RESEP
Research on Socio-Economic Policy

Stellenbosch University
Iyunivesithi
Universiteit

forward together
sonke siya phambili
saam vorentoe

RESEP WORKING PAPER

Department of Economics, Stellenbosch University

Working Paper No.

04/25

JANUARY 2025

This paper was produced as part of the MILAPS project, funded by Optima

Assessing the Relationship Between School-Based Assessment and Matric Performance in Mathematics: A Quantitative Analysis of Poor Schools in Limpopo

Author

Janke Claasen

OPTIMA

Assessing the Relationship Between School-Based Assessment and Matric Performance in Mathematics: A Quantitative Analysis of Poor Schools in Limpopo

Janke Claassen
January 2025

Abstract

This study investigates whether the quality of school-based assessment (SBA) influences matriculation performance in Mathematics within quintile 1 to 3 schools in Limpopo, South Africa. It builds on previous literature that identified poor schools in Limpopo excelling in Mathematics achievement at the matriculation level. The research aims to determine whether high-quality assessments contribute to improved outcomes in high-stakes examinations. Using a panel dataset with learner level SBA data, the paper evaluates assessment quality through two metrics: accuracy (correlation between SBA and NSC marks) and leniency (the gap between these marks). Findings reveal that both leniency and accuracy in assessment improved significantly from earlier grades in the Further Education and Training (FET) phase to matric. Additionally, the sampled period overlaps with years affected by COVID-19, during which changes in assessment quality were observed due to assessment policies implemented to mitigate the effects of school disruptions. Significantly higher leniency in assessment was detected in grades 10 and 11 between 2020 and 2022. Quintile 1 to 3 schools with consistently high assessment quality, even during COVID-19-affected years, were identified as "outlier" poor schools. These outlier schools achieved better matric Mathematics results than other poor schools in the province, highlighting a positive relationship between high-quality SBA and matric Mathematics performance. The study underscores the importance of meaningful assessment practices and identifies schools with high quality assessment as potential models for educational improvement in resource-constrained settings.

JEL Classification: I20, I21, I28

Keywords: South Africa, Education, School-Based Assessment, Mathematics Performance, Matric (NSC) Examinations

1 Introduction

The persistent underperformance of learners, particularly in Mathematics, represents a significant impediment to South Africa's economic and social development. Policymakers remain concerned about the limited number of black African learners attaining the necessary proficiency in Mathematics to pursue a university education that requires foundational mathematical skills. This challenge stems from inadequate academic achievement and notable disparities persisting along provincial and social demographic lines, even preceding matric (grade 12). Nevertheless, there are indications of progress, with no-fee schools producing an increasing number of black African learners attaining a Mathematics mark exceeding 60% in the National Senior Certificate (NSC) examination. Notably, Limpopo is one of two provinces that have disproportionately contributed to this rise in high-performing black African Mathematics learners (Gustafsson, 2016).

The underlying factors driving this increase in higher achieving Mathematics learners remain unclear. This paper seeks to explore whether quality school-based assessment (SBA) plays a role in fostering the growth of high-achieving Mathematics learners from poor schools in Limpopo. The research concern is not with how SBA is used in teaching, but rather about what feedback it provides for learners and its ultimate impact on their academic performance in matric. For the purposes of this research, the final matriculation examination is seen as the "correct" assessment of a learner's performance since it is standardised and externally moderated. Accordingly, the effectiveness of SBA is evaluated in relation to the NSC results.

The quality of assessment for different cohorts across grades 9 to 12 was evaluated using metrics such as correlations and gaps to measure assessment accuracy and leniency, respectively. Poor schools that demonstrated high-quality assessment in at least one cohort were classified as "outlier" poor schools. These outlier schools were subsequently compared to other quintile 1 to 3 schools to assess whether their matric Mathematics performance was significantly higher and whether this improved performance could be attributed to the higher quality assessments within these schools. The results of the quantitative analysis revealed a positive relationship between learners' matric Mathematics performance and the quality of assessment within the school.

The remainder of the paper is structured as follows. Section 2 of the paper provides a comprehensive literature review to contextualise the research question. It examines existing literature on the importance of SBA, the performance of South African learners in Mathematics, the performance of Limpopo learners in Mathematics, and the impact of COVID-19 on Mathematics assessment. The methodology employed to identify schools with high-quality assessment is outlined in Section 3. Section 4 describes the SBA and NSC data utilised in the analysis. In Section 5 the results stemming from a brief analysis of matric assessment data pre- and post-COVID-19 as well as the results of the analysis of assessment accuracy and leniency is presented. Section 6 identifies outlier schools based on the quality of assessment within these schools and explores whether assessment accuracy contributes to improved performance in the final NSC results of the learners in these schools. The final section concludes by discussing the findings and identifying areas for future research.

2 Literature Review

2.1 The Importance of School-Based Assessment

Internationally, two principal types of assessment in schools are well established within the field of education research: formative and summative assessments. The classification of an assessment as either formative or summative is determined by its purpose. These terms were originally introduced by Scriven (1967). In a more recent explanation, he defined the two types of assessment as follows:

> Formative evaluation... is typically conducted during the development or improvement of a program... and it is conducted, often more than once, for the in-house staff of the program with the intent to improve. (Scriven, 1991, p. 168)

> Summative evaluation of a program... is conducted after completion of the program... and for the benefit of some external audience or decision-maker. (Scriven, 1991, p. 340)

Newton (2007) challenges the clarity of the dichotomy, arguing that there are "many distinct assessment purposes". For instance, while diagnostic purposes are more commonly associated with formative assessments, summative results can also serve diagnostic functions. Summative assessments are further utilised for self-referenced judgments (e.g., assessing whether an individual has improved compared to a prior performance) and norm-referenced judgments (e.g., determining whether a learner has outperformed a certain percentage of their peers) (Newton, 2007).

There is a growing body of literature that explores the link between assessment and learning across all levels of education. In an experimental evaluation of the change in assessment feedback methods in a secondary school in Wales, Smith and Gorard (2005) found that learners did not value formative feedback on assessments more than simply receiving summative marks. Furthermore, the study could not find conclusive evidence that learners who received the formative feedback performed better over time than the learners who received only their summative marks. Pitt *et al.* (2020), in a study on students in higher education, highlighted that another layer of complexity is added to interpreting the value of feedback on assessment to learners. Their findings indicated that the perceived value of feedback varies depending on the learner's academic achievement, with lower-achieving students tending to extract less benefit from feedback.

O'Donovan *et al.* (2016) attempted to address the disappointing gap between the theoretical benefit of feedback practices and their practical success by identifying several solutions from the literature. However, they acknowledged that these proposed solutions to the feedback dilemma will only work if assessment literacy improve for both educators and learners.

Baird *et al.* (2017) argue that assessment serves as a communicative device, conveying what curriculum designers expect students to know and master. Wiliam (2010) supports this perspective noting that assessment tools operationalise these expectations and thus strongly influence teaching methods in the classroom. Through diverse assessment tools and materials, the education community establishes a framework for evaluating and validating learning outcomes. There is thus broad consensus that feedback from assessments is intrinsically linked to educational outcomes.

At a national level, Taylor *et al.* (2011) underscore the importance of meaningful assessment practices as well as valuable feedback to learners when choosing whether to take Mathematics or Mathematical Literacy in South Africa. Their study revealed a high degree of randomness within the historically black part of the schooling system when it comes to choosing between these two mathematics subjects. This is very concerning when considering that 5 out of 6 learners who fail matric also fail a mathematics subject.

Van der Berg and Shepherd (2015) investigated the different dimensions of weak signalling by continuous assessments (CASS) for different matric subjects in South Africa. They evaluated assessment quality using two criteria: (i) high assessment reliability (the CASS and matric examination marks were highly correlated), and (ii) low leniency (CASS marks are not significantly inflated compared to matric examination marks). CASS marks that were excessively inflated or poorly correlated with the final matric examination results were classified as 'weak', 'poor', or 'inaccurate' assessments (Van der Berg and Shepherd, 2015, p. 81).

2.2 Mathematics Performance in South Africa

The poor performance of South African learners in Mathematics is evident across national (Annual National Assessments (ANA)), regional (Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)) and international (Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS)) assessments. This persistent underachievement has become a significant concern for policymakers, particularly within no-fee schools where performance is notably weaker. Despite this, there are signs of improvement in the mathematical performance of South African learners (Gustafsson, 2016).

South Africa has participated in seven rounds of TIMSS, with the most recent round conducted in 2023. TIMSS is an international assessment that assesses the Mathematics performance of grade 8 learners, but in South Africa it is administered to grade 9 learners. Reddy and Hannan (2018) used the results from the different rounds of TIMSS assessments to show that there has been an improvement in the Mathematics performance of South African learners. The authors illustrated how the national Mathematics scores stagnated for the first three rounds (1995, 1999, and 2003) at a 'very low level', but improved from 2003 to 2015 by 87 points, almost one standard deviation. Leaving South Africa with a now 'low' national average. Van der Berg and Gustafsson (2019), generated a threshold for Mathematics from the 20th percentile of white matriculation candidates in Mathematics to analyse the changes in matric Mathematics performance. This specific threshold was used since it is assumed that the threshold will approximate a constant performance level (the threshold was also similar to the Mathematics entry requirement for engineering and other related disciplines at universities). Between 2002 and 2016, the number of black African learners achieving this benchmark increased by 65%. Notably, the bulk of this improvement was seen among schools that largely serve the poorer segments of the black African population. Limpopo and Mpumalanga yielded disproportionally large gains in high-level performance amongst the lower quintile schools. In the South African education sector, quintile rankings are determined by the average socioeconomic status of the learners in the school. Quintile 5 schools serve the wealthiest learners while quintile 1 schools caters to the poorest. Encouragingly, Van der Berg and Gustafsson (2019) reported that the most rapid improvement in Mathematics performance occurred within quintile 1 to 3 schools. Quintile 1 schools realised a 160% increase in the number of high-level achievers, while schools in quintiles 2 and 3 yielded an impressive increase of 90%.

2.3 Limpopo Schools' Performance in Mathematics

Gustafsson (2016) identified the provinces and districts responsible for producing the majority of black African learners who achieved at least 60% in Mathematics. His findings highlighted that several districts and schools in Limpopo were notable for generating a substantial number of such learners. Similarly, a report by Van der Berg *et al.* (2023) identified 'promising Mathematics schools' – defined as schools of which more than a third of the grade 9's in 2019 reached matric, elected to do Mathematics rather than Mathematical Literacy and at least one-third of those learners achieved a mark of more than 60% in Mathematics in matric. Limpopo emerged as the province with the largest number of promising schools. This result supported the finding of Gustafsson (2016) which illustrated that Limpopo produced a large number of promising schools.

However, the learners who achieved above 60% in Mathematics in the matric examinations represent less than 20% of their matric class who enrolled for Mathematics. This raises the question of whether the high performance of these learners reflects the level of functionality of the school they attend or not (Shepherd and Van der Berg, 2020). While focussing on primary schools, Wills (2017) also stressed the importance of identifying the outlier schools which "are exceptions to the norm" and "produce at least adequate levels of learning". Supporting these schools to ensure their sustained improvement is a key policy consideration.

Analysing the relationship between the performance of learners in SBA and their performance in the matric examinations provides an opportunity to evaluate the quality of SBA within schools. The correlation between the performance of a learner in these two types of assessments may offer greater insight into whether the quality of assessment may be an indicator of the functionality of the school, ultimately contributing to improved Mathematics results in matric.

2.4 The Effect of COVID-19 on Mathematics Assessment

The period for which the learner and school-level data is investigated stretches from 2018 to 2022, and thus overlaps with the onset and peak of the global COVID-19 pandemic, which had severe consequences for the education sector. The adverse impact of the COVID-19 pandemic on learning outcomes was especially dire for no-fee schools (Soudien *et al.*, 2022). As such, it is important to provide a brief overview of the pandemic's effects on educational outcomes, particularly the loss of learning time, as well as the policy changes implemented to mitigate the negative effects of this global disruption. Between 2020 and 2021, the majority of South African learners missed at least three-quarters of a school year due to the COVID-19 pandemic, which led to widespread school closures and rotational timetables (Wills and Van der Berg, 2022). The average number of school days missed in 2020 as a percentage of school days in 2019 was as follows: 42% for grade 9, 34% for grade 10, 32% for grade 11, and 22% for matric (Kotze, 2021). Daily school attendance only fully resumed in February 2022.

The reduction in contact time translated to a corresponding reduction in curriculum coverage (Bisgard *et al.*, 2022). In response, the Department of Basic Education (DBE) recommended the removal of certain topics from the syllabus for specific subjects to allow teachers to prioritise the essential core content for each grade within the available time. In Mathematics, topics on Statistics and Finance were removed from the Grade 10 and Grade 11 syllabi in 2020, while no topics were excluded from the matric syllabus.

Moreover, full examinations were substituted with controlled tests in 2020 as well as 2021. It was explicitly stated that "controlled tests should only be set on content taught, content not taught cannot be assessed." As a result, no common assessments were allowed in grades 10 and 11 during these years (Department of Basic Education, 2020*a*). Learners who failed Mathematics in grade 9 during 2020 to 2022, but met all other passing requirements and were condoned to grade 10, could still take Mathematics in grade 10 (Department of Basic Education, 2022).

In 2020, the weightings of SBA and final examinations in determining learners' final marks were adjusted. Prior to 2020, the weighting of SBA and final examinations in Mathematics in the FET phase was 25:75 (Department of Basic Education, 2011), but this was adjusted to 60:40 in grades 10 and 11 for 2020 and 2021, significantly increasing the weight of SBAs in determining the final progression mark. The weighting of SBA and final examinations in grades 10 and 11 returned to the original 25:75 split only in 2022. The weighting in matric remained unchanged at 25:75 throughout the COVID-19 pandemic.

3 Methodology

The econometric analysis of this paper is predominantly descriptive and correlational. The approach to measuring the quality of SBA is based on the same criteria employed by Van der Berg and Shepherd (2015), namely (i) high assessment reliability (the SBA and matric examination marks were highly correlated) and (ii) low leniency (SBA marks are not much higher than matric examination marks). Assessment reliability will be referred to as assessment accuracy throughout this paper.

To investigate the accuracy of assessment, Spearman rank correlations will be calculated between the term 4 report marks and the NSC Mathematics results for grades 9 to 11, as well as between the term 3 report mark in matric and the NSC Mathematics results. All these correlations will be calculated at the school level for each cohort. The Spearman rank correlation is deemed appropriate for the analysis since Mathematics assessment results tend to be non-normally distributed and continuous. This correlation method is also robust to outliers. The correlations will be compared across cohorts as well as to specific thresholds, to identify schools that consistently assessed well across the sampled cohorts. A low correlation value suggests inadequate reliability in assessment, indicating that the continuous assessment marks from the school do not align well with matric examination results.

The second criterion for quality assessment, namely low leniency, will be determined by calculating the average difference between the final NSC Mathematics mark and term 4 (term 3 for matric) marks for the respective grades at a school level. Based on both leniency and accuracy, the quality of assessment will then be analysed to determine whether there is a significant difference between schools that produce a higher proportion of learners achieving above 60% in Mathematics in matric and those that produce fewer learners with a final NSC Mathematics mark above 60%. If a difference in assessment quality is identified, it could indicate that the quality of assessment is an important indicator of a functional school and, more importantly, contributes to better matric Mathematics results.

Three key periods for providing feedback to learners regarding their Mathematics performance are identified:

1. At the conclusion of grade 9,
2. During grades 10 and 11,
3. During matric, as learners prepare for the NSC examination.

The first period is crucial because it marks the end of the Senior Phase, when learners in South Africa have to make subject choices between Mathematics and Mathematical Literacy for the Further Education and Training (FET) phase. If the signals regarding a learner's mathematical abilities are unclear, this may lead to a high degree of randomness when selecting the subsequent mathematics subject, as noted by Taylor *et al.* (2011). Between grades 10 to 11, the learners can switch from Mathematics to Mathematical Literacy if their performance suggests that they may struggle to pass matric. Lastly, the performance of a learner during the preliminary examinations gives the learner as well as the teacher a good indication of how well the learner can expect to perform in the NSC examination. The preliminary examination thus serves as a final signal to learners regarding their level of preparedness for the NSC examination.

Ultimately the quality of assessment within a school will be used to identify outlier poor schools. Outlier poor schools are defined for this purpose as quintile 1 to 3 schools with high-quality assessments, characterised by high accuracy and low leniency, across more than one cohort. To identify outlier schools, threshold values for both measures of assessment quality need to be specified. The threshold values for assessment accuracy will be specific values for the correlations between the different report marks of the respective grades and the final NSC marks of learners within a school. Threshold values for assessment leniency will be based on the differences between the report marks and the final NSC mark for each grade. This paper focuses on identifying outlier schools among quintiles 1 to 3. Therefore, outlier schools should outperform other quintile 1 to 3 schools.

