

1 Download files to desktop and open

- Data file and do file accessible from moodle page ([link](#))
  - Save files to desktop, then open

# MEd Research Methods Strand 16-17

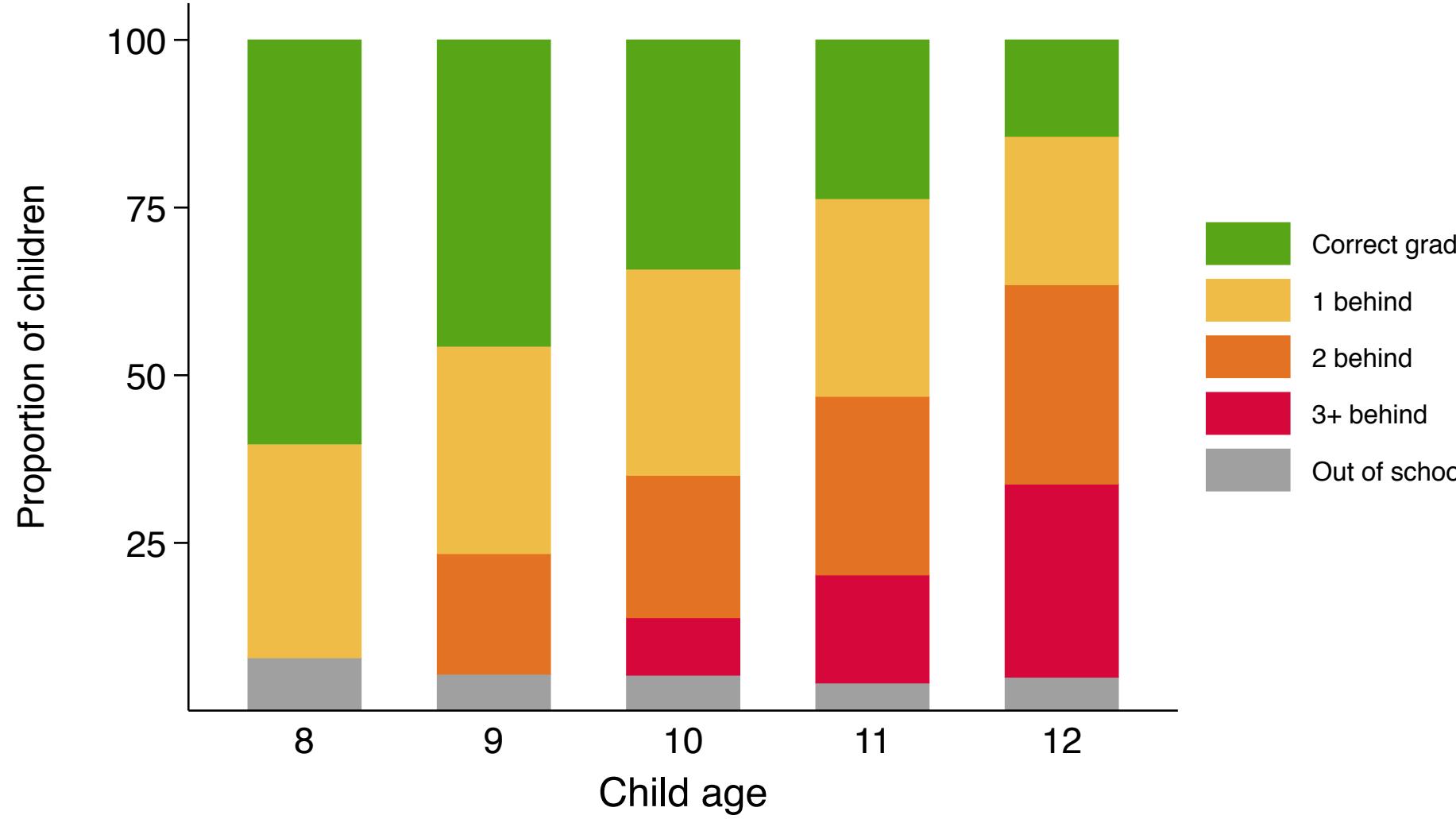
# Quantitative data analysis

An introduction using Stata

Ben Alcott

bma27@cam.ac.uk

## Uganda



Key	
frequency column percentage	

Overage	age					Total
	8	9	10	11	12	
0	5,849 68.11	3,710 51.10	3,718 39.44	1,571 27.78	1,771 19.36	16,619 41.47
1	2,739 31.89	3,550 48.90	5,708 60.56	4,085 72.22	7,377 80.64	23,459 58.53
Total	8,588 100.00	7,260 100.00	9,426 100.00	5,656 100.00	9,148 100.00	40,078 100.00

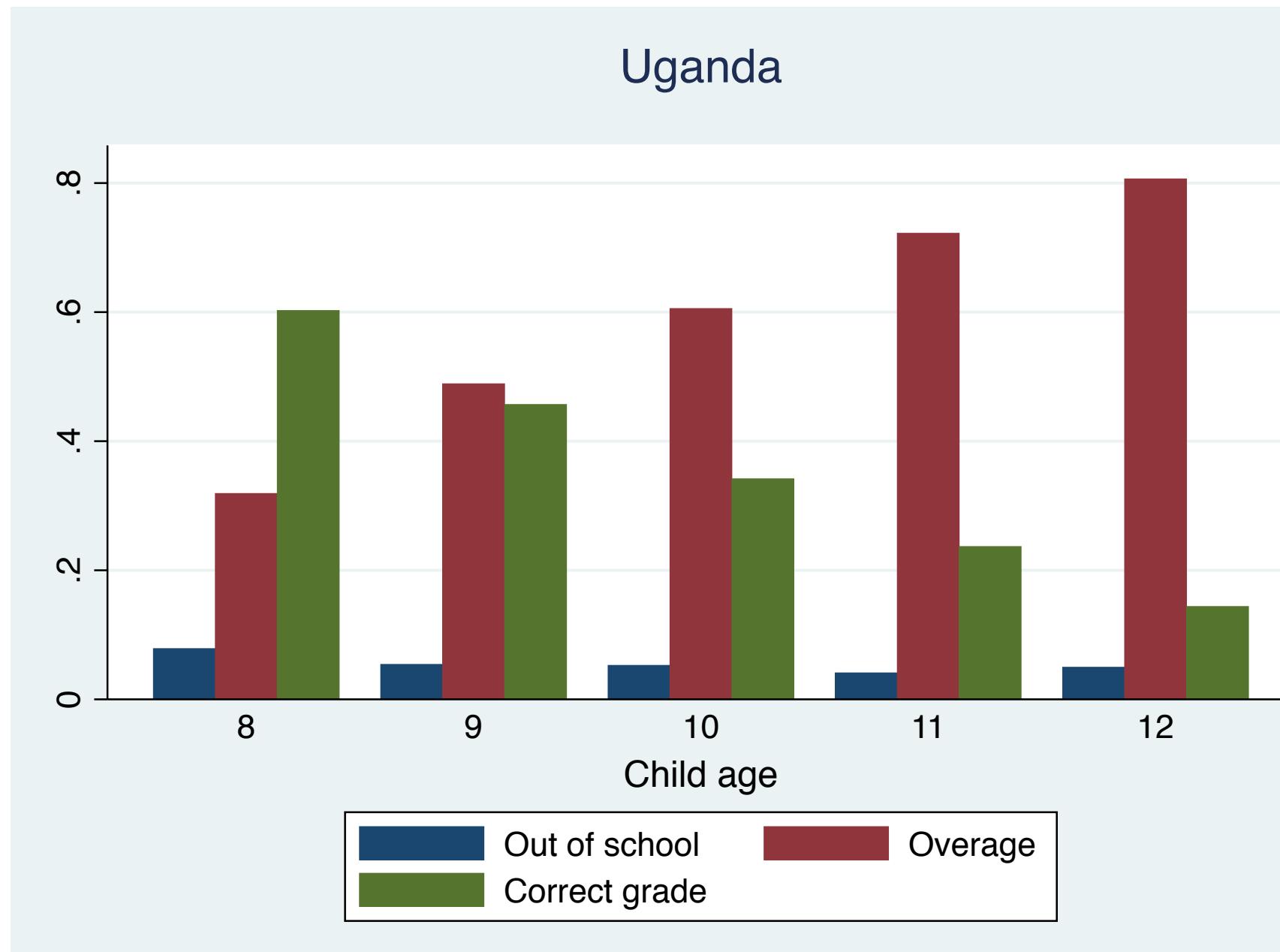


Table 1. Descriptive statistics, by age

	Age 8	Age 9	Age 10	Age 11	Age 12
Correct grade	0.60	0.46	0.34	0.24	0.14
Overage	0.32	0.49	0.61	0.72	0.81
Out of school	0.08	0.05	0.05	0.04	0.05
Observations	8588	7260	9426	5656	9148

