Private Beats Public:

A Flexible Value-Added Model with Tanzanian School Switchers

Kasper Brandt
Department of Economics
University of Copenhagen

June 2018

The Pitch

- What I do: Set up a flexible value-added model, and use it to estimate learning effects of private schools in Tanzania.

The Pitch

- What I do: Set up a flexible value-added model, and use it to estimate learning effects of private schools in Tanzania.
- What I expect: Better school inputs $\Rightarrow$ better performance. Costs as proxy for school inputs?

The Pitch

- What I do: Set up a flexible value-added model, and use it to estimate learning effects of private schools in Tanzania.
- What I expect: Better school inputs $\Rightarrow$ better performance. Costs as proxy for school inputs?
- Why I do it: Strong assumptions needed in existing literature and almost no evidence on Sub-Saharan countries.

The Pitch

- What I do: Set up a flexible value-added model, and use it to estimate learning effects of private schools in Tanzania.
- What I expect: Better school inputs $\Rightarrow$ better performance. Costs as proxy for school inputs?
- Why I do it: Strong assumptions needed in existing literature and almost no evidence on Sub-Saharan countries.
- How I do it: Compare secondary school GPA for students getting the same primary school GPA from the same primary school.

The Pitch

- What I do: Set up a flexible value-added model, and use it to estimate learning effects of private schools in Tanzania.
- What I expect: Better school inputs $\Rightarrow$ better performance. Costs as proxy for school inputs?
- Why I do it: Strong assumptions needed in existing literature and almost no evidence on Sub-Saharan countries.
- How I do it: Compare secondary school GPA for students getting the same primary school GPA from the same primary school.
- What I find: Private schools increase students' secondary school GPA by 0.40 of a standard deviation after two years of secondary schooling.

Why is it important to study?

- Huge increases in *quantity* of education, while *quality* of education remains weak or even worsens.

Why is it important to study?

- Huge increases in *quantity* of education, while *quality* of education remains weak or even worsens.
- Knowledge is good!

Why is it important to study?

- Huge increases in *quantity* of education, while *quality* of education remains weak or even worsens.
- Knowledge is good!
- Private schools tend to be cheaper to operate in developing countries.

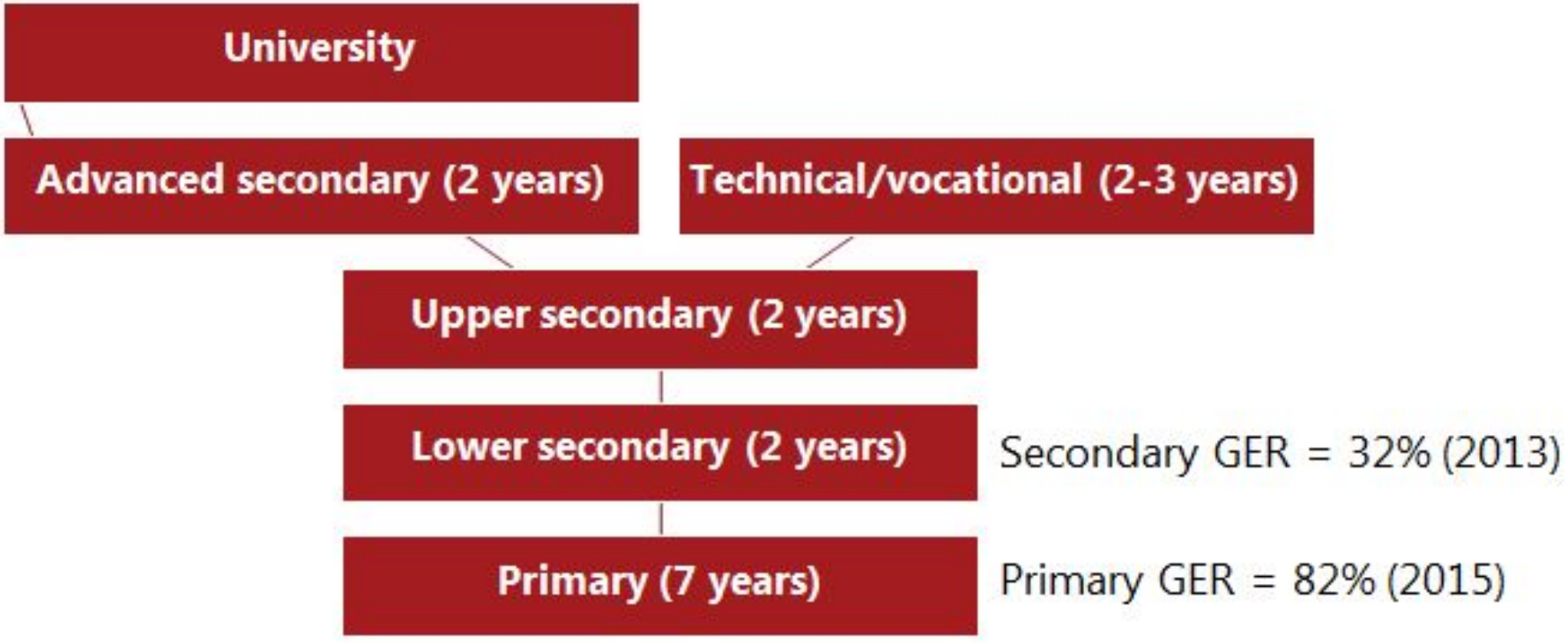
Why is it important to study?

- Huge increases in *quantity* of education, while *quality* of education remains weak or even worsens.
- Knowledge is good!
- Private schools tend to be cheaper to operate in developing countries.
- Tanzania has launched the programme "Big Results Now". This programme presents several ambitious goals for six key sectors, including the education sector.

Why is it important to study?

- Huge increases in *quantity* of education, while *quality* of education remains weak or even worsens.
- Knowledge is good!
- Private schools tend to be cheaper to operate in developing countries.
- Tanzania has launched the programme "Big Results Now". This programme presents several ambitious goals for six key sectors, including the education sector.
- Strong assumptions needed in the current literature estimating private school learning premiums.

Education in Tanzania



Private School Enrolment in East Africa

	Private school enrolment	
	Primary school	Secondary school
Burundi	1.2% (2013)	9.1% (2013)
Kenya	16.0% (2014)	No recent data
Rwanda	2.7% (2013)	18.0% (2013)
Tanzania	**2.4% (2013)**	**21.4% (2013)**
Uganda	16.2% (2013)	No recent data

Source: World Development Indicators.

High-quality studies (1)

- Singh (2015) (JDE) employs a value-added model to Indian students accounting for unobserved ability by including lagged Raven's test scores. Positive pivate school learning premium, but depends on rural/urban status, age of the child, and school subject.

High-quality studies (1)

- Singh (2015) (JDE) employs a value-added model to Indian students accounting for unobserved ability by including lagged Raven's test scores. Positive pivate school learning premium, but depends on rural/urban status, age of the child, and school subject.
- Andrabi et al. (2011) (AEJ: Applied) study the effects of measurement error and unobserved ability when estimating a private school learning premium in Pakistan. Accounting for these, they find a positive effect of 0.25 of a standard deviation per year.