An appropriate threshold for assessment accuracy and leniency will be the averages for quintile 5 schools in Limpopo over the relevant years and cohorts. However, since the number of quintile 5 schools in Limpopo is fewer than 25, the average values can be significantly influenced by individual outlier schools. A more stable alternative is to use the averages from quintile 5 schools in Gauteng, which has a larger number of quintile 5 schools, making the average values less susceptible to outlier effects. Quintile 1 to 3 schools in Limpopo that show higher accuracy and lower leniency than the quintile 5 schools in Gauteng will be identified as outlier schools. The characteristics of these outlier schools will then be compared to those of all the other quintile 1 to 3 schools in Limpopo to determine whether improved assessment quality in earlier grades contributes to better performance among learners in the final NSC Mathematics results.

4 Data Description

4.1 Data Sources

SA-SAMS

The learner-level school-based assessment (SBA) data utilised in this analysis is sourced from the South African School Administration and Management System (SA-SAMS), which was collected for the Data Driven Districts programme. This initiative is a collaboration between the Department of Basic Education (DBE) and the Michael & Susan Dell Foundation. The SA-SAMS data contains the learner-level subject data for every learner in a public school in the Limpopo province. The SBA results in Mathematics for each learner in Limpopo have been recorded at the learner level for the years 2018 to 2022. Additionally, the dataset includes unique identifiers for each learner (an anonymised identification number), the emiscode of the school where the learner was registered for a specific year, as well as the term examination marks and term report marks of learners from grades 9 to 12. It also contains the home language and the gender of the learner.

Final NSC Results

The final NSC Mathematics results for the years 2018 to 2022 were provided by the Department of Basic Education. This dataset includes the same unique anonymised identifiers as the SA-SAMS data, the emiscode of the school in which the learner was registered for matric as well as the final NSC mark of the learner. The final NSC mark in Mathematics and Mathematical Literacy consist out of 75% the final matric examination and 25% the year mark (SBA mark) of the learner. Further details on the components of the year mark will be elaborated upon later in the paper.

South African Schools Masterlist

The final dataset utilised in the analysis is the 2021 Schools Masterlist of South Africa. This dataset includes all public and private schools, but for this paper, only public secondary schools were considered. The Masterlist includes school-level data such as the school quintile, the school district, whether the school is located in a rural or urban setting, and the number of teachers and learners at the school. In total, 1293 public secondary schools in Limpopo were included in the dataset.

Data Cleaning Process

The SA-SAMS and NSC datasets were merged and matched at a learner-level by utilising the unique anonymised identifier. These results enabled the tracking of learners of different cohorts from grade 9 through to matric. Two complete cohorts from grade 9 to matric and three complete cohorts from grade 10 to matric were identified. The three from grade 10 to matric cohorts were for the years 2017 to 2020, 2018 to 2021 and 2019 to 2022. These three cohorts were chronologically named Cohort 1, Cohort 2, and Cohort 3. The two grade 9 to matric cohorts stretched from 2018 to 2021 and 2019 to 2022 respectively. These two four-year cohorts formed part of Cohort 2 and Cohort 3.

The distinction between these cohorts is important due to significant changes in the weighting of year-end examinations and SBA results in determining final report marks before and after the COVID-19 pandemic (Hoadley, 2020). Only the final report marks for grades 9, 10 and 11 were utilised in the analysis, since these are the marks that serve as signals to learners regarding their performance in Mathematics. Within the matric year, the term 3 report mark, which reflects the year mark of the learner, serves as the most prominent signal to the learners regarding their preparedness for the NSC examination.

A balanced panel was created for cohorts 1 to 3 for grades 10 to 12. For the analysis of the grade 9 marks a separate balanced panel was generated since a large number of learners were lost when using the unique identifier to track learners from grade 9 to 12, largely because of learners repeating grade 9 or switching schools. The sample for the analysis was restricted to learners who progressed from grade 9 to matric without repeating a year and without switching between Mathematics and Mathematical Literacy. Furthermore, only learners who completed their first full-scale NSC attempt (writing six subjects or more in the NSC examination) were included in the sample.

A further limitation was imposed on the sample to include only learners from schools with at least 15 matriculants enrolled in Mathematics for the relevant cohort. This restriction was necessary since a small number of observations make measures such as correlations very volatile. For the final analysis in Section 6, only the data from schools in quintiles 1 to 3 was included since the emphasis is on poor schools in Limpopo.

4.2 Limpopo Secondary Schools Landscape

The following simple descriptive analysis aims to highlight the distinctive secondary school landscape of Limpopo. Table 1 presents the composition of Limpopo's secondary school system according to quintiles. In Limpopo, there are 1,257 public secondary schools in quintiles 1 to 3, compared to only 36 schools in quintiles 4 and 5. This stark disparity indicates that poorer schools are disproportionately represented in the province, with quintile 1 to 3 schools serving about 95% of the secondary school learners in Limpopo. In contrast, nationally, only 75% of learners are in quintile 1 to 3 schools in the secondary school phase (Department of Basic Education, 2020*b*). Interestingly, the average school size in the lower quintiles is smaller than that of the wealthier quintiles. Quintile 5 schools have an average of 821 learners, whereas quintile 1 schools have an average of 420 learners.

Unsurprisingly, only 2.7% of schools in Limpopo are reported to be fee-paying institutions. Another significant characteristic of Limpopo's secondary schools, as shown in Table 2, is that 92% are located in rural areas.

Number of Secondary Schools in Limpopo by quintile			
Quintile	**Number of Schools**	**%**	**Average School Size***
1	511	39.52	420
2	535	41.38	472
3	211	16.32	758
4	14	1.08	947
5	22	1.70	821
Total	**1293**	**100**	

*Calculated using the 2021 Schools Masterlist

Table 1: Number of Secondary Schools in Limpopo by quintile

School Location (Urban or Rural)				
Quintile	**Unspecified**	**Rural**	**Urban**	**Total**
1	2	500	9	511
2	2	510	23	535
3	13	155	43	211
4	5	5	4	14
5	1	16	5	22
Total	**23**	**1,186**	**84**	**1,293**

*Calculated using the 2021 Schools Masterlist

Table 2: Location of Limpopo Secondary Schools

The large proportion of secondary schools classified as quintile 1, 2 or 3 schools, combined with the rural location of these schools, poses unique and severe challenges for the Limpopo Department of Education. Despite these critical challenges, Limpopo still produces a large number of bachelors passes in Mathematics (Gustafsson, 2016). Another unique characteristic of Limpopo is that a large proportion of learners select to take Mathematics in the FET phase rather than Mathematical Literacy. The percentage of learners who wrote Mathematics during their first full-scale NSC attempt in the years 2018 to 2022 fluctuated between 41% to 50% of the matric cohort each year. This percentage is higher than that of provinces like Gauteng and the Western Cape, which have proportionately more quintile 4 and 5 schools.

In the balanced panel that includes all three cohorts, learners who repeated a year or switched mathematics subjects were excluded. Since stronger learners - who are less likely to fail a year or switch subjects – are more inclined to take Mathematics, Mathematics learners are overrepresented in the final sample. However, the analysis focus exclusively on Mathematics; this overrepresentation of Mathematics learners will thus not bias the results.

5 Results

5.1 Analysis of Matric Term 3 Marks and NSC Results

The analysis of matric-specific data is conducted separately to examine differences in the relationship between SBA marks and NSC marks before and after the Covid-19 pandemic. Only the first term 3 report mark as well as the marks from the first full NSC examination attempt of students were considered. The term 3 report marks reflect the year marks for learners, which account for 25% of the final NSC Mathematics marks.

Prior to 2020, the year mark for Mathematics consisted of seven SBA tasks, while from 2020 onwards, this was reduced to six SBA tasks. The components of the year marks and their respective weightings during 2020 and 2021 were as follows: Term 1 included an assignment (10%), an investigation (20%), and a test (15%); Term 2 included a test (15%); Term 3 includes a test (15%) and the preliminary examination (25%). Prior to 2020, a full-scale June examination contributed 15% to the year mark, while individual tests each contributed only 10% (Department of Basic Education, 2011).

In 2020, the June examinations for all grades were substituted with tests due to the disruptions caused by the COVID-19 pandemic. As a result, the matriculants in 2020 wrote only one round of examinations – the preliminary examinations – before writing the NSC examinations. Full-scale June Mathematics examinations in matric remained optional until 2023 (Department of Basic Education, 2023). The number of learners for whom term 3 report marks were recorded in 2020 is significantly lower compared to the preceding and following years in the sample, as illustrated in Table 3.
Table 3 presents the sample sizes for the term 3 report marks, the final NSC marks, and the final sample size by year (2018–2022) for which the term 3 report marks and NSC marks could be linked using the unique individual identifier code in the dataset. It is evident that prior to 2020, a larger group of students completed the

Number of Matric Observations by Year					
	2018	**2019**	**2020**	**2021**	**2022**
Term 3 Report Mark	82 103	75 186	38 378	90 316	92 321
Final NSC Result	61 430	57 895	66 465	90 907	93 661
Final Sample	**54 214**	**52 994**	**37 255**	**85 616**	**90 546**

Table 3: Number of Matric Observations by Year

preliminary examinations and obtained a term 3 report mark, but did not write a complete NSC final examination. These discrepancies in the number of learners can be explained by the short-lived Multiple Examination Opportunities (MEO) policy implemented in 2016 and discontinued in 2019 (Wills *et al.*, 2024, p. 20). The years 2018 and 2019 are the only years in the sample that coincide with the period during which the MEO policy was in effect. Under this policy, learners were allowed to split their NSC examinations between the December examination session and the subsequent June session of the following year. In 2020, many schools did not report their term 3 report marks due to school closures and high levels of uncertainty regarding assessment policy caused by the COVID-19 pandemic. The lack of term 3 report marks in 2020 resulted in a much smaller sample for which the term 3 and NSC marks could be linked, despite having a larger group that wrote NSC in 2020 compared to the two preceding years.

The sharp increase in the number of first-time NSC candidates in 2021 (relative to 2018, 2019, and 2020), indicates that there might have been a large number of students who decided in term 4 of 2020 to postpone the writing of the NSC examinations to 2021. The increase in the term 3 report mark, as well as NSC candidate sample sizes, might be attributed to more lenient progression requirements implemented as a response to the global pandemic (Hoadley, 2020).
Despite significant differences in final sample sizes across years, the sample composition remains representative of the province in terms of school quintiles and the urban-rural distribution. As a result, sample selection does not pose a concern for bias in interpreting the results.

Table 4 reports the average term 3 report marks and final NSC results at the learner level across different quintiles for the respective years. The data highlights that learners in quintile 4 and 5 schools consistently outperform those in poor schools. Over the observed years, quintile 5 learners, on average, achieved term 3 marks

Averages of Matric Mathematics Assessment Marks at a Learner Level by Quintile												
	Term 3 Marks						NSC Final Marks					
Quintile	**2018**	**2019**	**2020**	**2021**	**2022**	**Average**	**2018**	**2019**	**2020**	**2021**	**2022**	**Average**
1	26.92	29.86	18.85	23.58	25.19	**24.88**	30.40	29.33	27.54	29.25	28.72	**29.05**
	(0.20)	(0.22)	(0.21)	(0.17)	(0.17)		(0.20)	(0.21)	(0.24)	(0.18)	(0.18)	
2	28.61	32.02	20.71	25.61	26.57	**26.70**	32.72	32.58	30.36	32.13	30.51	**31.66**
	(0.19)	(0.19)	(0.20)	(0.15)	(0.15)		(0.18)	(0.18)	(0.22)	(0.17)	(0.16)	
3	29.75	34.55	22.28	26.70	27.45	**28.15**	34.50	34.61	32.72	33.29	31.53	**33.33**
	(0.22)	(0.25)	(0.28)	(0.20)	(0.19)		(0.22)	(0.24)	(0.30)	(0.21)	(0.20)	
4	35.51	40.99	25.00	31.44	32.88	**33.17**	41.52	41.85	38.24	41.98	38.50	**40.42**
	(0.51)	(0.56)	(0.93)	(0.48)	(0.50)		(0.51)	(0.52)	(0.98)	(0.50)	(0.52)	
5	46.37	47.41	41.95	44.19	45.33	**45.05**	48.59	44.67	46.50	47.67	45.21	**46.53**
	(0.57)	(0.59)	(0.65)	(0.59)	(0.52)		(0.50)	(0.51)	(0.60)	(0.55)	(0.49)	
Year Average	**33.43**	**36.97**	**25.76**	**30.31**	**31.48**		**37.55**	**36.61**	**35.07**	**36.86**	**34.89**	

*Standard Errors are included in brackets

Table 4: Averages of Matric Mathematics Assessment Marks at a Learner Level by Quintile

20% higher and NSC marks 17% higher than quintile 1 learners. Additionally, the average performance gap between the bottom three quintiles is approximately 4% for both term 3 marks and NSC results. These findings underscore the substantial disparity in academic achievement between the lower and upper quintiles.

In 2020, learners across all quintiles performed worse than in 2019 in the preliminary examinations. However, this trend did not persist in the NSC examinations, where quintile 5 learners performed better than in 2019. The overall lower average in the NSC results for 2020 was driven by poorer performance among learners in the bottom four quintiles, consistent with the findings of Soudien *et al.* (2022). Notably, term 3 marks across all quintiles have yet to return to the levels observed before 2020, while NSC examination marks surpassed the 2019 average in 2021 before falling to below the 2019 average in 2022.

Figure 1 illustrates the gaps between the average final NSC results and the average term 3 report marks in Mathematics for the respective quintiles from 2018 to 2022. These gaps were calculated at the learner level and averaged by quintile. Across all cohorts, a decline in the gap was observed in 2019, and learners in quintiles 1 and 5 even exhibited negative gaps, meaning they performed better in their term 3 assessments than in the final NSC examinations during that year. In 2020, all quintiles experienced a sharp increase in the gap, as term 3 marks were lower across the board. The subsequent decline in the gap during 2021 and 2022 was largely driven by improvements in learners' term 3 marks. Quintile 4 schools were excluded from the graph due to their small sample size, which makes their mean results highly susceptible to outliers.

5.2 Measuring Assessment Accuracy

5.2.1 School-Level Mean Report Marks and NSC Examination Marks for Mathematics

Scatter plots and Lowess (locally weighted regression) curves illustrate the relationship between the signals received by the learners in the lower grades and preliminary examinations regarding their performance in Mathematics. To provide additional context, a diagonal line is included in each graph, indicating where the values on the two axes are equal.

Grade 9 Report Marks and NSC Performance

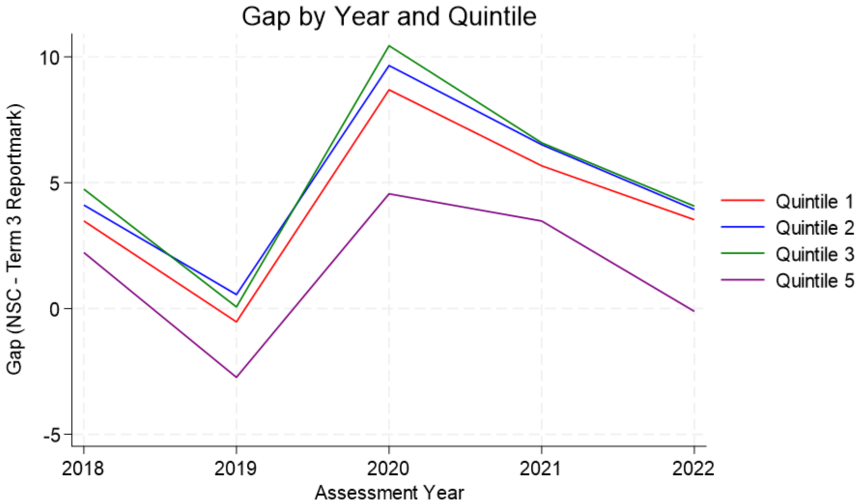


Figure 1: Gaps in Matric Mathematics Marks by Year and Quintile

Correlations between grade 9 and NSC marks were calculated for two grade 9 to grade 12 cohorts, spanning 2018 to 2021 and 2019 to 2022, respectively. These four-year cohorts includes Cohort 2 and Cohort 3. The number of learners tracked over four years was significantly smaller than those tracked from grade 10 to matric. This reduction is primarily due to learners either changing schools between grade 9 and grade 10 or repeating grade 9 or grade 10.

It is evident from Figure 2 that the number of schools that achieved a mean grade 9 report mark below 30% or above 70% is very low. The sparseness of data in these regions of the graph leads to outlier schools heavily influencing the Lowess curve for these values. Nevertheless, the trend of the Lowess curve is upward, indicating a positive relationship between the mean grade 9 report mark and the mean NSC examination mark in Mathematics for a school. Most schools lie below the diagonal, suggesting that grade 9 report marks may give learners an overly optimistic perception of their mathematical abilities. This misperception may lead learners to choose Mathematics over Mathematical Literacy in the FET phase. Such leniency in assessment could contribute to the observed “randomness” in the selection of mathematics subjects for grade 10 in Limpopo (Gustafsson, 2016).

