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.

05/25

AUGUST 2025

This paper was produced as part of the MILAPS project, funded by Optima and utilises Data Driven Districts data

Assessing assessment:

An Exploratory Analysis of School-based Assessment before, during and after the COVID-19 Pandemic in Gauteng

Author

Angela Euston-Brown

OPTIMA

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH

ASSESSING ASSESSMENT: AN EXPLORATORY ANALYSIS OF SCHOOL-BASED ASSESSMENT BEFORE, DURING AND AFTER THE COVID-19 PANDEMIC IN GAUTENG

Angela Euston-Brown

Abstract

School-based assessment (SBA) plays a critical role in signalling academic progress and shaping learner trajectories in secondary education. Yet, concerns around the accuracy and reliability of SBAs remain, particularly in contexts where external benchmarking is limited. This paper uses administrative data from South Africa's Gauteng province to examine how the reliability and validity of SBA changed during the Covid-19 pandemic. Focusing on Grades 11 and 12 in Mathematics, Mathematical Literacy, and English First Additional Language (EFAL), the analysis compares SBA marks with externally assessed National Senior Certificate (NSC) results from 2018 to 2022. SBA reliability is measured using Spearman rank correlations, while SBA validity is measured using the size of the gap between SBA and NSC scores. The results show that the pandemic disrupted SBA quality, notably so in Grade 11 compared to Grade 12. In addition, SBAs in no-fee schools are found to be less reliable than in fee-paying schools, especially for EFAL and Mathematical Literacy, though this gap narrows in Grade 12. Gender differences in assessment validity persist even after controlling for learner performance, pandemic years and school factors, with girls more likely to have smaller differences between their SBAs and final NSC results. While the analysis is exploratory and does not intend to establish causality, the findings contribute to a growing literature on assessment quality in South Africa and highlight the need to strengthen assessment moderation and support practices to ensure equitable and accurate assessment across schools.

JEL Classification: I21, I24, I28

Keywords: School-based assessment, assessment reliability and validity, Covid-19 and education, secondary schooling, National Senior Certificate, teacher-assigned grades, South Africa, educational inequality, Mathematics achievement, administrative data

CONTENTS

TABLES

FIGURES

1 INTRODUCTION

School-based assessment (SBA) plays a critical role in the learning and teaching process, the results of which send signals to various stakeholders and shape future decisions (Braun & Kanjee, 2006; Nitko & Brookhart, 2014; Van der Berg & Shepherd, 2009). As such, the quality and accuracy of school assessments are vital, as they significantly influence educational outcomes and young people's lives. While there is a growing body of literature on assessment accuracy - particularly focusing on identifying teacher bias as the explanation for discrepancies between school-based and external assessments (Black & de New, 2020; Burgess et al., 2022; Burgess & Greaves, 2013; Campbell, 2015; Lavy & Sand, 2015; Lavy, 2008; Lindahl, 2007; Ouazad, 2014) - this evidence base is scarce in developing countries, particularly those in Africa. Furthermore, while research has shown changes to repetition, enrolment and dropout rates during the pandemic in South Africa (Wills & Van der Berg, 2024; Van der Berg & Böhmer, 2025; Gustafsson, 2022), how learners' secondary school performance changed has yet to be explored.

This paper thus locates itself firmly in the South African context, using administrative data for the Gauteng province to better understand the factors that contribute to discrepancies between learners' SBA results and their external, school-leaving National Senior Certificate (NSC) results. Informed both by the international work on teacher bias and local studies on assessment accuracy (Van der Berg & Shepherd, 2009), the analysis examines the gaps and correlations between school assessment results and NSC results in Grade 11 and 12 and across three subjects, namely Mathematics, Mathematical Literacy and English First Additional Language (FAL). The Spearman rank correlations between SBA results and external NSC examination results are used as a measure of SBA *reliability*, while SBA *validity* is measured through the size of the gap between SBA and NSC results.

Two key research objectives underpin the analysis. The first research objective is to understand whether, and how, SBA results were affected by the Covid-19 pandemic in the Gauteng province. The distribution of school-based results for five balanced panels of Grade 11 and 12 learners are descriptively explored across years and grades. The second research objective is to investigate how the pandemic differentially affected the reliability and validity of SBA - when measured using correlations and gaps between SBA and NSC results - across subjects, grades, socioeconomic status, gender and levels of academic performance. The analytical approach does not aim to make causal claims, but rather focuses on identifying trends in assessment, as well as which student and school characteristics are significantly associated with higher levels of SBA quality. Particular attention is given to the differences in assessment accuracy for Mathematics.

Building on the work of Van der Berg and Shepherd (2009), the analysis finds that the Covid-19 pandemic is strongly associated with large, negative and lasting effects on assessment quality. Additionally, the effect of the pandemic on the reliability and validity of SBAs differs across subjects and grades. Grade 11 is consistently more affected than Grade 12, and Mathematics appears most affected in terms of validity, but least affected with respect to reliability. Academic performance and school fee-status are found to be strong predictors of both individual gaps and school-level correlations between Grade 11 report marks and the NSC results. Furthermore, gender emerges as a strong predictor of individual gaps. While no causal inferences are made, the exploratory analysis serves as a first step toward better understanding the trends in and predictors of SBA quality in a South African context.

The rest of the research paper is structured as follows. Section 2 reviews literature on secondary education and assessment, both internationally and in South Africa. Section 3 outlines South Africa's institutional setting and the pandemic-related assessment policy changes, while Section 4 describes the data and analytical strategies followed in the paper. The descriptive analysis and regression results are presented in Section 5 and Section 6, respectively. Section 7 discusses the results and limitations. Section 8 concludes.

2 LITERATURE REVIEW

2.1 AN UNEQUAL EDUCATION SYSTEM, LIMPING ALONG

> "[The South African education system] is, in its current form, a young system struggling to divest itself of its apartheid origins" – Soudien et al. (2022:306)

While South Africa has comparably high secondary school enrolment rates for a middle-income country, academic performance is characteristically low and highly unequal (Reddy et al., 2016; Van der Berg and Louw, 2006; Gustafsson, 2011). In the 2015 Trends in International Mathematics and Science Study (TIMSS), which tests Grade 9 learners in Mathematics, South Africa was amongst the five lowest performing countries (Reddy et al., 2016). Similarly, in the national school-leaving examinations Mathematics recorded one of the lowest averages (35%)[1] in 2021 and, of the learners who failed the NSC, 83% had failed one of the numeracy subjects (Van der Berg et al, 2023:19). However, this poor performance is not uniformly distributed throughout the system, with evidence suggesting that the likelihood of passing mathematics increases with the wealth of a learner's school (Taylor & Yu, 2009; Bridgman, 2020).

South Africa's education system is understood to operate as two distinct systems, divided by historical inequities in socioeconomics, race and geography (Fleisch, 2008; Shepherd, 2011; Spaull, 2013, 2019; Van der Berg, 2007, 2008, Van der Berg et al., 2016). One part of the system contains high-performing, well-resourced schools which serve the typically urban, white minority, while most of the system contains under-resourced, low-quality schools producing dismal learning outcomes (Fleisch, 2008; Spaull, 2013; Shepherd & Van der Berg, 2020; Spaull, 2019; Reddy et al., 2016; Gustafsson, 2011). In fact, Spaull (2019) finds that in 2018, the 200 top performing secondary schools had more students achieving over 80% in mathematics than the remaining 6,600 secondary schools combined. This poor performance at under-resourced schools reflects a plethora of systemic challenges, from internal inefficiencies such as high repetition rates, large class sizes and teacher absenteeism (Van der Berg, 2008; Van der Berg et al., 2016) to poor social climates (Winnaar, Arends & Beku, 2018) and high drop-out rates (Gustafsson, 2011; Spaull, 2015:36; Van Wyk et al., 2017). While many factors contribute to these outcomes, school functionality is of particular concern.

The role of school and teacher quality on student outcomes has been well-researched in post-Apartheid South Africa (Case & Deaton, 1999; Case & Yogo, 1999; Kimani & Bhorat, 2014; Moll,

[1] Amongst subjects written by over 10 000 learners (Van der Berg et al, 2023).

1998; Taylor, 2011; Van der Berg, 2007; 2008). Van der Berg et al. (2016) state both "weak institutional functionality" and "weak teacher content knowledge and pedagogical skill" as binding constraints in the education system. Using SACMEQ 2007 data, Venkat and Spaull (2015) find that 79% of grade 6 mathematics teachers have a content knowledge level below the level they teach, but these teachers are concentrated in the poorest schools. Furthermore, Taylor (2011) finds strong evidence linking improved student achievement with variables reflecting effective school resource management, such as curriculum planning, regularly updated assessment records and low teacher absenteeism, even after controlling for students' past performance and socioeconomic status. Similarly, Gustafsson and Taylor (2016) find that schools who were 'assigned' to a more functional province showed significant improvements in their final school-leaving results within five years. Evidently, assessment results can improve with improvements to educational inputs.

Despite considerable inequality and inefficiency, South Africa has witnessed improvements in mathematics performance, albeit from a low base. For instance, Grade 9 TIMSS scores steadily improved between 2003 and 2015, particularly at the lower end of the achievement distribution (Reddy et al, 2016). In addition, by analysing an adjusted set of matric results, Gustafsson (2016) finds that between 2008 and 2015 there has been an annual increase of 4.5% in the number of learners achieving over 60% in mathematics, with the largest increases occurring amongst previously disadvantaged learners. However, the low functionality of most schools raises concerns that the Covid-19 pandemic may have widened existing inequalities and undermined recent gains in mathematics achievement, particularly for the most disadvantaged learners.

2.2 THE IMPACT OF COVID-19

The Covid-19 pandemic profoundly disrupted education systems worldwide, with extensive evidence documenting learning losses (Ardington et al., 2021; Donnelly & Patrinos, 2022; Engzell, Frey & Verhagen, 2021) as well as shifts in repetition, enrolment, and dropout rates (Van der Berg et al., 2020). In South Africa, Ardington et al. (2021) found evidence of short-term learning losses in early grade reading among students in no-fee schools. Similarly, the PIRLS 2021 results report a 31-point decline in Grade 4 reading achievement compared to 2016, equating to the loss of 50–60% of a year's learning (Van der Berg & Böhmer, 2025). Furthermore, South Africa witnessed a decrease in repetition rates, no significant increase in secondary school dropouts, and an unprecedented rise in NSC candidates (Wills & Qvist, 2023; Wills & Van der Berg, 2024). The reduced repetition and dropout likely stemmed from more lenient promotion policies and flexible curricular demands during the pandemic (see Section 3.3). However, these changes were unevenly distributed across South Africa's education system.

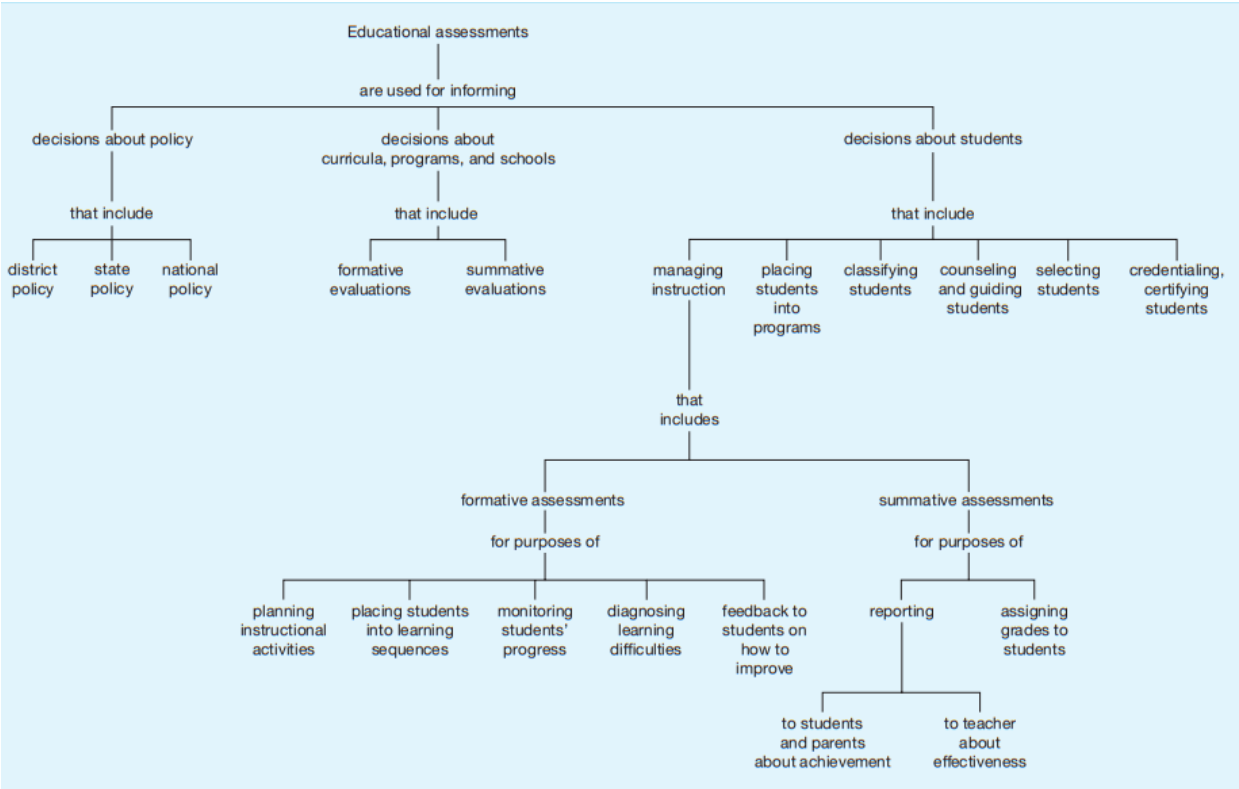
The pandemic had heterogeneous effects on South Africa's fragile and highly unequal education system, with Wills and Van der Berg (2024) identifying rising inequality as a defining educational trend of the Covid-19 years. Ardington et al. (2021) reported that learning losses in early-grade reading were more pronounced for girls and students with stronger initial proficiency. In contrast, the PIRLS 2021 results revealed a doubling of the proportion of extremely low performers compared to 2016, with socioeconomically disadvantaged students disproportionately represented in this group (Van der Berg & Böhmer, 2025). Despite these insights, and evidence of changes to assessment accuracy, limited research has explicitly explored the pandemic's effects on secondary school assessment, both internationally and in South Africa.

2.3 WHAT IS ASSESSMENT, WHY DO WE DO IT, AND HOW?

Nitko and Brookhart (2014) define assessment as "the process of obtaining information that is used for making decisions about students, curricula and programs, and educational policy". The purpose of assessment, which is well summarised in Figure 1 below, changes from feedback, to monitoring and evaluation as one moves from the classroom to the school and beyond (Nitko & Brookhart, 2014; Braun & Kanjee, 2006:10). Assessments are intended to inform teachers' decisions regarding what to teach, how to teach, and how to evaluate student progress (Nitko & Brookhart, 2014; Stiggins, 2008). In addition, assessments determine learners' promotion to subsequent grades, while simultaneously offering learners feedback to improve their mastery of content and skills (Braun & Kanjee, 2006; Kibble, 2017). Moving away from the classroom, assessment results provide valuable planning information to improve curricula, assess teaching effectiveness and evaluate schooling conditions and education quality (Braun & Kanjee, 2006; Postlethwaite & Kellaghan, 2008). Given the wide and varied purposes of assessment, policymakers are left with the challenge of balancing system-level monitoring and real-time feedback when developing assessment policies and practices.

Various arguments exist regarding the purpose of assessments, and the optimal weight between formative and summative evaluation. Black and William (1998) argue that countries at an earlier stage of educational development should administer fewer assessments and focus scarce resources only on assessments directly intended to improve learning and teaching. Despite this, literature suggests that many developing countries have inappropriate assessment policies that focus on examinations rather than formative classroom assessment (Kellaghan & Greaney, 2001). Ultimately, SBAs can act as valuable signalling devices that guide the future decisions of students, teachers, parents, schools and external stakeholders (Darling-Hammond & Wentworth, 2010).

Figure 1 Examples of Types of Educational Decisions made using Assessments. Source: Nitko and Brookhart, 2014



SBAs, which exist in some form in all education systems (Rosenkvist, 2010; Black & William, 1998), are usually set and administered by class teachers or other school staff (Department of Basic Education, 2019a). These assessments tend to align with nationally approved curricula and contain a greater variety of assessment formats, such as projects, orals or assignments (Department of Basic Education, 2019a). SBAs produce results that are immediately available to teachers, and potentially learners, and have the advantage of providing real-time feedback and influencing the course of instruction (Braun & Kanjee, 2006; Deliwe & Van der Berg, 2022). In South Africa, all assessments involved in producing a promotion result from Grade R to 11 are school based, yet not all the assessments are rigorously monitored (Hoadley, 2023). Thus, while the number and type of SBA tasks is specified for each subject and grade in South Africa's curriculum documents, there are concerns around the quality of SBAs across schools (Department of Basic Education, 2019a). Deliwe and Van der Berg (2022) argue that common SBAs[2] have the potential to improve learning and raise school-level accountability only if they are externally moderated and used in combination with standardised assessments, such as external, public examinations. However, public examinations serve very different purposes to SBAs.

[2] Common SBAs refer to assessments that are the same for a group of schools. For example, if the same test was used across all schools in a province, this test would be considered a common SBA.

Public examinations are usually designed and administered centrally. These examinations' results are intended to act as signals for high stakes decisions around selection into university or credentialling for work. Such assessments tend to be standardised, which means they are "administered under uniform conditions and graded according to a fixed set of rules" (Braun & Kanjee, 2006). South Africa administers only one public examination at the end of a learners' schooling career: the NSC. The NSC examination is universal, externally set, marked and moderated, and aligned to the Grade 12 curriculum (Deliwe & Van der Berg, 2022). It is thus considered more reliable than SBAs, as the exams assess everyone in the system against the same standards using the same marking guidelines. However, as Braun and Kanjee (2006) note, learners are seldom equally prepared due to inequalities in "opportunities to learn". This is particularly true of South Africa, and public exam results could be constructively used to influence policy and resource allocation, thereby improving educational quality and efficiency.

While assessments directly inform decisions, both at the school and national level, past assessment performance has equally been found to affect learners' future performance, independent of a learner's underlying ability. Azmat and Iriberri (2010) show that the provision of relative performance feedback increases students' grades. Similarly, Bandiera, Larcinese, and Rasul (2015) find that providing feedback has a positive effect on students' future grades, with the impact being greatest for more able students and students with less information about the academic environment. Past performance indicators have further been shown to be more beneficial to certain groups of students, such as high achievers, particularly if they come from low SES (Wyness et al., 2023; Murphy & Weinhardt, 2020). Similarly, evidence from professional tennis suggests that women may be more negatively affected by prior setbacks than men, highlighting potential gendered differences in responses to feedback (De Paola & Scoppa, 2017; Paserman, 2007). Black & de New (2020) find that harsher teacher assessments are associated with a reduction in future test performance and enjoyment in mathematics. Given the critical role of assessments in shaping educational outcomes, it is essential to design and implement them with a focus on equity and accuracy, particularly in contexts like South Africa where systemic inequalities exacerbate disparities in learning opportunities.

2.4 ASSESSMENT QUALITY

> "Regardless of how an assessment system is organised, student test results must be reliable, valid and fair… This increases the need for sophisticated assessments, as well as high quality gathering, analysis and use of the test results." – Rosenkvist, 2010

High quality assessments are usually considered to have two key principles: validity and reliability (Center on Standards and Assessment Implementation, 2018). Reliability refers to how consistent or replicable an assessment is in measuring student achievement. An assessment is reliable if, for example, the timing of when the student takes the test does not affect their result. Similarly, reliable assessments should be internally consistent: if a student gets a question correct, they should also get other questions based on the same skills or content correct. Reliability is commonly defined as the correlation between the results of a sample of individuals on two separate occasions (Braun & Kanjee, 2006). Validity, on the other hand, refers to how accurate an assessment is or how well it measures what it is intended to measure. High quality assessments should also be fair, in that they are free from bias and aligned with national curricula and assessment standards. At the school level, it is largely teachers who determine the quality of assessments.

A critical component of successful SBA is teacher expertise and assessment beliefs (Maxwell, 2006). Teachers must know, understand and be able to 1) use good assessment procedures to obtain accurate information and 2) use relevant performance standards to ensure accurate judgements. In-service teacher training, exemplars of good practice, and self-monitoring processes are thus essential to ensure reliable, valid and fair SBAs (Maxwell, 2006). This is particularly important given that teachers' attitudes and beliefs about the nature and purpose of assessment – how they conceptualise assessment – directly affect how they teach and assess, which in turn influences student engagement and learning (McMillan, 2013:xxiii). Furthermore, teacher conceptualisation of assessment may change in response to new policies (Fives & Buehl, 2012). For example, Brown and Harris (2009) find that when school-level analysis of assessment results was implemented in New Zealand, teachers moved away from believing that assessment was solely improvement-oriented to believing that assessment results reflected school quality. This highlights the critical role of teachers' beliefs about and expertise in assessment for understanding the accuracy of SBA in South African schools. However, grading accuracy has also been shown to be influenced by teachers' biases.

Most international literature exploring assessment accuracy focuses on teacher bias in grading behaviour, by gender (Lindahl, 2007; Lavy, 2008; Lavy & Sand, 2015), ethnicity (Burgess & Greaves, 2013; Ouazad, 2014; Campbell, 2015) and body size (Black & de New, 2020). The common identification strategy compares the results of different groups in a 'non-blind' classroom assessment and a 'blind' anonymously marked assessment (Lindahl, 2007; Lavy, 2008; Lavy & Sand, 2015; Lindahl, 2016; Burgess et al., 2022; Contreras, 2024). Burgess and Greaves (2013) find evidence that black and minority pupils in the United States are more likely to be under-assessed by teachers, while Lindahl (2007), Lavy (2008) and Contreras (2024) find evidence in Sweden, Israel and Chile that boys receive lower teacher assigned grades than girls when comparing to their grades in external

exams. Similarly, Burgess et al. (2022) show that semi-external assessment in mathematics at the end of high school in Denmark reduces the gender gap in graduation from post-secondary STEM degrees. The influence of teacher characteristics on grading behaviour has also been explored (Lindahl, 2016), with Cheng and Kong (2023) finding that female teachers are more likely to practice grade inflation than male teachers. Although no studies in South Africa have explored teacher bias in assessments, the stark socioeconomic and racial inequalities in the education system suggest that similar dynamics may exist. However, discrepancies between internal and external assessments have also been attributed to differences in learner characteristics.

Learners' behaviour, as opposed to teachers', has also been found to explain the gaps between 'blind' and 'non-blind' assessment results. Wyness et al. (2023) find that higher achievers are more accurately predicted compared to lower achievers, stating that "the higher-up the achievement distribution, the lower the probability of over-prediction". The authors further find that the accuracy of predictions varies by school type, socioeconomic status, and subject, and may arise due to certain groups of students having more noisy performance trajectories than others (Wyness et al, 2023). Additionally, observed teacher bias may instead be due to unobserved characteristics of a student (such as self-confidence and other non-cognitive skills) that are incorporated into grading by the teacher, but also correlated with the learners' gender, ethnicity or socioeconomic status. Alternatively, internal-external results gaps may be due to systematic differences in how students perform in externally marked, high-pressure exams relative to in-class performance, based on their gender, ethnicity or socioeconomics (Black & de New, 2020). While no existing research has explored the presence of teacher bias in South Africa, experts are sceptical about the accuracy of SBA results.

2.5 SCHOOL BASED ASSESSMENT QUALITY IN SOUTH AFRICA

In South Africa, concerns have been raised around the quality of SBAs. Teachers' limited understanding of assessment concepts, insufficient capacity to design and evaluate tasks and weak subject-specific content knowledge have been cited as drivers of low-quality assessment in schools (Umalusi, 2004; Van der Berg & Shepherd, 2009; Carnoy et al. 2012; Department of Basic Education, 2019a). The Department of Basic Education (DBE) (2019a) further mentions teacher and parent assistance during tasks and poor internal moderation and monitoring as challenges facing the quality of SBAs. Teacher bias has not directly been noted as a challenge, although a lack of professional capacity would undoubtedly lead teachers to produce subjective rather than objective assessments of learner ability. Given these concerns, it is surprising that very little research, barring that of Van der Berg and Shepherd (2009), has explored the accuracy of SBA results in South Africa.

Following a similar identification strategy to the research on teacher bias, Van der Berg and Shepherd (2009) investigate the accuracy of SBA in South Africa by comparing Grade 12 Continuous Assessment results (CASS) to the external Matric examination results between 2003 and 2005. In line with the components of a quality assessment, the authors investigate two dimensions of inaccurate assessment: 1) inflated CASS results, which they identify as the gap between the CASS mark and exam mark, and 2) unreliable assessment, which they identify as a weak correlation between CASS marks and exam marks. At the individual level, Van der Berg and Shepherd (2009) find that average CASS marks are consistently higher than the matric exam mark, for every subject and province. At the school level[3], assessment accuracy differed little amongst the poorest three quintiles but was high in Quintile 4 and 5 schools. Furthermore, only 7.5% of all schools assessing inaccurately were Quintile 5 schools, while 31.3% were in Quintile 1 schools. The results suggest that assessment accuracy improves with the socioeconomic status of a school in South Africa. However, the assessments analysed took place over 10 years ago, and more recent studies have provided updated insights into SBA quality.

Building on Van der Berg and Shepherd's (2009) foundational study, more recent research explores how SBA quality varies across grades and provinces. Van der Berg et al. (2023:20) present the correlations between earlier grade SBA marks and 2019 NSC results of an Eastern Cape student cohort. They find that correlations significantly increase in Grade 12, indicating greater alignment of SBA tasks with the NSC standards. In addition, when analysing the correlation for Mathematics by quintile in the Eastern Cape, Quintiles 4 and 5 have higher correlations, but by Grade 12 the quintile differentials become negligible. Furthermore, contrary to Van der Berg and Shepherd's (2009) findings, Van der Berg et al. (2023:22) find the NSC result is oftentimes higher than the SBA marks for Mathematics in 2022 Term 2. For example, in Gauteng, only 5% of students received an NSC result that was more than 10% below their SBA result. While Van der Berg et al. (2023) provide valuable insights into provincial and grade-level differences, their findings do not explicitly explore trends in assessment quality over time or the impact of the pandemic on these measures. The analysis presented in the following sections thus aims to update the findings of Van der Berg and Shepherd (2009) while offering a deeper understanding of the statistics presented by Van der Berg et al. (2023).

