

When Do Colonial Institutions Persist? Evidence from Two School Systems in South Africa*

Sarah Ferber[†]

Gabriel Brown[‡]

Johan Fourie[§]

Abstract

Research on colonial persistence rarely observes historical institutional quality. Its central assumption, that institutional quality persists, is therefore rarely tested. We test this assumption in South Africa’s Cape Colony. The 1896 and 1905 Education Reports record the attainment of every state-aided school on a common scale, covering public and poor schools that served predominantly white pupils and mission schools that served predominantly black pupils. We digitise these reports and the 1881 data on subjects and grants, interpolate recorded quality to modern school locations, and link it to schools’ 2018 mathematics results. Recorded quality predicts mathematics performance in both systems more than a century later, and it does so at different spatial scales. Our model attributes the pattern for public and poor schools to place and the pattern for mission schools to organisation. A formal test rejects equality of the two trajectories. A funding rule capped mission instruction near Standard IV. Applying this ceiling to the records of public and poor schools reproduces the mission pattern within one system. We provide a direct historical measure of institutional quality and a direct test of the assumption on which the persistence literature rests. We also show how proximity designs can confound mission density with operator type. In the model, persistence is a policy outcome: a colonial difference endures only while its transmission mechanism is allowed to continue.

Keywords: colonial education; mission schools; institutional quality; school quality; persistence; South Africa

JEL codes: I25; N37; O15; H75

*We thank Rulof Burger, Michiel de Haas, Ewout Frankema, Louis Henderson, Felix Meier zu Selhausen, Servaas van der Berg, Dieter von Fintel and participants at the World Economic History Congress for helpful comments. We are indebted to Servaas van der Berg for generously sharing data. All errors are our own. Replication data and code will be deposited in a public repository. Disclosure of AI use: OpenAI Codex (GPT-5) and Anthropic Claude Code were used to audit the R and L^AT_EX pipeline, assist with bounded corrections to code and the implementation of robustness checks, and support editorial quality control. The authors reviewed the code, estimates and text and remain responsible for all content. Cite this paper as: Ferber, Sarah, Gabriel Brown, and Johan Fourie. 2026. “When Do Colonial Institutions Persist? Evidence from Two School Systems in South Africa.” Working Paper, Department of Economics, Stellenbosch University.

[†]Department of Economic History, Lund University. Email: sarah.ferber@ekh.lu.se.

[‡]Department of Economic History, Lund University; and Department of Economics, Stellenbosch University. Email: gabriel.j.m.brown@gmail.com.

[§]Department of Economics, Stellenbosch University. Email: johanf@sun.ac.za.

1 Introduction

A large literature finds that the colonial past predicts present outcomes. It generally interprets this relation as evidence of institutional persistence, although that interpretation rests largely on an unobserved assumption. The canonical account links a historical shock to institutions in the past, institutions in the past to institutions today, and institutions today to present outcomes (Acemoglu, Johnson and Robinson, 2001). The middle link requires the *quality* of an institution to persist. It is rarely observed because historical institutional quality is rarely measured. Reduced-form regressions of modern outcomes on distant causes cannot distinguish this channel from geography, culture or path dependence. This paper measures historical institutional quality directly for two school systems assessed on a common administrative scale. We ask whether that quality persists and through which channel. We distinguish the school system from the individual school and from the administrative measure we observe. The *institution* is the school system, including its funding rule, inspectorate and certification standard. The *organisation* is the individual school that carries the system’s quality. *Recorded attainment* is the administrative measure of that quality, which the funding rule censored in one system.

A comparison from Wellington in the Cape Colony illustrates the question. In 1905, the Boys’ High School presented all 238 pupils to the inspector, of whom 182 passed. The mission school presented 125 of its 313 pupils, of whom 95 passed. Both schools therefore passed about 76 per cent of the pupils they presented. The remaining records differed sharply. The high school had 65 pupils in the Ex-Standard, above the highest ordinary Standard. The mission school had no pupils above Standard IV, while 191 of its 313 pupils were in the Sub-Standards. The same inspectorate used the same scale in the same town to record schools that operated under different systems.

The Cape records allow us to measure what became of this difference. The Cape Education Reports list every state-aided school in the colony. For 1896 and 1905, they record the Standards in which its pupils sat. For 1881, they record the subjects it taught and the grant it received. We digitise and geolocate the schools and link the historical records to school-level mathematics results from 2018. We can therefore test whether the recorded quality of an individual colonial school predicts performance in nearby schools today. Our measure is the pupil-weighted mean Standard, the average Standard in which a school’s pupils sat. It rises when a school advances and retains pupils rather than losing them in the early years.

Both systems reported to the same inspectorate on the same scale. We therefore use the same outcome, controls and inference to compare systems that colonial policy funded differently. The design asks at what spatial scale a colonial school’s recorded quality remains associated with present outcomes, and whether this scale differed between schools funded directly by the state and those it subsidised. Historical quality was not randomly assigned, and we do not estimate the causal effect of colonial schooling. The common scale instead permits a comparison that reduced-form persistence regressions cannot make. We compare two historical quality surfaces under identical controls while varying the degree of smoothing, applying the mission ceiling to the records of public and poor schools and examining linked successor schools. The model specifies in advance how each

transmission channel must respond to these exercises, so each is a test the data could fail.

The model uses one transmission technology under two funding regimes. It distinguishes transmission through the organisation from transmission through place. Schools hold organisational capital in teaching practice, recruitment and reputation. This capital persists while the organisation continues and is lost when it closes. Places hold demand for schooling, which can persist after an individual school closes. The funding regime enters through a certification ceiling. The mission grant financed instruction to about Standard IV, so the reports provide a censored measure of what mission schools could do. These channels imply different responses to spatial smoothing. If quality is transmitted through the organisation, averaging the historical measure over a wider area should weaken the association because the neighbours of a good school need not have been good. If quality is transmitted through place, the same averaging should strengthen the association. This prediction is falsifiable. Section 6 tests it.

Recorded quality predicts modern mathematics in both systems, but at different spatial scales. The model associates these patterns with different transmission channels. Measured per recorded Standard, the mission coefficients are larger at every grade. Measured per standard deviation of the treatment, the ordering reverses because the mission quality distribution is less than half as wide, and the mission estimates are less precise. A ceiling on recorded attainment predicts each element of this pattern, because censoring makes each recorded unit correspond to more underlying capital while reducing the variation available for estimation. The relevant evidence is therefore the joint pattern across the two systems, and it appears as the model predicts.

The spatial scale of the association provides the more informative result. As historical quality is averaged over a wider area, the mission association falls to zero and changes sign, most clearly for Grade 3 mathematics. The public and poor association roughly trebles. This contrast does not rely on comparing overlapping confidence bands. On the sample exposed to both systems, with both surfaces entered jointly, a formal endpoint test rejects equality of the changes in the two trajectories at Grade 3 ($p = 0.002$) and is marginal at Grades 6 and 9.

In the model, the mission trajectory requires organisational transmission combined with a negative smooth loading: mission schools transmitted quality through the school but were located in places whose other characteristics predict worse outcomes today. The public and poor trajectory requires a large positive smooth loading. This loading includes transmission through place and selection into favourable places. It also includes any spatially smooth component of organisational transmission, which a spatial design cannot separate from place.

A name-and-location crosswalk provides independent evidence of an organisational component. Among historical schools of either system matched to a plausible modern successor, a school's own historical quality predicts successor performance. A neighbourhood surface recomputed after removing the school does not predict performance when both measures enter jointly. A pure place channel would make both coefficients positive. The linked sample is selected on survival and pooled across the two systems, so its role is to corroborate the organisational channel, which it does. The cross-system comparison rests on the smoothing trajectories.

Two additional exercises examine the ceiling and date the cross-system difference. We first censor the public and poor schools' records at Standard IV, where mission schools concentrate. Within the public and poor system, this raises the coefficients and lowers the associations per treatment standard deviation. It also increases the standard errors. The coefficient increase is close to the value predicted by censoring alone, and a formal test does not reject that calibration.

The difference between the systems predates 1905. In 1881, public schools taught roughly twice as many subjects beyond the elementary curriculum and received three times the grant per pupil. These gaps are larger among schools that can be linked to a 1905 record.

We next examine whether the Cape pattern is visible elsewhere in South Africa. We use mission stations from the Protestant atlas of Beach (1903) and the Catholic atlas of Streit (1929), geolocated and linked across later editions (Dennis, Beach and Fahs, 1911; Beach and St. John, 1916; Beach and Fahs, 1925; Brown, 2025). Within provinces and within education districts, schools near a historical mission have higher school-leaving pass rates today. The national records also permit a test of whether operator quality, rather than mission presence, matters. We proxy operator quality by the educational orientation of the society that ran each mission (Brown and Martin, 2026). The exercise yields a methodological result for the missions literature. The usual coding classifies a school according to whether any mission of a given type lies within a radius. It therefore assigns the high-orientation label more often in places with dense mission coverage: schools in that category have more than three times as many missions nearby. Randomly permuting orientation labels across societies reproduces almost the same imbalance. The contrast is manufactured by the coding rule, not by where those societies operated, and assigning each school to its nearest mission removes it. Proximity designs can therefore attribute mission density to operator type. Claims about mission quality require school-level records rather than station atlases, and we accordingly draw no national conclusion about operator orientation.

The atlases record where mission stations stood, not what individual schools taught. We therefore confine the national analysis to comparisons the atlas data can support, and we retain foreign-staff presence in 1922 as an appendix diagnostic; its interaction with society educational orientation is statistically indistinguishable from zero.

The paper contributes first to the literature on institutional persistence. This literature shows that historical institutions predict present outcomes across many settings (Tabellini, 2010; Spolaore and Wacziarg, 2013; Guiso, Sapienza and Zingales, 2016; Nunn, 2020). Critics argue that these coefficients compress long periods, combine several channels and weaken under appropriate spatial inference (Austin, 2008; Abad and Maurer, 2021; Conley and Kelly, 2025; Fenske, Gupta and Mukhopadhyay, 2025). Other work concludes that persistence remains important even when particular estimates overstate it (Voth, 2020). These positions share an untested premise because they do not observe historical institutional quality. Moreover, Alesina et al. (2020) show that persistence can survive the direct destruction of institutions, which implies that other channels can carry it.

We observe historical institutional quality for two systems under one administration and on one

scale. The two systems persist at different spatial scales, in the pattern that the model assigns to different channels. Policy also set the margins along which those channels operated. A funding rule determined what one system could record, while later education acts specified the terms under which organisations continued. Both margins were objects of policy choice rather than fixed features of the past. Persistence, on this reading, is not evidence that the past constrains the present. It is evidence that a chosen transmission mechanism was allowed to operate.

The model also makes the spatial critique empirically testable. Equation (1) identifies the components combined in a spatial persistence coefficient. It turns dispersion accounting and kernel sensitivity into tests with predicted signs. The same approach can be applied elsewhere by reporting historical dispersion and the magnitude per standard deviation, adding a censoring diagnostic when an administrative ceiling constrained the historical measure, and showing the smoothing trajectory alongside the headline coefficient.

The second contribution is to the literature on missions and education. Most studies observe mission locations and, at best, the denomination or staffing of the operator. Gallego and Woodberry (2010) compare Protestant and Catholic missionary activity and competition between them. Becker and Meier zu Selhausen (2023) study the denomination and gender composition of mission staff, while Valencia Caicedo (2019) shows that operator identity mattered. Other studies use station locations alone (Wantchekon, Klačnja and Novta, 2015; Jedwab, Meier Zu Selhausen and Moradi, 2022; Okoye and Pongou, 2024; Brown, 2025). None observes the recorded quality of the individual school. We observe that quality for mission and state schools on a common inspectorate scale. Among linked successor schools, the school’s own recorded quality, rather than the quality of its surroundings, predicts present outcomes. The national analysis also shows how proximity coding can attribute mission density to operator type.

The third contribution is to the literature on education in developing countries. This literature often studies current differences in school quality. We show that some of these differences are inherited: within each historical system, quality recorded in 1905 predicts the performance of nearby schools in 2018, more than a century later. In South Africa, the administrative difference can also be dated. Existing work commonly traces educational inequality to Bantu education and its successors (Fiske and Ladd, 2004; Van der Berg, 2007; Spaull, 2013, 2019). Fedderke and co-authors interpret the period from 1910 as a failure to address an inherited inequality (Fedderke, de Kadt and Luiz, 2000). The records show that the legacy predates the Union. By 1881, the two systems differed by a factor of three in grants per pupil and by half in the number of subjects taught. By 1905, the average mission pupil sat more than one Standard below the average public pupil.

These records date administrative differences in funding and instruction rather than identifying their causal effect on attainment. The documented gap in 1881, together with the fact that recorded 1905 quality still predicts mathematics results in 2018 within each system, qualifies accounts in which colonial schooling generated broad mobility (Alesina et al., 2021) and locates the origins of an inequality that policy since 1994 has not eliminated (Von Fintel and Fourie, 2019; Carrillo,

Charris and Iglesias, 2023).

Section 2 describes the historical setting. Section 3 presents the model, and Sections 4 and 5 describe the data and empirical method. Section 6 reports the Cape estimates and tests the model’s predictions. Section 7 presents the national evidence. Section 8 concludes.

2 Historical setting

2.1 Two school systems in the Cape

The Cape Education Department was established in 1839 under a Superintendent General of Education (Malherbe, 1925). By the late nineteenth century, it supervised public, poor, farm and subsidised mission schools. Public schools were graded according to the size and level of their curricula. Poor and farm schools provided elementary education to less affluent and rural white communities. Mission societies provided most of the formal education available to black pupils (Cape Colony, 1906; Loram, 1917).¹

The two systems differed in administration and funding. Public schools received larger state grants and were increasingly consolidated into a publicly supported system. Mission schools relied on grants-in-aid, missionary resources and local contributions. Colony-wide accounts for 1881 imply state grants of approximately £1.74 per enrolled pupil in public schools and £0.42 in mission schools. These aggregates establish a difference in public support, but do not identify the effect of school-level spending.

This division was already present a generation before the main records used in this paper. The 1881 report lists each school’s subjects and grant rather than the Standards reached by its pupils. It therefore measures what a school offered, not attainment on the 1905 scale. The average public school taught roughly twice as many subjects beyond the elementary curriculum as the average mission school. It received three times the grant per pupil while teaching half as many pupils. Restricting the sample to schools that can be linked to a 1905 record widens each gap (Appendix Table B.5). The school-level grants reproduce the colony-wide difference, and the schools that survived to 1905 do not form a sample in which the systems had converged. Whatever the 1905 records measure, they record a difference that was already a generation old.

Mission schools did not follow the public-school curriculum without modification. Loram (1917, p. 94) describes their curriculum as a modified and abbreviated version of the public curriculum. In practice, the mission grant supported instruction to about Standard IV. Section 4 shows that mission schools concentrate at this level and have almost no enrolment above it, whereas public schools extend to Standard VII and the Ex-Standard. Section 3 represents this cap as a certification ceiling imposed by the funding regime. Both systems nonetheless reported pupils on the Cape

¹Under apartheid, the government classified the population into four racial groups: white, black/African, coloured and Indian. These categories were arbitrary, but they remain culturally significant and are still used by Statistics South Africa. We use “white” for the population of European descent and “black” for all other groups. Where relevant, “African” refers specifically to the Bantu-speaking population (e.g. Xhosa), and “coloured” refers to descendants of the Cape Malay, the indigenous Khoi-San and other people of mixed ancestry.

Standards, providing a common attainment scale. Inspectors visited each school annually and recorded enrolment, presentation for inspection and passes. They also recorded movement since the previous inspection and the number of pupils in each Standard.

The systems also followed different institutional paths during the twentieth century. Compulsory white education and public investment strengthened the public system. Mission schools continued under discriminatory funding until the state assumed their administration under the Bantu Education Act of 1953 for African education and the Coloured Persons Education Act of 1963 for schools serving pupils classified as coloured. The latter statute is more relevant to the present Western Cape (Horrell, 1964; Hyslop, 1999; Molteno, 1987). This history motivates the analysis of institutional continuity. It does not establish that later policy caused the estimated coefficient differences.

2.2 Mission education beyond the Cape

Outside the Cape, missionary societies and colonial education departments performed similar roles, but comparable school-level records are unavailable. Each of the four colonies that formed the Union of South Africa in 1910 ran its own education department. In each colony, missionary societies were the main providers of formal education for black pupils. The national mission data therefore come from missionary atlases, which record residence stations rather than the full network of outstations and schools. The Protestant series begins with Beach (1903) and continues in linked editions from 1911, 1916 and 1925 (Dennis, Beach and Fahs, 1911; Beach and St. John, 1916; Beach and Fahs, 1925). Catholic stations come from the 1913 and 1929 editions of *Atlas Hierarchicus* (Streit, 1913, 1929). Brown geolocates the stations and links records across editions, providing substantially more temporal coverage than a single atlas (Brown, 2025).

The atlases measure mission presence, operating society and foreign staffing more reliably than the educational quality of an individual station. We use the society-level educational-orientation score from Brown and Martin (2026) as the closest available national analogue to Cape school quality. The score is the first principal component of a society’s mean investment in five educational categories recorded worldwide in Beach (1903): primary schools, secondary schools, colleges, kindergartens and industrial schools. It is standardised across the worldwide distribution of missionary societies.

3 A model of school-quality transmission

This section models the creation, measurement and transmission of school quality over a century. The model uses one transmission technology under two funding regimes. It defines the regression coefficient estimated in the paper and derives predictions that the data could have rejected. Section 6 tests these predictions.

3.1 Environment

The two school systems, $s \in \{P, M\}$, are public and poor schools and mission schools. A historical school opens with *organisational capital* k_i (Assumption 1). Measured on the 1905 attainment scale, this is the mean Standard the school would record without a funding constraint. It summarises teaching practice, recruitment, norms and reputation, which are attributes of the organisation rather than the site.

