



UNEQUAL CAPACITY, UNEQUAL OUTCOMES? SUBNATIONAL CAPACITY AND JUNIOR-SECONDARY EDUCATION IN INDONESIA

Iik Nurulpaik^{*1}, Ajat Sudrajat², Lutfi Nur³

¹Universitas Pendidikan Indonesia, Indonesia.

²Universitas Terbuka, Indonesia.

³Universitas Pendidikan Indonesia, Indonesia.

Email: iik.nurulpaik@upi.edu ^{*1}, ajats@ecampus.ut.ac.id ², lutfinur@upi.edu ³

Abstract: This study examines whether unequal subnational capacity is associated with unequal educational participation and measured learning outcomes in a decentralised education system. It reconstructs a balanced panel of 27 districts and municipalities in West Java, Indonesia, from 2015 to 2019 (N = 135). A 15-indicator Structural Capacity Index combines educational resources, territorial accessibility, socioeconomic capacity, and population quality. Pooled ordinary least squares and district and year fixed-effects models are estimated with district-clustered standard errors; examination integrity is included in the achievement model. Capacity is positively associated with junior-secondary gross enrolment in pooled estimates, but the relationship disappears under fixed effects. No robust association is found between capacity and National Examination scores. These results distinguish capacity accumulation from capacity conversion: structural advantages separate jurisdictions, but short-term changes in capacity do not automatically generate improvements in access or learning. Capacity appears easier to convert into participation than into measured learning, which depends on instructional processes and assessment credibility. The findings support capacity-sensitive and asymmetric intergovernmental policy in decentralised systems.

Keywords: Subnational Capacity; Education Decentralisation; Educational Participation; Learning Outcomes; Indonesia.

INTRODUCTION:

Decentralisation is often expected to make public services more responsive because local governments possess better information about local needs and can adapt policy accordingly. Yet decentralisation also transfers responsibility to jurisdictions with unequal fiscal resources, administrative competence, infrastructure, and social conditions. When authority is devolved more evenly than capacity, the same policy mandate may produce uneven educational opportunities and outcomes.

This problem is especially consequential in education. Local governments do not merely finance schools; they shape the availability and distribution of teachers, classrooms, infrastructure, school quality assurance, and the social conditions under which families sustain participation. Subnational inequality therefore enters education through both the supply of services and households' capacity to use them.



Indonesia provides a theoretically relevant setting. Since the early 2000s, districts and municipalities have exercised extensive responsibility for basic education, while the central government retains regulatory, financing, and equalisation roles. Decentralisation was accompanied by major growth in education expenditure and enrolment, yet learning gains remained limited and subnational governments differed sharply in planning, budget execution, teacher management, and the use of non-salary spending (World Bank, 2020a, 2020b). This institutional arrangement creates a test of a broader question: when authority is devolved, under what conditions does local capacity translate into educational access and learning?

Existing research frequently examines school resources, poverty, teacher quality, infrastructure, or household conditions separately. Recent empirical work has also shown that the benefits of decentralisation are conditional rather than automatic. Administrative decentralisation in Colombia improved some teacher and student outcomes where local authority was accompanied by implementation capacity (Elacqua et al., 2021); historical local state capacity shaped divergent public-goods trajectories after decentralisation in Russia (Foa, 2022); and evidence from Indian states indicates that gradual fiscal decentralisation can support health and education delivery when institutional conditions are sufficiently developed (Singh et al., 2024). These studies suggest that decentralisation is best understood as a capacity-dependent reform, but they do not directly test a multidimensional measure of subnational capacity against both access and learning outcomes.

The article makes four contributions. First, it operationalises subnational capacity through a 15-indicator index while keeping educational outcomes outside the capacity construct. Second, it separates capacity accumulation from capacity conversion: districts may possess structural advantages without converting them efficiently into access or learning. Third, it distinguishes persistent between-place inequality from short-term within-place change by comparing pooled and fixed-effects estimates. Fourth, it extends emerging work on the educational “middle tier” by treating district-level capacity as a measurable system condition linking national policy to school-level implementation (Tournier et al., 2025).

The study addresses three questions: (1) Is structural subnational capacity associated with junior-secondary participation? (2) Is it associated with National Examination performance after examination integrity is taken into account? and (3) Is capacity more readily converted into participation than into measured learning outcomes?

THEORETICAL FRAMEWORK:

From state capacity to subnational capacity

State capacity refers broadly to the ability of public institutions to formulate collective goals, mobilise resources, implement decisions, and produce public goods. Fukuyama (2013) distinguishes government capacity from the scope of state activity: governments may be responsible for many functions yet remain unable to implement them effectively. Cingolani (2013) similarly treats capacity as a multidimensional set of institutional competences rather than a single fiscal or bureaucratic attribute.



Besley and Persson (2009, 2011) emphasise that capacity is accumulated through long-term investments in fiscal, legal, and collective capabilities. This perspective is important because it separates resources from the ability to use them. Budgets, personnel, and infrastructure are potential inputs; capacity is revealed when institutions organise, allocate, and convert those inputs into services and outcomes.

At the subnational level, capacity can be defined as a territory's structural and institutional ability to transform public resources and development conditions into effective public services. It includes, but is not limited to, local bureaucratic competence. Economic conditions, population education, health, and territorial accessibility affect what governments can deliver and what households can use.

This definition also separates capacity from performance. Capacity denotes the potential and ability to act; performance denotes the observed result. A district may possess relatively strong capacity but convert it inefficiently, while a district with moderate resources may perform well through effective coordination and implementation.

