Government Revenue as an Effective Measure to Increase Access to Education

By:
Hao Lin Alexander Chen
1007801130
haolin.chen@mail.utoronto.ca

Unurjargal Batsuuri
1007766703
unu.batsuuri@mail.utoronto.ca

ECO375: Applied Econometrics
University of Toronto
Department of Economics
11/21/2023

Abstract

Public education is currently the primary avenue to increase access to education for those who may not be able to otherwise. One proposed method to increase access to education is to increase government revenue, thereby increasing public education expenditure (UNESCO, 2014). To find whether this method is effective, we estimate the size of the robust and unbiased effect of total government revenue on public education expenditure. This is done with time/state fixed effects, the integration of possible confounding factors, and further explanatory variables. Our model predicts an overly small portion of additional government revenue being directed toward education, thus being inefficient, or even outright ineffective in increasing accessibility. This analysis provides statistical evidence against the aforementioned method and grounds for restructuring how access to education can be provided. However, the model is still limited by further potentially omitted factors, lack of historical data, and the inability to make a causal inference. In addition, the effectiveness in increasing academic results of such funding into public education will not be discussed, just access in and of itself.

Introduction

Access to education demonstratively increases the likelihood of an individual's success in modern society, as seen clearly with reduced average unemployment rates (BLS, 2023) and higher average wages for those with higher education attainment (NCES, 2021). Given such advantages, increasing access to education seems to be an important goal, especially in the United States. As such, it is paramount to consider the effectiveness of the U.S.' current method of providing education to those who may not be affordable otherwise: public education. Given the importance of access to education, the government should treat education expenditure as a major priority. This entails that an increase in total government revenue must also significantly increase education expenditure. If not, perhaps our primary method of providing opportunities for education for low-income households may be ineffective, inefficient, and requires reform. As the interpretations of our coefficients are along the vein of "On average, a $1000 increase in government revenue increases education expenditure by $X, we can interpret these coefficients as marginal propensities. Specifically, the marginal propensity to spend on education, given total government revenue. This allows for an assessment of the effectiveness and efficiency of this method.

Context and Data

Our data is drawn from Kaplan's (2018) review of government finances, where revenues and expenditures of each U.S. state from 1992-2018 are collected in a panel structure. This contains 1350 data points, for 50 states in each of the 27 years. The unit of observation is state-year, with our variables of interest being total government revenue (total revenue) and education expenditure, with our raw data both in thousands of U.S. dollars. We will present these figures in writing with singular dollar units. The mean total revenue is 31.4 billion USD, and the mean education expenditure is 9.1 billion USD. The standard deviation of government revenue is 42.2 billion USD, for the range of 1.75 billion to 412 billion USD, indicating a large difference in available funds for each state. This large variance extends towards education expenditure, with a standard deviation of 11.3 billion USD, for the range of 430 million USD to 10.2 billion USD. Further summary statistics are presented in Table 1. These variances can also be explained by the lack of inflation correction within our base dataset, as our focus is on the ratio between revenue and expenditure rather than the numbers' sizes themselves. The difference in means seems to suggest around 28% average allocation of total revenue to education expenditure, alluding to an economically significant result. This, however, is a biased assumption.

Regression Analysis

Base Case

We establish a simple linear regression between total government revenue and education expenditure to establish a baseline coefficient. This is shown in Table 2, specification (1). Our simple linear regression is modeled as:

$$Education\ Expenditure_{s,t} = \beta_0 + \beta_1 Total\ Revenue_{s,t} + e_{s,t}$$

Where e is the error term, s is state, and t is time in years since 1992. Our null hypothesis is that changes in total revenue have no explanatory effect on education expenditure. Our alternative hypothesis is that it does. We use robust standard errors here to account for heteroscedasticity. This specification has a coefficient of 0.259, interpreted as on average, an increase of 1000 USD in total revenue generates 259 dollars of increase in education expenditure. This result is highly statistically significant, even at a 0.1% significance level, with a t-statistic of 50.58. As such, we can easily reject the null hypothesis that TR has no influence on EE, an expected result. However, the true importance lies in the size of the coefficient and its 95% confidence interval, which is between 0.249 and 0.270. A coefficient of this size seems economically significant, as it suggests that the marginal propensity to spend on education is around 0.26. Given a government's other responsibilities, such as public welfare and infrastructure, the marginal propensity being more than a quarter demonstrates that increasing access to education is indeed a priority and that the method in question is effective.