The funding regime certifies attainment only up to a ceiling κ_s . Recorded quality is therefore $q_i = k_i - r_i$, where $r_i \geq 0$ (Assumption 2). The public ceiling lies above all observed attainment, so $q_i = k_i$. The mission ceiling, at about Standard IV, limits what mission schools could record. Mission quality is therefore a censored measure of what those schools could do. Censoring is one interpretation of the ceiling. The funding rule may instead have prevented the accumulation of capital itself, and the ceiling bound approximately rather than literally. Remark 1 distinguishes these interpretations and their implications for the empirical results.

Between 1905 and the present, a school survives as an organisation with probability $\bar{c}_s$. A contemporary outcome loads on the surviving capital of the nearest historical school through τ , the organisational channel. It also loads through π on *place*, defined as the surrounding area’s community demand and parental human capital, which can persist after an individual organisation closes (Assumption 3). Assigning surviving capital to the nearest historical school operationalises an exposure assumption. The surviving organisation affects contemporary schools in its catchment through its own continuation, its staff and the pupils it trained; proximity proxies for membership of that catchment. If this proxy fails, a coefficient that declines under smoothing distinguishes highly localised persistence from broader persistence. It does not distinguish organisational continuity from fine-scale site effects. The linked sample in Section 6.6 provides successor evidence that does not rely on this proxy.

Appendix A.2 derives $\tau > 0$ from an adjustment-cost problem. A continuing school can reproduce inherited practice at low cost, whereas rebuilding it is costly. Closure destroys organisational capital and leaves only the place component. Historical placement was not random. Compensatory placement, in which schools concentrate in otherwise disadvantaged places, is represented by $\zeta_s \leq 0$. The econometrician observes an inverse-distance surface $\hat{Q}_j(p)$ rather than q directly (Assumption 4). A high power p concentrates weight on the nearest historical school; a low power averages quality over a wider area. Appendix A states the environment, assumptions and disturbance structure formally. All variances and covariances below are within historical divisions and residualised on the control set.

3.2 What the persistence coefficient measures

The estimating equation regresses Y_j on $\hat{Q}_j(p)$ with the full control set. Let β denote the resulting coefficient. Its probability limit follows from Assumptions 2–4. Decompose a school’s measured quality into the local average and its deviation from that average, $q_{a(j)} = \bar{q}_{\ell(j)} + \xi_{a(j)}$. The observed

surface can then be written as $\widehat{Q}_j(p) = \bar{q}_{\ell(j)} + \omega(p) \xi_{a(j)} + m_j$. We define $\bar{q}_{\ell}$ as the linear projection of own quality on the location smooth. Thus $\text{Cov}(\xi, \bar{q}) = 0$ holds by construction.

The empirical inverse-distance surface only approximates this projection. Because the nearest school also receives the largest interpolation weight, the realised surface leaves a positive covariance between ξ and the remainder. This covariance produces a cross term in equation (1), whose direction depends on the loadings. Appendix A.1 gives the corrected term. Equation (1) and Proposition 2 use the orthogonal approximation. Under that approximation (Appendix A.4),

$$b_s(p) \equiv \text{plim } \hat{\beta} = \frac{\omega(p) A_s \sigma_{\xi,s}^2 + S_s \sigma_{\bar{q},s}^2}{\omega(p)^2 \sigma_{\xi,s}^2 + \sigma_{\bar{q},s}^2 + \sigma_m^2}, \quad (1)$$

with loadings

$$A_s = \tau \bar{c}_s \Lambda_s^{\circ}, \quad S_s = \pi + \zeta_s + \tau \bar{c}_s \bar{\lambda}_s, \quad (2)$$

where $\Lambda_s^{\circ} = \text{Cov}(k, \xi) / \text{Var}(\xi)$ is the censoring loading from Lemma 1 applied to own-school deviations. The corresponding loading on the local average is $\bar{\lambda}_s$. We use Λ_s for the ratio in levels, $\text{Cov}(k, q) / \text{Var}(q)$. The estimand depends on Λ_s° , the deviation-level loading, rather than the level ratio. Lemma 1 implies $\Lambda_s^{\circ} \geq 1$ only if its ordering condition continues to hold after the same residualisation. The *organisational* loading A_s is positive only when capital is transmitted through surviving organisations. The *smooth* loading S_s combines place transmission and placement selection. It also includes the smooth component of organisational transmission, which a spatial design cannot separate from place.

Two objects determine whether A_s is positive. The *signal* ξ is the quality difference recorded at a school after removing the component predicted by its surroundings. The *carrier* $\bar{c}_s$ is the share of schools whose organisations survived to the present. Organisational persistence requires both. A recorded difference cannot be transmitted through this channel if the organisation did not survive. If organisations survived but did not differ from their neighbours, there is no recorded difference to transmit.

The carrier enters the loading directly through $A_s = \tau \bar{c}_s \Lambda_s^{\circ}$. The signal enters through its dispersion in the organisational term $\omega(p) A_s \sigma_{\xi,s}^2$ of equation (1). If no organisation survived ($\bar{c}_s = 0$), the loading is zero. If schools did not differ from their neighbours ($\sigma_{\xi,s}^2 = 0$), the organisational term vanishes. In that case, Λ_s° is undefined rather than zero because it divides by $\text{Var}(\xi)$. Both cases remove the organisational contribution to the estimand. Assumption 2 acts on the signal by compressing what a mission school could record. Twentieth-century policy acted on the carrier by determining which schools continued.

The censoring loading captures the effect of the funding ceiling. A mission school funded only to Standard IV records its highest-attaining pupils at Standard IV even if they would otherwise have reached Standard V or Standard VII. Recorded quality q therefore understates capital k , with a larger understatement for schools whose pupils would have advanced further above the ceiling.

This compression has two opposing effects. Each recorded unit corresponds to more than one unit of underlying capital. The level loading $\Lambda_s = \text{Cov}(k, q) / \text{Var}(q)$ is therefore at least one

(Lemma 1), and the same mechanism raises the regression coefficient through the deviation-level loading Λ_s° (Appendix A.4). At the same time, compression reduces the dispersion of the recorded measure. The outcome association for a one-standard-deviation difference in recorded quality is therefore smaller, and the coefficient is estimated less precisely. In the model, the ceiling affects measurement only and leaves τ and π unchanged. Section 6.3 tests this mechanism within one system by imposing the mission ceiling on the public and poor records.

Equation (1) defines the object estimated or tested in each empirical exercise. The coefficient β is a weighted combination of the organisational and place channels, so it is a persistence parameter rather than a policy elasticity. The weight on the organisational channel depends on the interpolation power through $\omega(p)$ and on the dispersion of the own-school signal $\sigma_{\xi,s}^2$. The funding ceiling affects both Λ_s° and $\sigma_{\xi,s}^2$.

3.3 Predictions

The model gives three predictions that the data could have rejected. It also gives a fourth prediction that the available national data cannot test. Appendix A.1 states each prediction formally, and the remainder of Appendix A provides the proofs. We state the empirical content of each prediction here.

Censoring (Lemma 1, Proposition 1). A ceiling prevents the reports from recording the attainment of the highest-performing mission schools. It therefore compresses recorded quality and makes its persistence less precisely estimated. More formally, the variance of the recorded measure falls below the variance of capital. The coefficient per recorded unit rises by the loading $\Lambda^\circ \geq 1$, while the association per standard deviation falls and the standard error increases. If the systems share the same transmission parameters and differ only in the ceiling, the mission coefficient should be larger per recorded unit but smaller per standard deviation than the public coefficient. It should also be estimated less precisely. This prediction can be tested within one system. Imposing the mission ceiling on the public and poor schools' records should reproduce the pattern, with an increase in the per-unit coefficient close to the censoring loading measured directly from the capped records (Section 6.3, Appendix A.7).

Smoothing (Proposition 2). Lowering the inverse-distance power averages historical quality over a wider area. It therefore shifts weight in the estimand from the own-school signal ξ to the local average $\bar{q}_\ell$. When organisational transmission is positive but the smooth loading is negative, the gradient falls and may change sign as the measure is smoothed. This configuration describes an organisation that transmits quality but is located in an otherwise disadvantaged place. When transmission operates through place, the smooth loading is large and positive, and the gradient rises with smoothing. The trajectory across powers therefore orders the two channels without separately identifying them (Section 6.4).

Linking (Proposition 3). For historical schools matched by name and location to a modern successor, the analysis maintains organisational continuity ($C = 1$). The crosswalk motivates but does not establish this assumption. Own historical quality is observed without interpolation. In a

joint regression of successor outcomes on own quality and a leave-own-out neighbourhood surface, a positive coefficient on own quality combined with a non-positive neighbourhood coefficient requires an organisational channel. A pure place channel would make both coefficients positive (Section 6.6).

The fourth prediction concerns operator orientation (Proposition 4). If missions operated by high- and low-orientation societies were located in comparable places, the difference in modern outcomes would identify the organisational channel through $\tau\bar{c}\phi$. Missions operated by societies with a stronger educational orientation should produce better modern outcomes, conditional on comparable placement. The national atlas data do not satisfy this condition because proximity to a high-orientation society also proxies for local mission density. The prediction therefore has no successful test in this paper. Section 7.2 and Appendix D report the exercise as a diagnostic rather than a result.

Table 1 maps each prediction to the corresponding empirical exhibit.

Table 1: Model predictions and corresponding empirical exhibits

Model implication	Empirical exhibit	Rejection criterion
P1(i): compressed mission measure	Tables B.4, B.3	No concentration at Standard IV
P1(ii)–(iii) at the deviation level: per-unit coefficient rises; per-SD association falls	Tables 4, 5	Applying the ceiling to public records leaves coefficients unchanged
P2 with $\text{sign } A_M \neq \text{sign } S_M$: mission gradient falls; public gradient rises	Tables B.13, B.15	Ordering of the two systems’ trajectories reversed
P3 under $0 \leq S_M^L \leq A_M^L$: own quality predicts; neighbourhood quality does not	Table 6	Neighbourhood coefficient positive and own coefficient at zero
P4: orientation contrast	Table D.2	No valid test: the high- and low-focus groups are not comparable

3.4 What the model does not identify

The model leaves four objects unidentified, and the empirical analysis makes no claims about them. Organisational continuity $\bar{c}_s$ is not observed outside the linked sample. The parameters τ and π are not separately identified. Under stated assumptions, the linked sample supports a regression decomposition of the own-quality share of its observed association; it does not bound the structural contributions of τ and π . Any spatially smooth component of organisational transmission enters S_s and is observationally equivalent to place persistence in this and every other spatial design. The model and data also do not identify the effect of twentieth-century policy on mission schools. The model takes the ceiling and continuity processes as given and derives their implications for measurement; it does not explain their origins.

Table 2: Notation

Symbol	Defined in	Meaning
k_i	A1	organisational capital, in 1905 mean-Standard units
o_i, ϕ	A1	operator education orientation; its pass-through
κ_s	A2	regime ceiling on certifiable attainment
q_i, r_i	A2	measured 1905 quality; censored component
$\Lambda_s, \Lambda_s^\circ$	Lemma 1, §3.2	censoring loadings, in levels and in own-school deviations
$C_i, \bar{c}_s$	A3	organisational continuity; its share
τ, π	A3	organisational and place transmission
$\bar{q}_\ell, \xi_i$	§3.2	local quality average; own-school deviation
$\omega(p), m_j$	A4	own-signal weight of the power- p surface; interpolation noise
ζ_s	A4	placement-selection projection coefficient
A_s, S_s	eq. (2)	organisational loading; smooth loading
$b_s(p)$	eq. (1)	probability limit of the regression coefficient β

4 Data

4.1 The Cape Education Reports

The Cape Education Department published an annual report by the Superintendent General of Education. From 1881 through 1905, the reports list every state-aided school by name and district and give the results of its annual inspection. The published volumes contain no school-level detail after 1905. We digitised district-level information for every year from 1881 to 1905 and complete school-level information for 1881, 1896 and 1905.

For each school in 1896 and 1905, the report gives the number of pupils in each Cape Standard and the numbers presented for and passing the inspection examination. It records whether pupils were placed in a higher, unchanged or lower Standard than at the previous inspection. It also reports quarterly enrolment and attendance and the number of pupil-teachers in training. The 1881 report predates this level of detail. It lists each school’s subjects, enrolment and grant (Section 2). School-level expenditure was not published. The finance annexures report spending only by school class at the colony level.

We georeferenced schools from their names, districts and place descriptions in the reports. Appendix Table B.1 reports coverage. We locate 575 of 619 listed schools in 1881, 1,301 of 1,457 in 1896 and 1,639 of 1,788 in 1905 after removing competing duplicate geocodes. We exclude farm schools because they were small and short-lived. They taught 5.8 per cent of pupils in 1896 and one per cent by 1920.

Located schools enrolled more pupils than unlocated schools. Among public and poor schools, the located subset also records somewhat higher quality. We find no significant quality difference among mission schools. Of the georeferenced 1905 locations, 902 public and poor records and 640

mission records match report entries for the stated system and provide quality measures. The other 97 locations comprise 84 geocodes with no matching digitised inspection row, for which the report contains no usable inspection entry that year, and 13 industrial schools outside the two systems compared in this paper.

Mean Standard is available for 777 of the 902 public and poor records and 603 of the 640 mission records. These schools define the historical quality surfaces. The analytical sample is therefore narrower than the administrative universe: of 1,788 listed schools, 1,639 are georeferenced and 1,542 match one of the two systems; 777 public and poor schools and 603 mission schools have the headline measure. Appendix Table B.2 reports measure-specific coverage for all six measures. Figure 1 shows the overlap between historical and contemporary schools.

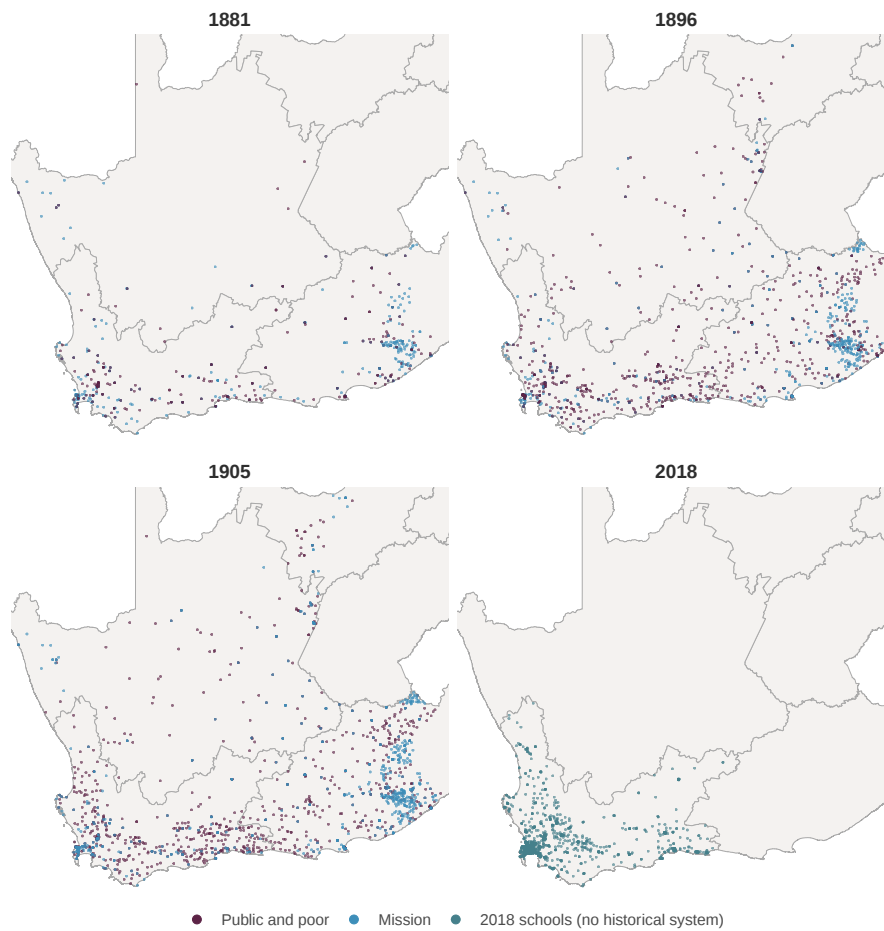


Figure 1: Georeferenced schools by year and system. The 1881, 1896 and 1905 panels show schools from the Cape Education Reports, coloured by historical system. The small industrial-school category is grouped with mission schools. The 2018 panel shows Western Cape schools with coordinates in the modern register; these schools have no historical system classification. Coverage in 2018 is limited to the province because the modern outcomes come from the Western Cape systemic tests. One 2018 school with an inconsistent coordinate is excluded.

The Education Reports cover schools that standard sources on African missions do not. Jedwab,

Meier Zu Selhausen and Moradi (2022) show that missionary atlases record mission stations rather than mission schools and omit most local schooling. The atlases contain 135 mission stations in the Cape Colony. The 1905 Education Report lists 706 mission schools and 13 industrial schools, of which we georeference 680. Its 1,069 public and poor schools do not appear in any missionary atlas. The reports also place both systems under a common measurement regime. The same inspectorate examined public, poor and mission schools and recorded their attainment on the same scale. All cross-system comparisons of recorded quality rely on this common scale.