Decentralisation, unequal capacity, and education policy

Fiscal federalism theory argues that decentralised provision can improve allocative efficiency when local governments match services to local preferences and conditions (Oates, 1972). However, these benefits depend on administrative competence, accountability, fiscal adequacy, and intergovernmental coordination. Decentralisation can therefore increase responsiveness and inequality at the same time.

In education, local capacity shapes policy implementation through planning, budgeting, teacher management, infrastructure, data use, and accountability. Evidence from Indonesia has shown substantial variation in local education governance, including differences in planning, transparency, management information, and teacher deployment (World Bank, 2013). This suggests that decentralisation does not create a uniform implementation environment.

Unequal capacity also creates an evaluative problem. Jurisdictions are often compared using common outcome targets even though they face different starting conditions and implementation constraints. A capacity-sensitive analysis asks not only whether outcomes differ, but whether differences reflect unequal stocks of resources, unequal conversion efficiency, or both.

Recent empirical evidence on decentralisation and local capacity

Recent empirical research rejects the assumption that decentralisation uniformly improves public-service outcomes. Elacqua et al. (2021) exploit an administrative population threshold in Colombia and show that greater municipal decision-making authority altered teacher quality and student outcomes, demonstrating that local control can matter when subnational institutions are capable of acting. Their findings support the proposition that authority must be accompanied by organisational capacity.



Foa (2022) reaches a complementary conclusion from post-Soviet Russia. Decentralisation produced increasingly unequal public-goods provision because jurisdictions entered reform with different stocks of historical state capacity, cohesion, and governing capability. This evidence is directly relevant to educational decentralisation: formal mandates may be common, but inherited capacity shapes what local governments can implement.

Singh et al. (2024) find that fiscal decentralisation in Indian states is more effective as a gradual, institutionally embedded process than as a discrete policy shift. Their results reinforce the view that decentralisation is not a binary institutional treatment; its consequences depend on the accumulation of fiscal, administrative, and implementation capabilities.

Indonesia-specific evidence points in the same direction. The World Bank (2020b) reports that many districts plan around input quantities rather than learning outputs, face uneven budget execution, and devote limited fiscal space to non-salary programmes that may improve learning. The broader sector review concludes that expanded spending and enrolment have not been matched by equivalent improvements in learning (World Bank, 2020a). Together, these studies motivate an analysis that distinguishes the stock of subnational capacity from the efficiency with which that capacity is converted into participation and achievement.

A related contemporary literature focuses on the education system's "middle tier"—district, regional, and other intermediate actors that translate national policy into support for schools. Tournier et al. (2025) argue that these actors require clear mandates, financing, professional capability, data use, coordination, and organisational support. The present study complements that institutional framework by constructing a territorial index of the structural conditions within which middle-tier actors operate.

Educational production and capacity conversion

Educational production-function research conceptualises achievement and participation as outputs generated by school, family, peer, and institutional inputs (Hanushek, 1979). At the district level, subnational capacity represents a macro-level bundle of conditions that shape the production environment of schools.

For participation, conversion pathways are relatively direct. Budgets, teachers, classrooms, school accreditation, and roads influence service availability and accessibility. Poverty and purchasing power affect direct costs and opportunity costs, while adult education and health influence household aspirations and students' ability to remain enrolled.

For learning outcomes, conversion is more demanding. Qualified and competent teachers must be distributed effectively; resources must reach classrooms; school leadership must organise instruction; and students must experience sustained opportunities to learn. Consequently, the same stock of capacity may generate different learning outcomes across jurisdictions.



Assessment credibility further complicates the interpretation of examination scores. If the integrity of test administration varies across places or over time, higher scores do not necessarily indicate stronger learning. Examination integrity is therefore treated as a control in the achievement model rather than as part of structural capacity.

Hypotheses

*H*₁. Structural subnational capacity is positively associated with junior-secondary gross enrolment.

*H*₂. Structural subnational capacity is positively associated with National Examination scores after examination integrity is controlled.

*H*₃. The association between capacity and enrolment is stronger and more robust than the association between capacity and examination performance.

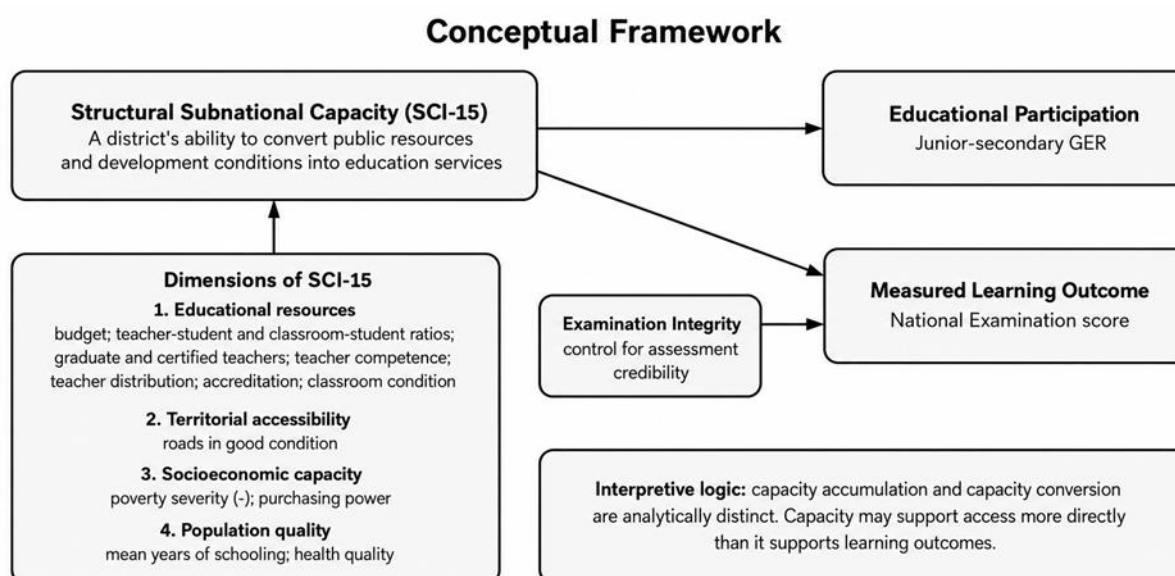


Figure 1. Conceptual framework

METHODOLOGY:

Data and unit of analysis

The analysis uses a reconstructed balanced panel of all 27 districts and municipalities in West Java for 2015-2019, yielding 135 district-year observations. The source archive was compiled for a 2021 research project from official administrative and statistical series. During reconstruction, district names, year labels, numeric formats, and indicator direction were harmonised, and each district-year was checked for completeness before estimation. The final panel contains no missing district-year outcome observations.

West Java is analytically useful because it combines large metropolitan municipalities, industrial districts, rural and mountainous areas, and substantial variation in poverty, infrastructure, educational provision, and human development. The study is historical by design: the period precedes the discontinuation of Indonesia's National Examination and therefore provides a consistent cross-jurisdictional outcome measure together with an integrity index.



Construction of the Structural Capacity Index

The Structural Capacity Index (SCI-15) contains 15 indicators organised into four dimensions. Each indicator is standardised within the full district-year panel as a z score. Indicators whose higher values represent lower capacity—poverty severity and unfavourable teacher-student and classroom-student ratios—are reverse-coded. Dimension scores are calculated as the mean of their standardised indicators, and SCI-15 is the unweighted mean of the four dimension scores. This dimension-balanced procedure prevents the educational-resource dimension, which contains more indicators, from mechanically dominating the index. A second index that assigns equal weight to all 15 indicators is used as a robustness check.

Table 1. Dimensions and Indicators of SCI-15

Dimension	Indicators	Expected pathway
Educational resources	Education budget; teacher-student ratio; classroom-student ratio; graduate teachers; certified teachers; teacher competence; distribution of examination-subject teachers; distribution of subject teachers; school accreditation; classrooms in good condition	Service availability and instructional conditions
Territorial accessibility	Roads in good condition	Physical access and travel costs
Socioeconomic capacity	Poverty severity (reversed); purchasing power	Household affordability and opportunity costs
Population quality	Mean years of schooling; health quality	Educational aspirations, support, and readiness

Outcomes and control

Junior-secondary gross enrolment rate (GER) is the participation outcome. GER can exceed 100 because it includes students outside the official age range and may also reflect cross-district mobility; it is therefore interpreted as system-level participation capacity rather than a precise cohort enrolment probability.

The National Examination score is the historical proxy for measured learning. The Examination Integrity Index is included as a control because changes in monitoring and test administration can alter observed scores independently of underlying learning.

Estimation strategy

Pooled ordinary least squares (OLS) estimates the combined cross-sectional and temporal association between SCI-15 and each outcome. District and year fixed-effects models then absorb time-invariant district characteristics and common annual shocks. Standard errors are clustered at the district level to allow arbitrary serial correlation and heteroskedasticity within jurisdictions. The fixed-effects coefficient is identified only by within-district change, whereas the pooled estimate also reflects persistent differences between districts.



The enrolment model is $GER_{it} = \alpha + \beta SCI_{it} + \mu_i + \lambda_t + \varepsilon_{it}$. The examination model is $Score_{it} = \alpha + \beta SCI_{it} + \gamma Integrity_{it} + \mu_i + \lambda_t + \varepsilon_{it}$. Because only five years are available and several capacity indicators change slowly, the estimates are interpreted as associations rather than causal effects.

Robustness checks use an equal-weight indicator index and a one-year lag of SCI-15. These checks assess whether the substantive conclusion depends on the dimension-balanced weighting scheme or contemporaneous timing.

Random-effects models are estimated as a complementary specification with year indicators. A Hausman specification test compares the capacity and integrity coefficients from fixed- and random-effects estimators. Because persistent district characteristics may correlate with capacity, fixed effects are retained as the conservative benchmark even when the Hausman test does not reject random effects.

Index validity is assessed in two additional ways. First, an equal-weight indicator index gives every standardised indicator the same weight. Second, principal component analysis (PCA) generates a data-driven index from the 15 oriented indicators. The first component is sign-adjusted so that higher scores represent greater capacity. Agreement among the dimension-balanced, equal-weight, and PCA indices is evaluated through correlations and model replication.

Additional analyses disaggregate SCI-15 into educational resources, territorial accessibility, socioeconomic capacity, and population quality. Heterogeneity is examined through an interaction between capacity and municipality status. One-year lagged models assess whether short delays alter the estimated relationship. These analyses are treated as robustness and mechanism tests rather than independent causal estimates.

Analytical transparency and scope of inference

The analysis was conducted on the reconstructed district-year panel rather than on five-year district averages. The principal estimates report coefficients, district-clustered standard errors, p values, observations, and model R-squared values. Because the sample covers the complete set of West Java districts but only five annual observations per district, statistical inference is treated cautiously, particularly for slowly moving structural indicators.

The models estimate conditional associations, not causal effects. Fixed effects reduce bias from stable unobserved district characteristics, but they do not remove time-varying confounding, measurement error, or reverse causality. The empirical contribution is therefore to distinguish cross-jurisdictional structural association from within-jurisdiction short-run change.

The study uses secondary aggregate data and contains no individual-level personal information. For reproducibility, the analytic data structure, variable dictionary, index-construction rules, and model specifications should accompany the submitted article as online supplementary material, subject to the reuse conditions of the original administrative sources.