[3] The analysis at the individual school level was not weighted by the number of candidates per school.

3 INSTITUTIONAL SETTING AND CONTEXT

3.1 THE SOUTH AFRICAN SCHOOL SYSTEM

In South Africa, the school system is organised into three levels: pre-primary, primary and secondary education. Secondary education consists of five grades separated into two phases, the General Education and Training (GET) phase (Grade 8 and 9) and the Further Education and Training (FET) phase (Grade 10, 11 and 12). Schooling is mandatory until the completion of Grade 9. In the FET phase, which is the focus of this paper, students are required to take seven subjects: a Home Language (HL), a First Additional Language (FAL), Mathematics or Mathematical Literacy, Life Orientation, and three elective subjects (Republic of South Africa, 2009:7).

South Africa has both public and private school providers. Private schools are administered by for-profit or non-profit private organizations, while public schools are administered by the provincial governments and financed through national government. Public schools are divided into five quintiles based on the socioeconomic status of the surrounding community, with Quintile 1 schools being the most poor and Quintile 5 being the least (Republic of South Africa, 2006:27). While private schools may choose which curriculum they follow, all public schools follow the National Curriculum and Policy Statement (CAPS), which provides guidelines for the coverage of content and programme of assessment for each grade and subject.

3.2 PROMOTION AND ASSESSMENT POLICIES AND PRACTICE

Formal assessment in South African Secondary School is separated into two components: "School-Based Assessment (SBA)" and "End of the Year Examinations" (Republic of South Africa, 2011:54). While the term "School-Based Assessment" would imply that the end of year examinations are externally set, this is only true in Grade 12. Otherwise, from Grades R to 11, all assessments involved in producing a promotion result, including the end-of-year examinations, are school based (Hoadley, 2023). Thus, the SBAs discussed in the CAPS documents refer to continuous assessments that have a greater range of types, such as projects or assignments, and are set, administered, marked and formally recorded by the school teacher for feedback and promotion purposes (Republic of South Africa, 2011). In contrast, the end of year examinations are summative, rarely used to provide learners with feedback and make up the larger proportion of the promotion result. In addition, these examinations are often "common", meaning that they are externally set by the provincial education department for core subjects such as Mathematics. While common papers provide some assurance that assessments are of reasonable quality and standardised across schools, there is no publicly

available information about which schools, grades and subjects participate in common assessments, and the practice seems to vary across years.

Assessment in Mathematics or Mathematical Literacy is clearly outlined for the FET phase. In Grade 10 and 11, the Term 4 "report mark" acts as the promotion result and determines a learner's progression to Grade 11 or 12, while a Grade 12's final school mark is their Term 3 "report mark". In the FET phase, end of year report marks are composed of 75% the end of year examination and 25% a weighted average of SBA tasks. However, in Grade 12, learners' continuous assessments are rigorously moderated by the DBE in June/July and September/October, and the end-of-year examinations – often referred to as Trial Examinations – are carefully set in line with the NSC examination guidelines and tend to be common papers. In addition to school assessment, Grade 12s write externally set, marked and moderated NSC examinations (discussed in Section 2.3). The final NSC result for Mathematics and Mathematical Literacy is comprised of 75% the external NSC examination result and 25% a weighted average of SBA tasks (but this includes the trial examinations) (Republic of South Africa, 2009:16). However, learners' NSC examination and SBA results are standardised and statistically moderated by Umalusi[4] before producing the final NSC result (Republic of South Africa, 2014:72; Umalusi, 2016). Thus, the NSC results are not a raw indication of a learner's performance in the NSC examination or across the Grade 12 school year.

Promotion requirements and policies are important to understand in order to interpret the results presented in the next sections. In the FET phase, promotion to the following grade requires that a learner achieve 40% in three subjects, one being an official language at Home Language level, and 30% in three subjects, provided the SBA component is submitted (Republic of South Africa, 2013:33). However, South Africa's grade repetition policy states that students may not be held back more than once in a school phase (Republic of South Africa, 2013:34). For example, if a student is held back in Grade 10, they must be automatically promoted to Grade 11 and 12 over the next three years, even if they do not meet the promotion requirements. In order to aid the progression of learners, mark adjustments and condonations may also be operationalised in Mathematics for Grades 8 and 9[5]. During the pandemic, many of these assessment policies were altered to ensure learners were not unfairly disadvantaged.

[4] According to the National Policy Pertaining to the Conduct, Administration and Management of the National Senior Certificate Examination (Republic of South Africa, 2014), the formula for the statistical moderation process of SBAs is provided by Umalusi, while the DBE is responsible for ensuring that the computer system is programmed to process the marks accordingly.

[5] An explanation of mark adjustment and condonation policy is found in the appendix.

3.3 SCHOOLING AND THE PANDEMIC

3.3.1 Lost Contact Time

Like many countries around the world, South Africa experienced widespread disruptions in response to the outbreak of Covid-19 in March 2020. Schools were quickly identified as high-risk areas for Covid-19 transmission and were closed on the 18 March, but remained closed after the national lockdown began on the 26 March (Department of Basic Education, 2020a). In mid-May, the Minister of Basic Education announced that schools would re-open on 1 June 2020, but only for Grades 7 and 12 (Department of Basic Education, 2020b). Thus, in addition to the days of school lost in the initial lockdown, many schools were forced to delay opening due to an inability to manage social distancing, with most schools following a phased return to schooling and adopting rotational attendance systems[6]. Contact time was further limited in 2021, as a second wave of infections forced schools to start a month late and continue with rotational attendance (Hoadley, 2023). Gustafsson (2022) thus estimates that, across Grades R to 12 in both public and independent schools, there was a 54% loss of contact time in 2020 and 17% loss in 2021 Term 3, on average. However, this loss of instructional time would have varied by grade and socioeconomic factors such as school size, local politics and economic and social trauma in households.

3.3.2 Curriculum and Assessment Policy Changes

In addition to school closures, the DBE made various changes to Annual Teaching Plans and assessment programmes to address the low curriculum coverage during 2020 and 2021. Circular S2 of 2020 (Department of Basic Education, 2020b) released revised ATPs which reduced curriculum content in every grade, except Grade 12[7], and every subject. For example, in Grade 10 Mathematics, the topic on Statistics and Financial Mathematics was removed from the curriculum. However, these ATPs did not anticipate the delayed and staggered reopening, rotational timetabling and high learner and teacher absenteeism, and the Department acknowledged that the curriculum may require further school-based trimming and reorganisation (Department of Basic Education, 2020d). Similarly, changes were made to assessments that placed a greater focus on formative assessment and increased teaching time (Department of Basic Education, 2020b).

During the pandemic, greater emphasis was placed on SBAs in the GET and FET phases. To maximize instructional time, June examinations were cancelled, end-of-year exams were replaced

[6] The rotational attendance system involved certain classes or grades alternating between staying at home and attending school and would have caused some grades to lose more teaching days than others.
[7] The Grade 12 curriculum was not trimmed, but may have been reorganised to make more effective use of time (Department of Basic Education, 2020b).

with controlled tests focused on covered content, and common examinations were suspended (Department of Basic Education, 2020b; 2020c; 2020f). In addition, in Grades 10 and 11, the examination (or control test) component of the promotion result was reduced from 75% to 40%, and the SBA component was increased from 25% to 60% (Department of Basic Education, 2020f). The 75% examination and 25% SBA split was retained for Grade 12s (as summarised in Table 1 below), and Grade 12 learners were still expected to write the Trial Examinations and final NSC examinations (Department of Basic Education, 2020b). Thus, while the curriculum and assessment programme underwent changes in Grades 10 and 11 in 2020, the Grade 12 curriculum and assessment programme remained largely unaffected. However, in 2020, the national external moderation of Grade 12 SBAs only took place once during September/October (Department of Basic Education, 2020e) (rather than in both June/July and September/October).

Table 1. SBA Weight in Promotion Results

	Grade 10 & 11		**Grade 12**	
	SBA	**Exam**	**SBA**	**Exam**
2018	25	75	25	75
2019	25	75	25	75
2020	60	40	25	75
2021	60	40	25	75
2022	60	40	25	75
2023	40	60	25	75

While all learners and teachers returned to full-time school attendance in January 2022 (Department of Basic Education, 2022), the assessment and curriculum changes implemented in 2020 were retained until the end of 2022. It was only in 2023 that mid-year examinations were reinstated and the SBA:exam ratio shifted back in favour of examinations, but only to a 40:60 split (Hoadley, 2023) (See Table 1 below). This lack of formal examinations and the increased weighting of SBAs between 2020 and 2022 likely increased the lack of reliability in promotion decisions (Hoadley, 2023). In addition, given variation in the quality of teachers across schools in South Africa (Venkat & Spaull, 2015), curriculum coverage and assessment practices would have varied considerably between schools and classrooms. As Hoadley (2023) argues, while the devolution of responsibility to the school and teacher level was necessary, it placed somewhat unrealistic expectations on the average teacher, requiring them to "function as autonomous, highly-skilled individuals, able to exercise just-in-time professional judgements regarding content selection, in-person and remote pedagogies and appropriate assessments". The 2020-2022 period likely exacerbated educational inequalities along lines of socioeconomics, race, gender and academic performance. These inequities form the basis for the analyses in the following sections.

4 DATA AND METHODOLOGY

The main dataset used in this paper is a panel dataset, compiled by the Data Driven Districts (DDD) programme, containing anonymised, student-level data for learners attending school in South Africa's Gauteng province. The data is drawn from a school administrative system updated by teachers, the South African School Administration and Management System (SA-SAMS), and thus contains easy-to-collect student attributes and academic records. The full, unbalanced dataset includes students' Grade 10, 11 and 12 end-of-year report marks between 2018 and 2023 for Mathematics and Mathematical Literacy, and between 2018 and 2022 for English FAL. In order understand how assessment quality differs by school attributes, the 2023 Master List of Schools (South African Department of Basic Education, 2023) is matched to individuals in the SA-SAMS data. In addition, NSC results from 2018 to 2022 are merged with SA-SAMS data to compare learners' internal and external assessment performance. The different samples and measures used in the analysis are expanded on below.

4.1 ESTIMATION SAMPLES

Four variations of the data are used in the analysis. While all four estimation samples are limited to ordinary, public schools[8], they do differ in terms of balance, scope, and level of aggregation, as summarized in Table 2 below. The first version of the data, explored in Section 5.1, contains an unbalanced panel of learners for Grades 10, 11 and 12 between 2018 and 2023, which adds breadth in the descriptive analysis. In order to focus on performance trajectories, the second version of the data is a balanced panel of Grade 11 and 12[9] learners who have progressed on track from Grade 11 to 12. This second dataset contains the SBA results of 5 matric cohorts (2019 to 2023), including both a pre-Covid cohort (2018/19) and a post-Covid cohort (2022/23) cohort. NSC results are deliberately not linked to learners' SBA results to avoid eliminating the 2023 Matric cohort. The third dataset, which is most frequently used in the paper, is a balanced panel of Grade 11–12 learners that links learners' school report marks to their NSC results, enabling a comparison of internal and external assessments. While this dataset does contain a pre-pandemic cohort (2018/19), the lack of NSC results for 2023 means that this subset of the data only contains 4 matric cohorts, and comparisons of internal and external results cannot be conducted on a post-pandemic cohort (2022/23). However, the

[8] Due to the focus on assessment and the various policy changes that took place during the Covid-19 pandemic, the dataset is limited to schools within the public sector, as the weighting of assessment tasks and assessment standards should be consistent across public schools, thus ensuring a greater degree of comparability across schools.
[9] Grade 10 results were not included in the analysis as no SBA data is available for 2017. As a result, it would not be possible to have a Grade 10 to Grade 12 balanced cohort completely unaffected by Covid-19. This motivated the choice of Grade 11-12 cohorts, rather than Grade 10-12 cohorts.

inclusion of both internal and external assessments is critical for enabling a detailed analysis of assessment quality. The fourth dataset aggregates the third dataset to the school level, but only retains schools with more than 15 learners per year and subject.

Table 2. Description of Estimation Samples

Sample	Used In	Sources	Description	Restrictions
Version 1	Section 5.1	SA-SAMS Masterlist of Schools 2023	Cross-section of learners, containing their school results for Grades 10 to 12 between 2018 and 2023.	Limited to learners who have repeated no more than once in the phase.
Version 2	Section 5.2	SA-SAMS Masterlist of Schools 2023	Balanced panel of learners, containing their school results for Grades 11 and 12 between 2018 and 2023.	Limited to learners who progress on track from Grade 11 to Grade 12 i.e. no repeaters.
Version 3	Section 5.3 onwards	SA-SAMS Masterlist of Schools 2023 NSC Results	Balanced panel of learners, containing their school results for Grades 11 and 12 and their NSC results between 2018 and 2022.	Limited to learners who write the NSC examination in the same year they complete the Grade 12 SBAs.
Version 4	Section 5.3 onwards	SA-SAMS Masterlist of Schools 2023 NSC Results	Version 3 of the data aggregated down to the school, year, grade and subject level.	Limited to schools who have more than 15 learners per year, grade and subject.

The sample sizes of each version of the dataset are summarised in Table 14 and Table 15 in the appendix[10]. Table 15 shows that version 2 and 3 of the data follows between 300 and 400 schools[11] in each cohort, excluding the 2019/20 cohort in which fewer schools are captured. From 2019/20 to 2021/22, for all three subjects, almost 90% of learners are retained in the data with the addition of the school master list, removal of independent schools and Special Needs Education Centres, and the matching of NSC results. In contrast, only around 78% of learners in the 2018/19 cohort are retained during this data merging and cleaning process; potentially due to the Mixed Examination Opportunities [12] (MEO) dispensation which was discontinued in 2020 (Department of Basic Education, 2019b). The sample sizes further reflect the large increase in the number of learners in the FET phase nationally between 2021 and 2023 (Gustafsson, 2022:4; Wills & Qvist, 2023). However, the estimation samples only cover the Gauteng province. It is thus necessary to be aware of Gauteng's unique position nationally when interpreting the descriptive statistics.

[10] Given the various dimensions of year, grade and subject, these tables are kept in the appendix.
[11] To put these values into perspective, there are 763 public schools in Gauteng (1 intermediate, 655 secondary, 104 combined, 3 Special Needs Education Centres) according to the 2023 school master list. Additionally, the SBA data contains 740 Schools in 2022 for Grade 11 Mathematics prior to excluding independent and SNEC schools. When excluded, only 574 schools remain – prior to removing duplicates, learners who switch subjects, or match to grade 12 set of SBA results.
[12] The MEO dispensation offered learners the choice to write some NSC examinations in October/November 2019, and some in May/June of the following year. Given that the NSC results were matched to ensure that the learner wrote their Grade 12 assessments in the same year that they wrote their NSC examinations, the learners who used the MEO dispensation and deferred subjects to the following year would not have been retained in the data.

4.1.1 Provincial Characteristics

It is critical to contextualise the sample data within the national context. The Gauteng province, which is predominantly urban, is one of the richest provinces in South Africa and had the lowest proportion of households relying on social grants in 2023 (36.6%) (Statistics South Africa, 2023). The 2023 General Household Survey indicated that Gauteng had the highest proportion of individuals over 20 years of age with a Grade 12 or post-school qualification (60.8%) (Statistics South Africa, 2023). Additionally, Gauteng ranked 2nd out of the 9 provinces in terms of NSC pass rates in both 2021 and 2022 (Department of Basic Education, 2023). This is relevant given the focus on performance distributions, NSC results and the 2021/22 period in the analyses. Furthermore, common examinations are administered in Gauteng[13], although there is no publicly available list of which years the examinations were administered, nor to which grades or schools in the system (Gauteng Department of Education, 2022). This limits the extent to which school attributes can be used to explain assessment quality. Ultimately, the results presented in subsequent sections reflect the unique characteristics of Gauteng's education system and the specific samples analysed, and should not be generalized to the national context without caution. Considering this, the following section delves into the characteristics of the third dataset to briefly examine how changes during the pandemic period shaped the structure and composition of the data.

4.1.2 Sample Characteristics

Table 3 presents the sample size and characteristics of the balanced Grade 11/12 cohort (referred to as version 3 above[14]), including key demographics, academic performance and school attributes. These descriptive statistics provide a snapshot of learner demographics and academic performance trends across cohorts and subjects. *Overage* is an indicator variable that takes the value of 1 if the student is one or more years older than they would be if they were "on-track" in terms of age-for-grade. Home language is recoded as an indicator variable, referred to as *Eng/Afr HL*, that takes the value of one if a learner speaks either English or Afrikaans as their home language, and zero otherwise. Grade 11 and 12 report marks refer to learners' end-of-year marks, while *NSC result* is learners' adjusted NSC result for each subject. *Female* is an indicator variable that takes the value of one if a student is female. School-level variables are also captured in the data, such as the indicator variable *No-Fee School* and the *Quintile 1 - 5* variables.

[13] Refer to Table 12 in the appendix for details regarding the GDE's communication around common examinations.
[14] Sample characteristics of the sample of learners without the NSC result data merged to the SBA data can be found in the appendix.

Table 3. Learner-level Descriptive Statistics by Cohort and Subject for Version 3 of the Data.

	Mathematics				Mathematical Literacy				English FAL			
	2018/19	**2019/20**	**2020/21**	**2021/22**	**2018/19**	**2019/20**	**2020/21**	**2021/22**	**2018/19**	**2019/20**	**2020/21**	**2021/22**
Total Learners	12253	8552	13244	14036	21507	14681	26582	41986	32728	30113	35961	58504
Results												
Gr.11 Mark	37.15	36.28	43.96	43.40	40.00	37.69	41.75	40.80	54.92	53.98	56.29	53.29
< 30%	0.37	0.38	0.21	0.21	0.18	0.23	0.13	0.13	0.00	0.00	0.00	0.00
Gr.12 Mark	38.01	32.11	33.67	32.06	43.33	41.75	43.34	45.00	55.47	56.08	50.73	50.84
< 30%	0.38	0.52	0.47	0.52	0.15	0.23	0.15	0.10	0.00	0.01	0.03	0.02
NSC Result	38.55	38.42	40.17	36.94	46.41	46.66	45.42	46.52	57.46	58.85	57.38	57.51
Demographics												
Female	0.60	0.60	0.61	0.60	0.55	0.55	0.55	0.54	0.56	0.57	0.57	0.57
Over-Age	0.20	0.18	0.18	0.17	0.45	0.45	0.44	0.45	0.39	0.38	0.39	0.39
Eng/Afr HL	0.23	0.20	0.20	0.20	0.20	0.13	0.16	0.14	0.15	0.14	0.11	0.09
School												
No-Fee School[15]	0.60	0.67	0.63	0.63	0.66	0.79	0.73	0.77	0.85	0.86	0.89	0.91
Quintile 1	0.11	0.09	0.13	0.11	0.13	0.14	0.17	0.14	0.15	0.15	0.20	0.16
Quintile 2	0.11	0.12	0.11	0.09	0.12	0.12	0.13	0.15	0.16	0.15	0.18	0.18
Quintile 3	0.16	0.20	0.17	0.15	0.17	0.21	0.18	0.20	0.23	0.27	0.25	0.26
Quintile 4	0.21	0.23	0.17	0.25	0.26	0.27	0.21	0.25	0.29	0.27	0.24	0.27
Quintile 5	0.41	0.36	0.42	0.39	0.33	0.26	0.31	0.26	0.17	0.17	0.13	0.12

Notes: Characteristics of the sample without NSC results can be found in the appendix.

The substantial increase in the number of learners in the 2021/22 cohort, particularly in Mathematical Literacy and English FAL, aligns with evidence of improved retention rates in the FET phase (Wills & Qvist, 2023). This trend may reflect both policy changes during the pandemic and demographic factors, such as larger birth cohorts for these years (Gustafsson, 2022). Interestingly, while gender composition stays fairly stable across cohorts within specific subjects, the proportion of English or Afrikaans Home Language speakers falls across each consecutive cohort for the Mathematical Literacy and English FAL subjects. At the same time, the proportion of learners attending no-fee schools in each cohort rises over time. This suggests that the reduced repetition and dropout, observed nationally, may be predominantly occurring in no-fee schools. While the proportion of over-age learners seems unaffected across cohorts in Mathematical Literacy and English FAL samples, the proportion declines over time in the Mathematics sample. This, too, may speak to reduced repetition during the pandemic. The trends in report marks, which are explored in the descriptive analyses, ultimately raise questions about the reliability and consistency of assessments during the pandemic.

[15] It is thus worth noting that while all Quintile 1 – 3 schools are no-fee schools, most Quintile 4 schools and some Quintile 5 schools in Gauteng are also reported as no-fee schools based on the 2023 Master List of Schools.

4.2 MEASURES OF ASSESSMENT QUALITY

In order to explore the quality of SBAs in Gauteng, the analysis adopts two key measures of discrepancies between internal and external results, as put forward by Van der Berg and Shepherd (2009). *Reliability*, the first measure of assessment quality used in the paper, is defined as the Spearman rank correlation between internal school results and external NSC results. It thus indicates how well the relative ordering of school marks matches that of the NSC results. As seen in Figure 2 below, in the context of this paper, reliability can be visualised as the closeness of observations around a regression line. As such, reliability measures how consistently an assessment task measures what it is intended to measure across learners.

The second measure of assessment quality will be referred to as *validity* and is captured as the gap between the internal school result and the external NSC result, i.e. the school result subtracted from the NSC result for each learner. Validity, or a lack thereof, can be visualised as the deviation of the observations from the diagonal, as shown in Figure 2, and reflects how accurately an assessment task is measuring what it is intended to measure. The direction of the deviation from the diagonal can indicate leniency in internal assessment, if school results are consistently higher than NSC results, or strictness in internal assessment, if school results are consistently lower than NSC results.

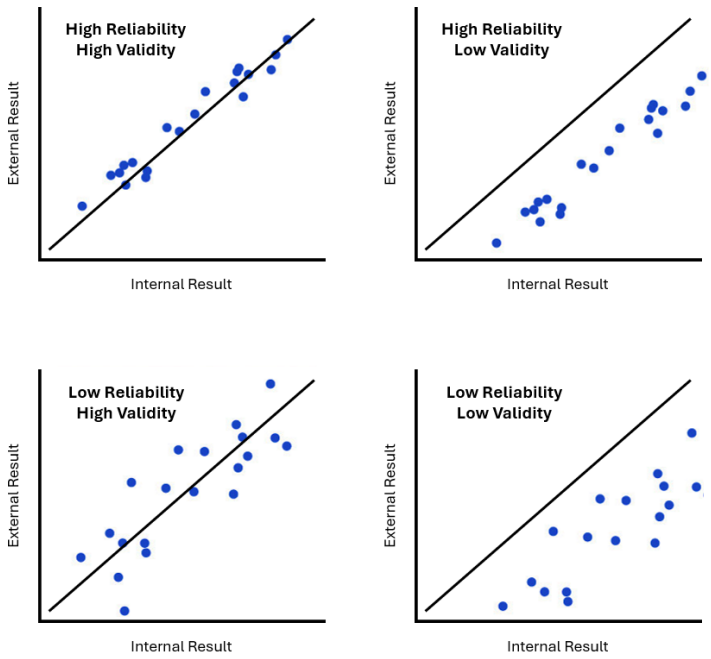


Figure 2. Graphical Representation of Reliability and Validity

The measures of assessment quality are contingent on the assumption that the NSC result is a counterfactual for the result that the student would have received in an externally set and marked SBA task. Or rather, the NSC result needs to be a consistent and accurate measure of a learner's true

ability in a given subject in any given year. Given this, the assessment quality measures are likely more reliable for Grade 12, in which the curriculum content is most closely aligned to that covered in the NSC examinations. Ultimately, this strong assumption may not always hold, particularly as NSC results are subject to statistical moderation and adjustments that vary across cohorts and subjects (see Section 3.2). These adjustments could influence the alignment between internal and external results, potentially affecting the reliability and validity measures. While these limitations highlight the complexity of quantitatively measuring assessment quality, the descriptive analysis of gaps and correlations offers valuable insights that are further strengthened through a multivariate regression analysis.

4.3 MULTIVARIATE ESTIMATION

Several Ordinary Least Square (OLS) regressions are used to determine which time period, individual factors and institutional factors are associated with larger gaps and lower correlations between SBA and NSC results. All regression output is disaggregated by grade and subject. The first set of models, presented in Section 6.1, are conducted on the school-level data. The dependent variable measures the Spearman rank correlation between a school's mean internal report marks and mean external NSC results, disaggregated by subject and year. These multivariate regressions intend to explore which school factors strengthen or weaken the correlation between a school's average report mark and NSC mark. The OLS regression models presented in Section 6.2 explore the factors determining the magnitude of the gaps between report marks and NSC results at the individual level. A second set of models presented in Section 6.1 explores the heterogenous effects of the pandemic years on assessment gaps by socioeconomic status, academic ability and gender. Importantly, in order to analyse the extent of misalignment, internal-external gaps are converted to absolute values, ensuring the focus is on the size of the discrepancy rather than its direction[16]. In addition, the standard errors are clustered at the school level to account for school-level correlation in the residuals. Although the regression analysis is exploratory and not intended to infer causality, it does identify key predictors and trends that shape assessment quality, providing a foundation for future research and policy interventions.

[16] While direction of the gap is important, the focus on assessment accuracy means that any deviation from the NSC examination could potentially reflect inaccuracies in SBAs. The absolute value of the gap additionally improves the models' interpretability.

5 DESCRIPTIVE ANALYSIS

5.1 SCHOOL-BASED ASSESSMENT IN A TIME OF UNCERTAINTY

Before exploring the performance trajectories of a balanced panel of learners, a snapshot of SBA results between 2018 and 2023 is explored for an unbalanced sample of FET phase learners. Figure 3 provides an overview of the mean end-of-year Mathematics report marks for Grades 10 to 12. Grade 10 average report marks are consistently lower than those in Grade 11. This is unsurprising given that weaker learners in Grade 10 may drop out entirely or switch to Mathematical Literacy in Grade 11, leaving an academically stronger cohort in Grade 11 Mathematics. Additionally, progression policies may explain the differences between Grade 10 and 11 averages, as Grade 10 is the first year in the FET phase[17]. Despite the difference, the Grade 10 and 11 averages follow an almost identical trend over time, with these parallel trends similarly observed for Mathematical Literacy and English FAL report marks (see Figure 13 and Figure 14 in the Appendix). The same cannot be said for the Grade 12s.