4.2 Measuring school quality in 1905

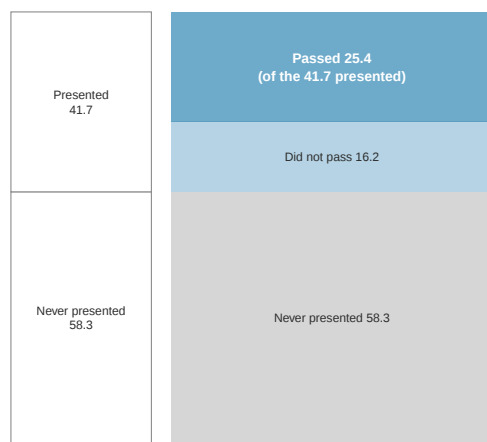
The headline measure is mean Standard, defined as the pupil-weighted average of the reported Standard distribution. We code the Sub-Standards as zero and Standard Ex as eight. The measure captures the realised level to which a school advanced and retained its pupils using the full attainment distribution rather than one threshold. It is not a pure measure of school supply because continued enrolment, and hence local demand for schooling, also affects the level pupils reach.

The reports provide several alternatives that place more weight on supply. The inspection pass rate is the number passing divided by the number presented. Annual progression is the share placed in a higher Standard than at the previous inspection. Readiness is the number presented divided by average enrolment. We also use the shares of pupils in Standard II or above and in Standard IV or above. Appendix Table B.2 summarises all six measures. The inspection pass rate and annual progression are weakly correlated with mean Standard, at 0.11 to 0.15, and therefore contain distinct information. Readiness has correlations of 0.60 to 0.72. The two threshold shares closely track mean Standard, with correlations of 0.82 to 0.96.

The inspection pass rate requires a qualification because schools chose its denominator. The rate describes the pupils presented for inspection, and the systems presented different shares of their pupils. Mission schools presented about two pupils in five, compared with about two in three in public and poor schools. Conditioning on presentation therefore selects a smaller share of mission pupils. When the pass rate is calculated over enrolment, so that a pupil not presented is counted as not passing, the cross-system gap roughly doubles. Conditional pass rates that are nearly equal correspond to a mission pass rate per enrolled pupil that is half the public rate. We report the conditional rate to match the reports, but interpret it only as the pass rate among pupils presented, not as annual school performance. Figure 2 compares the two denominators. Appendix Table B.7 reports the values.

The systems occupy different parts of the attainment scale in the pattern formalised by Assumption 2. Figure 3 shows the pupil-level distributions. Nearly two thirds of mission pupils were in the Sub-Standards, twice the public and poor share. Mission enrolment above Standard IV was almost zero. Appendix Table B.3 reports the corresponding school-level distribution. When schools are classified by the highest Standard containing any pupil, 44.4 per cent of mission schools end at Standard IV. Public and poor schools are distributed from Standard IV through Ex-Standard. Above the ceiling, mission enrolment is close to zero, whereas roughly one public or poor pupil

Mission schools



Of 100 pupils enrolled

Inspection outcome

Public and poor schools



Of 100 pupils enrolled

Inspection outcome

Figure 2: Pupils presented to the inspector and pupils passing per 100 pupils enrolled in 1905, by system. Values are school-level means from the 1905 report. The ratios shown in the figure are ratios of these means. They differ in the first decimal from the school-level mean conditional rates in Appendix Table B.7 because the mean of ratios differs from the ratio of means. The conditional rates, 25.4 of 41.7 presented and 49.8 of 65.0 presented, are similar. The pass rates per pupil enrolled differ by a factor of two because mission schools presented a substantially smaller share of their pupils.

in six is recorded. The mission grant financed a modified and abbreviated version of the public curriculum (Loram, 1917, p. 94). The concentration of mission schools at its upper funded level is consistent with a certification ceiling. Section 6.3 imposes this ceiling in the public and poor system.

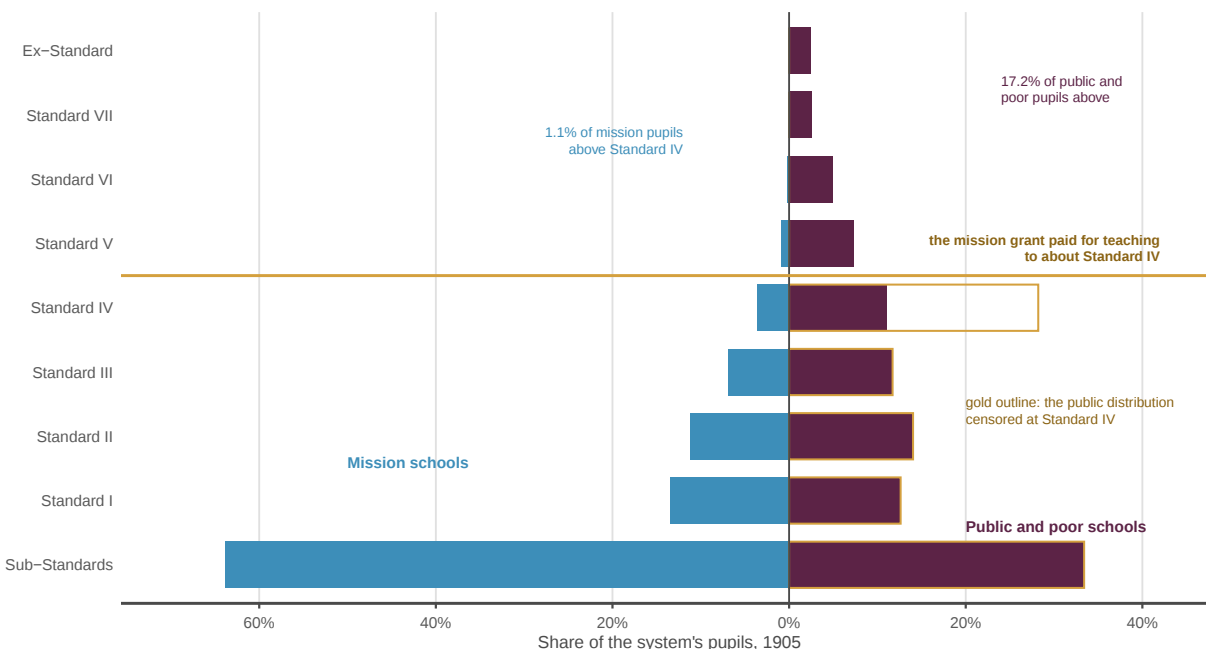


Figure 3: Distribution of pupils across the Cape Standards in 1905, by school system. Shares are calculated over pupils in each system listed in the 1905 report. Mission schools appear on the left and public and poor schools on the right. The gold horizontal line marks Standard IV, the level financed by the ordinary mission grant. The ceiling bound approximately rather than literally; 1.1 per cent of mission pupils appear above it (Remark 1). The gold outline on the public side shows the public and poor distribution after censoring at Standard IV and reassigning all pupils above the ceiling to that level. Section 6.3 uses this censored distribution.

4.3 Modern school outcomes

The Cape outcomes are school-level mean mathematics and language scores from the Western Cape systemic tests, which are administered annually in Grades 3, 6 and 9. We use the 2018 round. School locations and characteristics come from the Department of Basic Education's Education Management Information System. Appendix Table B.4 reports summary statistics for the estimation samples defined in Section 5.

The national outcomes are school-level matric (Grade 12) pass rates and bachelor pass rates averaged over 2021–2023. The sample contains 6,777 secondary schools with results in all three years. A bachelor pass is the higher National Senior Certificate outcome that permits application to a bachelor's degree, subject to university requirements. We classify contemporary schools by proximity to historical mission stations, not as present-day mission schools.

4.4 The national mission records

The national mission data use Gabriel Brown’s geolocation of two comprehensive atlases of South African mission stations: the Protestant atlas of Beach (1903) and the Catholic *Atlas Hierarchicus* of Streit (1929). The Beach atlas contributes 870 stations across the four historical colonies. Streit adds 41 Catholic stations located more than two kilometres from any Beach station, yielding 911 stations. Linking the Beach stations to later Protestant atlases (Dennis, Beach and Fahs, 1911; Beach and St. John, 1916; Beach and Fahs, 1925) produces a station-level panel of foreign staffing in 1900, 1908, 1915 and 1922 (Brown, 2025). The Beach atlas also reports educational statistics for 361 stations, including 1,050 village and day schools operating around 1900.

The atlases record stations rather than individual schools. In the Cape, where the atlas and Education Report can be compared, 71.6 per cent of georeferenced mission schools in 1905 lie within 10 kilometres of a Beach station and 93.6 per cent lie within 25 kilometres. Even so, the atlases identify fewer than one-fifth as many Cape mission locations as the Education Reports. The national analysis therefore measures proximity to a mission station, not the presence or quality of a particular school.

The closest national analogue to Cape school quality is the society-level education-orientation score constructed by Brown and Martin (2026). It uses the worldwide station records in Beach (1903), not only the South African stations. For each missionary society, the authors calculate the mean number of institutions in five educational categories: primary schools, secondary schools, colleges, kindergartens and industrial schools. They then extract the first principal component of these five means across societies. The component is standardised to mean zero and unit variance across the worldwide distribution of societies. One unit is therefore one standard deviation among missionary societies globally, not within South Africa.

A society that operated many schools per station receives a high score at all of its stations. The score measures the educational orientation of the organisation rather than the quality of a particular South African station. Because it is not divided by total mission activity, it measures the level of educational investment rather than education’s share of the society’s activities.

Sixty societies operated Beach stations in South Africa. Thirty-nine have an orientation score, covering 672 of the 870 stations. We classify a society as high-focus when its score exceeds the median of the South African distribution weighted by the number of stations operated by each society. This weighting produces station groups of similar size: 343 stations operated by ten high-focus societies and 329 operated by twenty-nine low-focus societies. The weighting is used for statistical power, not identification; it avoids a comparison between a large group and a small number of stations. Schools near missions operated by unscored societies form a separate category in the national analysis.

5 Empirical method

5.1 The Cape specification

For each historical system, we interpolate 1905 quality to contemporary school locations using inverse-distance weights with power four. We restrict the sample to contemporary schools within 10 kilometres of a historical school in that system. In the notation of Section 3, a high power makes $\omega(p)$ close to one, so the nearest historical schools receive most of the weight. We estimate the smoothing trajectory in Proposition 2 by repeating the specification across powers.

We choose a high power to match the estimand, not to minimise prediction error. The object of interest is transmission from a particular historical school, so the surface should place most weight on the nearest school. A low power substitutes a regional average for that school and therefore estimates a different object. Appendix Table B.17 shows that leave-one-out prediction error favours smoother surfaces. This criterion answers a different question. If organisations rather than places transmit the legacy, neighbouring historical schools contain little information about one another. A surface selected to predict a school omitted from its construction need not be the appropriate surface for measuring the school that remains in the construction. We therefore interpret the sweep across powers as the model’s smoothing trajectory rather than select the best-predicting surface as the baseline.

Figure 4 shows the assignment underlying the exposure measures. At power four, the nearest historical school receives most of the interpolation weight. Links from each 2018 school to its nearest 1905 school in each system therefore approximate the assignment encoded by the surfaces, and the sets of links around each 1905 school approximate its modern catchment. The networks of the two systems overlap throughout settled areas. Most 2018 schools are within ten kilometres of both a historical public or poor school and a historical mission school, allowing both exposure measures to enter the same regression.

The headline samples are nevertheless system-specific. Appendix Table B.14 repeats the analysis on the common-support sample of schools within 10 kilometres of both systems. It enters each surface separately and then jointly. Appendix Table B.15 reports the formal trajectory tests on the same sample.

The preferred specification is

$$Y_{ig} = \beta Q_{i,1905} + \mathbf{X}_i' \theta + \alpha_{d(i)} + \eta_{e(i)} + \varepsilon_{ig},$$

where Y_{ig} is the 2018 score for school i in grade g , and $Q_{i,1905}$ is the interpolated historical quality surface $\hat{Q}_i(p)$ at the baseline power. The terms $\alpha_{d(i)}$ and $\eta_{e(i)}$ denote 1891 census-division and historical ethnic-group fixed effects. The controls are distance to the coast, rainfall and temperature; altitude and local variation in altitude; and soil suitability. They also include distance to Cape Town, Port Elizabeth and Kimberley, distance to the historical railway and explorer routes, and an indicator for a cell containing urban population in 1900.

Urbanity is the main placement confound in the Cape design. Both systems operated schools

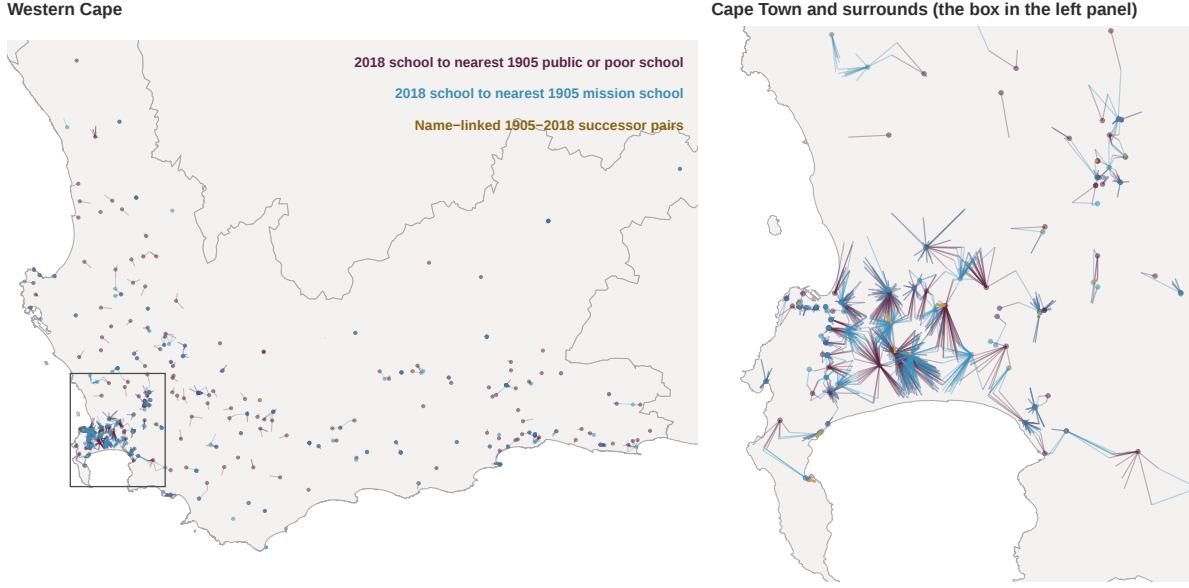


Figure 4: Links from each 2018 school to its nearest georeferenced 1905 school in each system: 1,363 public and poor links and 1,269 mission links. The 60 name-linked successor pairs from 1905 to 2018 are shown in gold. For legibility, links are drawn using edge-path bundling (Waller et al., 2022); links shorter than 250 metres are drawn as straight lines. One school with an inconsistent coordinate in the modern register is excluded. The right panel enlarges the boxed area in the left panel.

in towns and villages. Towns contained mission schools, and by 1905 the department’s third-class public schools served many villages. Within both systems, however, better-resourced schools were concentrated in towns, and urban status directly predicts modern outcomes.

The baseline measure comes from the HYDE reconstruction of historical population in 1900 (Klein Goldewijk et al., 2017). We classify a school as urban if its five-arc-minute grid cell, approximately nine kilometres across at this latitude, contained any urban population in 1900. Contemporary urban status lies between the 1905 record and the 2018 outcome. Replacing the 1900 measure with contemporary status raises the public and poor coefficients by as much as half and reduces their standard errors.

The timing of the contemporary control does not determine the direction of the change. An on-path control may remove a mediated component of persistence. Suppression or a more accurate proxy for persistent place characteristics could produce the same movement. The specifications therefore estimate different conditional associations; one cannot be labelled the true estimate and the other biased on timing alone. The 1900 indicator precedes both the records and the outcome, but earlier schooling may have contributed to a place becoming urban by 1900. The baseline therefore retains this concern as an assumption rather than eliminating it. Organised schooling was present by at least 1881, so the 1900 indicator is not fully predetermined. Appendix Table B.9 makes this choice explicit. Appendix Table B.9 compares the baseline with no urban control, continuous population density in 1900 and contemporary urban status. The first three specifications agree;

the contemporary control does not. The coefficient β estimates the composite $b_s(p)$ in equation (1). The regression compares contemporary schools within broad historical geographies and does not randomly assign historical quality.

The Cape contains 24 historical divisions, with fewer effective degrees of freedom in some grades. The main regressions and compiled appendix tables use Bell–McCaffrey CR2 standard errors clustered by division with Satterthwaite degrees of freedom (Bell and McCaffrey, 2002; Imbens and Kolesár, 2016). We report Conley calculations only in a clearly labelled sensitivity table because none of the covariance matrices reported there is positive semidefinite (Conley and Kelly, 2025).

5.2 Linked schools

The crosswalk begins with candidate historical–modern pairs that share a distinctive, non-generic name or place token and are no more than two kilometres apart. A candidate must be unique on the modern-school side, and we review each pair manually. We exclude six false matches based on generic place names. Confidence tier A further requires a distance of no more than one kilometre and a near-exact occurrence of the shared token in the modern school name or suburb. The final crosswalk contains 60 links, of which 53 have both historical mean Standard and a 2018 outcome.

The linked outcome is the average of language and mathematics scores after standardisation within grade. Own historical quality is also standardised. The neighbourhood measure is a power-two historical quality surface recalculated after removing the linked historical school. This prevents the school’s own quality from mechanically dominating its local interpolation.

Proposition 3 interprets this design. Within the linked sample, name-and-location continuity holds by construction. It proxies for organisational continuity but does not establish that the organisation continued unchanged. Own historical quality is observed without interpolation. Given the small sample, these regressions use heteroskedasticity-robust HC3 standard errors (MacKinnon and White, 1985).

5.3 The national specification

The national proximity specification regresses matric outcomes on an indicator for a historical mission station within 10 kilometres, an urban indicator and either province or education-district fixed effects. The education district provides the sub-provincial fixed effect. The district-municipality field in the school register does not identify the education district because 51 of its 74 values occur in more than one province. Education districts are consistent with province in all cases.

