Grade promotion, repetition and dropping out 2018 to 2021

Data Report

12 MAY

2023

Make Every Child A National Asset

basic education
Department:
Basic Education
REPUBLIC OF SOUTH AFRICA

Read to Lead
A Reading Nation is a Leading Nation

Second Chance Matric Support Programme

2030 NDP

STAY SAFE
PROTECT SOUTH AFRICA

Grade promotion, repetition and dropping out 2018 to 2021

12 MAY 2023

DG Foreword

Grade promotion, repetition and dropping out 2018 to 2021

For proper management of the schooling system, we need to know how many learners leave the schooling system before Grade 12, and how many repeat their grades. This is because we want to improve survival to Grade 12, and reduce grade repetition. In fact, we have seen improvements in both these areas in recent years. Today, just over 60% of our young population complete Grade 12 successfully, in the sense that they obtain the National Senior Certificate. This is similar to what is found in other middle income countries. Details in this regard are provided in our annual National Senior Certificate results report.

The current report provides promotion, repetition and dropping out rates across all grades during the last five years, but also how the challenges and rates of change differ across provinces. It makes important points on interpreting drop-out rates. This report is made possible by many years of hard work, at the national, provincial, district and school levels, devoted to improving the quality of our data. While a decade ago we essentially only had learner totals per grade and gender from schools, we now have a wealth of data at the level of each individual learner. Without learner-level data, it is not possible to study flows across grades properly. My hope is that this report will contribute to a more informed national debate.

HM MWELI

Director-General: Department of Basic Education

welcome to
DATA **REPORT**

Analysis of Flows Through Various Grades Sumamary Report

SUMMARY

ANALYSIS OF FLOWS THROUGH VARIOUS GRADES

The analysis makes use of four years of learner-level EMIS data, for the years 2018 to 2021, to produce year-on-year flow statistics relating to promotion, repetition and dropping out. For the latter phenomenon, the report frequently uses the term 'departing', given that certain forms of departing are not problematic, though the term 'dropping out' is generally considered to refer to a problem. Departures when learners emigrate, die, move to a college, or complete the Grade 12 National Senior Certificate, are unavoidable.

This report provides the most comprehensive set of flow statistics to date drawing from learner-level data. While the statistics have limitations, which are discussed in the report, on the whole they are sufficiently reliable to inform the policy debates, down to the level of province and grade (or age). Limitations relate to the fact that for around 17% of learners who stay from one year to the next, national ID numbers cannot link the learners across the two years. This non-linkage problem reduces substantially to just 3% if variables other than the ID number are used, but even this presents challenges for the calculation of the flow statistics. A key question is whether apparent drop-ins from one year to the next are actually the same learners as apparent drop-outs from the previous year. The present analysis uses ages across both years to guide the answering of this question, and to effect the necessary adjustments.

The statistics presented here confirm that the pandemic had a major effect on flows. When the 2019 to 2020 transition is compared to the 2020 to 2021 transition, it is clear that grade repetition declined by around two percentage points at the primary level, and to even greater extent at the secondary level. Dropping out clearly declined in grades 9 and 10. These changes are associated with large improvements in the percentage of learners who are promoted to the next grade, particularly at the secondary level. While 67% of Grade 11 learners moved to Grade 12 between 2019 and 2020, this figure rose to 80% between 2020 and 2021.

CONTENT

1 Introduction

This report uses a newly compiled five-year learner-level dataset[1], spanning 2017 to 2021, to produce statistics relating to promotion, repetition and dropping out. Due to issues with the linking of learners in the 2017 data to learners in 2018, issues which are explained below, the analysis is limited to the 2018 to 2021 data. This means that three year-on-year flows can be described: 2018 to the next year; 2019 to the next year; and 2020 to the next year.

The ability to report on year-on-year learner flows has improved considerably in recent years, as administrative data handled by the Department of Basic Education and the nine provincial departments have improved. These flows are important for planning purposes. Debates around whether grade repetition is too widespread in the schooling system, and should be further curtailed, need good information on the repetition phenomenon. Reliable information on dropping out must inform debates around whether the system is complying with legal imperatives relating to compulsory schooling. Moreover, dropping out *beyond* the upper limit for compulsory schooling, currently age 15, is a matter of considerable public interest.

Previously, attempts have been made to use administrative data to estimate flow statistics. An unpublished 2017 report of the Department of Basic Education (2017) provides an account of the methodologies available to estimate these statistics when data are non-ideal, as is often the case (and is the case in the current analysis). That report uses 2013 and 2014 learner-level data to calculate repeater and dropping out ratios by province and grade. As in the current analysis, age was used to link groups of learners across years where unique identifiers were unable to link individual learners.

A 2021 analysis by Van der Berg *et al* (2021) used 2017 to 2019 learner-level data in order to calculate national repeater ratios by grade. The report moreover reports on apparent drop-outs and drop-ins at the national level by grade but does not explicitly produce adjusted dropping out ratios that take into account the fact that apparent drop-outs and apparent drop-ins are often the same people. That analysis did not use learner age to assist in the interpretation of drop-ins (as is done here).

2 The data

Table 1 below illustrates what percentage of learners in each year and province could be found in the next year's data, using a hybrid three-step linking procedure. In this procedure, described in the earlier metadata report, linking first occurs using the 13-digit national ID number, where these identifiers are unique in both years in question. These learners are put to one side as definitively matched. Thereafter, remaining learners are linked if they are unique in each year with respect to four variables, none of which should contain missing values: date of birth, first name, surname, and gender. These learners are put to one side. Finally, the system accession number of each learner is used. As can be seen in Table 1, linking between 2017 and 2018 is relatively low, even if just learners in grades 1 to 6 in the initial year are considered. For this group, one can expect very few learners to actually leave the system between two years. This explains why the analysis that follows uses only the years 2018 to 2021. Critically, this allows for at least one transition which was clearly not influenced by the COVID-19 pandemic, namely that between 2018 and 2019. The transition from 2019 to 2020 is also unlikely to have been influenced by the pandemic, but this relies on the assumption that data compiled at the national level reflected only the situation before March 2020. This assumption should be true, based on descriptions of how the data are collected.

For Table 1, learners are only considered linked if their grades in both years are in line with policy. All learners promoted to the next grade and remaining in the same grade are accepted. However, learners moving up *two* grades or moving down *one* grade are also accepted in the linking process, based on the earlier conclusion, in the metadata report, that it appears that such movements are likely to be real. Such somewhat irregular movements represent around 0.6% of linked learners. However, all other grade movements, for instance up three or more grades, are rejected. Where grade is missing in either or both years, the link is also rejected.

1 Described in 'An anonymised five-year learner-level dataset for 2017 to 2021' (dated 15 December 2022).

Table 1: Linking to next year by province

	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
All										
2017	85	82	65	74	83	73	89	80	78	76
2018	90	91	86	88	91	88	91	91	88	89
2019	90	92	88	91	90	90	92	91	92	90
2020	89	92	88	90	92	91	91	91	91	90
Only grades 1 to 6										
2017	92	89	71	81	90	79	96	86	84	83
2018	97	98	93	96	98	95	97	96	95	96
2019	96	98	95	97	97	97	98	97	98	97
2020	95	98	95	97	98	98	97	97	97	97

Table 2 provides details on what percentage of linked learners relied on each of the three methods. Each linkage to the next year, for 2018, 2019 and 2020, was considered. Thus a learner present and linked across all three pairs of years would be taken into account three times. Gauteng and Western Cape had relatively low success rates with respect to the ID. In the case of Western Cape, this is due to the fact that no linking using the ID number between 2018 and 2019 was possible, for reasons given in the metadata report. 'Combination' in the table refers to the combination of the four variables referred to previously. The fact that well over 90% of linking relied on the ID number across seven provinces bodes well for the reliability of the linking.

Table 2: Types of linking employed 2018 to 2021

	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
ID no	96	95	83	94	97	93	98	95	64	90
Combination	3	5	16	5	3	6	2	5	35	9
Accession no	0	0	1	0	0	1	0	0	1	0
Total	100	100	100	100	100	100	100	100	100	100

3 Age-specific flow ratios at the national level

As described in the metadata report, year of birth was present for virtually all learners for the years 2018 to 2021. This facilitates the calculation of reliable flow statistics by age, before these statistics are calculated by grade. Flow statistics by age are important in themselves. For instance, for monitoring purposes it is important to know that dropping out is very close to zero for children aged 7 to 13 at the end of the year. In all the analysis that follows, age is the learner's age at the end of the calendar year. But the analysis by age presented here is also important for understanding the subsequent analysis by grade, as the latter analysis will use age to provide more reliable assumptions relating to the previous grade of drop-*ins.*

Figure 1 provides a simple analysis of the number of learners present in the dataset, if in the age range of 5 to 25. Clearly, learners exist up to at least age 24. For this reason the analysis that follows generally goes up to age 25.

Figure 1: National age-specific enrolments 2018-2021

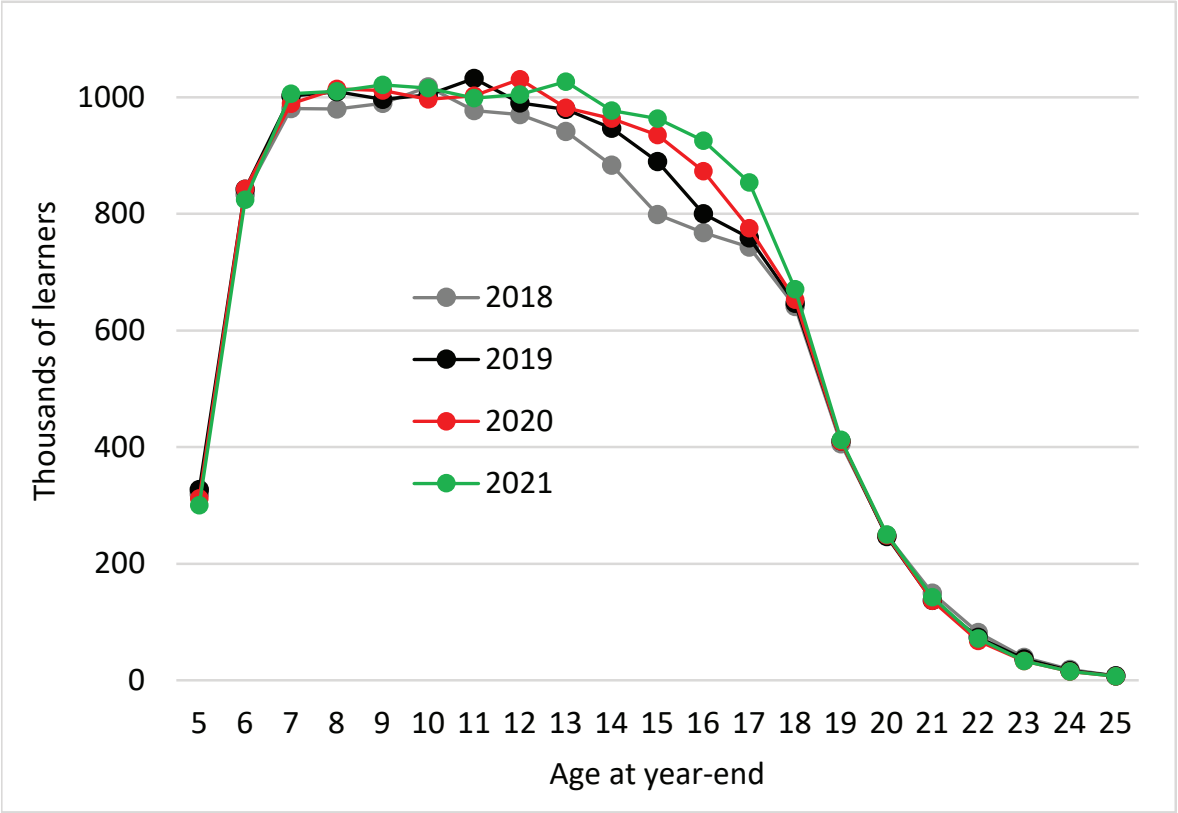


Table 3 below explains how drop-ins influence the three key flow statistics. The table reflects actual values for learners aged 10 to 13 in 2018. Asterisked promotion and repeating values reflect linking to the the next grade or same grade respectively. For this analysis only learners with valid grades (R to 12) in the initial year were considered. As pointed out in the metadata report, around 0.3% of learners do not have a valid grade. In the case of some learners this is correct: some learners in special schools are not studying at a particular grade level. The grade filter means that learners in the base year for the age analysis are the same as those used for grade analysis in section 5.