# Aims

To gain familiarity in how to:

1. Generate foundational analysis
2. Get analysis into a publishable format (without too much effort)
3. Make graphs that people can (and want to) interpret

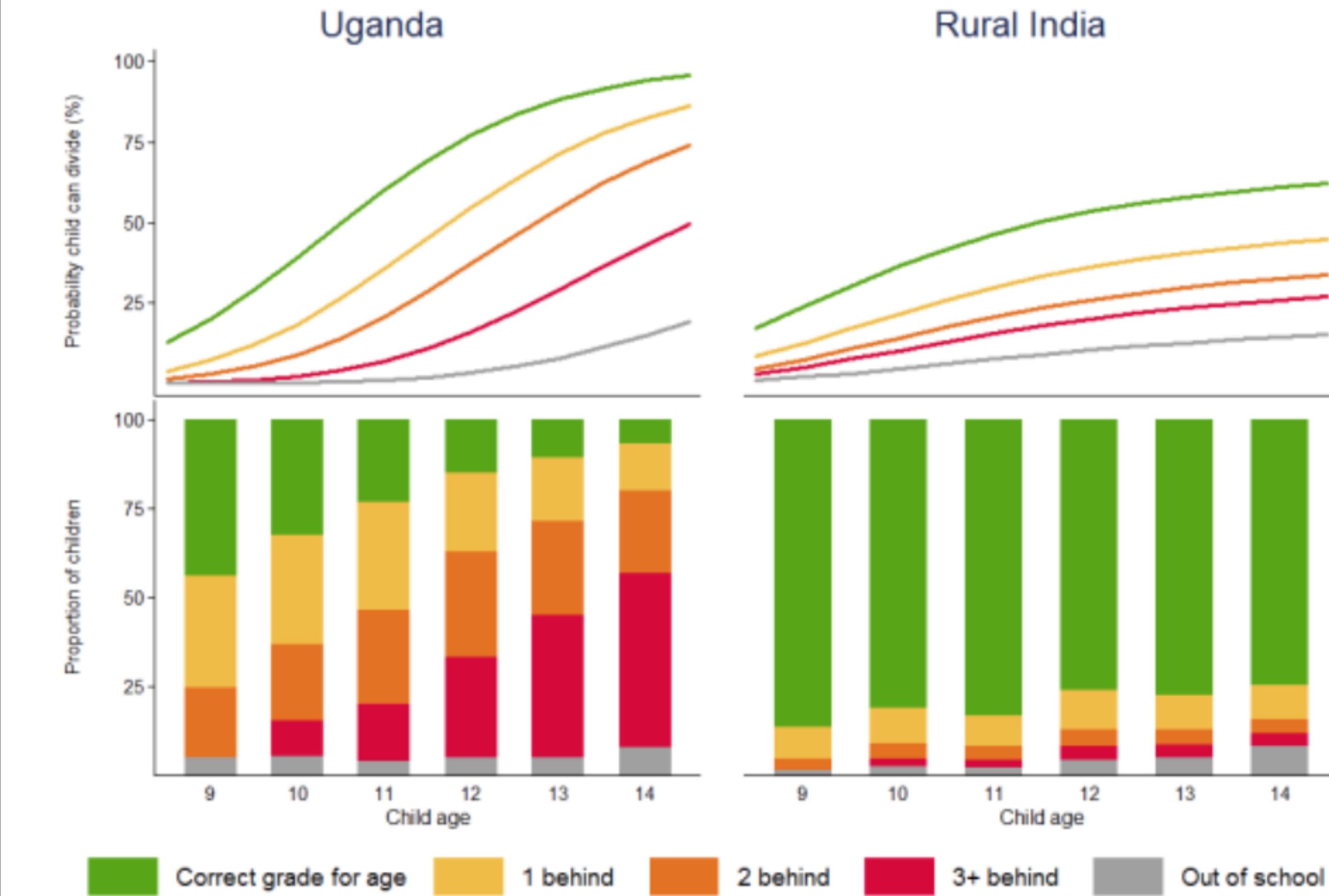
All with Stata!

Background Paper  
The Learning Generation

Overcoming Inequalities Within  
Countries to Achieve Global  
Convergence in Learning

Pauline Rose, Ricardo Sabates, Ben Alcott and Sonia Ilie  
Research for Equitable Access and Learning (REAL) Centre,  
University of Cambridge

Fig. 2.9. In Uganda, the poorest children are far more likely to be learning when progressing through the grades