Pre-Covid-19 (2018 and 2019), Grade 12s appear to do better than all other grades. This would be expected in the full sample, as the weak performers in Grade 10 would switch to Mathematical Literacy or dropout, leaving an academically stronger Grade 12 cohort (Van Wyk et al., 2017). In contrast, from 2020 to 2022 the Grade 12s perform worse than in previous years, and worse than other grades, with the Grade 10s and 11s seeing a considerable increase of around 10 percentage points in the mean Mathematics report mark from 2019 to 2020. The increase in report marks for Grades 10 and 11 is consistent across Mathematics, Mathematical Literacy and English FAL. While the Grade 10 and 11 Mathematics averages decrease slightly in 2021 and 2023, they remain higher than pre-pandemic averages. Grade 12 results steadily rise following 2020. These changes to the mean results over time, and in response to the pandemic, can be further understood when analysed across the performance distribution.

[17] Grade 10 is the first grade of the FET phase. Thus, marks in Grade 11 and 12 are potentially being upwardly adjusted to ensure learners who have already failed in the phase (i.e. in Grade 10) are progressed to the following grade. This is due to the progression policy discussed above.

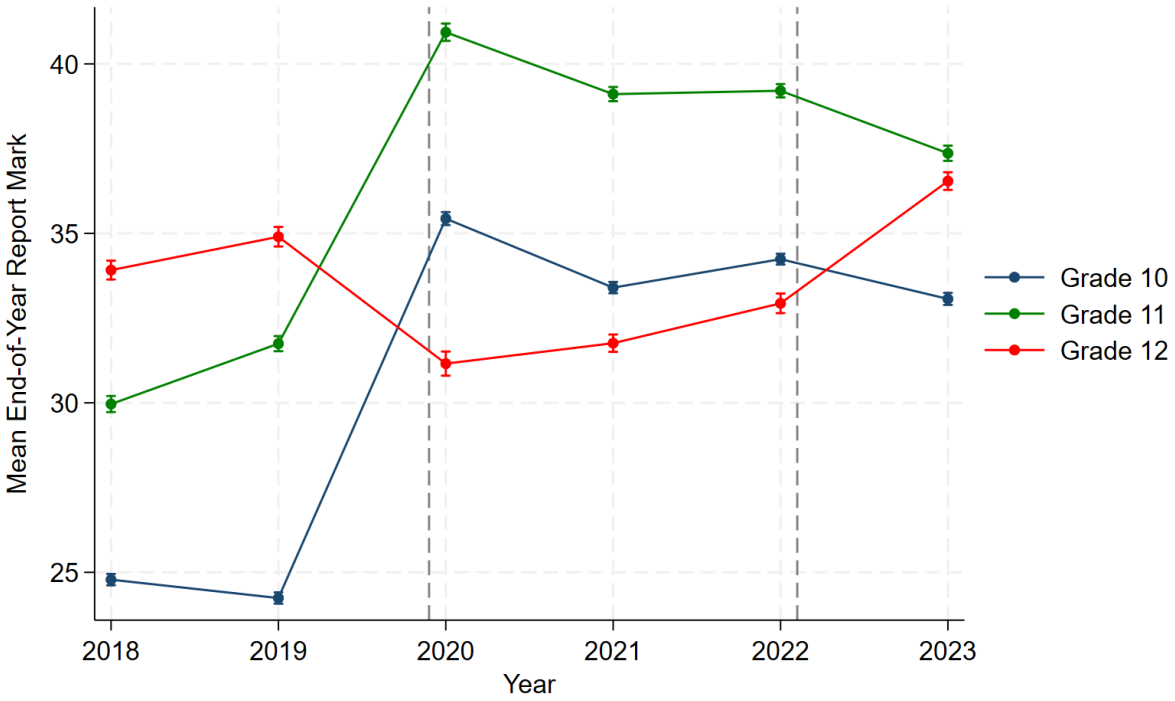


Figure 3. Mean End-Of-Year Mathematics Report Marks for Grades 10 - 12 between 2018 and 2023.

Figure 4 plots the kernel density estimates of Mathematics end-of-year report marks between 2018 and 2023 for Grades 10 – 12. As above, the Grade 10 and 11 distributions tell a similar story. The pre-Covid Grade 10 and 11 performance distributions (red and blue lines) are heavily skewed to the right, with the largest peak at the 10-20 percent mark range and a smaller peak at the 30 percent mark. However, in 2020, the distributions flatten out, with the peak at the 10-20 percent mark all but disappearing, while the density of learners achieving over 30 percent rises relative to pre-pandemic. This distributional shape remains in place in 2021 and 2022 but shows a slight shift in 2023 back toward the pre-pandemic distributional shape. Compared to Grades 10 and 11, the Grade 12 distributions are barely altered with the onset of the pandemic, although the 2020, 2021 and 2022 distributions do show a leftward shift relative to both the pre-pandemic (2018 and 2019) and post-pandemic (2023) distributions. These changes, which match those observed in the average report marks in Figure 3 above, are similarly observed in Mathematical Literacy.

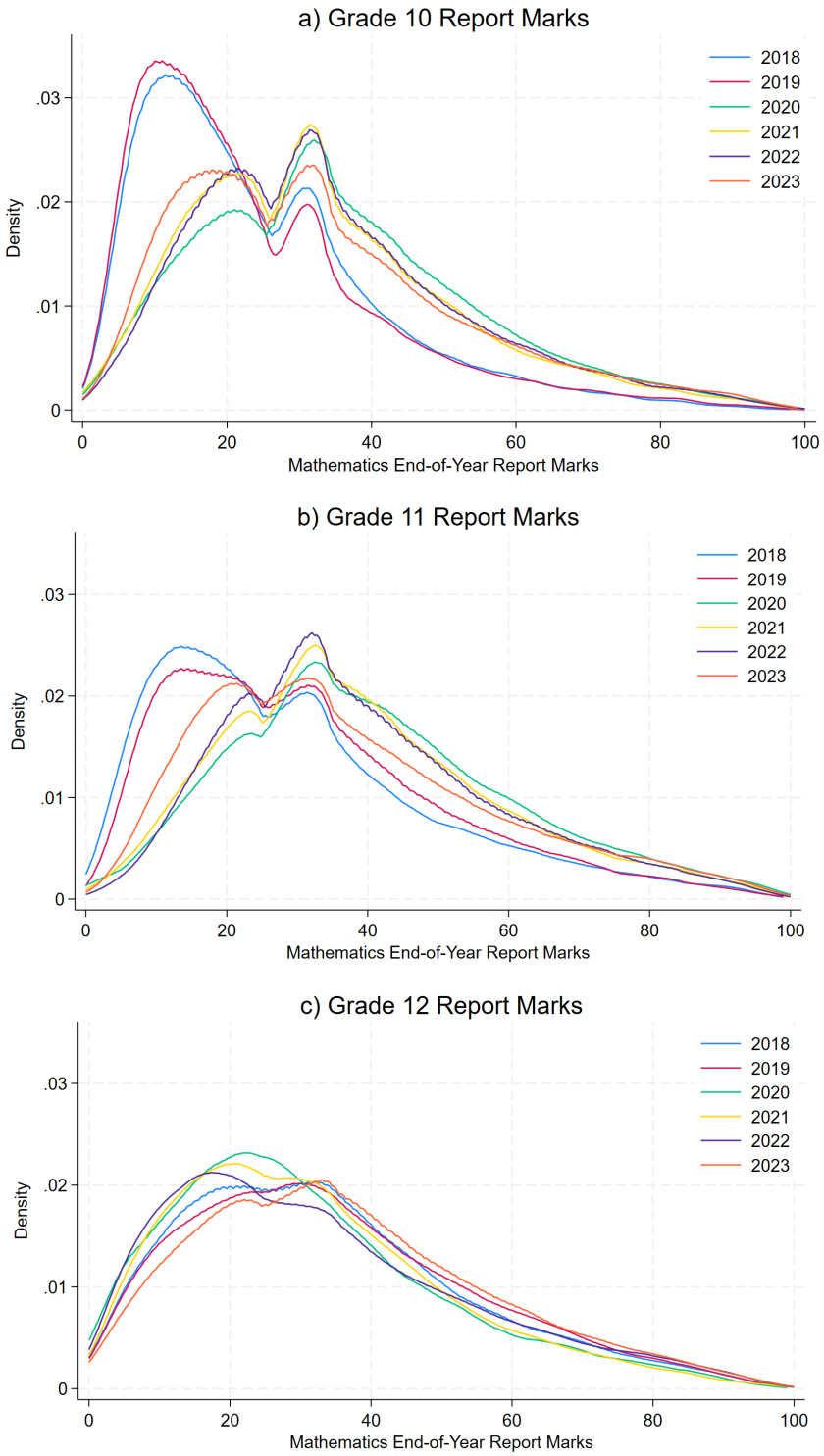


Figure 4. Kernel Density Estimates of Mathematics End-of-Year Report Marks across years for a) Grade 10, b) Grade 11 and c) Grade 12.

The Mathematical Literacy and English FAL report mark distributions are also altered by the pandemic, albeit in slightly different ways. The Mathematical Literacy report mark distributions, presented in the appendix, show that, between 2020 and 2022, the density of Grade 10 and 11 learners scoring below 30 percent declined, while the proportion scoring between 30 and 60 percent increased.

The 2023 distribution appears to have returned to the pre-pandemic distributional shape. The Grade 12 report mark distributions do not follow the distributional shape of the Grade 10 or 11 report marks, nor is there a consistent change in the distributions at the start of the pandemic. In contrast to both Mathematics subjects, the Grade 10 and 11 English FAL distributions, presented in the appendix, only change in 2020: shifting notably rightward across the performance distribution. The general increase in results in 2020 is observed in the Grade 12 English FAL distributions, while the 2021 and 2022 distributions lie to the left of the pre-pandemic distributions.

Figure 3 and 4 bring to light several preliminary dynamics relating to assessment between 2018 and 2023. The rightward shift of the Grade 10 and 11 Mathematics and Mathematical Literacy distributions in 2020 provides credence to Hoadley's (2023) concern that SBA results were inflated during the pandemic, as "SBAs" were given more weight in the promotion result. In addition, the peak and bunching at 30 percent (and 40 percent for Mathematical Literacy during the pandemic) in the Grade 10 and 11 distributions suggests that the FET phase pass requirements[18] influence final report marks. In fact, Figure 4 intuitively suggests that it was the Grade 10 and 11 learners who would have ordinarily achieved less than 30 percent who were being pushed into higher achievement brackets, perhaps in a particular effort to promote these learners to Grade 12 during the pandemic. This finding, which is in line with existing evidence that repetition and dropout rates declined in South Africa (Wills & Qvist, 2023; Wills & Van der Berg, 2024), is explored further in the following section.

The Grade 12 means and distributions across the years do not follow the same trends as those of Grade 10 and 11. The leftward shift of the 2020-2022 Grade 12 distributions, relative to the 2018 and 2019 distributions, may suggest a decline in performance in 2020. If assessment standards remained the same as in pre-pandemic years, a decline in performance would be expected given school closures and rotational attendance, little prior exam practice for 2021 and 2022 matric learners, and widespread evidence of learning losses (Ardington et al., 2021; Donnelly & Patrinos, 2022). In addition, the 2021 and 2022 Grade 12 cohorts may have been academically weaker, given the leniency in assessment (Van der Berg & Shepherd, 2009) and increased rates of progression (Wills & Qvist, 2023) during the pandemic. However, despite a leftward shift, the Grade 12 distributions show considerably less pandemic-related changes compared to the Grade 10 and 11 distributions. This may reflect 1) the lack of promotion and progression policies, 2) the high level of external moderation and standardisation of assessments in Grade 12, and 3) academically stronger learners in Grade 12, given high levels of dropout at the Grade 10 and 11 level. Evidently, many factors could explain the dynamics observed

[18] A learner passes the academic year if they receive at least 40% for three subjects, including an official home language, and at least 30% for the remaining three subjects.

in the unbalanced sample. Thus, while the descriptive trends illustrated in Figure 3 and 4 provide valuable insights, they further raise important questions about the signalling quality of SBA results, which can only be addressed by analysing the performance trajectories of the same learners over time.

5.2 GRADE 11 TO 12: POOR, PASSED AND PLUMMETED

> "Policymakers need to understand that assessment is a movie, not a snapshot. What we really want to know is how students are progressing over time, not where they stand on a particular day." – Pellegrino (2002)

Grade 11 assessment results play a pivotal role in determining progression to the final year of school and serve as early indicators of a learner's ability to obtain the NSC. They are also crucial for tertiary institutions, which often use these results to grant conditional admission to certain programmes. As such, it is critical to understand how learners progress from Grade 11 to 12, and if this was affected by the pandemic. Figure 5 below presents the kernel density estimates of the same[19] learners' report marks, in Grade 11 and 12, across 5 cohorts. The similarity between the Grade 11 and 12 distributions for the fully pre-Covid 2018/19 cohort (the first graph in each column) is a reassuring feature across all subjects. This potentially suggests that learners' Grade 11 marks pre-pandemic were good quality signals of their performance in Grade 12 the following year. In contrast, the Grade 11 and 12 distributions diverge from one another with the onset of the pandemic. Specifically, the Grade 11 report mark distribution is shifted to the right of the Grade 12 report mark distribution for Mathematics and English FAL for both the 2020/21 and 2021/22 cohorts. This divergence is most stark in Mathematics, as was seen in Figure 4 above.[20] However, the changes observed in SBA results were not uniformly distributed across the schooling system.

[19] These learners have progressed on track from Grade 11 to 12.
[20] The report mark distributions across cohort years for each grade, as reflected in Figure 4, can be found for the balanced panel of Grade 11 to 12 learners in the appendix.

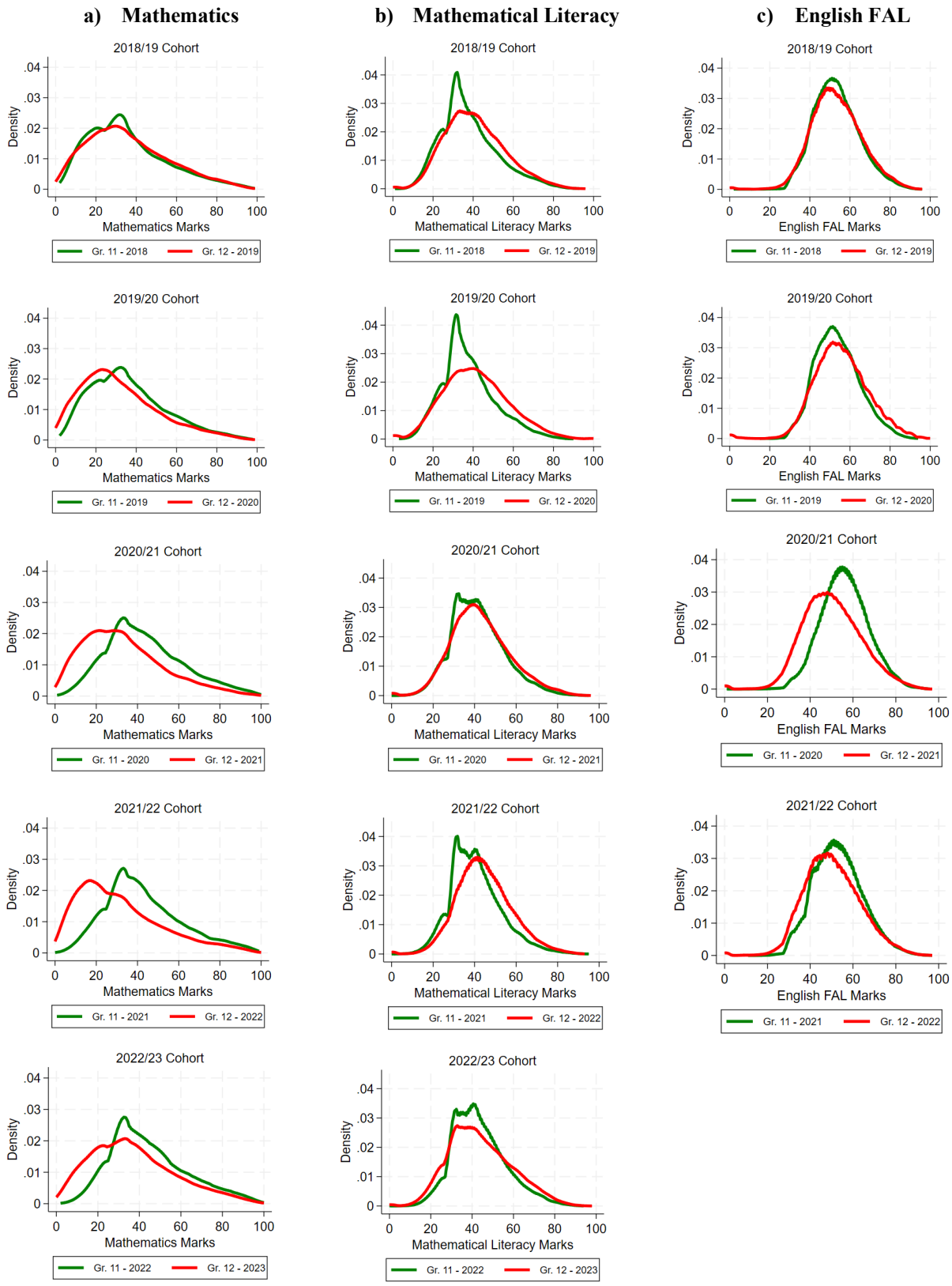


Figure 5. Kernel Density Estimates of End-of-Year Report Marks for Grade 11(green line) and Grade 12 (red line) across cohorts for a) Mathematics, b) Mathematical Literacy and c) English FAL

Existing evidence has highlighted the heterogenous impact of the pandemic on South Africa's unequal education system (Van der Berg & Böhmer, 2025; Wills & Van der Berg, 2024). Figure 6 below presents the kernel density estimates of Grade 11 and 12 Mathematics report marks across years by schools' fee status. Figure 6a shows that the changes in the report mark distribution observed in Figure 4 and 5 are predominantly driven by changes occurring in no-fee schools rather than fee-paying schools. While Grade 11 report marks shift toward the higher end of the performance distribution for both fee-paying and no-fee-paying schools, the shift is considerable amongst learners at no-fee schools. Mathematical Literacy results for Grade 11 (see appendix) similarly show that the change in the performance distributions at the onset of the pandemic occurred at no-fee schools. The Grade 12 performance distributions at no-fee schools shifts more toward the lower end of the performance distribution in 2020, 2021 and 2022 than the performance distributions at fee-paying schools, although the difference is almost negligible. It is concerning to see both how starkly the results of learners attending no-fee schools were influenced by the pandemic, and how at odds each cohort's Grade 11 and 12 results were during this time.

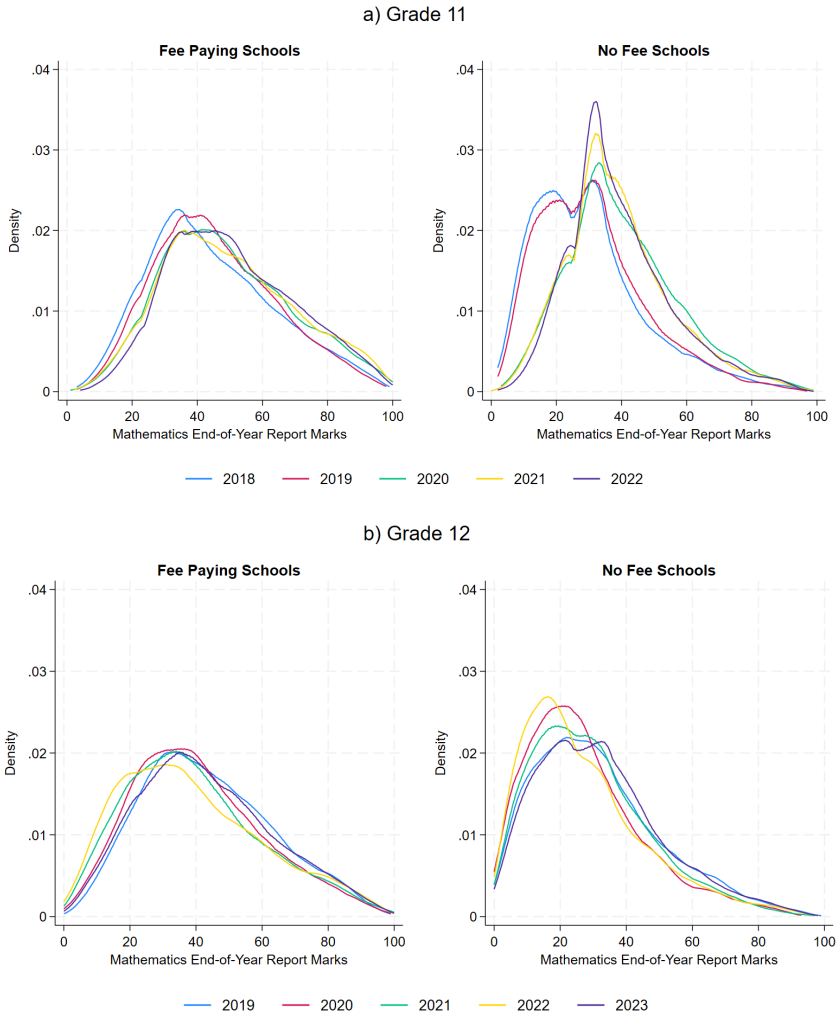


Figure 6. Mathematics Report Marks by Year and School Fee Status for a) Grade 11 and b) Grade 12

At face value, Figure 6 suggests that Grade 11s at no-fee schools were receiving more inaccurate signals regarding their Mathematics performance and likelihood of passing Grade 12 Mathematics, than their fee-paying counterparts. Figure 7 thus explores the Grade 12 Mathematics pass rates (using report marks) between fee and no-fee schools, amongst all those who achieved over 50 percent[21] in Grade 11. The plot shows that the proportion of learners failing Grade 12 Maths, after achieving over 50 percent in Grade 11, increased in the pandemic years, for learners at both fee and no-fee schools. However, while only 4% and 5% of learners at fee-paying schools who achieved over 50 percent in Grade 11 in 2020 and 2021 respectively went on to fail Grade 12, 16% and 17% of learners in no-fee schools who achieved over 50 percent in Grade 11 in 2020 and 2021 went on to fail Grade 12. Effectively, in 2020 and 2021, 1 in 6 learners at no-fee schools may have thought they would be eligible to apply to a tertiary degree requiring Mathematics, only to fail the subject the following year. Fortunately, this proportion seems to drop for the 2022 Grade 11s, suggesting that these inaccuracies were potentially isolated to the pandemic years.

Ultimately, Figures 2 - 4 cast doubt on the reliability of Grade 11 report marks as a predictor of Grade 12 performance, particularly in no-fee schools during the pandemic. These findings underscore the need for robust, equitable assessment practices to ensure that all learners receive accurate signals about their academic progress. To further investigate the misalignment between Grade 11 and 12, the following section uses learners' NSC (matric) results as an external benchmark of achievement.

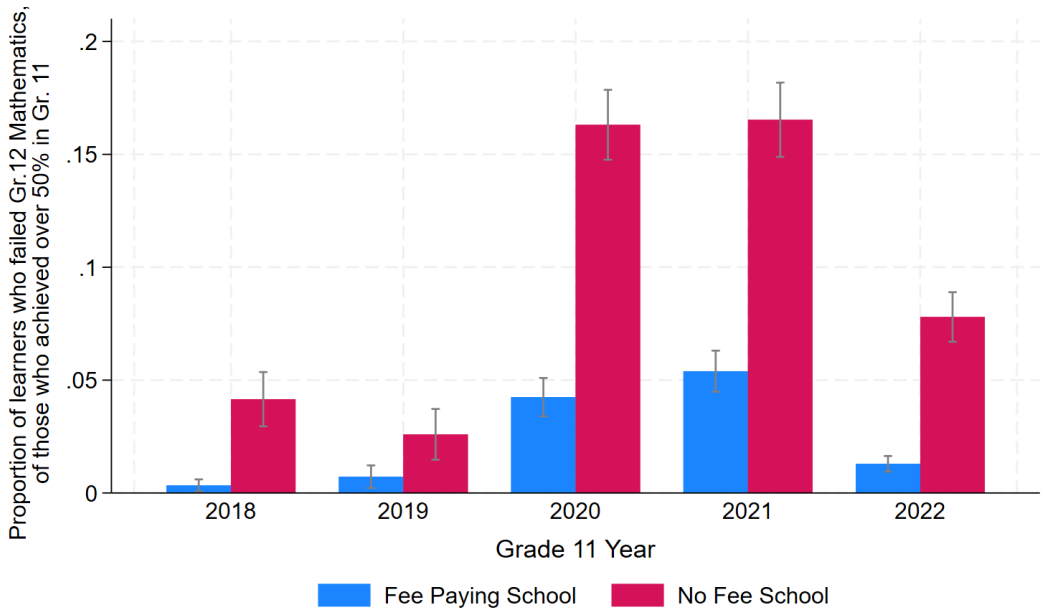


Figure 7. Proportion of Learners who Failed Grade 12 Mathematics after achieving over 50% in Grade 11 Mathematics

[21] A grade of over 50% in Mathematics makes one eligible to apply to some degrees that require Mathematics at certain tertiary institutions. It is thus significant for a Grade 11 learner to be incorrectly told that they had this potential.

5.3 WHAT CAN THE NSC RESULTS TELL US?

Given the notable divergence between the Grade 11 and 12 school results during the pandemic, the reliability and validity of SBA needs to be better understood. Figure 8 plots the mean Mathematics Grade 11 and 12 report marks in each school against the school's average NSC result for each year. Each dot represents a school, and lowess curves plot the linear relationship between schools' report marks and NSC results. The figure shows that, pre-pandemic, Grade 11 and 12 report marks show a near-perfect alignment with NSC results. However, for Grade 11 marks, the lowess line falls below the diagonal in 2020 and 2021, suggesting Grade 11 marks were inflated relative to the NSC marks. In addition, the dots seem less tightly clustered in 2020 and 2021 relative to 2018 and 2019, suggesting poorer correlation between Grade 11 report marks and NSC results. In contrast, the Grade 12 lowess curves for 2020-2022 lie above the diagonal line, suggesting that the Grade 12 report marks were "too strict", and consistently lower than NSC results during the pandemic years. Reassuringly, the dots for Grade 12 are tightly packed around the linear relationship, suggesting a high degree of correlation between the NSC result and Grade 12 report marks. The reliability of Grade 11 and 12 assessments is thus explicitly explored in Table 4 below.