High-quality studies (2)

- Angrist et al. (2002) (AER) study learning effects from a random allocation of private school vouchers in Columbia. Three years later, "lottery winners" were less likely to repeat grades, and they scored 0.21 of a standard deviation higher on tests.

High-quality studies (2)

- Angrist et al. (2002) (AER) study learning effects from a random allocation of private school vouchers in Columbia. Three years later, "lottery winners" were less likely to repeat grades, and they scored 0.21 of a standard deviation higher on tests.
- Muralidharan and Sundararaman (2015) (QJE) study learning effects from a random allocation of private school vouchers in India. Four years later, "lottery winners" scored 1.07 and 0.23 of a standard deviation higher in Hindi and English test scores, respectively. Insignificant effects on test scores in Telugu, mathematics, science, and social studies.

Cumulative learning production function

Todd and Wolpin (2003) (EJ) present a cumulative learning production function:

$$T_{ija} = T_a[\boldsymbol{F}_{ij}(a), \boldsymbol{S}_{ij}(a), \mu_{ij0}, \varepsilon_{ij}]. \quad (1)$$

T_{ija} is achievement for student i in household j at age a. $\boldsymbol{F}$ is a vector containing family inputs, $\boldsymbol{S}$ is a vector containing school inputs, and μ is unobserved ability for each student i.

Standard value-added model

$$T_{ija} = \boldsymbol{F}_{ija}\varphi_a + \boldsymbol{S}_{ija}\alpha_a + \gamma T_{ij,a-1} + \eta_{ija}, \tag{2}$$

Five assumptions needed for the standard value-added model:

1. The arguments in the cumulative learning production function are additively separable.

Standard value-added model

$$T_{ija} = \boldsymbol{F}_{ija}\varphi_a + \boldsymbol{S}_{ija}\alpha_a + \gamma T_{ij,a-1} + \eta_{ija}, \tag{2}$$

Five assumptions needed for the standard value-added model:

1. The arguments in the cumulative learning production function are additively separable.
2. The coefficients on inputs are non-age varying.

Standard value-added model

$$T_{ija} = \boldsymbol{F}_{ija}\varphi_a + \boldsymbol{S}_{ija}\alpha_a + \gamma T_{ij,a-1} + \eta_{ija}, \tag{2}$$

Five assumptions needed for the standard value-added model:

1. The arguments in the cumulative learning production function are additively separable.
2. The coefficients on inputs are non-age varying.
3. Learning effects from school and family inputs decay at the same rate over time.

Standard value-added model

$$T_{ija} = \boldsymbol{F}_{ija}\varphi_a + \boldsymbol{S}_{ija}\alpha_a + \gamma T_{ij,a-1} + \eta_{ija}, \tag{2}$$

Five assumptions needed for the standard value-added model:

1. The arguments in the cumulative learning production function are additively separable.
2. The coefficients on inputs are non-age varying.
3. Learning effects from school and family inputs decay at the same rate over time.
4. The impact of unobserved ability decays at the same rate as the effects from school and family inputs.

Standard value-added model

$$T_{ija} = \boldsymbol{F}_{ija}\varphi_a + \boldsymbol{S}_{ija}\alpha_a + \gamma T_{ij,a-1} + \eta_{ija}, \tag{2}$$

Five assumptions needed for the standard value-added model:

1. The arguments in the cumulative learning production function are additively separable.
2. The coefficients on inputs are non-age varying.
3. Learning effects from school and family inputs decay at the same rate over time.
4. The impact of unobserved ability decays at the same rate as the effects from school and family inputs.
5. Unobserved ability does not influence the return to school and family inputs.

A flexible value-added model

Including a *lagged school* $\times$ *lagged test score* $\times$ *cohort* fixed effect, I am able to loosen assumption 3, 4, and 5 from the previous slide.

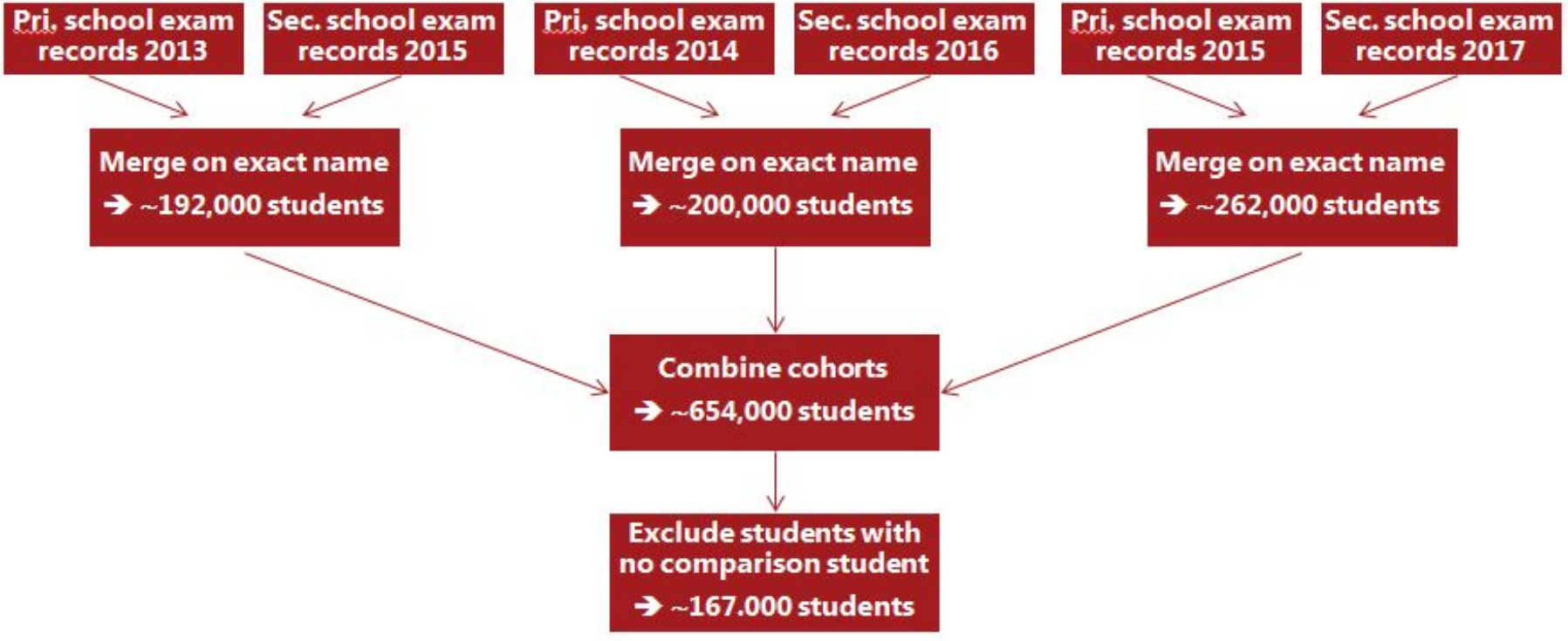
$$T_{isgc} = \boldsymbol{F}_i\varphi + \boldsymbol{S}_i\alpha + \mu_{i0}\beta + \theta_{sgc} + \eta_{isg} \tag{3}$$

T_{isga} is secondary school test score for student i, who attended primary school s, got primary school test score g, and belongs to cohort c.