Departing learners with an asterisk in Table 3 are all learners who could not be found in the following year. Drop-ins are learners found in the second year, 2019 in this instance, but not in the first year. Here one age up in the second year is used. Thus the 48,037 apparent drop-ins in the age 10 row would be age 11 in 2019. Drop-ins could be learners with an invalid grade in the first year. Given that around 99% of children in the age range 7 to 13 are attending school according to household data (see section 6), actual dropping out in these ages would be virtually zero. It is thus assumed that what appears as departing learners would nearly all be found among the drop-ins, if learner identifiers were more reliable. The column headed ‘Departing’, with no asterisk, reflects this assumption. Only for age 13 are there actual learners who depart, and the 1,001 who depart are the asterisked departures minus the drop-ins.

The term ‘departing’ is used here for what many would refer to as ‘dropping out’. Arguably, ‘departing’ is a better term as unlike ‘dropping out’, it does not carry a strongly negative connotation. Some departing is un-desirable, especially in the case of young children of compulsory school-going who leave school. However, several forms of departing do not represent an educational failure: learners can leave the system due to death, emigration, transfer to a college and of course because they have successfully completed Grade 12.

Table 3: How drop-ins are dealt with

Age	Total	Promoted*	Repeating*	Departing*	Drop-ins	Departing	Promoted	Repeating
10	1,017,305	908,736	69,063	39,506	48,037	0	945,452	71,853
11	976,163	870,422	69,137	36,604	54,054	0	904,333	71,830
12	969,147	874,423	47,579	47,145	66,746	0	919,135	50,012
13	940,544	824,424	52,139	63,981	62,980	1,001	883,658	55,885

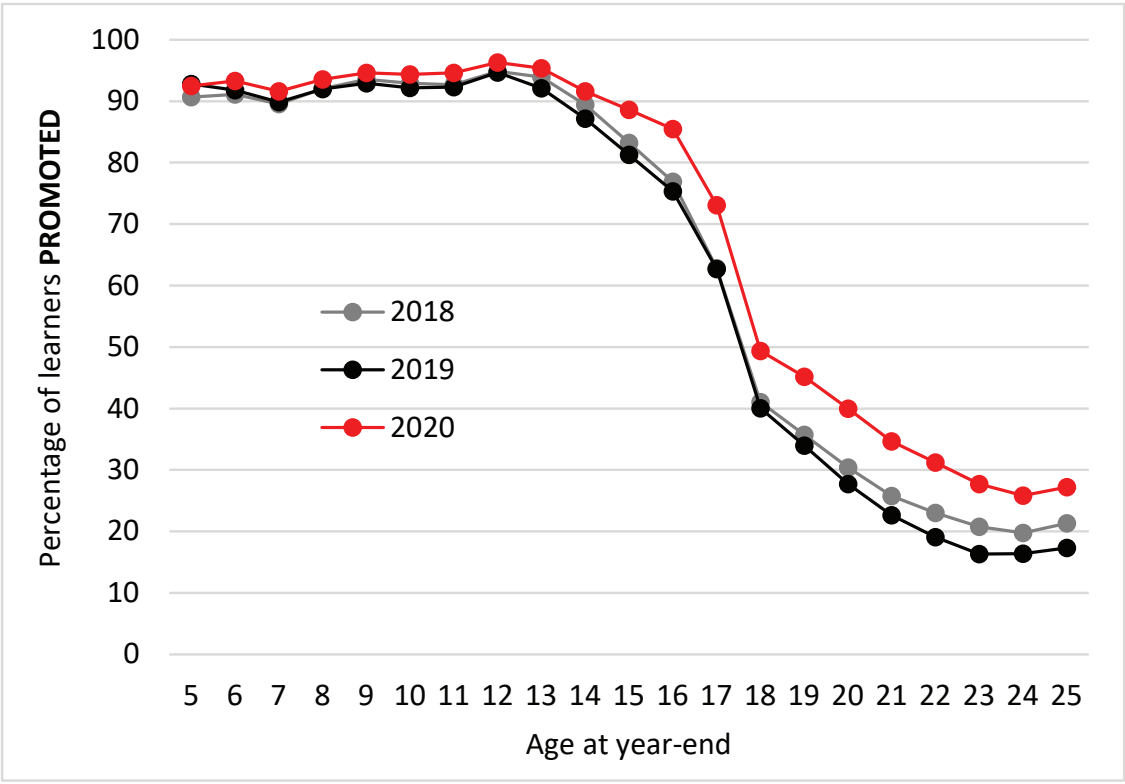
Finally, promoted learners (second-last column) are calculated as follows:

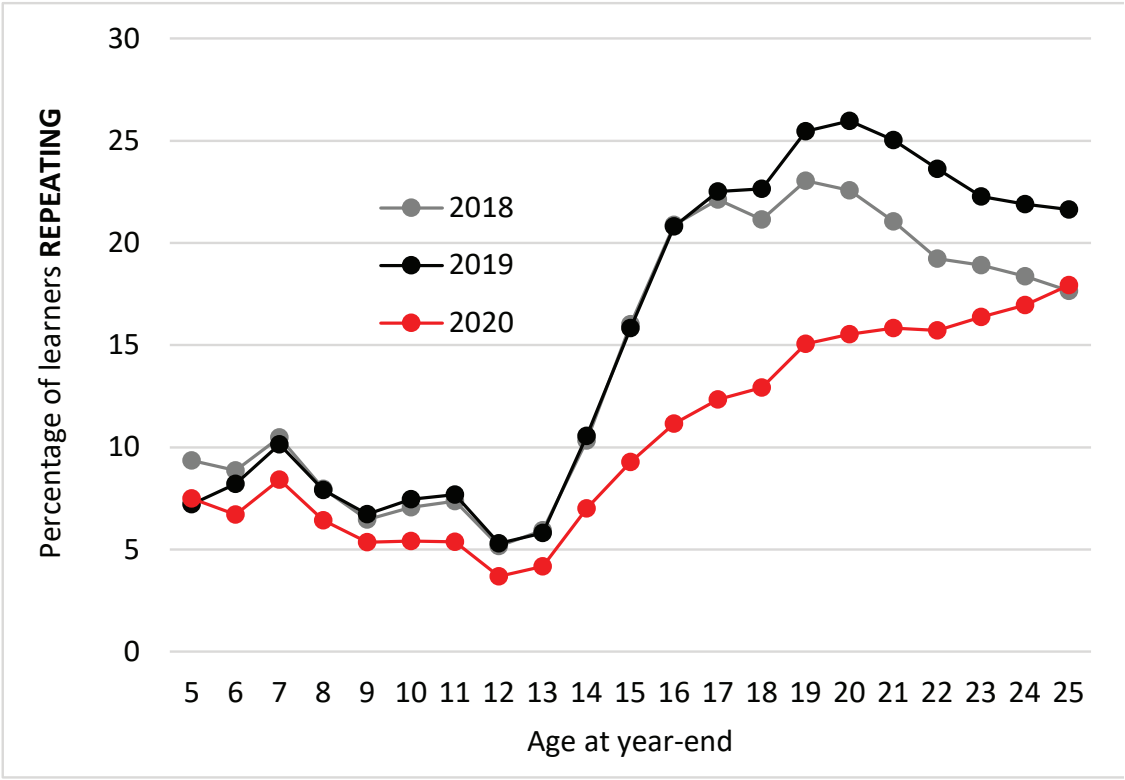
The initially estimated number of promoted learners (P^*) is expressed as a fraction of original promoted and repeating (R^*) learners, and this fraction is multiplied by total learners (T) who are not departing (D), using the number of departures after the drop-in adjustment. In other words, promoted learners are inflated using the assumption that unlinked learners are as likely to be promoted, as opposed to repeat, as unlinked learners. A similar approach is employed to obtain an inflated number of repeaters. The last three columns in Table 3 add up to the total number of learners, meaning the percentage of 2018 learners promoted, repeating and departing would add up to 100%.

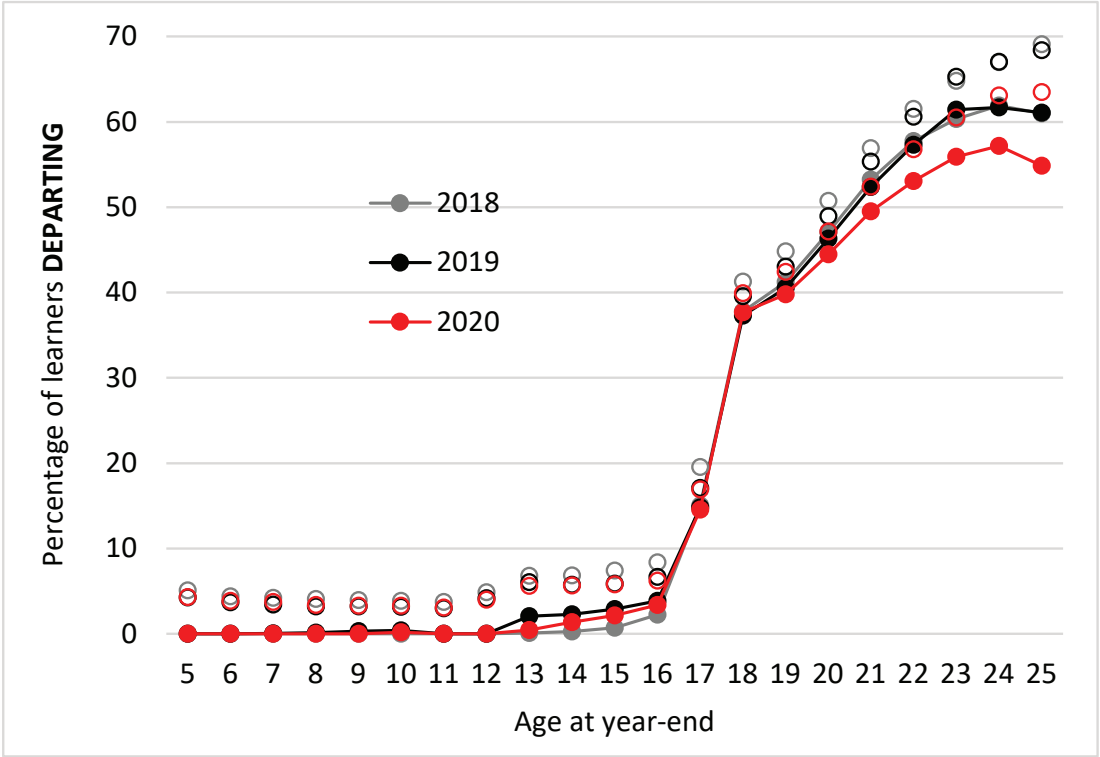
Table 3 suggests that net dropping *in* is substantial, at least for ages 10 to 12. For instance, it appears that 19,601 12-year-old learners actually joined between 2018 and 2019. This may reflect international in-migration. Closer analysis of the data, in particular of names, could shed more light on this.