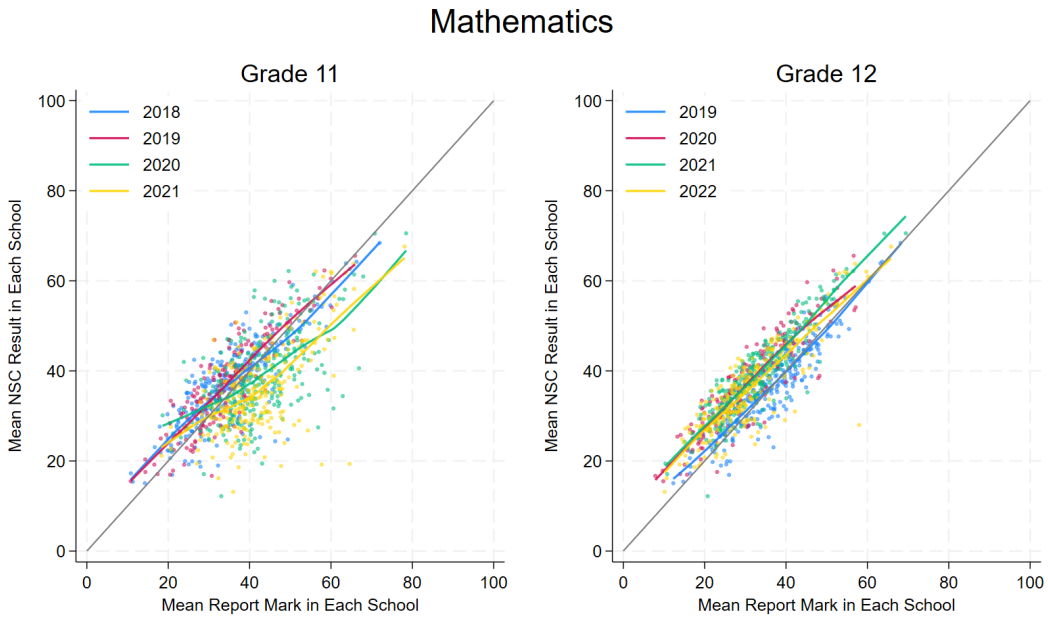


Figure 8. Scatterplot and Lowess Trend Line of School Level Grade 11 and Grade 12 Report Marks and NSC Results for Mathematics.[22]

Table 4 reports the Spearman Rank correlations between learners' end-of-year report marks and their final NSC result within each year and subject. In line with Van der Berg et al.'s (2023:20) findings for the Eastern Cape province, report mark correlations to NSC results are highest in Grade 12

[22] Only schools with more than 15 learners per year, grade and subject are included in the school-level analysis.

compared to Grade 11. This is to be expected due both to the higher degree of curriculum overlap between the NSC and Grade 12 curriculum and to the higher degree of assessment standardisation and moderation in all Grade 12 assessments. When comparing correlations across subjects, Mathematics school assessments are more highly correlated than Mathematical Literacy and EFAL marks to the NSC result, across years and grades. In fact, Grade 12 Mathematics report marks are considerably highly correlated to the NSC result, with correlations above 0.9 across all four years. This may be indicative of Mathematics being a more "exact discipline", with clear marking memorandums and past paper examples, and thus more consistency across years and grades. However, the pandemic disrupts this consistency across subjects, most notably in Grade 11.

The correlations between Grade 11 report marks and NSC results, shown in Table 4, weaken in 2020 and 2021 compared to pre-pandemic. The decline in assessment reliability in 2020 and 2021 is consistent across subjects in Grade 11. In contrast, the correlations of the Grade 12 report marks to the NSC results appear largely unaffected by the pandemic. This perhaps reiterates the dual argument that Grade 12 assessment programmes were not changed during the pandemic years and are relatively well moderated and standardised compared to Grade 11. However, the correlations for the Grade 12 Mathematical Literacy report marks to the NSC are slightly lower in 2021 and 2022, while the correlations of Grade 12 EFAL marks to the NSC weaken only in 2020. While there is evidence of a negative impact of the pandemic on school assessment reliability, with differences by grade and subject, the impact of the pandemic on assessment gaps is less clear.

Table 4. Individual Spearman Rank Correlations between Learners' Report Marks and Final NSC Results by Year, Grade and Subject.

	Year	Mathematics	Math. Literacy	English FAL
Grade 11	2018	0.85	0.72	0.74
	2019	0.88	0.73	0.75
	2020	0.76	0.66	0.67
	2021	0.76	0.66	0.68
Grade 12	2019	0.94	0.89	0.81
	2020	0.92	0.88	0.74
	2021	0.93	0.85	0.81
	2022	0.93	0.84	0.83

The gaps presented in Figure 9 below indicate the average difference, across years and subject, between a learner's NSC result and their report marks in Grade 11 and 12. A positive value indicates that, on average, school marks were lower than the NSC result, while a negative value indicates that school marks were higher than the NSC result. Figure 9(a) shows that, for Mathematics, the 2018 and 2019 Grade 11s had small gaps between their NSC results and report marks, of 1.40 and 2.14

respectively. However, in 2020 and 2021 the gaps are wider and turn negative, suggesting that, on average, learners' Grade 11 report marks during the pandemic years were higher than their NSC results the following year. In contrast to Mathematics, the Mathematical Literacy and English FAL Grade 11 report marks are lower, on average, than NSC results and the gap between these results is smallest in 2020. Thus, while there is potential evidence to support the argument that Grade 11 school-based marks were inflated during 2020 and 2021, this change appears isolated to Mathematics (at least at the mean).

The gaps between Grade 12 school report marks and NSC results are similarly presented in Figure 9(b). The gaps are positive in every year and subject, indicating that Grade 12 report marks are, on average, consistently lower than NSC results. These changes map onto the divergence in distributions observed in the kernel density estimates of school report marks and NSC results[23]. However, while the average gap between the Grade 12 report mark and NSC result for Mathematics is negligible in 2019, at 0.53, this gap widens considerably to approximately 6 percentage points in 2020 and 2021. Similarly, the gaps for English FAL widen to approximately 6.7 in 2021 and remain this large in 2022. In other words, Mathematics and English FAL school marks were lower than their NSC results, particularly during the pandemic. It seems unlikely that Grade 12 assessments would have been made 'stricter' during the pandemic years. Thus, while this finding is explored in more detail in the discussion, further research may be required to understand the underlying mechanisms behind this change.

This analysis of correlations and gaps reiterates key differences in assessment quality between grades and subjects. The consistently higher correlations for Mathematics may reflect the subject's standardized nature, while Mathematical Literacy and English FAL may be more susceptible to inconsistencies in assessment practices. The comparison to the external NSC examination strengthens the theory that the rightward shift in the Grade 11 Mathematics result distributions (see Figure 4) and large increases in the mean result (see Figure 3) were not driven by improved ability or selection bias, but rather pandemic related mark inflation. However, as shown in Figure 6 and Figure 7, this upward movement in Grade 11 Mathematics marks was predominantly occurring in no-fee-paying schools. The following section thus delves deeper into disparities in assessment quality, focusing on inequalities by school fee status, gender and academic performance.

[23] See appendix for detailed kernel density estimates of Grade 11 report marks, Grade 12 report marks and NSC results across cohorts and subjects.

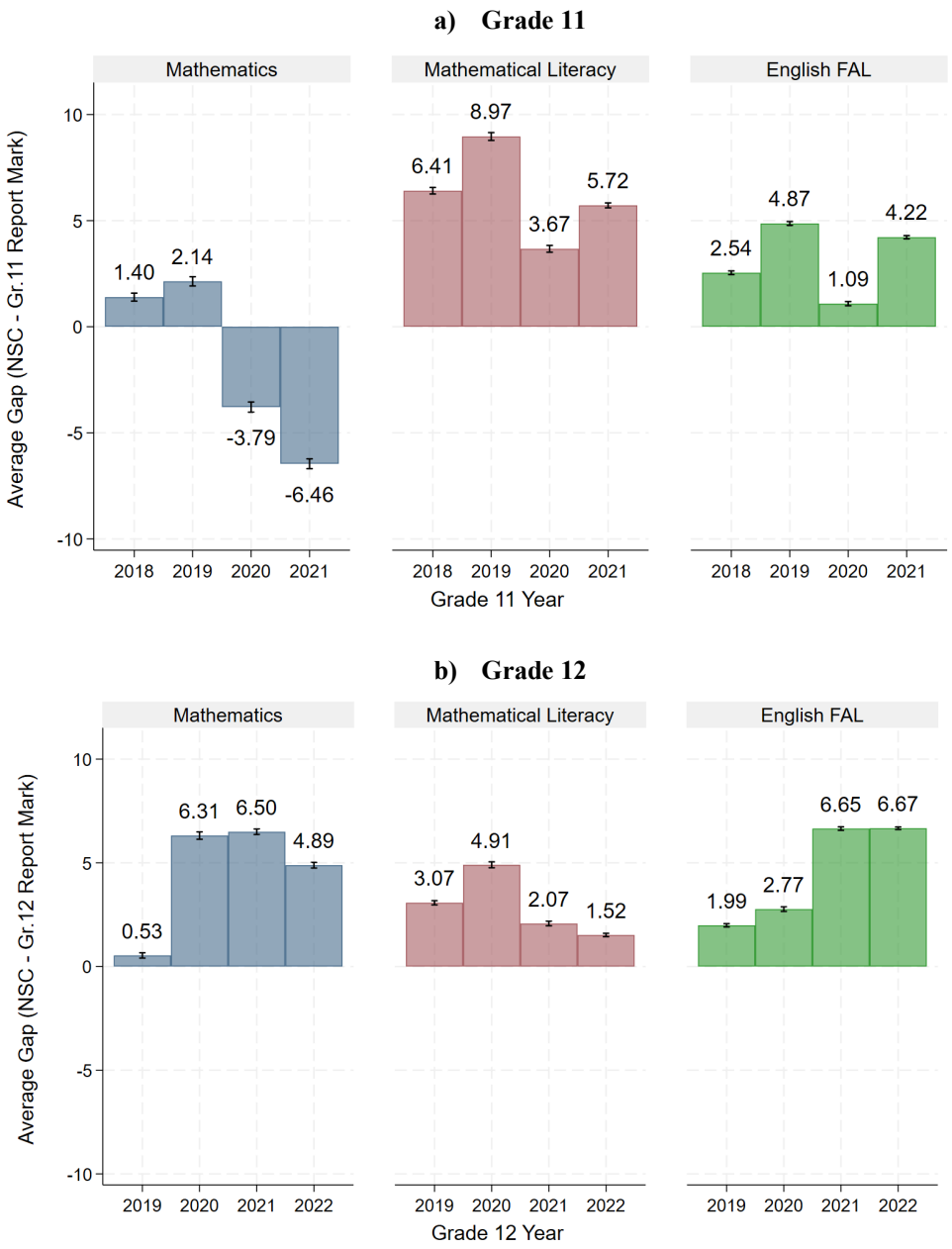


Figure 9. Average Gaps between NSC results and learners' report marks across years and subjects for a) Grade 11 and b) Grade 12.

5.3.1 Socioeconomics and Assessment Quality

Differences in report mark distributions between fee and no-fee schools during the pandemic raise concerns about the quality of assessments at no-fee schools. Table 5 presents the average intra-school correlation coefficients of report marks to NSC results by school fee status, subject, year and grade. The table shows that no-fee schools consistently have lower correlation coefficients between Grade 11 report marks and NSC results compared to fee-paying schools, and that these differences are statistically significant for all three subjects. Furthermore, the correlations in Grade 11 appear to worsen in the pandemic years, notably more so for no-fee schools than for fee-paying schools. This

corroborates Van der Berg et al.'s (2023:20) findings for the 2019 Eastern Cape matric cohort, in which Quintile 4 and 5 schools had higher correlations between NSC results and school marks in Mathematics in Grades 9 to 11. However, like this Eastern Cape cohort, differences in the correlations between fee and no-fee schools are of negligible magnitude in Grade 12 Mathematics (Van der Berg et al., 2023:20).

In contrast to Grade 11, the Grade 12 report mark correlations to the NSC are not statistically different between fee and no-fee schools for Mathematics. This may, once again, speak to the exactness of the mathematics discipline. On the other hand, English FAL Grade 12 report mark correlations to the NSC are statistically different across school fee-status in every year. These large and significant differences in correlations between no-fee and fee-paying schools for English FAL may be driven by differences in the subject-specific samples, which show that less than 15% of English FAL learners attend fee-paying schools, and of them 88% are English or Afrikaans Home Language (compared to around 50% in Mathematics and Mathematical Literacy samples at fee-paying schools) [24]. This means that most learners taking English FAL at fee-paying schools are Afrikaans, potentially suggesting that the sample of English FAL learners at fee-paying schools are more distinct from their no-fee school counterpart than is the case for Mathematics and Mathematical Literacy. While there are clear differences in the reliability of assessments across school fee-status, the differences in validity are less definitive.

Table 5. Means of Within-School Correlations between Report Marks and NSC marks by Subject and School Fee-Status.

		Mathematics			**Mathematical Literacy**			**English FAL**		
		Fee-Paying	No-Fee		Fee-Paying	No-Fee		Fee-Paying	No-Fee	
Grade 11	2018	0.868	0.823	***	0.804	0.695	***	0.875	0.761	***
	2019	0.878	0.837	*	0.812	0.747	***	0.885	0.789	***
	2020	0.832	0.760	***	0.793	0.657	***	0.878	0.685	***
	2021	0.847	0.749	***	0.793	0.670	***	0.867	0.722	***
Grade 12	2019	0.927	0.936		0.897	0.872	***	0.894	0.834	***
	2020	0.915	0.918		0.895	0.889		0.906	0.816	***
	2021	0.929	0.931		0.897	0.878	**	0.898	0.852	***
	2022	0.935	0.930		0.877	0.871		0.907	0.859	***

Note: Asterisks indicate whether the difference in the mean within-school correlations between fee and no fee schools is significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ based on two-tailed t-tests.

[24] Descriptive statistics by school-fee status are presented across cohorts and subjects in the appendix.

The average gaps between NSC results and school report marks across school fee status are given in Table 6 below. For Mathematics and Mathematical Literacy, the gaps between NSC results and Grade 11 results are statistically significantly larger at no-fee schools compared to fee-paying schools before the pandemic. This suggests lower validity, or perhaps large improvements from Grade 11 to the NSC for learners at no-fee schools, even during times of stable assessment practices. While the gaps in Grade 11 and Grade 12 Mathematics remain larger in absolute value for learners at no-fee schools in 2020 and 2021 (although not significantly so), the opposite is true in Mathematical Literacy. In 2020, fee-paying schools had significantly larger gaps in Grade 11 and 12 in 2020 and 2021 compared to no-fee schools. In contrast to the numeracy subjects, the average gap between the NSC and English FAL report marks was larger at fee-paying schools across both grades and in most years. The differences in the gaps by school fee-status do not clearly indicate which school type produces less valid SBAs, and this relationship is thus further explored in the regression analysis. However, in addition to school-related socioeconomic factors, learners' characteristics may play into how they perform across different assessments.

Table 6. Mean School-Level Gaps between Report Marks and NSC Results by Subject and School Fee-Status

		Mathematics			**Mathematical Literacy**			**English FAL**		
		Fee-Paying	No-Fee		Fee-Paying	No-Fee		Fee-Paying	No-Fee	
Grade 11	2018	0.937	**2.278**	*	5.695	**7.773**	**	**3.693**	2.781	
	2019	0.514	**3.407**	***	6.896	**10.026**	***	**5.630**	4.997	
	2020	-3.296	**-3.622**		**5.820**	3.255	***	**3.367**	1.034	**
	2021	-5.525	**-6.086**		**6.063**	5.994		**6.834**	4.174	***
Grade 12	2019	**0.602**	0.178		**4.874**	2.384	***	**3.112**	2.057	
	2020	4.705	**7.113**	***	4.645	**5.294**		**4.918**	2.722	*
	2021	5.739	**6.585**		**4.467**	1.469	***	5.908	**6.611**	
	2022	4.712	**5.014**		**3.153**	1.150	***	**7.322**	6.724	

Note: Bold indicates the larger gap in absolute value. Negative values indicate the report mark was inflated, positive values indicate the report mark was too "strict". Asterisks indicate whether the difference in the mean gap between fee and no fee schools is significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ based on two-tailed t-tests.

5.3.2 Performance and Assessment Quality in Mathematics

The hypothesis that schools with higher performers, on average, have more reliable assessments is explored for Mathematics in Figure 10 below. The scatterplots plot schools' average NSC result against the Spearman Rank correlation between their average NSC result and school report mark by

grade. The Lowess regression lines in Figure 10(a) indicate that the correlation between school report marks and NSC results in Grade 11 is positively related to schools' NSC results. For Grade 12, the NSC result is not as distinctly related to the correlation between report marks and NSC results, as shown by the flat lowess curves in 2019 and 2020 in Figure 10(b). This suggests that while better performing schools in the NSC produce higher correlations between their school assessment results and NSC results in Grade 11 (particularly for Grade 11s in 2020 and 2021), this is not the case in Grade 12, with most schools producing very similar correlation coefficients. This is in line with the findings in Figure 5 above. In addition to changes in reliability along the performance distribution, the validity of an assessment may vary depending on a learner's performance.

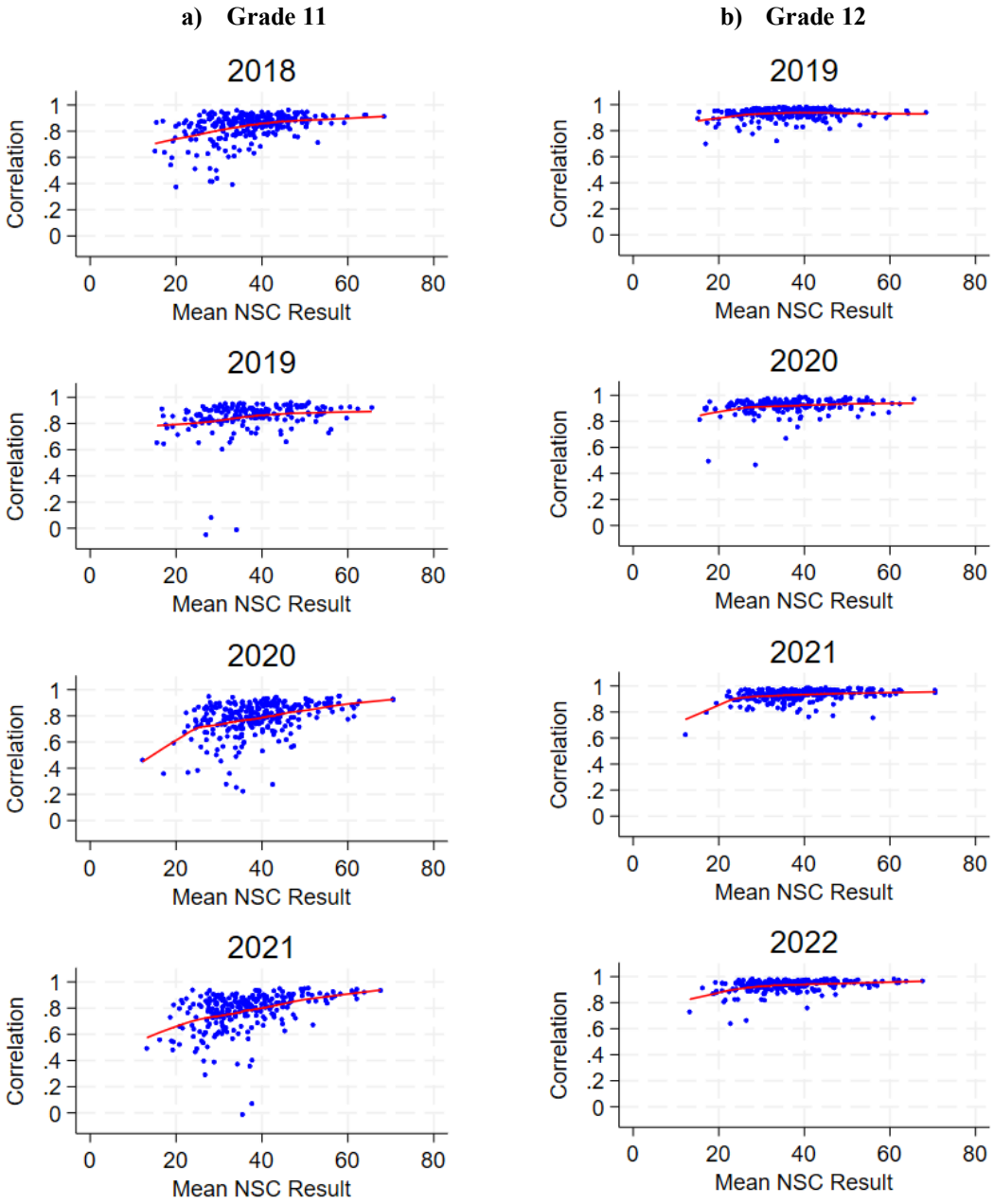


Figure 10. Relationship between Schools' Mean NSC Result and the Within-School Correlation between Report Marks and NSC Results for Mathematics.

The distribution of Grade 11 Mathematics report marks, as shown in Figure 4 indicated that it may be learners at the bottom end of the performance distribution who were being more leniently marked from 2020 onwards, compared to pre-pandemic. As such, distributions of the gaps between learners' Mathematics report marks and NSC results are presented in Figure 11 by learners' NSC performance tercile[25] and grade for the years before and during the Covid-19 pandemic. The distribution of the gaps pre-pandemic (blue line) follows an approximately normal distribution with a mean of 0, across all three performance terciles and in both Grade 11 and Grade 12. However, in Grade 11 2020 and 2021, the distributions of the gaps between report marks and NSC results are less normally distributed and differ across terciles. The lowest performing NSC tercile (Tercile 1) has more negative gaps than the highest performing tercile, suggesting that Grade 11 report marks were potentially inflated by more in the lowest tercile compared to the highest performing tercile. In contrast, Grade 12 report marks appear to be under-estimating the NSC results between 2020 and 2022, more so for the highest performing tercile. These results suggest that learners' academic ability - if one considers the NSC result a reliable reflection of academic ability - does influence both the reliabilty and validity of SBAs, particularly during the pandemic. However, given that research has found evidence of discrepancies between internal and external assessments by gender, both internationally (Lindahl, 2007; Lavy, 2008; Lavy & Sand, 2015) and in South Africa's Eastern Cape province (Van der Berg et al., 2023), the assessment accuracy measures explored here should be further investigated by gender.

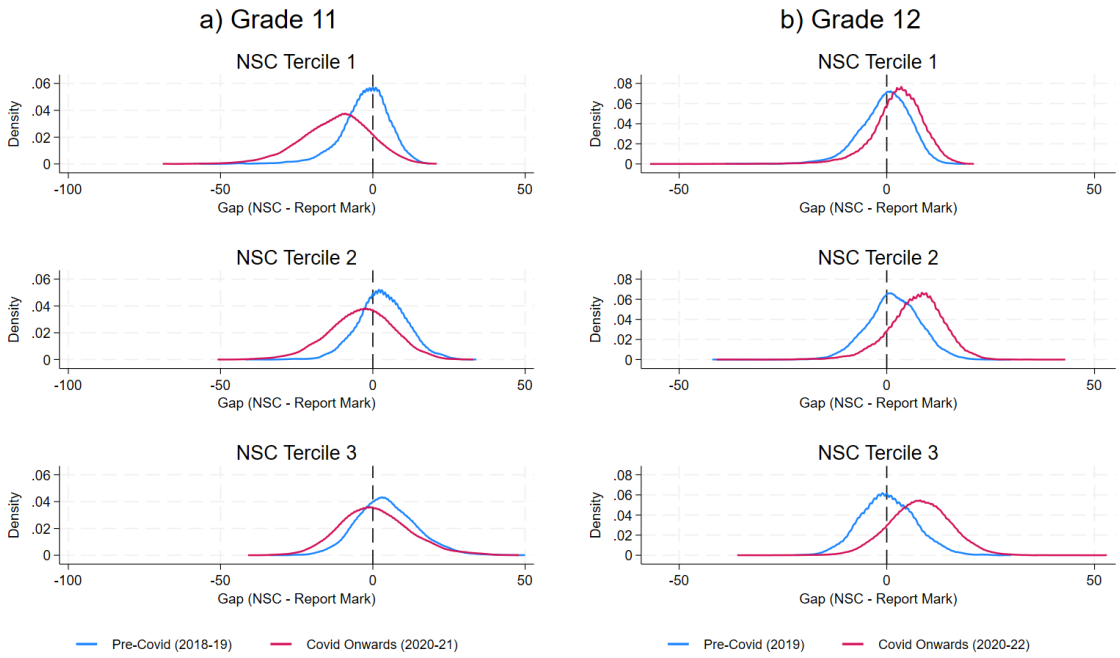


Figure 11. Kernel Density Estimates of the Individual Gaps between Learners' report marks and NSC results in (a) Grade 11 and (b) Grade 12 across NSC Performance Terciles.

[25] The NSC performance tercile splits each year, grade and subject group into three equal groups along the performance distribution. The cut offs for each group are summarized in Table 19 in the appendix.

5.3.3 Gender and Assessment Quality

While gender inequality is not the focus of this paper, it seems necessary to briefly explore given its prevalence in the literature (Lindahl, 2007; Lavy, 2008; Lavy & Sand, 2015; Van der Berg et al., 2023). Table 7 shows that before the pandemic, female learners had slightly higher correlations between their Grade 11 report marks and NSC results than male learners, across all three subjects. While this stays in place for English FAL in both 2020 and 2021, it does not hold for Mathematics in 2021 nor for Mathematical Literacy in 2020. Similarly, while female learners have higher correlations between their Grade 12 marks and NSC results in English FAL across all four years, the opposite is true for Grade 12 Mathematics. These correlations suggest that female learners' SBA results are more consistent across time and grades for English FAL. In contrast, while female learners produce more reliable report marks in Grade 11 Mathematics, they are less reliable than Male learners' in Grade 12. The observed gender differences may reflect various biases that disadvantage female learners in mathematics while disadvantaging male learners in language subjects.