The national specification uses contemporary urban status, unlike the 1900 measure in the Cape specification, because the data contain no historical population measure at school level. Missions were founded near accessible places, making urbanity the main placement confound (Jedwab, Meier Zu Selhausen and Moradi, 2022). A contemporary control absorbs more variation than a historical measure would. The estimates are therefore conditional descriptive associations, not conservative bounds on either the unconditional association or a causal effect.

The society-orientation specification uses the same controls and classifies mission-proximate schools by the educational orientation of the operating society, as in Proposition 4. When several scored missions lie within 10 kilometres, the continuous exposure is the maximum society score among them. We exclude contemporary school quintile, pupil–teacher ratios and enrolment because they may be outcomes of the historical process. National school-level analogues of the detailed Cape geographic and historical controls are unavailable in these data.

The national coefficients are descriptive spatial associations. We report 50-kilometre Conley standard errors, and education-district fixed effects provide the tighter geographic comparison. We retain Conley inference here although it is demoted in the Cape analysis. The Cape objection arises from its 24 historical divisions and spatially concentrated treatment, which produce covariance matrices that are not positive semidefinite. The national sample has no comparable cluster structure.

6 Cape results

6.1 Which records were stable

We first assess whether the 1905 measures were stable over time. The reports permit a direct comparison. We match schools within system between the 1896 and 1905 reports, requiring the identifier to be unique on both sides and within each report. This procedure yields 828 one-to-one links: 438 public and poor schools and 390 mission schools. For each measure, we regress its 1905 value on its 1896 value after standardising within system and year and include 1896 report-district fixed effects. The coefficient measures repeatability over nine years.

Table 3: Persistence of school records between 1896 and 1905

System	Matched schools	Pearson	Spearman	Adjusted slope	CR2 SE	<i>p</i> -value
<i>Mean Standard</i>						
Public/poor	388	0.720	0.689	0.697	0.047	< 0.001
Mission	365	0.312	0.351	0.282	0.096	0.017
Public/poor minus mission	753	–	–	0.415	0.121	0.003
<i>Inspection pass rate</i>						
Public/poor	384	0.039	0.056	0.013	0.064	0.840
Mission	362	0.223	0.214	0.220	0.043	< 0.001
Public/poor minus mission	746	–	–	-0.206	0.074	0.010
<i>Annual progression rate</i>						
Public/poor	330	0.022	0.025	0.016	0.078	0.843
Mission	315	0.132	0.140	0.094	0.071	0.207
Public/poor minus mission	645	–	–	-0.078	0.105	0.463

Notes. The sample contains one-to-one school links that remain in the same institutional system. The adjusted slope is from a regression of the within-system standardised 1905 measure on its standardised 1896 value, with 1896 report-district fixed effects. Standard errors use CR2 and are clustered by report district. The difference row reports the public/poor slope minus the mission slope from a stacked model with system-specific district fixed effects. Mean Standard records attained level; inspection pass and progression record annual administrative performance.

Attained level is substantially more repeatable in the public and poor system. Its slope is 0.70, compared with 0.28 for mission schools. The ordering reverses for the inspection pass rate. Repeatability is essentially zero among public and poor schools and modestly positive among mission schools.

The low repeatability of the pass rate is consistent with the construction of the measure rather than an error in the records. Schools chose which pupils to present for inspection, so the pass rate combines attainment with selection into the examination. If presentation choices changed across years, the measure would not be repeatable. Selection alone does not imply this result, because a stable presentation rule would reproduce. The interpretation is therefore consistent with the slopes but not implied by them. It depends on the stability of presentation rules and cohort composition and on unobserved year-specific measurement error, which the records do not separately identify.

Annual progression is defined over all enrolled pupils rather than only those presented. It produces the smallest difference between systems and is the within-school measure least exposed to both presentation selection and the ceiling in mean Standard. Figure 5 shows the school-level patterns. For attained level, public and poor schools form a narrow band around the diagonal, whereas mission schools are more dispersed. The pass-rate panels reverse this contrast.

A mission school’s recorded attainment in one year is therefore a weak predictor of its recorded attainment nine years later. This pattern is consistent with a measure compressed at a ceiling and recalculated over a changing enrolment. It also provides a reason to expect the mission estimates below to be less precise. The reversed ordering for the pass rate shows that the systems were not simply measured with different amounts of noise. Each system has a different margin on which its

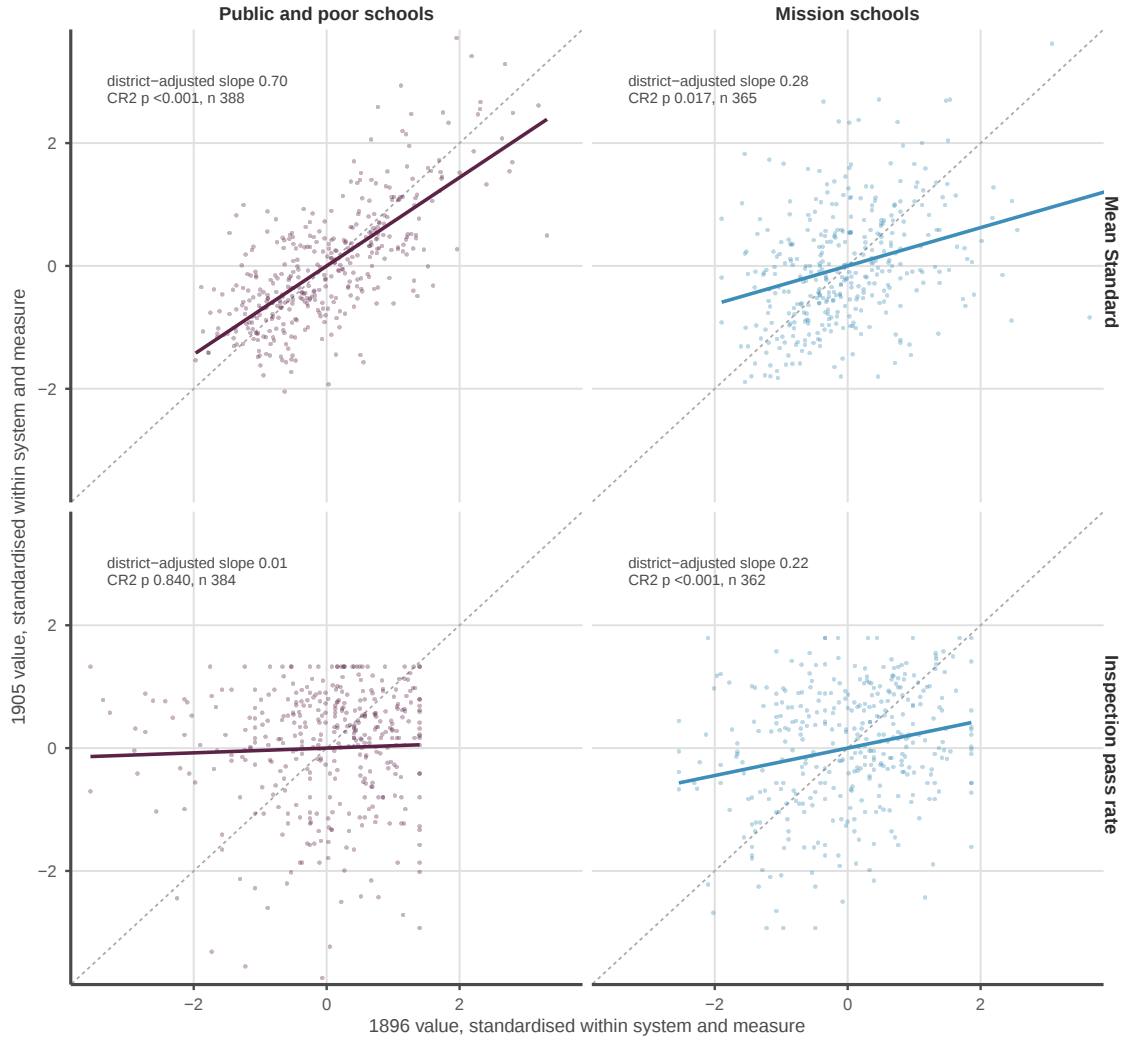


Figure 5: Matched schools in 1896 and 1905. Each point is one of the 828 one-to-one links, standardised within system and measure. Panel-specific samples require finite values in both years and a recorded district in 1896, so each panel corresponds to the relevant column of Table 3. The dashed line has slope one and denotes perfect repeatability. The solid line is the unadjusted fit. The reported slope and p -value are the district-fixed-effects estimates in Table 3.

records are more stable.

6.2 Historical quality and modern mathematics

Table 4 reports the main estimates. Recorded performance in 1905 is positively associated with mathematics performance in 2018 in both systems, and the association increases with the grade of the outcome. Precision is limited. Two of the six estimates differ from zero at the ten per cent level, none differs at the five per cent level, and the standard error of the Grade 9 mission estimate exceeds its coefficient.

The table alone is therefore suggestive rather than conclusive. The interpretation depends on whether the cross-system pattern recurs in the ceiling experiment, smoothing trajectory, alternative measures and linked schools. Some imprecision also appears to reflect year-specific measurement error. Averaging the 1896 and 1905 measures for the same schools increases every coefficient (Appendix Table B.8).

The comparison across panels is more informative than whether individual coefficients cross a significance threshold. At every grade, the mission coefficient is larger per recorded unit. The ordering reverses when effects are expressed per standard deviation of the treatment. Mission standard errors are between two and three times larger.

The difference between the two scalings reflects dispersion. Across historical schools, the standard deviation of recorded mission quality is less than half that of public quality. After interpolation to contemporary schools, it is less than one-third as large; Table 4 reports both values. A one-standard-deviation difference in the mission treatment therefore represents a smaller difference in recorded attainment than the corresponding public treatment.

These dispersions refer to recorded quality and the interpolated surface, not to the own-school deviation ξ used in Proposition 1. The reversed ordering is therefore consistent with the proposition’s deviation-level result, but does not estimate it directly. The larger mission standard errors are also consistent with lower treatment dispersion. Under CR2 inference, however, differences between samples can also affect precision. Proposition 1(iv) guarantees the ordering only under a common sample and common residual variance.

The headline panels use system-specific samples, so a coefficient could in principle absorb the geography of the other system. Appendix Table B.14 tests this possibility on the common-support sample. The two surfaces have a correlation of 0.26 and are therefore nearly orthogonal. When both enter the same regression, each association remains positive at every grade, and the mission association becomes somewhat stronger. The separate-sample estimates are not driven by overlapping historical geography.

This pattern corresponds to parts (ii)–(iv) of Proposition 1 at the level of the interpolated treatment; Appendix A.7 derives the surface-level result. A certification ceiling raises the coefficient per recorded unit because each unit represents more underlying capital. It reduces dispersion, so the association per standard deviation does not increase and precision falls. This result does not imply that mission quality mattered less. It implies that the reports measured mission quality on a

Table 4: Historical recorded performance and 2018 mathematics performance

Term	Grade 3 mathematics	Grade 6 mathematics	Grade 9 mathematics
Public/poor-school performance			
1905 mean Standard	0.74	1.08*	2.31*
Standard error	(0.40)	(0.49)	(0.89)
Satterthwaite d.f.	5.9	5.9	4.5
Historical-school SD	0.99	0.99	0.99
SD of interpolated treatment	0.72	0.72	0.72
Association per interpolated-treatment SD	0.53	0.78	1.66
Observations	981	948	427
Division clusters	24	24	24
R^2	0.205	0.192	0.209
Mission-school performance			
1905 mean Standard	1.78	2.58	2.63
Standard error	(0.94)	(1.37)	(3.02)
Satterthwaite d.f.	5.7	5.6	4.7
Historical-school SD	0.46	0.46	0.46
SD of interpolated treatment	0.23	0.23	0.23
Association per interpolated-treatment SD	0.40	0.58	0.59
Observations	905	874	411
Division clusters	24	24	24
R^2	0.201	0.175	0.205
Historical-division fixed effects	Yes	Yes	Yes
Mapped ethnolinguistic-group fixed effects	Yes	Yes	Yes
Geographic and historical controls	Yes	Yes	Yes
Urban in 1900	Yes	Yes	Yes
Maximum historical-school distance	10 km	10 km	10 km
Inference	Division-clustered CR2	Division-clustered CR2	Division-clustered CR2

Notes. The outcome is the school-level mean 2018 mathematics score. The treatment is the power-4 inverse-distance-weighted 1905 mean Standard for the named historical system. The historical-school SD is calculated across georeferenced 1905 schools; the interpolated-treatment SD is calculated across contemporary schools in the regression sample. The association per interpolated-treatment SD multiplies the coefficient by the latter SD. Standard errors use CR2 clustered by 1891 census division with Satterthwaite degrees of freedom. Because the design has only 24 division clusters and the treatment is spatially concentrated, the Satterthwaite degrees of freedom are small (about 4 to 6), so the two-sided critical values are well above the normal: roughly 2.0 at ten per cent, 2.5 at five per cent and 4.3 at one per cent. Stars are assigned against those values, not against coefficient-to-standard-error ratios read off a normal table. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

compressed scale. The cross-system comparison also does not identify the effect of twentieth-century policy because the systems already differed in clientele, resources, measurement and location.

6.3 The ceiling experiment

Proposition 1 interprets the cross-system pattern as a measurement effect of the ceiling. We test whether a Standard IV cap can produce this pattern within the public and poor system, holding other cross-system differences fixed. For every historical public or poor school, we recalculate 1905 quality after limiting each pupil’s recorded Standard to Standard IV, where mission enrolment ends in Appendix Table B.3. We then interpolate the censored records using the same power-4 procedure and re-estimate the headline specification.

The ceiling mechanism predicts that censoring will compress treatment dispersion and raise the coefficient per recorded unit. It should lower the association per treatment standard deviation. Under the common-sample and common-residual-variance conditions of Proposition 1(iv), it should also increase standard errors, although CR2 inference may affect them through other channels. Censoring could instead have left the coefficients unchanged, in which case the experiment would have rejected the mechanism.

Table 5: Imposing the mission ceiling on public-school records

Term	Grade 3 mathematics	Grade 6 mathematics	Grade 9 mathematics
Panel A: observed public/poor recorded performance			
Coefficient	0.74	1.08*	2.31*
Standard error	(0.40)	(0.49)	(0.89)
SD of interpolated treatment	0.72	0.72	0.72
Association per treatment SD	0.53	0.78	1.67
Panel B: public/poor recorded performance capped at Standard IV			
Coefficient	0.84	1.44*	3.03*
Standard error	(0.44)	(0.62)	(1.33)
SD of interpolated treatment	0.50	0.50	0.49
Association per treatment SD	0.42	0.72	1.49
Ratio of per-unit coefficients (B/A)	1.14	1.34	1.31
Censoring loading $\hat{\Lambda}_g$ (partialled surfaces, grade sample)	1.35	1.35	1.39
Calibration test p ($\beta_B - \hat{\Lambda}_g \beta_A = 0$)	0.45	0.96	0.59
Observations	981	948	427

Notes. Both panels use the public/poor estimation sample, outcome 2018 mathematics, and the headline specification (historical-division and mapped ethnolinguistic-group fixed effects, geographic and historical controls, division-clustered CR2 inference). Panel A uses the observed power-4 interpolated 1905 mean Standard. Panel B replaces each historical public school’s Standard distribution with the same distribution censored at Standard IV — the level at which mission enrolment concentrates — before interpolation. The censoring loading is the coefficient from projecting the uncapped surface on the capped surface after partialling out the control set (including the 1900 urbanisation indicator), computed on each grade’s exact estimation sample. If the additional orthogonality condition in Appendix A.7 held exactly and sampling noise were negligible, the Panel-B per-unit coefficient would exceed Panel A by this factor. The calibration test stacks the two equations on the same observations, fully interacted, with CR2 division-clustered inference, and tests that equality; it is conditional on the realised surfaces and $\hat{\Lambda}_g$, whose sampling error it does not propagate. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5 reports the results. Censoring reduces the dispersion of the interpolated public and poor treatment by about one-third. Every coefficient per recorded unit increases, while every association per treatment standard deviation falls. Every standard error also increases. The experiment therefore reproduces the mission pattern within the public and poor system.

The size of the coefficient increase provides a stronger test. The realised coefficient ratios are 1.14, 1.34 and 1.31. The corresponding grade-specific censoring loadings are $\hat{\Lambda}_g = 1.35, 1.35$ and 1.39. These loadings are obtained independently by projecting the uncapped surface on the capped surface within each grade’s estimation sample; the regressions do not use them. A conditional calibration test of $\beta_B = \hat{\Lambda}_g \beta_A$ stacks the two equations on the same observations and uses division-clustered CR2 inference. It rejects in no grade ($p = 0.45, 0.96$ and 0.59). The test conditions on the realised surfaces and $\hat{\Lambda}_g$ and does not propagate sampling error in the loading.

The decline in the association per standard deviation provides additional information. Under the orthogonality and non-negative-loading conditions in Appendix A.7, it favours the benchmark in which outcomes respond to underlying capital over the pure certified-attainment benchmark. The latter would weakly increase the standardised association under censoring. Mixtures of the two benchmarks are not separately identified. The ceiling mechanism therefore reproduces the full cross-system pattern within one system, holding clientele, resources, location and twentieth-century policy constant by construction. The cap alone can generate the observed scaling at approximately the magnitude predicted by the censoring loading. The exercise does not show that censoring was the only source of the mission distribution.

6.4 The smoothing trajectory

Proposition 2 predicts how the coefficient changes when the interpolation power falls and the treatment places less weight on the nearest historical school. Figure 6 traces both systems from power 0.5 to power 6 in quarter-step increments. Appendix Table B.13 reports integer powers from one to six.