RESULTS AND DISCUSSION:

Descriptive patterns

At the provincial level, the capacity index increased from 60.52 in 2015 to 66.09 in 2019. Junior-secondary GER remained close to 90, while examination integrity rose markedly. Examination scores, however, declined after 2016. The opposing movement of integrity and scores underscores the need to distinguish changes in learning from changes in assessment credibility.

Across districts and municipalities, capacity, enrolment, and examination performance varied substantially. Municipalities tended to have stronger infrastructure and human-development conditions, but higher capacity did not translate uniformly into higher participation or examination performance.

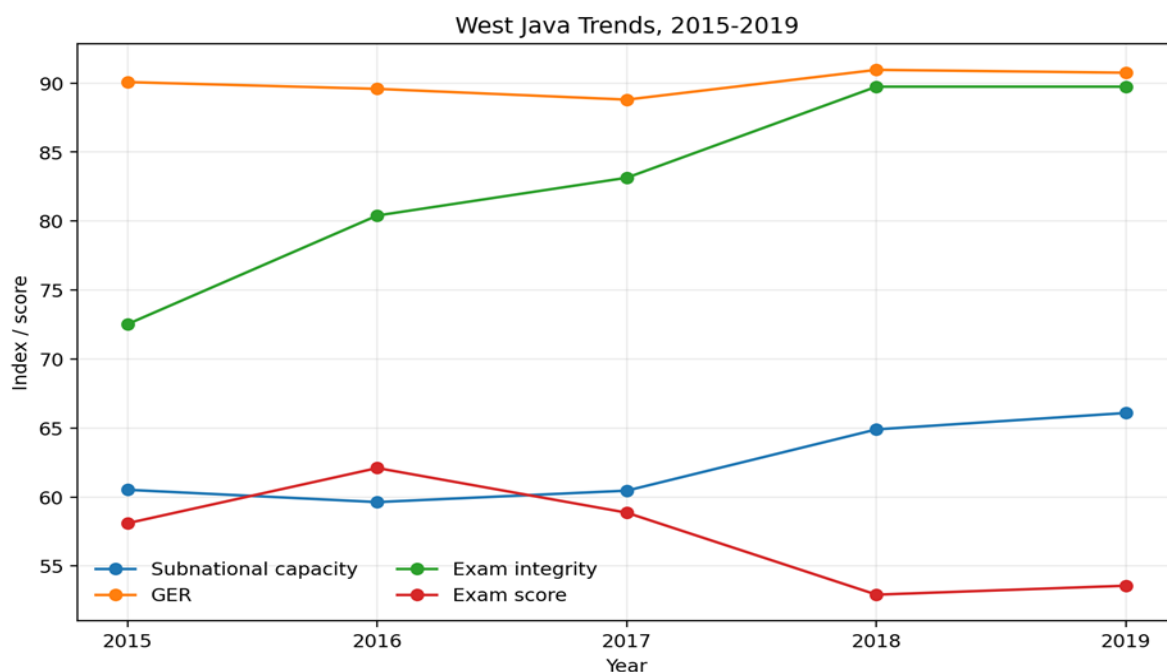


Figure 2. Provincial trends in capacity, participation, examination integrity, and scores

Main estimates

Table 2. Associations Between Structural Capacity and Educational Outcomes (N = 135)

Outcome	Model	SCI-15 β	SE	p	Integrity β	p	R ²
GER	Pooled OLS + year FE	2.90	1.04	.010	—	—	.100
GER	Random effects + year FE	1.45	0.78	.064	—	—	.209
GER	District + year FE	1.12	1.16	.342	—	—	.960
Exam score	Pooled OLS + year FE	0.85	0.64	.199	-0.27	.024	.481
Exam score	Random effects + year FE	0.43	0.70	.544	-0.15	.070	.588
Exam score	District + year FE	-0.43	1.55	.783	-0.09	.586	.754



In the pooled enrolment model, a one-standard-deviation increase in SCI-15 is associated with a 2.20-point increase in GER ($p = .011$). This supports H1 when variation across jurisdictions and years is analysed together. The coefficient falls to 1.12 and becomes statistically insignificant in the district and year fixed-effects model.

For examination performance, SCI-15 is not statistically significant in either specification. Examination integrity is negatively associated with scores in pooled OLS, but the association disappears after district and year fixed effects are introduced. H2 is therefore not supported.

Taken together, the results provide limited support for H3. Capacity is related to participation in the pooled comparison, while no robust association is found for examination scores. Yet the disappearance of the enrolment association under fixed effects means that the evidence concerns persistent differences between places rather than short-term within-place changes.

Robustness checks

PCA provides a meaningful but not dominant common factor: the first component explains 36.6% of total variance, and the first five components explain 77.7%. The PCA index correlates at .748 with the dimension-balanced SCI-15 and at .772 with the equal-weight index, indicating substantial agreement without implying that the indices are interchangeable.

The alternative indices reproduce the central result. Both the equal-weight and PCA indices are positively associated with GER in pooled models ($p = .003$ and $p = .009$, respectively), but neither remains statistically significant under district and year fixed effects. Neither alternative index is robustly associated with examination scores. One-year lagged SCI-15 is positively associated with GER only in the pooled model ($\beta = 2.47$, $p = .022$); the fixed-effects lag coefficient is not significant.