Table 7. Individual Spearman Rank Correlations between Learners' Report Marks and Final NSC Results across Cohort Years and Subject by Gender

		Mathematics		Mathematical Literacy		English FAL	
		Male	Female	Male	Female	Male	Female
Grade 11	2018	0.843	0.860	0.722	0.726	0.707	0.748
	2019	0.869	0.888	0.720	0.737	0.727	0.751
	2020	0.749	0.765	0.670	0.665	0.629	0.673
	2021	0.781	0.753	0.652	0.670	0.643	0.686
Grade 12	2019	0.940	0.930	0.880	0.890	0.790	0.820
	2020	0.930	0.920	0.880	0.870	0.730	0.730
	2021	0.940	0.930	0.850	0.840	0.790	0.810
	2022	0.940	0.920	0.840	0.840	0.810	0.830

Plots of the gaps between school report marks and NSC results by gender are shown in Figure 12 below. In Mathematics, male learners had larger gaps than female learners' pre-pandemic. However, from 2020 onwards, in both Grade 11 and 12, female learners have statistically significantly larger gaps between their school result and NSC result than male learners. In contrast, male learners consistently have larger gaps than female learners for Mathematical Literacy, across grades and years, except in Grade 12 2020. Given evidence of gender differences in assessment accuracy at the mean, it is worth investigating whether gender remains a significant correlate of assessment reliability and validity in a multivariate context.

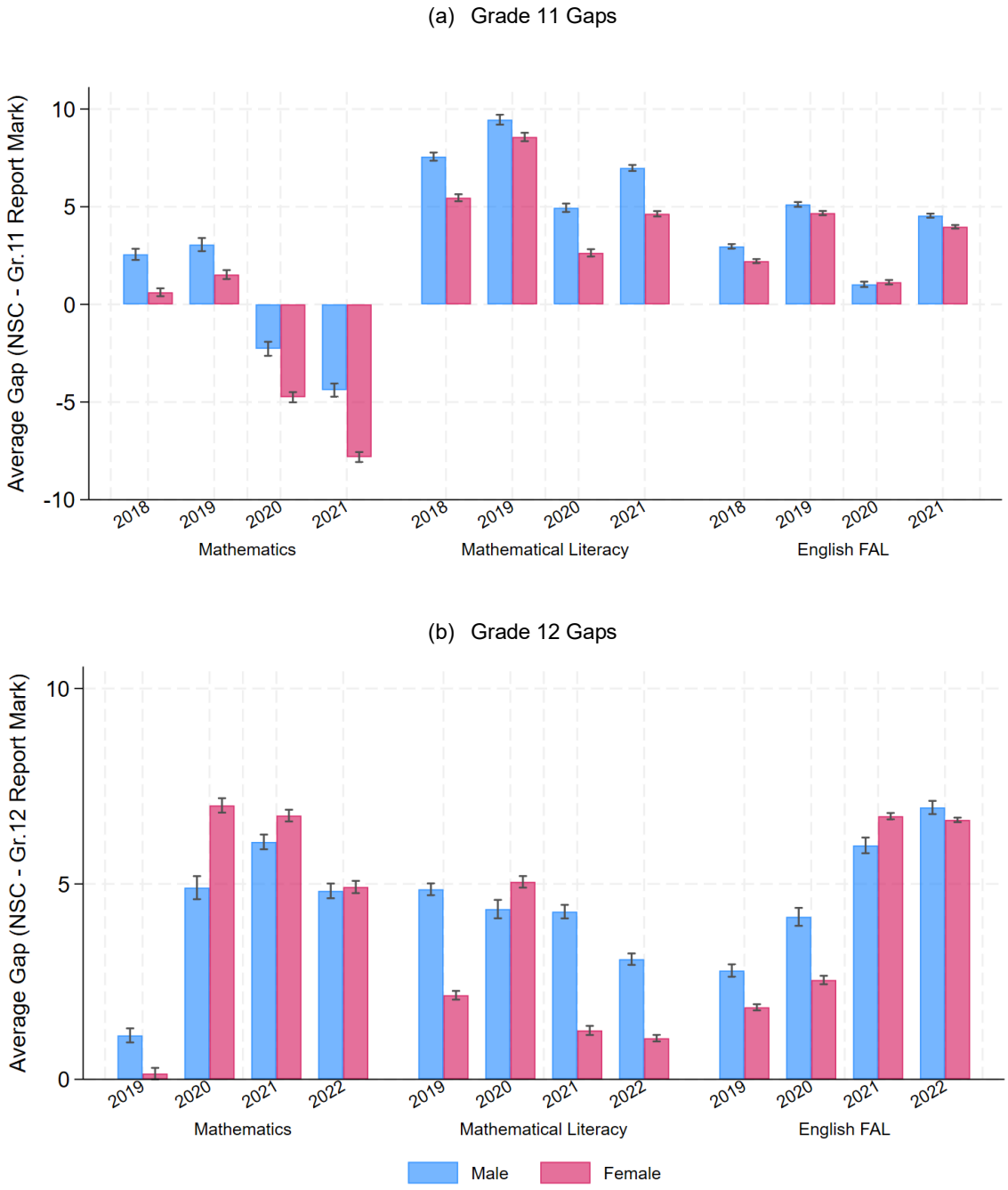


Figure 12. Gaps between Report Marks and NSC Results by Gender for (a) Grade 11 and (b) Grade 12 across Subjects.

6 REGRESSION ANALYSIS

6.1 UNDERSTANDING INTRA-SCHOOL CORRELATIONS

Given the strong differences in school-level correlations by school fee status and NSC performance, the factors influencing intra-school correlations between a report marks and NSC results are explored in Table 8 below. Even in a multivariate context, no-fee schools show significantly lower correlations between school and NSC results in Grade 11, across all three subjects. However, there is no significant relationship between school fee status and correlations for Grade 12. This implies that in no-fee schools, internal report marks in Grade 11 are less consistent with NSC results compared to fee-paying schools, but not in Grade 12.

The year of the internal assessment influences assessment reliability. The correlations between Grade 11 marks and NSC results are unaffected for Mathematics if the school assessment was in 2019 relative to 2018, while the correlations are stronger for Mathematical Literacy and EFAL in 2019. However, if the Grade 11 report marks were captured during the pandemic years of 2020 and 2021, this significantly decreased the correlation to the NSC result relative to that in 2018, across all three subjects, but particularly for Mathematics. This reflects the pandemic's impact on the consistency of internal grading practices in Grade 11. Additionally, if the Grade 12 report marks were captured in 2020, this is associated with a decrease in the correlation of the school's grade 12 mark with the NSC result, but only for Maths and EFAL. Furthermore, while the 2021 and 2022 Grade 12 years do not show a large or statistically significant effect on the correlations in the numeracy subjects, the Grade 12 correlations for EFAL strengthen during the post-2020 years.

A higher proportion of learners scoring below 60% in the NSC is strongly associated with weaker correlations across all subjects and grades. In Mathematics, correlations are significantly lower for Grade 11 (-0.326) and Grade 12 (-0.129), indicating that schools with more low-performing learners exhibit less consistent internal grading practices. Similar patterns are observed in Mathematical Literacy and English FAL, although the magnitude of the coefficients is smaller for English FAL. Additionally, a larger proportion of learners scoring below 60% has a greater negative effect on the Grade 11 correlations than the Grade 12 correlations. This reiterates that Grade 11 school assessment is less consistent or accurate, than that in Grade 12. The proportion of learners in a school who are English or Afrikaans home language, and the school's pupil-teacher ratio, has no significant effects on the correlations between school and NSC results. While this analysis of correlations highlights discrepancies in assessment reliability, the magnitude of gaps between internal and external marks provides further insights into assessment validity in a multivariate context.

Table 8. OLS Regression Results on the Correlations between Grade 11 and 12 Report Marks and NSC Results.

	Mathematics		Math. Literacy		English FAL	
	(1) Gr. 11 Corr	(2) Gr. 12 Corr	(3) Gr. 11 Corr	(1) Gr. 12 Corr	(2) Gr. 11 Corr	(3) Gr. 12 Corr
No-Fee School	-0.040***	0.003	-0.041***	0.002	-0.080***	-0.038
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
2019/20	0.005	-0.020***	0.035**	0.008	0.018*	-0.019*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
2020/21	-0.068***	-0.008*	-0.031**	0.002	-0.067***	0.015**
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
2021/22	-0.065***	-0.004	-0.027**	-0.009*	-0.040***	0.021***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
% NSC < 60	-0.326***	-0.129***	-0.264***	-0.124***	-0.114***	-0.059***
	(0.03)	(0.01)	(0.04)	(0.02)	(0.02)	(0.02)
% Eng/Afr HL	-0.018	-0.016	0.023	-0.010	0.005	0.001
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
Pupil-teacher ratio	-0.001	0.000	-0.000	0.001*	-0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.187***	1.040***	0.968***	0.947***	0.930***	0.909***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
N	982	982	1163	1163	1270	1270
Adjusted R^2	0.186	0.093	0.200	0.072	0.243	0.082

Notes: "2019/20" must be read as Grade 11 in 2019 and Grade 12 in 2020. Robust standard errors are in parentheses. Significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

6.2 MINDING THE GAP

Given preliminary evidence of systematic differences in the gaps between NSC results and report marks by school fee status, NSC performance terciles and gender, the multivariate OLS regression results are presented in Table 9 below. The results explore which factors remain significant in a multivariate context, and whether they increase or decrease the magnitude of the gaps. The coefficients on the year variables indicate the potential influence of Covid-19 on assessment validity. Being in Grade 11 in 2019, relative to 2018, has a negligible and insignificant effect on the size of the gap between a learners' Grade 11 report mark and NSC result. However, being in Grade 11 in 2020 or 2021, relative to 2018, is associated with significant and large gaps between Grade 11 report mark and NSC results in Mathematics, of 2.868 and 3.666 respectively. Similarly**,** being in Grade 12 in 2020 and 2021 relative to 2019 is associated with significant increases in the gap between a learner's Grade 12 report mark and NSC result. Furthermore, for Grade 12s, the post-pandemic year

of 2022 is still associated with a wider school-NSC result gap relative to 2019. While the coefficients on the year variable indicate greater misalignment of internal marks compared to NSC marks during the pandemic years for Mathematics, this picture is not as clear cut for Mathematical Literacy and English FAL.

Learners' academic performance in the NSC significantly predicts the assessment gaps in both Grade 11 and 12. For Mathematical Literacy and English FAL, and in both Grade 11 and 12, top-performing learners (Tercile 3) exhibit significantly larger gaps compared to the lowest performing learners (Tercile 1). This is similarly the case for Grade 12 Mathematics. The dynamics in Grade 11 Mathematics are thus an anomaly, with learners in Tercile 2 and Tercile 3 being associated with significantly smaller gaps compared to learners in Tercile 1. The coefficients suggest that, in Grade 11 Mathematics, mid-and top- performers' internal marks are 1.961 and 0.884 marks closer to their NSC results than low performers' internal marks, after controlling for year and demographic factors. Thus, while internal marks align more closely with external results for top performers in Grade 11 for Mathematics, the opposite is true of Grade 12 Mathematics and in all other subjects.

While learners' academic ability remains a strong predictor of their assessment gaps in both Grade 11 and 12, attending a no-fee school only remains significant in Grade 11. Learners from no-fee schools have statistically significantly larger gaps compared to those from fee-paying schools, across all three subjects in Grade 11. In contrast, the effect of being at a no-fee school on the Grade 12 gap is either small or statistically insignificant across subjects. Additionally, and perhaps most intriguingly, the coefficient on being female is negative and statistically significant across all subjects and grades, except for Grade 12 English FAL. This indicates that, after implementing controls, female learners tend to have smaller gaps between school marks and NSC results compared to male learners, and the magnitude of the difference is greater in Grade 11 than in Grade 12.

Similarly, being home language English or Afrikaans significantly reduces the gaps between school marks and NSC results in Grade 11 and 12 Mathematics and Grade 11 Mathematical Literacy. Language has no significant effect on the gaps in English FAL, likely due to the very low proportion of English or Afrikaans learners taking the subject. Learners who are over-age, and thus more likely to have repeated a previous grade, are associated with larger gaps between their NSC result and school report mark for Grade 11 and 12 English FAL, while associated with smaller gaps in Grade 12 Mathematics. While these results are valuable, the descriptive analysis further suggested that the effect of Covid-19 may be heterogenous across lines of socioeconomics, gender and performance.

Table 9. OLS Regressions of Gaps between Report Marks and NSC Results

	Mathematics		Math. Literacy		English FAL	
	(1)	(2)	(3)	(4)	(5)	(6)
	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap
2019/20	-0.207	2.892***	1.036***	1.298***	0.989***	1.654***
	(0.23)	(0.22)	(0.30)	(0.26)	(0.17)	(0.26)
2020/21	2.868***	2.761***	0.0520	0.803***	0.634***	2.698***
	(0.28)	(0.20)	(0.24)	(0.16)	(0.14)	(0.21)
2021/22	3.666***	1.943***	-0.331	0.249	1.447***	2.436***
	(0.38)	(0.20)	(0.23)	(0.14)	(0.15)	(0.16)
No-Fee School	1.595***	0.519*	2.668***	0.217	1.619*	1.629
	(0.34)	(0.22)	(0.30)	(0.20)	(0.70)	(0.94)
Tercile 2	-1.961***	2.446***	2.360***	0.135	0.944***	0.779***
	(0.24)	(0.12)	(0.17)	(0.11)	(0.10)	(0.08)
Tercile 3	-0.884**	3.219***	7.815***	1.934***	2.044***	0.680***
	(0.30)	(0.15)	(0.27)	(0.16)	(0.15)	(0.11)
Female	-0.612***	-0.385***	-0.647***	-0.268***	-0.476***	-0.0510
	(0.12)	(0.07)	(0.06)	(0.04)	(0.05)	(0.04)
Eng/Afr HL	-0.896**	-0.923***	-1.949***	-0.0982	0.961	1.371
	(0.33)	(0.20)	(0.27)	(0.21)	(0.75)	(1.04)
Overage	0.147	-0.209**	-0.00665	-0.0333	0.160**	0.371***
	(0.15)	(0.08)	(0.07)	(0.04)	(0.05)	(0.06)
Constant	7.746***	3.321***	5.175***	5.261***	3.764***	3.243***
	(0.34)	(0.23)	(0.32)	(0.23)	(0.72)	(0.97)
N	47857	47857	104504	104504	157069	157069

Notes: "2019/20" must be read as Grade 11 in 2019 and Grade 12 in 2020. Robust standard errors are in parentheses and clustered at the school level. Significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Hierarchical Linear Model presented in the appendix as a robustness check.

Heterogeneity in the effect of years on assessment gaps by students' school fee status, academic ability and gender is considered in Table 10 below. While there is reassuringly little evidence of heterogeneous effects of the pandemic years by school fee-status (except perhaps for Grade 12 Mathematical Literacy), it does appear that the pandemic had heterogeneous effects on assessment gaps across learners' academic ability and gender. On average, being in Tercile 1 relative to Tercile 2 and 3 is associated with smaller assessment gaps across grades and subjects. However, the interactions indicate that being in Tercile 1 and in Grade 11 in 2019, relative to 2018, is associated with smaller gaps across all subjects (by approximately -0.953, -1.412 and 1.125). In contrast, being in Tercile 1 and in Grade 11 in 2020, relative to 2018, is associated with larger assessment gaps across all subjects. This is also true of Grade 11 assessment in 2021, but only for the numeracy subjects.

This highlights how the dynamics observed with the onset of the pandemic in the Grade 11 report mark distribution appears to differ across the performance distribution.

Interestingly, there appears to be evidence of heterogeneous effects of the pandemic years by gender, but only in Mathematics. While being in Grade 11 in 2019 (relative to 2018) and being female has no significant effect on Mathematics assessment gaps, being in Grade 11 in 2020 and being female is associated with larger assessment gaps, of 0.833 points. This interaction effect persists into 2021, with Grade 11s in 2021 having significantly larger gaps of 1.850 if they are also female. These findings corroborate Figure 12(a) and suggest that the pandemic amplified existing inequities, disproportionately affecting low-performing learners and females in Mathematics. Consequently, these dynamics, which are discussed in the following section, highlight the need for differentiated support strategies in schools during periods of disruption.

Table 10. Regression Results of Heterogenous Effects of Years on the Gaps between Report Marks and NSC Results

	Mathematics		Math. Literacy		English FAL	
	(1)	(2)	(3)	(4)	(5)	(6)
	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap
A: By School Fee Status						
2019/20	0.156	2.474***	0.518	-0.183	1.129*	1.769**
	(0.22)	(0.39)	(0.62)	(0.45)	(0.46)	(0.66)
2020/21	2.672***	2.419***	1.293**	0.269	0.255	2.049***
	(0.37)	(0.33)	(0.48)	(0.30)	(0.32)	(0.47)
2021/22	3.178***	1.635***	0.234	-0.242	2.081***	2.762***
	(0.52)	(0.33)	(0.42)	(0.29)	(0.35)	(0.51)
No-Fee School	1.360***	0.0995	3.378***	-0.522	1.817**	1.640
	(0.35)	(0.24)	(0.44)	(0.29)	(0.69)	(0.93)
2019/20 x No-Fee	-0.522	0.662	0.547	1.997***	-0.165	-0.134
	(0.40)	(0.47)	(0.71)	(0.53)	(0.49)	(0.72)
2020/21 x No-Fee	0.319	0.558	-1.766**	0.799*	0.417	0.728
	(0.54)	(0.41)	(0.54)	(0.35)	(0.36)	(0.52)
2021/22 x No-Fee	0.780	0.502	-0.832	0.740*	-0.709	-0.357
	(0.73)	(0.41)	(0.50)	(0.33)	(0.39)	(0.54)
B: By Tercile						
2019/20	0.127	4.235***	1.667***	1.272***	1.391***	1.663***
	(0.21)	(0.27)	(0.41)	(0.28)	(0.20)	(0.25)
2020/21	1.496***	3.513***	-0.645	0.610**	0.311	2.649***
	(0.23)	(0.23)	(0.33)	(0.20)	(0.16)	(0.22)
2021/22	1.352***	2.805***	-0.476	-0.205	1.719***	2.486***
	(0.26)	(0.24)	(0.31)	(0.18)	(0.18)	(0.18)
Tercile 1	-1.517***	-0.740***	-5.298***	-1.636***	-1.054***	-0.715***
	(0.36)	(0.19)	(0.30)	(0.15)	(0.17)	(0.13)
2019/20 x Tercile 1	-0.953*	-3.943***	-1.412***	0.194	-1.125***	-0.0279
	(0.48)	(0.31)	(0.40)	(0.25)	(0.19)	(0.24)
2020/21 x Tercile 1	4.082***	-2.163***	2.111***	0.600*	0.918***	0.142
	(0.52)	(0.28)	(0.45)	(0.30)	(0.21)	(0.18)
2021/22 x Tercile 1	6.919***	-2.517***	0.729*	1.391***	-0.780***	-0.145
	(0.69)	(0.32)	(0.35)	(0.24)	(0.21)	(0.17)
C: By Gender						
2019/20	-0.117	3.229***	0.894**	1.130***	0.991***	1.472***
	(0.28)	(0.25)	(0.33)	(0.28)	(0.18)	(0.25)
2020/21	2.368***	2.971***	-0.0402	0.728***	0.583***	2.696***
	(0.30)	(0.23)	(0.25)	(0.17)	(0.15)	(0.21)
2021/22	2.553***	2.195***	-0.226	0.147	1.494***	2.404***
	(0.37)	(0.25)	(0.25)	(0.16)	(0.15)	(0.17)
Female	-1.358***	-0.0647	-0.647***	-0.421***	-0.465***	-0.135*
	(0.16)	(0.11)	(0.13)	(0.08)	(0.08)	(0.06)
2019/20 x Female	-0.147	-0.562**	0.256	0.305*	-0.00339	0.322**
	(0.21)	(0.18)	(0.19)	(0.14)	(0.11)	(0.11)
2020/21 x Female	0.833**	-0.350*	0.167	0.136	0.0894	0.00381
	(0.26)	(0.16)	(0.20)	(0.12)	(0.09)	(0.09)
2021/22 x Female	1.850***	-0.421*	-0.194	0.187	-0.0814	0.0571
	(0.26)	(0.21)	(0.15)	(0.10)	(0.10)	(0.08)
N	47857	47857	104504	104504	157069	157069

Notes: All models are estimated with the full set of controls, namely indicators for Female, No-Fee School, Tercile, Eng/Afr HL and Overage. Robust standard errors are in parentheses and clustered at the school level. Significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7 DISCUSSION

The descriptive analysis highlights the notable impact of the Covid-19 pandemic on the quality of Mathematics SBA, particularly in Grades 10 and 11. The considerable increases in the mean Mathematics results and rightward shifts in the distributions for Grade 10 and Grade 11 in 2020 likely reflect greater leniency in assessment during the pandemic, as noted by Hoadley (2023). This is further supported by a weakening of the correlations and an increase in the gaps between 2020 and 2021 Grade 11 school report marks and NSC results. Additionally, the direction of the gaps suggest report marks were inflated relative to learners' NSC performance, particularly at the lower end of the distribution. Thus, while there was no explicit directive to pass learners during the pandemic, considerable variation in instructional time across schools and limitations in assessment practices may have incentivized leniency in grading and the promotion of weaker learners. This aligns with existing evidence of lower repetition and dropout rates during the pandemic (Gustafsson, 2022; Wills & Qvist, 2024; Wills & Van der Berg, 2023). However, these trends persist into the post-pandemic period.

Despite the return to normal teaching conditions in 2022 and the reinstatement of most pre-pandemic assessment policies in 2023, the Grade 10 and 11 Mathematics distributions do not revert to pre-pandemic patterns. This may underscore the unreliability of SBA tasks, as opposed to examinations, given that the SBA component of the final report mark was reduced from 60% to 40% in 2023, rather than restored to the pre-2020 weighting of 25% (Umalusi, 2004; Van der Berg & Shepherd, 2009; Carnoy et al. 2012; Department of Basic Education, 2019a). Furthermore, the consistency of Grade 10 and 11 assessment results pre-pandemic potentially provide evidence in favour of Deliwe and Van der Berg's (2022) argument that using common, standardised examinations in combination with SBAs can improve the quality of assessments across schools. However, further research should specifically investigate whether assessment reliability is primarily driven by the use of common examinations or other assessment practices, such as the thorough external moderation practices used in Grade 12.

School-based Mathematics assessment in Grade 12 show encouraging and stable indications of reliability. The correlations between school report marks and NSC results are higher in Grade 12 compared to Grade 11, across all 4 cohorts, suggesting greater reliability in Grade 12 assessment in Gauteng. This is in line with Van der Berg et al.'s (2023) findings for the Eastern Cape province. In addition, these correlations and the Mathematics result distributions remain largely unaffected by the pandemic, in contrast to Grade 10 and 11. A combination of external moderation, no pressure to progress and promote learners, and a lack of Covid-19 related changes to curriculum and assessment

policy may explain the observed reliability in Grade 12 Mathematics SBA, even during times of instability. However, while reliable, the Grade 12 Mathematics results do show indications of low validity from 2020 onwards, in both the descriptive and regression analysis.

The Grade 12 report marks during the pandemic seem to be under-scoring learners relative to their later performance in the NSC, particularly in Mathematics and English FAL. It seems unlikely that Grade 12 assessments would have been made more challenging or marked more strictly during the pandemic years. It would, however, be likely for learners to be performing worse in these years due to learning losses, as documented in other grades (Ardington et al., 2021) and countries (Donnelly & Patrinos, 2022; Engzell, Frey & Verhagen, 2021). Given this, I would tentatively postulate that these gaps are indicative of upward NSC mark adjustments from 2020 to 2022, rather than stricter Grade 12 SBA, particularly as the NSC standardisation process makes use of the result distributions of the preceding three to five years (Umalusi, 2016:13) which were unaffected both by Covid-19 and the larger birth cohorts (Gustafsson, 2022). Under this line of thought, the lower Grade 12 Mathematics report marks in 2020, 2021 and 2022 (see Figure 3) may in fact be more accurate indicators of learners' skill and content knowledge at the time.

A concerning finding is that assessment quality systematically differs by the socioeconomic status of the school. The multivariate analyses show that Grade 11 assessments at no-fee schools are significantly less reliable and less valid than those at fee-paying schools, and this is consistent across subjects. Furthermore, the reliability of Grade 11 assessments worsened by more at no-fee schools, compared to fee-paying schools, with the onset of the pandemic. While the poor content knowledge of teachers at no-fee schools (Venkat & Spaull, 2015) could be driving these differences, particularly with the devolution of assessment management to schools during the pandemic (Hoadley, 2023), there are a multitude of other issues at under-resourced schools which could explain the distinction. Reassuringly, however, there are no significant differences in the reliability of Mathematics assessments between fee-paying and no-fee schools at the Grade 12 level, across all 4 cohorts. This may suggest that the assessment policies and practices adopted at the Grade 12 level, particularly in Mathematics, could be used to inform policymakers on how to improve equality in assessment quality in lower grades. However, learner characteristics, which are less simply affected by policy, were also found to significantly influence assessment quality measures.

Learners' performance in the NSC is found to be significantly related to both the gap and correlation between school marks and NSC results. Similar to Van der Berg and Shepherd's (2009) finding, the reliability of assessment is positively correlated to average NSC performance, particularly for Grade 11, and schools with a higher proportion of low-performing learners consistently show weaker alignment between school and NSC assessments. This may reflect noisier performance trajectories

amongst low performing students and is something worth exploring in future research. Similarly, gaps between school Mathematics marks and NSC results are smaller in Grade 11 for top performers relative to lower performers, while in Grade 12 Mathematics, being a top performer widens the gap between your school mark and NSC result. In all other subjects and across grades, being in the top 66% of the distribution is associated with larger gaps between your school result and NSC result. However, the impact of your position in the NSC distribution shows heterogeneity across years. Thus, further research should explore this relationship in more depth using a quantile regression analysis, particularly given that the analysis presented in this paper is limited by relatively arbitrary cut-offs in the NSC performance distribution.

A final interesting result relates to gender. While gender was not the focus of this paper, it emerges as a significant correlate of assessment validity. In the multivariate regressions in Section 6.2, being female significantly reduces the size of the gap between learners' school report marks and NSC results, across grades and subjects. Furthermore, the effect of gender appears to differ across cohort years, suggesting that the pandemic may have differentially affected assessment validity by gender. These findings need to be more thoroughly explored and better understood in future research.