Figure 2 below illustrates all the age-specific flow rates, within the age 5 to 25 range, at the national level. The last graph in Figure 2 illustrates the number of departing learners seen before and after the drop-in adjustment. Hollow markers indicate the statistics before the adjustment. At the end of 2019 fewer learners were promoted and more repeated than one year earlier. However, 2020 saw a large increase in promotion and a decline in repetition. Nationally, the change in the departure ratios was minor, despite fears that the pandemic would lead to a higher number of learners disengaging permanently from schooling. Most of the change occurred for exceptionally old learners, where the departure ratio *declined* considerably, meaning learners stayed on in the system.

Figure 2: National age-specific flow statistics 2018-2021







The following three graphs reflect the widely known problem of the under-performance of boys in the schooling system. Boys are more likely to repeat, and this pattern begins at an early age – see Figure 3 and Figure 4. Figure 5 indicates females begin departing earlier than males. This is mostly because females attain specific levels of schooling, and the Grade 12 National Senior Certificate (NSC), earlier than males. The difference between the percentage of females and males achieving the NSC has been over five percentage points *in favour of* females in recent years (Department of Basic Education, 2020: 99).

Figure 3: National age-specific repeater ratios by gender 2018

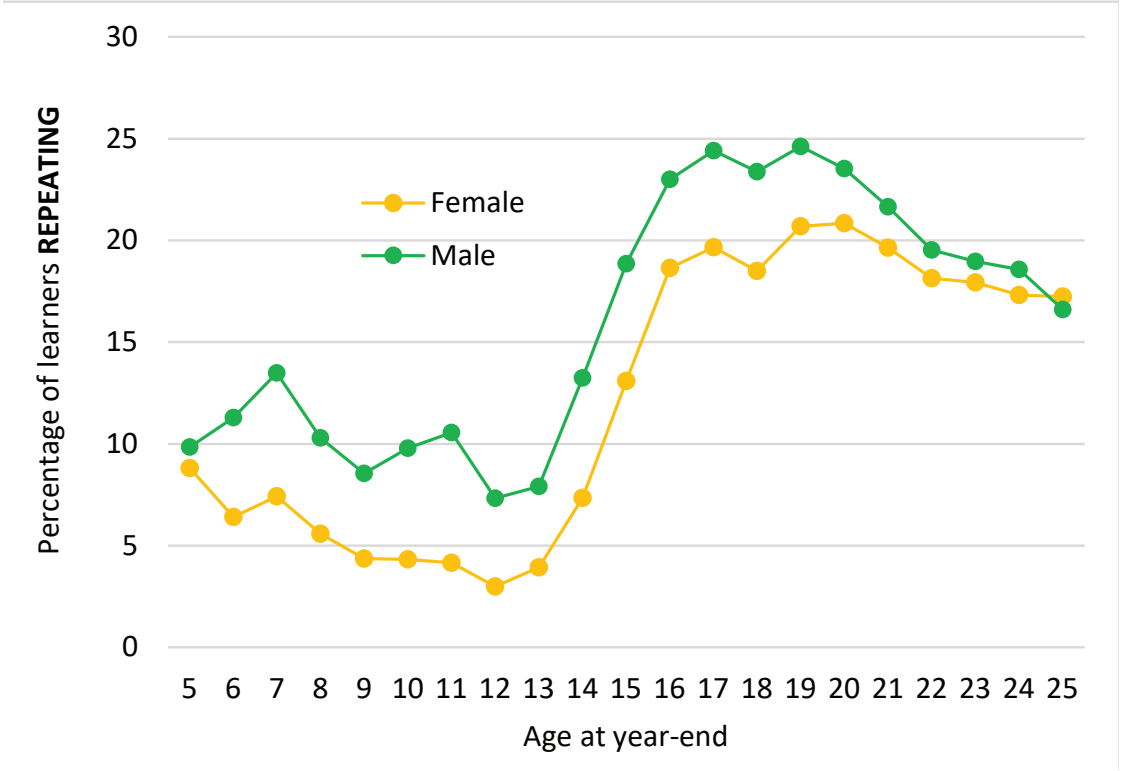


Figure 4: National age-specific repeater ratios by gender 2020

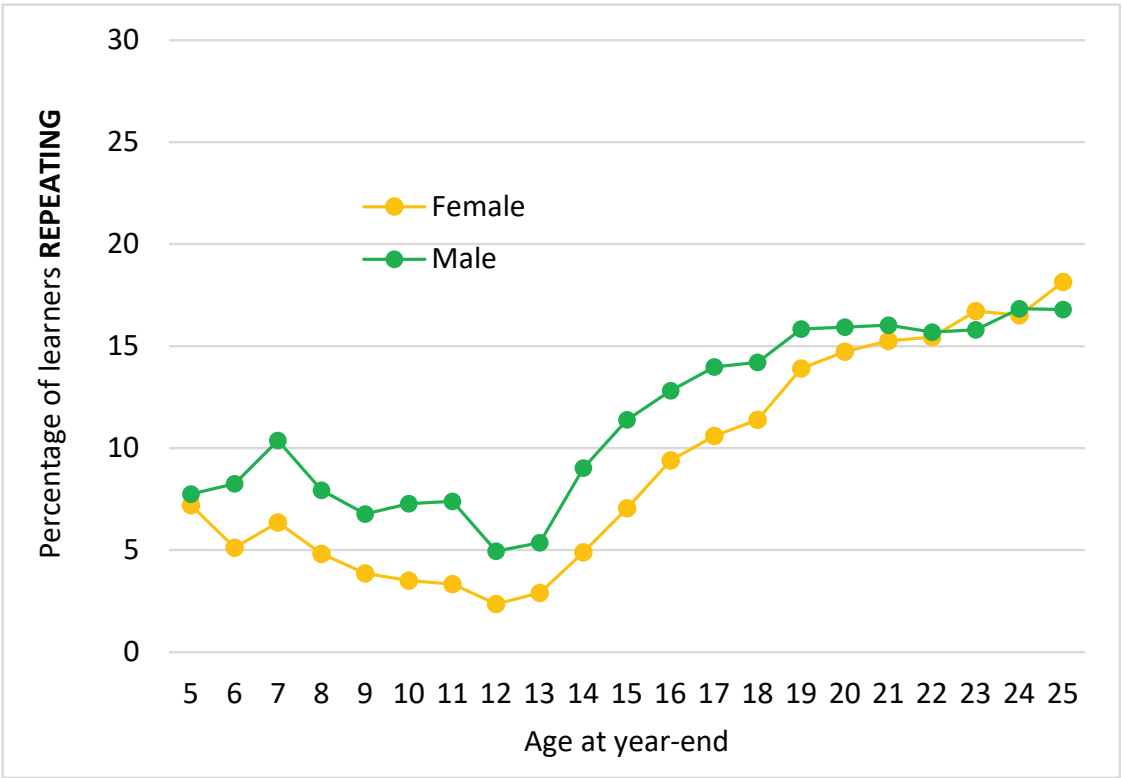
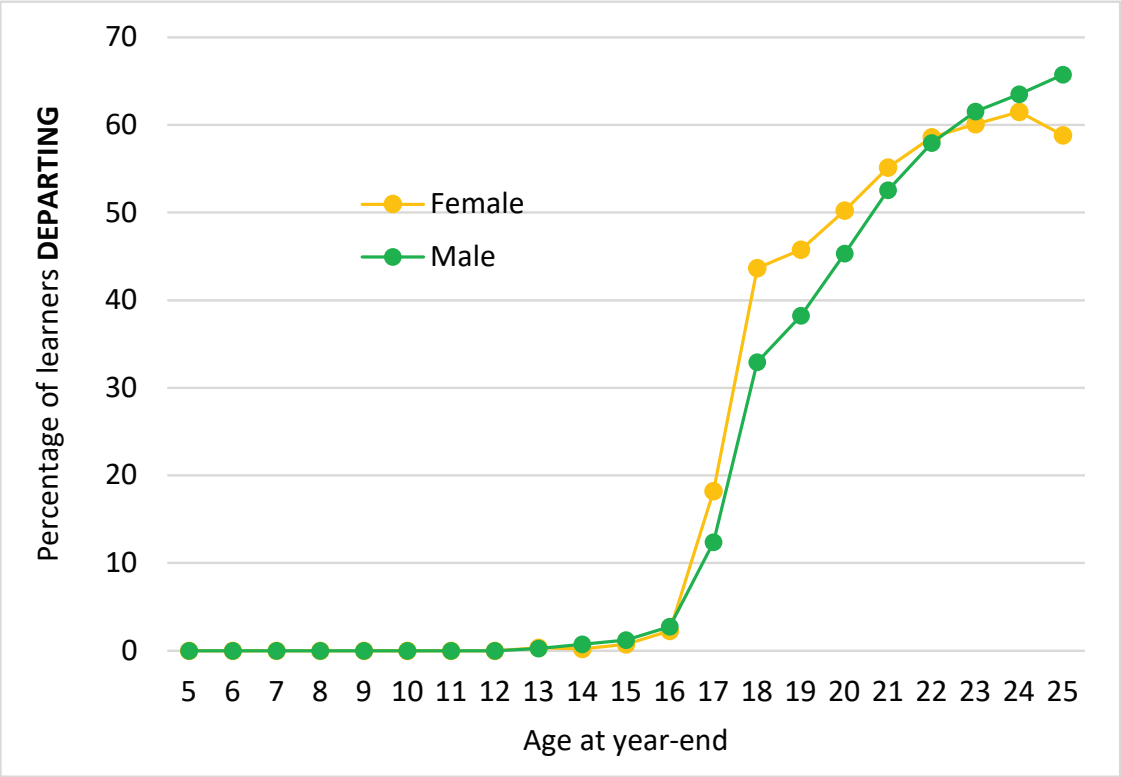


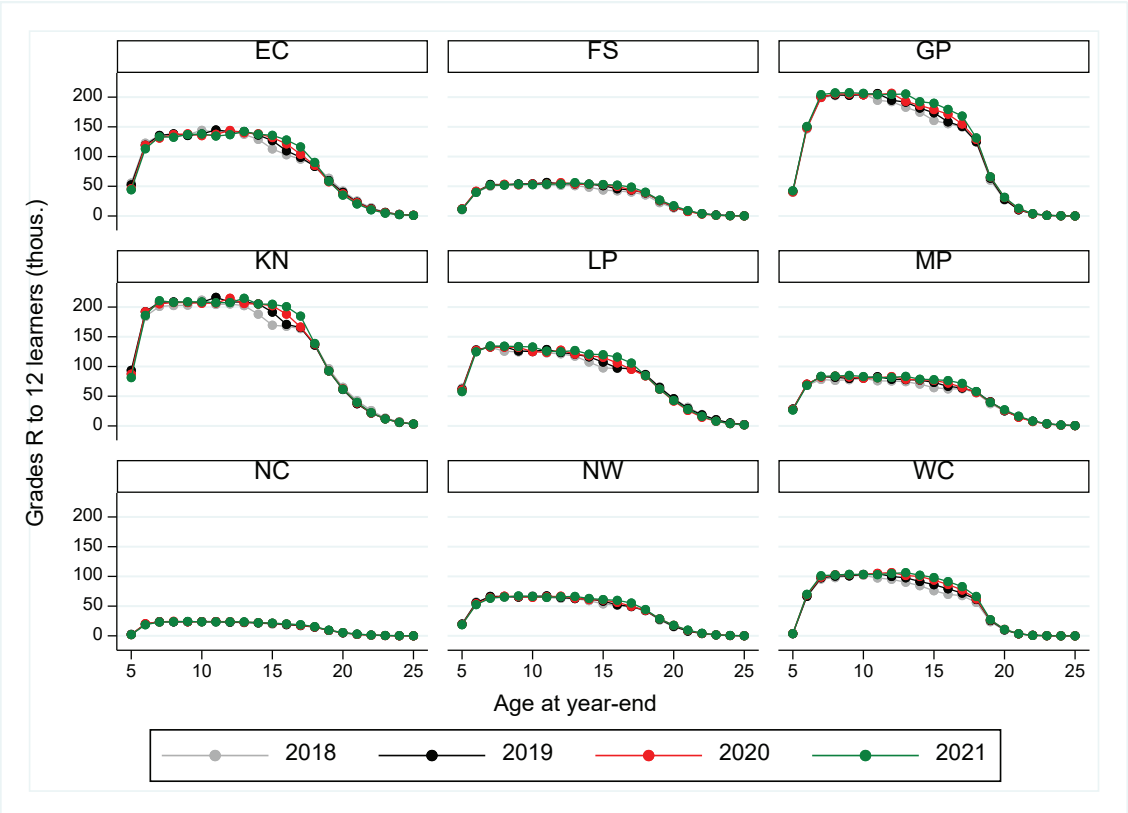
Figure 5: National age-specific departure ratios by gender 2018



4 Age-specific flow ratios and migration patterns at the provincial level

Figure 6 illustrates enrolment numbers across the four years in the nine provinces. Especially for KwaZulu-Natal, the higher level of enrolment in 2021 for ages 15 to 17, relative to earlier years, is visible.