The trajectories move in opposite directions and are each monotone for the Grade 3 outcome. The mission association increases as the measure places more weight on the nearest school. It is positive at every power above 1.4 and changes sign only below that value, when quality is averaged more broadly. Under Proposition 2, this pattern requires an organisational loading and a smooth loading with opposite signs. The model therefore interprets the result as transmission through mission schools located in places whose other characteristics predict lower modern outcomes. Appendix Table B.10 provides related evidence: town fixed effects remove the disadvantaged-place component by comparing schools within towns.

The public and poor trajectory has the opposite slope. The association roughly trebles as the measure broadens from the nearest school towards the district and remains above zero. In the model, this pattern corresponds to a large positive smooth loading. Public and poor schools were located in places whose characteristics independently predict higher modern outcomes, so broader averaging raises rather than attenuates the association.

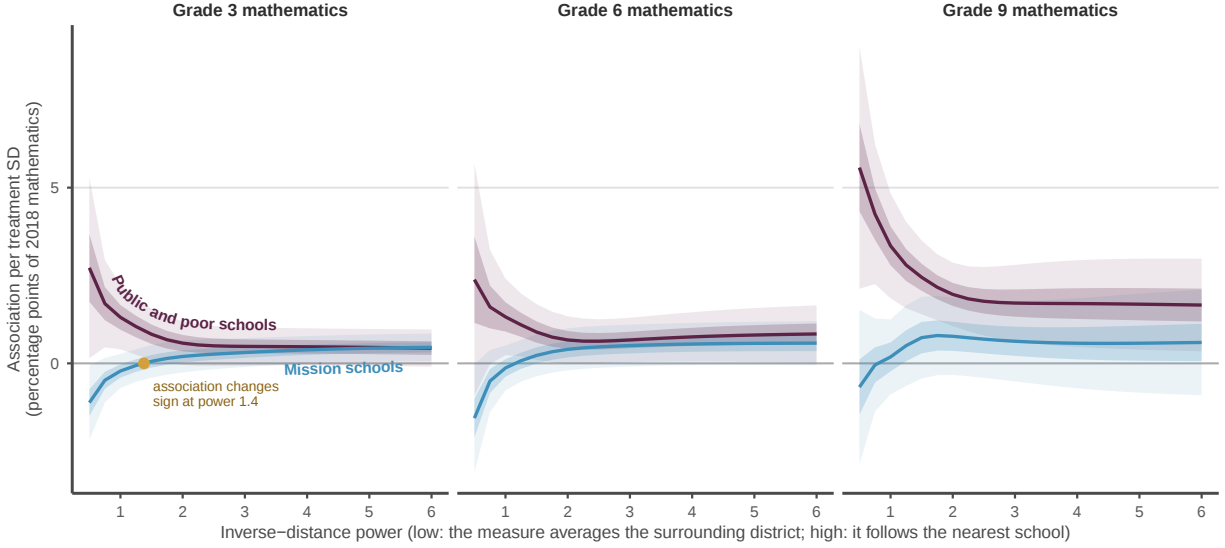


Figure 6: Association between recorded performance in 1905 and mathematics performance in 2018, expressed per treatment standard deviation, across inverse-distance powers. A low power averages historical performance over the surrounding district; a high power places more weight on the nearest school. Lines show point estimates at quarter-step increments from 0.5 to 6 using the samples and specification in Table 4. A single harmonised interpolation engine is used throughout; its power-4 surface has a correlation of 0.994 with the precomputed headline surface. Bands show 50 and 90 per cent pointwise intervals based on CR2 standard errors clustered by 1891 division with Satterthwaite degrees of freedom. They are not simultaneous intervals, and estimates at neighbouring powers are highly correlated. The figure therefore does not support joint inference across powers. Appendix Table B.15 reports formal endpoint contrasts at powers 1 and 4 and the cross-system difference between them. Appendix Table B.12 reports trajectories excluding Cape Town. The gold point marks the Grade 3 mission association’s sign change at power 1.4. Appendix Table B.13 reports values at powers one through six.

The smoothing comparison distinguishes the spatial organisation of the two legacies. An association transmitted locally through a surviving institution is attenuated when the measure is averaged over the surrounding area. An association transmitted through place becomes stronger under the same averaging. The movement of the mission coefficient at power one is therefore the pattern implied by organisational transmission combined with compensatory placement, not a separate specification failure.

Three additional exercises assess the contrast beyond the pointwise confidence intervals. Appendix Table B.15 reports a formal endpoint test. We stack joint two-system equations at powers 1 and 4 on the common-support sample and enter both surfaces together. Between the two endpoints, the public and poor gradient falls and the mission gradient rises in every grade. The system-by-smoothing difference-in-differences is -1.77 per treatment standard deviation in Grade 3 ($p = 0.002$), -1.56 in Grade 6 ($p = 0.06$) and -3.22 in Grade 9 ($p = 0.07$), using CR2 inference clustered by division. These contrasts compare powers 1 and 4 and do not test monotonicity over the full grid.

Appendix Table B.12 excludes the City of Cape Town. The two trajectories continue to move in opposite directions and are, if anything, steeper.

The power sweep also has a purely spatial interpretation because it traces the covariance scale of each quality surface. Fine-scale quality is averaged away, whereas broadly organised quality remains. Appendix Table B.16 and Figure B.1 show that recorded quality in the two systems was organised at the corresponding scales. Mission quality is strongly correlated among schools within two kilometres, and the correlation declines with distance. Public and poor quality has a weaker short-range correlation and a broader spatial reach. It is almost entirely a between-district pattern and disappears after removing district means. The trajectories are therefore consistent with the mechanisms' predicted spatial scales. The model interprets the fine-scale mission signal as organisational and the broad public structure as place. The spatial evidence alone does not prove this interpretation because highly localised place persistence would produce the same scale pattern. Section 6.6 addresses this remaining ambiguity with linked successor schools.

6.5 Alternative measures

Appendix Table B.6 estimates the six quality measures using the same sample definition, controls and CR2 inference. Effects are reported per standard deviation of the interpolated treatment. Mission pass rates and progression have positive coefficients across grades, and most public and poor coefficients are positive. Readiness is weaker in both systems. The public and poor threshold measures closely follow mean Standard, whereas the mission threshold estimates are less stable.

The pass-rate and progression results are informative about the ceiling. Both are defined within the Standards a school taught, so they are less exposed to the certification ceiling than mean Standard. Both produce a positive mission gradient. Progression is the more informative measure because it covers all enrolled pupils and is not affected by selection into presentation. It is not fully independent of the ceiling, since the cap could still restrict advancement above Standard IV. Even

so, a positive gradient for a measure without presentation selection and with less exposure to the ceiling is consistent with persistent mission quality and with mean Standard understating it.

Language results are weaker, particularly for mission schools (Appendix Table B.19). CR2 addresses inference with few spatial clusters, but not historical placement or measurement error. Averaging the 1896 and 1905 observations for the same schools increases every coefficient and makes the Grade 3 and Grade 6 mission estimates significant (Appendix Table B.8). This pattern is consistent with attenuation from year-specific measurement error in the single-year estimates, especially for the less stable mission records.

6.6 Institutional continuity

Table 6 reports the linked-school estimates associated with Proposition 3. The full sample contains 53 links, pools the two historical systems and controls for system. In a regression containing only own historical quality, its coefficient is 0.53 standard deviations. The coefficient is 0.54 when the neighbourhood surface is added. The point estimate is similar in confidence tier A and in samples restricted to one kilometre, although these smaller samples are imprecise. In the joint regression, leave-own-out neighbourhood quality has a coefficient of -0.09 and a standard error of 0.14.

Table 6: Linked historical and successor schools: own quality and neighbourhood quality

Specification	Coefficient	HC3 standard error	Observations	R^2
All retained links	0.526*	(0.291)	53	0.146
Confidence tier A only	0.571	(0.365)	38	0.190
Distance no greater than 1 km	0.507	(0.362)	42	0.157
All links: own quality, joint with neighbourhood	0.543*	(0.292)	53	0.153
All links: leave-own-out neighbourhood	-0.089	(0.143)	53	0.153

Notes. The crosswalk contains 60 retained name-and-location links; 53 have both 1905 mean Standard and a 2018 test outcome. Candidate pairs had to share a distinctive non-generic name or place token, lie within 2 km and be unique on the modern-school side. Tier A additionally requires distance no greater than 1 km and near-exact occurrence of the shared token in the modern name or suburb. Six generic-place false matches were manually excluded. The outcome averages language and mathematics scores standardised within grade. Own-quality specifications control for historical system. The final two rows report the own-quality and neighbourhood coefficients from a single joint regression, where the neighbourhood measure is a power-2 neighbourhood-quality surface recomputed after removing the linked historical school. HC3 standard errors; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This coefficient configuration requires an organisational channel in Proposition 3. With a pure place channel, both coefficients would be positive because own quality would proxy for location when the leave-own-out surface is measured with error. A positive own coefficient combined with a non-positive neighbourhood coefficient requires a positive organisational loading at least as large as the smooth loading.

The decomposition in Proposition 3 gives a point-calibrated share, not a confidence bound. The decomposition is degenerate at the point estimates because the neighbourhood coefficient is

negative. Evaluating it instead at the neighbourhood coefficient’s upper 95 per cent confidence limit of 0.19, and using the joint own coefficient of 0.54, implies an own-quality share of at least 65 per cent in the linked sample. This calculation does not propagate the joint sampling uncertainty in the two coefficients and the projection weight.

The linked evidence remains corroborative and pooled. Because the 53 links span both systems, it supports transmission through surviving schools in general rather than through mission schools specifically. The matching procedure selects on name and locational continuity, and an unobserved attribute of both school and site could also persist. Under quality-monotone survival, the own-quality coefficient understates transmission. It overstates transmission if twentieth-century policy removed the strongest mission lineages (Appendix A.5). The table therefore does not establish that institutional continuity caused the broader Cape coefficients.

7 National evidence

7.1 Historical mission proximity

The atlases measure only the extensive margin: whether contemporary outcomes differ near a historical educational institution. Table 7 shows that they do. Within both provinces and education districts, schools within 10 kilometres of a historical mission have higher school-leaving pass rates. The difference is about two percentage points for the ordinary pass rate and approximately twice as large for the bachelor-level pass rate.

Table 7: Historical mission proximity and national secondary-school outcomes

Term	Matric pass	Matric pass	Bachelor pass	Bachelor pass
	Province fixed effects	District fixed effects	Province fixed effects	District fixed effects
Historical mission within 10 km	2.35***	2.05***	4.31***	3.58***
Standard error	(0.51)	(0.48)	(0.63)	(0.57)
Observations	6,777	6,777	6,777	6,777
R^2	0.144	0.188	0.131	0.167
Urban indicator	Yes	Yes	Yes	Yes
Fixed effects	Province	Education district	Province	Education district
Contemporary quintile control	No	No	No	No
Inference	Conley, 50 km	Conley, 50 km	Conley, 50 km	Conley, 50 km

Notes. Outcomes are school-level averages for 2021–2023. The matric pass rate is the share of candidates passing the National Senior Certificate; the bachelor pass rate is the share obtaining the stronger pass that permits application to bachelor’s-degree study. The treatment indicates a historical Beach or Streit mission station within 10 km. Contemporary school quintile is excluded because it may be post-treatment. Standard errors allow spatial correlation within 50 km. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Replacing province fixed effects with education-district fixed effects reduces the estimates by roughly one-sixth, consistent with part of the proximity association reflecting mission placement. Neither specification eliminates that selection. Missions initially chose accessible and economically

promising places before expanding to less connected areas (Jedwab, Meier Zu Selhausen and Moradi, 2022). The national coefficients are therefore associations between historical mission geography and present outcomes. They show that the extensive-margin association extends beyond the Cape, but do not identify a mission effect. There is no national counterpart to the Cape quality comparison because no national source records what individual schools taught.

7.2 Missionary-society educational orientation

Proposition 4 predicts different outcomes near missions operated by societies with different educational orientations, provided the place characteristics that attracted missions are comparable across orientation groups. The atlas permits an attempted test, but the comparability condition fails.

The failure appears for both the continuous score and a high-versus-low classification. Each produces a large apparent orientation association. The association is substantially attenuated and becomes specification-sensitive once local mission density is held constant. Using the maximum society score among nearby missions treats the exposure as an upper envelope, so a nearby low-orientation operator does not dilute a high-orientation one. This measure predicts matric and bachelor pass rates at the one per cent level, but controlling for the number of nearby missions removes more than two-thirds of the association.

The categorical specification defines a school as high-focus if any mission within 10 kilometres belongs to a high-focus society, with high-focus taking precedence over low-focus. This rule produces a large and apparently precise high-minus-low difference. It also assigns high-focus status more often where missions are dense. Schools in the high-focus category have more than three times as many missions nearby.

Randomly permuting orientation labels across societies reproduces almost the same density imbalance. The imbalance is therefore generated by the coding rule rather than by the locations chosen by high-orientation societies. At the station level, these societies operated in sparser neighbourhoods. Assigning each school symmetrically to its nearest scored mission removes the imbalance and produces a province-level contrast that is statistically indistinguishable from zero.

The nearest-mission assignment does not eliminate every contrast. A smaller contrast remains within education districts and for the bachelor-level pass rate. The operating society’s orientation may affect outcomes, but the national data do not establish such an effect. The specification that appears to do so is also the specification in which the coding rule mechanically creates the imbalance.

We therefore draw no conclusion about operator orientation, and Proposition 4 has no successful national test. Appendix D reports the full analysis. We retain the negative result because the design is common in research on historical institutions and the problem is not specific to these atlases.

When treatment is defined as the presence of any unit of a type within a radius, and one type takes precedence over another, the treatment indicator also measures local unit density. If density predicts the outcome conditional on the specification, the estimated type difference will partly reflect the density difference. Permuting type labels at the level at which they are assigned,

in this case the missionary society, reveals this problem without additional data. Studies that compare historical operators by proximity should report this diagnostic. The result also explains why operator-quality claims require school-level quality measures of the kind available in the Cape reports but not in the atlases.

7.3 What the national data do not establish

The atlas panel does not provide a credible national measure of school closure or administrative takeover by the state. Appendix Table E.1 classifies missions with foreign staff recorded in 1900 according to whether foreign staff are again recorded in 1922. Neither the ordering of these categories nor the interaction between the continuous education score and 1922 staffing produces a consistent continuity gradient. Missing foreign staff may indicate closure, changes in atlas coverage or transfer to African staff. We therefore do not use this measure to explain the Cape public-mission comparison.

For the same reason, we do not interpret atlas school counts as a dose, founding year as cumulative exposure, former apartheid education department as a quality regime, or the Land Act as an event of mission destruction. These variables may be relevant to other questions. In the available data, they do not improve the paper’s common comparison between historical educational quality and modern outcomes.

8 Conclusion

Research on institutional persistence generally assumes that institutional quality is transmitted over time, but rarely observes that quality in the past. We observe it for two school systems assessed by the same inspectorate on the same administrative scale and ask whether recorded quality persists. It does when an organisation or a local condition remains to transmit it. In the Cape, quality recorded for a school in 1905 predicts the performance of surrounding schools in 2018. The relevant spatial scale differs between the two systems, in the pattern that the model assigns to different transmission channels.

For public and poor schools, the association remains and becomes stronger as historical quality is averaged across the surrounding district. The persistent component is therefore spatially smooth. It may reflect place, the characteristics on which school placement selected, or a smooth component of organisational transmission. No spatial design can distinguish among these mechanisms. For mission schools, the association is present when the measure places most weight on the nearest historical school and disappears when the measure is smoothed. The persistent component is local to the school. The model interprets this fine-scale signal as organisational, and the linked successors corroborate an organisational component. Highly localised place persistence could nevertheless generate the same scale pattern. The same colonial state, inspectorate and administrative scale thus produced two legacies operating at different spatial scales.

A single administrative record from 1905 carries measurable information about school performance more than a century later. A one-standard-deviation difference in historical quality corresponds to between eight and twenty-one per cent of a standard deviation of mathematics performance in 2018. The systems differ more in what that record could contain than in the strength of its association with modern outcomes. Mission schools were funded to teach to about Standard IV, and their recorded attainment concentrates at that level; whether the ceiling also constrained what they taught cannot be determined from these records (Remark 1). Censoring the public and poor schools' records at the same point reproduces the mission pattern within that system at approximately the magnitude predicted by censoring alone.

These findings change the interpretation of persistence. Evidence that history predicts the present is often taken to imply that the past constrains what policy can achieve. The two systems support the opposite interpretation. They diverged because one system was resourced to do less. Its grant rule capped what it could certify, paid a fraction of the public grant per pupil and supported a substantially less trained teaching force (Loram, 1917). Whether each legacy could persist through organisations also depended on which organisations continued, a margin the education acts of the following century placed under direct policy control.

Measurement matters equally for the missionary literature. Counts of missions and comparisons between nearby operators record where institutions stood, not what they did. The national exercise shows that such measures can confound local density with operator type. The Cape records add the recorded quality of the individual school. Among linked successors, conditional on presence and continuity, the school's own historical quality, rather than the quality of its surroundings, predicts modern outcomes.

The next step is a panel that the published reports cannot provide, because their school-level detail ends after 1905. Following the schools measured here through the 1905 School Board Act, the 1953 Bantu Education Act and the 1963 Coloured Persons Education Act would require linking them to education-department and archival records at the dates of transfer and closure. Such a panel would observe what this cross-section cannot: closures and changes in administration, school by school and year by year, and the fate of the mission network beyond the Western Cape schools for which modern mathematics outcomes are available. That design would turn the comparison in this paper into an estimate of the effect of policy. The Cape records already establish the premise such an estimate requires: institutional quality can be measured in the past, it persists, and the channels that carry it were set by policy choices that could have been made differently.

Online Appendix

A Model derivations

This appendix states and proves the formal results summarised in the main text. Section 3 presents the estimand, the two loadings and the intuition. The environment, assumptions, lemma and

propositions are stated in full below.