Overall, the analysis underscores the importance of addressing systemic biases in internal SBAs to ensure equity and alignment with external standardized metrics. However, while the results may suggest that standardised testing is a valuable tool for equalising assessment accuracy, it is by no means the silver bullet. For one, SBAs that are designed, administered and marked by teachers, while potentially less reliable, are important for assisting teachers in improving their professional practice and their ability to support student learning (Darling-Hammond & Wentworth, 2010). Secondly, while more standardised SBAs or the widespread use of common assessments would improve the value of SBA results for feedback at the systems level, SBAs should not become overly focused on providing a measure for high-stakes accountability, as this will limit their acceptance in schools and their value for learners (Deliwe & Van der Berg, 2022). Thus, strengthening the external moderation and monitoring of SBA quality is likely the better solution to differential assessment accuracy between schools, subjects and grades.

7.1 LIMITATIONS

While the panel nature of the dataset presents a novel opportunity to compare internal and external results, several limitations need to be acknowledged. First, the lack of SBA data prior to 2017 and NSC data beyond 2022, restricts the analysis to balanced cohorts in only Grade 11 and 12. Consequently, the results cannot tell us how assessment accuracy in Grade 10 differs to that in Grade

11 and 12. Additionally, none of the balanced cohorts contain Grade 11s in 2023, which limits the extent to which we can attribute assessment inaccuracies to the assessment policies implemented during Covid-19. Second, the NSC data is technically not directly comparable across years, nor is it directly comparable to the SBA data for Grade 11 and 12, as this contains different types of assessment and different content compared to that in the NSC examinations. Third, the NSC results are not the "raw" examination results but rather a combination of adjusted examination results and transformed SBA results[26], thus making them less reliable as a measure of students' true ability. Fourth, there is a lack of data capturing which schools and grades in the Gauteng province write common examinations. This restricts the degree with which we can attribute assessment reliability to school and teacher characteristics. Fifth, the data is captured by teachers and schools and thus may be subject to measurement and capturing errors, particularly during the Covid-19 period. Lastly, the datasets only contain basic student and school attributes, which reduces the potential to control for student attributes, such as individual socio-economic status or parental education. Ultimately, these data limitations preclude casual inference and restrict the paper to an exploratory analysis.

8 CONCLUSION

In this paper, SBA results are explored across time in a sample of Grade 11 and 12 Gauteng learners. Four key findings emerged. First, the pandemic notably altered SBA results across all three subjects, most notably in Grade 10 and 11 compared to Grade 12. Second, SBA results emerge as less reliable in no-fee schools compared to fee-paying schools, particularly for Mathematical Literacy and English FAL. However, in Grade 12 this systematic difference largely disappears in Mathematics and is less pronounced in Mathematical Literacy and English FAL. Third, learner and school academic performance in the NSC are strongly correlated to measures of assessment reliability and validity, and there is evidence of an interaction effect between the pandemic and academic ability on assessment validity. Fourth, gender emerges as a predictor of assessment accuracy, with female learners more likely to have smaller gaps between their school assessments and NSC results across grades and years, after controlling for other factors. While causality cannot be inferred from these associations, the results constitute an important first step in understanding how school assessment results differentially changed in response to the pandemic by subject, grade, socioeconomics, performance and gender. Overall, the analysis highlights the need for further research, particularly

[26] Please refer to Umalusi's (2016) *Requirements and Specifications for Standardisation, Statistical Moderation and Resulting*. This resource is available at: https://www.umalusi.org.za/docs/reports/2016/REQUIREMENTS%20AND%20SPECIFICATIONS%20FOR%20THE%20STANDARDISATION%20STATISTICAL%20MODE.pdf

qualitative in nature, to identify which factors explain the changes observed during the pandemic and are responsible for better quality SBAs in Grade 12, in Mathematics and in fee-paying schools.

REFERENCES[27]

Ardington, C., Wills, G., & Kotze, J. 2021. COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development* [Online], *86*:102480. Available: https://www.sciencedirect.com/science/article/pii/S0738059321001334?casa_token=sgtryZc-SMcAAAAA:hBq0P2FC1LHSm0o1BCNepyPbl8kM0aPrCQ2C4QLsnw1VyD2VngMFmGmPKYj8XGu76ktK2CyII2Q [2024, January 1]

Azmat, G., & Iriberri, N. 2010. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics* [Online], *94*(7-8):435-452. Available: https://www.sciencedirect.com/science/article/pii/S0047272710000411 [2024, January 1]

Bandiera, O., Larcinese, V., & Rasul, I. 2015. Blissful ignorance? A natural experiment on the effect of feedback on students' performance. *Labour Economics* [Online], 34:13-25. Available: https://www.sciencedirect.com/science/article/pii/S092753711500010X?ref=pdf_download&fr=RR-2&rr=8fcbb169482b0714 [2024, November1]

Black, N., & de New, S. C. 2020. Short, heavy and underrated? Teacher assessment biases by children's body size. *Oxford Bulletin of Economics and Statistics* [Online], 82(5):961-987. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1111/obes.12370 [2024, December 16]

Black, P. & Wiliam, D. 1998. Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice,* 5(1):7- 74

Braun, H. & Kanjee, A. 2006. Using assessment to improve education in developing nations, in Braun, H., Kanjee, A., Bettinger, E. & Kremer, M (eds.). *Improving education through assessment, innovation, and evaluation.* Cambridge: American Academy of Arts and Sciences.

Bridgman, G. 2020. Correspondence between mathematics and mathematical literacy scores: an analysis from 2010 to 2018. Stellenbosch Economic Working Papers no. WP03/2020.

Brown, G. T. & Harris, L. R. 2009. Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of MultiDisciplinary Evaluation*, 6(12):68-91.

Burgess, S. & Greaves, E. 2013. Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3):535-576.

[27] Referencing style: Harvard (Stellenbosch).

Burgess, S., Hauberg, D. S., Rangvid, B. S., & Sievertsen, H. H. 2022. The importance of external assessments: High school math and gender gaps in STEM degrees. *Economics of Education Review* [Online], *88*:102267. Available: https://www.sciencedirect.com/science/article/pii/S0272775722000437?ref=pdf_download&fr=RR-2&rr=8fcbb3cc497e0714 [2024, December 10]

Campbell, T. 2015. Stereotyped at seven? Biases in teacher judgement of pupils' ability and attainment. *Journal of Social Policy* [Online], 44(3):517-547. Available: https://www.cambridge.org/core/journals/journal-of-social-policy/article/stereotyped-at-seven-biases-in-teacher-judgement-of-pupils-ability-and-attainment/B6907C36F39D0476DB795A9EE7D7D6F7 [2024, December 15]

Carnoy, M., Chisholm, L., & Chilisa, B. 2012. The low achievement trap: Comparing schooling in Botswana and South Africa*. Human Sciences Research Council Press.*

Case, A. & Deaton, A. 1999. School inputs and educational outcomes in South Africa. *The Quarterly Journal of Economics*, 114(3):1047-1084.

Case, A. & Yogo, M. 1999. Does School Quality Matter? Returns to education and the characteristics of schools in South Africa. NBER Working Paper no. 7399

Center on Standards and Assessment Implementation (CSAI). 2018. *Valid and Reliable Assessments.* Available: https://files.eric.ed.gov/fulltext/ED588476.pdf [2024, April 18]

Cheng, Y. & Kong, D. 2023. Are female teachers more likely to practice grade inflation? Evidence from China. *China Economic Review*, 80:101987.

Contreras, D. 2024. Gender differences in grading: teacher bias or student behaviour? *Education Economics* [Online], 32(6):762-785. Available: https://www.tandfonline.com/doi/epdf/10.1080/09645292.2023.2252620?needAccess=true [2024, December 20]

Darling-Hammond, L. & Wentworth, L. 2010. Benchmarking learning systems: Student performance assessment in international context. Stanford Center for Opportunity Policy in Education. Stanford University, CA.

Deliwe, C.N. & Van der Berg, S. 2022. The role of assessment in Foundation Phase improvement: The Annual National Assessments and beyond. In: H. Venkat and N. Roberts (eds.), *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press Southern Africa.

Department of Basic Education. 2019a. Colloquium presentation on Policy and School Based Assessment [Online]. Available: https://www.umalusi.org.za/wp-content/uploads/2019/09/DBE_Dr-Poliah-SBA-Colloquium.pdf [2024, September 10]

Department of Basic Education. 2019b. *CIRCULAR E29 OF 2019: THE DISCONTINUATION OF THE MULTIPLE EXAMINATION OPPORTUNITIES (MEO) DISPENSATION WITH EFFECT FROM 2020*, October 4 [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Publications/Circular%20E29%20of%202019.pdf?ver=2019-10-28-142703-507 [2024, December 3].

Department of Basic Education. 2020a. *CIRCULAR NO.3 OF 2020: CONTAINMENT/ MANAGEMENT OF COVID 19 FOR SCHOOLS AND SCHOOL COMMUNITIES,* March 15 [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Publications/Circular%20on%20change%20of%20school%20dates%20COVID19.pdf?ver=2020-03-16-132144-460 [2024, December 3].

Department of Basic Education. 2020b. *CIRCULAR NO S2 OF 2020: RELEASE OF THE REVISED ANNUAL TEACHING PLANS (ATPs) FOR GRADE 7 AND GRADE 12 RESPECTIVELY*, May 23 [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Recovery%20plan%20page/Links%20for%20schools/circular-s2-2020-curriculum-recovery.pdf?ver=2020-06-15-091313-790 [2024, December 10]

Department of Basic Education. 2020c. *NATIONAL ASSESSMENT CIRCULAR 02 OF 2020 IMPLEMENTATION AND QUALITY ASSURANCE OF THE AMENDED 2020 ASSESSMENT PROGRAMME IN THE GENERAL EDUCATION AND TRAINING (GET) BAND (GRADES R-9),* July 9 [Online]. Available: https://irp-cdn.multiscreensite.com/c0cc1c10/files/uploaded/Circular%2002%20of%202020%20-%20National%20Assessement.pdf [2024, December 10]

Department of Basic Education. 2020d. *CIRCULAR S3 OF 2020: DISTRIBUTION OF THE TEACHER GUIDELINES FOR THE IMPLEMENTATION OF ANNUAL TEACHING PLANS (ATPs) AND THE MINIMUM CORE CONTENT AND SKILLS PER SUBJECT AND PER GRADE*, July 9 [Online]. Available: https://wcedeportal.co.za/eresource/146801 [2024, December 10]

Department of Basic Education. 2020e. *CIRCULAR E11 OF 2020: IMPLEMENTATION AND QUALITY ASSURANCE OF 2020 SCHOOL BASED ASSESSMENT: GRADES 10-12,* July 13 [Online]. Available: https://wcedonline.westerncape.gov.za/circulars/circulars20/circ18-20-

Circular%20E11%20%20Implementation%20%26%20QA%20of%202020%20SBA%20Grades%2010-12%20.pdf [2024, December 10]

Department of Basic Education. 2020f. *CIRCULAR S7 OF 2020: REVISED PROMOTION REQUIREMENTS FOR GRADE 10 AND 11 FOR THE 2020 ACADEMIC YEAR,* September 12 [Online] Available: https://www.education.gov.za/Portals/0/Documents/Publications/Circular%20S7%20of%202020%20Revised%20Promotion%20Requirements%20for%20Grade%2010-11%20Guidelines.pdf?ver=2020-10-13-093808-787 [2024, December 10]

Department of Basic Education. 2022. *CIRCULAR NO. S1 OF 2022: MEASURES FOR THE RESUMPTION OF FULL TIME DAILY SCHOOL ATTENDANCE,* February 5 [Online]. Available: https://www.kzneducation.gov.za/images/documents/Circulars/GeneralCirculars/Circular_S1_of_2022.pdf [2024, December 10]

Department of Basic Education. 2023. *2023 National Senior Certificate (NSC) Exam Results: Technical Report*. Pretoria, South Africa: Department of Basic Education

Donnelly, R. & Patrinos, H. A. 2022. Learning Loss during Covid-19: An Early Systematic Review. *Prospects*, 51(4), 601–609

Engzell, P., Frey, A., & Verhagen, M. D. 2021. Learning Loss Due to School Closures during the Covid-19 Pandemic. *Proceedings of the National Academy of Sciences*, 118(17).

Fives, H., & Buehl, M. M. 2012. Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? in K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (eds.). *APA educational psychology handbook, Vol. 2. Individual differences and cultural and contextual factors*. American Psychological Association. 471–499.

Fleisch, B. 2008. *Primary education in crisis: Why South African School children underachieve in reading and mathematics*. Cape Town: Juta and Company Ltd.

Gauteng Department of Education [GDE]. 2022. Department of Education: Province of Gauteng Vote 05 Education Annual Report 2021/2022 Financial Year [Online]. Available: https://education.gauteng.gov.za/GDE%20Annual%20Reports/Gauteng%20Department%20of%20Education%20Annual%20Report%202021%20-%202022.pdf [2024, September 10]

Gustafsson, M. 2011. The when and how of leaving school: The policy implications of new evidence on secondary schooling in South Africa. Stellenbosch Economic Working Papers no. 09/2011.

Gustafsson, M. 2016. Understanding trends in high-level achievement in Grade 12 mathematics and physical science [Online]. Available: https://resep.sun.ac.za/wp content/uploads/2016/06/Correcting-racial-imbalances-2016-01-25.pdf [2024, March 28]

Gustafsson, M. 2022. The COVID-19 pandemic, enrolments, dropping out and attendance explained [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Reports/The%20pandemic%20and%20attendance%202022%2003%2008.pdf [2024, January 3].

Gustafsson, M. & Taylor, S. 2016. Treating schools to a new administration: Evidence from South Africa of the impact of better practices in the system-level administration of schools. Stellenbosch Economic Working Papers no. 05/16.

Hoadley, U. 2023. Covid-19 and the South African curriculum policy response. Research on Socio-Economic Policy (RESEP): Stellenbosch University [Online]. Available: https://resep.sun.ac.za/wp-content/uploads/2023/05/Hoadley_2023_COVID-19_curriculum_policy_response_SA_15May2023.pdf [2024, November 7]

Kellaghan, T. & V. Greaney. 2001. *Using Assessment to Improve the Quality of Education*. Paris: UNESCO. Available: https://unesdoc.unesco.org/ark:/48223/pf0000126231 [2024, September 10]

Kibble, J.D. 2017. Best practices in summative assessment. *Advances in physiology education*, 41(1):110-119.

Kimani, E.M. & Bhorat, H. 2014. The effects of pupil-teacher ratio and expenditure per pupil on educational attainment in South Africa [Online]. Available: https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/AEC_2014_-_the_effects_pupil-teacher_ratio_and_expenditure_per_pupil_on_educational_attainment_in_south_africa_-_11_2014.pdf [2024, March 28]

Lavy, V. 2008. Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of public Economics* [Online], 92(10-11):2083-2105. Available: https://www.sciencedirect.com/science/article/pii/S0047272708000418 [2024, December 19]

Lavy, V. & Sand, E. 2015. On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases. National Bureau of Economic Research Working Paper No. 20909.

Lindahl, E. 2007. Gender and ethnic interactions among teachers and students: Evidence from Sweden. Institute for Labour Market Policy Evaluation (IFAU), Uppsala. Working Paper No. 2007:25.

Lindahl, E. 2016. Are teacher assessments biased?–evidence from Sweden. *Education economics* [Online], 24(2):224-238. Available: https://www.tandfonline.com/doi/epdf/10.1080/09645292.2015.1014882?needAccess=true [2024, December 20]

Maxwell, G. 2006. Quality management of school-based assessments: Moderation of teacher judgments. *32nd IAEA Conference, Singapore* [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=be9e45f1bc3614e8bf5af0dbc28d1cb938ab13d1 [2024, March 1]

McMillan, J.H. 2013. *SAGE Handbook of Research Classroom Assessment*. SAGE, Thousand Oaks.

Moll, P.G. 1998. Primary schooling, cognitive skills and wages in South Africa. *Economica,* 65(258):263-284.

Murphy, R. & Weinhardt, F. 2020. Top of the class: The importance of ordinal rank. *The Review of Economic Studies* [Online], 87(6):2777-2826. Available: https://academic.oup.com/restud/article/87/6/2777/5831842 [2025, January 1]

Nitko, A.J. & Brookhart, S.M. 2014. *Educational assessment of students*. 6th edition. Harlow: Pearson Education Limited.

Ouazad, A. 2014. Assessed by a teacher like me: Race and teacher assessments. *Education Finance and Policy* [Online], 9(3):334-372. Available: https://direct.mit.edu/edfp/article/9/3/334/10193/Assessed-by-a-Teacher-Like-Me-Race-and-Teacher [2024, December 10]

Paserman, M. D. 2007. *Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players.* IZA Discussion Papers 2834, Institute of Labor Economics (IZA).

Postlethwaite, T.N. & Kellaghan, T. 2008. *National assessments of educational achievement.* Education Policy Series. UNESCO, Paris: International Institute for Educational Planning.

Reddy, V., Visser, M., Winnaar, L., Arends, F., Juan, A and Prinsloo, C.H. & Isdale, K. 2016. *TIMSS 2015: Highlights of Mathematics and Science Achievement of Grade 9 South African Learners.* Human Sciences Research Council.

Republic of South Africa. Department of Education. 2006. *South African Schools Act 84 of 1996: Amended National Norms and Standards for School Funding*. Government Gazette no. 29179, 31 August [Online]. Available: https://www.gov.za/sites/default/files/gcis_document/201409/29179.pdf [2024, April 25]

Republic of South Africa. Department of Basic Education. 2009. *National Senior Certificate: A Qualification at Level 4 on the National Qualifications Framework (NQF).* Government Gazette no.32678, 3 November [Online]. Available: https://www.gov.za/sites/default/files/gcis_document/201409/doethe-national-senior-certificate-qualification-level-4-national-qualifications-framework.pdf [2024, December 3].

Republic of South Africa. Department of Basic Education. 2013. *National Policy Pertaining to the Programme and Promotion Requirements of the National Curriculum Statement Grades R – 12*. Government Gazette no. 36465, 17 May [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Policies/PolicyProgPromReqNCS.pdf [2024, December 3].

Republic of South Africa. Department of Basic Education. 2014. National policy pertaining to the conduct, administration and management of the National Senior Certificate examination. Government Gazette no.37652,16 May [Online]. Available: https://www.education.gov.za/Portals/0/Documents/Policies/PolCondAdminManageNSC.pdf?ver=2015-02-03-154737-197 [2024, December 3].

Republic of South Africa. Department of Basic Education. 2011. *Curriculum and Assessment Policy Statement (CAPS) Grades 10-12: Mathematics* [Online]. Available: https://www.education.gov.za/Portals/0/CD/National%20Curriculum%20Statements%20and%20Vocational/CAPS%20FET%20_%20MATHEMATICS%20_%20GR%2010-12%20_%20Web_1133.pdf?ver=2015-01-27-154314-253 [2024, December 3].

Rosenkvist, M.A. 2010. Using Student Test Results for Accountability and Improvement: A Literature Review. OECD Education Working Paper No. 54 [Online]. Available:

https://www.oecd.org/en/publications/using-student-test-results-for-accountability-and-improvement_5km4htwzbv30-en.html [2024, March 16]

Shepherd, D. L. 2011. Constraints to school effectiveness: What prevents poor schools from delivering results. Stellenbosch Economic Working Papers no. 05/11.

Shepherd, D. L. & Van der Berg, S. 2020. Analysing matric data to identify 'promising' schools in mathematics performance. Stellenbosch Economic Working Papers no. 16/2020

Soudien, C., Reddy, V., & Harvey, J. 2022. The impact of COVID-19 on a fragile education system: The case of South Africa. *Primary and secondary education during COVID-19: Disruptions to educational opportunity during a pandemic* [Online], 303-325. Available: https://link.springer.com/chapter/10.1007/978-3-030-81500-4_12 [2024, December 1]

South African Department of Basic Education. Master List of Schools 2023 [dataset]. Version 1. Pretoria: DBE [producer], 2023. Cape Town: DataFirst [distributor], 2024. DOI: https://doi.org/10.25828/1f7t-hn28

Spaull, N. 2013. Poverty & privilege: Primary school inequality in South Africa. *International journal of educational development*, 33(5):436-447.

Spaull, N. 2015. Schooling in South Africa: How low-quality education becomes a poverty trap. *South African Child Gauge*, 12(1):34-41.

Spaull, N. 2019. Chapter 1: Equity: A price too high to pay in N. Spaull and J.D. Jansen (eds). *South African Schooling: The Enigma of Inequality.* Cham: Springer Nature Switzerland AG. 1-24.

Statistics South Africa. 2023. General Household Survey 2023. Pretoria, South Africa: Statistics South Africa.

Stiggins, R.J. 2008. *Assessment Manifesto: A Call for the Development of Balanced Assessment Systems* [Online]. Available: https://famemichigan.org/wp-content/uploads/2018/06/Stiggins-Assessment-Manifesto-A-Call-for-the-Development-of-Balanced-Assessment-Systems.pdf [2024, March 13]

Taylor, S. 2011. Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study. Stellenbosch Economic Working Papers no.10/11.

Taylor, S. & Yu, D. 2009. The importance of socio-economic status in determining educational achievement in South Africa. Stellenbosch Economic Working Papers no.01/09.

Umalusi. 2004. Investigation into the Standard of the Senior Certificate Examination: A Report on Research conducted by Umalusi. [Pretoria]: Umalusi [Online]. Available: https://umalusi.org.za/ur/research/SC%20Research%20Summary.pdf [2024, September 10]

Umalusi. 2016. REQUIREMENTS AND SPECIFICATIONS FOR STANDARDISATION, STATISTICAL MODERATION AND RESULTING. [Pretoria]: Umalusi [Online]. Available: https://www.umalusi.org.za/docs/reports/2016/REQUIREMENTS%20AND%20SPECIFICATIONS%20FOR%20THE%20STANDARDISATION%20STATISTICAL%20MODE.pdf [2024, September 10]

Van der Berg, S., Spaull, N., Wills, G., Gustafsson, M., & Kotzé, J. 2016. *Identifying binding constraints in education: synthesis report for the Programme to support Pro-poor Policy development (PsPPd)* [Online] Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2906945 [2024, August 20]

Van der Berg, S. & Louw, M. 2006. Lessons learnt from SACMEQII: South African student performance in regional context. Stellenbosch Economic Working Papers no. 16/07

Van der Berg, S. 2007. Apartheid's enduring legacy: Inequalities in education. *Journal of African Economies*, 16(5):849-880.

Van der Berg, S. 2008. How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, 34(3):145–154.

Van der Berg, S. & Shepherd, D. 2009. *Signalling performance: An analysis of continuous assessment and matriculation examination marks in South African schools*. Pretoria: Umalusi.

Van der Berg, S., van Wyk, C., Gustafsson, M., Meyer, H., Chari, A., van Biljon, C., Lilenstein, A., Selkirk, R. & McCallum, J. 2023. *WHAT RICH NEW EDUCATION DATA CAN TELL US: New insights into learner flows, assessment, learner mobility and the subjects teachers teach* [Online]. Available: https://resep.sun.ac.za/what-rich-new-education-data-can-tell-us/sample-post/ [2024, March 28]**.**

Van der Berg, S. & Böhmer, B. 2025. South Africa: COVID-19 Learning Losses and Attempts at Recovery in a Poorly Performing and Unequal Education System, in *Improving National Education Systems After COVID-19: Moving Forward After PIRLS 2021 and PISA 2022*. Cham: Springer Nature Switzerland. 153-169.

Van der Berg, S., Van Wyk, C., & Selkirk, R. 2020. Schools in the time of COVID-19: Possible implications for enrolment, repetition and dropout. Stellenbosch Economic Working Papers no. 20/20.

Van Wyk, C., Gondwe, A., & De Villiers, P. 2017. Learner flow through patterns in the Western Cape using CEMIS datasets from 2007 to 2014: A longitudinal cohort analysis Stellenbosch. Working Paper No. 02/2017.

Venkat, H. & Spaull, N. 2015. What do we know about primary teachers' mathematical content knowledge in South Africa? An analysis of SACMEQ 2007. *International Journal of Educational Development*, 41:121-130.

Wills, G., & Qvist, J. 2023. *Repetition and dropout in South African before, during, and after COVID-19: Covid-Generation Research Project*. RESEP, Stellenbosch University [Online]. Available: https://resep.sun.ac.za/wp-content/uploads/2024/01/2023-12-21-Wills_Qvist_covid_dropout_repetition_report_v4_upload.pdf [2024, November 25]

Wills, G. & Van der Berg, S. 2024. COVID-19 disruptions and education in South Africa: Two years of evidence. *Development Southern Africa*, 41(2): 446–465.

Winnaar, L., Arends, F. & Beku, U. 2018. Reducing bullying in schools by focusing on school climate and school socio-economic status. *South African Journal of Education* [Online], *38*(1). Available: https://www.sajournalofeducation.co.za/index.php/saje/article/view/1596/797 [2024, December 1]

Wyness, G., Macmillan, L., Anders, J. & Dilnot, C. 2023. Grade expectations: how well can past performance predict future grades? *Education Economics* [Online], 31(4):397-418. Available: https://www.tandfonline.com/doi/epdf/10.1080/09645292.2022.2113861?needAccess=true [2024, December 15]

APPENDIX

8.1 APPENDIX FOR SECTION 3: INSTITUTIONAL SETTING

Table 11. Brief Outline of the Statistical Moderation of SBAs by Umalusi

The statistical moderation of SBA results is outlined by Umalusi in the document titled, *Requirements and Specifications for Standardisation, Statistical Moderation and Resulting* (Umalusi, 2016). The SBA moderation process makes use of the following measures at a subject, centre level: mean of external exam mark (ME), mean of the SBA mark (MS), standard deviation of the exam mark (SDE), standard deviation of the SBA mark (SDS) and the number of candidates. Thus, the SBA statistical moderation practice calibrates candidate's marks on the SBA component(s) using marks from the NSC exam component(s) and, as such, calibrates the SBA component to a new scale that is common to all candidates, that is, the exam component.

Table 12. Gauteng Education Department's Communication around Common Examinations

The Gauteng Education Department is involved in producing common examinations, particularly for Grade 12 Trial Examinations. While the Gauteng Department postponed all common examinations in 2020, exemplars for Grade 10 and 11 were emailed to districts for schools to administer and schools were assisted with sourcing past common examination papers. In addition, The Gauteng Department of Education stated in their 2021/22 Annual Report that the "preparatory examination" was offered to Grade 12 candidates at full-time schools, with schools obtaining a pass rate lower than 80% being "compelled" to write the GDE preparatory examination (Gauteng Department of Education [GDE], 2022).