Random-effects estimates lie between pooled and fixed-effects results. The GER coefficient is positive but marginal ($\beta = 1.45$, $p = .064$), whereas the examination coefficient is not significant. Hausman tests do not reject random effects for GER ($\chi^2 = 0.84$, $p = .361$) or examination performance ($\chi^2 = 3.29$, $p = .193$). Nevertheless, fixed effects remain the more conservative substantive benchmark because unobserved district characteristics are plausibly correlated with accumulated capacity.

Dimension-specific models show that population quality is positively associated with GER in the pooled specification ($\beta = 5.08$, $p = .019$), while educational resources are only marginally related ($p = .093$). None of the four dimensions is statistically significant in the fixed-effects GER model, although socioeconomic capacity is marginal ($p = .081$). No capacity dimension is robustly associated with examination performance.

The municipality interaction provides suggestive evidence of heterogeneity. In the fixed-effects GER model, the capacity-by-municipality interaction is negative ($\beta = -3.22$, $p = .019$), implying that within-place changes in SCI-15 are less positively associated with GER in municipalities than in districts. Because the interaction is estimated from only nine municipalities and a short panel, it is interpreted as exploratory rather than definitive. No corresponding heterogeneity is detected for examination scores.



Table 3. Selected Robustness, Dimension, and Heterogeneity Estimates

Analysis	Outcome	Term / index	β	SE	p
PCA robustness	GER pooled	PCA index	4.03	1.43	.009
PCA robustness	GER FE	PCA index	2.85	1.75	.115
Equal-weight robustness	GER pooled	Equal-weight index	4.49	1.35	.003
Equal-weight robustness	GER FE	Equal-weight index	0.23	0.96	.810
Lagged model	GER pooled	SCI-15 (t-1)	2.47	1.01	.022
Lagged model	GER FE	SCI-15 (t-1)	-0.81	1.48	.586
Dimension model	GER pooled	Population quality	5.08	2.03	.019
Dimension model	GER FE	Socioeconomic capacity	1.92	1.06	.081
Heterogeneity	GER FE	SCI-15 $\times$ municipality	-3.22	1.29	.019
PCA robustness	Exam FE	PCA index	-3.99	4.42	.375

DISCUSSION

Interpreting the two objectives together

The study pursued two linked objectives: to assess whether structural subnational capacity is associated with junior-secondary participation and to assess whether it is associated with measured learning after examination integrity is considered. Read together, the results reveal an important asymmetry. Capacity is positively associated with gross enrolment in pooled comparisons, but the relationship does not remain statistically distinguishable from zero after district and year fixed effects are introduced. For examination performance, no robust association is found in either specification. The evidence therefore does not support a simple proposition that more capacity automatically produces better educational outcomes.

The contrast between pooled and fixed-effects estimates is central rather than incidental. Pooled estimates capture persistent differences between jurisdictions as well as temporal change, whereas fixed effects isolate change within the same jurisdiction. The results suggest that SCI-15 operates primarily as a marker of accumulated territorial advantage. Districts with stronger educational resources, infrastructure, purchasing power, human development, and health tend to occupy a more favourable structural position, but annual changes in those conditions over a five-year period are not consistently followed by measurable changes in enrolment or examination scores.

This interpretation directly answers the research objectives. Structural capacity matters for understanding why places differ, particularly in participation, but the data do not show that short-run improvements in capacity are sufficient to generate short-run improvements in either outcome.

Capacity and participation: a structural rather than immediate relationship

The pooled association between SCI-15 and gross enrolment is consistent with the supply-and-demand logic of educational participation. On the supply side, budgets, teachers, classrooms, accreditation, and road access determine whether services are available and reachable. On the demand side, poverty severity, purchasing power, adult schooling, and health influence whether families can sustain attendance and whether students are able to remain in school.



This result is also consistent with international evidence that decentralised authorities can improve enrolment when autonomy is matched by the ability to allocate resources and recruit or deploy personnel effectively. Elacqua et al. (2021), for example, found that greater municipal autonomy in Colombia improved enrolment and achievement partly through better teacher quality and more efficient allocation. The important mechanism was not simply a larger volume of resources, but the capacity to use those resources more effectively.

At the same time, the fixed-effects result tempers that conclusion. The absence of a significant within-district association implies that changes in annual capacity are not equivalent to changes in participation. Many components of SCI-15 are slow-moving stocks: teacher distribution, adult education, road quality, accreditation, and health conditions accumulate over time. A five-year panel may therefore capture territorial stratification more effectively than policy-induced transformation.

This finding also cautions against interpreting gross enrolment as a direct behavioural response. GER may exceed 100 because it includes over-age and under-age students and can be affected by cross-district mobility. The outcome is therefore best understood as a measure of system-level participation capacity rather than an individual probability of enrolment.

Why structural capacity does not translate automatically into learning

The absence of a robust association between SCI-15 and examination scores is theoretically plausible. Structural capacity creates enabling conditions, but learning depends on a longer chain of organisational and instructional conversion. Teacher qualifications must translate into effective pedagogy; teacher availability must translate into equitable subject allocation; budgets must reach instructional priorities; and school leaders must convert formal resources into coherent teaching and learning practices.

This distinction echoes the broader literature on educational production. Inputs matter, but their effects are mediated by governance, incentives, allocation, professional practice, and school organisation. OECD (2022) similarly argues that the level of funding is important, but beyond a threshold, the governance, distribution, and effective use of resources are at least as consequential for learning. The World Bank's review of Indonesian subnational education spending reaches a related conclusion: local planning remains heavily input-oriented, while the link between spending decisions and learning outcomes is often weak (Afkar et al., 2020).

The present results therefore point to a conversion gap. SCI-15 captures whether a jurisdiction possesses favourable structural conditions, but not whether the local education administration and schools convert them into instructional quality. This gap may explain why capacity is more visible in participation than in achievement. Increasing access can be achieved through additional places, teachers, transport, and affordability; improving learning requires sustained changes in classroom practice, instructional leadership, teacher support, and accountability.



