RESEP
Research on Socio-Economic Policy

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

forward together
sonke siya phambili
saam vorentoe

RESEP POLICY BRIEF

Department of Economics, Stellenbosch University

Policy Brief No.
06/25

NOVEMBER 2025

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

Addressing spatial inequalities in early grade literacy

Author

Julian Golmann

OPTIMA

Policy brief: Addressing spatial inequalities in early grade literacy

Date: November 2025

Author: J.G. Goldmann, RESEP, Department of Economics, Stellenbosch University

EXECUTIVE SUMMARY

Despite near-universal primary school enrolment (97%), South Africa's education system faces a foundational literacy crisis: By Grade 4, 81% of learners cannot read for meaning and only 1.6% reach the PIRLS high benchmark. These weak outcomes limit future learning and depress labour-market prospects, perpetuating intergenerational poverty. Yet this "national crisis" is not experienced equally across space.

The evidence presented here shows that where a child lives is nearly as predictive of literacy outcomes as socioeconomic status. The urban-rural gap of 61.7 points is equivalent to roughly 1.5 years of learning. A decomposition analysis shows that language accounts for 55%, with school quality and peer socioeconomic status also making large and significant contributions to the observed difference.

This spatial divide in educational outcomes does not simply reflect geography but reflects the compounding effects of multiple disadvantages. Paradoxically, rural schools exhibit striking equality, but it is an equality of shared disadvantage, with most learners clustered far below basic proficiency. Urban schools, by contrast, contain both the system's strongest performers and its most disadvantaged, resulting in much wider inequality.

Quantile regression analysis highlights that factors shaping achievement differ across the performance distribution, and between urban and rural contexts. High-performing urban learners face steep penalties associated with language, background and age, while rural learners experience uniformly low outcomes regardless of characteristics. This reflects deep, systemic dysfunction.

These findings make one conclusion unavoidable: rural and urban schools face fundamentally different constraints. Meaningful progress in literacy will require targeted, spatially differentiated interventions, not one=size-fits all national solutions.

1 UNDERSTANDING SPATIAL INEQUALITIES IN EDUCATIONAL ATTAINMENT

South Africa's literacy outcomes rank among the worst globally despite significant education expenditure (Van Staden, 2023). For the purposes of this analysis the PIRLS 2016 data was utilised, as the PIRLS 2021 data had a number of issues and quirks as a result of COVID-19 disruptions, particularly for the rural subset of the sample. The PIRLS 2016 data shows that by Grade 4, only 21.2% of learners reach the PIRLS Low Benchmark, meaning the vast majority cannot read for meaning at the point where they must shift from "learning to read" to "reading to learn". Figure 1 illustrates how South African learners' performance is distributed across PIRLS literacy scores; only 7.4% reach the Intermediate Benchmark, and only 1.6% reach the High Benchmark. This has significant implications for:

- **Economic Impact:** Illiteracy limits labour market participation and productivity. (Hanushek, 2012)
- **Social Mobility**: Education remains the primary mechanism for escaping poverty.
- **Intergenerational Transmission**: Low parental literacy perpetuates disadvantage, creating a vicious cycle of poverty and disadvantage.