A.1 Formal statements

Assumption 1 (Founding). *School i opens with organisational capital*

$$k_i = \mu_s + \phi o_i + v_i, \quad (3)$$

where o_i denotes the education orientation of the founding operator, $\phi \geq 0$, and v_i is an idiosyncratic component satisfying $\mathbb{E}[v_i \mid o_i] = 0$. Organisational capital is measured on the 1905 attainment scale: k_i is the pupil-weighted mean Standard that school i would record in 1905 without a regime constraint. It comprises teaching practice, teacher recruitment, internal norms and reputation. These assets belong to the organisation rather than its site.

Assumption 2 (Ceiling). *The funding regime in system s supports instruction and certification up to Standard κ_s . Recorded quality in 1905 is the mean Standard after each pupil's recorded position is limited to κ_s :*

$$q_i = k_i - r_i, \quad r_i \geq 0, \quad (4)$$

where r_i is the attainment above the ceiling that the report cannot record. School attainment distributions are stochastically increasing in k_i . Hence both q_i and r_i are nondecreasing in k_i . The public ceiling is slack: κ_P is at the top of the scale, so $r_i = 0$ and $q_i = k_i$. The mission ceiling binds at approximately Standard IV.

Remark 1 (Censoring and prevented accumulation). Assumption 2 interprets the funding ceiling as censoring: capital accumulated beyond Standard IV, but the report could not record it. The historical rule also admits a second interpretation. The grant may have prevented that capital from accumulating, in which case the records are accurate and mission capital itself was compressed.

The implications differ. Under censoring, $\Lambda^\circ \geq 1$, the per-unit coefficient is inflated and the association per standard deviation falls (Proposition 1). Under prevented accumulation, $q = k$ and the loading is one. The compressed mission dispersion then reflects compressed capital rather than compressed measurement, with no per-unit inflation.

The ceiling experiment in Section 6.3 shows that censoring is sufficient to reproduce the observed cross-system pattern within one system. It cannot determine whether censoring or prevented accumulation operated in the mission system because uncensored mission capital is not observed. The ceiling was also approximate rather than literal: 1.1 per cent of mission pupils and 10.7 per cent of mission schools appear above Standard IV (Table B.3). The alternative measures defined within the Standards a school taught, especially progression, are informative because they are less exposed to either interpretation.

Assumption 3 (Transmission and outcomes). *Between 1905 and the outcome years, school i continues as an organisation with probability $\bar{c}_s$, indicated by $C_i \in \{0, 1\}$. The baseline assumption is that continuity is mean-independent of capital, operator orientation and location: $\mathbb{E}[C_i \mid$*

$k_i, o_i, \bar{q}_{\ell(i)}, \xi_i] = \bar{c}_s$ conditional on the control set. Appendix A.5 signs the consequences of relaxing this condition.

A contemporary school j whose nearest historical school in system s is $a(j)$ has outcome

$$Y_j = \alpha_s + \tau C_{a(j)} k_{a(j)} + \pi \bar{q}_{\ell(j)} + \mathbf{X}_j' \theta + \alpha_{d(j)} + \eta_{e(j)} + \nu_{\ell(j)} + \varepsilon_j, \quad (5)$$

where $\bar{q}_{\ell(j)}$ is the average measured quality in 1905 of historical schools around location $\ell(j)$. The parameter $\tau \geq 0$ captures transmission through the surviving organisation. The parameter π captures transmission through place: community demand for schooling and parental human capital accumulated through local historical schooling and persisting whether or not any single organisation survived.

The disturbance has two components. The location component ν_ℓ contains outcome-relevant unobservables of the place on which historical placement selected. Its projection on $\bar{q}_\ell$, net of the control set, defines ζ_s in Assumption 4. The idiosyncratic component satisfies $\mathbb{E}[\varepsilon_j \mid k, o, C, \mathbf{X}, \alpha_d, \eta_e, \nu_\ell] = 0$.

Equation (5) assigns the surviving capital of the nearest historical school to every contemporary school nearby. This is an exposure assumption: the surviving organisation affects schools in its catchment, and proximity proxies for catchment membership. Where that proxy fails, declining coefficients under smoothing distinguish highly localised persistence from broader persistence. They do not distinguish organisational continuity from fine-scale site effects. The microfoundation in Appendix A.2 covers inheritance within a continuing organisation; its catchment reach is assumed rather than derived.

Separating these two disturbance components is necessary. A single disturbance required to be mean-independent of k would impose $\zeta_s = 0$ and exclude the compensatory placement allowed here. The research design assumes that the controls and fixed effects absorb ν_ℓ except for the ζ_s term included explicitly in the model. The model states this identification condition; it does not derive it.

Assumption 4 (Observation). The econometrician does not observe $q_{a(j)}$ directly. The observed object is the interpolated surface

$$\hat{Q}_j(p) = \omega(p) q_{a(j)} + (1 - \omega(p)) \bar{q}_{\ell(j)} + m_j, \quad (6)$$

where p is the inverse-distance power, $\omega(p) \in (0, 1)$ is increasing in p , and m_j is interpolation noise. A high power places more weight on the nearest historical school. A low power averages measured quality over a wider area.

Historical placement was not random. Let ζ_s denote the coefficient from projecting the outcome-relevant unobserved component of a location on $\bar{q}_\ell$, net of the control set. Compensatory placement, in which schools from system s concentrated in otherwise disadvantaged places, corresponds to $\zeta_s \leq 0$.

Lemma 1 (Censoring compresses the measure and inflates the slope). *Under Assumption 2, suppose q and r are nondecreasing functions of a common scalar index and are therefore comonotone. Suppose also that $\text{Var}(q) > 0$. Then (i) $\text{Var}(q) \leq \text{Var}(k)$; (ii) $\Lambda \equiv \text{Cov}(k, q) / \text{Var}(q) \geq 1$; (iii) $\text{Cov}(k, q) / \sigma_q = \text{Corr}(k, q) \sigma_k \leq \sigma_k$.*

The proof in Appendix A.3 uses only the association inequality for monotone functions of a common random variable. It requires no distributional assumption. Comonotonicity is essential: monotonicity of the conditional means $\mathbb{E}[q | k]$ and $\mathbb{E}[r | k]$ alone is insufficient, and both (i) and (ii) can fail. The same ordering condition must hold after the residualisation used to construct the deviations ξ below, because the estimand loads on those deviations. Appendix A.3 states this condition separately.

Proposition 1 (Compression without irrelevance). *Suppose the two systems share $(\tau, \pi, \bar{c}, \sigma_k)$ and differ only in their ceilings, with κ_M binding and κ_P slack. Relative to the public system, the mission system then has: (i) a smaller measured-quality dispersion, $\sigma_{q,M} \leq \sigma_{q,P} = \sigma_k$, with strict inequality when the censored component varies across schools ($\text{Var}(r) > 0$); (ii) a weakly larger per-unit coefficient, $A_M \geq A_P$, with $A_M/A_P = \Lambda_M^\circ \geq 1$ whenever $\tau\bar{c} > 0$; (iii) a weakly smaller association per standard deviation of the regressor on which the estimand loads, $A_M \sigma_{\xi,M} = \tau\bar{c} \text{Corr}(k, \xi) \sigma_k^\circ \leq A_P \sigma_{\xi,P}$, where σ_k° is the dispersion of k after the same residualisation; and (iv) a larger sampling variance of the estimated gradient under homoskedastic independent residuals of common variance and a common sample.*

A binding ceiling therefore reduces the variation available to persist and the precision with which its legacy can be recovered, without changing the transmission parameters. Parts (ii) and (iii) refer throughout to deviations. It would be invalid to combine the deviation-level loading Λ° with the total dispersion σ_q : they are computed on different objects, and their product is unbounded.

Proposition 2 (Smoothing re-weights the estimand toward place). *Under equation (1), as p falls, the organisational contribution to the numerator of $b_s(p)$, $\omega(p)A_s\sigma_{\xi,s}^2$, declines. Its share relative to the smooth contribution $S_s\sigma_{q,s}^2$ also declines. The coefficient $b_s(p)$ approaches a limit with the sign of S_s . A sign change along the trajectory requires A_s and S_s to have opposite signs and occurs when $\omega^* = -S_s\sigma_{q,s}^2/(A_s\sigma_{\xi,s}^2)$ lies within the range of powers considered. Because the numerator alone determines the sign of $b_s(p)$, this result does not depend on how interpolation noise varies with p .*

The multiplier on A_s in $b_s(p)$ is not monotone in ω . It rises and then falls, reaching its maximum where $\omega^2\sigma_{\xi,s}^2 = \sigma_{q,s}^2 + \sigma_m^2$. Smoothing therefore changes the relative weight of the two channels, not the absolute multiplier.

The association per standard deviation of the interpolated treatment is monotone increasing in $\omega(p)$ whenever $S_s \leq 0 < A_s$. It is monotone decreasing in $\omega(p)$ when S_s is positive and sufficiently large relative to A_s , provided $\sigma_m^2(p)$ is held fixed. Appendix A.4 gives the additional term when noise varies with the power. If σ_m^2 is nonincreasing in ω , that term is non-negative wherever the covariance numerator is non-negative. The increasing result for $S_s \leq 0 < A_s$ therefore extends to

variable noise on that segment. The same condition does not preserve the decreasing result for a large positive S_s , which requires fixed noise or the exact derivative bound stated there.

The trajectory of $b_s(p)$ across interpolation powers therefore orders the two channels without identifying either. Like equation (1), this proposition uses the orthogonal approximation $\text{Cov}(\xi, \bar{q}) = 0$.

Proposition 3 (Own quality and neighbourhood quality separate the channels in the linked sample). *For historical schools linked by name and location to a modern successor, the analysis maintains that the link identifies an organisational successor. Thus $C = 1$ in the linked sample. This is an assumption about the crosswalk, not a fact established by it. Own q is observed without interpolation.*

Let A_M^L and S_M^L denote the linked-sample loadings. They do not contain the $\bar{c}$ dilution because continuity is maintained (Appendix A.5). Suppose the variances of the local average and the own deviation are positive. Write the leave-own-out neighbourhood surface as $\bar{q}_\ell + u$, where u is uncorrelated with $\bar{q}_\ell$ and the outcome residual, and let $c_u = \text{Cov}(u, \xi_a)$. Maintain (i) $c_u \leq 0$ and (ii) $\sigma_{\bar{q}}^2 + c_u \geq 0$.

In the joint regression of the successor outcome on own quality and this surface, the neighbourhood coefficient has the sign of $S_M^L - A_M^L$. The maintained conditions are substantive. The term u averages the deviations of neighbouring schools. A leave-own-out construction gives $c_u < 0$ under spatially independent quality, but fine-scale spatial correlation in school quality can make c_u positive. A sufficiently large positive value reverses the sign result. Conditions (i) and (ii) together keep the projection coefficient used in the decomposition bound between zero and one.

Organisational transmission with a weak place channel ($A_M^L > 0$ and $0 \leq S_M^L \leq A_M^L$) therefore predicts $b_{own} > 0$ and $b_{nbhd} \leq 0$. A pure place channel ($A_M^L = 0 < S_M^L$) predicts positive coefficients on both regressors whenever the leave-own-out surface is measured with error. The observed configuration $b_{own} > 0$ and $b_{nbhd} \leq 0$ requires $A_M^L > 0$. In addition, if $b_{own} > 0$ and the total association is positive, the own-quality share of that association is bounded below by $1 - \max\{b_{nbhd}, 0\}/b_{own}$.

Proposition 4 (Operator orientation nets out place). *Among modern schools near a historical mission, the difference in expected outcomes between missions run by operators with high and low education orientation is $\tau \bar{c} \phi(o_H - o_L)$. This result requires operator assignment to locations to be unrelated to location unobservables and local historical quality, conditional on the control set. Under that condition, the place terms are common across orientation groups and cancel.*

A positive high-minus-low contrast then indicates an organisational channel even when the place channel cannot be measured. The condition is substantive. If the orientation groups are not compared across comparable places, the place terms do not cancel and the contrast is not informative about ϕ . Section 7.2 and Appendix D show that the atlas data do not provide such a comparison: proximity to a high-orientation society also proxies for mission density. The proposition therefore has no successful test in this paper.

Throughout the remainder of this appendix, every random variable is the residual from a projection on the controls, historical-division fixed effects and ethnic-group fixed effects. Assumptions

1–4 apply to these residualised objects.

Two orthogonality conditions recur. First, $\bar{q}_\ell$ is defined as the linear projection of own quality on the location smooth. The decomposition $q_{a(j)} = \bar{q}_{\ell(j)} + \xi_{a(j)}$ is therefore orthogonal by construction: $\text{Cov}(\xi, \bar{q}) = 0$. This is a definition, not a consequence of Assumption 4.

The realised inverse-distance surface is not exactly that projection. Let the weights w_i sum to one and let the nearest school a receive the largest weight. The surface then implies $\text{Cov}(\xi, \bar{q}) = (w_a - \sum_i w_i^2) \sigma_q^2 \geq 0$, with strict inequality unless all weight is placed on a . If $c = \text{Cov}(\xi, \bar{q})$, retaining this term adds $(A_s + \omega(p) S_s) c$ to the numerator of equation (1) and $2\omega(p) c$ to its denominator. The direction of the change depends on the signs and relative magnitudes of A_s , S_s and $\omega(p)$ and is not generally signed. Equation (1) and Propositions 1 and 2 therefore use the orthogonal case $c = 0$. The projection definition delivers that case exactly; the realised surface only approximates it.

The leave-own-out surface in the linked sample has the opposite construction. Under spatially independent quality it implies $-\sum_{i \neq a} w_i^2 \sigma_q^2 < 0$. Fine-scale spatial correlation can nonetheless make the corresponding covariance positive. Proposition 3 therefore treats its covariance conditions as maintained assumptions.

Second, placement selection operates at the location level, not among individual schools within a location. For the location component of the disturbance, $\text{Cov}(\nu, \bar{q}) = \zeta_s \text{Var}(\bar{q})$ and $\text{Cov}(\nu, \xi) = 0$. The idiosyncratic component ε is uncorrelated with both. Interpolation noise m_j is uncorrelated with all other objects, and its variance $\sigma_m^2(p)$ depends on the interpolation power.

A.2 Microfoundation of organisational transmission

A school that continues as an organisation inherits capital k_1 and chooses current capital k_2 to maximise

$$(r_0 + r_1 h_\ell) k_2 - \frac{\psi}{2} (k_2 - (1 - \delta) k_1)^2 - \frac{\chi_0}{2} (k_2 - \bar{k}_s)^2, \quad (7)$$

where $r_0 + r_1 h_\ell$ is the marginal return to reputation and teaching quality in enrolment and fees. It increases with place capital h_ℓ , or community demand. The parameter ψ captures the cost advantage of reproducing inherited practice rather than building new practice, with δ denoting depreciation. The parameter χ_0 measures regulatory pressure towards the system norm $\bar{k}_s$.

The first-order condition gives

$$k_2 = \frac{(r_0 + r_1 h_\ell) + \psi(1 - \delta)k_1 + \chi_0 \bar{k}_s}{\psi + \chi_0}, \quad \frac{\partial k_2}{\partial k_1} = \frac{\psi(1 - \delta)}{\psi + \chi_0} \equiv \varrho \in (0, 1). \quad (8)$$

Here $\psi > 0$, $\chi_0 > 0$ and $\delta \in [0, 1)$, which imply $\varrho \in (0, 1)$. Inherited capital persists at rate ϱ . This rate increases with the reproduction advantage ψ and decreases with depreciation and regulatory anchoring. The level term increases with h_ℓ : maintaining a good school yields a higher return where community demand is stronger.

A successor founded after closure is a new organisation with no inherited practice to reproduce.

Its problem therefore contains no reproduction-cost term. It maximises $(r_0 + r_1 h_\ell)k_2 - \frac{\chi_0}{2}(k_2 - \bar{k}_s)^2$, which gives

$$k_2 = \bar{k}_s + \frac{r_0}{\chi_0} + \lambda h_\ell, \quad \lambda = \frac{r_1}{\chi_0}, \quad (9)$$

with the constant absorbed into the system intercept.

Setting $k_1 = 0$ in equation (8) would not yield this expression: it would retain the coefficient $\chi_0/(\psi + \chi_0) < 1$ on $\bar{k}_s$. A new entrant faces no adjustment cost relative to inherited practice; it does not inherit capital equal to zero.

When modern outcomes load on current capital, transmission through a surviving organisation is $\tau = \varrho \times$ the outcome loading on capital. Closure removes the dependence on k_1 , as Assumption 3 imposes. The two dates represent a century of adjustment, so ϱ is a reduced-form persistence rate rather than a one-period parameter. The parameters ψ , δ , χ_0 , r_0 , r_1 and λ appear only in this subsection.

A.3 Proof of Lemma 1

Index schools by a scalar random variable T whose realisation orders the within-school attainment distributions by first-order stochastic dominance, as in Assumption 2. Write pupil attainment as $a \sim F_T$ and define

$$k(T) = \mathbb{E}_{F_T}[a], \quad q(T) = \mathbb{E}_{F_T}[\min\{a, \kappa\}], \quad r(T) = \mathbb{E}_{F_T}[(a - \kappa)^+],$$

so that $k = q + r$. Both $\min\{a, \kappa\}$ and $(a - \kappa)^+$ are nondecreasing in a . Under first-order stochastic dominance, their expectations are therefore nondecreasing in T . Chebyshev's association inequality states that $\text{Cov}(f(T), g(T)) \geq 0$ for nondecreasing functions f and g of a common random variable. Hence $\text{Cov}(q, r) \geq 0$.

(i) Since $\text{Var}(k) = \text{Var}(q) + \text{Var}(r) + 2\text{Cov}(q, r) \geq \text{Var}(q)$, part (i) follows. The inequality is strict if the ceiling binds for a set of schools with positive variance in r .