Evidence from remote Indonesian schools reinforces this mechanism. Bjork and Susanti (2023) show that community-based accountability can alter teacher behaviour, but its success depends on locally appropriate design, training, and relationships. Structural resources alone do not produce instructional improvement unless organisational arrangements activate them.

Capacity accumulation, conversion efficiency, and the role of the middle tier

The results support a sharper conceptual distinction between capacity accumulation and conversion efficiency. Capacity accumulation refers to the long-term development of fiscal resources, infrastructure, qualified personnel, human capital, and socioeconomic conditions. Conversion efficiency refers to the ability of district institutions to coordinate those assets, target them to need, support schools, and generate educational outcomes.

This distinction places district education authorities within the 'middle tier' of the education system. The middle tier connects national policy to schools, supports teacher management, uses data, coordinates professional learning, and translates general policy into locally workable action. UNESCO IIEP (2025) describes this tier as the connective tissue of decentralised education systems and notes that its potential is often constrained by fragmented mandates, insufficient resources, weak data, and limited management capacity.

SCI-15 captures several enabling conditions surrounding the middle tier, but it does not directly measure planning quality, data use, instructional supervision, organisational routines, or coordination with schools. The null fixed-effects results may therefore indicate that these conversion mechanisms—not merely structural stocks—are the missing explanatory link. Future models should pair the structural index with direct measures of district education management and school-support practices.

This reframing also resolves an apparent paradox: a district can possess relatively strong structural capacity and still record modest learning outcomes if conversion efficiency is weak. Conversely, a district with moderate structural resources may perform better if it allocates teachers effectively, prioritises high-impact programmes, and provides coherent instructional support.

Positioning the findings within conflicting international evidence

International evidence on decentralisation is mixed, and the present study helps explain why. Elacqua et al. (2021) found positive effects of municipal education autonomy in Colombia under a design that strengthened local managerial authority and improved teacher allocation. By contrast, Kameshwara et al. (2020), using data from 65 education systems, found no general achievement advantage from decentralisation and warned against treating it as a universally effective reform. These apparently conflicting findings are compatible with a capacity-conditional interpretation. Decentralisation can improve outcomes when local authorities possess the resources, autonomy, information, and organisational competence needed to act. It may have weak or unequal effects when authority is transferred without commensurate capacity, when accountability is fragmented, or when local institutions reproduce existing territorial inequalities.



The West Java evidence aligns more closely with this conditional view than with either an unqualified endorsement or rejection of decentralisation. Capacity differentiates jurisdictions, but it does not by itself guarantee improvement. Decentralisation creates an opportunity structure; its consequences depend on both accumulated capacity and the institutional mechanisms that convert capacity into service quality.

Recent evidence on Indonesian public expenditure efficiency also underlines the territorial dimension. Sugiyarto et al. (2025) document persistent differences in expenditure efficiency between lagging and more developed districts and show that infrastructure, urbanisation, literacy, local revenue, and grant dependence shape performance. Their findings support the argument that local outcomes reflect not only spending levels but also structural and institutional context.

Examination integrity and the interpretation of measured quality

The negative pooled association between examination integrity and examination scores must not be interpreted as evidence that integrity harms learning. A more plausible explanation is that stronger monitoring reduces score inflation. In systems where test administration varies in credibility, observed achievement combines student learning with the integrity of measurement. The loss of statistical significance after district and year fixed effects suggests that much of the pooled relationship reflects persistent local differences or system-wide changes in examination governance. Even so, including integrity remains conceptually necessary because it prevents the model from treating all score differences as equivalent indicators of learning.

This finding contributes to a broader debate about policy indicators. High-stakes measures can become organisational targets and can be affected by the conditions under which they are produced. Comparing districts solely on examination scores may therefore reward weak assessment governance or penalise jurisdictions that improve integrity. Educational accountability should distinguish learning performance from the credibility of the measurement process.

The historical design of this study is therefore a strength as well as a limitation. The 2015-2019 period permits the simultaneous observation of a nationally standardised examination and an integrity measure, offering an unusual opportunity to examine how governance conditions shape the interpretation of educational quality.

Implications for territorial equity and intergovernmental policy

The results indicate that territorial inequality in education is rooted partly in unequal development capacity. Jurisdictions begin with different stocks of teachers, classrooms, infrastructure, purchasing power, adult education, and health. Applying identical policy instruments to unequal territories can therefore reproduce rather than reduce disparities.

OECD (2021) emphasises that decentralised public-service systems require coordination, equalisation, and attention to differences in fiscal and institutional capacity. Equalisation transfers and common standards remain important, but they should be paired with differentiated technical assistance and needs-sensitive allocation. Afkar et al. (2020) similarly recommend identifying low-capacity Indonesian districts and tailoring capacity-building support to planning and execution weaknesses.



The policy implication is not that central government should reverse decentralisation. Rather, central and provincial authorities should differentiate support according to the type of constraint. Resource-poor districts may need fiscal equalisation; geographically constrained districts may need transport and network solutions; districts with weak conversion efficiency may need planning, data, teacher-management, and instructional-support capacity.

This approach shifts the policy question from 'How much authority should be decentralised?' to 'What combinations of authority, capacity, support, and accountability enable decentralised systems to produce equitable outcomes?'

What the study contributes theoretically

The study makes four theoretical contributions. First, it conceptualises subnational capacity as a multidimensional territorial condition rather than a narrow measure of budget size or bureaucratic staffing. Second, it separates capacity inputs from educational outcomes and assessment integrity, avoiding circularity in measurement. Third, it distinguishes between-place inequality from within-place change, showing that the meaning of capacity depends on the temporal and statistical comparison being made.