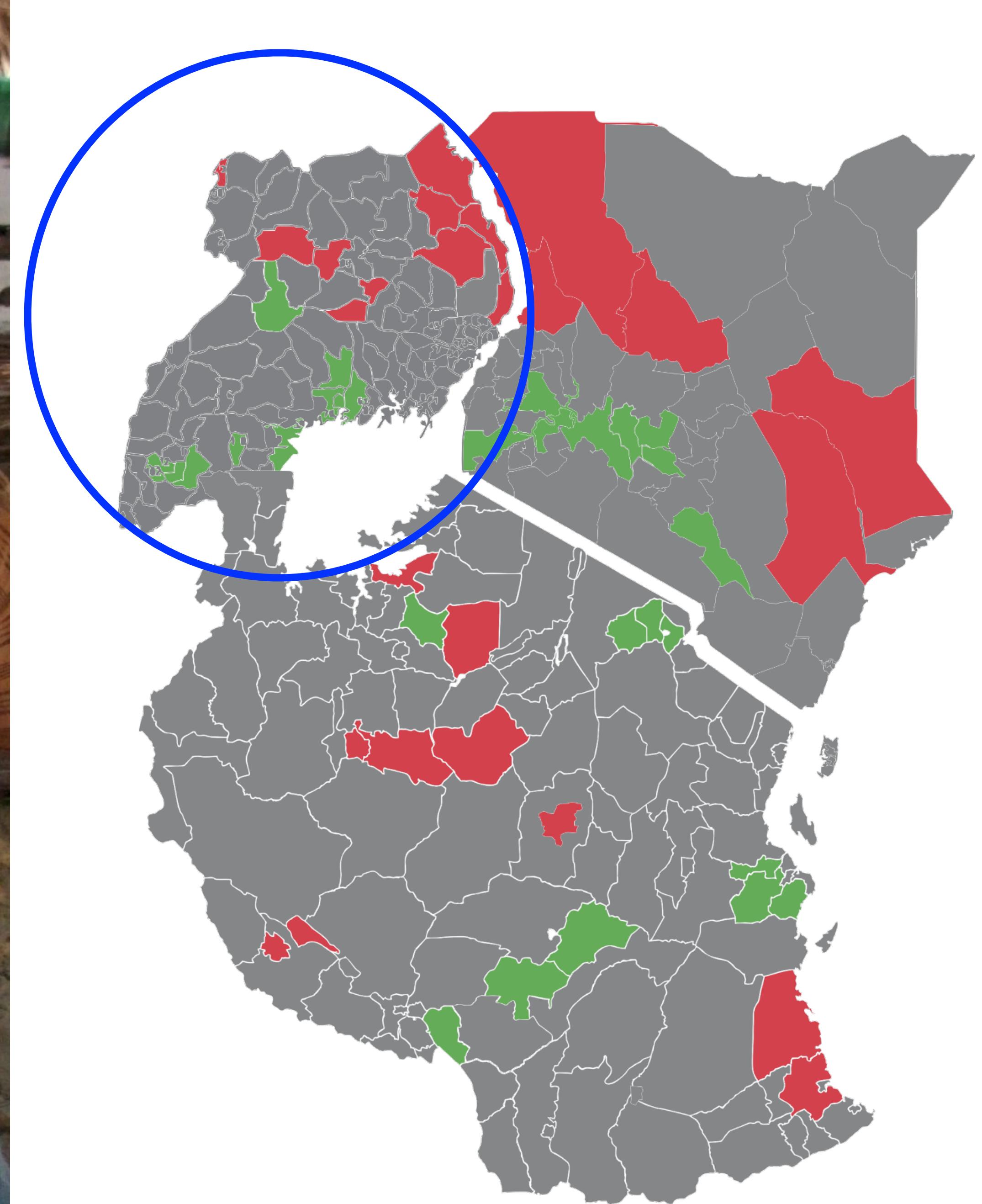


# Today's dataset

Uwezo Uganda

More information available at:  
[uwezo.net](http://uwezo.net)





## MDG 2

### Number of children of primary school age