The pattern between the mean grade 9 report marks and mean NSC examination percentage is consistent across the second and third cohorts. The Lowess curve

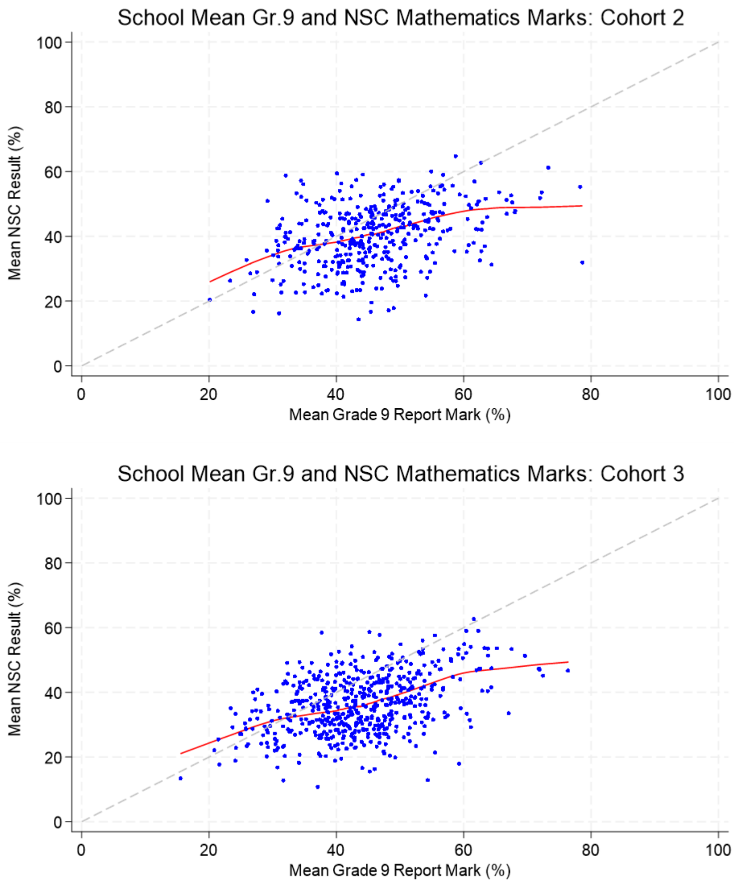


Figure 2: Grade 9 Report Marks and NSC Results Correlations

maintains an upward trajectory for all mean grade 9 report mark values, indicating a stable positive relationship between these metrics.

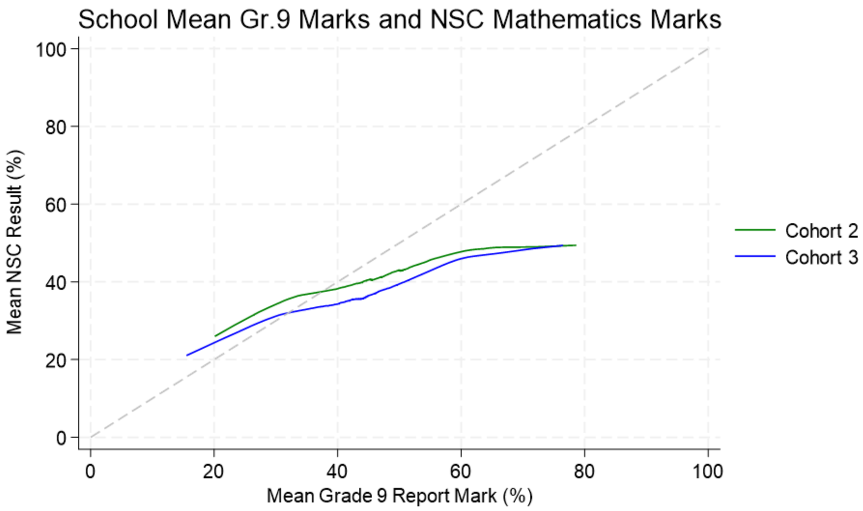


Figure 3: Lowess Curves of Grade 9 Correlations

The similar shape of the two Lowess, as shown in Figure 3, illustrates a positive relationship between the mean grade 9 report marks and the mean NSC marks in Mathematics for schools. Both curves also reveal evidence of assessment leniency in schools where the average performance of learners in grade 9 Mathematics is above 40%.

Grade 10 Report Marks and NSC Performance

The correlations between the grade 10 report marks and NSC results could be calculated for all three three-year cohorts as set out in Section 4.1. The mean NSC results and mean grade 10 report marks of schools with their corresponding Lowess curves are illustrated in Figure 4 for the respective cohorts. Schools with weaker performing learners who select to take Mathematics in grade 10 and continue with the subject till matric have a lower average performance in grade 10 than in the final NSC examination. This is illustrated by the Lowess curves that all lie above the diagonal line until just above 35%. A positive relationship between the mean grade 10 report mark of the learners in a school and the school's mean NSC examination is illustrated by the upward trajectory of the Lowess curves in Cohort 1 and 2. For the second cohort, the Lowess curve lies predominantly above the

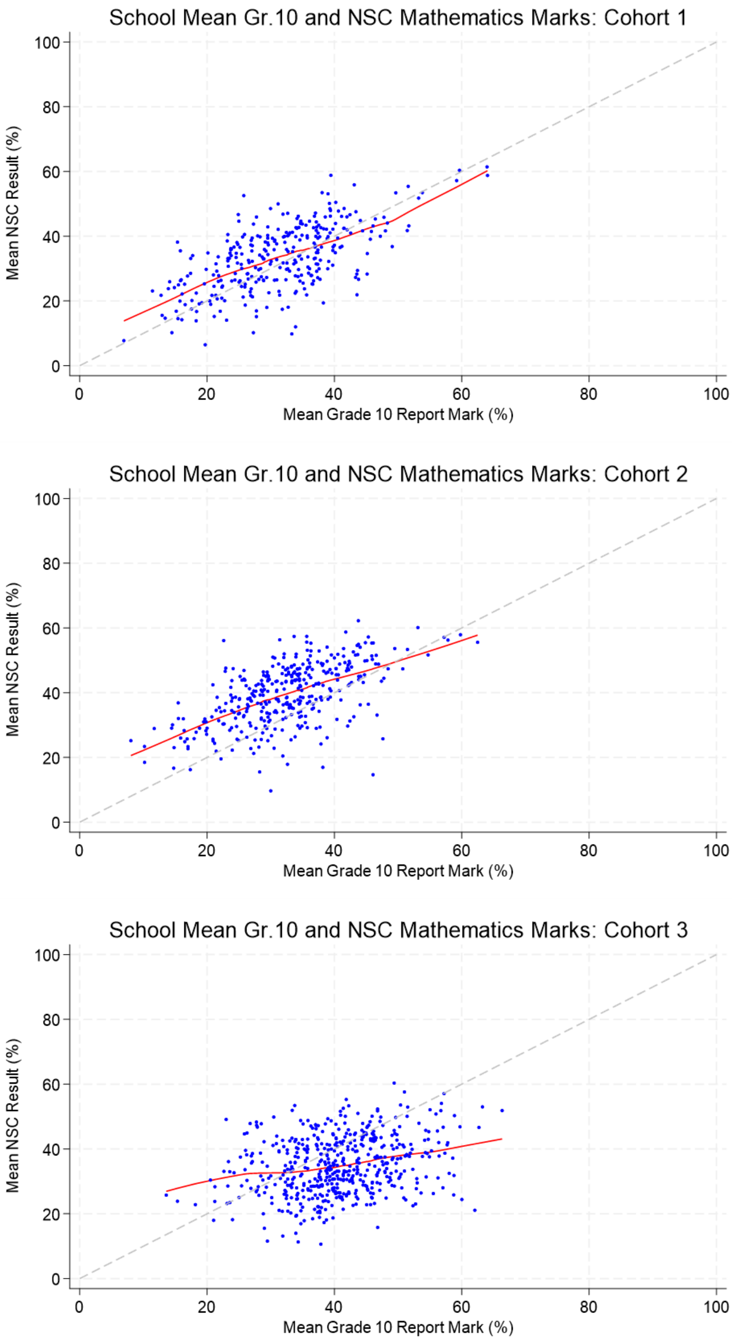


Figure 4: Grade 10 Report Marks and NSC Results Correlations

diagonal. This positioning of the Lowess curve indicates that for this cohort most schools achieved higher average marks in the final NSC results compared to the final grade 10 report marks. This trend could reflect that matric learners have successfully adjusted to the subject's demands, or that teachers adopted a more long-term approach in preparing learners for the matric examination.

The Lowess curve for the third cohort is notably flatter compared to that of the previous cohorts. Additionally, most of the observations for this cohort lie below the diagonal, indicating that, on average, learners from these schools performed better in the final grade 10 report marks than in the final NSC results. This discrepancy may stem from the change in the weighting of the year mark to end-of-year examinations, which shifted from 25:75 to 60:40 in 2020 - the year in which the learners in this cohort were in grade 10 (Department of Basic Education, 2020*a*).

Furthermore, Cohort 3 exhibits significantly more variance in the relationship between the mean NSC mark and the mean grade 10 mark of a school. This variation may result from some schools opting to administer full-scale end-of-year examinations in 2020, despite the official national policy recommending only tests (Department of Basic Education, 2020*a*). The broader distribution of school-level observations may also reflect disparities in schools' ability to maintain continuity in teaching and learning during 2020 (Soudien *et al.*, 2022).

A comparison of the Lowess curves in Figure 5 reveals that the strongest correlation between the mean term 3 mark and mean NSC result of schools was observed in Cohort 1, as this Lowess curve lies closest to the diagonal. A notable shift occurred between Cohort 2 and Cohort 3. In Cohort 2, schools generally achieved higher mean NSC results than mean grade 10 report marks. Conversely, schools in Cohort 3 with a mean grade 10 report mark above 30% predominantly achieved worse mean NSC results than mean grade 10 report marks. This trend supports the argument that policy changes implemented in 2020 in response to the COVID-19 pandemic weakened the relationship between the mean grade 10 report marks and mean NSC marks at the school level.

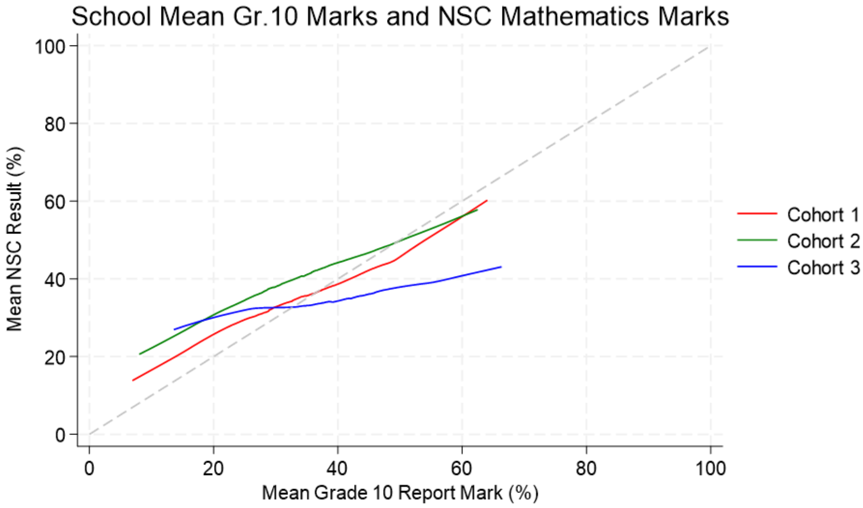


Figure 5: Lowess Curves of Grade 10 Correlations

Grade 11 Report Marks and NSC Performance

Figure 6 illustrates the relationship between the mean NSC results and mean grade 11 results for the respective cohorts. For the first cohort, the average grade 11 examination marks of a school are lower than the average performance of the schools' learners in the NSC examinations. This pattern mirrors the trend observed in Figure 4 for the grade 10 marks of Cohort 1. The Lowess curve lies above or on top of the diagonal across all values of the mean Grade 11 report mark, underscoring this relationship.

However, this picture changes for the second cohort, where the learners reached grade 11 in 2020. The significant difference can be attributed to the same factors discussed for learners who reached grade 10 in 2020. The change in assessment weights led to the mean final grade 11 report marks being lower than the mean NSC marks for schools with weaker-performing learners and higher than the mean NSC mark for schools with stronger learners. The less tight fit of observations to the Lowess curve reflects the pattern seen in the grade 10 group of Cohort 3. This greater variance may stem from the inconsistencies in how schools adhered to the education policies implemented in response to the COVID-19 pandemic in 2020.

The last cohort shows a clear shift away from the first cohort where most learners in schools achieved higher NSC examination marks than their average grade 11

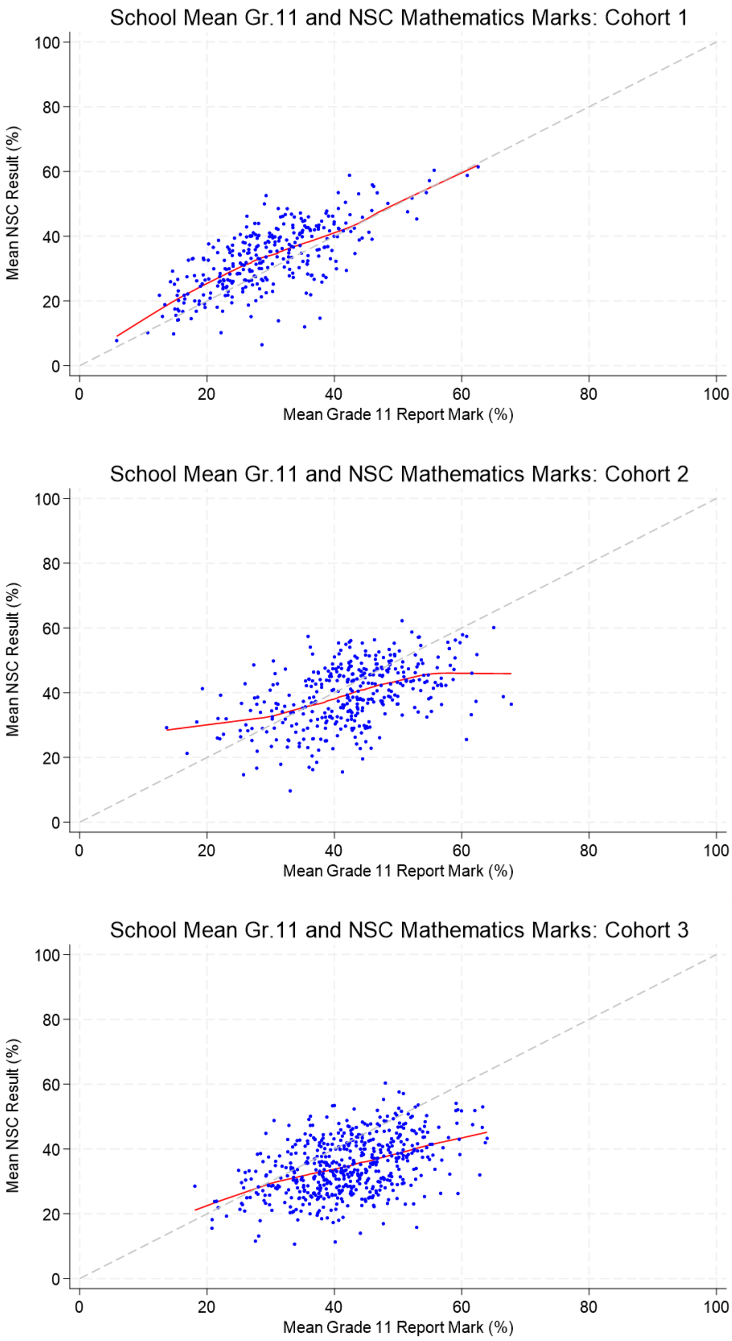


Figure 6: Grade 11 Report Marks and NSC Results Correlations

report marks. In contrast, schools with stronger learners in the third cohort, on average, performed worse in the final NSC results than in grade 11.