Figure 6: Provincial age-specific enrolments 2018-2021



The approach taken for age-specific flow statistics for provinces was to ignore the province of the second year. Thus, a learner in 2018 in Eastern Cape found in the next grade in KwaZulu-Natal would be counted the same as a promoted learner who remained in Eastern Cape. Departures in this analysis thus means not found anywhere in the country in the next year. With fully reliable linking across years, it would be optimal, for instance, to separate promoted in another province from promoted in the same province. However, given that there is an adjustment process to deal with linkage gaps, this was avoided to reduce complexity. Instead, inter-provincial migration is first analysed to see what impact it may have on the flow statistics, and then flow statistics are calculated without any accounting for inter-provincial migration.

Table 4 presents the percentage of learners present in one year moving to different provinces in the next year, using only learners who could be linked across the two years. The sending province is the row heading and the receiving province the column heading. The percentages in each *row* add up to 100%. To illustrate, 0.9% of learners in Eastern Cape in 2018 were found in Western Cape in 2019. Perhaps surprisingly, it is not Eastern Cape but Gauteng which displays the highest number of out-migrating learners between 2018 and 2019: this figure is 45,050 learners (see the last column). Yet the number of *in*-migrating learners, at 54,699 (see row 'In') exceeds this, meaning that Gauteng displays a high *net* in-migration of 9,649 learners.

Table 4: Inter-provincial movements

2018 to 2019										
	EC	**FS**	**GP**	**KN**	**LP**	**MP**	**NC**	**NW**	**WC**	**Out**
EC	97.8	0.1	0.5	0.5	0.0	0.0	0.0	0.1	0.9	36,399
FS	0.1	98.6	0.7	0.1	0.0	0.1	0.1	0.2	0.1	8,740
GP	0.3	0.2	97.9	0.4	0.5	0.3	0.0	0.3	0.1	45,050
KN	0.2	0.0	0.4	99.1	0.0	0.1	0.0	0.0	0.0	21,437
LP	0.0	0.0	0.9	0.0	98.6	0.3	0.0	0.1	0.0	21,425
MP	0.1	0.1	0.7	0.3	0.3	98.5	0.0	0.1	0.0	14,055
NC	0.1	0.2	0.3	0.0	0.0	0.0	98.2	0.6	0.3	4,775
NW	0.1	0.2	1.1	0.1	0.2	0.1	0.2	98.0	0.0	15,323
WC	0.7	0.0	0.1	0.0	0.0	0.0	0.1	0.0	99.0	10,450
In	21,669	9,302	54,699	21,497	16,262	15,890	4,577	13,343	20,415	177,654
Net in	-14,730	562	9,649	60	-5,163	1,835	-198	-1,980	9,965	
% in	-0.9	0.1	0.5	0.0	-0.3	0.2	-0.1	-0.3	1.0	
2019 to 2020										
EC	98.1	0.1	0.4	0.4	0.0	0.0	0.0	0.1	0.8	31,378
FS	0.1	98.6	0.6	0.1	0.1	0.1	0.1	0.2	0.1	8,939
GP	0.2	0.1	98.1	0.3	0.5	0.3	0.0	0.3	0.1	41,710
KN	0.2	0.0	0.4	99.2	0.0	0.1	0.0	0.0	0.0	19,984
LP	0.0	0.0	1.0	0.0	98.5	0.3	0.0	0.1	0.0	24,494
MP	0.1	0.0	0.7	0.2	0.3	98.6	0.0	0.0	0.0	13,537
NC	0.1	0.2	0.2	0.0	0.0	0.0	98.5	0.6	0.3	4,201
NW	0.1	0.1	1.0	0.1	0.2	0.1	0.2	98.2	0.0	14,441
WC	0.2	0.0	0.1	0.0	0.0	0.0	0.1	0.0	99.5	5,988
In	14,208	7,563	53,890	18,506	18,028	17,352	4,866	12,270	17,989	164,672
Net in	-17,170	-1,376	12,180	-1,478	-6,466	3,815	665	-2,171	12,001	
% in	-1.0	-0.2	0.6	-0.1	-0.4	0.4	0.2	-0.3	1.1	

2020 to 2021										
	EC	FS	GP	KN	LP	MP	NC	NW	WC	Out
EC	98.1	0.1	0.5	0.4	0.0	0.0	0.0	0.1	0.8	31,460
FS	0.2	98.7	0.6	0.1	0.0	0.1	0.1	0.2	0.1	8,506
GP	0.3	0.2	97.5	0.5	0.7	0.3	0.0	0.3	0.1	56,072
KN	0.2	0.0	0.4	99.2	0.0	0.1	0.0	0.0	0.0	21,547
LP	0.0	0.0	0.9	0.0	98.7	0.2	0.0	0.1	0.0	20,840
MP	0.1	0.0	0.6	0.2	0.3	98.6	0.0	0.1	0.0	13,734
NC	0.1	0.2	0.2	0.1	0.0	0.0	98.6	0.5	0.2	3,806
NW	0.1	0.1	0.8	0.1	0.2	0.1	0.2	98.4	0.0	12,265
WC	0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	99.3	7,642
In	20,741	9,237	50,659	22,582	21,594	15,534	3,719	13,886	17,920	175,872
Net in	-10,719	731	-5,413	1,035	754	1,800	-87	1,621	10,278	
% in	-0.6	0.1	-0.2	0.0	0.0	0.2	0.0	0.2	0.9	

Note: Row headings indicate which province learners are from, column headings indicate which province learners go to. Province-specific percentages in rows add up to 100%.

Across all three periods, Western Cape displays the highest net in-migration expressed as a percentage of all learners in the base year: this figure is around 1.0%.

What is remarkable is that between 2020 and 2021 Gauteng switched from being a major receiver of migrating learners to a major sender of such learners. This would be in line with anecdotal evidence that in response to the lockdowns of the pandemic, but also economic hardship, households in Gauteng sent learners to neighbouring provinces to stay with family. The evidence here suggests that a fair proportion of these learners ended up enrolled in schools in the receiving provinces.

The next two graphs illustrate the migration patterns by age for two of the three periods. Migrating learners as a proportion of all learners is roughly similar across age. Migration is thus not something that affects younger or older learners to a much greater degree. Specifically, up to age 15 around 0.15% of learners nationally migrate across provinces in a year, with the figure dropping to around 0.10% for learners aged 16 to 20. Above age 20 the percentage rises, but the absolute number of learners involved is small.

Figure 7: Provincial migration gains 2018 to 2019

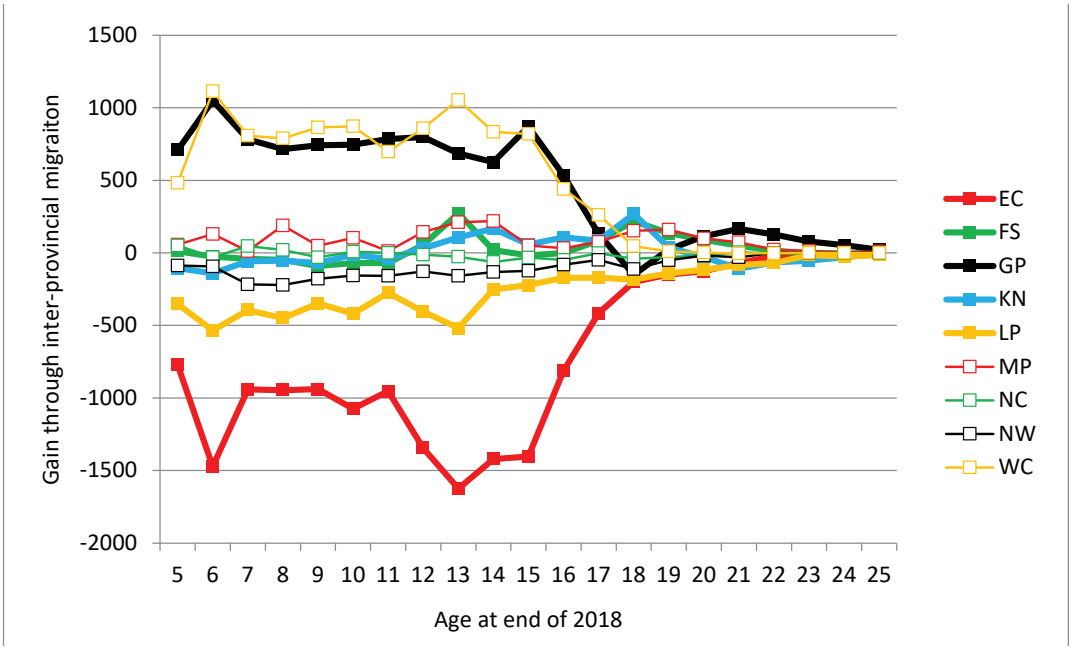
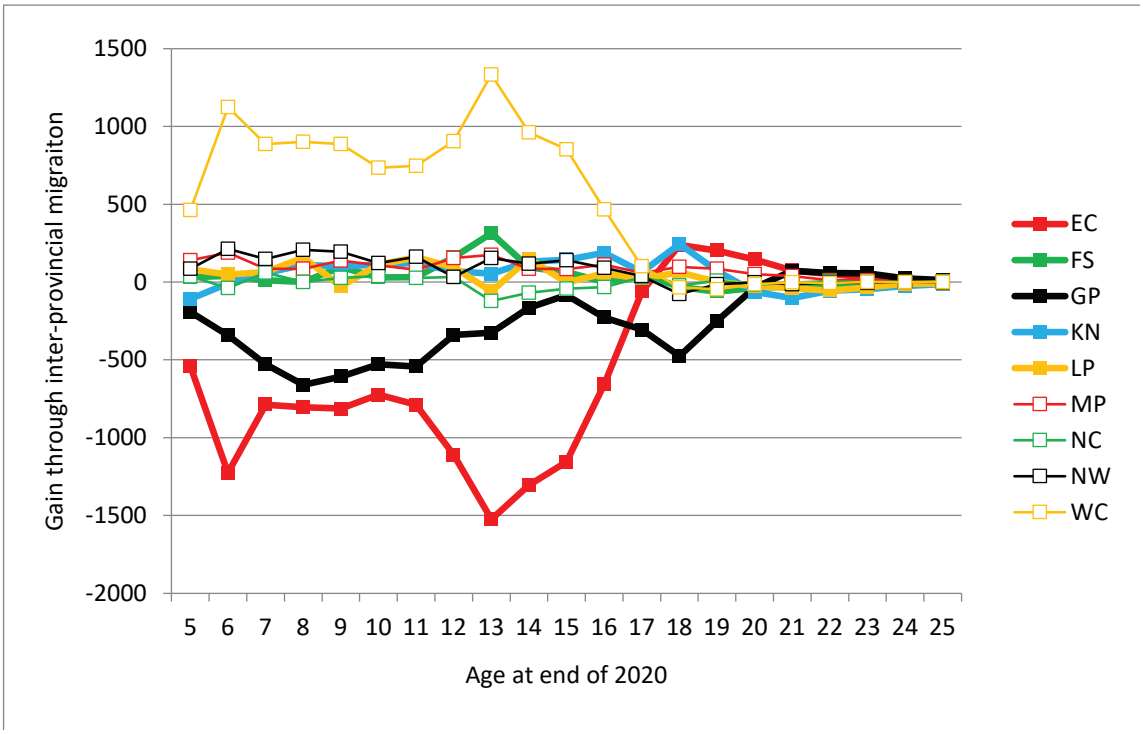