not in school



### Primary school net enrolment rate in sub-Saharan Africa

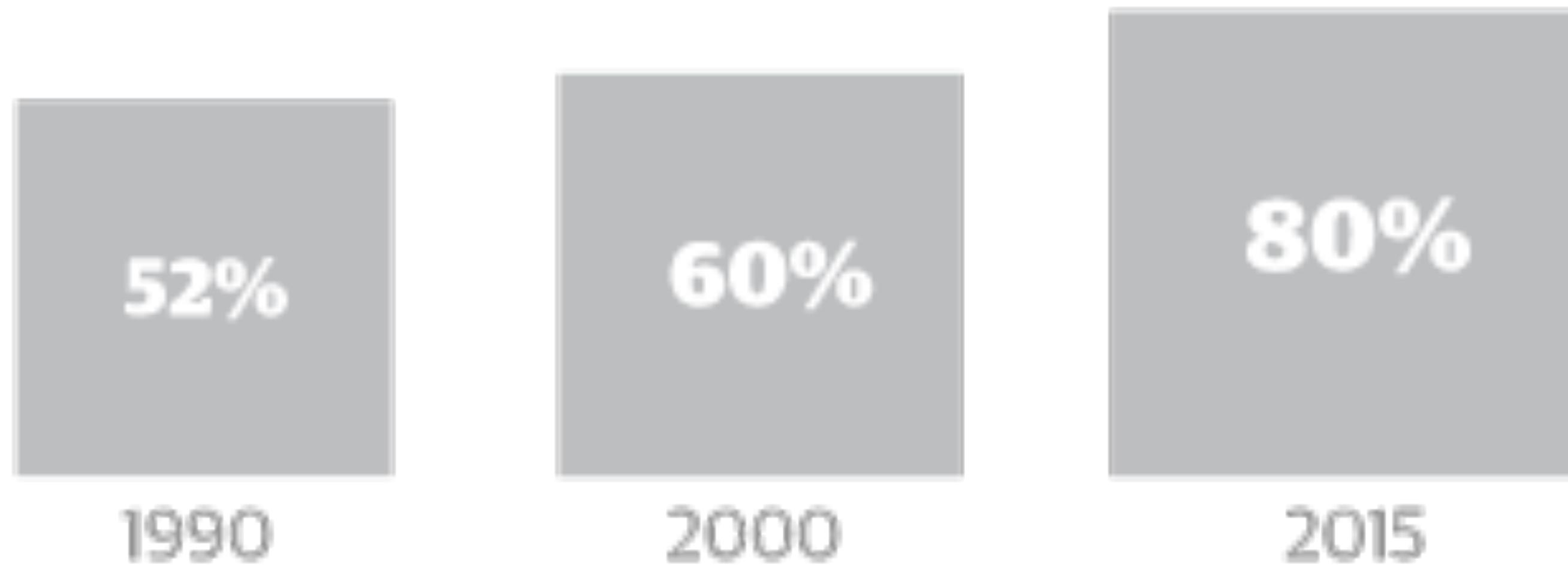
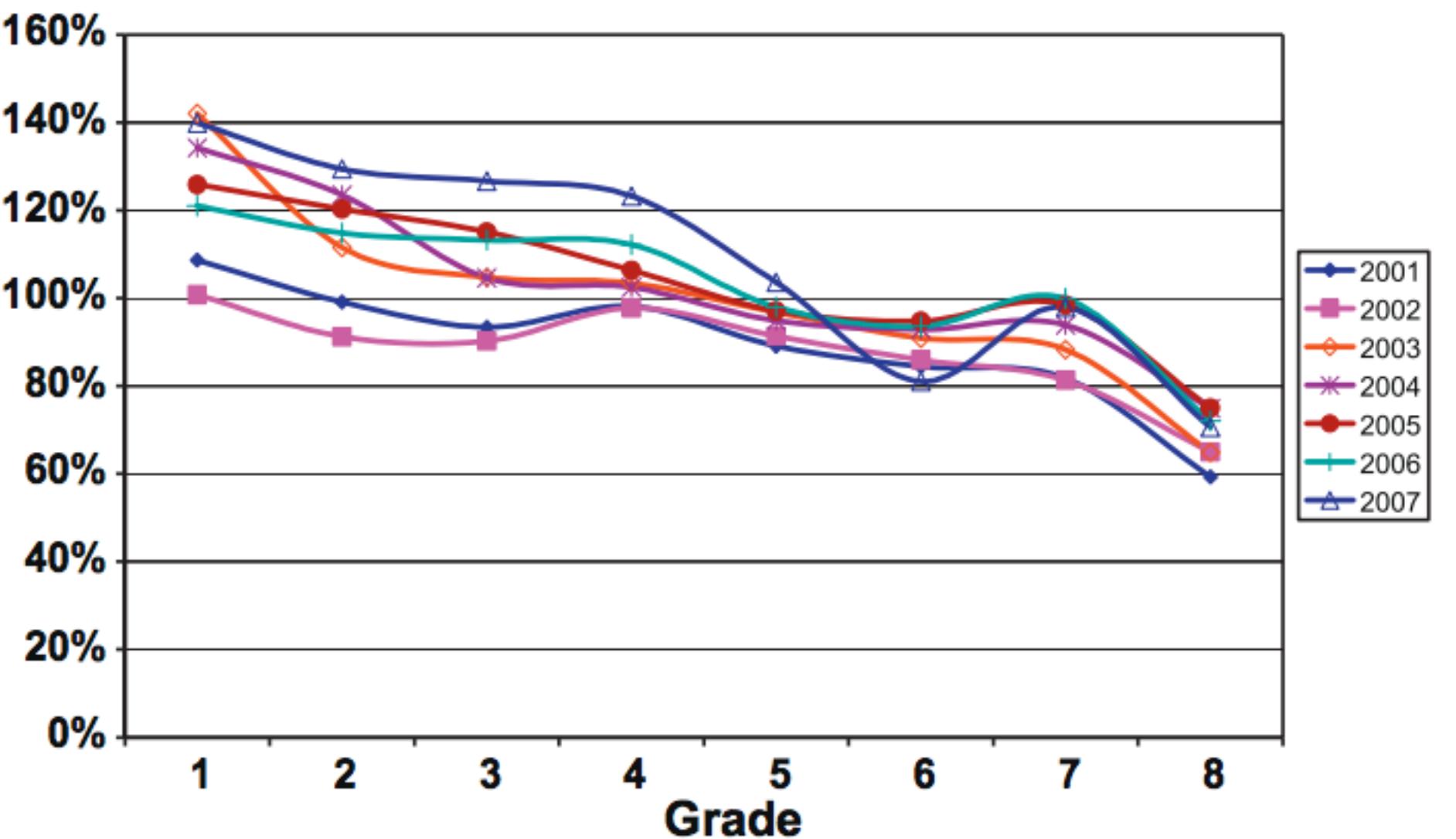
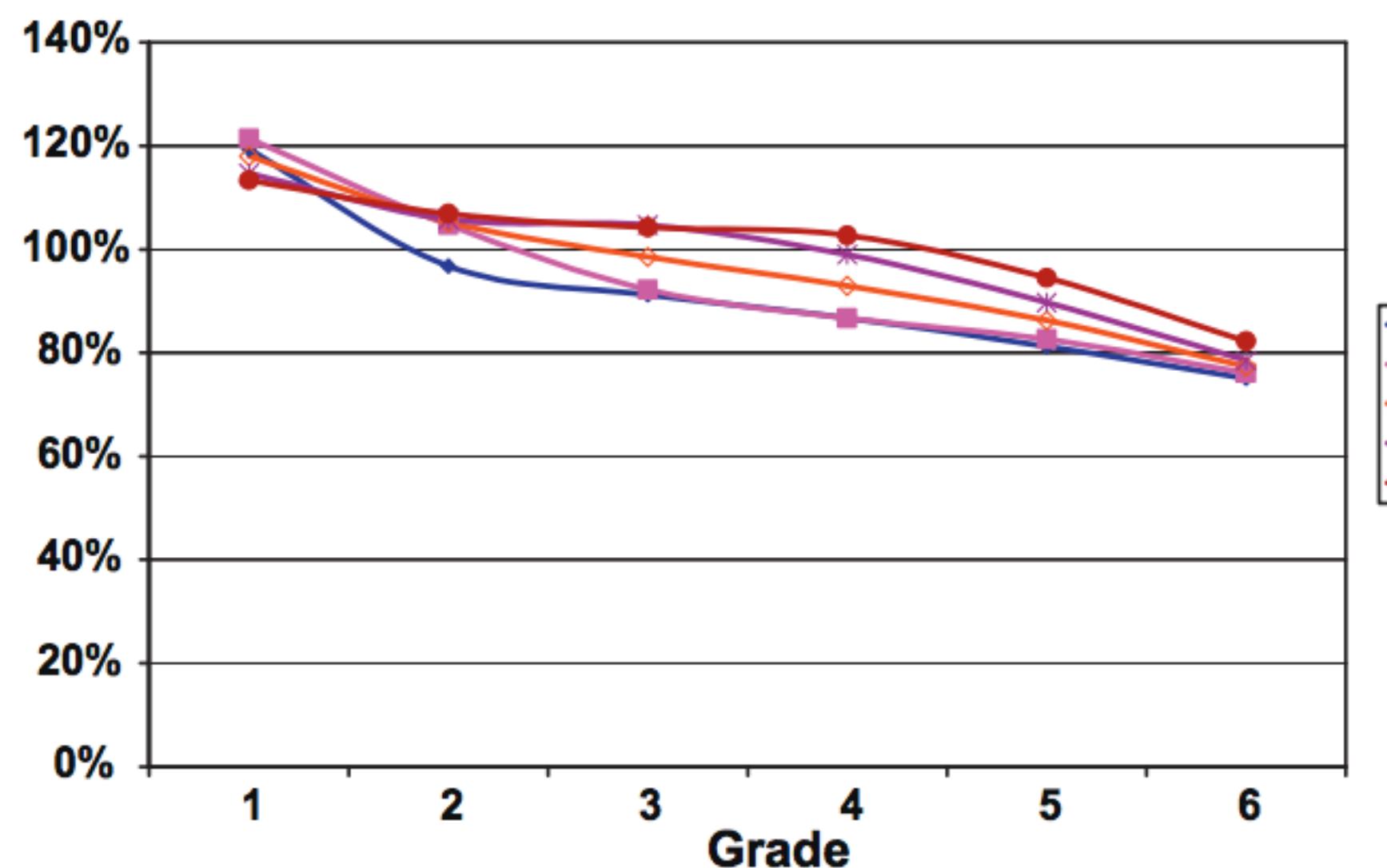