However, given that our data is collected in a field setting and is panel data, it may not adhere to LSA2. Panel data contains repeated observations on the same entity, and our data was not collected in a random experiment setting. As such, the data cannot be guaranteed to be independent and identically distributed. LSA 1.1, the homoscedasticity assumption, is also under threat. However, this is rather inconsequential as we merely need some estimation of the effect for our purposes, rather than the best as the Guass-Markov theorem prescribes. This is further mitigated with our use of robust standard errors. Our errors may not be normally distributed either, a violation of LSA 1.2, but this hindrance can be mitigated by our large sample size. Furthermore, given the straightforwardness of a simple linear regression, a glaring issue arises with omitted variable bias (OVB). This is in violation of LSA1, and our current model's conditional distribution of the error term given the treatment may not have the mean of zero. Our estimated coefficient of 0.26 is biased and cannot correctly capture the size of the effect given the lack of inclusion for, for example, debt and government expenditure excluding education. Both variables relate to our treatment variable, being determined in some part by revenue, and our outcome variable, by influencing how much can be spent in the total budget. Not including such terms in the specification leads to the error correlating with the regressor, hence the OVB. Although this makes our base case ineligible for an unbiased interpretation, this issue is addressed in the following specifications.

Extension Analysis

Specification (2) adds debt, expenditure excluding education, as a new "baseline" without overt OVB by accounting for such omitted variables. We also add clustered standard errors, clustered by state, to prevent underestimation of true uncertainty. This is done as our error terms may not be uncorrelated over time within a state, threatening assumption FE6. The resulting coefficient is 0.106, more than a 50% decrease in the effect of total revenue on education expenditure. After considering debt and other expenditures' effect on the relationship, on average, an increase of 1000 dollars in total revenue only increases education expenditure by 106 dollars. Other expenditures excluding education, however, play a bigger explanatory role than even total revenue in (1), where an increase of 1000 dollars in other expenditures seems to increase education expenditure by, on average, 365 dollars. Both coefficients are statistically significant, with t-statistics of 2.08 and 5.2, respectively. This indicates that government revenue is not as important as how much governments are willing to spend to determine education expenditure. This result suggests policies that encourage larger overall spending benefit access to education more than simply giving the government more revenue. Variables and controls added for specification (2) also remain for the rest of the specifications (3-5).

For specifications (3 and 4), either state or time fixed effects, respectively, are added to specification (2). For (3), this is to account for state-to-state revenue and expenditure variability, leading to a more precise estimate of the effect that only compares in-state changes in total revenue and its effect on education expenditure. The coefficient decreases, now at a statistically significant (at 10%) 0.0288. Questions for economic significance seem pertinent now, as an average increase of 28.8 dollars in education per 1000 dollars of total revenue when controlling for the aforementioned effects seems inefficient. The reason for our inability to introduce both state and time fixed effects simultaneously, and therefore need (4), stems from the fact that there is only one observation per state, per year. As such, we would lack the variability in our treatment variable that would be required for a meaningful coefficient. (4) compares within years, bypassing a possible confounding factor of changes in attitudes towards access to education over time. The coefficient here is 0.0408, but statistically insignificant.

Finally, with specification (5), we add the political position of the states as both a static control and an interaction variable. This specification allows us to observe coefficient differences dependent on political leanings. The political leanings are collected from the results of the 2004 presidential election, the closest presidential election to the median year of 2004.5. This, in conjunction with all explanatory effects discussed above (aside from state effects for the aforementioned reasons), provides the most accurate and informative coefficient. The coefficient is 0.0283 for total revenue and education expenditure and 0.115 for the interaction variable. Interestingly, even counterintuitively, this positive coefficient means that red states have a higher marginal propensity to spend on education. This appears contradictory to the general message of the political left, the blue states. Although this variable is inconsequential when it comes to our verdict about the efficacy of this method, it does control for political stances of the states and provides interesting insight.