Figure 8: Provincial migration gains 2020 to 2021

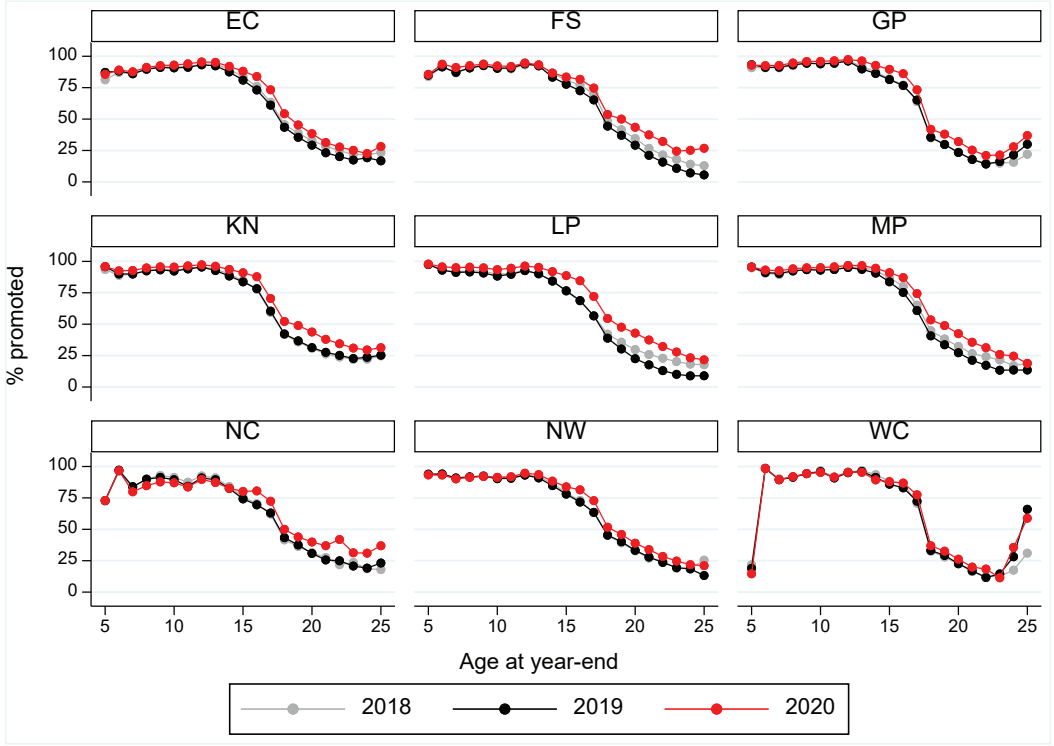


The change in Gauteng from being a net receiver to being a net sender of migrating learners is clear from the two graphs.

Figure 9 to Figure 11 below should be understood in the light of the migration patterns. Specifically, dropping out will be under-estimated somewhat in provinces receiving many migrating learners, and over-estimated in provinces sending many such learners. For instance, many drop-ins in Western Cape in 2020 are actually unlinked learners, rather than drop-ins, but what is not known is how many would be in Western Cape and how many would be in another province in 2019. Learners who were in another province in 2019 would be used to reduce the dropping out numbers in Western Cape, even if they should not be used for this. This would lead to some under-estimation of the dropping out phenomenon in Western Cape. The opposite would occur in Eastern Cape. But the margin within which the problem occurs is small, judging from what has been presented above. Around 3% of Western Cape's learners of compulsory age are not linked (Table 1), and perhaps 1.0% of these learners would have been in another province in the previous year (see Table 4). This translates to 0.03% of Western Cape's learners being unlinked drop-ins from another province. Even if one considers the linking problems would be aggravated when learners move from one province to another, the distortion to the provincial flow statistics caused by across-province migration seems low enough not to change the statistics substantively.

From Figure 9 it is clear that all provinces contributed to the increase in the promotion ratios during the pandemic, in other words between 2020 and 2021. However, the trend was particularly pronounced in Limpopo.

Figure 9: Provincial age-specific promotion ratios



From the next two graphs it is clear that a reduction in grade repetition was the major driver of the increase in the promotion ratios in all nine provinces between 2019 and 2020. Departure ratios moved in different directions: Western Cape and Northern Cape appear to have experienced an increase in departures, while in for instance Limpopo and North West it declined.

The Gauteng spike for 13-year old departures in 2018 and 2019 is striking. A closer examination of the data behind Figure 6 reveals that in 2018 Gauteng's ratio of 13-year-olds to 8-year-olds was the lowest among all provinces, and that in 2019 Gauteng's ratio was the second-lowest (after Limpopo). This suggests that the departure spike for age 13 could be indicative of an actual phenomenon, as opposed to a manifestation of data problems, specifically linking problems. The matter could be examined in more detail using the data. For instance, could this be an indication of problems in the transition from primary to secondary schools?

Figure 10: Provincial age-specific repeater ratios

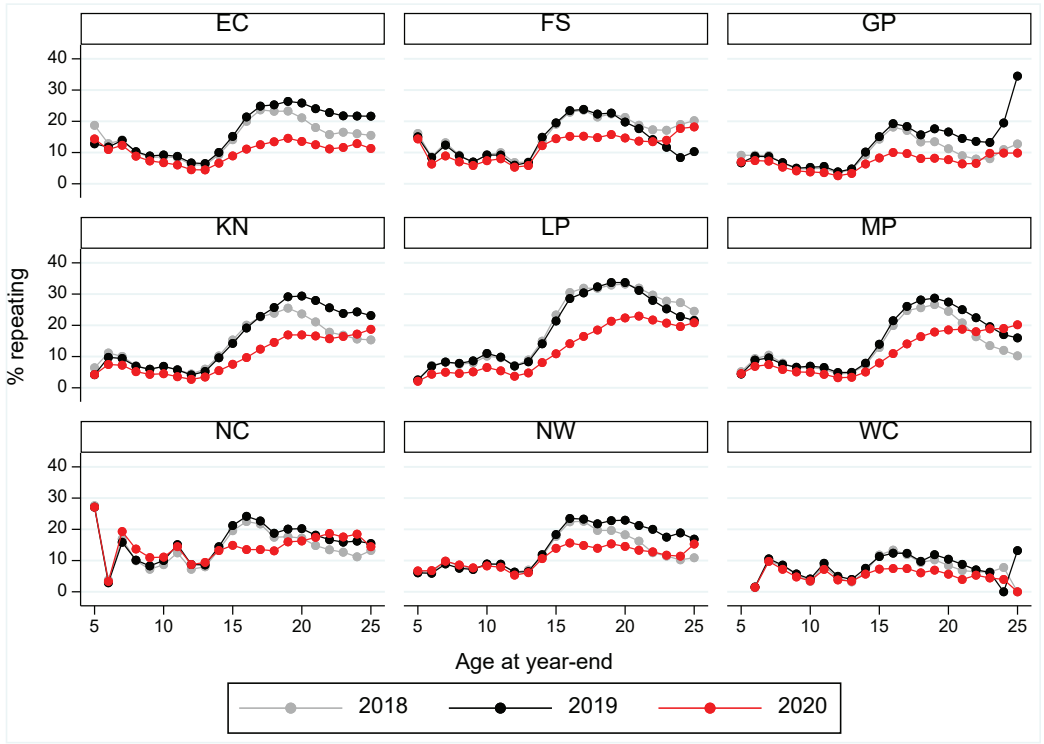
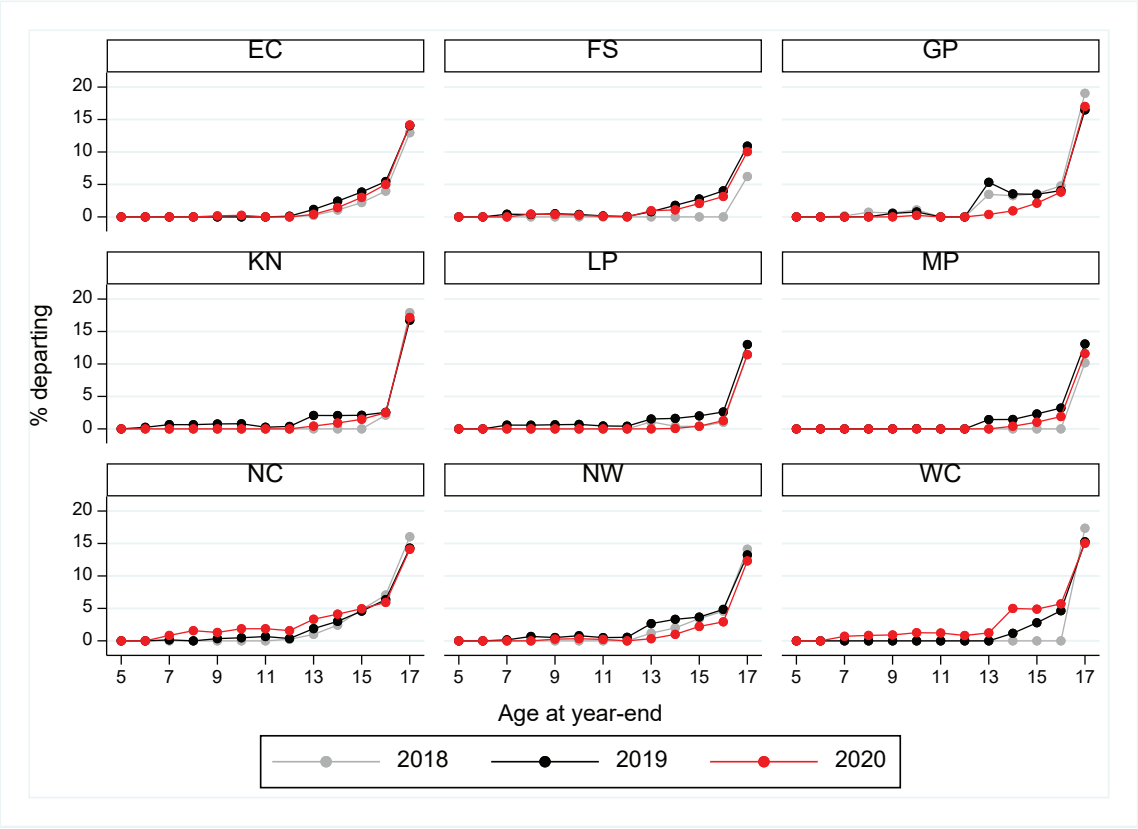