Current school inputs include private school enrolment, peer effects, and school size.

Family inputs are excluded. Fortunately, the literature agrees they are irrelevant when controlling for lagged achievement and peer effects.

Data source



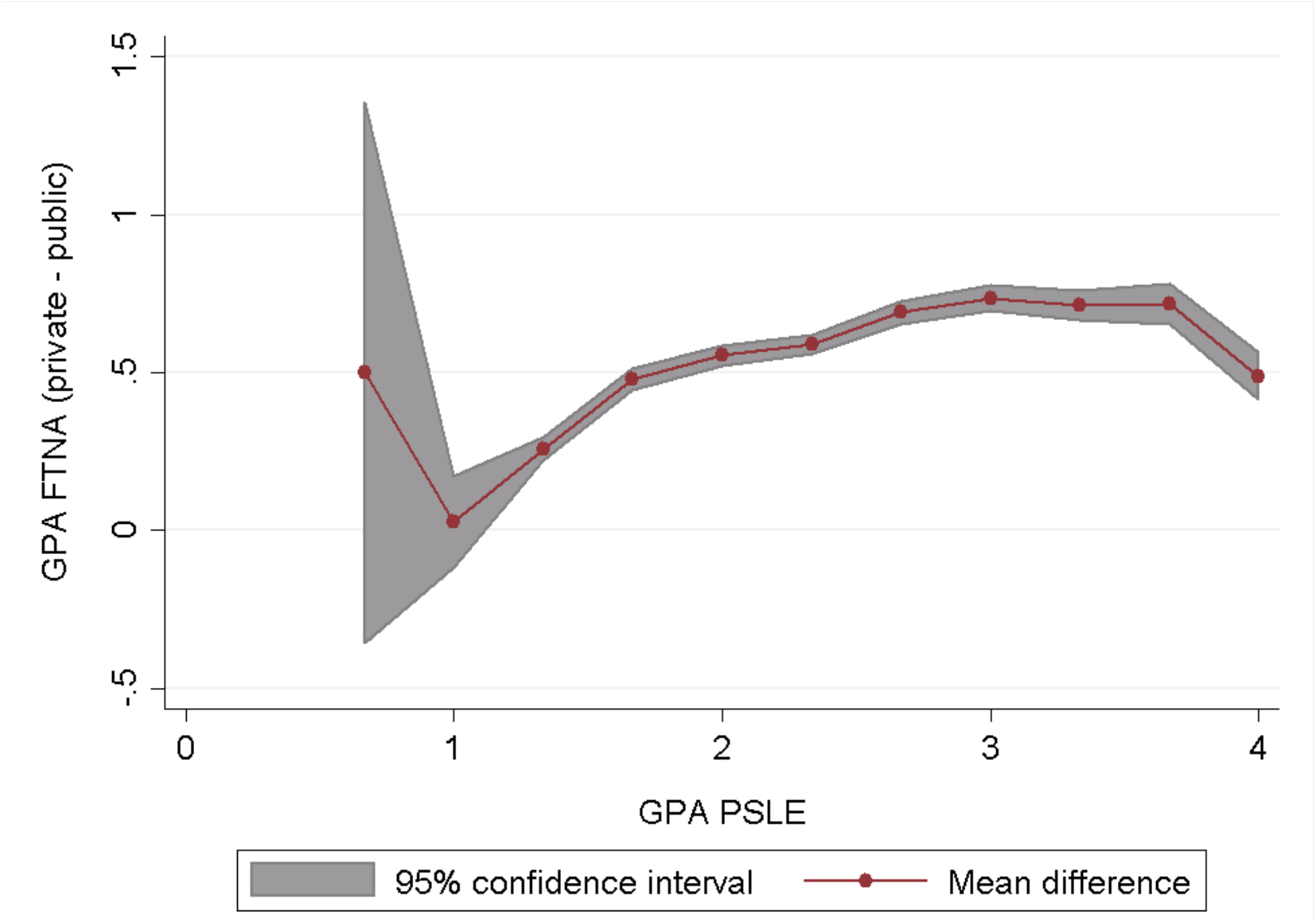
NOT representative sample, but...

Descriptive statistics

	Pop. Mean	Pop. Std.	Sample Mean	Sample Std.	Private Mean	Public Mean
GPA (secondary)	1.308	0.881	1.661	0.953	2.347	1.397
GPA (primary)	1.664	0.832	2.411	0.739	2.718	2.293
Proxy for unobs. ability	1.713	0.780	2.316	0.632	2.439	2.268
Private (primary)	0.034	0.180	0.167	0.373	0.390	0.081
Private (secondary)	0.182	0.386	0.278	0.448	1.000	0.000
Secondary school size	146	95	180	106	123	202
Peers' GPA (primary)	2.224	0.426	2.353	0.492	2.721	2.211
Female	0.516	0.500	0.539	0.498	0.557	0.532
Cohort 2016	0.325	0.468	0.315	0.464	0.319	0.313
Cohort 2017	0.388	0.487	0.388	0.487	0.379	0.391
N	See notes		167,334	167,334	46,560	120,774

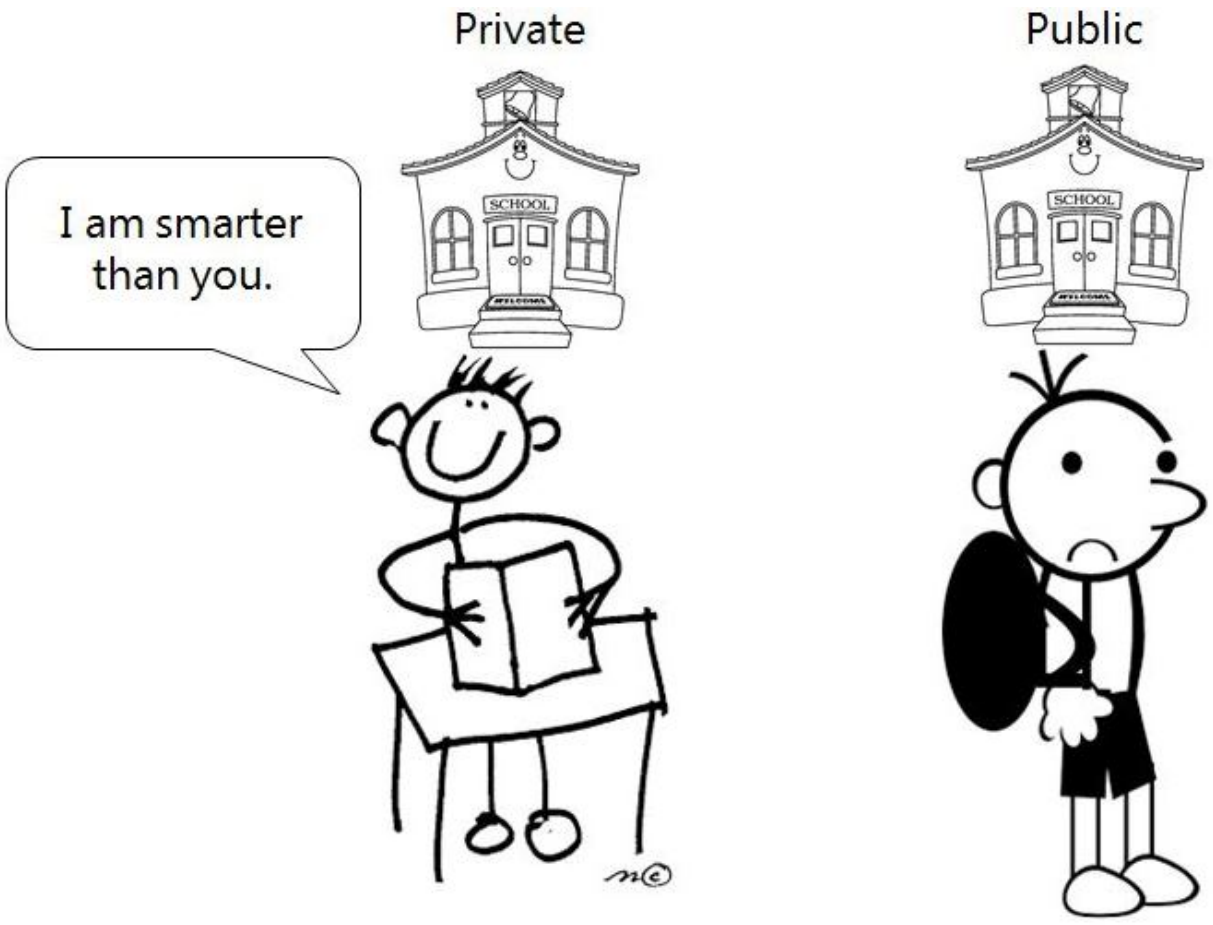
Source: Author's own calculations.
Notes: Population means of *GPA PSLE*, *GPA PSLE other*, and *Private Primary* are based on 2,314,638 primary school students. The population means of the remaining variables are based on 1,246,267 secondary school students. The two last columns provide mean values for sample students attending private and public secondary school, separately.

