected by changes in GDP per capita, changes in urbanization, aging, change in labor force participation? (panel data)
    a. And as a corollary, is the edexp/GDP affected by current GDP per capita, current urbanization, current age proportions, and current labor force participation after controlling for omitted variables?
    b. H_0: None of these independent variables predict the change in edexp/GDP
    c. H_a: One or more of these independent variables predict the change in edexp/GDP

$$\Delta(Ed/GDP) \;=\; \beta_0 \;+ \beta_1 g_{GDP} + \beta_2 \Delta medage \;+\; \beta_3\, \Delta prop_{retired} + \beta_4\, \Delta LFP \;+ \beta_5\, \Delta prop_{urban}$$

IV. Data

For the first dataset, we found time series for the United States. The main data source for this project was IES NCES, the National Center for Education Statistics. Specifically, we used the datasets 106.60 for Gross Domestic Product, State and Local Direct General Expenditures, and US Population, 1970-2019 (T = 49). We took the growth rate of GDP per capita because it grows exponentially and the variation is proportional to the value. Similarly, educational expenditure per capita tends to increase over time and vary with the value. For the explanatory variables, we found the US median age from Our World in Data, the proportion of the population over 65 and 85 from the Federal Interagency Forum on Aging-Related Statistics, the percentage of the population that was urban using data from the World Bank, and the Labor Force Participation from the Federal Reserve. We transformed the Labor Force Participation rate from the original monthly data to annual averages using the Fed's options. We then took the first difference of median age, urban population share, the population over 65, the population over 85, and the labor force participation rate to match the growth rate and render the time series stationary. Finally, we inspected any series that had an evident trend and removed the trend.

We also used a two-period panel (N = 50, T = 2) dataset on educational expenditures to test additional hypotheses, mostly derived from Table 106.40 from NCES. Specifically, we used the data in the panel together with CPI data from the St. Louis Fed. Since all real educational expenditures were in 2021-2022 terms, we adjusted all GDP numbers to mid-year 2021-2022 terms. We obtained GDP data from 2018, 2019, and 2020 and averaged to school-year terms to match the educational expenditures in this data set. We also obtained a state-by-state median age dataset from the US Census Bureau and processed it similarly. We also took 1/10 of the change in the percentage of over 65s from 2010-2020 as a proxy. We obtained the actual change in the labor force participation rate based on labor force participation data from the St. Louis Fed. Finally, we calculated 1/10 the change in the urbanization index from 2010-2020 as a proxy, getting data in both years from the US Bureau of Economic Analysis.

Note that for both datasets, we converted all numbers to decimal, including percent as p/100.

V. Empirical Application

There are two main regressions that we carried out to check the relationship between the dependent variable of Education Expenditure and the other independent variables. The data taken was first cleaned into CSV files and then the OLS regressions were taken as followed from Fig V.I and V.II.

Fig V.I

OLS Regression Results

Dep. Variable:	edexp_gdp	R-squared:	0.576
Model:	OLS	Adj. R-squared:	0.548
Method:	Least Squares	F-statistic:	27.04
Date:	Tue, 14 May 2024	Prob (F-statistic):	3.75e-10
Time:	01:10:38	Log-Likelihood:	230.62
No. Observations:	49	AIC:	-453.2
Df Residuals:	45	BIC:	-445.7
Df Model:	3		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	0.0506	0.001	54.379	0.000	0.049	0.052
dtr_durban	-0.7743	0.229	-3.376	0.001	-1.224	-0.325
d_over_85	2.9963	3.179	0.943	0.346	-3.234	9.226
d_LFP	-0.6764	0.086	-7.901	0.000	-0.844	-0.509

Omnibus:	2.914	Durbin-Watson:	0.566
Prob(Omnibus):	0.233	Jarque-Bera (JB):	2.736
Skew:	-0.516	Prob(JB):	0.255
Kurtosis:	2.477	Cond. No.	9.93e+03

In **fig V.I**, we find that the urban index as well as the labor force participation had effects on the Education expenditure. Here we find that a 1 percentage-point increase in urbanity % (urban index/100), decreases Education Expenditure as a % of GDP by 0.77 percentage points, holding all other variables constant ($p = 0.001$). A 1% increase in the labor force participation decreases Education Expenditure as a % of GDP by 0.67 percentage points, holding all other variables constant ($p < 0.001$). However, the percentage of over 85s did not have a statistically significant effect. There is a negatively skewed distribution, indicating more extreme observations on the left-hand side of the regression. The excess kurtosis of 2.417 indicates many more extreme values than that of a normal distribution.