(ii) Since $\text{Cov}(k, q) = \text{Var}(q) + \text{Cov}(r, q) \geq \text{Var}(q)$, $\Lambda = \text{Cov}(k, q)/\text{Var}(q) \geq 1$.

(iii) Finally, $\text{Cov}(k, q)/\sigma_q = \text{Corr}(k, q)\sigma_k \leq \sigma_k$ because a correlation cannot exceed one. $\square$

The argument requires q and r to be comonotone: both must be nondecreasing functions of the common index T . It is not enough for their conditional means to increase in k . Under that weaker condition, the result can fail. Suppose k takes two values and an independent shock u enters as $q = k/2 + u$ and $r = k/2 - u$. Both conditional means increase in k , but $\text{Var}(q) > \text{Var}(k)$ and $\Lambda < 1$. Parts (i) and (ii) are then false.

The estimand loads on the within-location deviations ξ . The ordering condition must therefore hold after residualisation as well; comonotonicity in levels does not itself imply $\Lambda^\circ \geq 1$. We maintain this condition. The ceiling experiment in Appendix A.7 verifies the corresponding surface-level condition for the artificially capped public and poor records: $\text{Cov}(\hat{Q}^c, \hat{Q}^u - \hat{Q}^c) \geq 0$ in the data. The mission system's deviation-level condition remains maintained because mission capital above the ceiling is not observed.

For Proposition 1, with κ_P slack in the public system, $q = k$, $\sigma_{q,P} = \sigma_k$, and $\Lambda_P = \Lambda_P^\circ = 1$. Lemma 1(i), applied to the mission system, gives part (i).

For part (ii), the per-unit slope of the outcome on the deviation regressor is $\tau \bar{c} \text{Cov}(k, \xi) / \text{Var}(\xi) = \tau \bar{c} \Lambda_s^\circ$. The ratio is therefore Λ_M° .

For part (iii), apply Lemma 1(iii) to the same deviations: $A_M \sigma_{\xi,M} = \tau \bar{c} \text{Cov}(k, \xi) / \sigma_\xi = \tau \bar{c} \text{Corr}(k, \xi) \sigma_k^\circ \leq \tau \bar{c} \sigma_k^\circ = A_P \sigma_{\xi,P}$. Both parts concern deviations. Pairing Λ° with the total dispersion σ_q would be invalid because $\Lambda^\circ \sigma_q$ is not bounded by σ_k .

For part (iv), suppose the homoskedastic regression residual has common variance σ_e^2 and the sample is common. The asymptotic variance of the coefficient is $\sigma_e^2 / (n \text{Var}(x))$ for residualised regressor x . Part (i) shows that this regressor has less variance under a binding ceiling. Without equal residual variance, the ordering need not hold; the proposition therefore states a ceteris-paribus comparison. $\square$

A.4 The estimand and the smoothing trajectory

Substitute $k_a = \bar{q}_\ell + \xi_a + r_a$ into equation (5) and define the deviation-level and smooth-level censoring loadings as

$$\Lambda_s^\circ = \frac{\text{Cov}(k_a, \xi_a)}{\text{Var}(\xi_a)}, \quad \bar{\lambda}_s = \frac{\text{Cov}(k_a, \bar{q}_\ell)}{\text{Var}(\bar{q}_\ell)}.$$

Lemma 1(ii) applies to Λ_s° when the ordering condition in Assumption 2 holds for deviations within locations.

The mean-independence condition in Assumption 3 allows the continuity indicator to be replaced by its mean. For $Z \in \{\bar{q}_\ell, \xi_a\}$, the equality $\text{Cov}(Ck_a, Z) = \bar{c} \text{Cov}(k_a, Z)$ requires $\mathbb{E}[C \mid k, \bar{q}, \xi] = \bar{c}$, not only independence of C and k . With $\hat{Q}(p) = \bar{q}_\ell + \omega(p)\xi_a + m$,

$$\begin{aligned} \text{Cov}(Y, \hat{Q}(p)) &= \tau \bar{c} [\bar{\lambda} \sigma_{\bar{q}}^2 + \omega(p) \Lambda^\circ \sigma_\xi^2] + \pi \sigma_{\bar{q}}^2 + \zeta \sigma_{\bar{q}}^2 \\ &= \omega(p) A \sigma_\xi^2 + S \sigma_{\bar{q}}^2, \end{aligned}$$

where $A = \tau \bar{c} \Lambda^\circ$ and $S = \pi + \zeta + \tau \bar{c} \bar{\lambda}$. Also, $\text{Var}(\hat{Q}(p)) = \omega(p)^2 \sigma_\xi^2 + \sigma_{\bar{q}}^2 + \sigma_m^2$. Equation (1) is the ratio of these expressions.

Sign change (Proposition 2(i)). The numerator $N(\omega) = \omega A \sigma_\xi^2 + S \sigma_{\bar{q}}^2$ is linear in ω , with intercept $S \sigma_{\bar{q}}^2$ and slope $A \sigma_\xi^2$. It changes sign on $(0, 1)$ if and only if A and S have opposite signs and $\omega^* = -S \sigma_{\bar{q}}^2 / (A \sigma_\xi^2) \in (0, 1)$. The observed trajectory changes sign when ω^* lies between $\omega(1)$ and $\omega(6)$.

Monotonicity per treatment standard deviation (Proposition 2(ii)). The association per standard deviation of the interpolated treatment is $g(\omega) = N(\omega) / \sqrt{D(\omega)}$, where $D(\omega) = \omega^2 \sigma_\xi^2 + \sigma_{\bar{q}}^2 + \sigma_m^2(\omega)$. The interpolation-noise variance may depend on the power. Differentiation gives

$$F(\omega) = A \sigma_\xi^2 [\sigma_{\bar{q}}^2 + \sigma_m^2(\omega)] - S \omega \sigma_\xi^2 \sigma_{\bar{q}}^2, \quad \text{sign } g'(\omega) = \text{sign} \left[F(\omega) - \frac{1}{2} N(\omega) \frac{d\sigma_m^2}{d\omega} \right].$$

If σ_m^2 is fixed, the second term vanishes and the sign is that of $F(\omega)$. When $S \leq 0 < A$, $F(\omega) > 0$ for every ω , so g is strictly increasing in ω and hence in the interpolation power. When $S > 0$ and $S/A > [\sigma_q^2 + \sigma_m^2]/(\omega_{\min}\sigma_q^2)$ over the range considered, F is negative throughout and g is strictly decreasing in the power.

When σ_m^2 varies, the additional term has the sign of $-N(\omega) d\sigma_m^2/d\omega$. The exact conditions are

$$g'(\omega) > 0 \iff 2F(\omega) > N(\omega) \frac{d\sigma_m^2}{d\omega}, \quad g'(\omega) < 0 \iff 2F(\omega) < N(\omega) \frac{d\sigma_m^2}{d\omega}.$$

If noise is nonincreasing in ω , the additional term is non-negative wherever $N(\omega) \geq 0$. This extends the increasing result ($S \leq 0 < A$) to variable noise on that segment. It works against the decreasing result for a large positive S , which under variable noise requires the second displayed inequality and is therefore stated for fixed noise. Thus Proposition 2(ii)'s increasing claim holds under fixed noise, and under nonincreasing noise on the segment where $N(\omega) \geq 0$. Its decreasing claim holds under fixed noise. $\square$

A.5 The linked sample: probability limits, bound and selection

In the linked sample, $C = 1$ is maintained rather than established. The crosswalk uses continuity in name and location; the derivations below assume that this criterion identifies organisational continuity. If some links are refoundings, mergers or breaks in lineage, the coefficient on own quality partly captures highly local place persistence omitted by the power-two neighbourhood surface. Own measured quality q_a is observed without interpolation.

Write the projection of the outcome on the linked sample's structural components as $Y = A_M^L \xi_a + S_M^L \bar{q}_\ell + e$, where $A_M^L = \tau \Lambda_M^\circ$ and $S_M^L = \pi + \zeta + \tau \bar{\lambda}$. These linked-sample loadings do not contain the $\bar{c}$ dilution because continuity is maintained. The residual e is orthogonal to the regression-relevant objects.

Consider the joint regression of Y on $q_a = \bar{q}_\ell + \xi_a$ and the leave-own-out surface $\widehat{Q}^{-a} = \bar{q}_\ell + u$. The term u is uncorrelated with $\bar{q}_\ell$ and the outcome residual but may be correlated with ξ_a . Let $B = \sigma_q^2$, $X = \sigma_\xi^2$, $U = \sigma_u^2$ and $c_u = \text{Cov}(u, \xi_a)$. Maintain the conditions (i) $c_u \leq 0$ and (ii) $B + c_u \geq 0$ from Proposition 3. Direct calculation gives

$$\Delta = (B + X)(B + U) - (B + c_u)^2 = B(X + U - 2c_u) + XU - c_u^2 > 0$$

and

$$\text{plim } b_{\text{own}} = \frac{A_M^L [B(X - c_u) + XU - c_u^2] + S_M^L B(U - c_u)}{\Delta}, \quad \text{plim } b_{\text{nbhd}} = \frac{(S_M^L - A_M^L) B(X - c_u)}{\Delta}.$$

Setting $c_u = 0$ gives the uncorrelated-noise case.

If $B > 0$, $X > 0$ and $c_u \leq 0$, all bracketed terms are positive: $X - c_u > 0$, $U - c_u \geq 0$, and $XU - c_u^2 \geq 0$ by Cauchy-Schwarz. The neighbourhood coefficient therefore has the sign of $S_M^L - A_M^L$. Because the own regressor is q_a rather than the deviation ξ_a , the neighbourhood surface

corrects the weight on the smooth component; it does not estimate that smooth loading directly.

The observable implications in Proposition 3 follow. Under organisational transmission with a weak place channel ($A_M^L > 0$ and $0 \leq S_M^L \leq A_M^L$), both terms in the numerator of b_{own} are non-negative and the first is positive. Thus $b_{\text{own}} > 0$ and $b_{\text{nbhd}} \leq 0$. Under a pure place channel ($A_M^L = 0 < S_M^L$), both coefficients are strictly positive provided $U > 0$ or $c_u < 0$: own quality proxies for place through its smooth component whenever the leave-own-out surface is measured with error.

Conversely, suppose $b_{\text{own}} > 0$ and $b_{\text{nbhd}} \leq 0$. Then $S_M^L \leq A_M^L$. If $A_M^L = 0$, $b_{\text{nbhd}} \leq 0$ would imply $S_M^L \leq 0$ and therefore $b_{\text{own}} \leq 0$, a contradiction. Hence $A_M^L > 0$. Without condition (i), this sign result can fail. For example, $c_u > \text{Var}(u)$ together with $A_M^L = 0$ and $S_M^L < 0$ produces $b_{\text{own}} > 0$ and $b_{\text{nbhd}} < 0$ through the covariance alone.

Bound. A regression of Y on own quality alone gives $b_{\text{tot}} = b_{\text{own}} + \hat{\gamma} b_{\text{nbhd}}$, where $\hat{\gamma} = \text{Cov}(\hat{Q}^{-a}, q_a) / \text{Var}(q_a) = (\sigma_q^2 + c_u) / (\sigma_q^2 + \sigma_\xi^2)$. This coefficient lies in $[0, 1]$ exactly under maintained conditions (i) and (ii); condition (ii) excludes $\hat{\gamma} < 0$.

If $b_{\text{own}} > 0$ and $b_{\text{tot}} > 0$, the own regressor's share of b_{tot} in this decomposition is at least $1 - \max\{b_{\text{nbhd}}, 0\} / b_{\text{own}}$. To see this, when $b_{\text{nbhd}} \geq 0$ and $t = b_{\text{nbhd}} / b_{\text{own}}$, the share is $1 / (1 + \hat{\gamma}t) \geq 1 - t$. When $b_{\text{nbhd}} < 0$, the share exceeds one.

This is a regression-decomposition bound on the share of the association attributed to own quality, not a point estimate of a structural parameter. The own coefficient itself combines A_M^L with a smooth term that vanishes only as $\sigma_u^2 \rightarrow 0$ or $S_M^L \rightarrow 0$. With the point estimates in Table 6, the bound is degenerate at one because $b_{\text{nbhd}} < 0$. Section 6.6 therefore evaluates the bound at the upper confidence limit of b_{nbhd} .

Selection. Linking conditions on organisational survival to 2018. Under a threshold selection rule in which a school survives when a single increasing index of its quality and outcome-relevant unobservable exceeds a threshold, conditioning on survival induces negative dependence between quality and the unobservable among survivors. Low-quality schools require favourable unobservables to survive. This result follows from the threshold structure; it does not follow from monotonicity alone. Under this structure, b_{own} is attenuated relative to the case without selection.

If twentieth-century policy instead removed the best-matched mission lineages, the dependence reverses and b_{own} overstates transmission. In either case, b_{own} also contains a smooth term when $\sigma_u^2 > 0$. We therefore do not interpret it as a point estimate or strict bound for $\tau\Lambda_M^\circ$. The data cannot distinguish these cases, so the linked-sample evidence is treated as corroborative.

A.6 The orientation contrast

Substituting Assumption 1 into equation (5), the expected outcome near a mission operated by a society with orientation o is

$$\mathbb{E}[Y \mid o] = \alpha_M + \tau\bar{c}(\mu_M + \phi o) + (\pi + \zeta) \mathbb{E}[\bar{q}_\ell \mid o].$$

Suppose operator assignment to locations is unrelated to location unobservables and local quality, conditional on the control set. Then $\mathbb{E}[\bar{q}_\ell \mid o_H] = \mathbb{E}[\bar{q}_\ell \mid o_L]$, and the high-minus-low contrast is $\tau\bar{c}\phi(o_H - o_L)$. The place terms are common and cancel.

Replacing $\mathbb{E}[Ck \mid o]$ with $\bar{c}\mathbb{E}[k \mid o]$ uses the mean-independence condition in Assumption 3, which conditions on o . Independence of C and k alone is insufficient. The contrast between low orientation and no mission equals the common mission place shift plus $\tau\bar{c}(\mu_M + \phi o_L)$, the organisational contribution of a low-orientation operator. It approximates the place channel when that organisational contribution is small. $\square$

A.7 The ceiling experiment

Table 5 constructs, for each historical public school, the measured quality that the 1905 report would have recorded under the mission ceiling. Each pupil’s Standard is capped at Standard IV before the school mean is calculated, and the censored records are interpolated with the same power-4 procedure. Let $\hat{Q}^u$ and $\hat{Q}^c$ denote the uncensored and censored surfaces.

Suppose the outcome loads on the uncensored surface: $Y = \beta_A \hat{Q}^u + e$, with e orthogonal to both surfaces conditional on the controls. Then

$$\text{plim} \frac{\beta_B}{\beta_A} = \frac{\text{Cov}(\hat{Q}^u, \hat{Q}^c)}{\text{Var}(\hat{Q}^c)} = \hat{\Lambda},$$

the surface-level censoring loading after partialling out the controls. Lemma 1 does not automatically apply to the interpolated surfaces because interpolation combines schools. The sign condition $\text{Cov}(\hat{Q}^c, \hat{Q}^u - \hat{Q}^c) \geq 0$ must therefore hold in the data. It does. After partialling on the grade-specific estimation samples, the measured loadings are $\hat{\Lambda}_g = 1.35$ to 1.39 ; the raw loading is 1.41 . These values exceed one if and only if the covariance is positive.

Suppose instead that the outcome loads on certified attainment $\hat{Q}^c$. The predicted ratio is then $\text{Var}(\hat{Q}^u) / \text{Cov}(\hat{Q}^u, \hat{Q}^c)$, which is also at least one and is strictly greater than one under nondegenerate censoring. The two loadings differ in one observable implication. Under capital loading, censoring reduces the association per treatment standard deviation by the factor $\text{Corr}(\hat{Q}^u, \hat{Q}^c)$. Under certified-attainment loading, it weakly increases that association. A decline in the association per standard deviation therefore indicates that outcomes respond to underlying capital rather than only to the certified measure.

These directional results assume non-negative outcome loadings, positive treatment variances and residuals orthogonal to both surfaces. They distinguish the two pure-loading models; mixtures and other measurement-error structures are not separately identified by this comparison. The standard-error result requires the equal-residual-variance condition in Proposition 1(iv). Section 6.3 tests these implications.

B Cape measurement and robustness

B.1 Descriptive statistics

This subsection reports the descriptive tables cited in the main text. They cover the digitised reports (Table B.1), the six recorded quality measures and their overlap (Table B.2), school-level concentration at the mission ceiling (Table B.3), and summary statistics for the estimation samples (Table B.4).

Table B.1: The Cape Education Reports: coverage and georeferencing

Term	Schools listed	Georeferenced	Share georeferenced
Panel A: report coverage and georeferencing, by year			
1881	619	575	92.9%
1896	1,457	1,301	89.3%
1905	1,788	1,639	91.7%
Panel B: georeferenced 1905 schools, by system			
Public		718	
Poor		241	
Public and poor combined		959	
Mission (incl. industrial)		680	
Cape mission stations in the missionary atlases		135	

Notes. School-level information was digitised from the Cape Education Reports for 1881, 1896 and 1905. Listed counts exclude farm, special, evening and boarding schools, which were not georeferenced; the 1896 and 1905 counts cover public, poor, mission and industrial schools. The 1905 georeferenced count uses canonical locations after removing competing duplicate geocodes. The final row reports the number of Cape mission stations recorded in the standard missionary atlases (Beach, 1903; Streit, 1929), which document only a small fraction of the mission schools listed in the education reports.