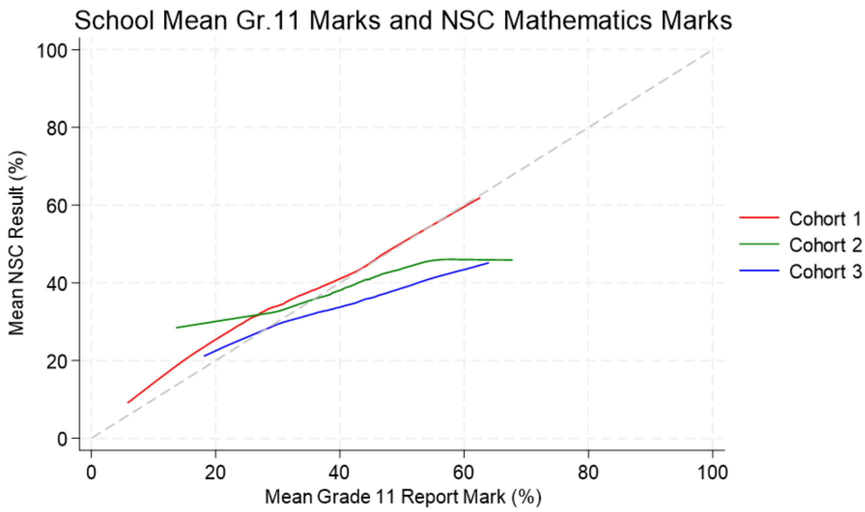


Figure 7: Lowess Curves of Grade 11 Correlations

The comparison between the three different Lowess curves in Figure 7 further underscores the shift in the relationship between the mean grade 11 marks and mean NSC marks of schools from Cohort 1 to the other two cohorts, in which grade 11 marks were affected by COVID-19 policies. In Cohort 1, the Lowess curve lies above or on top of the diagonal, while for both Cohort 2 and Cohort 3, the Lowess curve lies below the diagonal in the more data-dense regions.

Matric Term 3 Marks and NSC Performance

The relationship between the mean NSC results and mean grade 12 term 3 results for the respective cohorts is illustrated in Figure 8. Learners in the first cohort reached matric in 2020. Due to COVID-19 disruptions to the schooling system, fewer schools reported term 3 matric marks for this year. The graph for 2020 is thus less representative of all the schools in Limpopo than that of 2021 and 2022. The gap between the average term 3 marks and the average NSC marks is larger for the first cohort than for the subsequent cohorts, as illustrated by the Lowess curve lying further to the left of the diagonal line. This confirms what was illustrated in Figure 1 in Section 5. For the following cohorts, the same positive relationship between the average term 3 marks and the average NSC marks is observed. In all cohorts, the average term 3 report marks are lower than the average NSC exam-

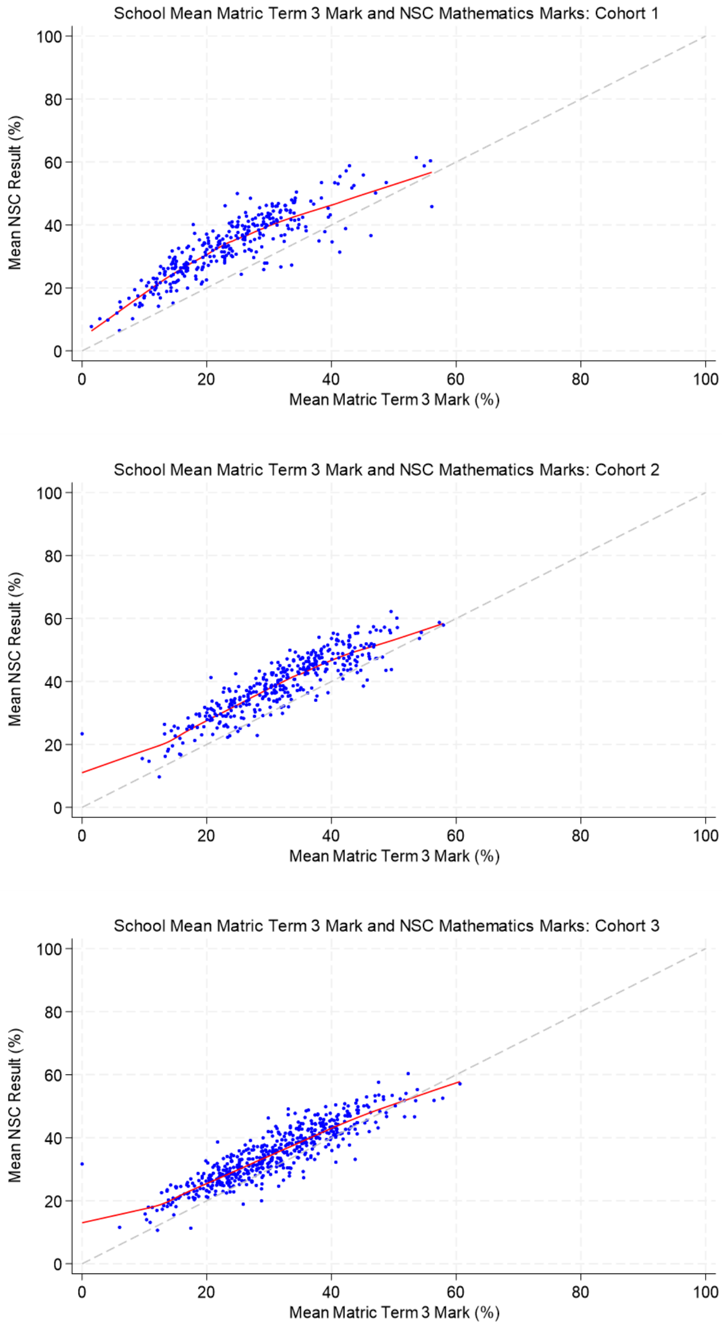


Figure 8: Matric Term 3 Report Marks and NSC Results Correlations

ination marks. This is plausible given that learners write the NSC examinations on the same content they prepared for in the preliminary examinations, which contributes to the term 3 report mark.

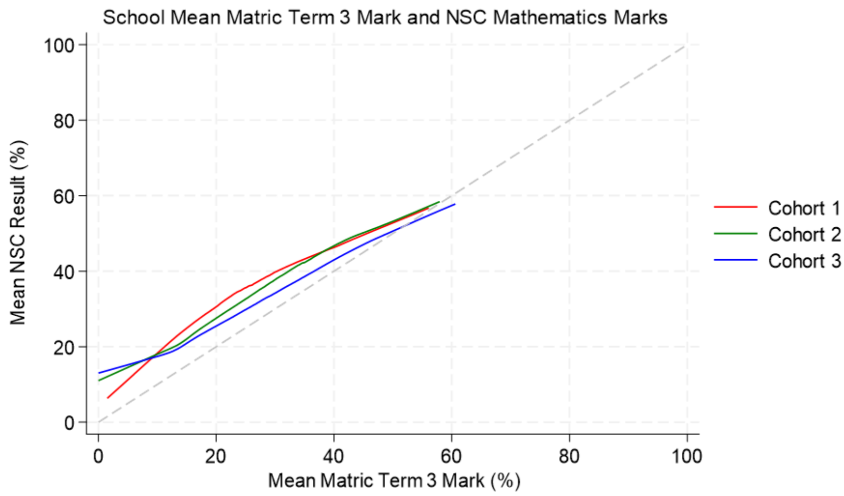


Figure 9: Lowess Curves of Matric Correlations

As illustrated in Figure 9, all three Lowess curves lie close to the diagonal, indicating that the mean term 3 mark of matriculants is a strong predictor of their final NSC results in Mathematics. The more robust positive relationship between the mean term 3 mark and the mean NSC result across all three cohorts is unsurprising, since learners are being assessed on largely the same work in the NSC examination as in the preliminary examination, which makes up a large part of the term 3 mark. However, the final matric mark can still deviate a lot from the term 3 mark if the mark obtained by a learner in the NSC examination differed significantly from the term 3 matric mark of the learner. If there is a more than 15% discrepancy between the term 3 report mark (year SBA mark) and the NSC mark obtained by a learner in the NSC examination, the SBA mark is omitted when calculating the final NSC result (Hoadley, 2020, p. 15). The Lowess curves for cohorts 1 and 2 lie above the Lowess curve for Cohort 3 at all values of the mean term 3 mark. The smaller distance between the Lowess curve of Cohort three and the diagonal line suggests that the improvement in average marks from the term 3 report mark to the final NSC result was less pronounced for Cohort 3 compared to the earlier cohorts.

5.2.2 Spearman Correlations by Mean NSC Mathematics Result

To illustrate what the Spearman correlation represents, two schools with a high and low Spearman correlation are plotted individually. The range of Spearman correlations spans from -1 to 1. The diagonal lines in the graphs represent perfect correlation (a value of 1) where the learners in the schools obtained the same final mark at the end of matric as at the end of grade 10. The Spearman correlation values used for illustration are the grade 10 final report marks and NSC marks in Cohort 1.

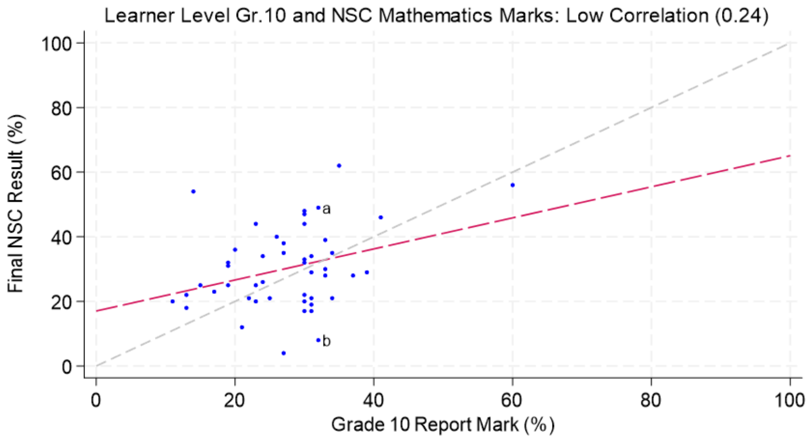


Figure 10: Illustration of Low Correlation Within a School

Figure 10 displays learners' marks in a school with a low correlation value of 0.24. The scatter plot shows that while there are a few observations close to the diagonal, the line that best fits the data is flatter, indicating a weaker positive relationship between the final NSC result and the grade 10 report marks of learners. Two learners, "a" and "b", are labelled in the plot. Both learner "a" and learner "b" achieved a mark of 32% at the end of grade 10, but learner "a" achieved a NSC pass mark of 49%, while learner "b" failed with a final NSC mark of 8%.

Figure 11 represents learners in a school with a high correlation value of 0.84. Most of the data points are clustered close to the diagonal line, indicating that the grade 10 final marks obtained by the learners were a strong predictor of the final NSC results they would achieve. The red dashed line, which is fitted to the

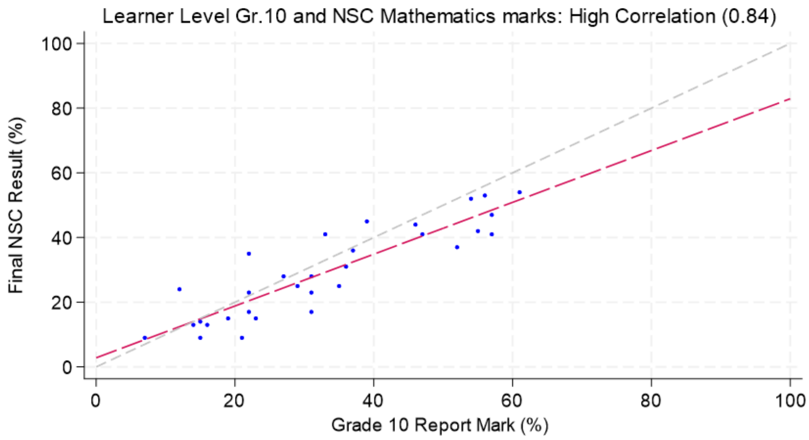


Figure 11: Illustration of High Correlation Within a School

scatter plot, closely resembles the diagonal, further illustrating the strong positive relationship between the grade 10 report marks and final NSC results obtained by the learners at this school. The Spearman correlations - illustrated at an individual level in Figures 10 and 11 – will now be investigated at the school level across different grades, to determine whether stronger correlations are associated with better average performance in the NSC within schools.

Grade 9 Spearman Correlations

Figure 12 illustrates that for both cohorts, the Spearman correlations at a school level between the mean grade 9 report mark and the mean NSC mark do not reveal a clear trend regarding the mean NSC result. Some schools exhibit a negative Spearman correlation, indicating extremely weak assessment accuracy. A negative correlation suggests that schools with higher grade 9 marks in Mathematics are associated with lower NSC marks in Mathematics and vice versa. Quintile 4 and 5 schools generally have Spearman correlation values above the Lowess curve, indicating that they show stronger correlation values at the same levels of mean NSC results relative to the quintile 1 to 3 schools.

Although the Spearman correlation values start from a very low point, they appear to increase for schools with higher mean NSC results. The Lowess curves in Figure 13 suggest that, for schools in Cohort 3, Spearman correlation values between the

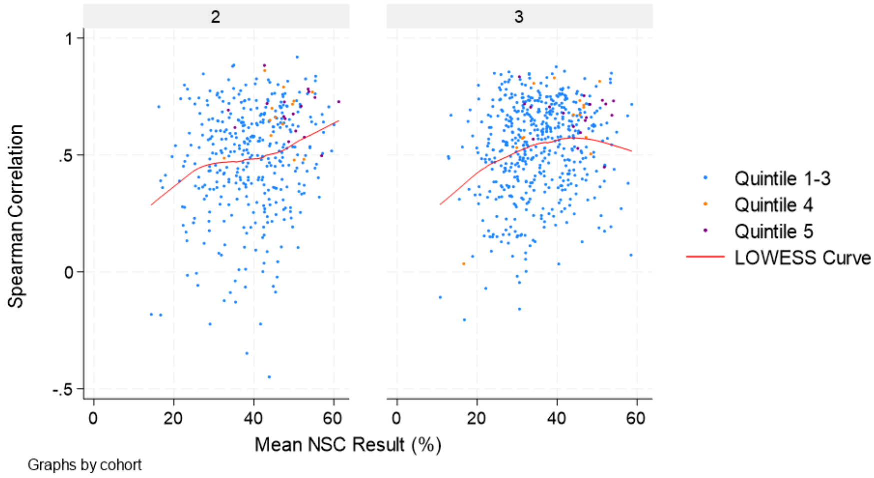


Figure 12: Spearman Correlation for Grade 9 Report Mark and NSC Result by Cohort

mean grade 9 report marks and mean NSC results are notably higher for schools with mean NSC results below 50%, when compared to schools in Cohort 2.

Grade 10 Spearman Correlations

As illustrated in Figure 14, for all three cohorts, the Spearman correlation between the average grade 10 report mark and the average NSC mark of the school seems to be larger for schools with higher average NSC marks. However, some schools had negative Spearman correlations, indicating extremely weak assessment accuracy. The less tight fit for Cohort 3 could be due to the fact that this cohort was in grade 10 in 2020, when the weightings of the SBA and end-of-year examinations in the calculation of the final year mark of learners were 60:40, as discussed in depth in Section 2 of the paper. Quintile 4 and 5 schools seem to largely achieve average NSC results above 30% (the pass requirement for Mathematics), while there are a few quintile 1 to 3 schools with low Spearman correlations achieving below this threshold. Interestingly, there seem to be a few quintile 1 to 3 schools that achieved average NSC results above 40%, even though their Spearman correlations between the average grade 10 report mark and the average NSC mark are below the value of 0.5, reflecting poor assessment accuracy.

Figure 15 shows that, similar to the Lowess curves for grade 9, there is a strong positive relationship between the Spearman correlation for the grade 10 report

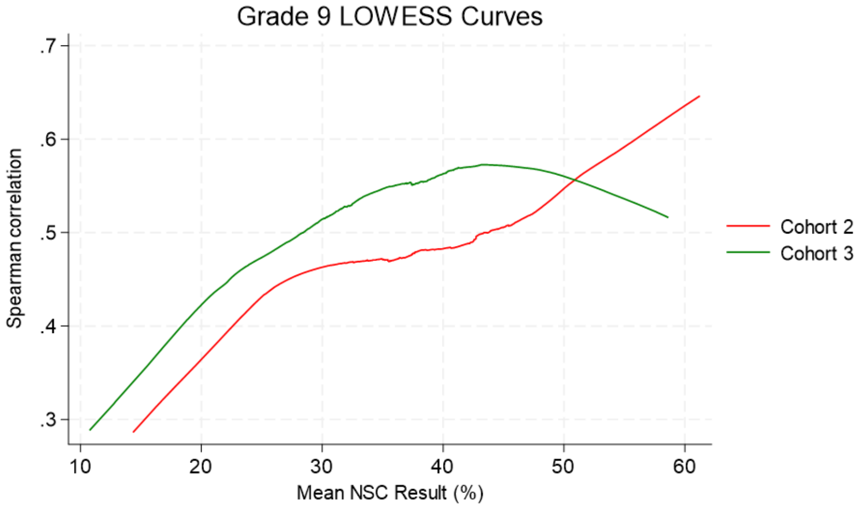


Figure 13: Grade 9 Lowess Curves of Spearman Correlations

mark and final NSC results of a school and the mean NSC result achieved by the learners in the school. However, for grade 10, Cohort 3 achieved higher mean NSC results than Cohort 2 for the same Spearman correlation value over the entire range of correlation values. This implies that the accuracy of grade 10 assessments for Cohort 3 may have played a smaller role in influencing learners' success in the final NSC examination compared to its impact on the earlier cohorts.