Fourth, it develops the distinction between capacity accumulation and capacity conversion. This distinction explains why resources may be associated with access but not with achievement, why pooled and fixed-effects estimates diverge, and why decentralisation produces heterogeneous results across settings.

The resulting framework suggests that future research should not ask only whether local capacity matters. It should ask which capacity dimensions matter for which outcomes, over what time horizon, and through which organisational mechanisms.

Methodological interpretation and robustness

The advanced specifications make the central conclusion more credible because it survives three different index constructions and alternative estimators. The first PCA component explains 36.6% of indicator variance, while its correlation with the theory-driven SCI-15 is .748. This level of agreement supports convergent validity but also confirms that subnational capacity is not reducible to a single empirical factor.

The Hausman tests do not statistically reject random effects, yet this result should not be read as proof that capacity is uncorrelated with district-specific unobservables. Historical urbanisation, administrative traditions, political economy, and geography are plausible common causes of both capacity and outcomes. Reporting pooled, random-effects, and fixed-effects estimates therefore makes the identifying assumptions transparent rather than selecting a model solely on the basis of statistical significance.

The robustness checks strengthen the main interpretation but do not eliminate uncertainty. The equal-weight index yields the same substantive conclusion, indicating that the findings are not an artefact of the dimension-balanced weighting scheme. The lagged-capacity models also remain insignificant under fixed effects, suggesting that a one-year delay is insufficient to reveal short-run conversion effects.



The high R-squared values in fixed-effects models require careful interpretation. They largely reflect district and year indicators rather than the explanatory contribution of SCI-15 alone. The relevant evidence lies in the capacity coefficient, its uncertainty, and its stability across specifications.

Multicollinearity among structural indicators is another concern because budget, teacher resources, accreditation, adult education, health, and purchasing power are developmentally interrelated. The index approach reduces the number of predictors, but it also compresses heterogeneous mechanisms. Future work should compare equal-weight, dimension-weighted, and principal-component indices and report the sensitivity of results to the exclusion of individual dimensions. Finally, reverse causality remains possible. Educational expansion can itself improve adult education, local productivity, and fiscal capacity. The models should therefore be read as estimates of association, not causal effects.

Boundaries of generalisation

The study is based on one province and a five-year historical window. West Java contains substantial urban-rural and socioeconomic variation, but it does not represent all Indonesian provinces. The findings are most directly generalisable to decentralised systems characterised by strong subnational inequality and slowly changing structural conditions.

The study also operates at the district level. Aggregate relationships may conceal inequalities among schools, neighbourhoods, and student groups. A district can have high average capacity while maintaining severe internal disparities. Multilevel research linking district capacity, school organisation, and student outcomes is therefore required.

Despite these boundaries, the analysis offers a transferable proposition: decentralised authority should not be assumed to generate equal outcomes when the capacity to exercise that authority is territorially unequal.

Policy Implications

First, intergovernmental education policy should be capacity-sensitive and asymmetric. Jurisdictions with weak educational resources require support for teacher deployment, classroom quality, school accreditation, and planning capacity; geographically constrained districts require transport and infrastructure solutions; and areas where poverty depresses participation require coordination with social-protection policy. Equal national standards should therefore be paired with differentiated instruments and technical assistance.

Second, access and learning should not be treated as the same policy problem. Expanding participation requires sufficient places and lower barriers; improving learning requires stronger conversion mechanisms inside schools and local education administrations.

Third, performance accountability should consider starting capacity and measurement credibility. Common outcome targets should be combined with transparent capacity diagnostics and credible assessment systems.



Fourth, structural capacity should be monitored over longer horizons through an integrated district dashboard. Annual programme evaluation is poorly suited to slowly changing conditions such as population education, teacher distribution, health, and infrastructure. Capacity indicators should be reported alongside outcomes so that accountability distinguishes low performance caused by weak conversion from low performance rooted in structural disadvantage.

Limitations and Future Research

The panel covers only five years and one province, limiting within-district variation and external generalisability. The capacity index also depends on indicator definitions, standardisation, and weighting choices, although the equal-weight robustness check produces similar conclusions.

Future research should extend the time series, test the index in other provinces, and combine district-level capacity measures with school-level organisational data. Quasi-experimental strategies, lag structures, and mediation analysis could help distinguish accumulation, conversion, and outcome pathways.

CONCLUSIONS

This study examined whether structural subnational capacity is associated with junior-secondary participation and measured learning outcomes across 27 districts and municipalities in West Java between 2015 and 2019. Its central conclusion is qualified but clear: capacity helps explain persistent differences between places more convincingly than it explains short-term change within the same place.

For the first objective, higher SCI-15 scores are associated with higher gross enrolment in pooled estimates, but the relationship becomes statistically insignificant when district and year fixed effects are introduced. Capacity therefore appears to be part of the structural foundation of participation rather than an annual policy lever with immediate effects. For the second objective, no robust association is found between SCI-15 and National Examination scores after examination integrity is considered. Structural resources are enabling conditions for learning, but they require organisational and instructional conversion.

The contrast between the two outcomes is the study's most important substantive result. Access has relatively direct pathways from resources, infrastructure, and household affordability. Learning depends on a longer chain involving teacher allocation, pedagogy, school leadership, accountability, and credible assessment. The study therefore rejects a mechanical resources-to-results model and proposes a distinction between capacity accumulation and conversion efficiency.

