What we can learn from the data underlying the District Driven Dashboards (DDD)?

Servaas van der Berg, Chris van Wyk,

Rebecca Selkirk & Kate Rich

Resep, Stellenbosch University
Quantitative Education Research Conference
Stellenbosch, 12 September 2019

RESEP
Research on Socio-Economic Policy

UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY

100
1918-2018

SA-SAMS, DDD and ASS ('Emis')

❑ **Longitudinal data** is needed to track individual learners through the education system from year to year

❑ With such longitudinal data one can determine exactly how many learners of a specific cohort:

  o progressed through the system without any repetition

  o are still in the system after repeating once or more

  o dropped out of the system

With unit-level learner records key questions can be answered such as

❑ "What is the **profile** of learners who dropped out of the system, or

❑ What is the **profile** of learners who progressed without any repetition?"

UNIQUE LEARNER IDENTIFIER

Importance of a unique learner identifier

Learner absence by days of the week in a sample of schools, 2018

| Day of the Week	Total	%
Monday	4 548 410	16.7
Tuesday	5 024 677	18.5
Wednesday	5 698 293	20.9
Thursday	5 331 536	19.6
Friday	6 616 986	24.3
Total	17 219 902	100.0

AGGREGATED DATA

VS

UNIT-LEVEL DATA

Enrolment in a GAUTENG sample of schools based on SA-SAMS data, 2015-2018

| year	Gr1	Gr2	Gr3	Gr4	Gr5	Gr6	Gr7	Gr8	Gr9	Gr10	Gr11	Gr12
2015	138 891	127 496	118 977	115 773	105 655	95 438	89 757	80 416	74 832	98 413	76 011	50 087
2016	141 364	136 297	127 701	125 628	115 702	109 076	103 333	90 961	79 770	103 471	76 726	56 508
2017	141 027	137 596	133 467	130 667	121 898	114 478	110 591	86 816	81 789	108 644	78 024	67 401
2018	146 652	136 511	136 169	137 371	126 541	121 489	118 305	96 865	84 505	111 826	78 864	65 157

Cross-sectional analysis of SA-SAMS data (2015-2018): Gauteng sample
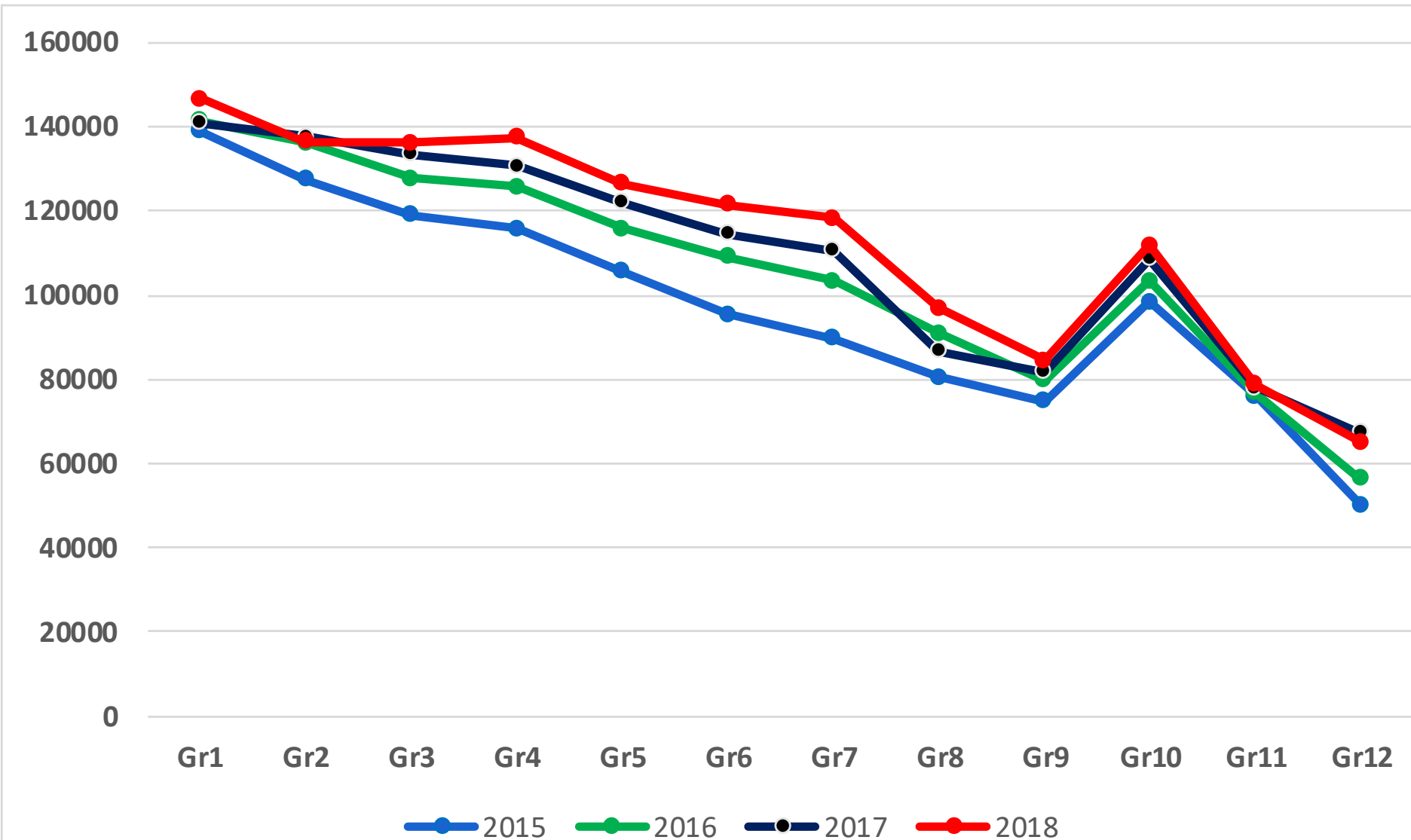
Using Pseudo-cohorts in analysing the SA-SAMS data (2015-2018): Gauteng sample
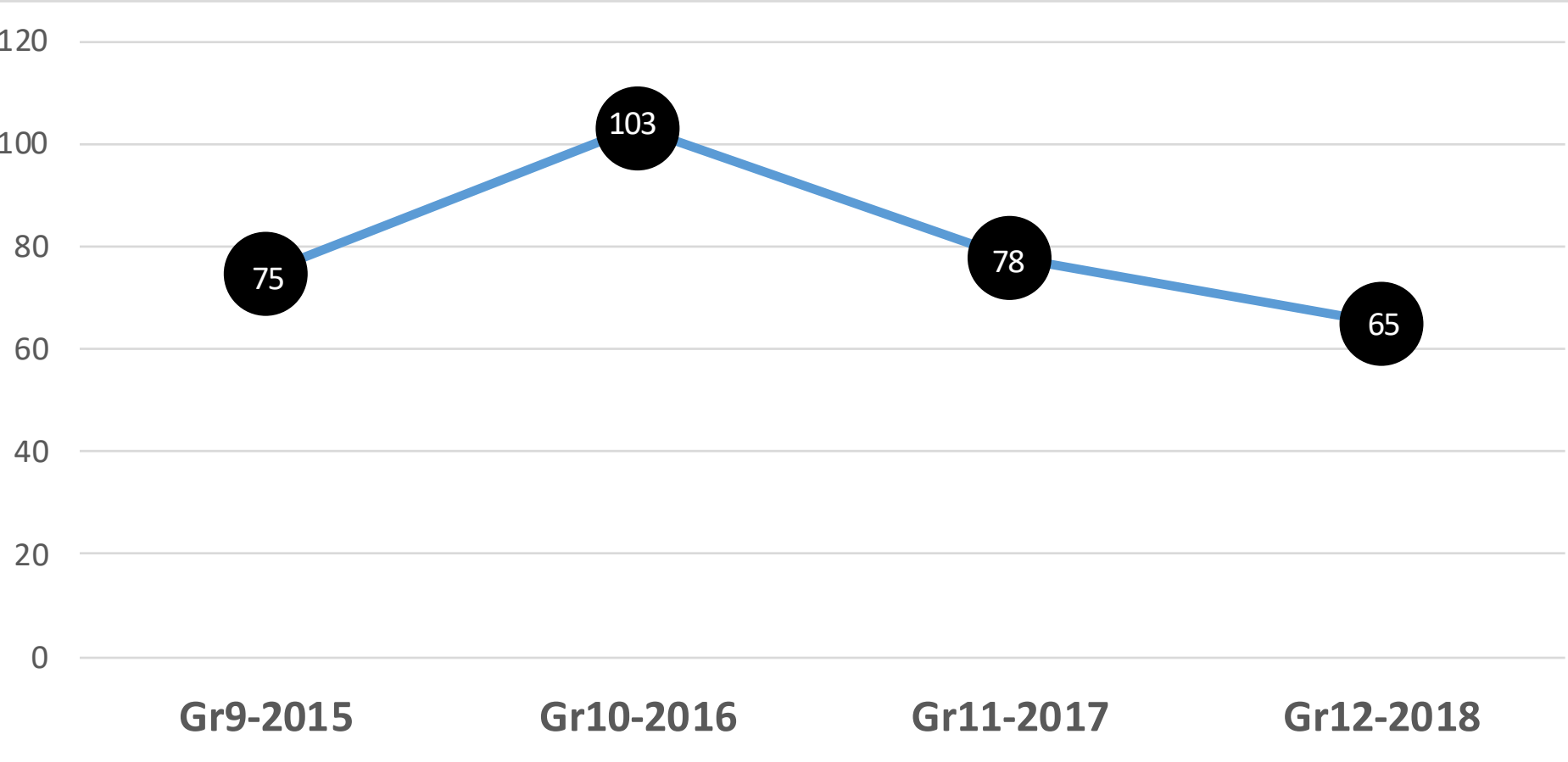
Pseudo vs actual 'on track' grade progression for the 2015 Gr9 cohort: Gauteng sample
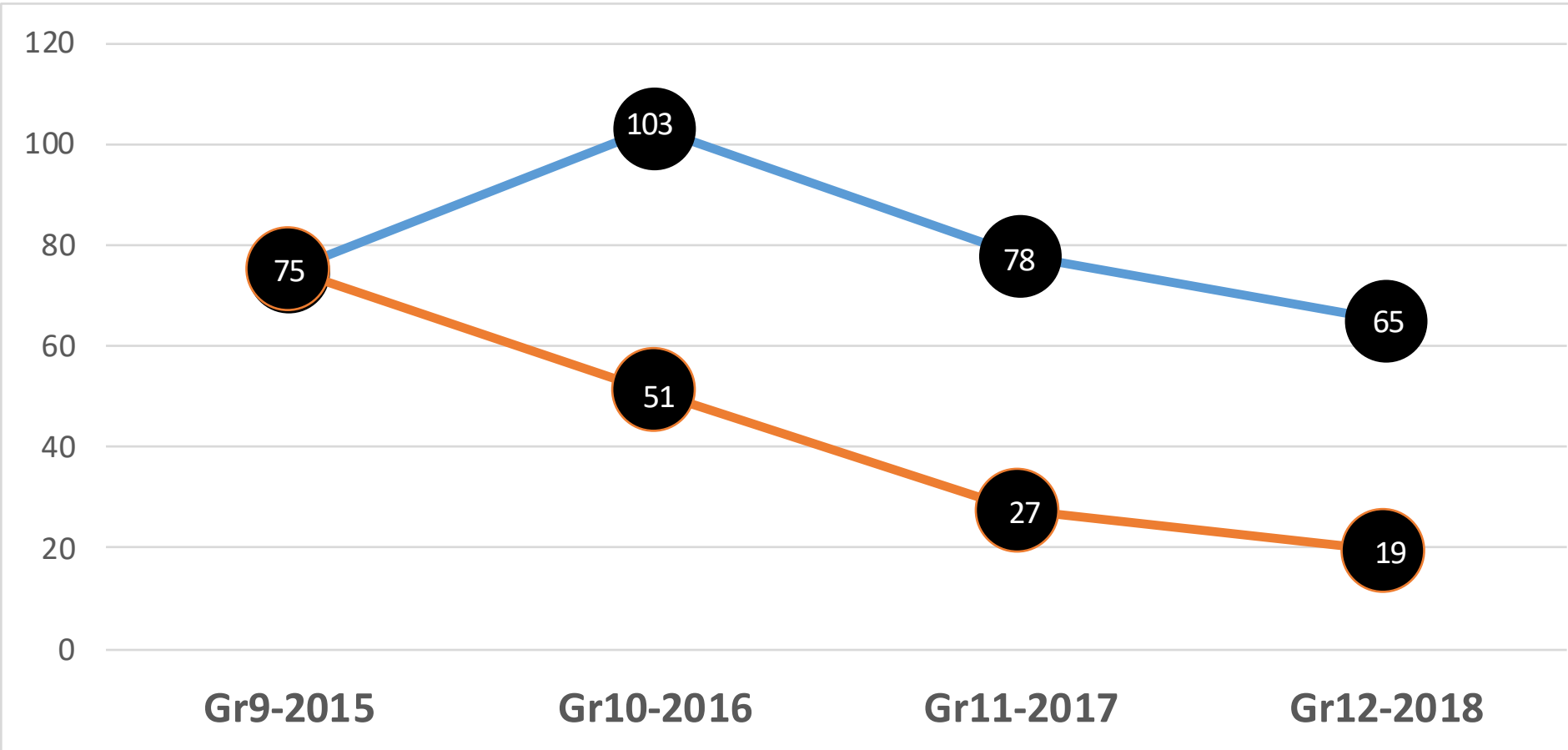
Progression of learners that enrolled in Grade 9 in 2015 for the period 2015-2018: Gauteng sample