Grade 11 Spearman Correlations

From Figure 16, it is evident that for all three cohorts the Spearman correlations of schools are higher in grade 11 compared to the lower grades. Consistent with the trends observed in earlier grades, the Lowess curves show an upward trajectory, suggesting that the average NSC results of schools and the average grade 11 marks for schools are more correlated for schools that achieve a higher average NSC result in Mathematics. This may indicate that more accurate assessment in grade 11 contributes to better performance in the NSC examinations. All quintile 5 schools across all three cohorts had a Spearman correlation value of at least 0.6, while several quintile 1 to 3 schools exhibited correlation values below 0.6, indicating poor assessment accuracy at the grade 11 level in these schools. Cohorts 2 and 3 were both affected by COVID-19 assessment policies, which could have led to the less tight fit around the Lowess curve compared to Cohort 1.

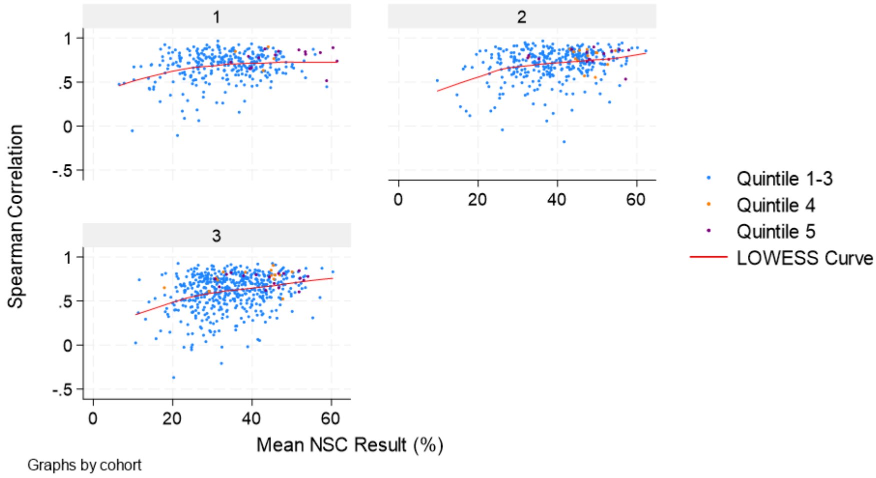


Figure 14: Spearman Correlation for Grade 10 Report Mark and NSC Result by Cohort

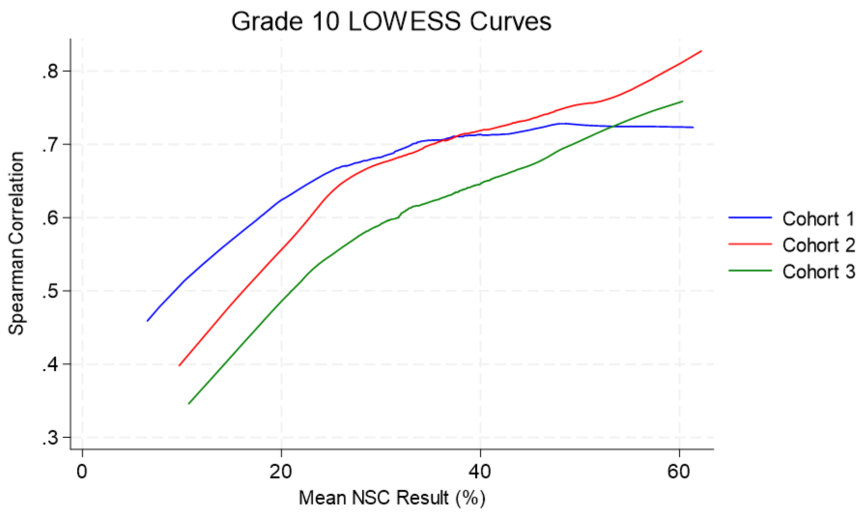


Figure 15: Grade 10 Lowess Curves of Spearman Correlations

Figure 17 presents the Lowess curves for grade 11 across all three cohorts. The Lowess curves for Cohorts 2 and 3 are closely aligned, with both lying below the curve for Cohort 1.

Matric Term 3 Spearman Correlations

As expected, the Spearman correlations between the average mark in term 3 of

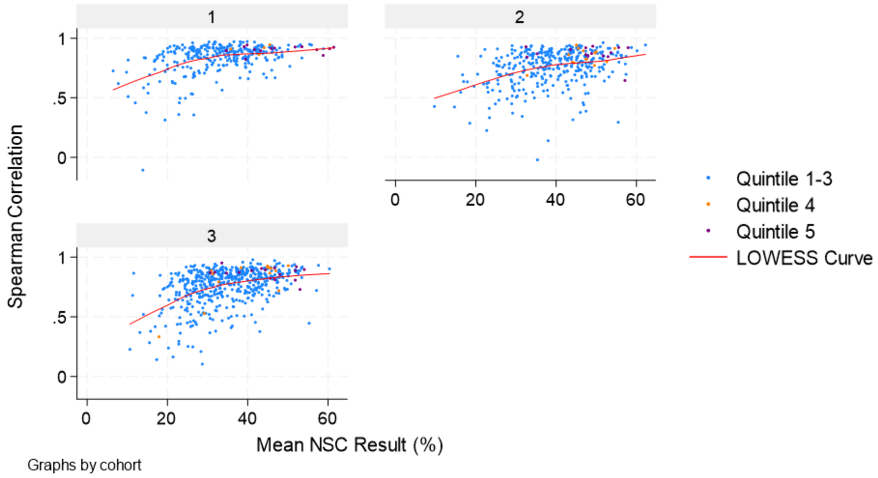


Figure 16: Spearman Correlation for Grade 11 Report Mark and NSC Result by Cohort

matric and the average mark in the NSC examinations are higher than those of lower grades and the corresponding NSC results. Since some matric assessments, such as the preliminary examination papers and memorandums, are externally set, higher Spearman correlation values are unsurprising. Figure 18 illustrate that, for all three cohorts, there is an initial increase in the Lowess line at lower mean NSC results after which it plateaus at correlation values above 0.9. This shape of the Lowess curve indicates that schools with lower mean NSC results tend to have less accurate assessments earlier in matric. What is particularly interesting, however, is the significant improvement in the correlation values from grade 9 to matric.

Figure 19 represents the Lowess curves for the matric term 3 marks over all three cohorts. The Lowess curve for Cohort 1 is strongly influenced by outliers, particularly schools with mean NSC results below 20%. For all three cohorts, the Spearman correlations between the matric term 3 results and the final NSC results stabilise at levels around 0.95, as the mean NSC result increases to above 40%.

5.2.3 Spearman Correlations by the Percentage of the School's Matric Mathematics Cohort Achieving Above 60%

For this part of the analysis on assessment accuracy, only the results from Cohort 1 are included. This is justified as this cohort was the least affected by the assess-

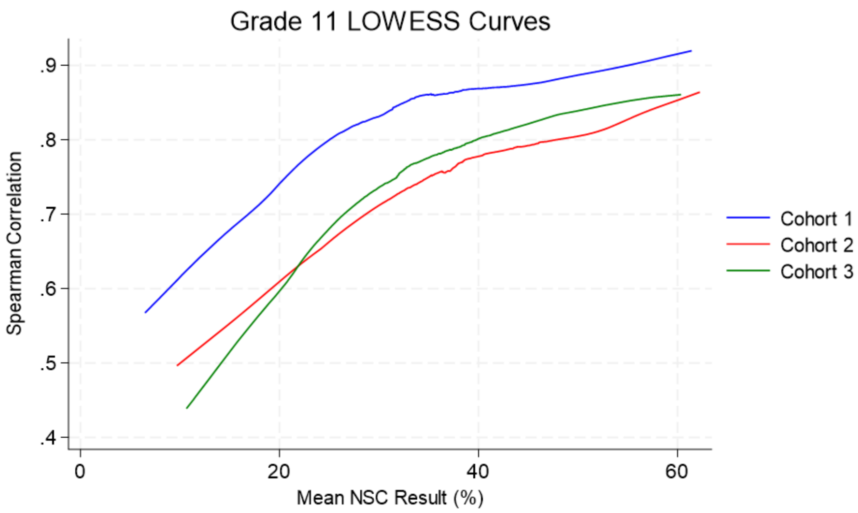


Figure 17: Grade 11 Lowess Curves of Spearman Correlations

ment policy changes implemented in response to the COVID-19 pandemic. Scatter plots and Lowess curves illustrate the relationship between assessment accuracy (measured by the Spearman correlation value) and the percentage of Mathematics learners in a school who achieve a mark higher than 60%. Only grade 10 to 12 results are considered due to the weak relationship between grade 9 report marks and NSC results. A large number of schools had no learners who achieved above 60% in Mathematics in matric. These schools have a large range of correlation values for all three respective grades, as seen on the left-hand side of the respective graphs. There are nine quintile 1 to 3 schools in which more than 30% of Mathematics learners achieved at least 60%.

Figure 20 includes the graphs for all three grades of Cohort 1. For grade 10, in the data-dense regions, there appears to be a positive relationship between the Spearman correlation and the percentage of the cohort of a school achieving above 60% in the final NSC result. However, the number of schools where more than 30% of learners achieve above 60% in Mathematics is limited, thus the Lowess curve in this section is heavily influenced by outliers, causing it to bend downwards.

For grade 11, the Lowess curve consistently increases across all values of mean performance in matric, indicating a positive relationship between the Spearman correlation and the percentage of Mathematics learners achieving above 60%. The

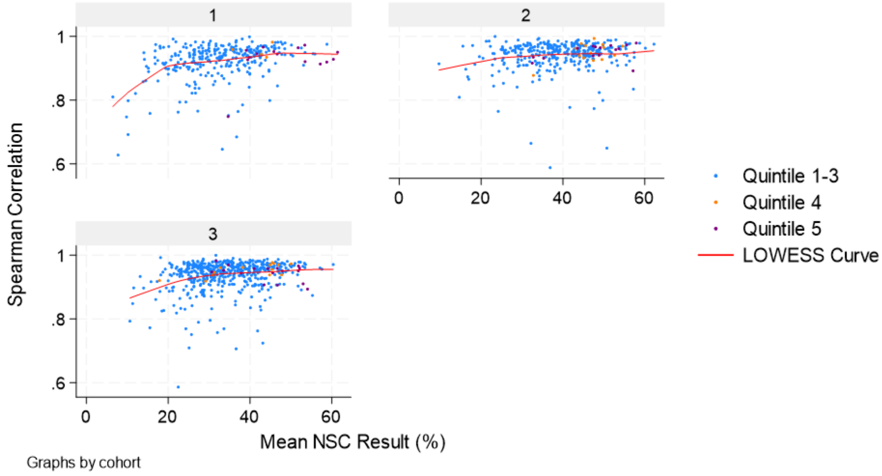


Figure 18: Spearman Correlation for Grade 12 Term 3 Report Mark and NSC Result by Cohort

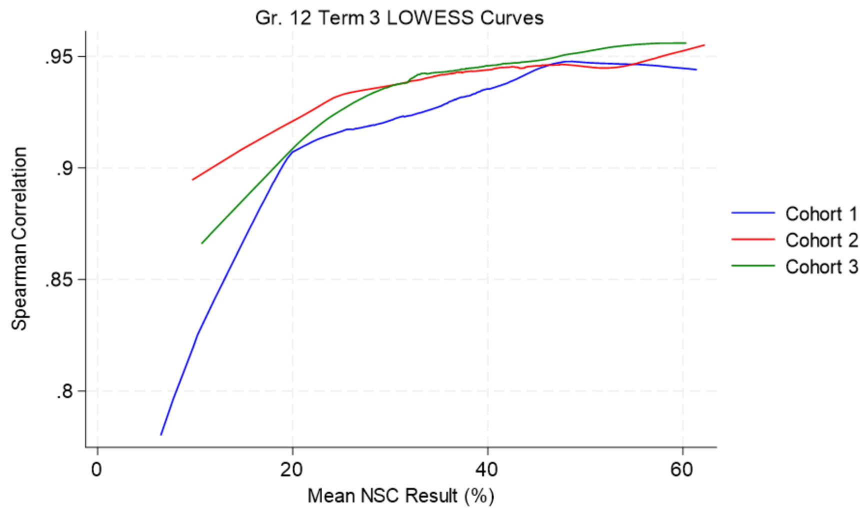
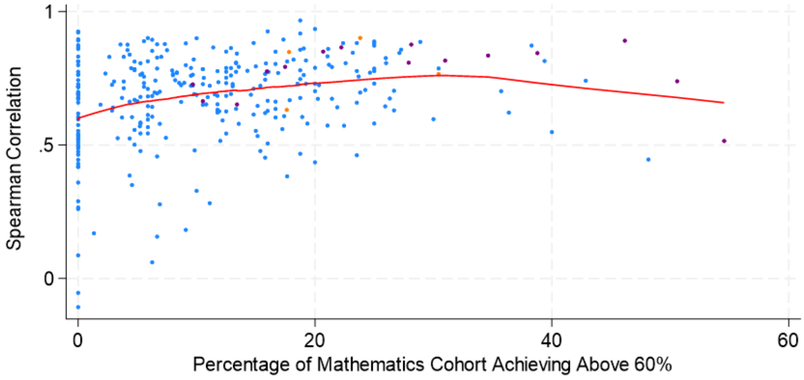


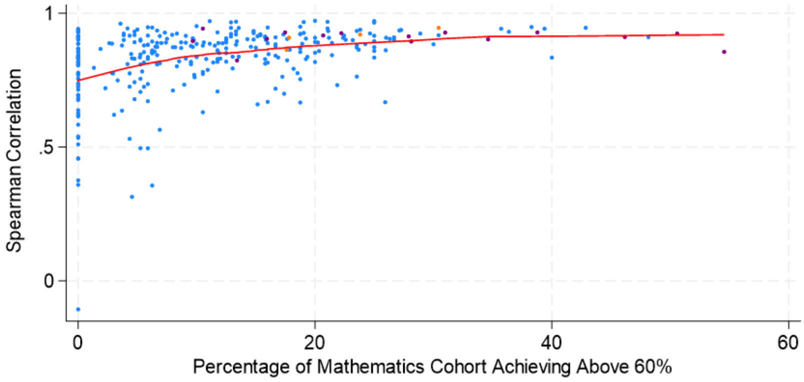
Figure 19: Grade 12 Lowess Curves of Spearman Correlations

observations in this graph are more concentrated around the Lowess curve compared to the grade 10 results.

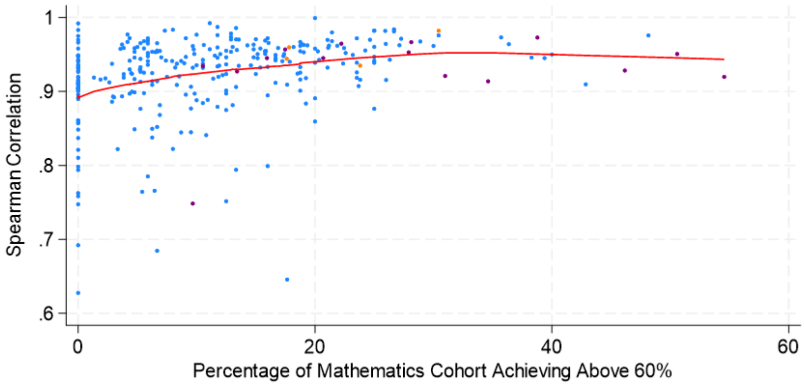
Figure 20 shows much larger correlation values for matric term 3 marks than for the preceding grades' report marks. Similar to the pattern observed for grade 11 marks, there is a continuous positive relationship between the Spearman correla-



(a) Grade 10



(b) Grade 11



(c) Grade 12 Term 3

Figure 20: Spearman Correlations and Percentage of Mathematics Cohort achieving above 60% in the NSC

tion and the percentage of Mathematics learners in a school achieving above 60%. For matric, all schools with at least 30% of learners achieving above 60% in the NSC results have a Spearman correlation value between 0.9 and 1.

In summary, across all grades, schools with a large percentage of learners achieving above 60% in the NSC examination tend to exhibit larger Spearman correlations. Conversely, schools where no learners achieve above 60% in the NSC results show a wide range of Spearman correlations, from negative values to those approaching one.