Differences in test scores conditional on lagged test scores



OLS

$$GPA_i = \beta_0 + \beta_1 private_i + \beta_2 school\ size_i + \beta_3 female_i + \beta_4 cohort16_i + \beta_5 cohort17_i + \varepsilon_i$$

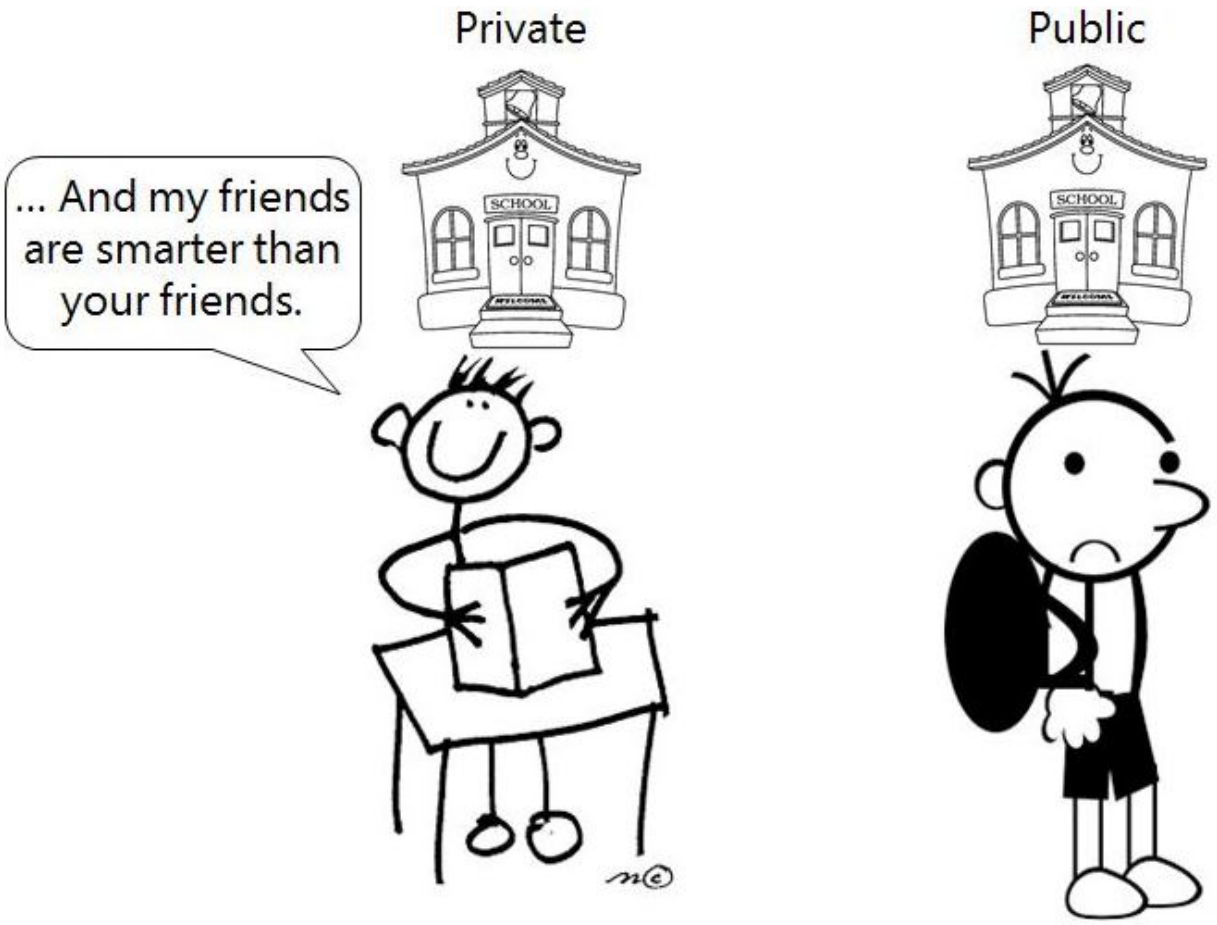


Standard value-added model

$$GPA_{i,t} = \beta_0 + \beta_1 private_{i,t} + \beta_2 school\ size_{i,t} + \beta_3 female_i + \beta_4 cohort16_i + \beta_5 cohort17_i + \beta_6 GPA_{i,t-1} + \varepsilon_{i,t}$$