The findings also refine the interpretation of decentralisation. Transferring authority to local governments does not equalise their ability to act. Decentralised systems require differentiated intergovernmental support, fiscal and technical equalisation, and stronger middle-tier institutions capable of translating resources into school improvement. Uniform mandates applied to unequal territories are unlikely to produce equal outcomes.

Methodologically, the analysis shows why pooled and fixed-effects estimates should not be treated as interchangeable and why examination integrity must be separated from measured achievement.



The results are associational, not causal, and should be tested with longer panels, alternative index construction, direct measures of district-management capacity, and multilevel school and student data.

More broadly, the study contributes a proposition relevant beyond Indonesia: subnational capacity matters, but what ultimately determines educational progress is not only the stock of capacity a territory possesses, but the institutional efficiency with which that capacity is converted into access, instruction, and credible learning outcomes.

References:

- [1] Asrijanty. (2018). Indeks integritas Ujian Nasional dan capaian Ujian Nasional. Pusat Penilaian Pendidikan, Kementerian Pendidikan dan Kebudayaan. <https://repositori.kemdikbud.go.id/>
- [2] Besley, T., & Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4), 1218-1244. <https://doi.org/10.1257/aer.99.4.1218>
- [3] Besley, T., & Persson, T. (2011). Pillars of prosperity: The political economics of development clusters. Princeton University Press. <https://press.princeton.edu/books/hardcover/9780691152684/pillars-of-prosperity>
- [4] Cingolani, L. (2013). The state of state capacity: A review of concepts, evidence and measures (UNU-MERIT Working Paper 2013-053). <https://ideas.repec.org/p/unm/unumer/2013053.html>
- [5] Fukuyama, F. (2013). What is governance? *Governance*, 26(3), 347-368. <https://doi.org/10.1111/gove.12035>
- [6] Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, 14(3), 351-388. <https://hanushek.stanford.edu/publications/conceptual-and-empirical-issues-estimation-educational-production-functions>
- [7] Hanushek, E. A., & Woessmann, L. (2015). The knowledge capital of nations: Education and the economics of growth. MIT Press. <https://mitpress.mit.edu/9780262029179/the-knowledge-capital-of-nations/>
- [8] Oates, W. E. (1972). Fiscal federalism. Harcourt Brace Jovanovich. <https://ideas.repec.org/a/eee/pubeco/v2y1973i3p322-325.html>
- [9] OECD. (2020). Policy framework on sound public governance: Baseline features of governments that work well. OECD Publishing. <https://doi.org/10.1787/c03e01b3-en>
- [10] Sen, A. (1999). Development as freedom. Oxford University Press. <https://global.oup.com/academic/product/development-as-freedom-9780192893307>
- [11] World Bank. (2013). Local governance and education performance: A survey of the quality of local education governance in 50 Indonesian districts. World Bank. <https://documents.worldbank.org/>
- [12] World Bank. (2018). World development report 2018: Learning to realize education's promise. World Bank. <https://doi.org/10.1596/978-1-4648-1096-1>



- [13] Elacqua, G., Munevar, I., Sánchez, F., & Santos, H. (2021). The impact of decentralized decision-making on student outcomes and teacher quality: Evidence from Colombia. *World Development*, 141, 105378. <https://doi.org/10.1016/j.worlddev.2020.105378>
- [14] Foa, R. S. (2022). Decentralization, historical state capacity and public goods provision in Post-Soviet Russia. *World Development*, 152, 105807. <https://doi.org/10.1016/j.worlddev.2021.105807>
- [15] Singh, R., Bhattacharjee, S., & Nandy, A. (2024). Fiscal decentralization for the delivery of health and education in Indian states: An ongoing process is more desirable than a policy shift. *Journal of Policy Modeling*, 46(2), 254-271. <https://doi.org/10.1016/j.jpolmod.2024.01.006>
- [16] Tournier, B., Godwin, K., Cameron, E., & Lugaz, C. (2025). Leveraging the potential of the middle tier to improve education outcomes: The role of a capacity assessment framework. UNESCO International Institute for Educational Planning. <https://www.iiep.unesco.org/en/publication/leveraging-potential-middle-tier-improve-education-outcomes-role-capacity-assessment-framework>
- [17] World Bank. (2020a). The promise of education in Indonesia. World Bank. <https://hdl.handle.net/10986/34807>
- [18] Afkar, R., Luque, J., Nomura, S., & Marshall, J. H. (2020). Revealing how Indonesia's subnational governments spend their money on education: Indonesia subnational education public expenditure review 2020. World Bank. <https://doi.org/10.1596/34831>
- [19] Bjork, C., & Susanti, D. (2023). Can community participation leverage changes in teacher behavior? Evidence from remote areas of Indonesia. *International Journal of Educational Development*, 102, 102840. <https://doi.org/10.1016/j.ijedudev.2023.102840>
- [20] Kameshwara, K. K., Sandoval-Hernandez, A., Shields, R., & Dhanda, K. R. (2020). A false promise? Decentralization in education systems across the globe. *International Journal of Educational Research*, 104, 101669. <https://doi.org/10.1016/j.ijer.2020.101669>
- [21] OECD. (2021). Delivering quality education and health care to all: Preparing regions for demographic change. OECD Publishing. <https://doi.org/10.1787/83025c02-en>
- [22] OECD. (2022). Value for money in school education: Smart investments, quality outcomes, equal opportunities. OECD Publishing. <https://doi.org/10.1787/f6de8710-en>
- [23] Sugiyarto, A., Lee, B. L., & Wilson, C. (2025). Narrowing the gap: Dynamics of Indonesia's public expenditure performance in economically lagging districts. *Economic Analysis and Policy*. <https://doi.org/10.1016/j.eap.2025.03.014>