Grade	2015	2016	2017	2018
Gr9	75 259	10 536	1 158	199
Gr10	0	51 001	27 421	10 021
Gr11	0	885	27 414	21 149
Gr12	0	19	625	19 453
Still in school	75 259	62 441	56 618	50 822
Repeaters	0	10 536	28579	31 369
Dropout	0	12 818	5 823	5 796
Cum Dropout	0	12 818	18 641	24 437

Data issues

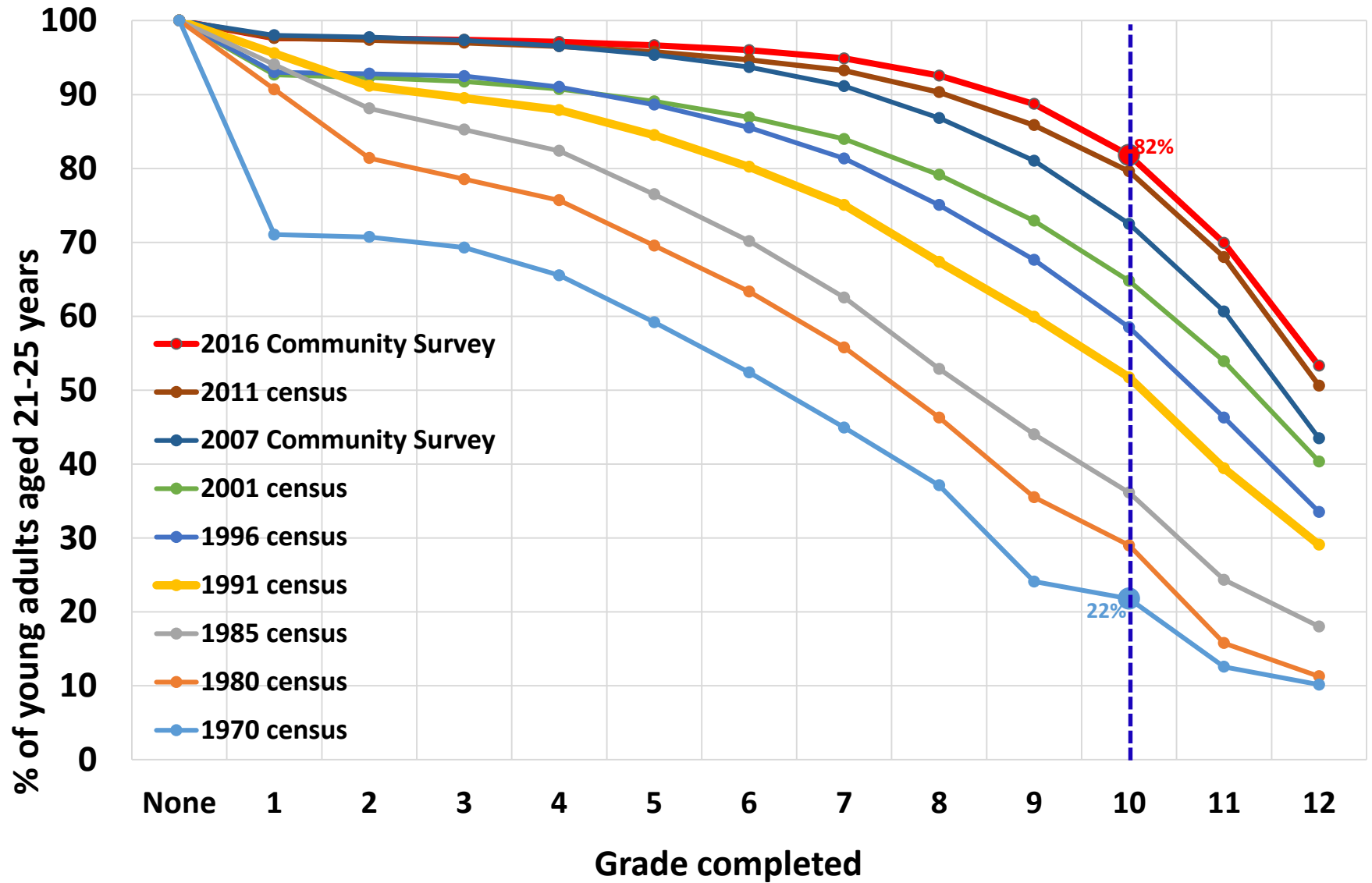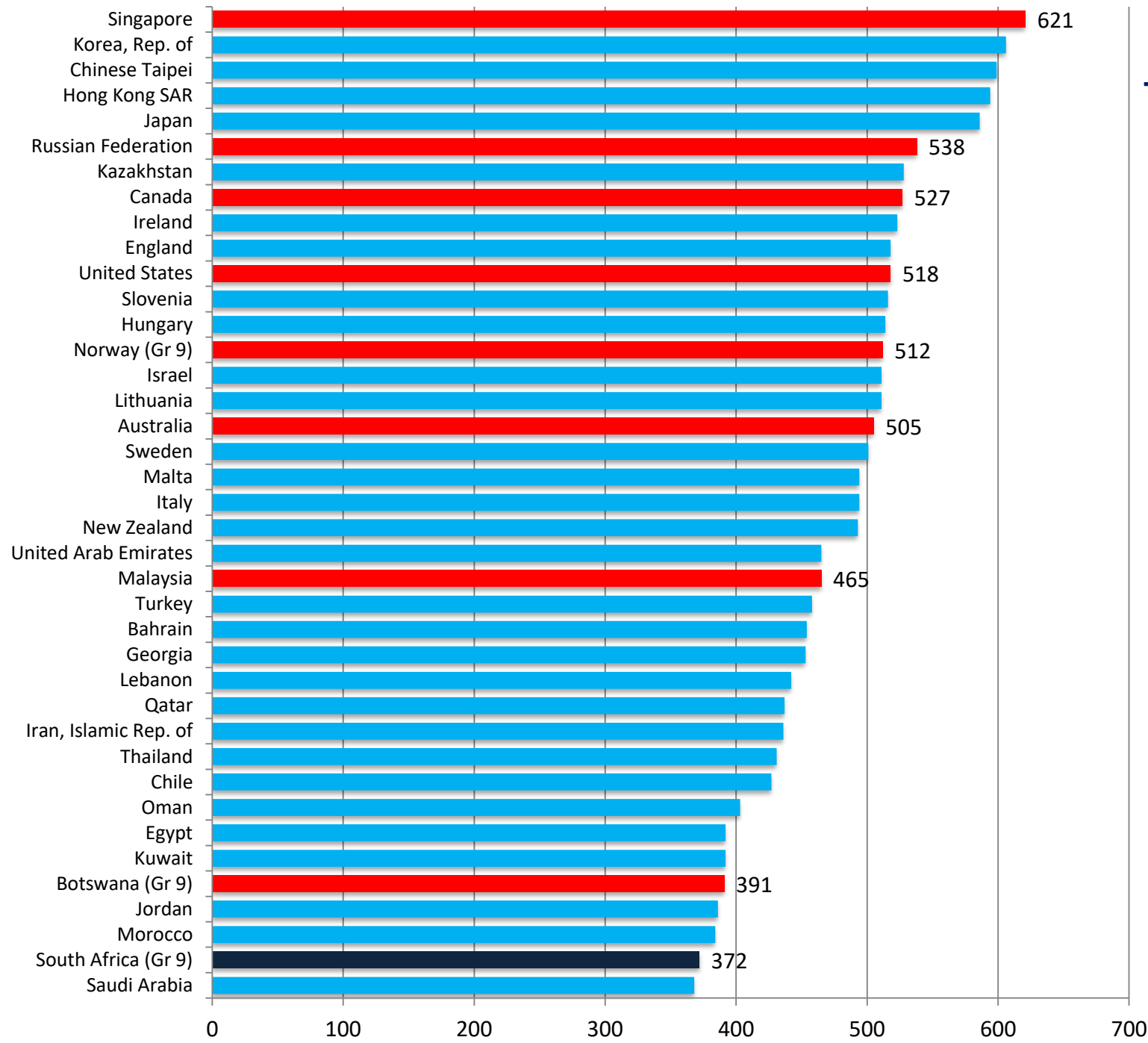**Data Inconsistencies**:

❑ Far fewer schools submitted SA-SAMS data in 2015 than in 2018

❑ It is important to track the same schools over time in the cohort analysis

❑ We only analyse time trends for schools that submitted data every year from 2015 to 2018. This leaves fewer schools in the analysis. The cohort followed from 2015 is much smaller than it would have been if we had followed a cohort beginning from 2016, but then the period would have been shorter (only 3 years).

   o Sample Selection Effect: A sample selection bias is introduced by the selection of schools based on schools that submitted data, as they are not randomly distributed.

   o Thus the sample obtained is not representative of schools in the particular provinces.

Schools by year and quintile: Gauteng, both actual (EMIS 2018) and for the DDD sample (2015-18)

Year	Q1	Q2	Q3	Q4	Q5	NA	Total
2015	231	202	422	273	161	102	1391
2016	290	249	499	349	385	292	2064
2017	289	245	500	368	539	445	2386
2018	297	257	509	377	605	485	2530
Same schools (2015-2018)	**225**	**199**	**410**	**270**	**155**	**93**	**1352**
EMIS:2018	**275**	**257**	**416**	**477**	**641**	**814**	**2880**

% of young adults aged 20-25 that have completed different grades, 1970-2016

Legend:
- 2016 Community Survey
- 2011 census
- 2007 Community Survey
- 2001 census
- 1996 census
- 1991 census
- 1985 census
- 1980 census
- 1970 census

Y-axis: % of young adults aged 21-25 years

X-axis: Grade completed (None, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)

82%

22%

TIMSS Gr8 Maths scores, 2015
Setpoint (original average) = 500, standard deviation = 100)

Country	Score
Singapore	621
Korea, Rep. of	
Chinese Taipei	
Hong Kong SAR	
Japan	
Russian Federation	538
Kazakhstan	
Canada	527
Ireland	
England	
United States	518
Slovenia	
Hungary	
Norway (Gr 9)	512
Israel	
Lithuania	
Australia	505
Sweden	
Malta	
Italy	
New Zealand	
United Arab Emirates	
Malaysia	465
Turkey	
Bahrain	
Georgia	
Lebanon	
Qatar	
Iran, Islamic Rep. of	
Thailand	
Chile	
Oman	
Egypt	
Kuwait	
Botswana (Gr 9)	391
Jordan	
Morocco	
South Africa (Gr 9)	372
Saudi Arabia	

Data underlying DDD offers insights into:

- Patterns of progression

- Who drops out in Gr10

- The usefulness of school-based assessment

- How current performance predicts progression and future performance

- Does school absence predict dropout, or performance?

- How subject choice for Maths/MathsLit in Gr10 and Gr12 is associated with earlier, and later, performance
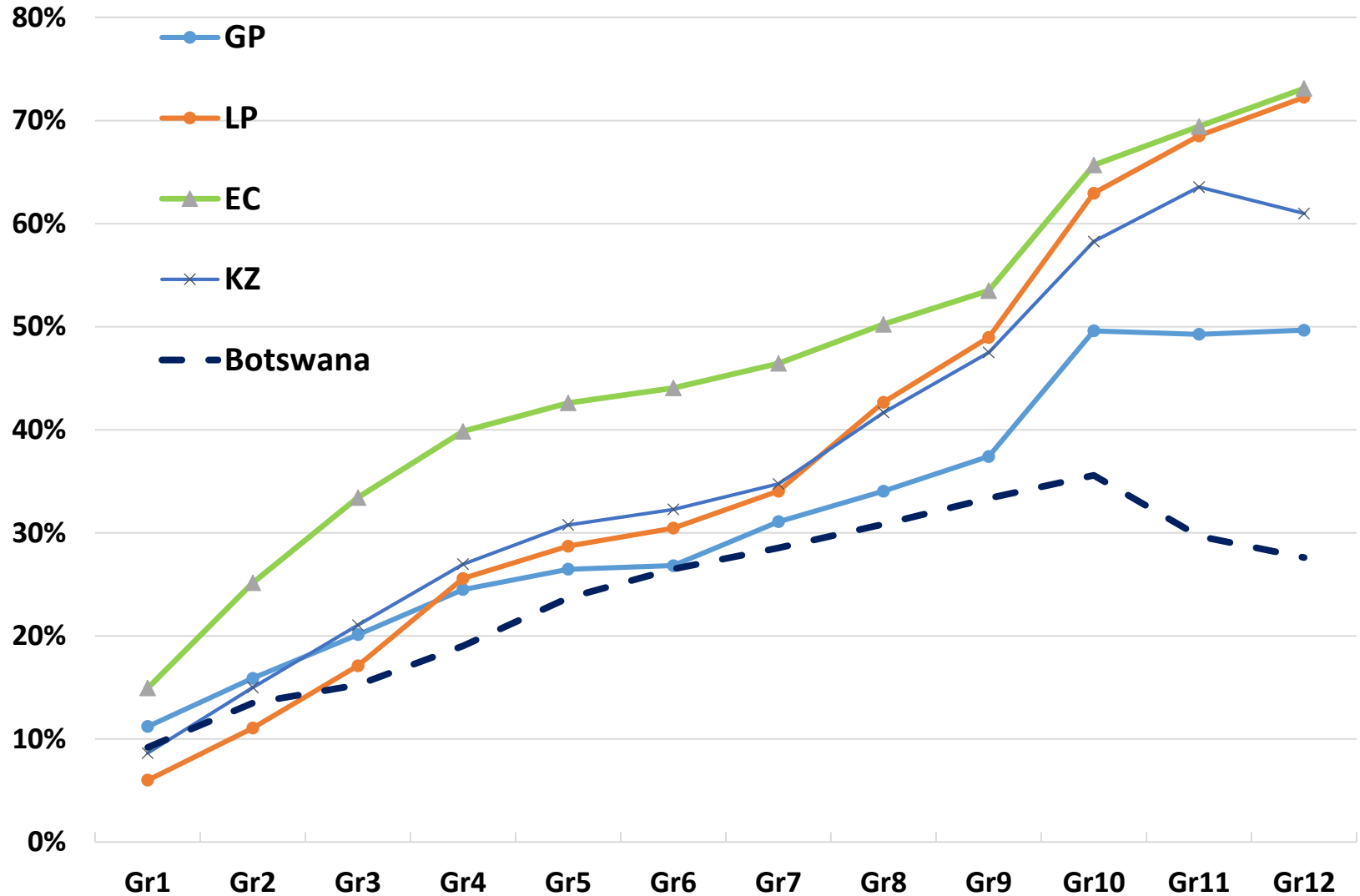
Age by grade in Eastern Cape sample, 2018