Figure 11: Provincial age-specific departure ratios



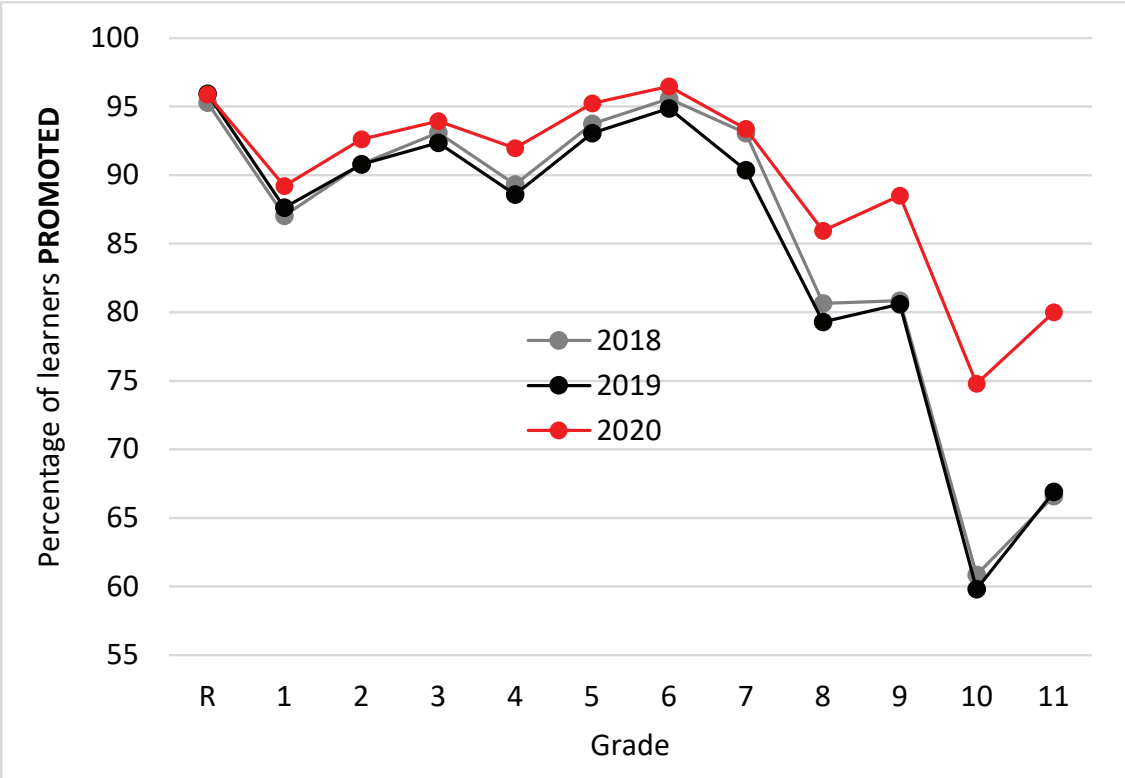
5 Grade-specific flow ratios

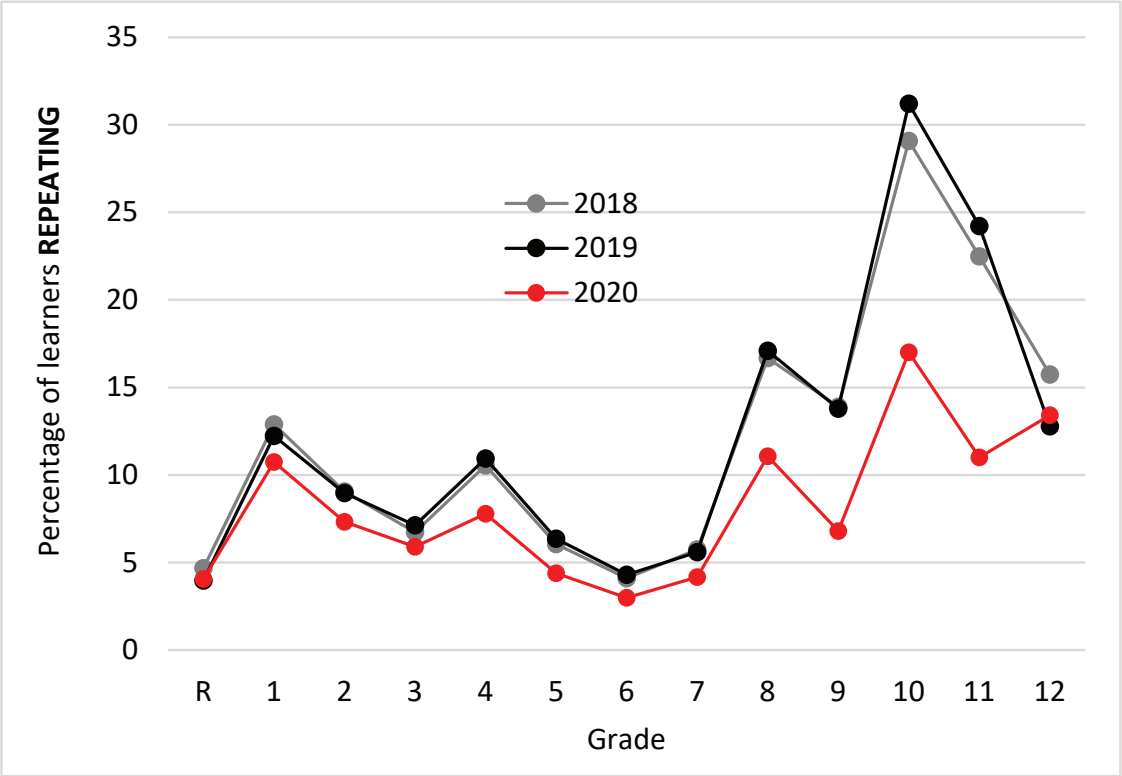
Figure 12 presents grade-specific flow ratios for the three periods. The methodology followed here was is based on the methodology for the age-specific statistics. The problem when calculating grade-specific statistics is that while a drop-in's age in the previous year can be determined with certainty, this learner's grade cannot, as we do not know whether this learner repeated or not. The number of drop-outs per grade, and age, was used to determine how to classify the drop-ins. An example dealing with learners aged 15 at the end of the base year illustrates the method. Assume that 50 of these 15-year-olds were apparent departures from Grade 8, while 100 were apparent departures from Grade 9. Assume too that 15-year-olds departed only from these two grades. If the number of drop-*ins* aged 16 at the end of the second year was 60, then regardless of their grade in the second year, 50 over 150 times 60 of them would be assumed to be unlinked learners in Grade 8 in the first year, and 100 over 150 times 60 of them would be assumed to be unlinked learners in Grade 9 in the first year. This translates to 20 and 40 learners, meaning the 15-year-old drop-*outs* would be reduced to 30 and 60 for grades 8 and 9 respectively. The process would be repeated for other ages.

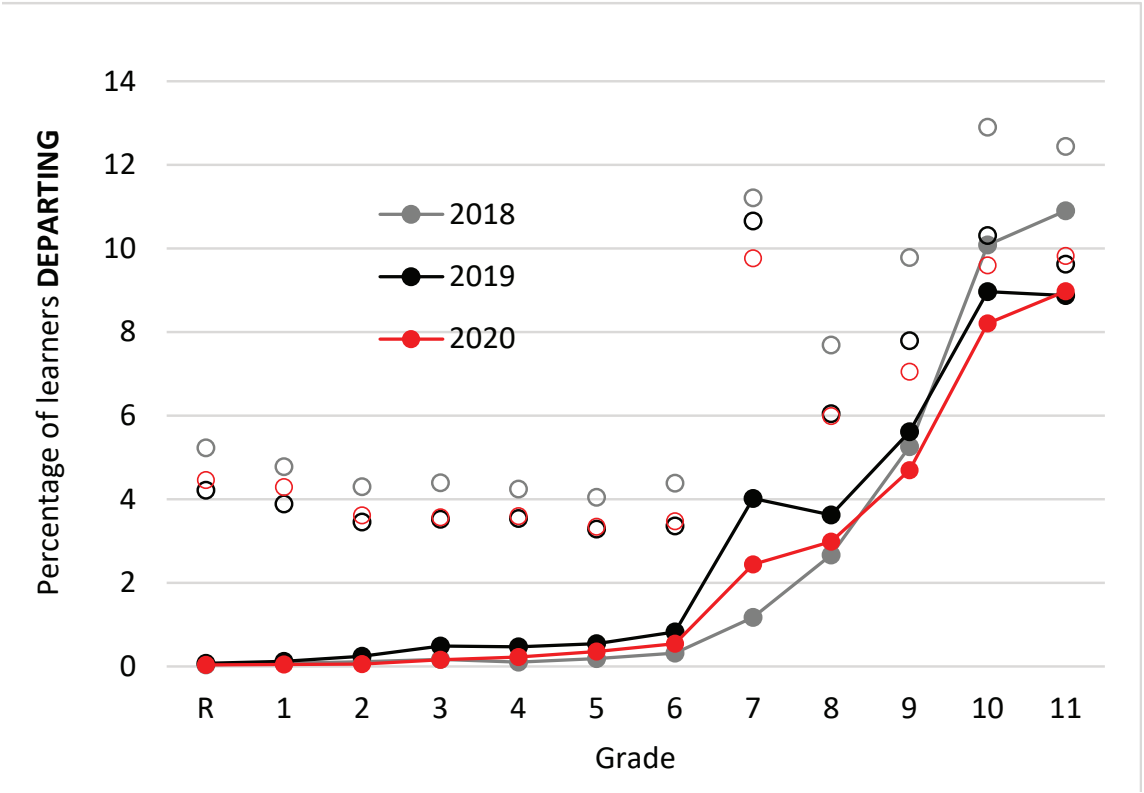
This method could be developed further to take into account the grade of a drop-in in the second year. However, it seems unlikely this would make a substantial difference to the flow statistics.

Figure 12 indicates that in the grades 1 to 12 range promotion rose during the pandemic across all grades except Grade 7, that repetition declined across all grades except Grade 12, and that dropping out from grades 9 and 10 was particularly low at the end of 2020.

Figure 12: National grade-specific flow statistics







The following three figures present the province-level grade-specific statistics. Northern Cape stands out as having experienced somewhat different patterns from the other provinces during the pandemic: in grades 1 to 3 repetition *increased*, as did dropping out. In Western Cape, dropping out also appears to have been worryingly high at the primary level following 2020. Gauteng's high dropping out from Grade 7 at the end of the two *pre*-pandemic years is the grade-specific manifestation of the age 13 issue already discussed in section 4, an issue which seems to warrant further interrogation of the data.