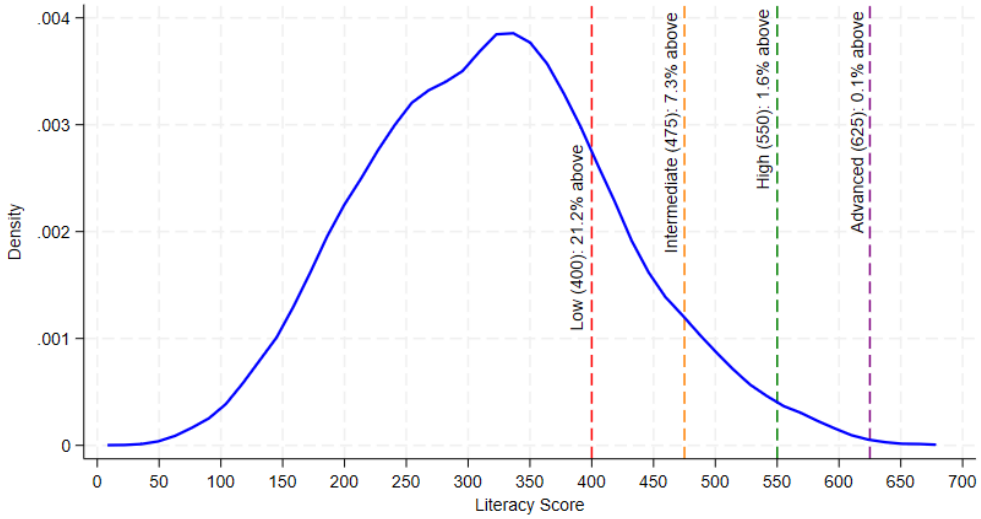

Figure 1 Distribution of PIRLS Reading Achievement Scores with International Benchmarks
Source: PIRLS 2016

While inequality in early grade literacy in South Africa has traditionally been studied along racial and socioeconomic lines, the geographical dimension remains critically understudied. Where a child lives — and therefore the type of school they attend — is as predictive of literacy outcomes as socioeconomic status or home background.

Table 1 illustrates that the gap in mean scores of 61.7 points reflects more than simple geographic differences. Rural learners are concentrated below basic proficiency, with close to 90% performing

below the Low International Benchmark. Rural learners almost entirely attend no-fee schools and are more likely to be over-age, both factors that correlate with weaker performance. Rural learners are also overwhelmingly African-language speakers, a marker of both linguistic and socioeconomic marginalisation.

Table 1 Summary Statistics by Geographic Location

Geography	Mean PIRLS Literacy Score	% Above Lower Int. Benchmark	% African Language Speakers	% Overage Learners	% Learners Attending No Fee Schools
Urban	356.38	35.48%	48.29%	26.70%	34.85%
Rural	294.70	11.15%	85.78%	32.24%	94.04%

Source: own calculations using PIRLS 2016

These characteristics do not operate in isolation but compound within rural schooling contexts: weaker school infrastructure, limited access to learning materials, greater poverty concentration, and the absence of high-performing peer environments all reinforce one another (Christie, 2007). The result is a schooling environment where even the most capable learners face systemic barriers to literacy acquisition.

2 KEY FINDINGS

2.1 Inequality Across the Test Score Distribution

Average literacy scores and traditional analyses conceal important variation in how different factors shape learners' outcomes. Given South Africa's exceptionally high levels of educational inequality, it is unlikely that the same determinants of literacy operate uniformly across all learners. Traditional regression approaches that estimate a single average effect obscure these nuances (Koenker, 2001).

To better understand these dynamics, this study uses quantile regression to examine how key demographic and school-related factors influence literacy at different points of the performance distribution. Figures 2(a) and 2(b) plot these effects separately for urban and rural learners. The results reveal two core insights:

a) **The effects of learner characteristics differ significantly across the distribution:** Among urban learners, the penalties associated with factors such as speaking an African home language or being over-age become substantially larger at the top end of the distribution. For example, high-performing urban learners who speak an African language face penalties of nearly 100 points, equivalent to several years of learning. This indicates that even capable learners struggle to convert their potential into high literacy outcomes under current conditions.
b) **Rural learners experience uniformly suppressed outcomes:** the slopes for rural learners are far flatter. This means that characteristics like language, age, or

household disadvantage matter less in differentiating performance, not because they are unimportant, but because structural barriers depress learning for nearly all rural children.