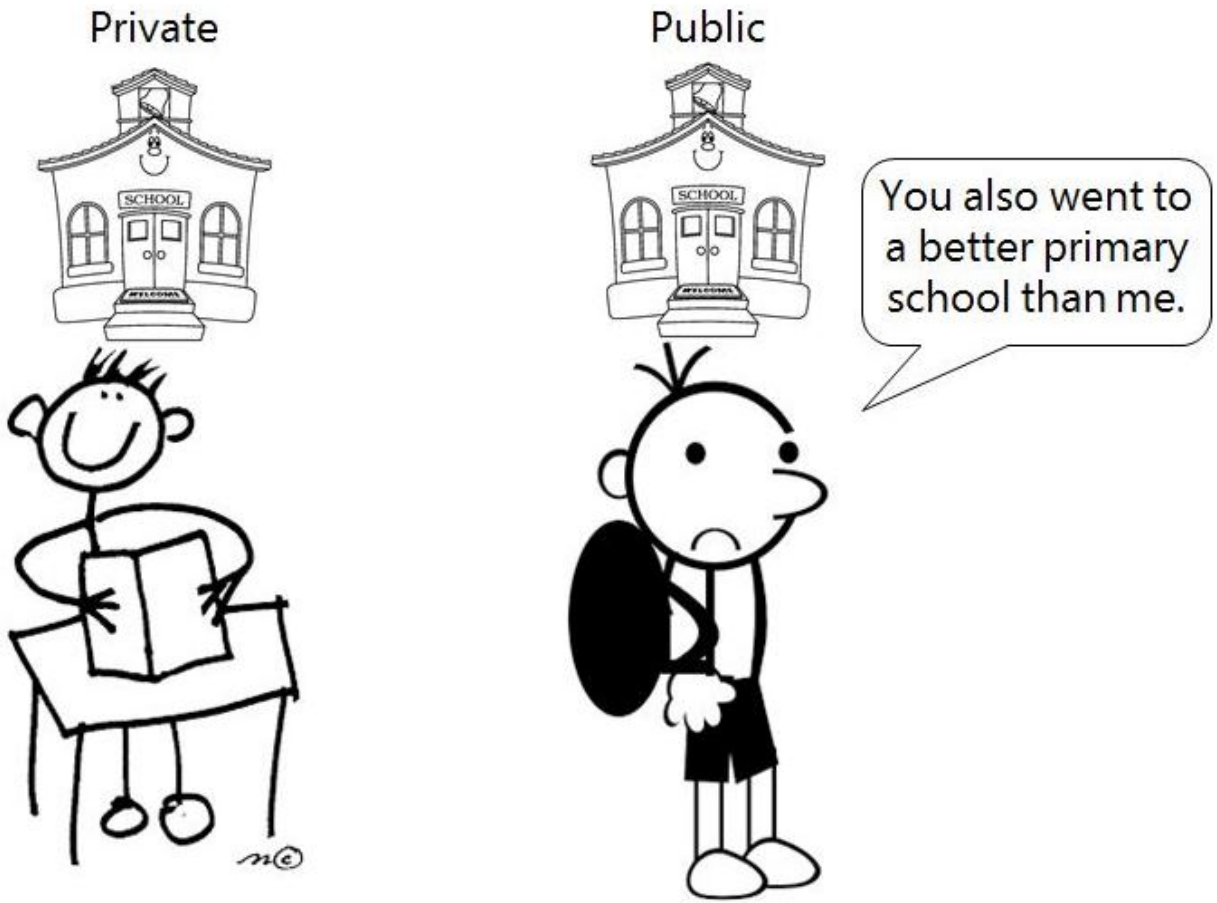
Standard value-added model including peer effects

$$GPA_{i,t} = \beta_0 + \beta_1 private_{i,t} + \beta_2 school\ size_{i,t} + \beta_3 female_i + \beta_4 cohort16_i + \beta_5 cohort17_i + \beta_6 GPA_{i,t-1} + \beta_7 peer\ effects_{i,t} + \varepsilon_{i,t}$$



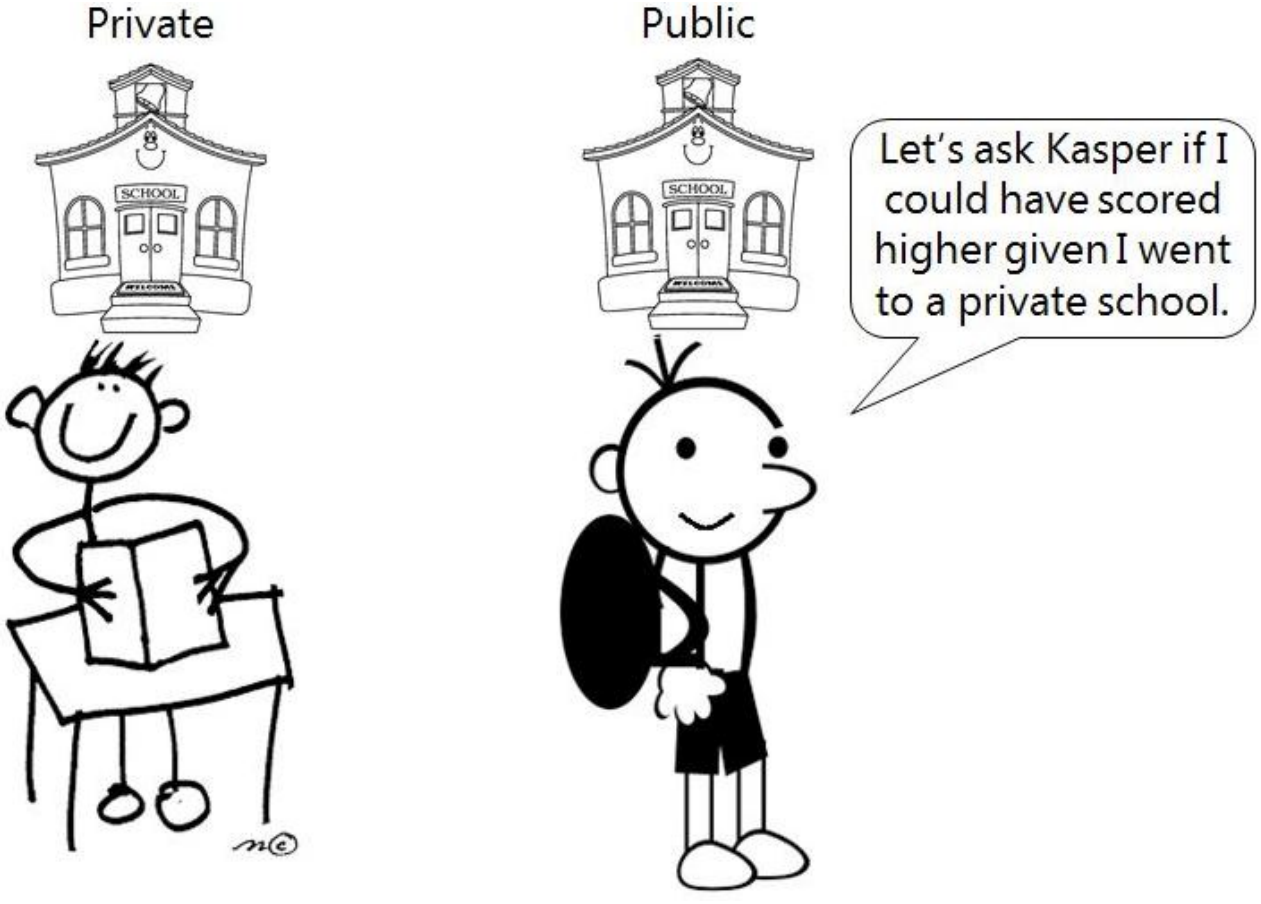
A flexible value-added model

$$GPA_{isgc,t} = \beta_0 + \beta_1 private_{i,t} + \beta_2 school\ size_{i,t} + \beta_3 female_i + \beta_6 peer\ effects_{isgc,t} + \theta_{sgc} + \varepsilon_{i,t}$$



Flexible value-added model including unobserved ability

$$GPA_{isgc,t} = \beta_0 + \beta_1 private_{i,t} + \beta_2 school\ size_{i,t} + \beta_3 female_i + \beta_6 peer\ effects_{i,t} + \theta_{sgc} + \beta_7 GPA\ other_{i,t-1} + \varepsilon_{isgc,t}$$


Private
Public
SCHOOL
SCHOOL
Let's ask Kasper if I could have scored higher given I went to a private school.