| Age	Gr1	Gr2	Gr3	Gr4	Gr5	Gr6	Gr7	Gr8	Gr9	Gr10	Gr11	Gr12
1	3		2	1					1			
2	2	1	6	2								
3	1	1	2	1		1	1					
4	32	3	2	4	3		1	2				
5	834	26	15	6	3		7	2	1	2		
6	58 950	520	22	22	11	7	3	3	3			
7	97 641	41 296	511	33	6	4	4	3	1			
8	22 388	86 238	32 450	458	20	14	4	8	3	1		1
9	3 163	32 137	79 253	29 270	517	30	21	6	2	2	2	1
10	854	7 443	39 429	75 856	27 473	598	31	8	5	3	3	1
11	421	1 688	10 960	42 650	66 368	25 762	657	16	4	4	1	1
12	261	670	3 175	17 448	39 869	61 874	24 888	690	25	8	4	
13	166	376	1 010	6 069	18 368	38 249	56 629	22 029	741	17	8	2
14	130	245	503	2 257	7 107	18 387	36 436	49 605	19 402	677	9	5
15	77	147	223	823	2 661	7 495	18 652	31 871	40 963	15 159	482	18
16	54	77	205	380	1 077	3 119	8 649	19 388	29 000	36 191	10 271	472
17	48	60	197	179	468	1 274	4 192	10 959	18 566	31 191	26 933	7 704
18	24	48	170	107	262	619	1 987	6 022	11 503	25 456	24 816	23 491
19	30	46	144	56	139	218	840	2 861	6 385	19 012	21 528	24 425
20	14	35	107	20	55	90	323	1 094	2 891	11 933	16 289	21 431
21	13	27	95	12	34	45	139	509	1 201	6 418	10 855	16 560
22	7	24	59	3	13	22	62	207	439	3 145	6 379	11 503
23	5	16	45	6	13	10	20	75	184	1 298	3 057	6 199
24	8	9	34	1	4	6	8	46	85	565	1 432	3 195
25	1	3	6	1	2	3	3	23	28	217	617	1 388
25+	10	3	10	4	5	-	8	24	56	412	646	1 514
Total	**185 137**	**171 139**	**168 635**	**175 669**	**164 478**	**157 827**	**153 565**	**145 451**	**131 489**	**151 711**	**123 332**	**117 911**
Over-age	27 674	43 054	56 372	70 016	70 077	69 537	71 319	73 079	70 338	99 647	85 619	86 215
% over-age	15%	25%	33%	40%	43%	44%	46%	50%	53%	66%	69%	73%

Correct age (6 or 7 in Gr1)   **Over-age (8 or older in Gr1)**

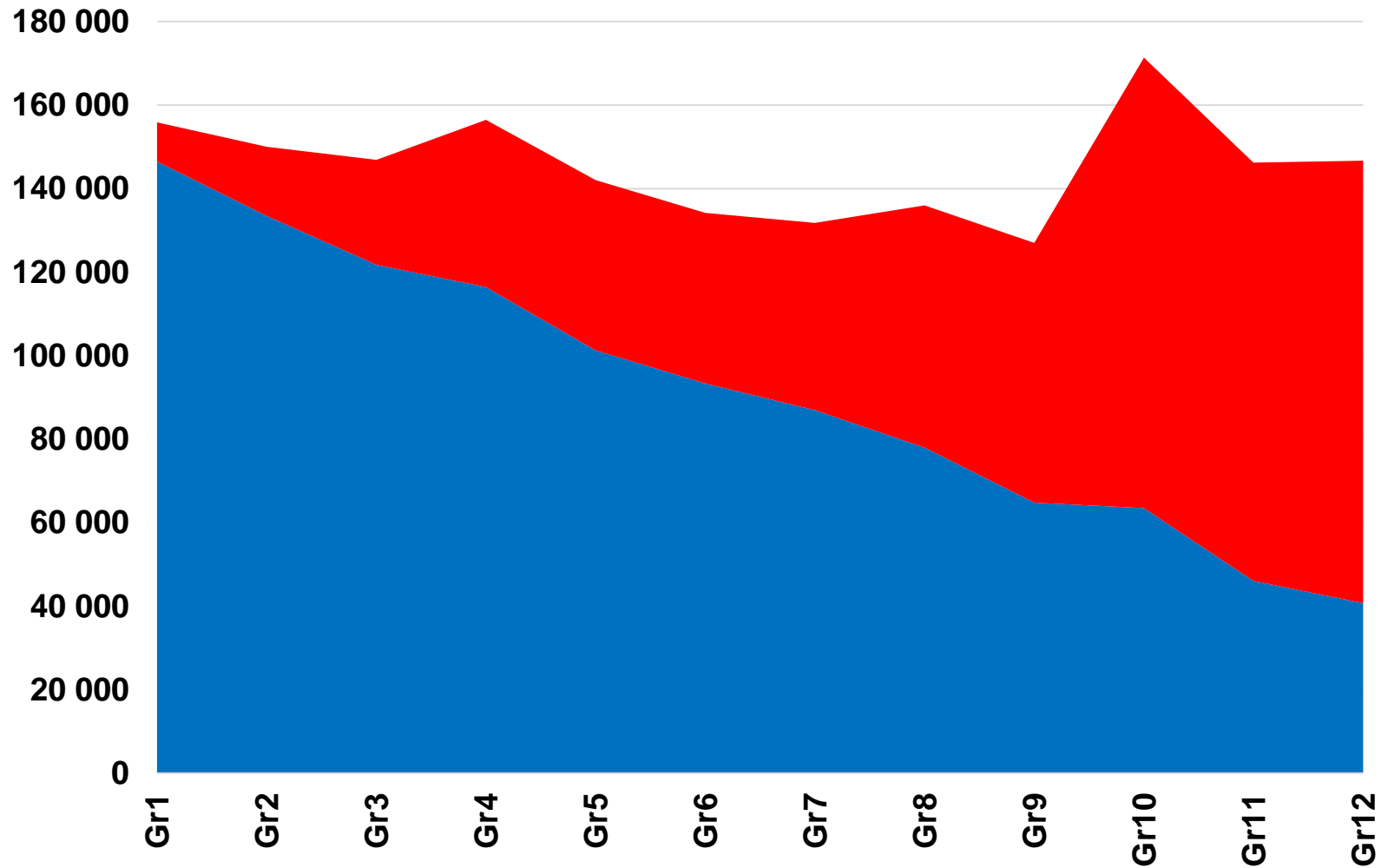# Over-age by grade and province, 2018
(with comparison to Botswana, 2014)