Fig V.II

OLS Regression Results

Dep. Variable:	edexp_gdp	R-squared:	0.579
Model:	OLS	Adj. R-squared:	0.541
Method:	Least Squares	F-statistic:	21.78
Date:	Tue, 14 May 2024	Prob (F-statistic):	5.78e-10
Time:	02:14:18	Log-Likelihood:	230.81
No. Observations:	49	AIC:	-451.6
Df Residuals:	44	BIC:	-442.2
Df Model:	4		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	0.0517	0.003	20.593	0.000	0.047	0.057
dtr_g_gdp_cap	-0.0112	0.023	-0.487	0.626	-0.056	0.034
dtr_durban	-0.8432	0.278	-3.034	0.002	-1.388	-0.299
d_over_85	2.7720	3.206	0.864	0.387	-3.513	9.057
d_LFP	-0.6572	0.089	-7.412	0.000	-0.831	-0.483

Omnibus:	3.601	Durbin-Watson:	0.540
Prob(Omnibus):	0.165	Jarque-Bera (JB):	3.154
Skew:	-0.530	Prob(JB):	0.207
Kurtosis:	2.349	Cond. No.	1.00e+04

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)
[2] The condition number is large, 1e+04. This might indicate that there are strong multicollinearity or other numerical problems.

In **Fig V.II**, we find how the change in GDP per capita plus the aforementioned independent variables affect the education expenditure as a % of GDP per capita. The current de-trended change in GDP per capita as well as the change in elderly individuals did not have a significant effect on education expenditure as a % of GDP per capita. This is likely because the GDP per capita is typically measured a year later, at which point it is too late to budget. However, a 1 percentage-point increase in the urban % predicts a 0.8432 percentage-point decrease in education expenditures as % of GDP, holding all other variables constant. A 1 percent-point change in labor force participation predicts a 0.6572 percentage-point decrease in education expenditures as % of GDP, holding all other variables constant.

In this regression there is a slightly more negative skewness of -0.530, meaning there is an even longer left tail to the residuals. The kurtosis 2.349 is not very different from Fig V.I, indicating that there are slightly fewer extreme values than the normal distribution.

The differential F-statistic is $[(R^2_{unres} - R^2_{res})/q]/[(1-R^2_{unres})/(N-k-1)]$

$= [(0.579-0.576)/1]/[(1-0.579)/(49-4-1)] = 0.314 << 1.96^2$. This of course also agrees with the conclusion that dtr_g_gdp_cap, essentially a proxy for the business cycle, does not affect the educational expenditure % in the same year. This model also has an adjusted R^2 of 0.541, while

the model in Figure V.I has a higher adjusted R^2 of 0.548. These factors show that the model in Figure V.I works better.

$edexp/GDP = 0.0506 - 0.7743\Delta(urban\%)_{detr} - 0.06764\Delta(LFP\%)$.

Note that we allowed for heteroskedasticity in both of the models as using heteroskedastic standard errors does not decrease validity when the variables are homoskedastic.

Fig V.III

ARDL Model Results

Dep. Variable:	edexp_gdp	No. Observations:	49
Model:	ARDL(2, 1, 5, 6)	Log Likelihood	272.282
Method:	Conditional MLE	S.D. of innovations	0.001
Date:	Tue, 14 May 2024	AIC	-512.563
Time:	01:47:51	BIC	-482.961
Sample:	6	HQIC	-501.423
	49		