5.3 Measuring Assessment Leniency

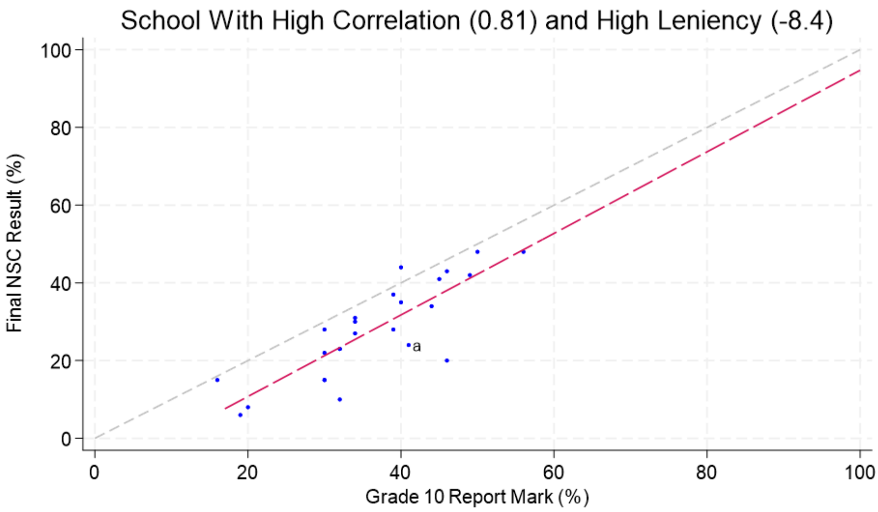


Figure 21: Illustrating High Leniency Within a School

Leniency in assessment is defined as the situation where the SBA marks of learners (in this case, the report marks) are much higher than the examination marks (in this case, the final NSC marks) (Van der Berg and Shepherd, 2015). This additional criterion for assessment quality is crucial, as schools can have large correlations between the report marks of previous years and the final NSC results of its learners, yet still exhibit lenient assessments. Figure 21 illustrates this, as the school for which the learners' marks are plotted shows a strong correlation of 0.81, but also a large negative average gap of -8.4% between the final NSC result and the grade 10 report mark. The strong positive correlation is shown by the red fitted line that is almost parallel to the diagonal. Learner "a" highlights the importance of measuring assessment quality through both accuracy and leniency. Learner "a" obtained a mark of 41% in grade 10, but failed mathematics in matric with a mark of 24%. Due to high assessment leniency, the incorrect message was conveyed to the learner in grade 10, suggesting they would pass Mathematics in matric rather than advising a switch to Mathematical Literacy.

5.3.1 Mean Report Mark and NSC Result Gaps in the FET phase

Using learner-level data across all quintiles the average gaps were calculated for each school year in the FET phase for the respective cohorts. These gaps at a learner level were computed as the final NSC mark minus the report mark, with larger gaps indicating less lenient assessment practices. For grades 10 and 11, the term 4 report mark was employed, while for matric, the term 3 report mark of learners was used, consistent with the correlation calculations in Section 5.1. Figure 22 illustrates the leniency at different grades for the respective cohorts for all quintiles. The red columns represent the mean NSC result for learners in Limpopo for the specific cohort. The red column remains unchanged for the different grades in each cohort, as the leniency of assessment in grades 10, 11 and 12 is measured against the mean NSC result in Mathematics for the same year. The mean NSC Mathematics results increased by over 5% from 2020 to 2021 but decreased again in 2022 to just above 35%. In Cohort 1, only the matric results were impacted by COVID-19 assessment policies. The matric group of 2020 had the largest gap—approximately 10%—between the mean NSC marks and the mean matric term 3 report marks. This large gap is largely due to the low term 3 report marks as the result of the extended school closures in 2020, as well as the lower number of learners who decided to write the NSC examination in that year (Hoadley, 2020). In Cohort 1, no leniency in any of the FET grades is detected, as the average NSC mark is higher than report marks of the respective grades.

In Cohort 2, despite higher NSC marks, some leniency in assessment is observed in grade 11, which corresponds to the year 2020 for this cohort. Learners performed on average 3% better in grade 11 than they did in the final NSC results. This leniency does not seem to persist into matric for this cohort, where matric learners had lower term 3 report marks compared to their final NSC results.

Cohort 3 exhibits leniency in grades 10 and 11, which fall within the years 2020 and 2021, respectively. This leniency might be attributed to the substitution of end-of-year examinations with tests during the 2020 -2021 period (Department of Basic Education, 2021). Additionally, the weighting of SBA-to-end-of-year was amended to 60:40, compared to the 25:75 weighting before 2020 (Department of Basic Education, 2020*a*). School-level data from Limpopo shows that some schools still opted to conduct examinations in both June and at the end of the year during 2020 and 2021. This resulted in lower leniency in the Mathematics marks of the learners in these schools, since full examinations are more challenging than control tests. Further investigation is needed to examine whether assessment leniency in lower quintiles is masked by better-performing higher quintiles.

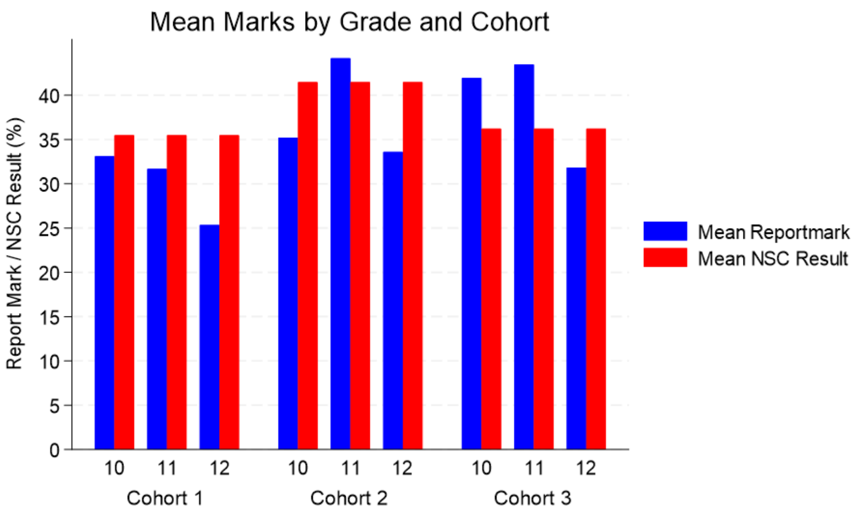


Figure 22: Mean Report Marks and NSC Results by Grade and Cohort

5.3.2 Mean Gaps in the FET Phase by Quintile

Figures 23 represent the average gaps within schools for each respective quintile. In Cohort 1, the only quintile group for which assessment leniency is observed is quintile 5 schools in grade 10, where the average grade 10 school level gap was negative, at -2.5%. However, the number of quintile 5 schools in the sample is limited, so less emphasis will be placed on the results for this quintile. Quintile 4 schools were excluded due to a too small number of schools in this quintile. Notably, Figure 23 shows that no leniency is observed in any grade for quintiles 1 to 3 in Cohort 1.

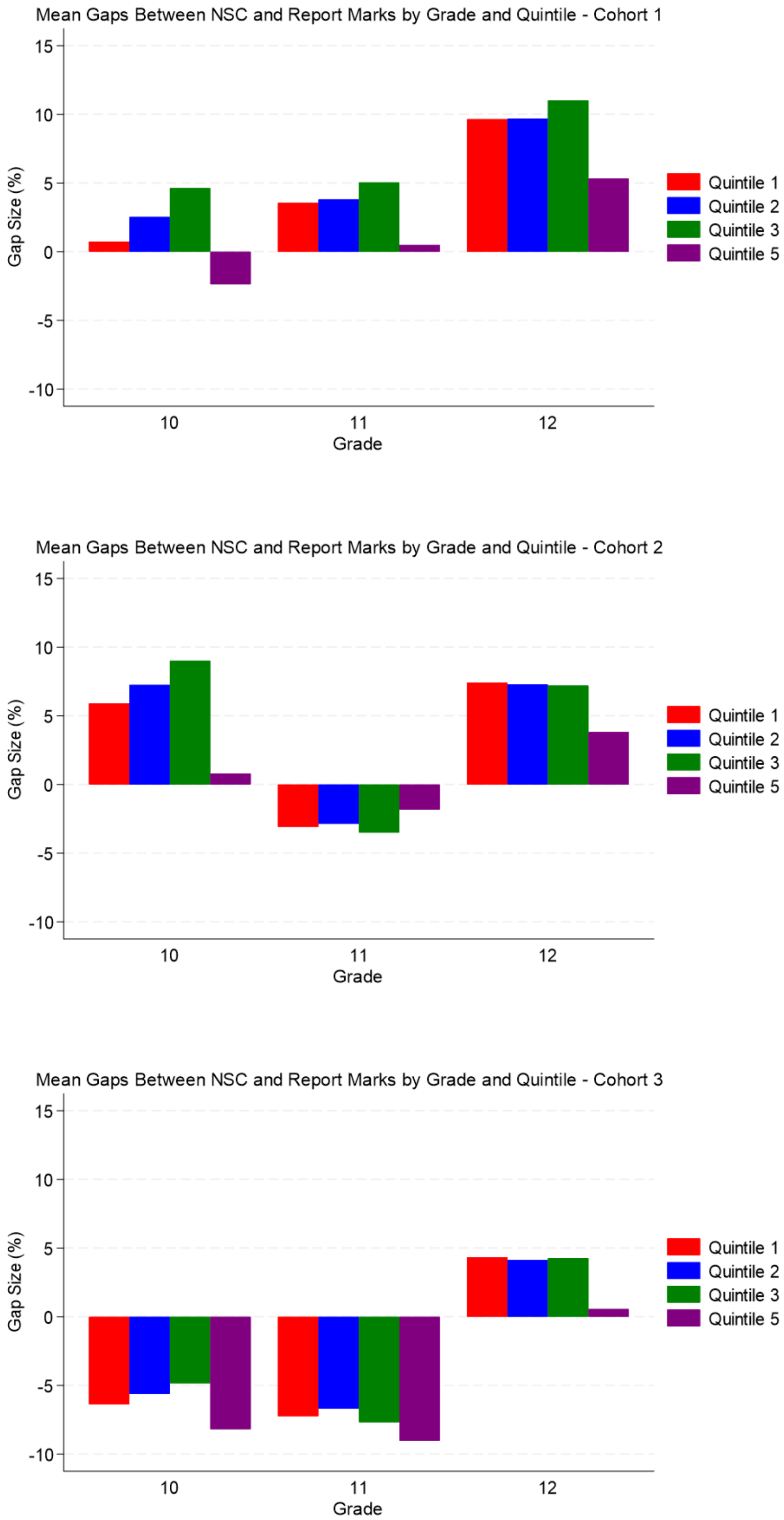


Figure 23: Leniency Within Quintile by Cohort

In Cohort 2, the only leniency observed across all quintiles was in grade 11, which coincided with 2020. All quintiles exhibited negative gaps in grade 11, meaning that the average school-level gap between the grade 11 report marks and final NSC results was negative across all quintiles. As a result, high grade 11 report marks have falsely signalled to learners that they were more capable in Mathematics when entering matric than they actually were. However, this leniency did not persist into matric, where gaps across all quintiles are positive.

Cohort 3, exhibits the highest prevalence of assessment leniency across all cohorts. For both grades 10 and 11 all quintiles had negative gaps, meaning that on average learners in the schools in the respective quintiles performed worse in the final NSC results than they did in the report marks of grades 10 and 11. What is unusual, however, is that the leniency within all quintiles increased from grade 10 to 11, but then changed to a positive gap in matric.

6 Schools That Consistently Assessed Well

6.1 Identifying Outlier Schools

The following section discusses the methods used to identify outlier quintile 1 to 3 schools, characterised by high assessment quality in Mathematics. As discussed in Section 3 of the paper, threshold values obtained from quintile 5 schools in Gauteng will be used to identify schools that portray high-quality assessments through both high accuracy and low leniency. The threshold values for both the mean Spearman correlation and gap values are reported in Table 5. The same threshold values are used for all three cohorts.

Threshold Values			
	Gr. 10	**Gr. 11**	**Gr. 12**
Correlations	0.65	0.85	0.85
Gaps	3.5	1	5

Table 5: Gauteng Quintile 5 Schools Threshold Values

Disappointingly, after applying these threshold values to the school-level data of quintile 1 to 3 schools in Limpopo, only one school was found to consistently assess well over all three cohorts. Nine schools were found to assess well over two cohorts, while 85 schools demonstrated high-quality assessment in only a single cohort. Figure 24 illustrates the relationship between the school-level mean NSC Mathematics marks and the number of cohorts in which a school was identified as assessing well, using a 95% confidence interval plot. The mean values, represented by the red nodes on the graph, are 33.95%, 40.48%, 41.49%, and 46.22%, respectively. Given that only one school consistently assessed well across all three cohorts, this group is represented by a single point on the graph. Due to the small number of schools that assessed well over only two cohorts, the confidence interval for this group is very wide and overlaps with the confidence intervals of all the other groups. However, there is a statistically significant difference in the mean NSC Mathematics marks of schools that assessed well in only one cohort and schools that did not assess well in any cohort.

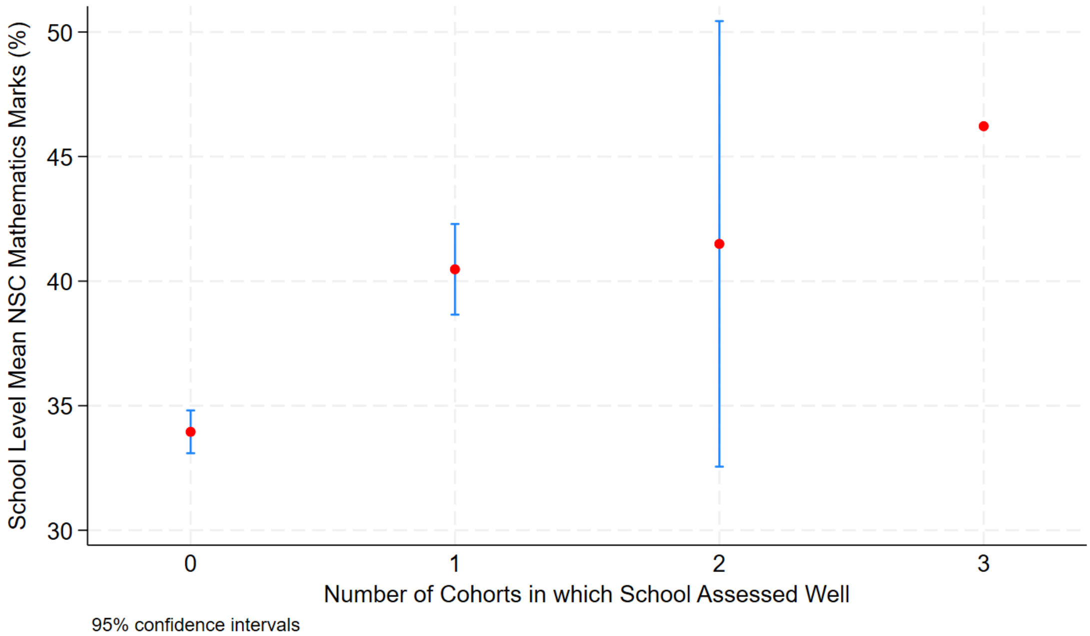


Figure 24: 95% Confidence Interval Plot of Mean NSC Results of Schools That Assessed Well

Due to the small number of schools meeting the definition of outlier schools as outlined in Section 3 of the paper, namely schools illustrating "high-quality assessments ... across **more** than one cohort," a school that demonstrated high-quality assessment **in at least** one full cohort will henceforth be referred to as an outlier school. The number of quintile 1 to 3 schools that exhibited high quality assessment according to the specified threshold values are presented by cohort in Table 6.

Quintile 1 to 3 Schools with High-Quality Assessment by Cohort		
Cohort	**Number of Schools**	**As a Percentage of All Q1-3 Schools (%)**
1	57	19.4
2	33	9.7
3	16	3.2

Table 6: Number of Quintile 1 to 3 Schools with High-Quality Assessment by Cohort

In Cohort 1, 57 schools were found to have high-quality assessment; representing 19.4% of the quintile 1 to 3 schools in the Cohort 1 sample. In Cohort 2, 33 schools met the criteria for high-quality assessment, accounting for 9.7% of the schools in that cohort. Cohort 3 had the fewest outlier schools, with only 16 schools, equating to a mere 3.2% of the sampled schools. The decline in the percentage of schools with high-quality assessment over the three consecutive cohorts reflects the decrease in assessment accuracy and an increase in assessment leniency, particularly in grades 10 and 11, as seen in Section 5 of the paper. The inconsistency in which schools achieved high-quality assessments across different cohorts highlights that very few quintile 1 to 3 schools consistently maintain high-quality assessments in Mathematics.

6.2 Characterising Outlier Schools

The remainder of this section will analyse the characteristics of the outlier schools in comparison to all other quintile 1 to 3 schools within the cohort. This analysis aims to determine whether schools with high-quality assessments achieve better NSC results than other quintile 1 to 3 schools in the same cohort and to assess whether this improved performance can be attributed to higher-quality assessments.

Academic Achievement

Figure 25 depicts the kernel density distributions of the mean NSC results of outlier schools and all other quintile 1 to 3 schools across the three respective cohorts.

For all cohorts, the density curves for the mean NSC results of outlier schools lie to the right of those of all other quintile 1 to 3 schools. This indicates that, on average, outlier schools outperform other quintile 1 to 3 schools in NSC results across all cohorts.