Comparing enrolment patterns in Limpopo sample and Zimbabwe, 2018

Note the peaks in the first year of each phase

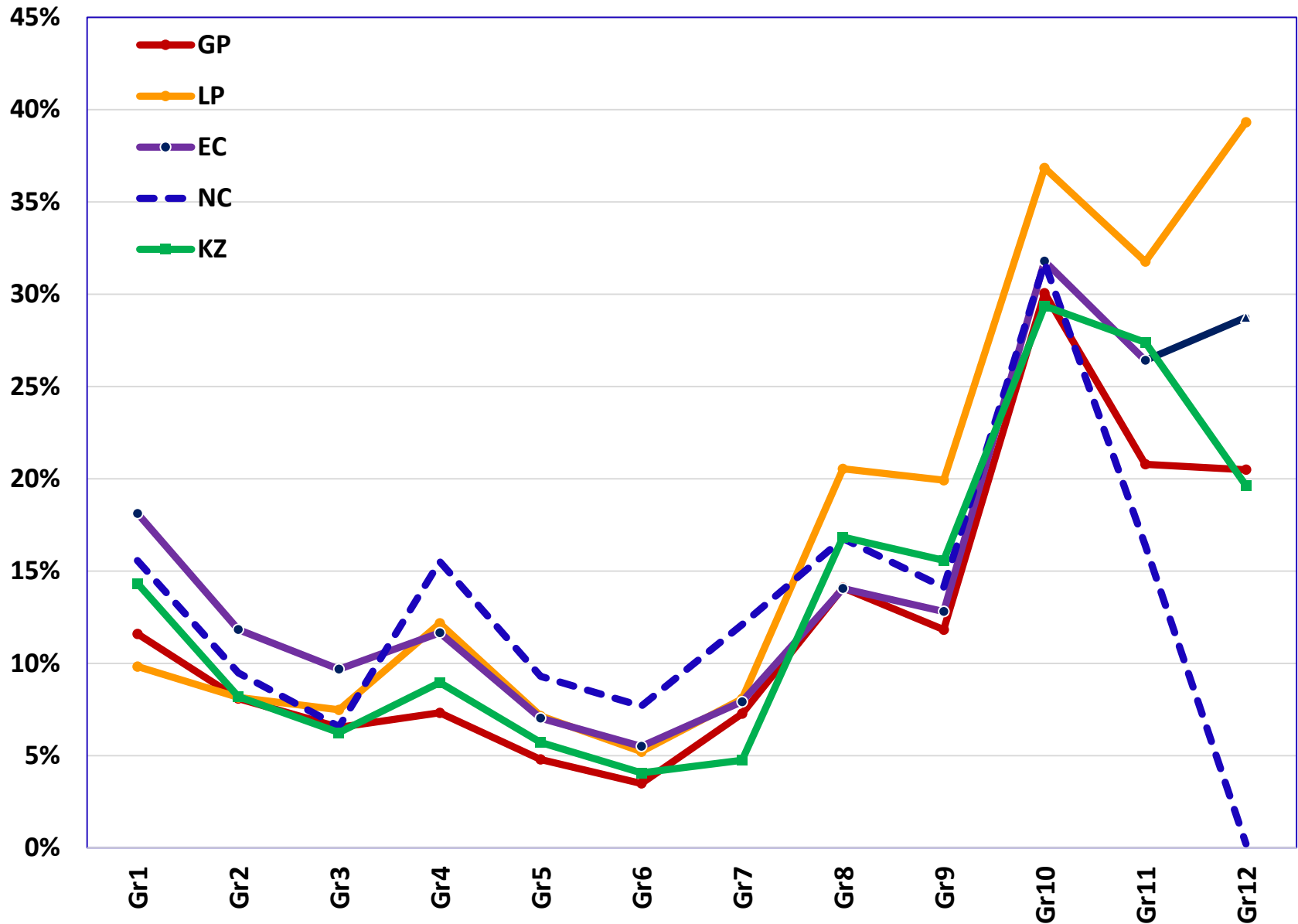Note the effect of the high stakes Form 4 exam
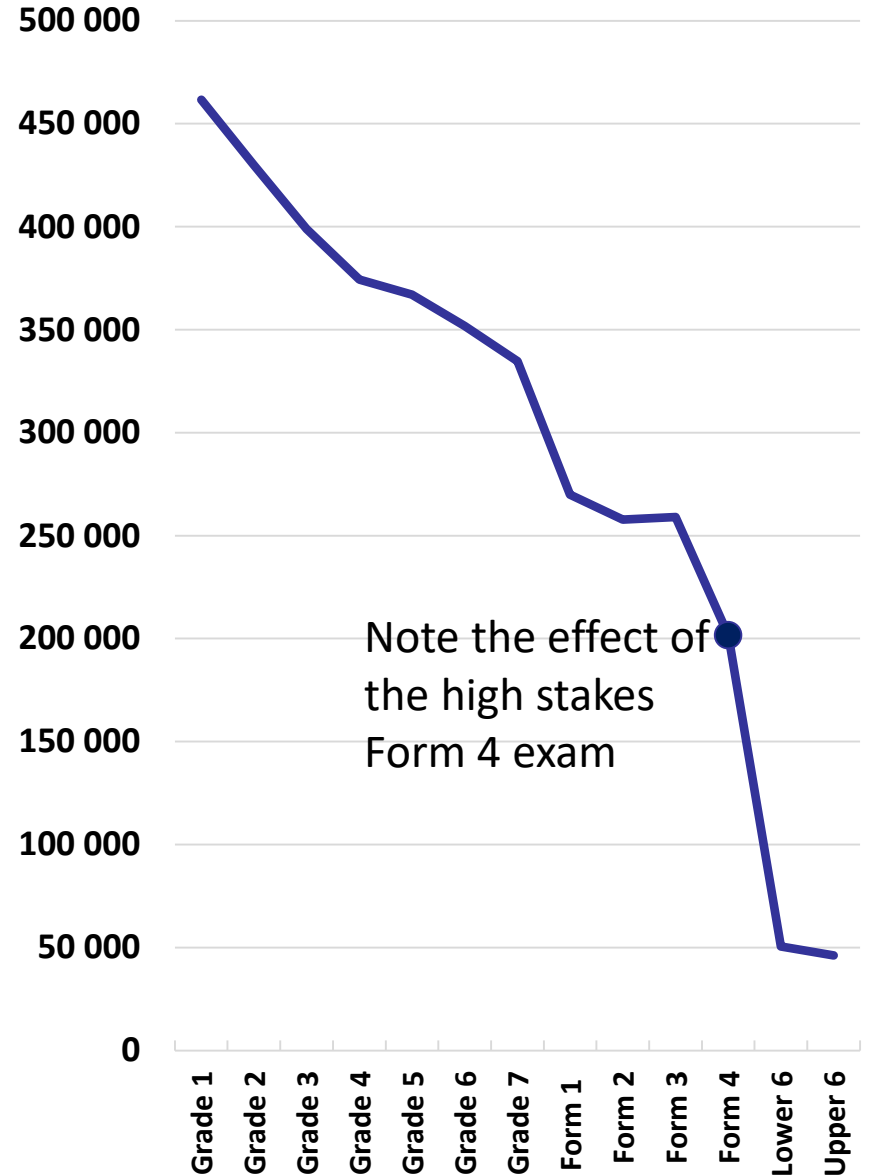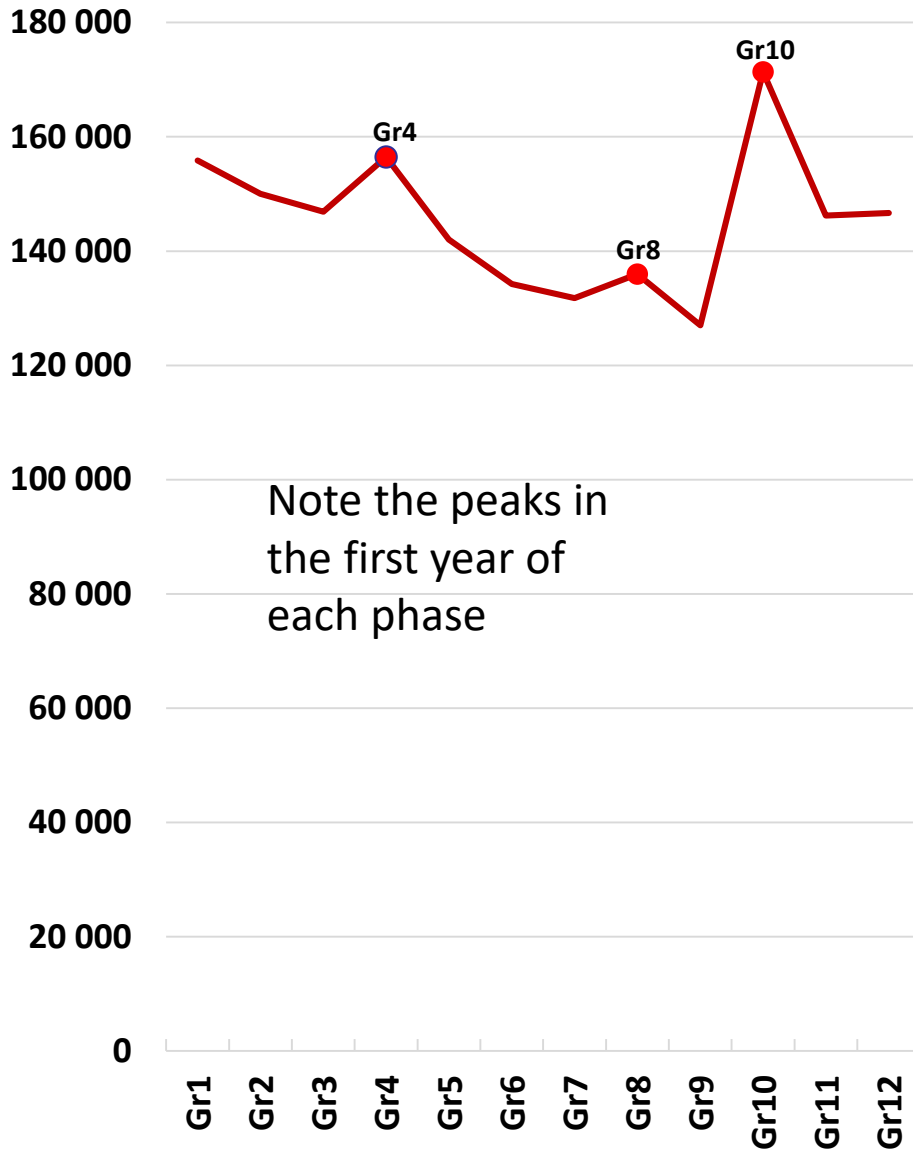
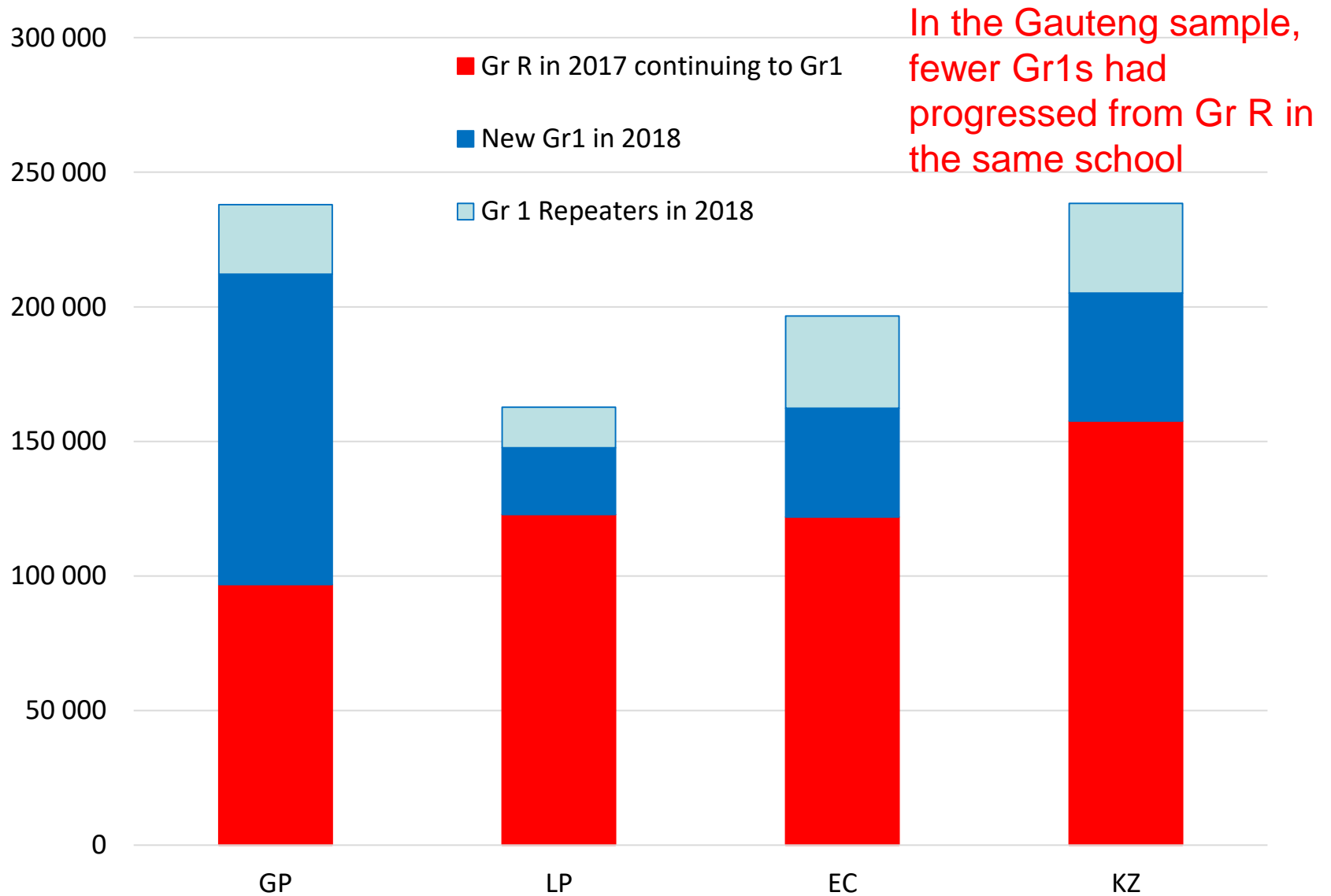# Total enrolment, showing over-age, for Eastern Cape sample, 2018