Results

Dependent variable: GPA_t (*FTNA*)	(1)	(2)	(3)	(4)	(5)
$Private_t$	1.004***	0.661***	0.527***	0.379***	0.396***
	(0.040)	(0.023)	(0.022)	(0.012)	(0.011)
$log(School\ size_t)$	0.006	-0.042***	-0.061***	-0.085***	-0.088***
	(0.023)	(0.013)	(0.012)	(0.008)	(0.008)
Female	-0.041**	0.029***	0.025***	0.059***	0.126***
	(0.017)	(0.008)	(0.007)	(0.005)	(0.005)
GPA_{t-1} *(PSLE)*		0.546***	0.450***		
		(0.008)	(0.005)		
$Peer\ effects_t$			0.173***	0.216***	0.188***
			(0.010)	(0.006)	(0.006)
$GPA\ other_{t-1}$ *(PSLE)*					0.228***
					(0.003)
Accounts for θ_{sgc}	No	No	No	Yes	Yes
N	167,334	167,334	167,334	167,334	167,334
R^2	.221	.491	.505	.690	.706

Source: Author's own calculations.
Notes: Standard errors are clustered at secondary school level. GPA_t is the grade point average of the subjects Kiswahili, English, and mathematics. $Peer\ effects_t$ is the average grade point average of the subjects Kiswahili, English, and mathematics in primary school for secondary school schoolmates. $GPA\ other_{t-1}$ is the grade point average of the subjects Community Knowledge and Science in primary school. GPA_{t-1}, $Peer\ effects_t$, $GPA\ other_{t-1}$, and GPA_t are standardized by their sample means and standard deviations. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

References

T. Andrabi, J. Das, A. I. Khwaja, and T. Zajonc. Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics*, 3(3):29–54, 2011.

J. Angrist, E. Bettinger, E. Bloom, E. King, and M. Kremer. Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review*, 92(5):1535–1558, 2002.

K. Muralidharan and V. Sundararaman. The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India. *The Quarterly Journal of Economics*, 130(3):1011–1066, 2015.

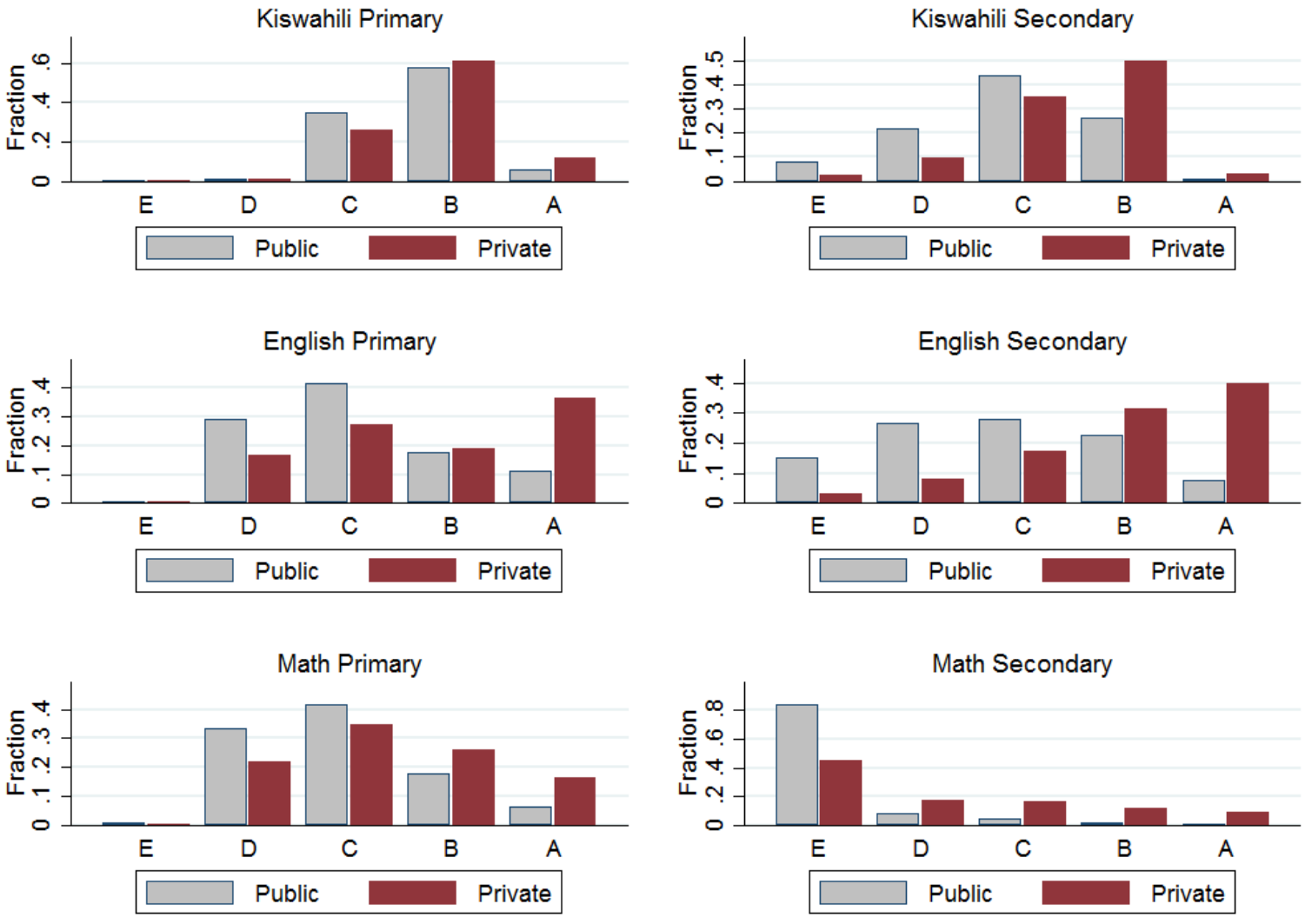
A. Singh. Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113:16–32, 2015.

P. E. Todd and K. I. Wolpin. On the Specification and Estimation of the Production Function for Cognitive Achievement*. *The Economic Journal*, 113(485):F3–F33, 2003.