Table 13. Policies on Mark Adjustments and Condonations in Senior Phase Mathematics

Mark adjustments were introduced in Assessment Circular 3 of 2015 for the Senior Phase. The policy stated that if a learner had obtained a subject mark within a 7% range of the pass requirement, marks could be adjusted by this 7% in a maximum of three subjects per learner. The mark adjustments could only be implemented if it would lead to a learner meeting the promotion requirements, and mark adjustments should prioritise the "fundamental subjects" of Home Language, First Additional Language and Mathematics. However, the number of subjects and the adjustment range was intended to decrease with each successive year of CAPS implementation. The policy stated that by 2018, only a 2% adjustment could be made in *one* subject (Assessment Circular 3 of 2015). It appears, however, that this was changed such that the 2% adjustment could be applied in a maximum of 3 subjects in 2019, not in one as previously stated by the circular (Assessment Instruction no 36 of 2019).
Condonations were introduced in Assessment Circular 3 of 2016 to be exclusively applied to the Senior Phase. A condonation in Mathematics occurs where a learner has met all the requirements in respect of promotion, except has not attained a level 3 (40%) in Mathematics. Provided the learner attains a minimum of 20% in Mathematics, the Mathematics mark is then condoned and the learner must be promoted to the next grade. Condonations were expected to take place after any mark adjustments were made. Furthermore, Grade 9 learners who were condoned in Mathematics would not be allowed to take Mathematics in Grade 10, as a minimum mark of 30% was required (Assessment Circular 3 of 2016). However, this later changed, allowing learners who were condoned to offer Mathematics, even if they attained a Grade 9 result below 30%.
During the pandemic, GET phase mark adjustments were increased from 2% in 2019 to 5% in 2020 in a maximum of three subjects (Circular 7 of 2020), and thereafter, condonation in Mathematics was to be applied.

8.2 APPENDIX FOR SECTION 4: DATA

Table 14. Sample Sizes for the Unbalanced Sample (Ver.1) across Grades, Years and Subjects.

Grade	Year	Mathematics	Mathematical Literacy	English FAL
	2018	39 353	55 407	72 194
	2019	42 374	72 464	87 391
	2020	39 580	66 181	78 397
10	2021	49 026	75 021	91 920
	2022	54 148	92 607	108 661
	2023	48 102	86 045	97 761
	Total	**272 583**	**447 725**	**536 324**
	2018	27 857	47 320	56 411
	2019	30 480	60 341	68 502
	2020	24 274	53 771	57 124
11	2021	34 059	74 338	80 236
	2022	37 723	77 307	83 652
	2023	33 864	73 995	76 481
	Total	**188 257**	**387 072**	**422 406**
	2018	20 758	40 889	59 417
	2019	19 845	42 006	59 596
	2020	11 945	23 050	44 793
12	2021	22 387	52 979	70 067
	2022	21 301	60 481	81 247
	2023	24 853	52 671	
	Total	**121 089**	**272 076**	**315 120**

Table 15. Structure of Attrition for Mathematics, Mathematical Literacy and English FAL as the school master list and NSC result data is matched to the balanced Gr.11/12 DDD data.

	2018/19		**2019/20**		**2020/21**		**2021/22**		**2022/23**	
	N	% left	N	% left	N	% left	N	% left	N	% left
a) Math. & M.Lit.										
Gr11/12 Results	42,985		26,548		44,719		62,617		64,459	
Schools	444		305		407		501		542	
School Masterlist	42,884	99.77	26,532	99.94	44,705	99.97	62,581	99.94	64,436	99.96
Schools	439	98.87	303	99.34	406	99.75	499	99.60	541	99.82
Cleaned Data *	39,369	91.59	24,207	91.18	41,186	92.10	58,102	92.79	59,083	91.66
Schools *	343	77.25	239	78.36	321	78.87	393	78.44	416	76.75
NSC Results ♦	33,760	78.54	23,233	87.51	39,826	89.06	56,022	89.47		
Schools ♦	340	76.58	237	77.70	320	78.62	392	78.24		
b) English FAL										
Gr11/12 Results	42,356		33,808		40,049		65,475			
Schools	398		338		339		455			
School Masterlist	42,287	99.84	33,772	99.89	40,006	99.89	65,469	99.99		
Schools	395	99.25	337	99.70	338	99.71	454	99.78		
Cleaned Data *	40,028	94.50	31,855	94.22	38,125	95.20	62,324	95.19		
Schools *	345	86.68	290	85.80	300	88.50	397	87.25		
NSC Results ♦	32,728	77.27	30,113	89.07	35,961	89.79	58,504	89.35		
Schools ♦	343	86.18	288	85.21	298	87.91	395	86.81		

Note: Gr. 11/12 Results refers to the balanced dataset of Grade 11 and 12 school assessment results from the DDD data. Cleaned data refers to the removal of independent schools, Special Needs Education Centres, and schools with no report mark data. * Version 2 of the data ♦ Version 3 of the data.

Table 16. Descriptive Statistics for Numeracy Subjects' School Based Assessment Sample at the Learner Level

	Mathematics					Mathematical Literacy				
	2018/19	**2019/20**	**2020/21**	**2021/22**	**2022/23**	**2018/19**	**2019/20**	**2020/21**	**2021/22**	**2022/23**
Total Learners	13 613	8 929	13 720	14 568	19 816	25 756	15 278	27 466	43 534	39 267
Results										
Gr.11 Mark	35.54	36.17	43.63	43.00	44.49	37.92	37.66	41.53	40.57	43.36
< 30%	0.41	0.38	0.21	0.21	0.18	0.26	0.24	0.13	0.14	0.09
Gr.12 Mark	36.07	31.87	33.17	31.53	37.08	40.81	41.35	42.93	44.59	43.91
< 30%	0.42	0.52	0.48	0.53	0.39	0.23	0.23	0.15	0.10	0.15
Demographics										
Female	0.59	0.60	0.61	0.60	0.62	0.54	0.55	0.55	0.54	0.55
Over-Age	0.24	0.19	0.18	0.18	0.13	0.51	0.45	0.45	0.46	0.36
Eng/Afr HL	0.22	0.20	0.19	0.20	0.25	0.18	0.13	0.16	0.14	0.18
School										
No-Fee School	0.63	0.66	0.63	0.64	0.55	0.69	0.79	0.73	0.77	0.69
Quintile 1	0.12	0.09	0.13	0.11	0.10	0.13	0.14	0.17	0.14	0.13
Quintile 2	0.12	0.11	0.12	0.10	0.09	0.13	0.12	0.13	0.15	0.13
Quintile 3	0.17	0.20	0.17	0.15	0.16	0.17	0.21	0.18	0.20	0.18
Quintile 4	0.21	0.23	0.17	0.24	0.18	0.26	0.27	0.21	0.25	0.23
Quintile 5	0.39	0.36	0.42	0.38	0.47	0.31	0.26	0.31	0.26	0.33

Notes:

Table 17. Descriptive Statistics for English FAL at the Learner Level

	a) English FAL (SBA only)				b) English FAL (SBA and NSC data)			
	2018/19	**2019/20**	**2020/21**	**2021/22**	**2018/19**	**2019/20**	**2020/21**	**2021/22**
Total Learners	40028	31855	38125	62324	32728	30113	35961	58504
Results								
Gr.11 Mark	53.25	53.80	56.07	53.02	54.92	53.98	56.29	53.29
< 30%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gr.12 Mark	53.38	55.50	50.23	50.34	55.47	56.08	50.73	50.84
< 30%	0.02	0.01	0.04	0.03	0.00	0.01	0.03	0.02
NSC Result	-	-	-	-	57.46	58.85	57.38	57.51
Demographics								
Female	0.55	0.56	0.56	0.56	0.56	0.57	0.57	0.57
Over-Age	0.46	0.38	0.40	0.40	0.39	0.38	0.39	0.39
Eng/Afr HL	0.14	0.15	0.11	0.09	0.15	0.14	0.11	0.09
School								
No-fee School	0.86	0.85	0.89	0.91	0.85	0.86	0.89	0.91
Quintile 1	0.15	0.15	0.20	0.16	0.15	0.15	0.20	0.16
Quintile 2	0.17	0.15	0.18	0.19	0.16	0.15	0.18	0.18
Quintile 3	0.23	0.26	0.25	0.26	0.23	0.27	0.25	0.26
Quintile 4	0.29	0.27	0.24	0.27	0.29	0.27	0.24	0.27
Quintile 5	0.16	0.17	0.13	0.12	0.17	0.17	0.13	0.12

Notes:

8.3 APPENDIX FOR SECTION 6: DESCRIPTIVE ANALYSIS

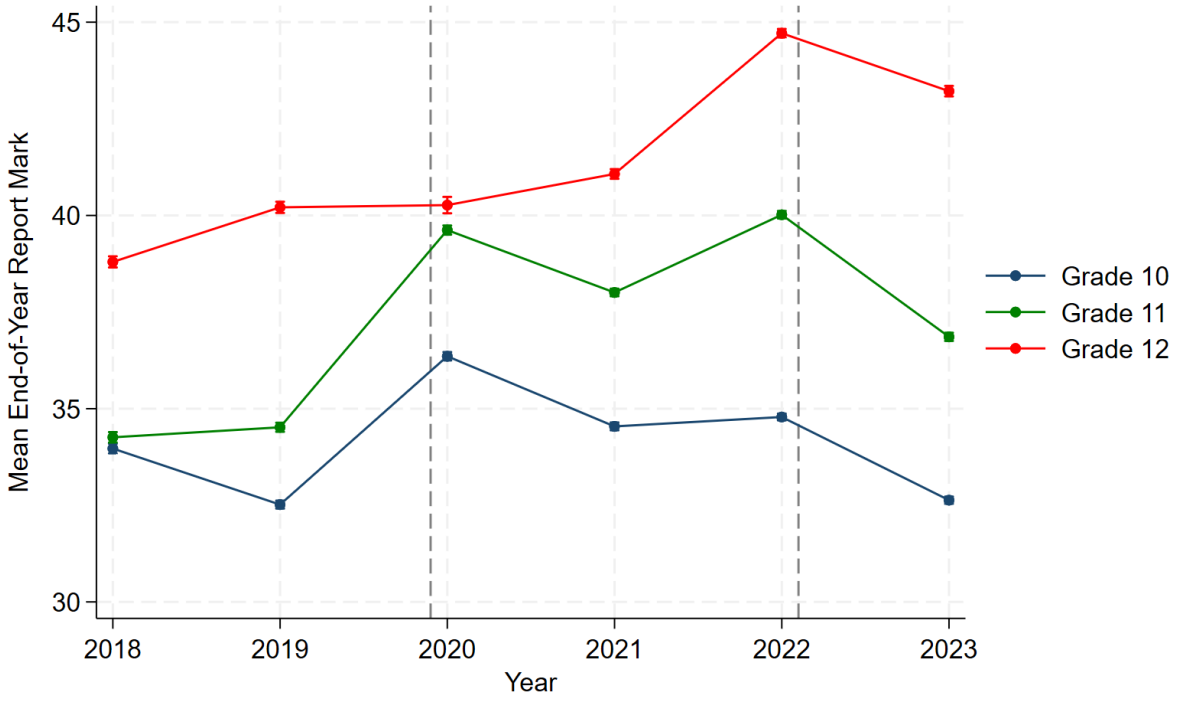


Figure 13. Mean End-of-Year Mathematical Literacy Report Marks for Grades 10-12 between 2018 and 2023.

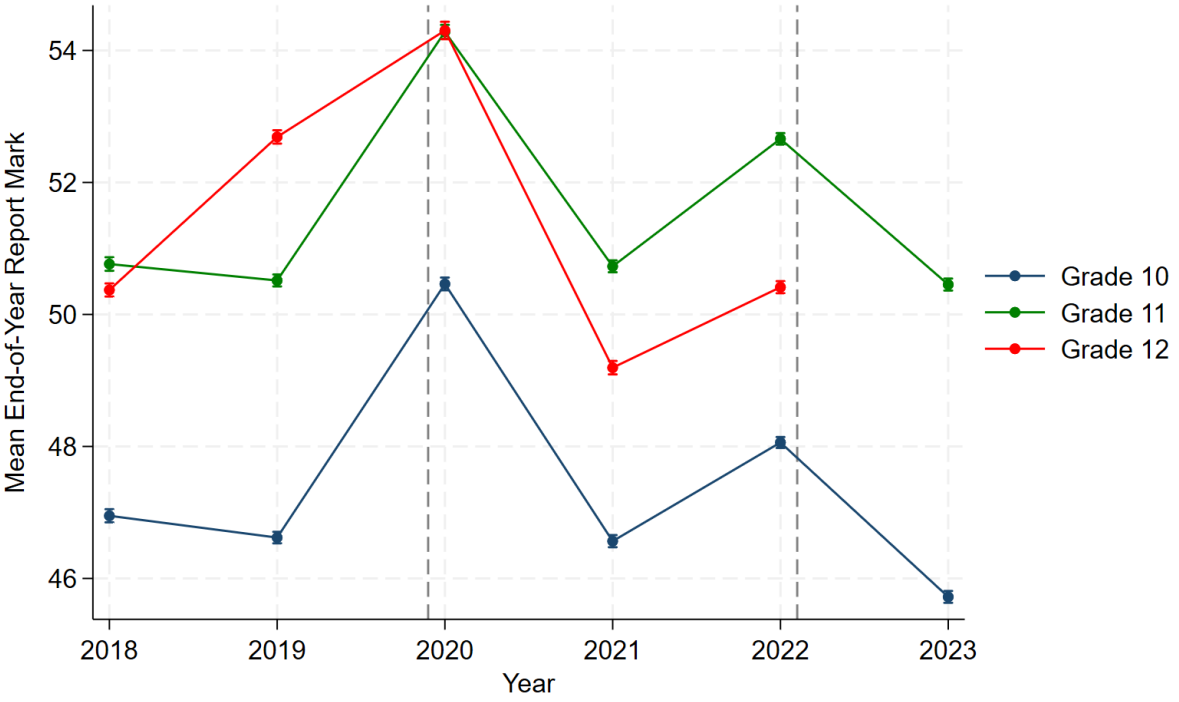


Figure 14. Mean End-of-Year English FAL Report Marks for Grades 10-12 between 2018 and 2023.

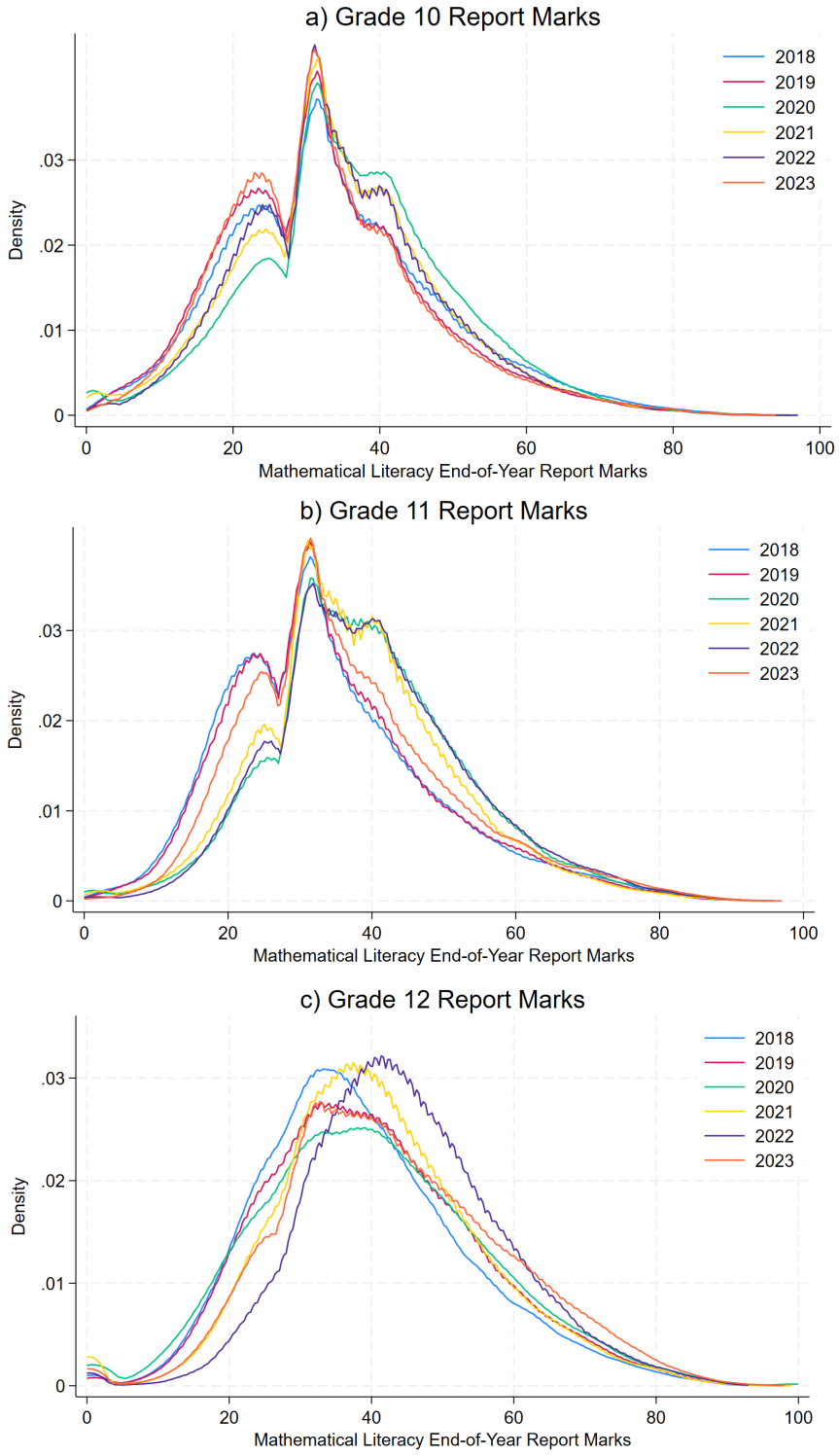


Figure 15. Kernel Density Estimates of Mathematical Literacy End-of-Year Report Marks across years for a) Grade 10, b) Grade 11 and c) Grade 12

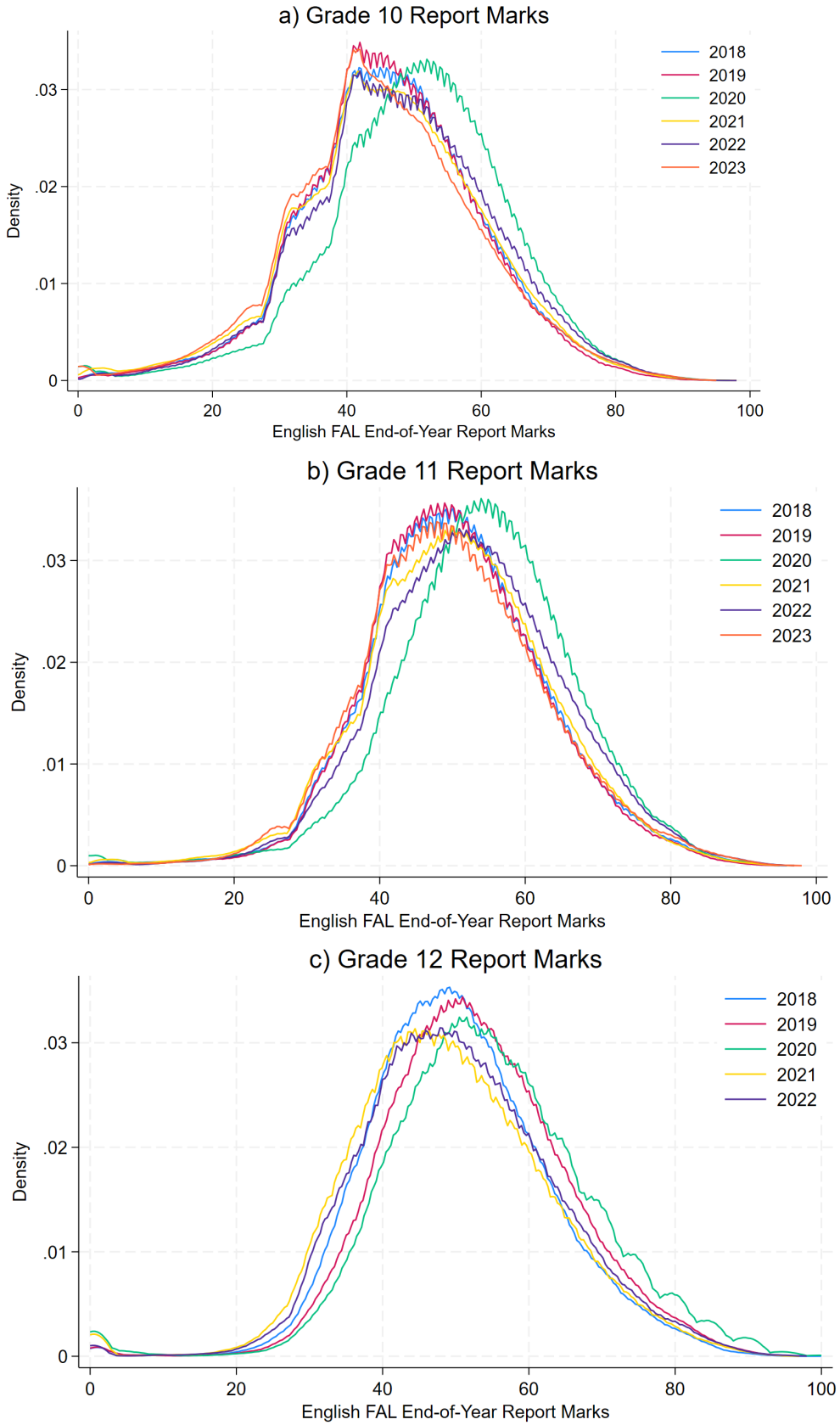


Figure 16. Kernel Density Estimates of English FAL End-of-Year Report Marks across years for a) Grade 10, b) Grade 11 and c) Grade 12.

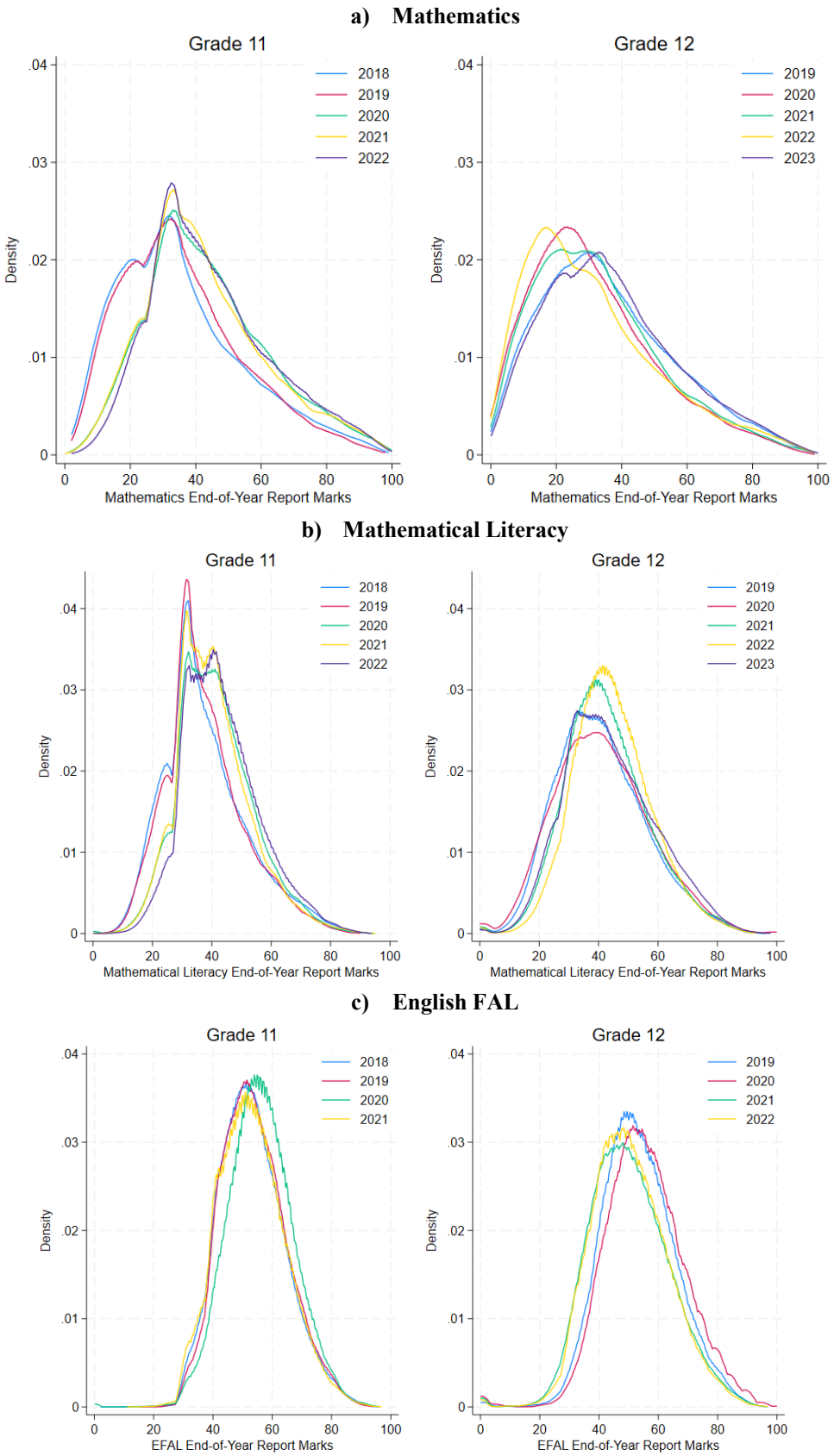


Figure 17. End-of-Year Report Marks for Grade 11 (2018 – 2022) and Grade 12 (2019 – 2023) across Cohorts for a) Mathematics, b) Mathematical Literacy and c) English FAL, using the balanced SBA sample.