Over-age in Gauteng and Limpopo samples, 2015 vs 2018
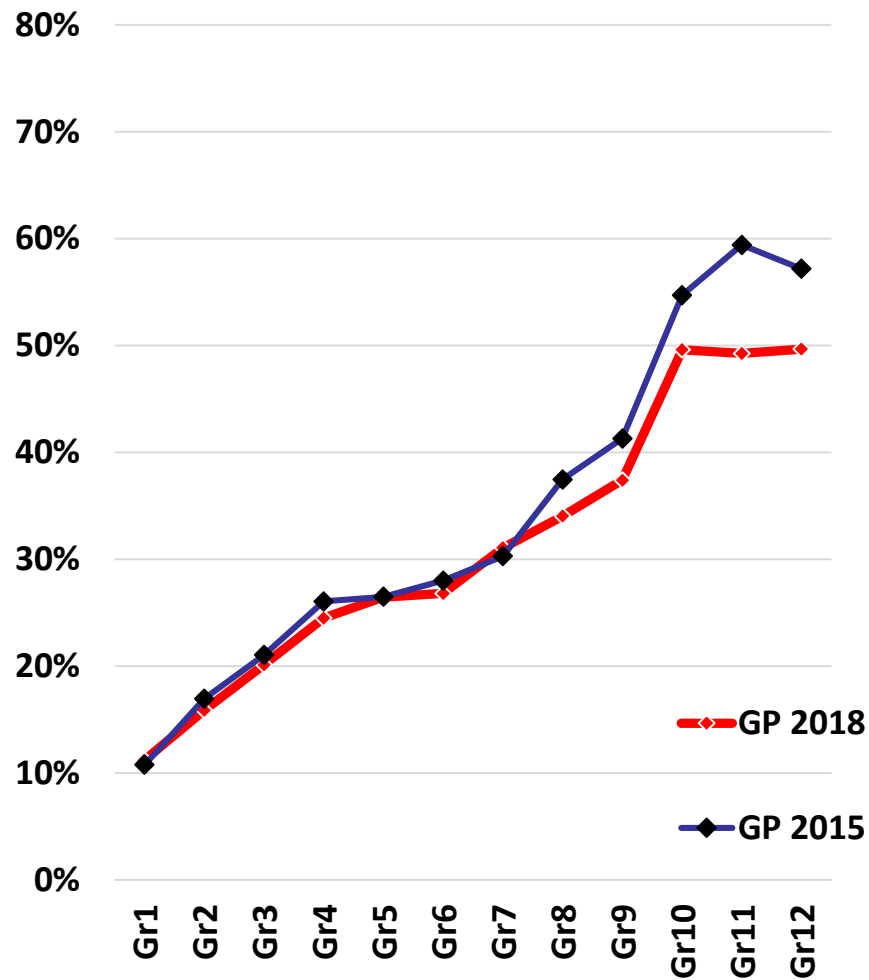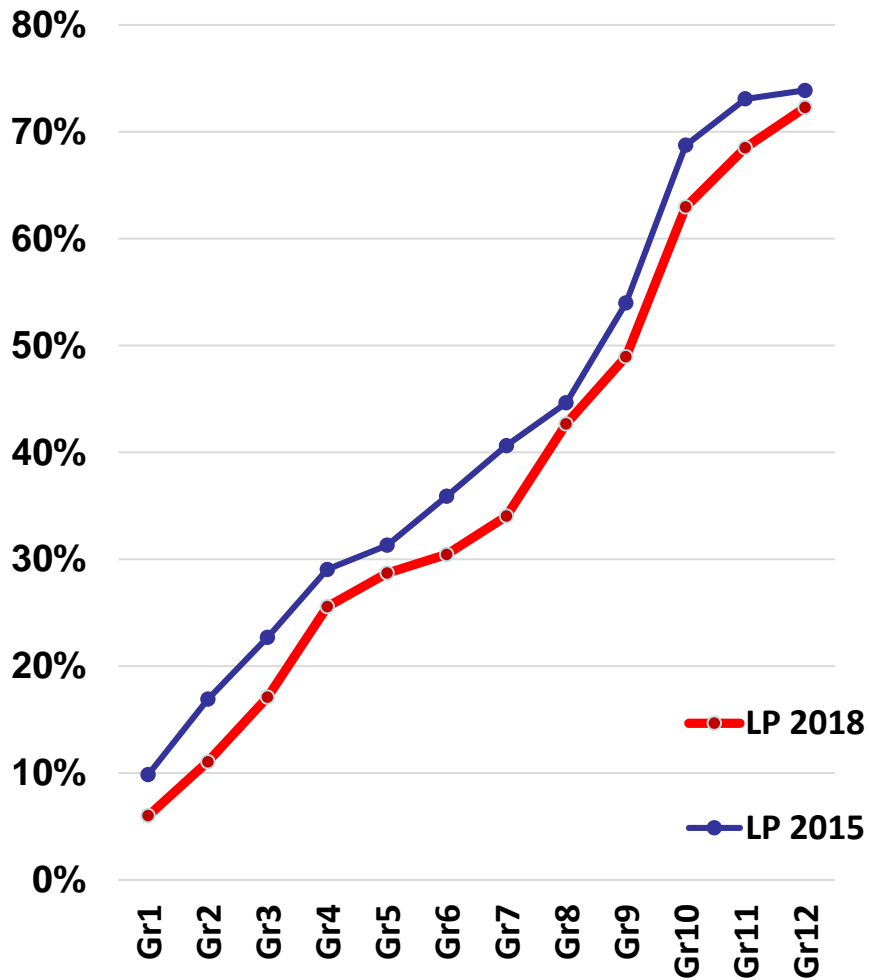
Repeaters by grade in five provincial samples, 2017