Image source: The Guardian

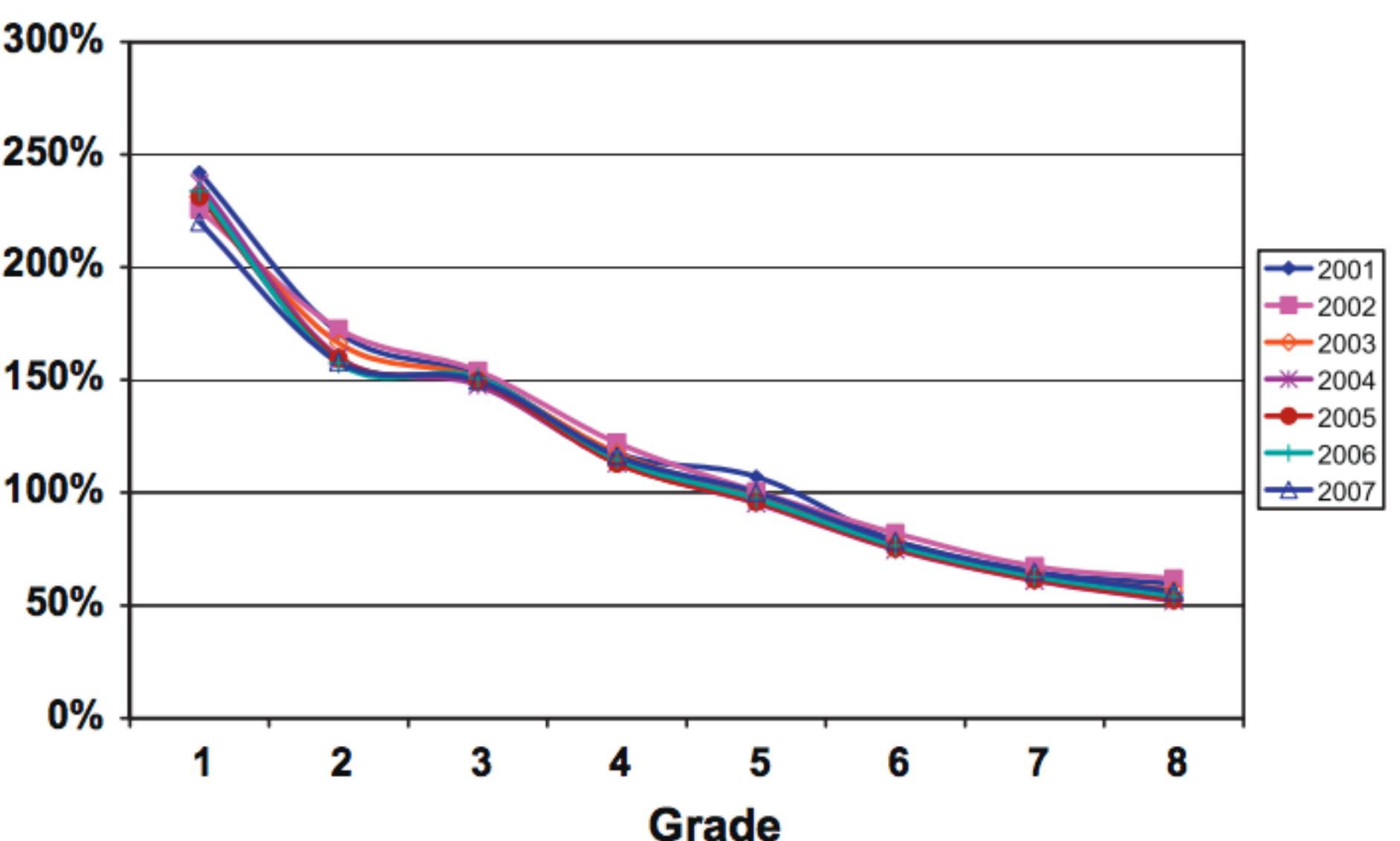
## Kenya



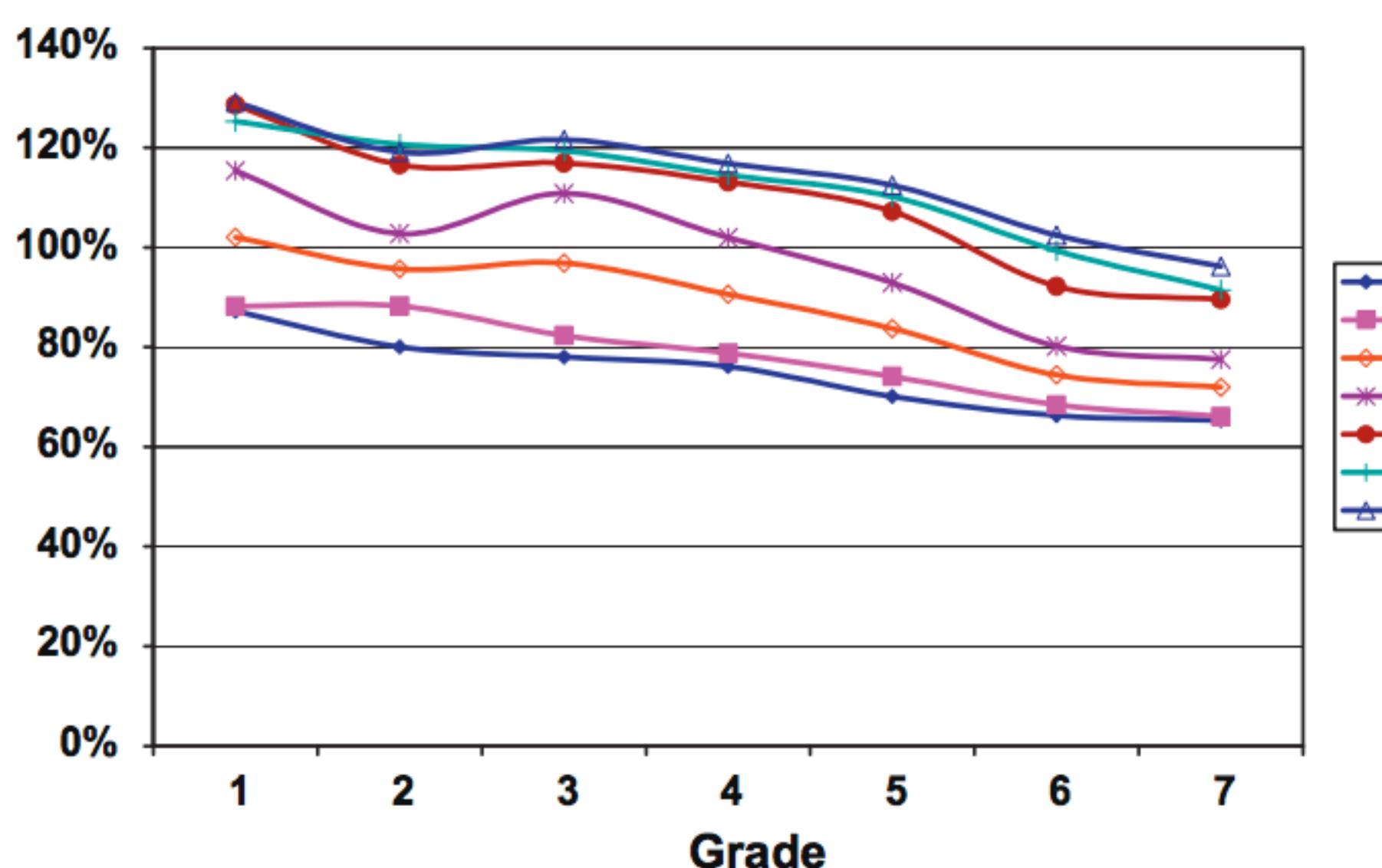
## Nigeria



## Malawi



## Zambia



Lewin, K. M., & Sabates, R. (2012). Who gets what? Is improved access to basic education pro-poor in Sub-Saharan Africa?. *International Journal of Educational Development*, 32(4), 517-528.

Today's research question

How prevalent  
is overage  
status in  
Uganda?



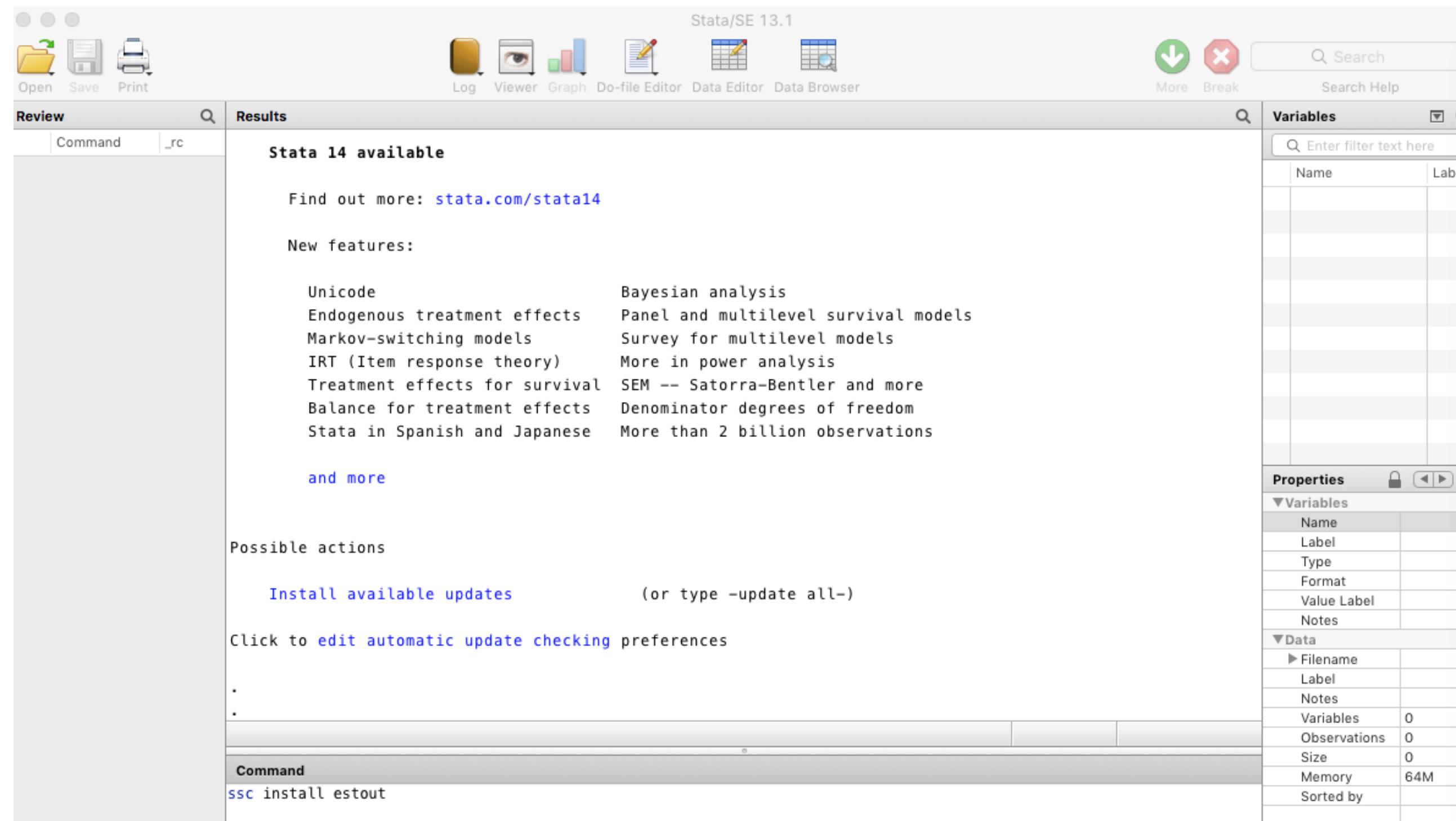
# 1 Download files to desktop and open

- Data file and do file accessible from moodle page ([link](#))
  - Save files to desktop, then open