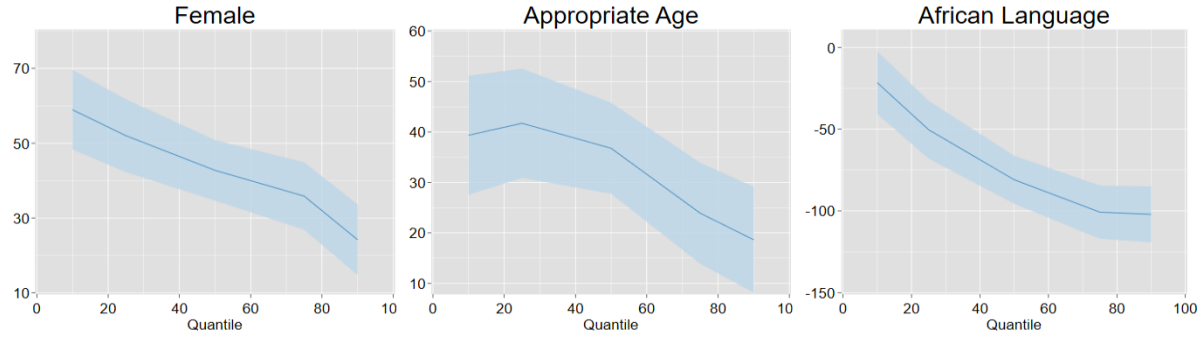

Figure 2(a) Quantile Regression Plots of Demographic Variables for Urban Learners

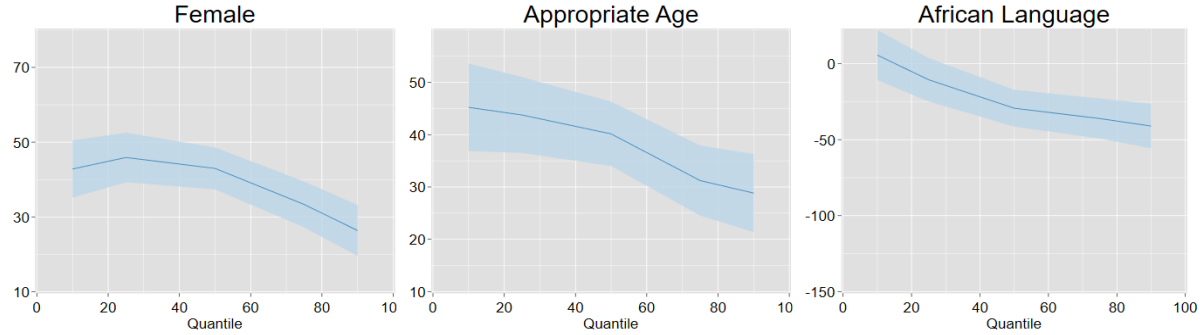

Figure 2(b) Quantile Regression Plots of Demographic Variables for Rural Learners

Taken together, the distributional patterns reinforce a central conclusion: Urban inequality is characterised by wide variation and steep penalties for certain groups, while rural inequality is characterised by the uniformity of disadvantage.

This analysis shows that there are **fundamental differences** between urban and rural schools, and their underlying conditions. Understanding these differences is crucial for designing **targeted policy responses that match the realities of each context.**

2.2 Language is a Key Driver

To understand *why* urban and rural learners perform so differently, this study uses a decomposition analysis. In simple terms, a decomposition analysis helps identify the reasons behind the urban–rural literacy gap. Instead of just showing that a gap exists, it separates the total difference into contributions from various factors; for example, how much of the gap is due to language, how much to school quality, and how much to

differences in poverty levels. This provides a clearer picture of the root causes and helps guide targeted policy responses.

The decomposition results show that language is the single largest contributor to the urban–rural literacy gap, accounting for 55% of the difference in scores. This reflects more than just the language spoken at home: It captures deep, structural inequalities in the quality of schooling available to African-language–speaking learners in rural areas. For example, a lack of teachers with strong training in early-grade literacy pedagogy in these languages, and an uneven or limited distribution of high-quality reading materials.