Composition of Gr1 enrolment in 4 provincial samples, 2018

Associations from regressions on learner flows in three provincial samples (GP, EC, LP) (controlling for home language and district)

	Repetition	Progression		Drop-out in 2017	Maths mark			Switch Gr10-12 to MLit
	Gr4-6	**Gr9-11**	**Gr10-12**	**Gr10**	**Gr 5-7**	**Gr9-11**	**Gr10-12**	**Gr10-12**
Initial Maths/ MLit mark	(not in regressions)	+	+	+ if repea-ting Gr10	++	+	+	-
Other marks	(not in regressions)	+ EFAL, LO, SocSc	+ EFAL	(not in regressions)	+ EFAL, NatSc, SocSc	+ EFAL, LO	+ EFAL, LO	- EFAL, LO
Female	-	- in GP, EC	- in EC, LP	+	+	--	-	+ GP, LP
White	- (Indians also)	+ GP	+ GP	- EC	+ in GP, EC	+ GP	no	no
Over-age	+	-	- in EC	++	-	-	-	+
SES	- Q4, Q5	Q5 + in EC, LP	Q5 + EC, LP	- Q5	no	+ (but Q5 in GP —)	+ Q5 in LP only	Inconsistent
Absence	++, esp Term3	- in EC, LP	-	++, esp Term3	- -	-, esp Term3	- -, esp Term3	no

Regressions on subject choice in Gr10

- Factors significantly 'affecting' likelihood of taking Maths in Gr10, controlling for being over-age (-), home language, district, Gr9 Maths performance (+), Gr9 FAL mark (+), Gr9 Science mark (+), % in class taking Maths (+)

|	Gauteng	Limpopo
Female	-7.8%	-5.3%
Quintile 5	-13.6%	-80%

- Factors significantly 'affecting' likelihood of taking Science in Gr10, controlling for being over-age (-), home language, district, Gr9 Science performance, % in class taking Science (+)

	Gauteng	Limpopo
Female	-10.0%	-7.5%
Quintile 5	-13.9%	

Factors significantly associated with Gr12 Maths & MathsLit marks in 2018

	Gr 12 Maths		Gr 12 MathsLit	
	Gauteng	Limpopo	Gauteng	Limpopo
Female	-6.8%	-8.2%	-3.7%	-3.7%
% taking Maths	-0.16	-0.08	+0.15	+0.07
1 year over-age	-6.1	-6.5%	-5.5	-5.0%
2 years over-age	-5.1	-8.4%	-8.1	-8.6%
3 years over-age	-12.6	-12.1%	-12.9	-11.8%
Quintile 4	-	-7.8%	-	+13.3%%
Quintile 5	-	-7.6%	-	+23.8%

Maths Gr12 mark by Gr9 Maths performance in 2015
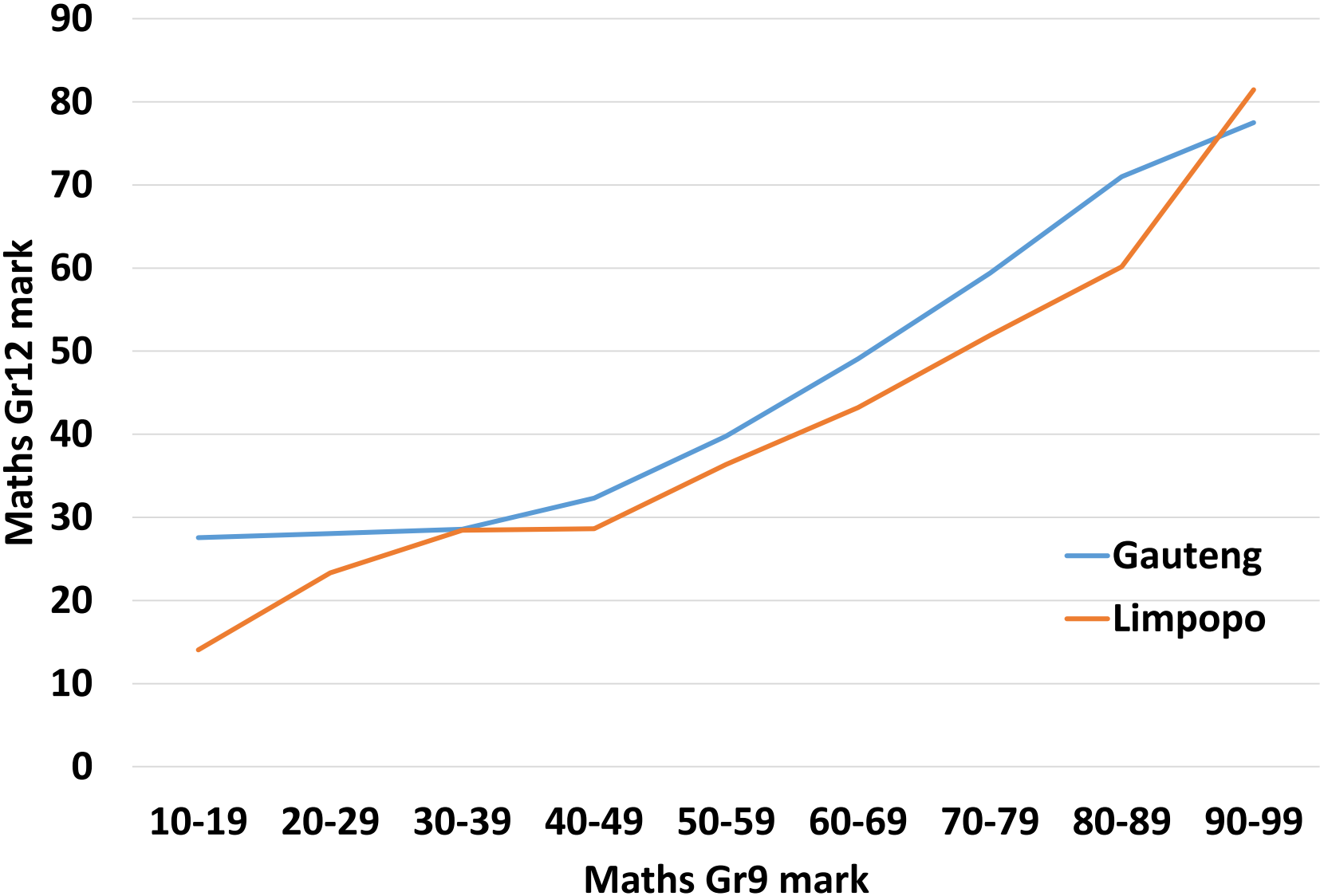
Gauteng sample Gr9 vs Gr12 Maths scores, 2015 Gr9 cohort