Figure 13: Provincial grade-specific promotion ratios

Figure 14: Provincial grade-specific repeater ratios

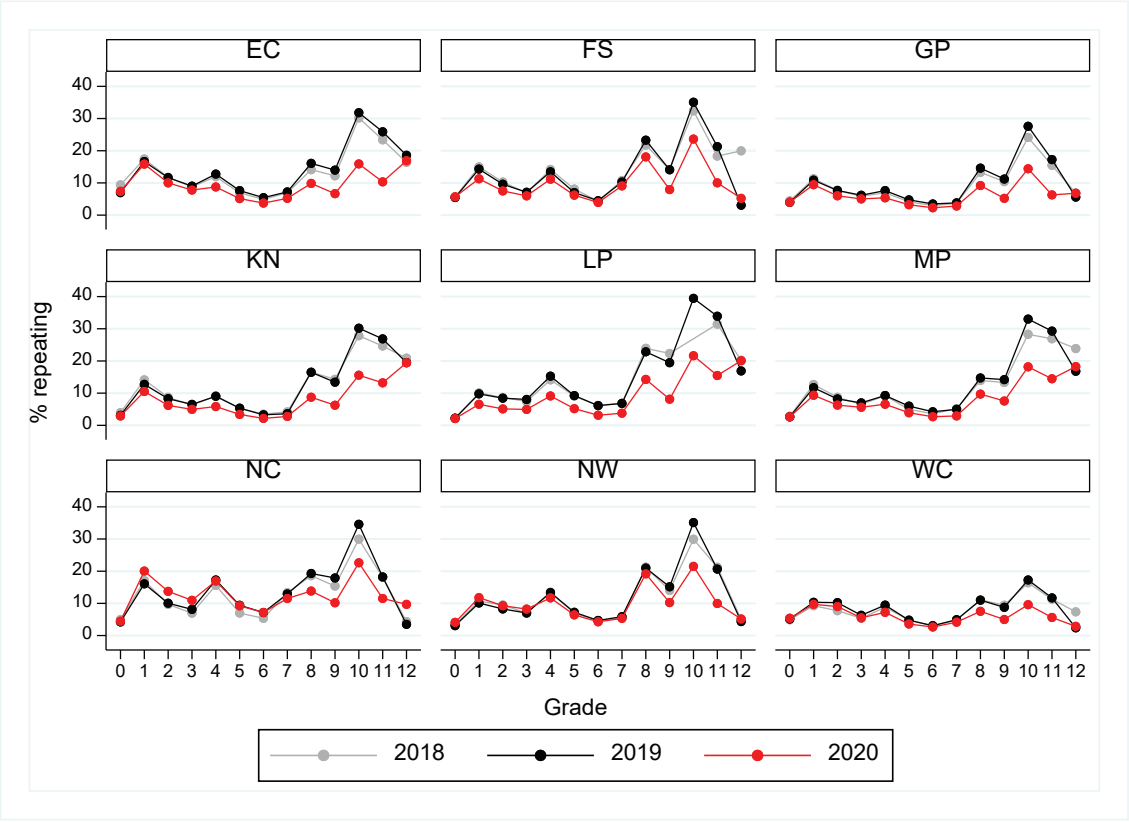
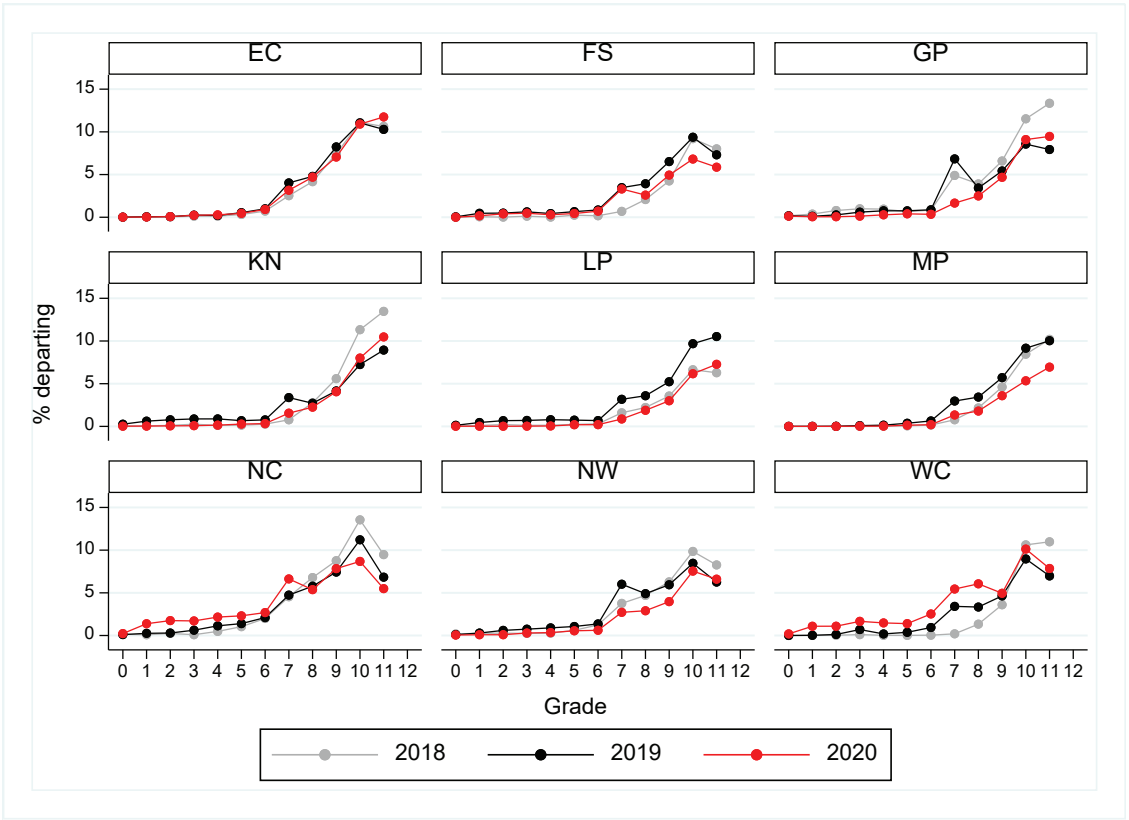


Figure 15: Provincial grade-specific departure ratios



The following four tables provide the statistics of Figure 12 to Figure 15. For each province and grade the percentages add up to 100% across Table 5, Table 6 and Table 7 (rounding means the total may not be precisely 100% in all instances).

Table 5: Grade-specific promotion ratios

Year	Grade	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
2018	R	91	94	95	96	98	97	95	96	95	95
2019	R	93	94	96	97	98	97	96	97	95	96
2020	R	93	94	96	97	98	97	95	96	94	96
2018	1	83	85	88	86	90	87	83	89	91	87
2019	1	83	85	89	87	90	88	84	90	90	88
2020	1	84	89	90	89	93	91	79	88	89	89
2018	2	88	90	91	91	91	91	90	90	92	91
2019	2	88	90	92	91	91	92	90	91	90	91
2020	2	90	92	94	94	95	94	85	91	90	93
2018	3	91	93	93	93	92	93	93	93	95	93
2019	3	91	92	93	93	91	93	91	92	93	92
2020	3	92	94	95	95	95	94	87	91	93	94
2018	4	88	86	92	91	86	91	84	86	91	89
2019	4	87	86	92	90	84	91	82	86	90	89
2020	4	91	89	94	94	91	93	81	88	91	92
2018	5	93	92	95	95	91	95	92	92	95	94
2019	5	92	92	95	94	90	94	89	92	95	93
2020	5	94	93	96	96	95	96	88	93	95	95
2018	6	94	95	96	96	94	96	93	94	97	96
2019	6	94	95	96	96	93	95	91	94	96	95
2020	6	95	95	97	98	97	97	90	95	95	96
2018	7	91	89	91	95	92	94	82	91	95	93
2019	7	89	86	89	93	90	92	82	88	92	90
2020	7	92	88	96	96	95	96	82	92	90	93
2018	8	82	76	83	81	74	84	75	74	88	81
2019	8	79	73	82	81	74	82	75	74	86	79
2020	8	85	79	88	89	84	88	81	78	86	86
2018	9	80	82	83	80	74	82	76	80	87	81
2019	9	78	79	83	82	75	80	75	79	87	81
2020	9	86	87	90	90	89	89	82	86	90	89
2018	10	59	58	64	61	52	63	56	60	73	61
2019	10	57	56	64	63	51	58	54	56	74	60
2020	10	73	70	76	76	72	76	69	71	80	75
2018	11	66	74	71	62	62	63	72	70	78	67
2019	11	64	71	75	64	56	61	75	73	81	67
2020	11	78	84	84	76	77	79	83	83	87	80
2018	12	0	0	0	0	0	0	0	0	0	0
2019	12	0	0	0	0	0	0	0	0	0	0
2020	12	0	0	0	0	0	0	0	0	0	0

Table 6: Grade-specific repeater ratios

Year	Grade	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
2018	R	9	6	4	4	2	3	5	3	5	5
2019	R	7	5	4	3	2	3	4	3	5	4
2020	R	7	6	4	3	2	3	4	4	5	4
2018	1	17	15	11	14	10	13	17	11	9	13
2019	1	17	14	11	13	10	12	16	10	10	12
2020	1	16	11	9	11	7	9	20	12	10	11
2018	2	12	10	8	9	9	9	10	10	8	9
2019	2	12	10	8	8	8	8	10	8	10	9
2020	2	10	8	6	6	5	6	14	9	9	7
2018	3	9	7	6	6	7	7	7	7	5	7
2019	3	9	7	6	7	8	7	8	7	6	7
2020	3	8	6	5	5	5	6	11	8	6	6
2018	4	12	14	7	9	14	9	16	13	9	11
2019	4	13	13	8	9	15	9	17	13	9	11
2020	4	9	11	5	6	9	7	17	12	7	8
2018	5	7	8	4	5	9	5	7	7	5	6
2019	5	8	7	5	5	9	6	9	7	5	6
2020	5	5	6	3	3	5	4	9	6	4	4
2018	6	5	4	3	3	6	4	5	5	3	4
2019	6	5	4	3	3	6	4	7	5	3	4
2020	6	4	4	2	2	3	3	7	4	3	3
2018	7	7	11	4	4	7	5	13	6	5	6
2019	7	7	10	4	4	7	5	13	6	5	6
2020	7	5	9	3	3	4	3	12	5	4	4
2018	8	14	22	13	17	24	14	19	21	11	17
2019	8	16	23	15	16	23	15	19	21	11	17
2020	8	10	18	9	9	14	10	14	19	8	11
2018	9	12	14	10	14	22	13	15	14	9	14
2019	9	14	14	11	13	19	14	18	15	9	14
2020	9	7	8	5	6	8	8	10	10	5	7
2018	10	30	32	24	28	41	28	30	30	16	29
2019	10	32	35	28	30	39	33	35	35	17	31
2020	10	16	24	14	16	22	18	23	22	10	17
2018	11	23	18	15	25	31	27	18	21	11	22
2019	11	26	21	17	27	34	29	18	21	12	24
2020	11	10	10	6	13	15	14	12	10	6	11
2018	12	16	20	7	21	20	24	4	5	7	16
2019	12	19	3	6	19	17	17	3	4	2	13
2020	12	17	5	7	19	20	18	10	5	3	13

Table 7: Grade-specific departure ratios

Year	Grade	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
2018	R	0	0	0	0	0	0	0	0	0	0
2019	R	0	0	0	0	0	0	0	0	0	0
2020	R	0	0	0	0	0	0	0	0	0	0
2018	1	0	0	0	0	0	0	0	0	0	0
2019	1	0	0	0	1	0	0	0	0	0	0
2020	1	0	0	0	0	0	0	1	0	1	0
2018	2	0	0	1	0	0	0	0	0	0	0
2019	2	0	0	0	1	1	0	0	1	0	0
2020	2	0	0	0	0	0	0	2	0	1	0
2018	3	0	0	1	0	0	0	0	0	0	0
2019	3	0	1	1	1	1	0	1	1	1	0
2020	3	0	0	0	0	0	0	2	0	2	0
2018	4	0	0	1	0	0	0	0	0	0	0
2019	4	0	0	1	1	1	0	1	1	0	0
2020	4	0	0	0	0	0	0	2	0	1	0
2018	5	0	0	1	0	0	0	1	1	0	0
2019	5	1	1	1	1	1	0	1	1	0	1
2020	5	0	0	0	0	0	0	2	1	1	0
2018	6	1	0	1	0	0	0	2	1	0	0
2019	6	1	1	1	1	1	1	2	1	1	1
2020	6	1	1	0	0	0	0	3	1	3	1
2018	7	3	1	5	1	2	1	5	4	0	1
2019	7	4	3	7	3	3	3	5	6	3	4
2020	7	3	3	2	2	1	1	7	3	5	2
2018	8	4	2	4	3	2	2	7	5	1	3
2019	8	5	4	3	3	4	3	6	5	3	4
2020	8	5	3	2	2	2	2	5	3	6	3
2018	9	7	4	7	6	4	5	9	6	4	5
2019	9	8	7	5	4	5	6	7	6	5	6
2020	9	7	5	5	4	3	4	8	4	5	5
2018	10	11	9	12	11	7	8	14	10	11	10
2019	10	11	9	9	7	10	9	11	8	9	9
2020	10	11	7	9	8	6	5	9	8	10	8
2018	11	11	8	13	13	6	10	9	8	11	11
2019	11	10	7	8	9	11	10	7	6	7	9
2020	11	12	6	9	10	7	7	5	7	8	9
2018	12	84	80	93	79	80	76	96	95	92	84
2019	12	81	97	94	80	83	83	95	96	97	87
2020	12	83	95	93	81	80	82	90	95	97	87