Regional distribution of secondary school students

	Population	Sample	Private	Public
Arusha	0.056	0.069	0.068	0.069
Dar Es Salaam	0.106	0.217	0.158	0.240
Dodoma	0.035	0.039	0.036	0.040
Geita	0.032	0.020	0.010	0.023
Iringa	0.035	0.038	0.038	0.038
Kagera	0.048	0.030	0.033	0.029
Katavi	0.008	0.005	0.002	0.006
Kigoma	0.031	0.025	0.032	0.022
Kilimanjaro	0.071	0.082	0.129	0.064
Lindi	0.016	0.007	0.004	0.008
Manyara	0.027	0.013	0.012	0.013
Mara	0.043	0.024	0.019	0.026
Mbeya	0.075	0.095	0.088	0.098
Morogoro	0.050	0.045	0.054	0.042
Mtwara	0.027	0.015	0.012	0.017
Mwanza	0.077	0.061	0.068	0.058
Njombe	0.022	0.025	0.025	0.025
Pwani	0.034	0.042	0.072	0.030
Rukwa	0.016	0.010	0.007	0.011
Ruvuma	0.032	0.022	0.026	0.021
Shinyanga	0.029	0.027	0.024	0.028
Simiyu	0.023	0.010	0.005	0.012
Singida	0.023	0.016	0.012	0.017
Songwe	0.001	0.001	0.004	0.000
Tabora	0.028	0.024	0.025	0.023
Tanga	0.055	0.038	0.037	0.038
N	1,246,267	167,334	46,560	120,774

Distribution of subject-specific exam scores



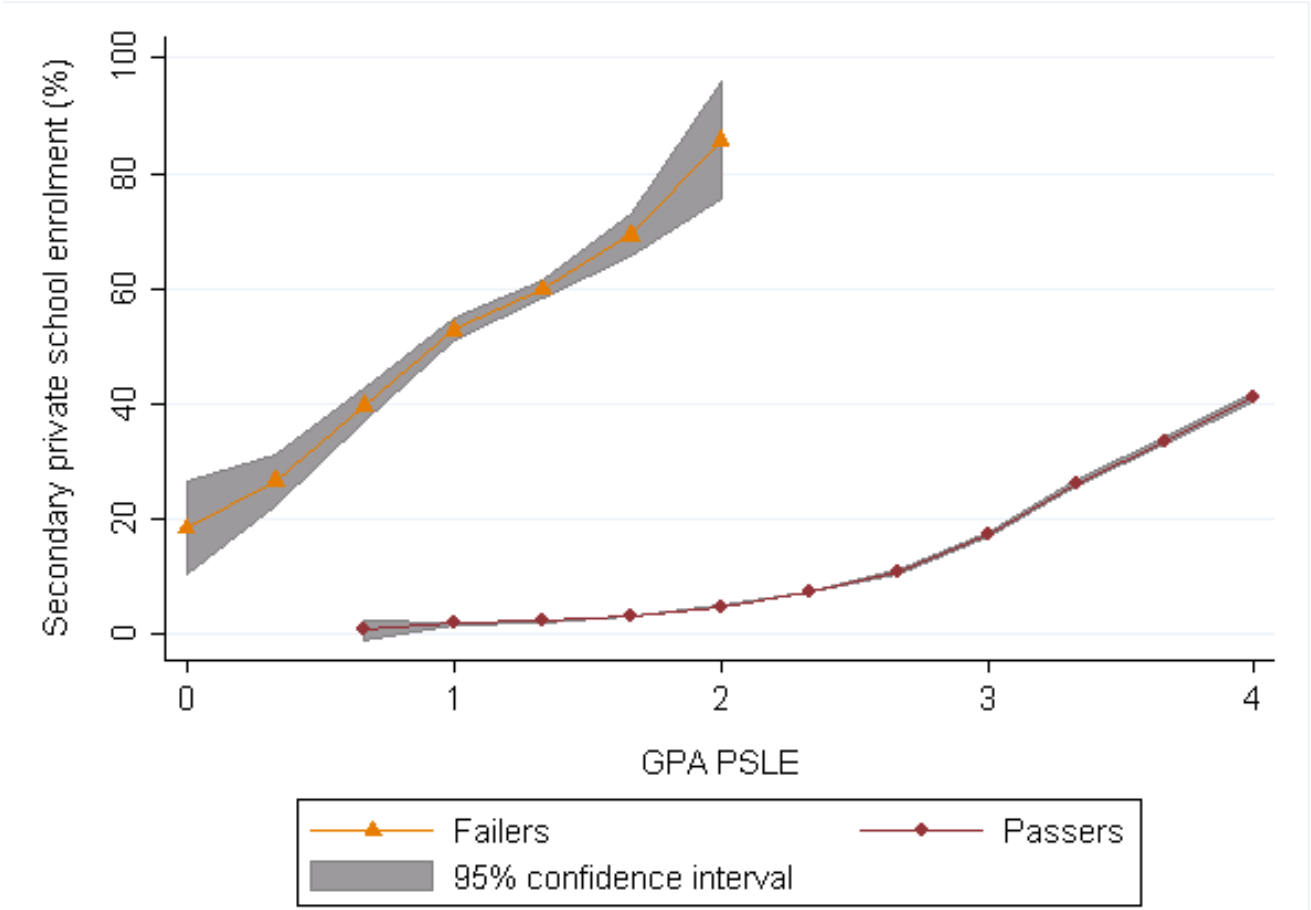
Value-added model versus Heckman-type correction and IV models

Dependent variable: GPA_t ($FTNA$)	Value-Added (1)	Heckman (2)	IV (3)	IV (4)	IV (5)
$Private_t$	0.719***	0.730***	0.728***	1.153**	0.733***
	(0.024)	(0.024)	(0.048)	(0.531)	(0.048)
$Nearby\ private\ schools_t$	0.023***	0.025***	0.023***	0.020***	0.023***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)
$Nearby\ private\ schools_t^2$	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$No\ nearby\ private\ schools_t$	-0.037***	-0.055***	-0.037***	-0.028	-0.037***
	(0.011)	(0.012)	(0.011)	(0.017)	(0.011)
Inverse Mills ratio		0.106***			
		(0.017)			
Instruments:					
- *Failing PSLE*			Yes	No	Yes
- *Private school share 10 km*			No	Yes	Yes
Cragg-Donald Wald F statistic			>32,000	401	>16,000
Hansen J statistic					0.633
N	592,499	592,499	592,499	592,499	592,499
R^2	.432	.432	.432	.421	.432

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. The *Inverse Mills ratio* is estimated in a first-step probit model, using the dummy "failing the overall PSLE" and the continuous variable "private secondary school share within 10 kilometres" as the exclusion restriction. The same two variables are used as instruments for $Private_t$ in the IV model. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. $Nearby\ private\ schools_t$ is the number of private secondary schools within 10 kilometres of a student's primary school, whereas $No\ nearby\ private\ schools_t$ is a dummy taking the value one if there is no private schools within 10 kilometres. The models further account for school size, peer effects, gender, test scores in primary school, private primary schooling, year fixed effects, and region fixed effects. The Hansen J statistic has a chi-squared distribution with one degree of freedom. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Secondary private school enrolment for PSLE passers and failers



Source: Author's own calculations.
Notes: *GPA PSLE* is the average of primary school test scores in English, Kiswahili, and mathematics. The figure is based on 652,405 secondary school students. Basing the figure on the main sample of 167,334 students, increases the shares of students attending a private school for all levels of primary school GPA and independent of whether a student fails or passes the overall PSLE.