Gr 12 by Gr 9 maths report mark for learners taking mathematics in Gr 12
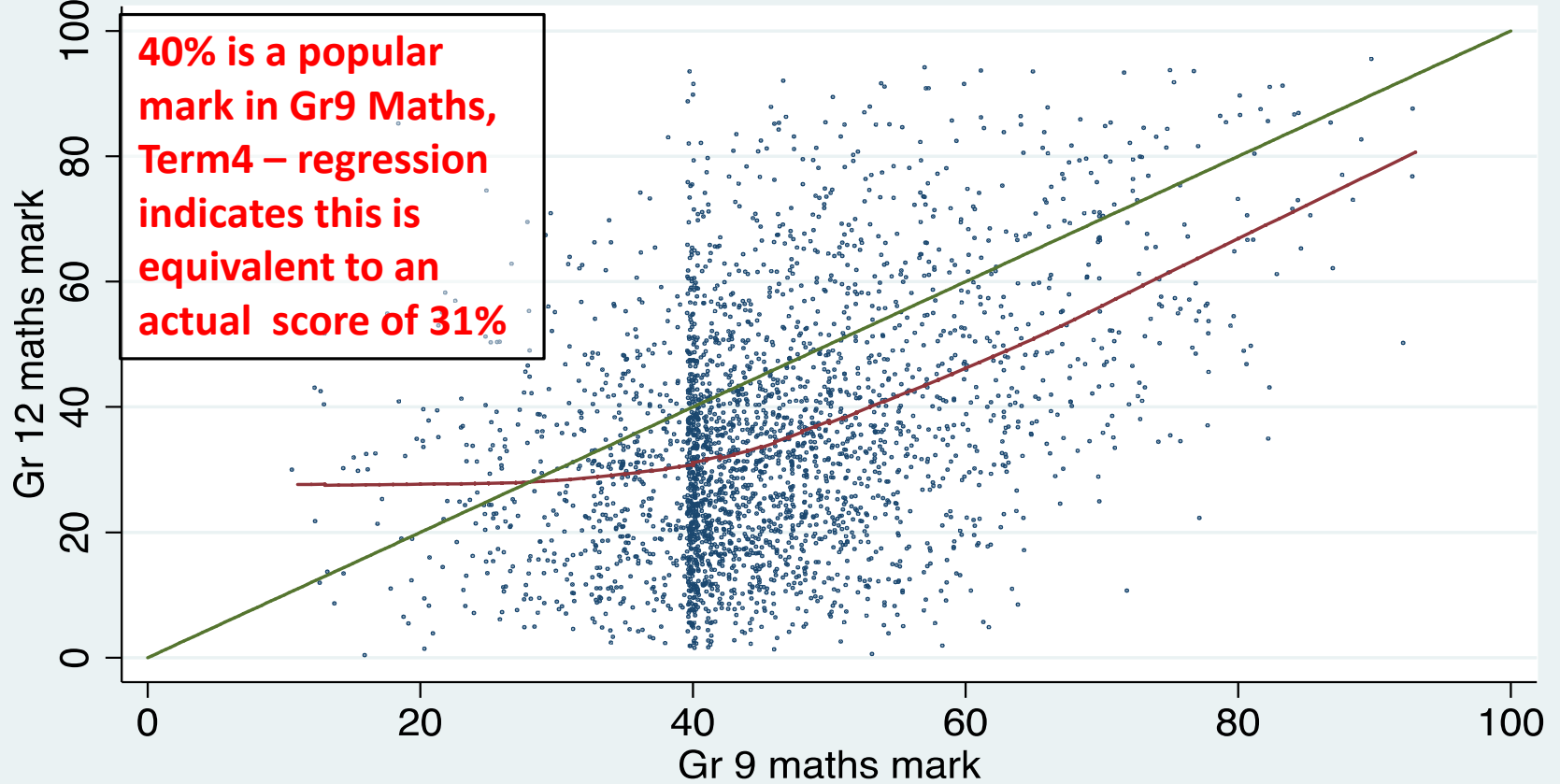
40% is a popular mark in Gr9 Maths, Term4 – regression indicates this is equivalent to an actual score of 31%

Gr 12 maths mark

Gr 9 maths mark

Mark   Lowess curve
45 degree line

bandwidth = .8

Gr9 English FAL mark by Gr 12 subject choice in Maths/MathsLit

Gauteng

Grade 9 English FAL mark by Grade 12 maths subject

Limpopo

Grade 9 English FAL mark by Grade 12 maths subject

Correlations in Gr9 marks, 2015

|	Gauteng			Limpopo		
	Maths	**English FAL**	**Natural science**	**Maths**	**English FAL**	**Natural science**
Maths	1			1		
English	0.31	1		0.39	1	
Natural science	0.46	0.48	1	0.57	0.43	1

School-based assessment (SBA) data in the SA-SAMS national warehouse

- In the absence of central assessments at primary level, how well can SBA data be used as managerial and accountability tool? What could feature in DDD dashboards?

- Martin Gustafsson found remarkable school-specific internal consistency across subjects in 2018 Eastern Cape data:

Table 1: Continuous assessment scores term 2 2018 in Grade 3

	Learners	Mean mark out of 100	10th p'tile	50th p'tile	90th p'tile
Mathematics	139,165	60	20	31	44
English FA	119,726	57	18	36	54
IsiXhosa HL	106,984	61	23	36	56
English HL	19,439	65	20	31	45
Afrikaans FA	14,512	61	20	36	54
Afrikaans HL	10,584	58	19	37	55

School-based assessment (SBA) data in the SA-SAMS national warehouse (contd.)
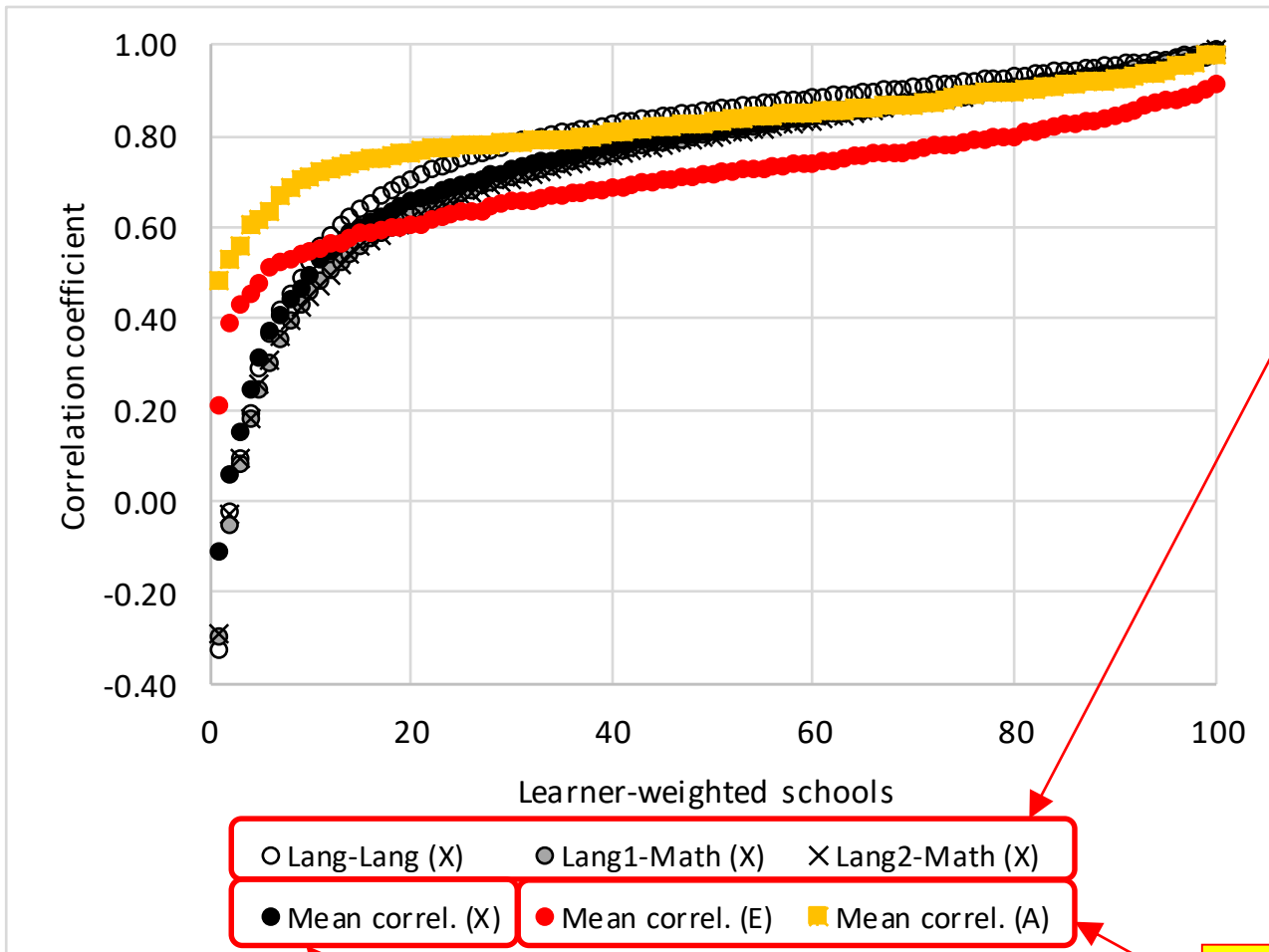
E.g. for EC schools with isiXhosa in Gr 3, three correlation coefficients calculated comparing (1) mathematics, (2) home language (isiXhosa), (3) first additional language (mostly English).

Conclusion: Often less consistency in learners' marks across subjects where isiXhosa is HL.

Mean of 3 correlation coefficients described above for schools with isioXhosa HL.

The mean correlation coefficient for schools with English and Afrikaans HL.

School-based assessment (SBA) data in the SA-SAMS national warehouse (contd.)

The Grade 3 indicator of consistency, colours showing 10th percentile per district.

Another indicator, gauging consistency across three points in time in 2018: Term 2 CASS, Term 4 CASS, final examination. All in Grade 6, five subjects considered.

☆Districts whose dummies returned statistically significant coefficients in a SCHOOL-LEVEL regression where each school's measure of consistency was the dependent variable, and independent variables were maths average score and district, quintile and home language dummies. Reference district was Graaff Reinet.

Across-subject Grade 3 correlations in schools: 10th percentile in district
- .40 to .49
- .50 to .59
- .60 to .72

Within-subject across-term Grade 6 correlations in schools: 10th percentile in district
- .45 to .51
- .52 to .60
- .61 to .68

The DDD data

- The dashboards have thus far been used to varying degrees in schools and districts

- Our focus was the <u>underlying data</u>

- The data holds immense promise for research, policy and management

  - but to fully unlock its potential would require attention to particular data issues, particularly better tracking of children across schools

- The broadening and lengthening sample is already increasing its representivity and usefulness

Thank you!
svdb@sun. ac.za /chrisvanwyk@sun.ac.za
Resep, University of Stellenbosch