# MEd Research Methods Strand 16-17

# 2 Install *estout*

```
5 //////////////////////////////////////////////////////////////////
6 /// 2 Install commands to export tables to Word. Only do this once
7 //////////////////////////////////////////////////////////////////
8 *ssc install estout
```



- *estout* will help us transfer tables into word processors
- Copy the command on line 8 (without the asterisk)
- Paste into Stata command box and run
- Only do this once
- The asterisk keeps the command in the do file but stops Stata from running it

# 3 Generate foundational analysis

```
13  
14 * Find the proportion of children who are in  
15 * the correct grade, overage, and out of school  
16 sum correct_age overage out_of_school  
17  
18 * Now do this by child age  
19 tab correct_age age, col  
20 tab overage age, col  
21 tab out_of_school age, col  
22 *alternative to achieve the same  
23 bysort age: sum correct_age overage out_of_school  
24
```

Variable	Obs	Mean	Std. Dev.	Min	Max	
correct_age	40078	.3585009	.4795662	0	1	
overage	40078	.5853336	.4926705	0	1	
out_of_school	40078	.0561655	.2302439	0	1	
Overage	8	9	10	11	12	Total
0	5,849	3,710	3,718	1,571	1,771	16,619
	68.11	51.10	39.44	27.78	19.36	41.47
1	2,739	3,550	5,708	4,085	7,377	23,459
	31.89	48.90	60.56	72.22	80.64	58.53
Total	8,588	7,260	9,426	5,656	9,148	40,078
	100.00	100.00	100.00	100.00	100.00	100.00

# 4 Get analysis into a publishable format

Descriptive statistics, by age

	Age 8	Age 9	Age 10	Age 11	Age 12
	mean	mean	mean	mean	mean
Correct grade	0.60	0.46	0.34	0.24	0.14
Overage	0.32	0.49	0.61	0.72	0.81
Out of school	0.08	0.05	0.05	0.04	0.05
Observations	8588	7260	9426	5656	9148

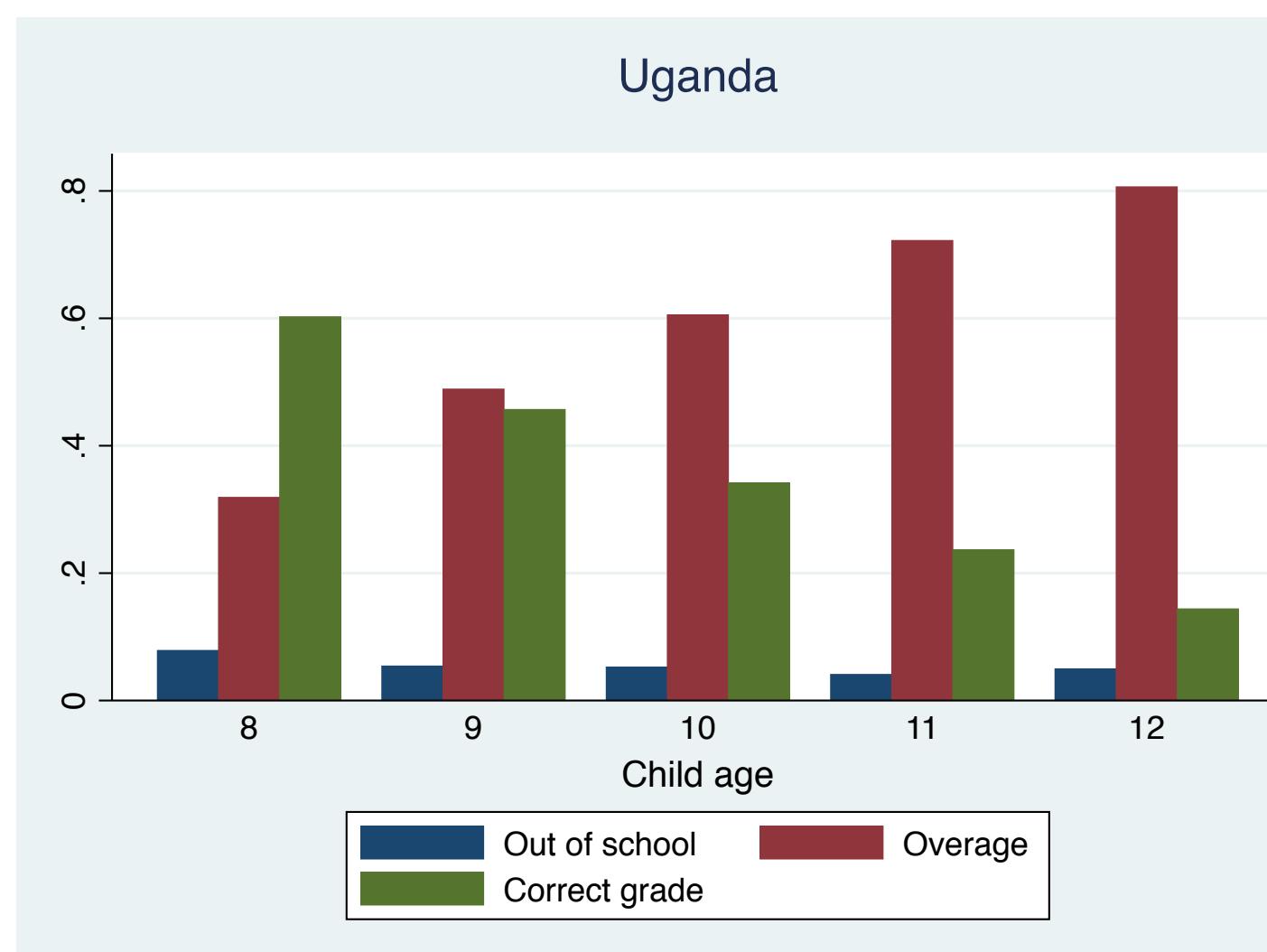
```
39 * Table showing averages by age
40 quietly bysort age: eststo: quietly estpost summarize ///
41 correct_age overage out_of_school, listwise
42 esttab using Table2.rtf, cells("mean (fmt(%9.2f))" ) label ///
43 nonumber mtitle("Age 8" "Age 9" "Age 10" "Age 11" "Age 12") ///
44 title(Descriptive statistics, by age) replace
45 eststo clear
```