Discussion of Results

To discuss the results of our regression, we will focus on our most robust and informative specifications, being specification (5), with specification (4) slightly less informative given the lack of political leaning as a variable. We consider specification (4) to be more informative than specification (3) because time fixed effects can control for the changes in attitude towards education funding or access to public education, both on a nationwide level. Given the U.S. combined national identity and general culture at any given time, we deemed state differences in marginal propensities less confounding than the changes in attitudes over time. Additionally, our red-state interaction variable can account for a major reason for state-to-state differences, while we have no variable to account for attitude changes over time.

So, our focus is on the coefficients prescribed by specification (4, 5), being a statistically insignificant 0.0408, and a statistically significant (at a 1% level) 0.0283, respectively. These are both surprising results; one fails to reject the null hypothesis, and the other indicates an economically insignificant effect. Failing to reject the null indicates that, when controlling for other expenditures excluding education, debt at the end of the financial year, and time fixed effects, change in total revenue cannot be said to affect education expenditure. As such, according to (4), trying to increase government revenue to increase education is not only inefficient, but wholly ineffective, as it simply finds no significant relationship between the two variables. On the other hand, specification (5) indicates that, when controlling for other expenditures excluding education, debt at the end of the financial year, time fixed effects, political leanings, and an interactive variable of political leanings and total revenue, 1000 dollar increases to government revenue only increase education expenditure by 28.3 dollars, on average. This is astoundingly low, though not statistically insignificant. There is an effect, but its small size renders this method inefficient. This can be seen in comparison to specification (6), which is (4) but utilizes public welfare expenditure as the variable of interest. A statistically significant (at a 1% level) coefficient of 0.153 heavily outshines the marginal propensity to spend on education. An average increase of $153 in expenditure per $1000 increase in revenue, with all the controls of (4), seems to align more with what would be required for increasing revenue to be an effective strategy for increasing specific expenditures. As such, we must conclude that UNESCO's 2014 policy paper supports an, at best inefficient and, at worst ineffective policy for increasing access to education.

Limitations of Results

As mentioned above, the nature of our data entails that several least squares assumptions would have been violated, even if these breaches were mitigated. For example, although LSA1.2 can be mitigated with our large sample size, these infringements constrain our result to be correlational rather than causational and threaten the validity of our inference. Furthermore, we were unable to find data for a number of possibly confounding variables, such as average high school GPA or SAT scores, a numerical representation of the level of support in widening access to education, and average income for each state and year within our time period. These variables would have improved the accuracy of our estimation of the true isolated effect, as they may also play an explanatory role. Standardized test scores may influence education expenditure by motivating spending. For example, a state with lower than nation-wide average GPAs may decide to invest more in education to increase this GPA. Levels of support for access to education may replace our time fixed effects as a clearer indication of how attitudes of those with decision-making power affect spending. Average incomes may indicate how necessary widening access to education is, with higher income states less needing government support for education. Higher average incomes indicate a smaller reliance on public education, as households are more likely to afford private education. Further studies with these explanatory effects can shed more light on the efficiency and effectiveness of government revenue-based ways of funding access to education.

Additionally, our lack of per-capita data may hinder our inference's accuracy. Higher revenue states are almost guaranteed, simply by the fact that their budgets are higher, to spend more on education as a state. This leads to an upward trend as government revenue increases, whereas per-capita values may not have this (Figure 1). Perhaps smaller revenue states can have higher per-capita education expenditures and disrupt this assumed positive trend, but this remains unclear due to our lack of data. However, this may not affect our conclusion by much. If total revenue was inefficient in increasing total education expenditure, we would likely reach the same conclusion for per-capita statistics.

As a further note, future studies should explore alternative methods to increase access to education, given the lackluster effect of total revenue on education expenditure as seen in this paper.