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0015	0.004	0.428	0.672	-0.006	0.009
edexp_gdp.L1	1.9328	0.210	9.208	0.000	1.505	2.360
edexp_gdp.L2	-1.0627	0.198	-5.380	0.000	-1.465	-0.660
dtr_g_gdp_cap.L1	0.0515	0.013	4.012	0.000	0.025	0.078
dtr_durban.L1	-0.0654	0.424	-0.154	0.878	-0.930	0.799
dtr_durban.L2	1.0422	0.659	1.581	0.124	-0.301	2.385
dtr_durban.L3	-1.1183	0.684	-1.634	0.112	-2.513	0.276
dtr_durban.L4	1.2405	0.669	1.855	0.073	-0.122	2.603
dtr_durban.L5	-0.8942	0.457	-1.958	0.059	-1.824	0.036
dtr_d_over_65.L1	0.2082	0.343	0.607	0.548	-0.490	0.906
dtr_d_over_65.L2	0.2395	0.399	0.601	0.552	-0.573	1.052
dtr_d_over_65.L3	0.0497	0.415	0.120	0.906	-0.796	0.896
dtr_d_over_65.L4	-0.7456	0.447	-1.666	0.105	-1.657	0.166
dtr_d_over_65.L5	0.5313	0.449	1.184	0.245	-0.383	1.445
dtr_d_over_65.L6	-0.7122	0.340	-2.092	0.044	-1.406	-0.019

In Fig. V.III, we display the ARDL model with the lowest Bayesian Information Criterion of -482.961, and thus the best fit, out of multiple models we tried.

Based on the statistically significant coefficients, the model predicts that:

$$Y_t = 1.9328Y_{t-1} - 1.0627Y_{t-2} + 0.0515X_{1,t-1} - 0.7122X_{2,t-6}$$

X_1 = detrended growth in GDP per capita, X_2 = detrended change in % of population over 65

The model predicts that a 1 percentage-point increase in education spending as a % of GDP last year predicts a 1.9328 percentage-point increase in the quantity this year, and a 1 percentage-point increase two years ago predicts a 1.0627 percentage-point decrease this year, holding all other variables constant. This model also predicts that holding all other variables

constant, a 1 percentage-point increase in the growth of GDP per capita last year is associated with a 0.0515 percentage-point education-budget increase (a very slight but still statistically significant effect), and a 1 percentage-point excess increase of the proportion of seniors six years ago is associated with a 0.71 percentage-point decrease in the education budget this year. Though the ARDL model itself does not establish causation, we find causation plausible as economic booms can lead to optimism and high spending even on a percent basis, including on education. Additionally, the people who are now over 71 will likely start drawing from government spending to support living and health expenses.

Figure V.IV

OLS Regression Results

Dep. Variable:	d_ed_gdp	R-squared:	0.149
Model:	OLS	Adj. R-squared:	0.053
Method:	Least Squares	F-statistic:	1.661
Date:	Tue, 14 May 2024	Prob (F-statistic):	0.164
Time:	01:10:46	Log-Likelihood:	313.07
No. Observations:	50	AIC:	-614.1
Df Residuals:	44	BIC:	-602.7
Df Model:	5		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-2.988e-05	0.000	-0.119	0.905	-0.001	0.000
g_gdp	0.0099	0.006	1.705	0.088	-0.001	0.021
d_medage	-0.0002	0.000	-0.472	0.637	-0.001	0.000
d_retired	-0.0358	0.041	-0.880	0.379	-0.116	0.044
d_lfp	-0.0003	0.018	-0.018	0.986	-0.035	0.034
d_urban	0.0070	0.086	0.081	0.935	-0.161	0.175

Omnibus:	11.797	Durbin-Watson:	1.717
Prob(Omnibus):	0.003	Jarque-Bera (JB):	22.541
Skew:	-0.562	Prob(JB):	1.27e-05
Kurtosis:	6.091	Cond. No.	893.

As for the two-period panel data difference regression in **Fig. V.IV**, we did not find any statistically significant coefficients, even the intercept. We tried the growth rate in education as a dependent variable as well, but even this did not give statistically significant results. We do not find this credible as the sample is very restricted in time. This could be due to the limited data in two periods and fifty states (total = 100) as well as internal correlation between states.