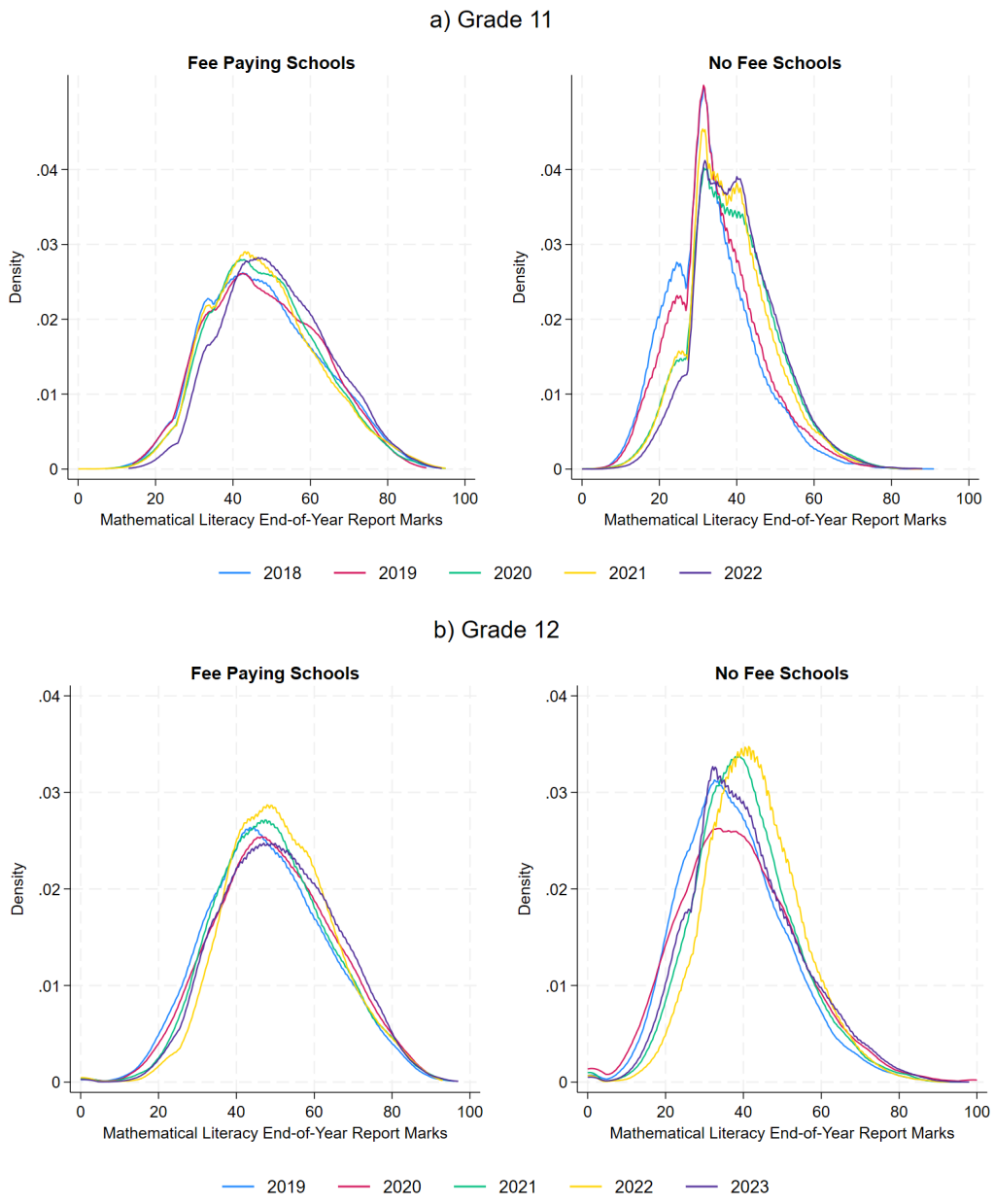


Figure 18. Mathematical Literacy Report Marks by Year and School Fee Status for a) Grade 11 and b) Grade 12

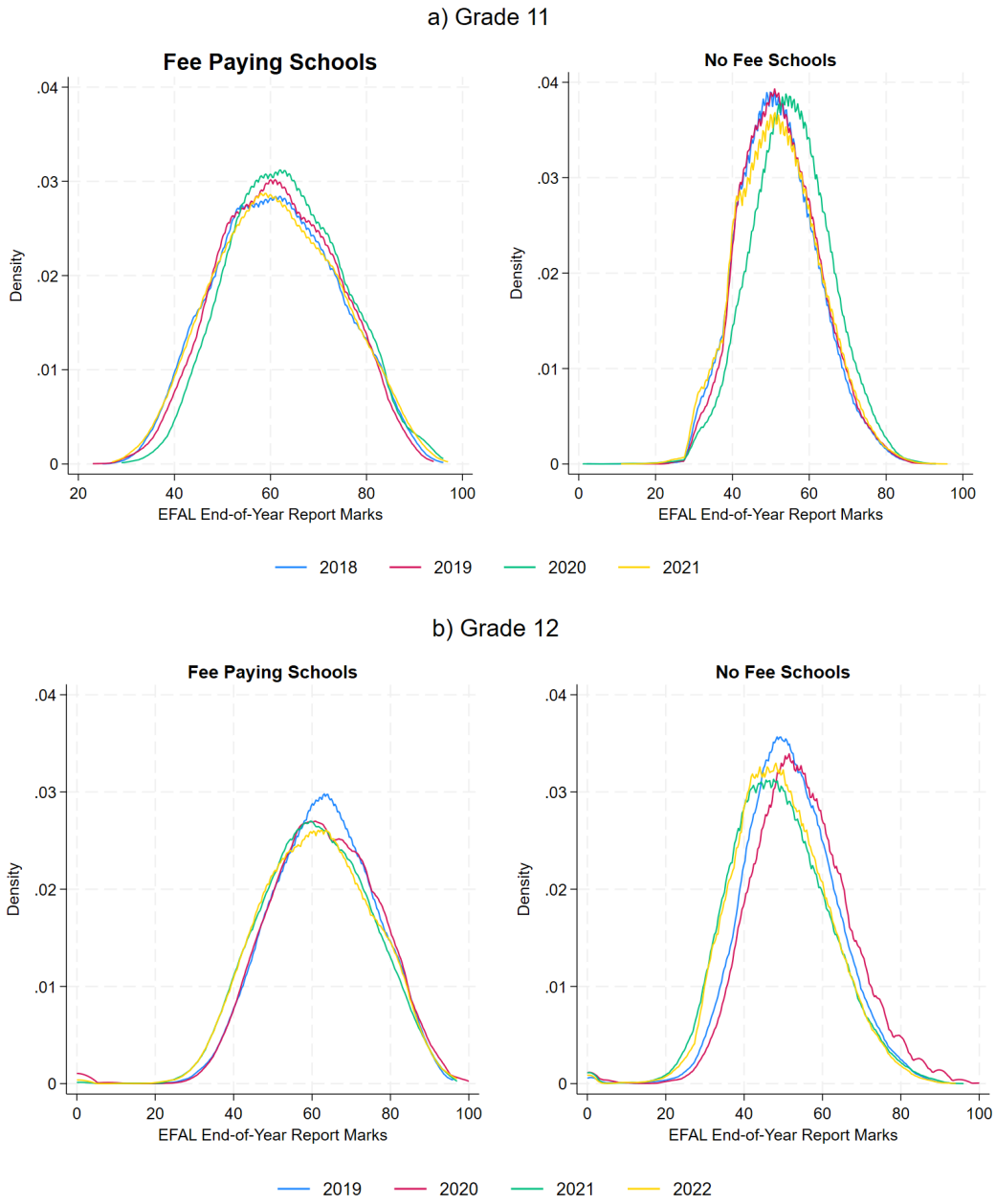


Figure 19. English FAL Report Marks by Year and School Fee Status for a) Grade 11 and b) Grade 12

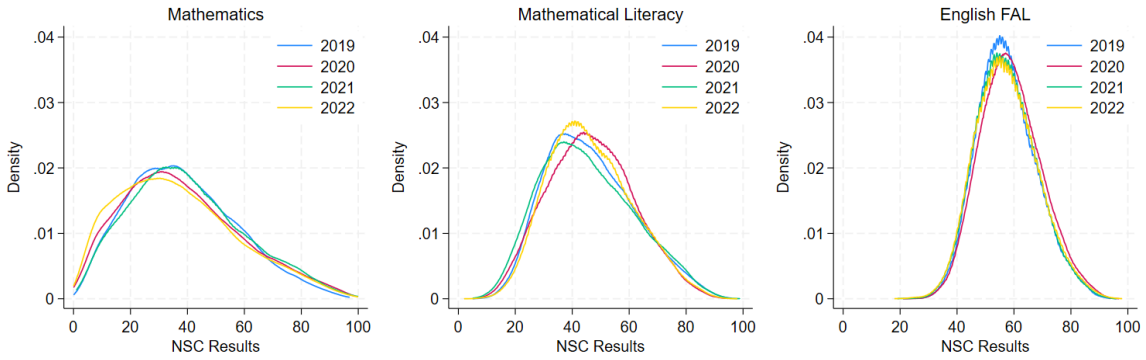


Figure 20. NSC Results across years for Mathematcis, Mathematical Literacy and English FAL.

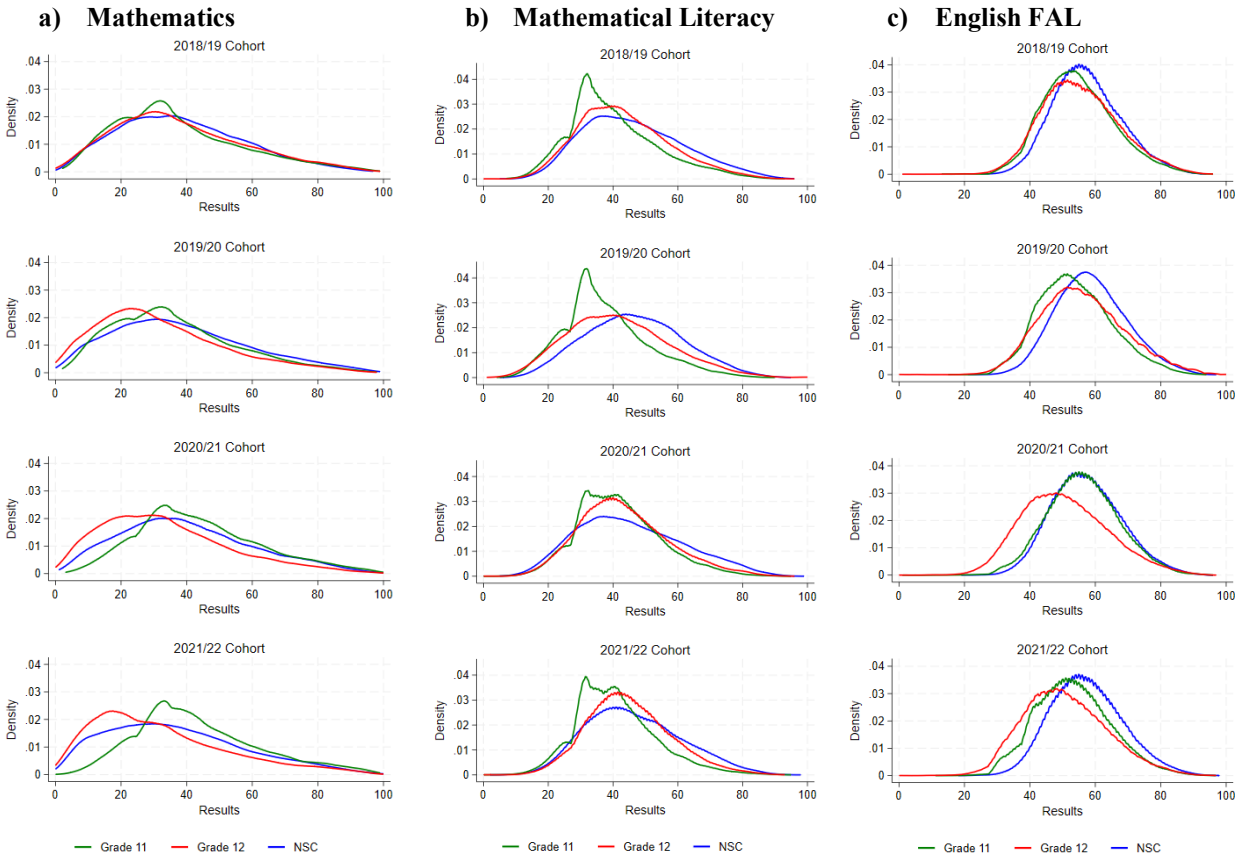


Figure 21. Kernel Density Estimates of NSC Results (blue line) and End-of-Year Report Marks for Grade 11(green line) and Grade 12 (red line) across cohorts for a) Mathematics, b) Mathematical Literacy and c) English FAL

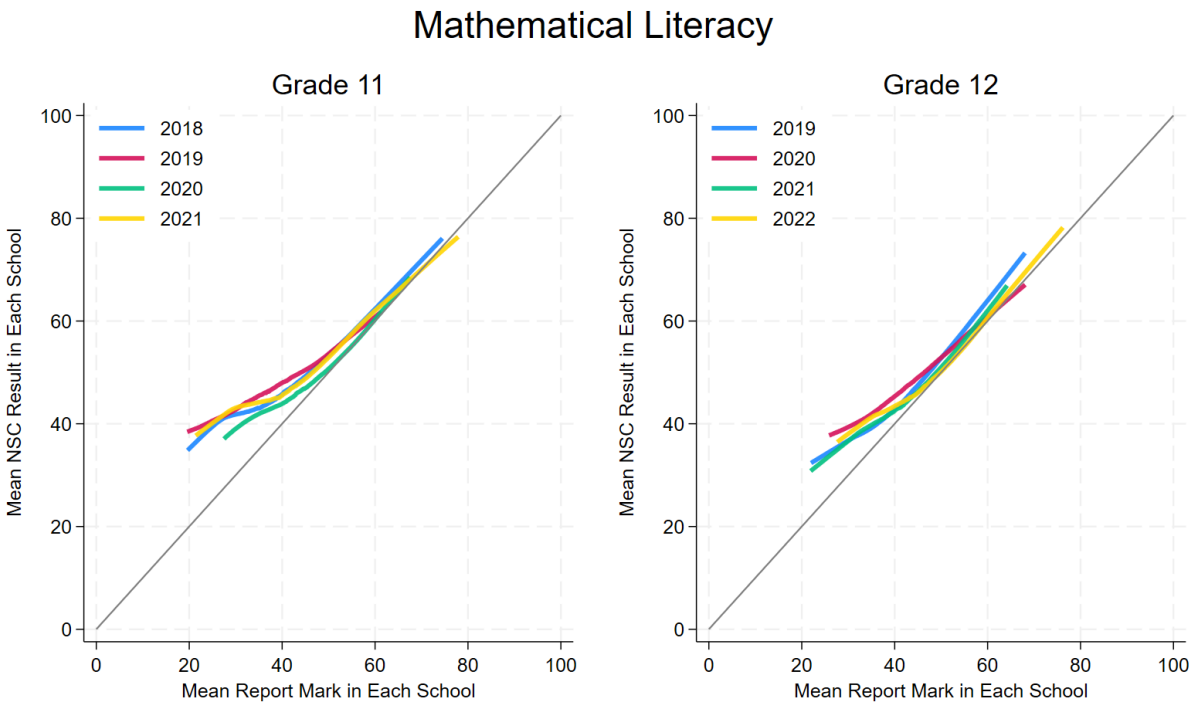


Figure 22. Lowess Curves of School Report Marks to NSC Results in Mathematical Literacy.

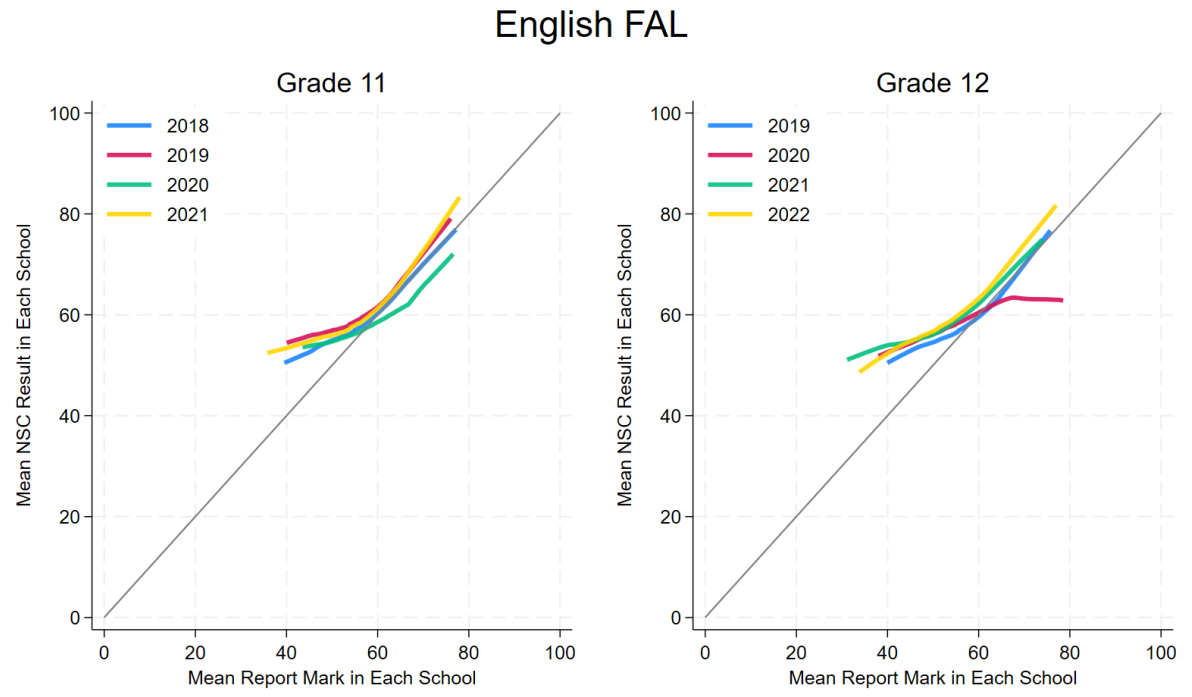


Figure 23. Lowess Curves of School Report Marks to NSC Results in English FAL.

Table 18. Descriptive differences between Mathematics learners at fee and no-fee schools, by Matric cohort

	2018/19			2019/20			2020/21			2021/22		
	Fee School	No-Fee School		Fee School	No-Fee School		Fee School	No-Fee School		Fee School	No-Fee School	
a) Mathematics												
Gr.11 Report Mark	44.90 (0.28)	32.07 (0.20)	***	46.12 (0.35)	31.44 (0.22)	***	49.52 (0.28)	40.63 (0.19)	***	50.44 (0.28)	39.33 (0.17)	***
<30%	0.22 (0.01)	0.47 (0.01)	***	0.17 (0.01)	0.48 (0.01)	***	0.14 (0.00)	0.25 (0.00)	***	0.13 (0.00)	0.25 (0.00)	***
Gr.12 Report Mark	44.02 (0.28)	34.07 (0.22)	***	41.16 (0.36)	27.65 (0.23)	***	40.11 (0.29)	29.83 (0.19)	***	39.45 (0.30)	27.79 (0.20)	***
<30%	0.25 (0.01)	0.46 (0.01)	***	0.31 (0.01)	0.62 (0.01)	***	0.34 (0.01)	0.55 (0.01)	***	0.37 (0.01)	0.61 (0.01)	***
NSC Result	45.15 (0.26)	34.23 (0.21)	***	46.06 (0.38)	34.66 (0.26)	***	46.18 (0.28)	36.58 (0.21)	***	44.27 (0.29)	32.71 (0.21)	***
Female	0.59 (0.01)	0.61 (0.01)	**	0.57 (0.01)	0.62 (0.01)	***	0.59 (0.01)	0.62 (0.01)	***	0.55 (0.01)	0.64 (0.01)	***
Over-Age	0.12 (0.00)	0.25 (0.01)	***	0.10 (0.01)	0.22 (0.01)	***	0.10 (0.00)	0.22 (0.00)	***	0.11 (0.00)	0.20 (0.00)	***
Eng/Afr HL	0.54 (0.01)	0.03 (0.00)	***	0.52 (0.01)	0.04 (0.00)	***	0.48 (0.01)	0.02 (0.00)	***	0.49 (0.01)	0.04 (0.00)	***
Proportion	40%	60%		33%	67%		37%	63%		37%	63%	
a) M. Literacy												
Gr.11 Report Mark	49.00 (0.17)	35.36 (0.09)	***	48.15 (0.26)	34.87 (0.10)	***	48.58 (0.16)	39.21 (0.08)	***	48.35 (0.14)	38.54 (0.06)	***
<30%	0.05 (0.00)	0.25 (0.00)	***	0.07 (0.00)	0.28 (0.00)	***	0.06 (0.00)	0.16 (0.00)	***	0.06 (0.00)	0.15 (0.00)	***
Gr.12 Report Mark	49.24 (0.17)	40.29 (0.11)	***	50.04 (0.27)	39.51 (0.14)	***	49.92 (0.17)	40.91 (0.09)	***	51.48 (0.14)	43.06 (0.07)	***
<30%	0.08 (0.00)	0.19 (0.00)	***	0.09 (0.01)	0.26 (0.00)	***	0.07 (0.00)	0.18 (0.00)	***	0.04 (0.00)	0.11 (0.00)	***
NSC Result	54.10 (0.18)	42.45 (0.12)	***	54.39 (0.25)	44.57 (0.14)	***	54.21 (0.18)	42.16 (0.11)	***	54.56 (0.15)	44.11 (0.08)	***
Female	0.57 (0.01)	0.54 (0.00)	***	0.57 (0.01)	0.55 (0.00)	*	0.55 (0.01)	0.55 (0.00)		0.54 (0.01)	0.54 (0.00)	
Over-Age	0.34 (0.01)	0.51 (0.00)	***	0.32 (0.01)	0.48 (0.00)	***	0.33 (0.01)	0.48 (0.00)	***	0.34 (0.00)	0.48 (0.00)	***
Eng/Afr HL	0.49 (0.01)	0.04 (0.00)	***	0.46 (0.01)	0.04 (0.00)	***	0.47 (0.01)	0.04 (0.00)	***	0.43 (0.01)	0.06 (0.00)	***
Proportion	34%	66%		21%	79%		27%	73%		23%	77%	
a) English FAL												
Gr.11 Report Mark	63.00 (0.17)	53.45 (0.06)	***	62.70 (0.18)	52.56 (0.06)	***	63.94 (0.19)	55.37 (0.06)	***	62.33 (0.18)	52.37 (0.05)	***
<30%	0.00 (0.00)	0.00 (0.00)	**	0.00 (0.00)	0.00 (0.00)	*	0.00 (0.00)	0.00 (0.00)	***	0.00 (0.00)	0.00 (0.00)	***
Gr.12 Report Mark	64.64 (0.17)	53.81 (0.06)	***	64.46 (0.20)	54.71 (0.08)	***	61.77 (0.22)	49.39 (0.07)	***	62.13 (0.19)	49.69 (0.05)	***
<30%	0.00 (0.00)	0.00 (0.00)	***	0.00 (0.00)	0.01 (0.00)	***	0.00 (0.00)	0.03 (0.00)	***	0.00 (0.00)	0.03 (0.00)	***
NSC Result	67.42 (0.15)	55.66 (0.06)	***	68.62 (0.17)	57.25 (0.06)	***	67.76 (0.18)	56.12 (0.05)	***	69.08 (0.16)	56.33 (0.04)	***
Female	0.57 (0.01)	0.56 (0.00)		0.54 (0.01)	0.57 (0.00)	***	0.57 (0.01)	0.57 (0.00)		0.57 (0.01)	0.57 (0.00)	
Over-Age	0.18 (0.01)	0.43 (0.00)	***	0.17 (0.01)	0.41 (0.00)	***	0.17 (0.01)	0.42 (0.00)	***	0.19 (0.01)	0.41 (0.00)	***
Eng/Afr HL	0.88 (0.00)	0.02 (0.00)	***	0.87 (0.01)	0.02 (0.00)	***	0.89 (0.00)	0.01 (0.00)	***	0.82 (0.01)	0.01 (0.00)	***
Proportion	15%	85%		14%	86%		11%	89%		9%	91%	

Source: Author's calculations using DDD and 2023 school master list data. Note: Standard deviations are in parentheses. Asterisks indicate statistically significant differences between fee and no-fee-paying schools at * p<0.1, ** p<0.05, *** p<0.001. p-value from a pooled t-test.

Table 19. Lower and Upper Bounds for NSC Performance Terciles by Year and Subject.

		Mathematics				Mathematical Literacy				English FAL			
		Year				Year				Year			
Tercile		2018	2019	2020	2021	2018	2019	2020	2021	2018	2019	2020	2021
1	Minimum	0	0	1	0	5	5	5	2	25	21	19	18
	Maximum	28	27	30	25	38	40	37	39	52	54	52	52
2	Minimum	29	28	31	26	39	41	38	40	53	55	53	53
	Maximum	46	46	47	44	52	53	52	52	61	63	62	62
3	Minimum	47	47	48	45	53	54	53	53	62	64	63	63
	Maximum	97	99	100	100	96	95	99	98	94	97	96	98

8.4 APPENDIX FOR SECTION 6: REGRESSION ANALYSIS

Table 20. Hierarchial Linear Models for Gaps between learners' Grade 11 and 12 Report Marks and NSC Results

	Mathematics		Math. Literacy		English FAL	
	(1)	(2)	(3)	(4)	(5)	(6)
	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap	Gr.11 Gap	Gr.12 Gap
2019/20	-0.420***	2.797***	1.073***	1.289***	1.130***	1.521***
	(0.11)	(0.07)	(0.08)	(0.06)	(0.04)	(0.04)
2020/21	2.707***	2.594***	-0.0445	0.897***	0.537***	2.483***
	(0.10)	(0.07)	(0.07)	(0.05)	(0.04)	(0.04)
2021/22	3.120***	2.156***	-0.395***	0.336***	1.478***	2.317***
	(0.10)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)
No-Fee School	1.965***	0.919***	3.436***	0.287	0.0325	0.301
	(0.35)	(0.19)	(0.28)	(0.16)	(0.28)	(0.35)
Tercile 2	-1.409***	2.521***	2.319***	0.155***	0.834***	0.682***
	(0.08)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)
Tercile 3	-0.0892	3.470***	7.963***	2.040***	1.916***	0.640***
	(0.09)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
Female	-0.557***	-0.384***	-0.594***	-0.233***	-0.485***	-0.0967***
	(0.07)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)
Eng/Afr HL	-0.217	-0.206*	-0.431***	0.115	-0.561***	0.453***
	(0.14)	(0.09)	(0.09)	(0.06)	(0.13)	(0.12)
Overage	0.231**	-0.145*	-0.127**	0.00390	0.0974***	0.219***
	(0.09)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)
Constant	7.127***	2.823***	4.551***	5.051***	5.503***	4.792***
	(0.31)	(0.18)	(0.25)	(0.15)	(0.27)	(0.34)
N	47857	47857	104504	104504	157069	157069

Notes: Standard errors are in parentheses. Random effects parameters are not reported above. Significant at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$