Conclusion

Given the coefficients of specification (4) and specification (5), which are the best controlled for specifications, we deem it fit to conclude that access to education is not treated as a spending priority for the government. As such, tax rate increases, or other forms of government revenue increases would not serve as an effective method in widening education access. It would be inefficient in the case of (5), but wholly ineffective in the case of (4). Advocates against the policy paper receive empirical support that this is not only an effective solution to the allocation of education, but likely needs to be fully reformed.

References:

270ToWin. 2023. “Presidential Election of 2004.” n.d. 270toWin.com. https://www.270towin.com/2004_Election/

Kaplan, Jacob. Annual Survey of State Government Finances 1992-2018. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2021-01-16. https://doi.org/10.3886/E101880V4

National Center for Education Statistics. 2022. “COE - Annual Earnings by Educational Attainment.” Nces.ed.gov. May 2022. https://nces.ed.gov/programs/coe/indicator/cba/annual-earnings

United Nations Educational Scientific Cultural Organization. 2014. “UNESCO - Increasing Tax Revenues to Bridge the Education Financing Gap. 2014. UNESCO. UNESCO. March, 2014. https://unesdoc.unesco.org/ark:/48223/pf0000227092

U.S. Bureau of Labor Statistics. 2009. “Unemployment Rates for Persons 25 Years and Older by Educational Attainment.” 2009. Bls.gov. 2009. https://www.bls.gov/charts/employment-situation/unemployment-rates-for-persons-25-years-and-older-by-educational-attainment.htm

World Bank. 2023. Review of *Inflation, Consumer Prices (Annual %) - United States*. World Bank. World Bank. 2023. https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?end=2018&locations=US&name_desc=true&start=1992

Tables and Figures

Table 1
Summary Statistics

Variable	N	Mean	Std.Dev.	Min	Max
Total Revenue	1350	31403451	42167534.4	1756015	412400000
Education Expenditure	1350	9079486.41	11348031.67	430517	102400000
Public Welfare Expenditure	1350	7494452.40	11629907.63	191468	141300000
Total Expenditure Excluding Education Expenditure	1350	20874966.98	28730349.98	1134365	273000000
Total Expenditure Excluding Public Welfare Expenditure	1350	22459617.42	28388063.91	1274427	234000000
Debt at the end of fiscal year	1350	16105420.02	23361510.09	485787	156800000

Note. Unit of observation is state-year.

Table 2
Dependent variable: education expenditure (thousands of USD)

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Total Revenue	0.260***	0.106*	0.0288*	0.0408	0.0283**	0.153***
	(0.00513)	(0.0508)	(0.0140)	(0.0218)	(0.0100)	(0.0160)
95%Confidence Interval	[0.249, 0.270]	[0.004, 0.208]	[0.001, 0.056]	[-0.003, 0.085]	[0.008, 0.049]	[0.121, 0.185]
Expenditure Excluding Education Expenditure		0.365***	0.357***	0.333***	0.319***	
		(0.0700)	(0.0257)	(0.0218)	(0.0174)	
Debt at the End of Fiscal Year		-0.182***	-0.0275	-0.0446	-0.0115	0.0101
		(0.0479)	(0.0297)	(0.0292)	(0.0165)	(0.0547)
Red State * Total Revenue					0.115***	
					(0.0237)	
Expenditure Excluding Public Welfare Expenditure						0.237***
						(0.0550)
State Effects	No	No	Yes	No	No	No
Time Effects	No	No	No	Yes	Yes	Yes
Clustered Standard Errors	No	Yes	Yes	Yes	Yes	yes
F-Statistics & *p*-Values:						
Time Effects				4.42	5.57	7.21
				(0.0000)	(0.0000)	(0.0000)
Rest Expenditure		27.14	192.81	233.64	336.16	18.57
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Debt		14.43	0.86	2.33	0.48	0.03
		(0.0004)	(0.3551)	(0.1331)	(0.4896)	(0.8541)
R^2	0.930	0.952	0.917	0.927	0.946	0.888
Observations	1350	1350	1350	1350	1350	1350

Note. Unit of observation is state-year. Standard errors in parentheses are either robust or, if indicated, clustered at the state level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$