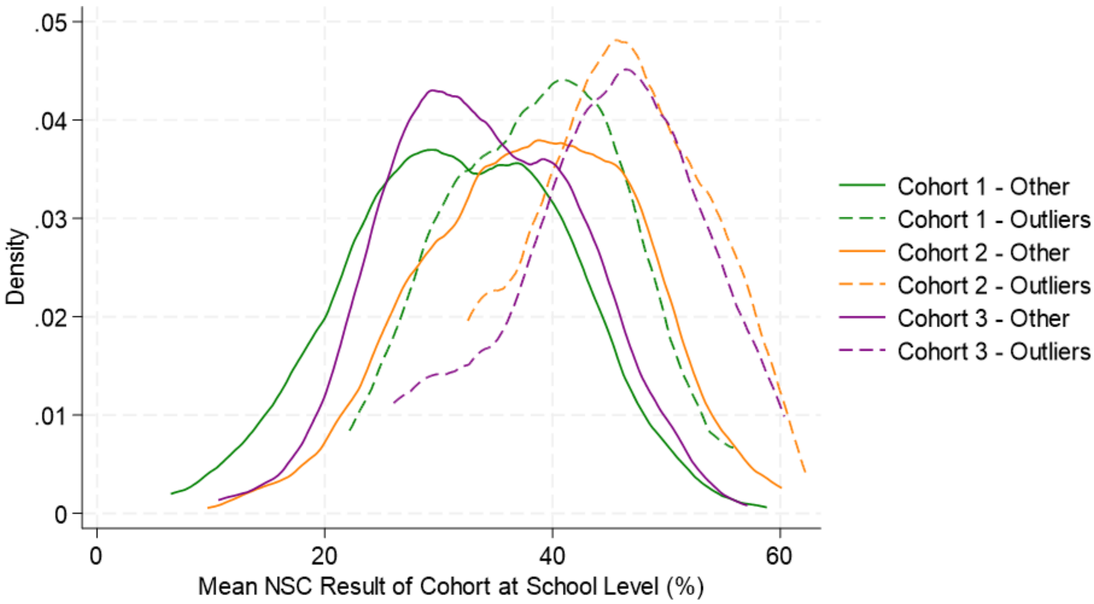


Figure 25: Kernel Densities of Mean NSC Results of Quintile 1 to 3 Schools by Cohort

To determine whether the superior performance of learners in outlier schools arises from school-specific characteristics rather than these schools simply attracting stronger learners, the mean grade 10 report marks were calculated for both outlier schools and the remaining quintile 1 to 3 schools. Table 7 presents these mean values for grade 10 report marks alongside the mean NSC results, with 95% confidence intervals included to assess whether the differences between outlier schools and other quintile 1 to 3 schools are statistically significant at the 5% level.

The table reveals that, for the mean grade 10 report marks, none of the differences between outlier schools and other quintile 1 to 3 schools are statistically significant. However, this contrasts sharply with the NSC results, where the differences are statistically significant across all cohorts. Outlier schools had on average better

Mean NSC Mathematics results than all the other quintile 1 to 3 schools in all three cohorts. This finding suggests that the superior NSC performance of learners in outlier schools is not attributable to these schools having stronger learners that take Mathematics in grade 10. Other school-specific characteristics besides assessment quality that may benefit outlier schools that needs further investigation are: learners-to-teacher ratios (Köhler, 2022), the percentage of female learners in the schools (Spaull and Makaluza, 2019), and the location of the schools (Moloi and Chetty, 2010; Du Plessis and Mestry, 2019).

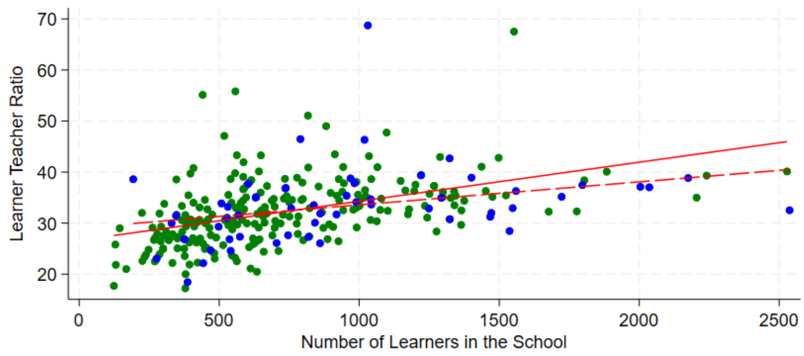
	Mean NSC Results (%)		**Mean Gr. 10 Report Marks (%)**	
Cohort	**Other Q1-3**	**Outlier Schools**	**Other Q1-3**	**Outlier Schools**
1	31.12 [29.91 ; 32.33]	38.87 [36.81 ; 40.93]	30.55 [29.49 ; 31.61]	29.55 [27.41 ; 31.70]
2	37.96 [36.92 ; 39.00]	47.16 [44.73 ; 49.60]	31.45 [30.57 ; 32.32]	33.90 [31.12 ; 36.68]
3	34.01 [33.25 ; 34.77]	46.72 [42.99 ; 50.46]	40.25 [39.53 ; 40.97]	36.19 [32.06 ; 40.32]

Table 7: Mean Gr.10 Report Marks and Mean NSC Results of Quintile 1 to 3 Schools

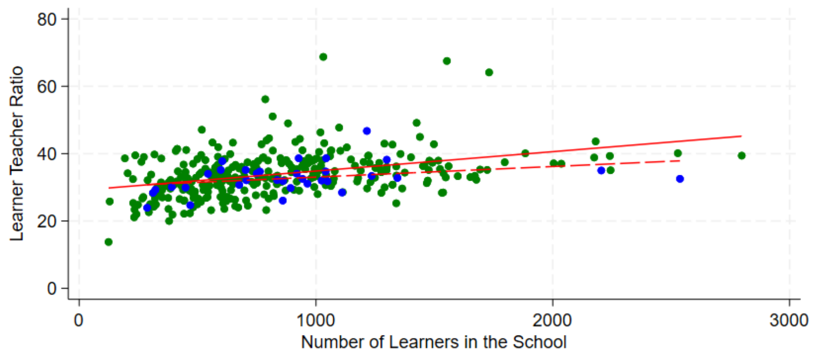
Learners-to-Teacher Ratio

Figure 26 demonstrates that across all three cohorts, there is no discernible difference between outlier schools and other quintile 1 to 3 schools in Limpopo regarding the number of learners per school or the learners-to-teacher ratio. The number of learners in outlier schools varies from 250 to 2,800, indicating that the quality of assessment within a school does not appear to be associated with its size.

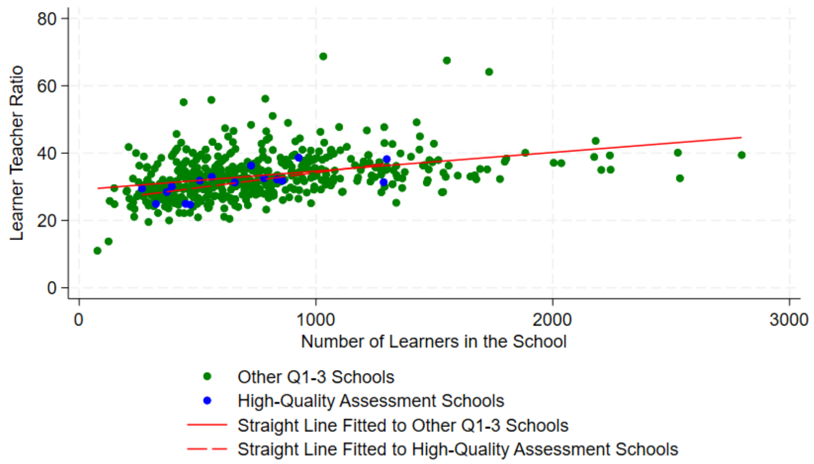
When examining the learners-to-teacher ratio, minimal differences are observed between schools with high-quality assessments and other quintile 1 to 3 schools. In Cohort 1, the fitted line for outlier schools lies slightly below the line for other schools once the school size exceeds 750 learners. However, the gap between the two lines remains narrow across the range of school sizes. For Cohort 2, the dashed line representing outlier schools consistently lies just below the line for other schools, and while the gap widens with school size, it never exceeds 10 learners. Cohort



(a) Cohort 1



(b) Cohort 2



(c) Cohort 3

Figure 26: Learners-to-Teacher Ratio and School Size by Cohort

3 reveals the closest alignment between the learners-to-teacher ratios of the two groups, with the dashed line for outlier schools lying slightly below the solid line for other schools when the school size is below 1,000 learners. These findings suggest that the superior NSC performance observed in outlier schools is not attributable to smaller learners-to-teacher ratios in schools with high-quality assessments.

Gender Composition

The gender composition of the matric Mathematics cohort of a school can also affect the average level of performance in the NSC results since male learners on average outperform female learners in Mathematics in the FET phase, when learner drop-out is not taken into consideration (Spaull and Makaluza, 2019). Table 8 presents the mean percentage of female learners in schools identified as having high-quality assessment in the respective cohorts and the other quintiles 1 to 3 schools.

Percentage Female Mathematics Learners in Quintile 1 to 3 Schools (%)		
Cohort	**Other Q1-3 Schools**	**Outlier Schools**
1	59.98	60.29
2	61.83	61.10
3	63.20	59.05

Table 8: Percentage Female Mathematics Learners in Schools (%)

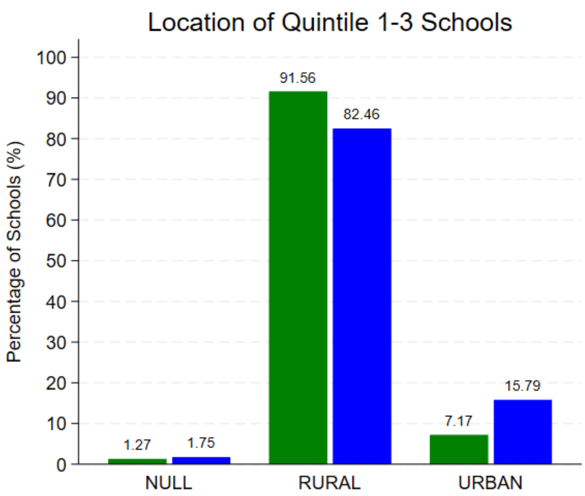
For Cohort 1 the outlier schools have on average a small percentage more female Mathematics learners in the school than other quintiles 1 to 3 schools. However, this is different for the outlier schools in cohorts 2 and 3 with less female Mathematics learners on average than the other quintiles 1 to 3 schools in the cohort. The possibility that outlier schools perform better due to a higher percentage of male learners in the matric Mathematics cohort is thus discarded. All three cohorts had a school-level average of females constituting about 60% of all learners writing Mathematics in matric. The larger percentage of female learners taking Mathematics from grade 10 to Matric relative to male students can be explained by the higher drop-out rate among male learners in the FET phase (Spaull and Makaluza, 2019).

School Location
Lastly, the geographical location of outlier schools is also an important factor to investigate, as urban schools generally outperform rural schools and are more likely to attract higher-quality teachers (Moloi and Chetty, 2010; Du Plessis and Mestry, 2019). Figure 27 illustrates the distribution of all quintile 1 to 3 schools, including outlier schools, across rural, urban, and unspecified (null) categories, expressed as a percentage. In Cohort 1, outlier schools show a notably higher proportion of urban schools, with 15.79% located in urban areas compared to only 7.17% of the other quintile 1 to 3 schools. However, this urban-rural discrepancy does not persist across cohorts. In both Cohort 2, outlier schools exhibit lower proportions of urban schools than other quintile 1 to 3 schools. In Cohort 3, similar to what was seen in Cohort 1, outlier schools have a slightly higher percentage of schools in urban areas than other quintile 1 to 3 schools. For all three cohorts the large majority of outlier schools are located in rural areas. These findings suggest that the superior performance of outlier schools in the NSC results cannot be attributed to the location of these schools.

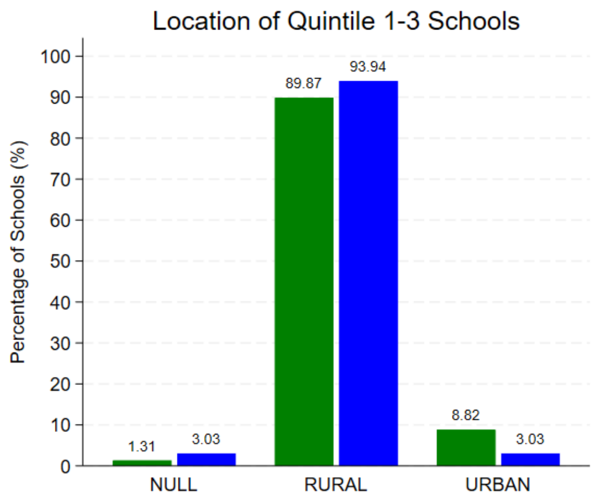
6.3 Selection into Mathematics in Outlier Schools

Percentage of Learners in Outlier Schools Taking Mathematics
Another potential explanation for the superior performance of outlier schools' in the final NSC Mathematics results is that these schools enforce stricter requirements for learners to enroll in or to continue with Mathematics in the FET phase, thereby engaging in a form of in-school cream-skimming. To investigate this, the average percentage of learners at the school level in each cohort who took Mathematics in grade 10 and continued with the subject until the NSC exam was calculated.

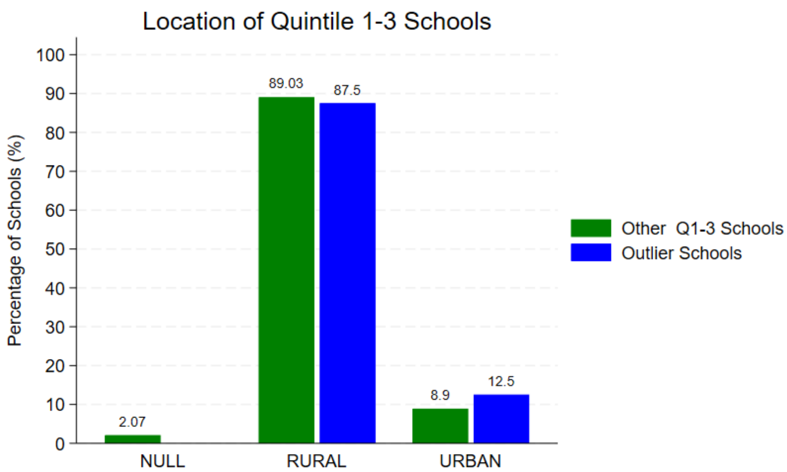
Figure 28 illustrates the average percentage of learners who wrote Mathematics in the NSC at outlier schools compared to other quintile 1 to 3 schools across the three cohorts. For Cohort 1, 80% of learners in outlier schools wrote Mathematics in the NSC. However, this figure declined for both outlier and other quintile 1 to 3 schools over the subsequent cohorts. In Cohort 2, there was still on average slightly more learners taking Mathematics in outliers schools than in other quintile 1 to 3 schools. In Cohort 3, there is a significant difference with an average of 10% less learners taking Mathematics in outlier schools than in other quintile 1 to 3 schools.