VI. Conclusion

Based on our findings, we reject our null hypothesis for both the simultaneous regression and ARDL on the time series data, but we fail to reject the null hypothesis based on panel data. We can conclude that the empirical evidence suggests that a few factors are associated with the education budget as a percentage of GDP:

1. The change in urbanization is negatively associated with the education budget %.
    a. This result is highly statistically significant in a counterintuitive direction.
    b. This is likely because urban jobs pay better, and spending on education only increases the efficacy of education up to a certain extent. Therefore, the spending as % of income and thus production must decrease.
2. Changes in labor force participation are negatively associated with the education budget %.
    a. This is again likely because of the fixed-expense effect. As labor force participation increases, there is more income per capita, but the efficacy of spending on education diminishes with the amount spent.
    b. Therefore, the % of optimal spending of GDP on education decreases.
3. The % of GDP dedicated to education is positively associated with the amount *last year* at a beta coefficient greater than 1. This implies that above average % of GDP spent last year will lead to an even more extreme percentage this year. Yet the lag-2 coefficient is negative, suggesting that if an above average % of GDP was spent on education two years ago, a below-average % of GDP will be spent on education this year.
4. The % of GDP dedicated to education is positively associated with the change in GDP per capita in excess of the trend last year. This suggests that economic upswings tend to increase the percentage of GDP allocated to education, but only slightly.
5. It takes six lags for the change in the senior (>65) population to have a statistically significant negative effect on the % of GDP spent on education. Furthermore, the change in the >85 population did not have a significant effect on the % of GDP spent on education This suggests that senior expenses only divert funds from education spending at about 71 years, confirming the results in prior papers.

Overall, note that the stationarized time series of education expenditures as a percentage of GDP may decrease for two reasons – increasing GDP or decreasing expenditure on education.

VII. Bibliography

Data Sources:

1. United States Census Bureau: Media Age of US population, via https://data.census.gov/table/ACSST1Y2017.S0101?q=state%20median%20age&g=
2. National Center of Education Statistics: Education Data, via https://nces.ed.gov/programs/digest/current_tables.asp
3. Bureau of Economic Analysis: GDP, via https://www.bea.gov/news/2020/gross-domestic-product-state-4th-quarter-and-annual-2019 https://www.bea.gov/news/2021/gross-domestic-product-state-4th-quarter-2020-and-annual-2020-preliminary
4. World Bank: Urbanization Population %, via https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=US
5. Fred: CPI Adjustments, via https://fred.stlouisfed.org/series/CPIAUCSL
6. Fred, BLS: Labor Force Participation, via https://fred.stlouisfed.org/series/CIVPART#, https://fred.stlouisfed.org/release/tables?eid=784070&rid=446, https://www.bls.gov/opub/ted/2023/labor-force-participation-rate-for-people-ages-25-to-54-in-may-2023-highest-since-january-2007.htm,
7. Visual Capitalist: Urbanization Index, via https://www.visualcapitalist.com/sp/mapping-us-urbanization-by-state/
8. Federal Interagency Forum on Aging Statistics, KFF: Aging Statistics, via

https://agingstats.gov/data.html,

https://www.kff.org/other/state-indicator/distribution-by-age/

Works Cited:

Appiah, Elizabeth. “The Effect of Education Expenditure on Per Capita GDP in Developing Countries.” *International Journal of Economics and Finance*, vol. 9, no. 10. *Canadian Center of Science and Education*.

Grob, Ueli, and Stefan Wolter. “Demographic Change and Public Education Spending: A Conflict between Young and Old? Ueli Grob &Stefan C. Wolter.” *Education Economics*, vol. 15, no. 3, 2007, pp. 277-292, https://www.tandfonline.com/doi/abs/10.1080/09645290701273467.

Harris, Amy, et al. “Education spending in an aging America.” *Journal of Public Economics*, vol. 81, no. 3, 2001, pp. 449-472, https://www.sciencedirect.com/science/article/abs/pii/S004727270000133X#BIB18.

“Largest Year-to-Year Increase for Public School Spending Per Pupil.” *Census Bureau*, 25 April 2024, https://www.census.gov/newsroom/press-releases/2024/public-school-spending-per-pupil.html. Accessed 13 May 2024.