Table 2 Summary of Oaxaca Blinder Decomposition Results

What Explains the Gap?	Literacy Score Points	% of Gap	What This Means
Language	-32***	55%	A much higher concentration of African speaking learners in rural areas. Language acts as a proxy of socioeconomic disadvantage.
School Quintile	-20***	35%	The quality of one's school plays a significant role in learners' literacy achievement.
Peer Effects of Wealth	-7**	12%	Rural students attend schools where learners are much poorer on average than urban schools.
TOTAL GAP	**-58***		

Source: own calculations using PIRLS (2016).
Note that decomposition analysis has positive and negative coefficients which interact to determine the total effect, therefore when selecting only a few of the variables it is possible to have more than 100% of the gap or literacy score points because this will be counterbalanced by some slight positive effects.

Furthermore, school quality (proxied by school quintile) accounts for 20 points of the gap, equivalent to roughly half a year of learning. Peer effects of wealth that capture the relative socioeconomic status of classmates also play a meaningful role, contributing about 12% of the total gap. Learners in rural classrooms are surrounded by peers facing much deeper poverty, which lowers the overall learning environment and further depresses literacy outcomes.

3 RECOMMENDATIONS

The evidence presented in this brief points to clear policy implications and areas for future research:

3.1 Address Systemic Inequalities in No Fee Schools

With language accounting for 55% of the gap between urban and rural schools, it is important to recognize that, in the absence of a variable denoting population group or race, language serves as a proxy for deep historical disadvantages, and this is indicating

that despite progress, the legacy of historically disadvantaged schools endure. Policy interventions, therefore, need to focus on: (1) more progressive resource allocation, rather than equal allocation which simply maintains the status quo; (2) high-quality learning materials and more equitable language policy which will foster literacy development in all 11 official languages; (3) teacher development and training, and incentives for teachers to work in rural areas.

3.2 Differentiated Support for Urban vs. Rural Schools

The finding that rural and urban schools demonstrate such different effects across the literacy score distribution indicates that policy and interventions designed to tackle problems at a national level are likely to be inefficient, slow and ineffective. Both urban and rural schools require targeted interventions, such as intensive investment in teacher development focused on pedagogy in rural schools, and differential remuneration packages for both rural schools and selected (no fee) urban schools in order to secure and retain high quality teachers.

3.3 Targeted Interventions by Achievement Level

The analysis in this study reveals that different students across the performance distribution and across geographic areas face markedly different barriers in literacy attainment. Low-performing learners need to be shielded from systemic disadvantages and avoid grade repetition, while high-performing or high-potential learners are constrained by limited resources and inadequate teaching quality, resulting in ceilings which learners are not able to break through. Effective policy needs to be aligned according to the specific issues it is attempting to address, and it is therefore necessary to recognize that what is needed for high-potential rural learners will likely be completely irrelevant to low-performing urban learners (Banerjee, 2007).

3.4 Areas for Future Research

A number of critical research questions remain unanswered and hold great potential for informing policy decisions:

- Disentangling language from disadvantage: Does the current language policy contribute independently to achievement gaps, or is it simply a proxy for historical disadvantage?
- Rural school interventions: Which specific interventions have been proven to have high returns in low-resource and socioeconomically disadvantaged areas?

- Spillovers in Education: Is there any evidence of spillover effects or negative externalities in early grade literacy outcomes in South Africa? Are these inequalities consistent across different levels of geography, i.e. do the same effects hold when we compare urban and rural within one province, or within one district, or even within one municipality?

South Africa's literacy crisis is not an inevitable consequence of poverty or complexity, this study shows that spatial inequalities in early grade literacy are structured, and measurable, and therefore we are able to address them.

REFERENCES

Banerjee, A. V. (2007). Remedying education: Evidence from two randomized experiments in India. . *The quarterly journal of economics, 122(3)*, 1235-1264.

Christie, P. B. (2007). *Schools that work.* Pretoria: Government Printer: Report of the Ministerial Committee.

Hanushek, E. A. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of economic growth, 17(4)*, 267-321.

Koenker, R. &. (2001). Quantile regression. *Journal of economic perspectives*, 143-156.

Van Staden, S. R. (2023). *PIRLS 2021: South African children's literacy achievement.* Centre for Evaluation and Assessment.