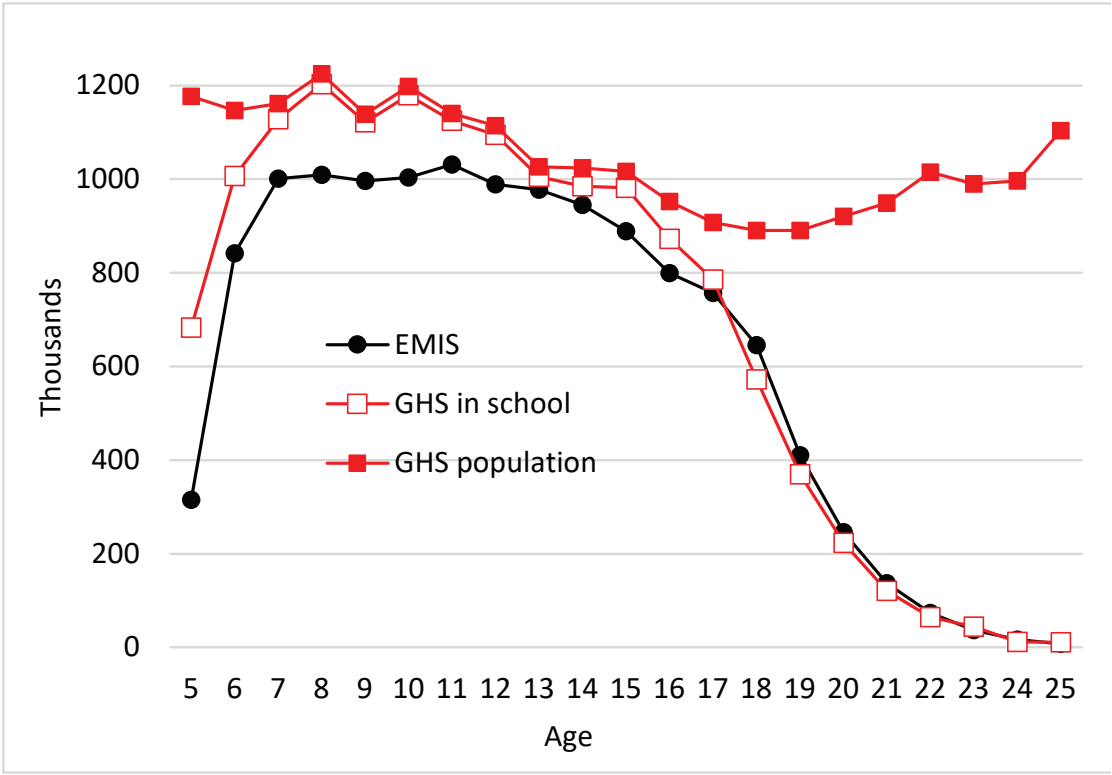
Table 8: Grade-specific departure ratios BEFORE DROP-IN ADJUSTMENT

Year	Grade	EC	FS	GP	KN	LP	MP	NC	NW	WC	SA
2018	R	5	3	10	5	3	5	2	4	6	5
2019	R	5	2	7	4	3	3	3	3	2	4
2020	R	6	3	8	4	3	3	4	4	3	4
2018	1	4	3	8	5	3	5	2	4	6	5
2019	1	5	2	6	3	3	3	2	3	3	4
2020	1	6	3	7	4	2	3	4	4	4	4
2018	2	3	2	7	4	3	5	2	4	5	4
2019	2	4	2	6	3	3	3	2	3	2	3
2020	2	5	2	6	3	2	2	3	3	3	4
2018	3	3	2	7	4	2	5	2	3	6	4
2019	3	4	2	6	3	3	3	2	3	3	4
2020	3	5	2	6	3	2	3	3	3	3	4
2018	4	3	2	7	4	3	6	3	3	5	4
2019	4	4	2	6	3	3	3	3	3	2	4
2020	4	5	2	6	3	2	3	4	3	3	4
2018	5	3	3	6	4	2	6	3	4	5	4
2019	5	4	2	5	3	3	3	3	3	2	3
2020	5	5	2	5	3	2	2	4	3	3	3
2018	6	3	3	6	4	2	7	4	4	6	4
2019	6	4	2	5	3	2	3	3	3	4	3
2020	6	5	2	5	3	2	3	4	3	4	3
2018	7	7	7	20	9	9	12	8	10	12	11
2019	7	10	6	19	9	10	9	7	11	6	11
2020	7	10	7	15	8	8	9	10	10	7	10
2018	8	7	6	10	8	5	7	9	7	10	8
2019	8	8	5	7	5	6	6	7	6	6	6
2020	8	8	4	8	5	4	4	7	5	8	6
2018	9	10	8	12	10	6	11	10	8	11	10
2019	9	11	7	9	6	7	8	8	7	6	8
2020	9	10	6	9	7	5	6	9	5	6	7
2018	10	13	12	15	15	8	12	14	11	16	13
2019	10	13	10	10	9	11	10	12	9	10	10
2020	10	13	7	11	10	7	6	9	8	11	10
2018	11	12	10	15	16	7	12	10	9	13	12
2019	11	11	7	9	10	11	11	7	6	7	10
2020	11	13	6	10	12	8	8	6	7	8	10
2018	12	91	94	99	93	86	91	99	99	99	94
2019	12	89	99	99	92	88	90	98	99	99	94
2020	12	91	98	98	92	86	89	96	98	98	93

6 Validation against other sources

How well do the statistics presented above agree with previously calculated statistics? Figure 16 below suggests that participation by age in 2019 is reliable in the data used for the current analysis. The same finding would arise if another year were chosen. Here the 2019 curve from Figure 1 above has been reproduced, and two additional curves drawing from the 2019 General Household Survey (GHS) data of Stats SA are included. Given that absolute numbers in the GHS rely strongly on sampling and weighting procedures, which might have limitations, it is not surprising that there is far more uniformity across ages 7 to 14 in the EMIS data, which is censal, than in the GHS data. What *can* be considered reliable in the GHS data is the percentage of children who are enrolled in school, in other words the ratio produced by the two GHS curves. This in turn suggests that the GHS weights over-estimate the number of children in the population, and in schools, up to around age 12 by a considerable margin. These types of discrepancies have been noted in previous analyses. In view of this, discrepancies between the EMIS and GHS curves in the graph are not concerning.

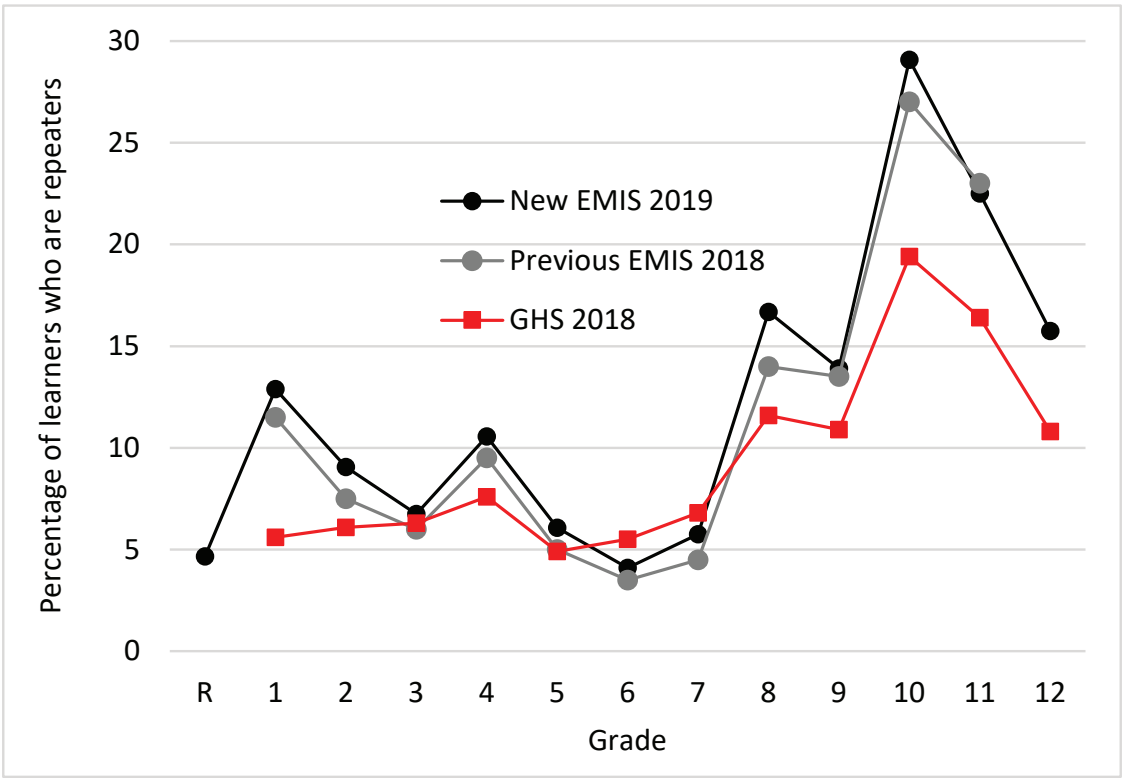
Figure 16: Participation in 2019 in EMIS and the GHS



Source: The GHS curves are based on own analysis of the 2019 GHS microdata.

For Figure 17, data behind the analysis of the earlier sections was used to calculate the percentage of learners in a grade who are repeaters. This is different from what is presented in, for instance, Table 6, where learners who will repeat in the *next* year are divided by the total learners in the *current* year. For the purposes of Figure 17, the numerator and denominator are from the *same* year, for all three series of statistics. The other two sources only had statistics calculated in this manner. The patterns seen in 2019 using the new EMIS dataset analysed in the current report are very similar to the 2018 EMIS-based patterns presented by Van der Berg *et al* (2021). The GHS statistics are a bit different, but generally follow a similar pattern. In particular grade repetition appears to under-reported in the household data, which is perhaps to be expected, given the stigma attached to grade repetition. Proportionally, the greatest degree of under-reporting is in Grade 1, where the GHS statistic is less than half that seen in the EMIS data. This discrepancy has been noted in previous analyses, but why it should exist is not clear.

Figure 17: Extent of grade repetition from three analyses



Source: 'Previous EMIS 2018' is read from Figure 11 in Van der Berg et al *(2021). 'GHS 2018' is from Table 62 of Department of Basic Education (2019).*

References

Department of Basic Education (2017). *Use of LURITS data to arrive at transition rates and enrolment scenarios*. Pretoria. [Unpublished report.]

Department of Basic Education (2019). *General Household Survey (GHS): Focus on schooling 2018*. Pretoria.

Department of Basic Education (2020). *Action Plan to 2024: Towards the realisation of Schooling 2030*. Pretoria.

Van der Berg, S., Van Wyk, C., Selkirk, R., Hofmeyr, H. (2021). *Learner flows through schools: Using high quality administrative data to understand education system performance.* Stellenbosch: University of Stellenbosch.

Make Every Child A National Asset

Published by the Department of Basic Education
222 Struben Street
Private Bag X895, Pretoria, 0001
Telephone: 012 357 3000 Fax: 012 323 0601
© Department of Basic Education