Sequential sample selection and weighting

Dependent variable: GPA_t ($FTNA$)	(1)	(2)
$Private_t$	1.119***	0.435***
	(0.040)	(0.012)
Inverse Mills ratio, λ_1	-1.575***	
	(0.056)	
Inverse Mills ratio, λ_2	-0.323***	
	(0.026)	
N	167,334	167,334
R^2	.734	.706

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. In column (1), λ_1 and λ_2 origin from two first-stage probit models explaining whether a student's PSLE records have been identified and whether the student is in the sample, respectively. In column (2), sample weights are applied to get a representative sample in regard to student ability, gender, private schooling, year of exam, ability of peers, and school size. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. The models further account for school size, gender, peer effects, "Primary school $\times$ Primary school GPA $\times$ CohortÂŽÂŽ fixed effects, and GPA of the subjects Community Knowledge and Science in primary school. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis of subject-specific exam scores

Dependent variable:	Kiswahili FTNA score	English FTNA score	Math FTNA score
$Private_t$	0.338***	0.371***	0.391***
	(0.015)	(0.014)	(0.017)
$log(School\ size_t)$	-0.115***	-0.080***	-0.069***
	(0.012)	(0.009)	(0.012)
Female	0.182***	0.049***	0.012
	(0.010)	(0.007)	(0.010)
$Peer\ effects_t$	0.031***	0.026***	0.091***
	(0.009)	(0.007)	(0.010)
"*Primary school* × *PSLE score* × *Cohort*" fixed effects	Yes	Yes	Yes
N	66,291	76,594	62,958
R^2	.49	.665	.637

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. *PSLE score* is the subject-specific exam score in primary school, and it is instrumented by the exam scores in all other primary school subjects. The dependent variables are standardized by their sample means and standard deviations. The sample sizes in columns (7), (8), and (9) are smaller than the full sample due to more lagged achievement possibilities and the requirement of only comparing students with the same lagged achievement. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis with different geographical sub-samples

Dependent variable: GPA_t ($FTNA$)	(1)	(2)	(3)	(4)
$Private_t$	0.450*** (0.015)	0.391*** (0.012)	0.386*** (0.013)	0.379*** (0.014)
"*Primary school* × *Primary school GPA* × *Cohort*" fixed effects	Yes	Yes	Yes	Yes
N	101,874	148,701	142,995	124,362
R^2	.726	.707	.71	.71

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. Column (1) excludes urban areas with more than 100,000 inhabitants, column (2) excludes the regions of Singida and Mbeya, column (3) excludes the regions of Iringa, Njombe, and Kilimanjaro, and column (4) excludes the regions of Singida, Mbeya, Iringa, Njombe, and Kilimanjaro. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. The models further account for school size, gender, peer effects, and GPA in Community Knowledge and Science. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis of public and private primary school students separately

Dependent variable: GPA_t ($FTNA$)	(1)	(2)
$Private_t$	0.433***	0.215***
	(0.012)	(0.022)
"*Primary school* × *Primary school GPA* × *Cohort*" fixed effects	Yes	Yes
Sample	Public primary school students	Private primary school students
N	137,449	27,500
R^2	.627	.626

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. The models further account for school size, gender, peer effects, and GPA in Community Knowledge and Science. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis with private school and peer effects interaction

Dependent variable: GPA_t ($FTNA$)	(1)
$Private_t$	0.387***
	(0.011)
$Peer\ effects_t$	0.152***
	(0.008)
$Private_t \times Peer\ effects_t$	0.072***
	(0.010)
"*Primary school* × *Primary school GPA* × *Cohort*" fixed effects	Yes
N	167,334
R^2	.707

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. GPA_t is the grade point average of the subjects Kiswahili, English, and mathematics. $Peer\ effects_t$ is the average grade point average of the subjects Kiswahili, English, and mathematics in primary school for secondary school schoolmates. The model further accounts for school size, gender, and GPA in Community Knowledge and Science. $Peer\ effects_t$ and GPA_t are standardized by their sample means and standard deviations. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis of secondary schools offering religious courses

Dependent variable: GPA_t ($FTNA$)	(1)	(2)	(3)
$Private_t$	0.410***	0.391***	0.414***
	(0.014)	(0.013)	(0.013)
$Religious\ courses_t$	-0.000		
	(0.010)		
$Private_t \times Religious\ courses_t$	-0.037*		
	(0.019)		
$Bible\ course_t$		-0.010	
		(0.015)	
$Private_t \times Bible\ course_t$		0.024	
		(0.023)	
$Islamic\ course_t$			-0.008
			(0.011)
$Private_t \times Islamic\ course_t$			-0.128***
			(0.024)
"*Primary school* × *Primary school GPA* × *Cohort*" fixed effects	Yes	Yes	Yes
N	167,334	167,334	167,334
R^2	.706	.706	.707

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. *Religious courses* is an indicator for whether the school offers either Bible Knowledge or Islamic Knowledge as elective courses. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. The models further account for school size, gender, peer effects, and GPA in Community Knowledge and Science. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Analysis of same gender secondary schools

Dependent variable: GPA_t ($FTNA$)	(1)	(2)	(3)
$Private_t$	0.381***	0.397***	0.380***
	(0.012)	(0.011)	(0.012)
$Same\ gender\ school_t$	0.079***		
	(0.029)		
$Private_t \times Same\ gender\ school_t$	0.101***		
	(0.032)		
$Boys\ school_t$		0.066	
		(0.053)	
$Private_t \times Boys\ school_t$		0.052	
		(0.060)	
$Girls\ school_t$			0.038
			(0.024)
$Private_t \times Girls\ school_t$			0.120***
			(0.030)
"$Primary\ school \times Primary\ school\ GPA \times Cohort$" fixed effects	Yes	Yes	Yes
N	167,334	167,334	167,334
R^2	.708	.706	.707

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. GPA_t is the standardized grade point average of the subjects Kiswahili, English, and mathematics. The models further account for school size, gender, peer effects, and GPA in Community Knowledge and Science. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Analysis of cohorts separately

Dependent variable: GPA_t ($FTNA$)	(1)	(2)	(3)	(4)
$Private_t$	0.376***	0.349***	0.457***	0.375***
	(0.017)	(0.014)	(0.017)	(0.014)
$Private_t \times Cohort16$				-0.011
				(0.015)
$Private_t \times Cohort17$				0.071***
				(0.015)
$log(School\ size_t)$	-0.082***	-0.096***	-0.078***	-0.085***
	(0.014)	(0.010)	(0.011)	(0.008)
$Female$	0.137***	0.108***	0.132***	0.126***
	(0.008)	(0.007)	(0.008)	(0.005)
$Peer\ effects_t$	0.184***	0.203***	0.177***	0.187***
	(0.008)	(0.008)	(0.007)	(0.006)
$GPA\ other_{t-1}$ $(PSLE)$	0.200***	0.213***	0.261***	0.228***
	(0.005)	(0.005)	(0.005)	(0.003)
"$Primary\ school \times Primary\ school\ GPA \times Cohort$" fixed effects	Yes	Yes	Yes	Yes
FTNA cohort	2015	2016	2017	All
N	49,803	52,680	64,851	167,334
R^2	.702	.714	.692	.706

Source: Author's own calculations.

Notes: Standard errors are clustered at the secondary school level. GPA_t is the grade point average of the subjects Kiswahili, English, and mathematics. $Peer\ effects_t$ is the average grade point average of the subjects Kiswahili, English, and mathematics in primary school for secondary school schoolmates. $GPA\ other_{t-1}$ is the grade point average of the subjects Community Knowledge and Science in primary school. $Peer\ effects_t$, $GPA\ other_{t-1}$, and GPA_t are standardized by their sample means and standard deviations. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.