(a) Cohort 1

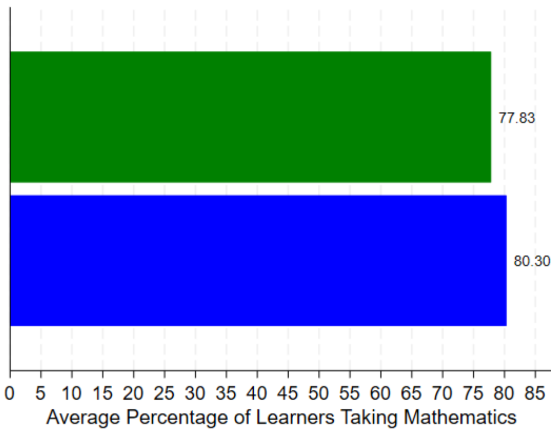


(b) Cohort 2

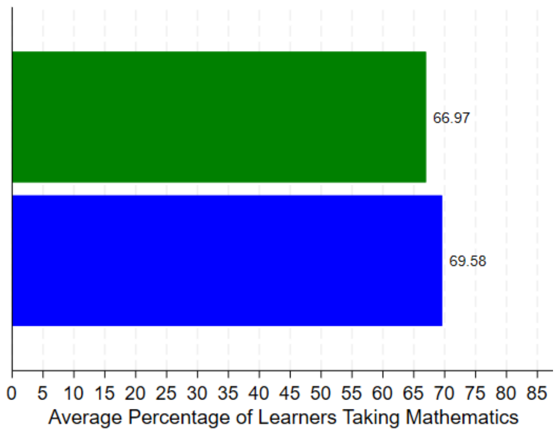


(c) Cohort 3

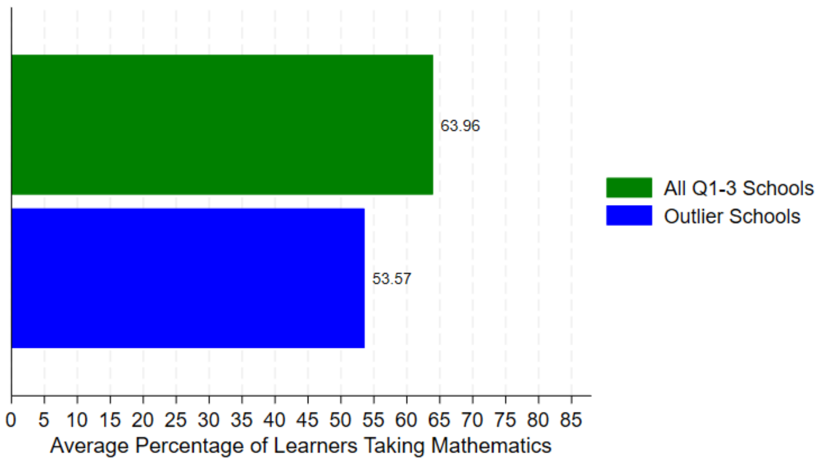
Figure 27: Location of Outlier and Other Quintile 1 to 3 Schools by Cohort



(a) Cohort 1



(b) Cohort 2



(c) Cohort 3

Figure 28: School-Level Average Percentage of Learners taking Mathematics

It is important to interpret these percentages in the context of the sample, which consists of learners who did not repeat a year or switch Mathematics subjects during the FET phase. This subgroup is likely to include academically stronger learners, introducing an upward bias toward Mathematics enrolment. Consequently, the percentages in Figure 28 are higher than those reported in Section 4, which focuses on first-time NSC attempts without excluding learners based on grade repetition or subject switching.

Table 9 reports the average percentages of grade 9 learners in outlier schools and other quintile 1 to 3 schools who continued with Mathematics as a subject in grade 10 without omitting those who dropped out before matric or switched subjects. For Cohorts 2 and 3, the percentage of learners taking Mathematics in outlier schools was slightly higher than in other quintile 1 to 3 schools, but this difference is not statitically significant. While the percentage of learners taking Mathematics increased in both groups from Cohort 2 to Cohort 3, this increase was statistically significant only for the other quintile 1 to 3 schools.

Percentage of Grade 9 learners continuing with Mathematics (%)		
Cohort	**Other Q1-3 Schools**	**Outlier Schools**
2	51.08 [49.63 ; 52.52]	54.22 [45.31 ; 63.14]
3	54.33 [52.95 ; 55.72]	63.67 [50.95 ; 76.40]

*95% confidence intervals are included in the brackets

Table 9: Percentage of Grade 9 Learners Continuing With Mathematics

Average Grade 9 Mathematics Marks of Learners Taking Mathematics
Outlier schools are defined as schools with high quality assessment and thus also schools that place greater emphasis on the signals conveyed by assessments. It would be expected that these schools would be stricter in permitting learners to continue with Mathematics in the FET phase, based on their performance in Mathematics in grade 9. Table 10 presents the mean grade 9 marks of learners in outlier and other quintile 1 to 3 schools for the respective cohorts for which grade 9 marks were available. Importantly, the values reported in the table includes all learners, without omitting those who repeated a year, switched Mathematics subjects, or dropped out before matric. For both Cohort 2 and 3 there is no sta-

tistically significant difference between the mean grade 9 Mathematics marks of grade 9 learners in outlier and other quintiles 1 to 3 schools. This is evident from the overlap in the 95% confidence intervals for both cohorts. Therefore, it appears that outlier schools are not more stringent in selecting learners who may continue with Mathematics in the FET phase compared to other quintile 1 to 3 schools.

Mean Gr. 9 Mathematics Marks of Learners taking Mathematics in Gr. 10 (%)		
Cohort	**Other Q1 - 3 Schools**	**Outlier Schools**
2	37.90 [37.32 ; 38.49]	36.11 [33.55 ; 38.66]
3	36.79 [36.22 ; 37.36]	38.89 [33.52 ; 44.27]

*95% confidence intervals are included in the brackets

Table 10: Mean Gr. 9 Mathematics Marks of Learners taking Mathematics in Gr. 10 (%)

Prevalence of Subject Switching in Mathematics Subjects

The percentage of students that switch mathematics subjects during the FET phase can also be examined for the respective groups. These percentages are calculated only for learners who progressed from grade 9 to matric without repeating a year. For Cohort 2, only 5.9% of mathematics subject switches during the FET phase were from Mathematical Literacy to Mathematics. This is similar in Cohort 3, where these switches made up only 6.05% of mathematics subject switches. The vast majority of mathematics subject switches is thus as expected from Mathematics, the more challenging subject, to Mathematical Literacy. In Cohort 2, the percentage of learners that switched mathematics subjects were 10.08% for the outlier schools and 8.50% for the other quintile 1 to 3 schools. In Cohort 3, similar to Cohort 2, the percentage of learners who switched subjects was slightly higher for the outliers than for the other quintile 1 to 3 schools, with 8.33% in the outlier schools and 7.55% in the other schools.

6.4 Matric Mathematics Performance of Learners in Outlier Schools

Probability of Failing Matric Mathematics
Lastly, a basic probit model is run to see whether learners in schools that assess accurately in Mathematics in the FET phase have a lower probability of failing Mathematics in the final NSC results. The probit model is set up as follow:

$$Fail = \beta_1(Grade9reportmark) + \beta_2(Outlier) + \beta_3(Female) + \beta_4(Overaged)$$

A probit regression is employed to examine the variable "Fail," which takes the value of 1 if the learner scored below 30% in Mathematics in matric and 0 otherwise. This regression models the relationship between the probability of a learner failing Mathematics in matric, the grade 9 report mark of the learner, the outlier status of the school, the gender of the learner as well as whether the learner is overaged or not. The variable "Outlier" takes the value of 1 if the school is identified as an outlier based on assessment quality for the cohort and 0 otherwise. "Female" is a binary variable that takes the value of 1 if the learner is female and 0 if the learner is a male. The final control variable, "Overaged," is a binary variable that takes the value of 1 if the learner is 19 years or older in their matric year and 0 otherwise.

Since probit models are non-linear, only the average marginal effects over all individuals are reported and interpreted. The marginal effects for the grade 9 marks and outlier status of the school are specifically reported. To control for heteroskedasticity, robust standard errors are included. Table 11 presents the marginal effects of the probit model for the two respective cohorts.

The marginal results over both cohorts indicate that an increase in the grade 9 report mark of a learner decreases the probability of failure. More importantly, the marginal effects indicate that if the learner attends an outlier school, the probability of obtaining a fail mark in Mathematics in matric decreases. For Cohort 2, a one percentage point increase in the grade 9 mark of a learner decreases the probability of failing matric Mathematics by 1.1 percentage points, while if the learner is in an outlier school, this decreases the probability of failing by 18.8 percentage points. The results for Cohort 3 differ only marginally from that of Cohort 2. An increase of 1 percentage point in the grade 9 report mark decreases the probability of failing Mathematics in matric by 1.4 percentage points, while being in an outlier

Marginal Effects of Probit Models on Failure in NSC						
	dy/dx	**std. err.**	**z**	**P>z**	**[95% conf.**	**interval]**
Cohort 2						
Gr.9 Report Mark	-0.011	0.000	-45.32	0.000	-0.012	-0.011
Outlier	-0.188	0.015	-12.85	0.000	-0.216	-0.159
Cohort 3						
Gr.9 Report Mark	-0.014	0.000	-71.80	0.000	-0.015	-0.014
Outlier	-0.214	0.026	-8.16	0.000	-0.265	-0.163

Table 11: Marginal Effects of Probit Models on Failure in NSC

school decreases the probability of failing by 21.4 percentage points.

Probability of achieving above 60% in Mathematics in Matric
Similarly, a probit model controlling for grade 9 report marks and outlier status can be used to investigate whether being in an outlier school improves the probability of achieving above 60% in Mathematics in the NSC. This new probit regression is:

$$Above60 = \beta_1(Grade9reportmark) + \beta_2(Outlier) + \beta_3(Female) + \beta_4(Overaged)$$

The variable “Above 60” takes on the value of 1 if the learner achieved a final NSC Mathematics mark of 60% or higher and 0 if otherwise. The marginal effects of the probit models for the respective cohorts are reported in Table 12.

In Cohort 2, a one percentage point increase in the Grade 9 report mark increases the probability of achieving 60% or more in the NSC by 0.8 percentage points. Being in an outlier school raises this probability by 10.2 percentage points. In Cohort 3, a one percentage point increase in the Grade 9 report mark increases the probability by 0.7 percentage points, while being in an outlier school raises it by 9.6 percentage points. All effects are statistically significant at the 1% level. While being in an outlier school has a notable positive impact on a learner achieving a final NSC Mathematics mark above 60%, this effect is smaller than its effect on reducing the probability of failing Mathematics in the NSC examination.

Although outlier schools look similar to other quintiles 1 to 3 schools concerning school characteristics like learner-to-teacher ratios, school size, percentage of fe-

Marginal Effects of Probit Models on Achieving 60% or more in the NSC						
	dy/dx	**std. err.**	**z**	**P>z**	**[95% conf.**	**interval]**
Cohort 2						
Gr.9 Report Mark	0.008	0.000	46.55	0.000	0.008	0.008
Outlier	0.102	0.008	12.10	0.000	0.085	0.118
Cohort 3						
Gr.9 Report Mark	0.007	0.000	54.18	0.000	0.007	0.008
Outlier	0.096	0.011	8.64	0.000	0.074	0.118

Table 12: Marginal Effects of Probit Models on Achieving 60% or more in the NSC

male students, and the number of learners electing to take Mathematics, there are clear signs of better NSC performance among these schools for the cohort in which they demonstrated high-quality assessment.

7 Discussion and Conclusion

The quality of assessment in Mathematics in the FET phase for the three respective cohorts was determined by calculating both the level of accuracy and leniency of assessment within schools. A positive relationship between higher assessment quality and better learner performance was identified. Unfortunately, only a single poor school in Limpopo illustrated consistently high assessment quality across all three cohorts. Consequently, the definition of an “outlier” poor school had to be revised to include those that demonstrated high-quality assessment for at least one cohort.

A concerning finding was that there is a decrease in the percentage of quintile 1 to 3 schools in Limpopo that demonstrate high-quality assessment in Mathematics. The deterioration in assessment quality coincides with the implementation of assessment policies aimed at mitigating the impact of school disruptions due to COVID-19 on learners’ educational outcomes. The effect of more lenient assessment policies, coupled with more relaxed progression requirements was especially evident in grades 10 and 11. As a result, the signals sent to learners regarding their Mathematics ability through assessments were thus distorted by these policy changes.

Since high-quality assessment reduces the likelihood of a learner failing Mathematics in matric and increases the probability of achieving an above 60% pass, the weaker signals regarding learners' abilities in Mathematics could have a negative impact on their final NSC Mathematics marks. Further research needs to be done in order to investigate the impact of the weaker signals through assessment, particularly in Mathematics, on grade repetition.

Data limitations, particularly the inability to track a complete cohort unaffected by COVID-19 education policies, restricted the study's ability to isolate the effect of high-quality assessment on learners' performance in Mathematics in matric. An area for future research is examining cohorts prior to the COVID-19 pandemic to determine how the prevalence of schools with consistently high-quality assessments compares to the patterns observed in the post-COVID-19 period. As new data becomes available, further research should also focus on investigating how long the effects of increased leniency and reduced accuracy in assessment persist after the suspension of COVID-19 related education policies.

References

Baird, J., Andrich, D., Hopfenbeck, T.N. and Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy Practice*, vol. 24, no. 3, pp. 317–350.

Bisgard, J. *et al.* (2022). Consolidated report on covid-19 research. final. Tech. Rep., Khulisa, Johannesburg.

Department of Basic Education (2011). Curriculum and assessment policy statement grades 10 - 12: Mathematics.

Department of Basic Education (2020 September*a*). Circular s7. revised promotion requirements for grades 10 and 11 for the 2020 year.

Department of Basic Education (2020 March*b*). Wq 434 na learner to educator ratio in quintiles 1, 2 and 3 schools.
Available at: `https://www.education.gov.za/Newsroom/ParliamentaryQuestions`

Department of Basic Education (2021 May). Circular s10 of 2021. release of the revised subject weightings for fet, grades 10 -12.

Department of Basic Education (2022 October). National assessment circular 5 of 2022. special mark adjustment and condonation dispensation for learners in grades 4-9.

Department of Basic Education (2023 March). Circular s8. release of revised subject weightings relating to the programme of assessment for the further education and training (fet) band, grades 10 -12.

Du Plessis, P. and Mestry, R. (2019). Teachers for rural schools–a challenge for south africa. *South African Journal of Education*, vol. 39.

Gustafsson, M. (2016). *Understanding trends in high-level achievement in matric mathematics and physical science*. Department of Basic Education, Pretoria.

Hoadley, U. (2020). Schools in the time of covid-19: Impacts of the pandemic on curriculum. *RESEP Non-Economic Working Paper*.

Kotze, J. (2021). Research note: Effects of the pandemic on early grade reading outcomes.

Köhler, T. (2022). Class size and learner outcomes in south african schools: The role of school socioeconomic status. *Development Southern Africa*, vol. 39, no. 2, pp. 126–150.

Moloi, M. and Chetty, M. (2010). *The SACMEQ iii project in South Africa: A study of the conditions of schooling and the quality of education*. SACMEQ Educational Policy Research Series. Department of Basic Education.

Newton, P.E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, vol. 14, no. 2, pp. 149–170.

O'Donovan, B., Rust, C. and Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment Evaluation in Higher Education*, vol. 41, no. 6, pp. 938–949.

Pitt, E., Bearman, M. and Esterhazy, R. (2020). The conundrum of low achievement and feedback for learning. *Assessment Evaluation in Higher Education*, vol. 45, no. 2, pp. 239–250.

Reddy, V. and Hannan, S. (2018). *Timss in South Africa: Making global research locally meaningful*, vol. 23. Human Sciences Research Council, Pretoria.

Scriven, M. (1967). The methodology of evaluation. In: *American Educational Research Association*. Washington, DC.

Scriven, M. (1991). *Evaluation Thesaurus*. 4th edn. Sage, Newbury Park.

Shepherd, D.L. and Van der Berg, S. (2020). Analysing matric data to identify "promising" schools in mathematics performance. *Working Papers 16/2020, Stellenbosch University, Department of Economics*.

Smith, E. and Gorard, S. (2005). 'they don't give us our marks': The role of formative feedback in student progress. *Assessment in Education: Principles, Policy Practice*, vol. 12, no. 1, pp. 21–38.

Soudien, C., Reddy, V. and Harvey, J. (2022). The impact of covid-19 on a fragile education system: The case of south africa. In: *Primary and secondary education during COVID-19: Disruptions to educational opportunity during a pandemic*, pp. 303–325.

Spaull, N. and Makaluza, N. (2019). 'girls do better: The pro-female gender gap in learning outcomes in south africa 1995–2018. *Agenda*, vol. 33, no. 4, pp. 11–28.

Taylor, S., Van der Berg, S., Reddy, V. and Janse van Rensburg, D. (2011). How well do south african schools convert grade 8 achievement into matric outcomes? *Working Paper 13/2011, Stellenbosch University, Department of Economics*.

Van der Berg, S. and Gustafsson, M. (2019). Educational outcomes in post-apartheid south africa: Signs of progress despite great inequality. In: Spaull, N. and Jansen, J. (eds.), *South African Schooling: The Enigma of Inequality*, pp. 25–45. Springer, Cham.

Van der Berg, S. and Shepherd, D. (2015). Continuous assessment and matriculation examination marks-an empirical examination. *South African Journal of Childhood Education*, vol. 5, no. 2, pp. 78–94.

Van der Berg, S., van Wyk, C., Gustafsson, M., Meyer, H., Chari, A., van Biljon, C., Lilenstein, A., Selkirk, R. and McCullum, J. (2023). What rich new education data can tell us: New insights into learner flows, assessment, learner mobility and the subjects teachers teach. Tech. Rep., Research on Socio-Economic Policy, Stellenbosch University, Stellenbosch.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In: Andrade, H.L. and Cizek, G.J. (eds.), *Handbook of formative assessment*, pp. 18–40. Routledge, New York, NY.

Wills, G. (2017). What do you mean by 'good'? the search for exceptional primary schools in south africa's no-fee school system. *Working Papers 16/2017, Stellenbosch University, Department of Economics*.

Wills, G., Selkirk, R. and Kruger, J. (2024). School completion, the matric and post-school transitions in south africa. Tech. Rep., Research on Socio-Economic Policy, Stellenbosch University, Stellenbosch.

Wills, G. and Van der Berg, S. (2022). Covid-19 disruptions and education in south africa: Two years of evidence. *SSRN*.