Table 1. *Descriptive statistics, by age*

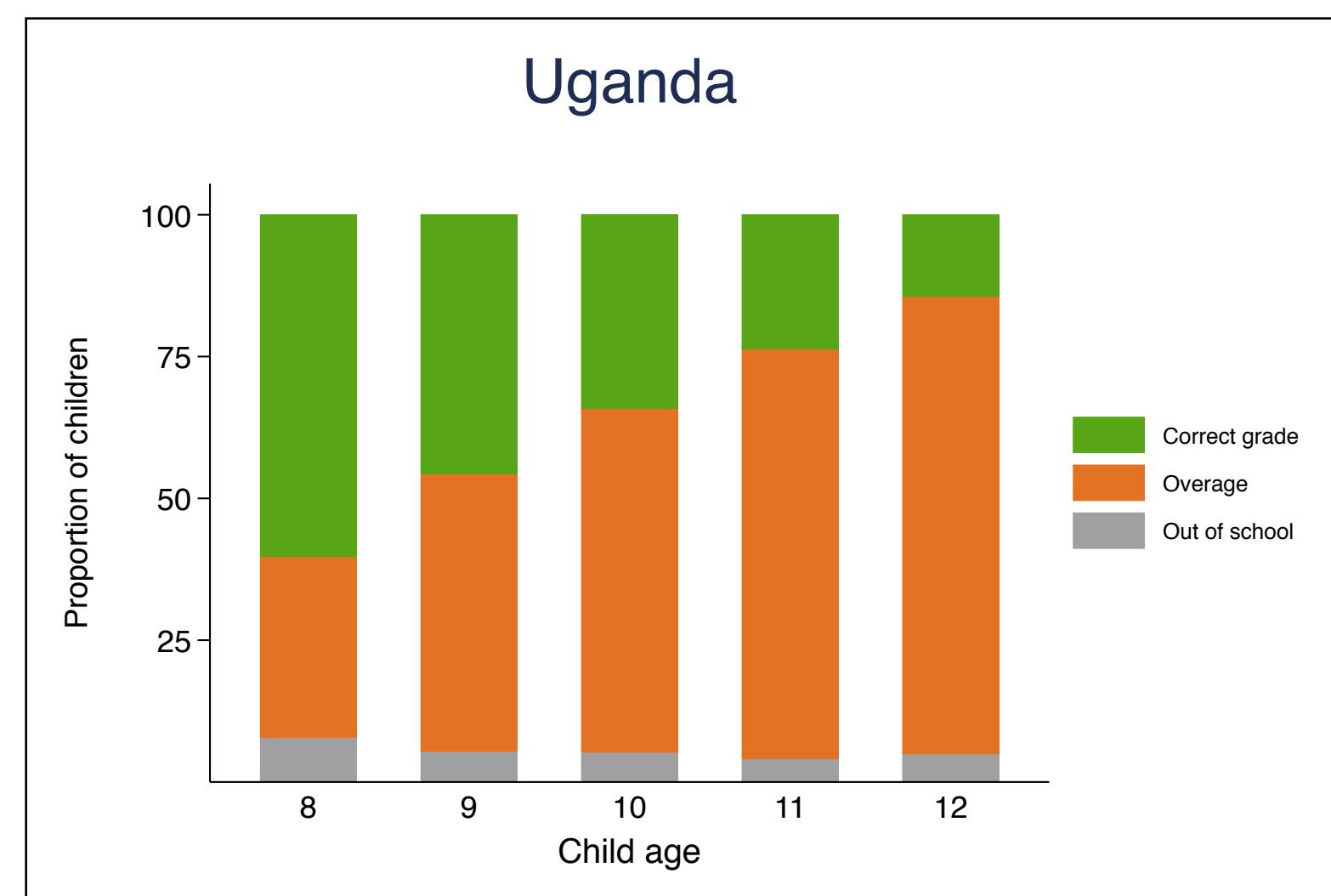
	Age 8	Age 9	Age 10	Age 11	Age 12
Correct grade	0.60	0.46	0.34	0.24	0.14
Overage	0.32	0.49	0.61	0.72	0.81
Out of school	0.08	0.05	0.05	0.04	0.05
Observations	8588	7260	9426	5656	9148

# 5 Make graphs that people can interpret

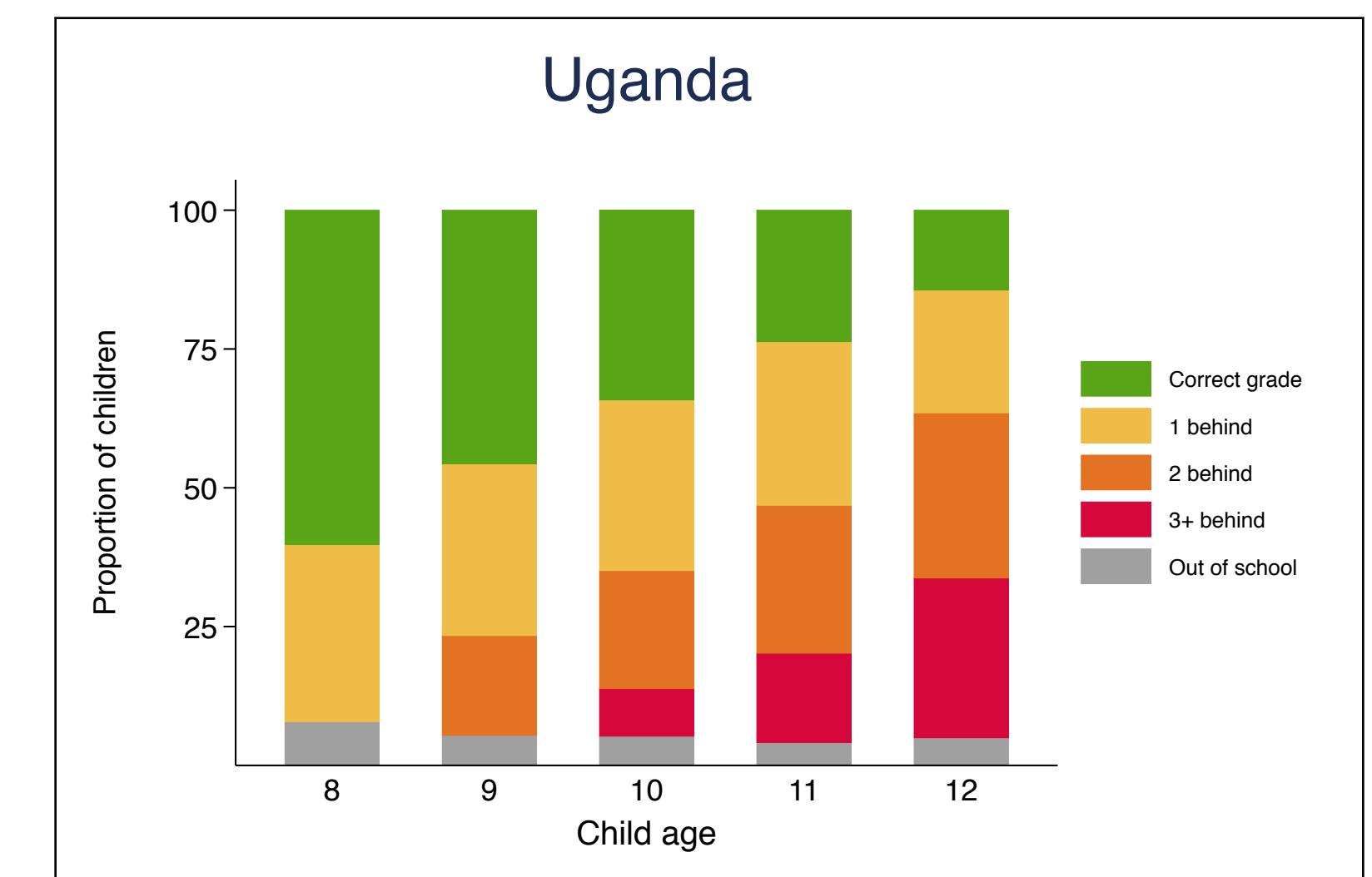
Basic graph



Same graph reformatted  
to ease interpretation



Same graph reformatted  
with different levels of overage



## Uganda

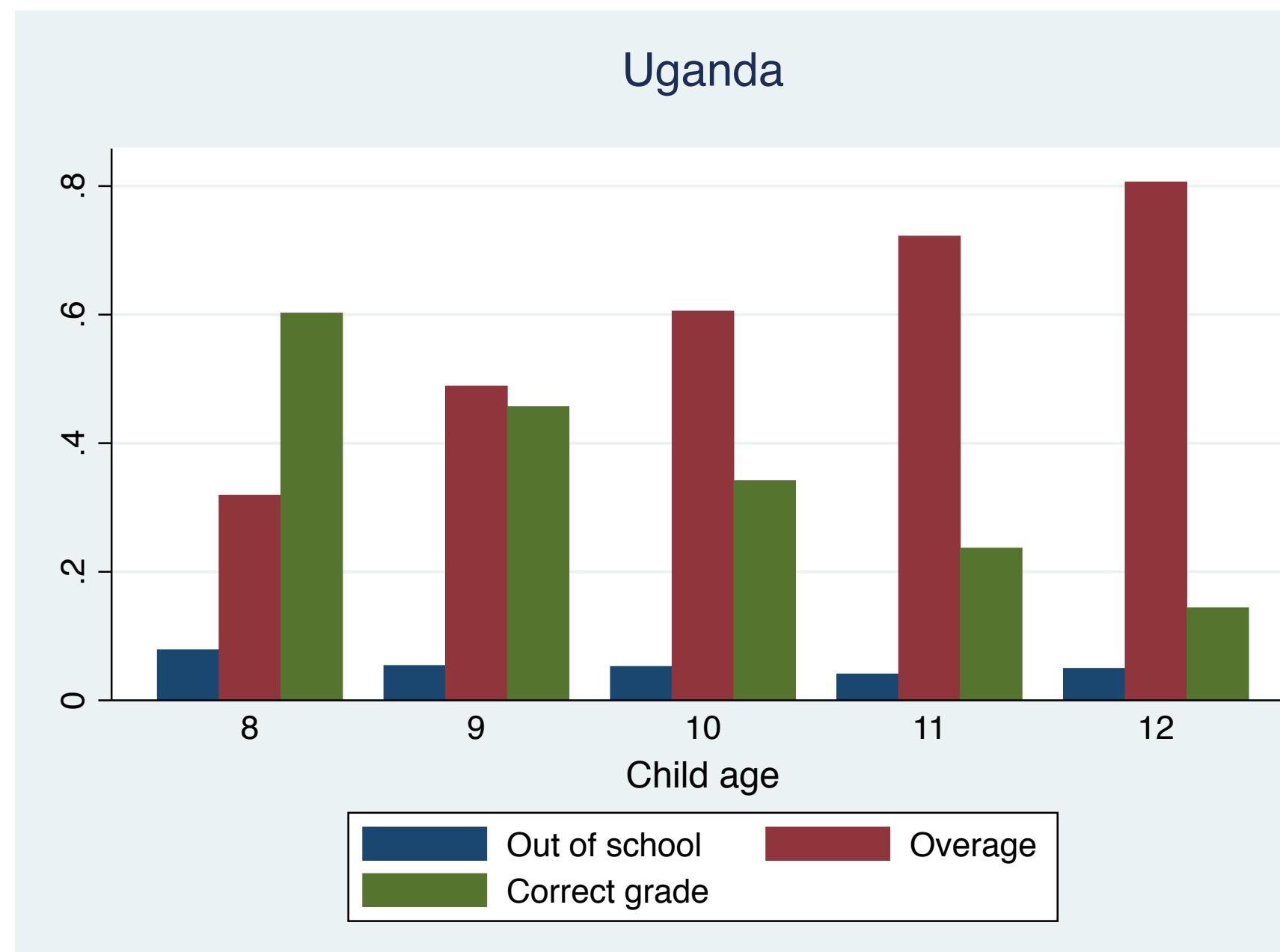
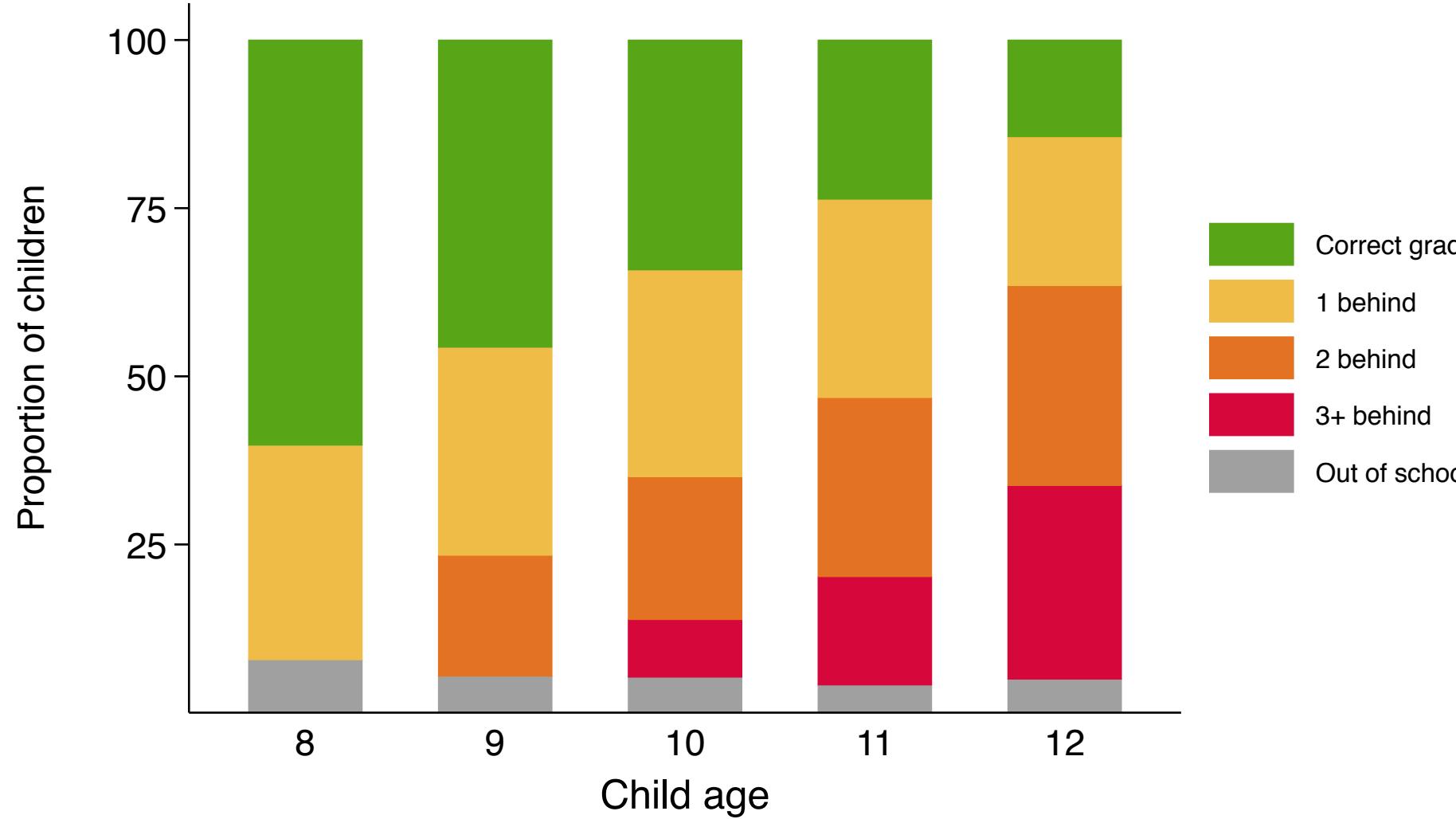


Table 1. Descriptive statistics, by age

	Age 8	Age 9	Age 10	Age 11	Age 12
Correct grade	0.60	0.46	0.34	0.24	0.14
Overage	0.32	0.49	0.61	0.72	0.81
Out of school	0.08	0.05	0.05	0.04	0.05
Observations	8588	7260	9426	5656	9148