Table B.2: Historical school-performance measures in 1905

Measure	System	Mean	SD	N	Correlation with mean Standard
Mean Standard	Mission	0.727	0.460	603	1.000
Mean Standard	Public/poor	1.834	0.987	777	1.000
Inspection pass rate	Mission	0.617	0.226	599	0.129
Inspection pass rate	Public/poor	0.766	0.186	773	0.154
Annual progression rate	Mission	0.609	0.223	524	0.105
Annual progression rate	Public/poor	0.795	0.179	678	0.151
Examination readiness (presented/enrolled)	Mission	0.376	0.191	637	0.721
Examination readiness (presented/enrolled)	Public/poor	0.571	0.302	890	0.600
Share in Standard II or above	Mission	0.219	0.137	603	0.958
Share in Standard II or above	Public/poor	0.480	0.207	777	0.925
Share in Standard IV or above	Mission	0.039	0.072	603	0.821
Share in Standard IV or above	Public/poor	0.209	0.176	777	0.950

Notes. Statistics are calculated across the canonical georeferenced historical-school locations used to construct the treatment surfaces. Mean Standard is the pupil-weighted mean position in the Cape Standards distribution, coded from 0 (Sub-Standard) to 8 (Standard Ex). Inspection pass rate is pupils passed divided by pupils presented at the annual inspection; annual progression is the share placed in a higher Standard than at the previous inspection; readiness is pupils presented divided by average enrolment. Correlations use pairwise complete observations.

Table B.3: The highest Standard with enrolled pupils, 1905

Term	Public/poor	Mission
Standard III or below	16.3%	44.9%
Standard IV	23.9%	44.4%
Standard V	19.0%	7.8%
Standard VI	17.6%	2.7%
Standard VII or Ex-Standard	23.2%	0.2%
Any pupil above Standard IV (share of schools)	59.8%	10.7%
Pupils above Standard IV (share of pupils)	17.2%	1.1%
Number of schools	898	664

Notes. Each school is classified by the highest Cape Standard in which the 1905 report records at least one pupil. The universe is all digitised schools assigned to the named report system with positive enrolment, whether or not georeferenced. Distribution rows and the ‘any pupil’ row are shares of schools in the named system with positive enrolment; the ‘pupils above’ row is a share of pupils. Mission schools concentrate at exactly Standard IV, the highest level the mission grant and curriculum ordinarily supported; public and poor schools spread across Standards IV to Ex-Standard.

Table B.4: Summary statistics

Term	Mean	SD	N
Panel A: 2018 systemic test scores, public/poor-system sample			
Grade 3 mathematics	20.64	6.26	992
Grade 6 mathematics	23.25	7.70	958
Grade 9 mathematics	16.99	7.94	429
Grade 3 language	19.81	6.80	992
Grade 6 language	29.83	8.24	958
Grade 9 language	35.97	7.90	429
Panel B: 2018 systemic test scores, mission-system sample			
Grade 3 mathematics	21.13	5.96	912
Grade 6 mathematics	23.78	7.63	881
Grade 9 mathematics	16.93	7.90	413
Grade 3 language	20.25	6.67	912
Grade 6 language	30.25	8.23	881
Grade 9 language	35.92	7.96	413
Panel C: 1905 recorded performance (mean Standard)			
Historical public/poor schools	1.83	0.99	777
Historical mission schools	0.73	0.46	603
Interpolated public/poor treatment (power 4)	1.96	0.72	1,364
Interpolated mission treatment (power 4)	0.56	0.23	1,273
Panel D: national secondary schools, 2021–2023			
Matric pass rate	78.94	14.30	6,777
Bachelor pass rate	37.15	18.60	6,777
Historical mission within 10 km	0.51	0.50	6,777

Notes. Panels A and B describe school-level mean 2018 Western Cape systemic-test scores in the two Cape estimation samples (contemporary schools within 10 km of a georeferenced 1905 school of the named system). Panel C describes the pupil-weighted 1905 mean Standard across historical schools with quality data, and the power-4 inverse-distance-weighted treatment across contemporary schools in each sample. Panel D describes the national sample of secondary schools with matric results in all of 2021–2023.

B.2 The two school systems in 1881

Table B.5: The two school systems in 1881: subjects taught and grant per pupil

System	Schools	Pupils	Higher subjects taught	Higher-subject enrolments per pupil	Grant per pupil (£)
<i>All 1881 schools</i>					
Public	238	54.0	4.40	2.06	1.87
Mission	368	109.5	2.33	0.58	0.60
<i>Linked to a 1905 school</i>					
Public	134	64.5	5.43	2.54	2.22
Mission	186	126.2	2.45	0.56	0.58

Notes. The 1881 report records the subjects a school taught and the grant it received rather than the Standards its pupils reached, so it cannot measure attained level on the 1905 scale. Higher subjects are arithmetic beyond the simple rules, geometry, algebra, higher mathematics, physics, Latin, Greek, English grammar, history and geography. The lower panel keeps schools linked to a 1905 record on the composite key of school type and 1881 identifier, which is verified unique on the 1881 side; the identifier is a crosswalk column carried on the georeferenced 1905 layer, and where several 1905 points carry the same key the first is retained. The link does not track renames, relocations or changes of system between 1881 and 1905, and a school reclassified from poor to public breaks the link by construction, so the lower panel describes schools traceable by this key rather than all survivors. Grant per pupil is the school's grant divided by pupils enrolled and is a school-level analogue of the colony-wide figures quoted in the text.

B.3 Alternative historical quality measures

Table B.6: Alternative historical performance measures and 2018 mathematics performance

Term	Grade 3 mathematics	Grade 6 mathematics	Grade 9 mathematics
Public/poor-school performance			
Mean Standard	0.53	0.78*	1.66*
Standard error	(0.29)	(0.36)	(0.64)
Inspection pass rate	0.30	0.52	1.86**
Standard error	(0.31)	(0.24)	(0.34)
Annual progression rate	0.13	0.38	1.66**
Standard error	(0.30)	(0.25)	(0.32)
Examination readiness (presented/enrolled)	0.06	0.29	1.30**
Standard error	(0.27)	(0.33)	(0.44)
Share in Standard II or above	0.48*	0.77*	1.46
Standard error	(0.22)	(0.33)	(0.81)
Share in Standard IV or above	0.52	0.74*	1.68*
Standard error	(0.29)	(0.31)	(0.62)
Mission-school performance			
Mean Standard	0.40	0.58	0.59
Standard error	(0.21)	(0.31)	(0.68)
Inspection pass rate	0.94**	1.17*	1.30*
Standard error	(0.23)	(0.38)	(0.41)
Annual progression rate	0.92**	0.92*	1.02
Standard error	(0.13)	(0.21)	(0.40)
Examination readiness (presented/enrolled)	0.24	0.48	0.58
Standard error	(0.26)	(0.32)	(0.40)
Share in Standard II or above	0.29	0.33	0.40
Standard error	(0.19)	(0.34)	(0.68)
Share in Standard IV or above	0.37	0.78*	1.26
Standard error	(0.24)	(0.30)	(0.79)
Historical-division fixed effects	Yes	Yes	Yes
Mapped ethnolinguistic-group fixed effects	Yes	Yes	Yes
Geographic and historical controls	Yes	Yes	Yes
Urban in 1900	Yes	Yes	Yes
Maximum historical-school distance	10 km	10 km	10 km
Inference	Division-clustered CR2	Division-clustered CR2	Division-clustered CR2

Notes. Entries report the association with a one-standard-deviation increase in the interpolated historical treatment; standard errors appear below. Each estimate comes from a separate regression using the same sample definition, fixed effects, controls and CR2 inference as Table 4. The performance measures are defined in Table B.2. Standardising the treatment makes the magnitudes comparable across differently scaled measures but does not equalise their reliability. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Inspection pass rates counted per pupil presented and per pupil enrolled, 1905

System	Schools	Share of pupils presented	Pass rate, per presented		Pass rate, per enrolled	
			Mean	(SD)	Mean	(SD)
Mission	664	0.417	0.615	(-)	0.254	(-)
Public/poor	898	0.650	0.764	(-)	0.498	(-)

Notes. The inspection pass rate divides passes by pupils presented for examination. Schools that presented a smaller share of their pupils therefore report a pass rate computed over a more selected group. The final columns divide passes by pupils enrolled, which counts a pupil who was never presented as not passing. School-level means over the 1905 report.

Table B.8: Averaging 1896 and 1905 recorded quality: a measurement diagnostic

Term	Public 1905	Public 1896+1905	Mission 1905	Mission 1896+1905
Grade 3 mathematics	0.27	0.47	0.04	0.52*
standard error	(0.35)	(0.31)	(0.20)	(0.25)
Grade 6 mathematics	0.44	0.76	0.37	0.80**
standard error	(0.47)	(0.45)	(0.29)	(0.32)
Grade 9 mathematics	1.73***	2.09***	0.79	0.96
standard error	(0.43)	(0.51)	(0.55)	(0.62)
Matched historical schools	388	388	365	365

Notes. The power-4 interpolated treatment is built from the schools matched one-to-one between the 1896 and 1905 reports, using the 1905 mean Standard alone and the average of the 1896 and 1905 mean Standard. Because the two columns within each system use the same matched schools and the same contemporary sample, the only change is the measurement: if each year carries independent measurement error, averaging should attenuate it and, other things equal, raise the coefficient or tighten its standard error. Coefficients are per standard deviation of the interpolated treatment, with CR2 standard errors clustered by 1891 division. The matched sample is smaller and more survival-selected than the full 1905 cross-section, so this is a measurement check, not the headline. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The 1905 report provides one noisy measure of school quality. Table B.8 averages that measure with the 1896 value for schools observed in both reports. The matched historical and contemporary samples are held fixed, so only measurement changes.

Averaging increases the coefficient per standard deviation in every cell. The increase is largest in the mission system, where the Grade 3 and Grade 6 estimates move from statistically indistinguishable from zero to significant. This is the pattern produced by attenuation from year-specific measurement error, provided the two reports contain independent year-specific errors around the same underlying quality. It is also consistent with the noisier mission records benefiting most from a second observation. The observations refer to the same schools nine years apart; they are not mechanical duplicates.

The matched sample is smaller and more selected on survival than the full 1905 cross-section. We therefore treat this result as a measurement diagnostic consistent with the headline estimates understating the mission gradient, not as a replacement for those estimates.

B.4 Urbanisation controls

Table B.9 compares the baseline with three alternatives: no urbanisation control, log population density in 1900, and the contemporary urban indicator. For public and poor schools, the baseline estimates, the estimates without an urban control, and those using 1900 density are similar. The contemporary indicator is the exception. It changes the estimates in the direction expected from a control measured between treatment and outcome.

The contemporary indicator records a school’s urban status in 2018. Every school located in an urban cell in 1900 is in an urban place today, but more than half of today’s urban schools were not urban then. The modern indicator therefore reflects a century of urbanisation after the historical records were compiled.

The comparison also shows the consequence of using contemporary geographic controls in persistence regressions. Such controls are often included to hold constant the setting in which the modern outcome is observed. Here, that choice would increase the Grade 3 and Grade 6 coefficients by about 40 and 50 per cent, while reducing their standard errors by two fifths and one quarter. The historical association would therefore appear both larger and more precisely estimated than under the design based on 1900 conditions.

Timing alone does not determine the direction of the difference. The modern indicator may block a place-mediated component of persistence, absorb confounding, or proxy more precisely for persistent place characteristics. This comparison cannot distinguish among those explanations.

We also do not use continuous 1900 density as the baseline. Doing so imposes a log-linear relationship on the coarse HYDE population raster. In the mission sample, this control behaves differently from the other three specifications. We therefore report it as a sensitivity check rather than use it as the preferred control.

Table B.9: The baseline and three urbanisation controls

	Baseline (urban in 1900)	No urbanisation control	Log density 1900	Urban in 2018 (indicator)
<i>Public/poor schools</i>				
Grade 3 mathematics	0.74 (0.40)	0.66 (0.39)	0.67 (0.37)	1.05*** (0.24)
Grade 6 mathematics	1.08* (0.49)	1.07* (0.50)	1.05* (0.51)	1.62*** (0.36)
Grade 9 mathematics	2.31* (0.89)	2.16** (0.81)	2.31** (0.82)	2.64** (0.97)
<i>Mission schools</i>				
Grade 3 mathematics	1.78 (0.94)	2.42* (0.98)	3.05** (1.08)	2.28* (0.92)
Grade 6 mathematics	2.58 (1.37)	2.83* (1.35)	3.80** (1.29)	2.59* (1.31)
Grade 9 mathematics	2.63 (3.02)	2.86 (2.87)	4.56 (3.11)	2.86 (2.88)

Notes. Each cell is a separate regression of the 2018 school mean mathematics score on the power-4 interpolated 1905 mean Standard, within 1891 division and ethnolinguistic-group fixed effects and the predetermined geographic control set. The columns differ only in the urbanisation control. The baseline, the no-control estimates and the 1900 density agree closely for public and poor schools. The 2018 indicator is the exception, where it raises the coefficient and lowers the standard error, the pattern produced by a control that lies between the treatment and the outcome: it records a school's status a century after the treatment is measured. The 1900 indicator precedes both the records and the outcome and is therefore the baseline. The continuous 1900 density imposes a log-linear functional form on a coarse historical population raster, so it is reported beside the baseline as a sensitivity rather than used as the control. Division-clustered CR2 standard errors with Satterthwaite degrees of freedom. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.5 Spatial construction and fixed effects

Table B.10: Specification sensitivity of the Cape performance gradient per treatment standard deviation

Term	Grade 3 mathematics	Grade 6 mathematics	Grade 9 mathematics
Public/poor-school performance			
Baseline: power 4, 10 km	0.53	0.78*	1.66*
Standard error	(0.29)	(0.36)	(0.64)
Tighter match: power 4, 5 km	0.65	1.21*	2.01*
Standard error	(0.35)	(0.54)	(0.87)
Smoother surface: power 1, 10 km	1.31**	1.32*	3.34***
Standard error	(0.50)	(0.59)	(0.77)
Town fixed effects, power 4	0.62	0.93	3.11
Standard error	(0.32)	(0.57)	(1.03)
Mission-school performance			
Baseline: power 4, 10 km	0.40	0.58	0.59
Standard error	(0.21)	(0.31)	(0.68)
Tighter match: power 4, 5 km	0.57*	0.76*	0.59
Standard error	(0.25)	(0.35)	(0.83)
Smoother surface: power 1, 10 km	-0.22	-0.13	0.18
Standard error	(0.27)	(0.36)	(0.59)
Town fixed effects, power 4	1.42***	1.15*	1.83
Standard error	(0.20)	(0.37)	(1.64)

Notes. Entries report the association with a one-standard-deviation increase in the interpolated historical mean Standard; standard errors appear below. All specifications include geographic and historical controls and use division-clustered CR2 inference. Baseline, 5-km and power-1 specifications include historical-division and ethnic-group fixed effects. The final specification adds town fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: The public and mission quality surfaces around Cape Town

Term	Value
Corr(public surface, mission surface)	0.261
within City of Cape Town	0.287
outside City of Cape Town	0.165
Cape Town share of public/poor 10 km sample	0.58
Cape Town share of mission 10 km sample	0.64
Grade 6 public coefficient, full sample	1.08*
standard error	(0.49)
Grade 6 public coefficient, excl. Cape Town	1.44**
standard error	(0.50)
Grade 6 mission coefficient, full sample	2.58
standard error	(1.37)
Grade 6 mission coefficient, excl. Cape Town	3.04
standard error	(2.07)

Notes. The public/poor and mission power-4 quality surfaces are correlated across the schools of the 10 km estimation samples that carry both. A correlation near 0.26 shows the two systems' surfaces are close to distinct; because the systems are estimated separately, this describes the distinctness of the exposures rather than a multicollinearity diagnostic. The City of Cape Town supplies most of each regression sample. The Grade 6 mean-Standard coefficient at power 4 is reported for the full sample and with Cape Town removed; both systems' point estimates keep their sign and order of magnitude. Table B.12 re-estimates the smoothing trajectories without Cape Town for all three grades. Standard errors are CR2, clustered by 1891 division. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The modern outcomes cover Western Cape schools, and both historical systems were concentrated in and around Cape Town. The separate public and mission surfaces could therefore be difficult to distinguish there. If the two quality measures represented the same local exposure, their separation would be spurious.

Table B.11 shows that they remain distinct. The two interpolated surfaces have a correlation of 0.26 in the overlap of the 10-kilometre estimation samples and 0.29 within the City of Cape Town. These figures concern the distinctness of the exposures, because the systems are estimated separately; they are not a multicollinearity diagnostic.

Cape Town accounts for most observations in each estimation sample. Excluding it nevertheless leaves the Grade 6 coefficients for both systems with the same sign and order of magnitude, and the public estimate becomes larger. Table B.12 re-estimates the complete smoothing trajectories without Cape Town. The trajectories continue to move in opposite directions and, if anything, do so more steeply.

Table B.12: Smoothing trajectories excluding the City of Cape Town

Power	Grade 3	Grade 6	Grade 9
<i>Public/poor performance, per treatment SD</i>			
Power 1	3.03*** (0.77)	3.82*** (0.99)	9.66*** (2.10)
Power 2	1.25** (0.52)	1.80** (0.69)	5.71*** (1.40)
Power 4	0.48 (0.54)	0.99 (0.60)	3.52*** (0.94)
Power 6	0.32 (0.53)	0.83 (0.57)	2.82*** (0.81)
<i>Mission performance, per treatment SD</i>			
Power 1	-0.33 (0.41)	-0.37 (0.57)	0.32 (1.13)
Power 2	0.38 (0.40)	0.62 (0.60)	1.77 (1.25)
Power 4	0.55 (0.37)	0.93 (0.63)	2.11 (1.23)
Power 6	0.61 (0.38)	1.03 (0.65)	2.28 (1.26)

Notes. Per-treatment-SD associations at inverse-distance powers 1, 2, 4 and 6, estimated on the common-support sample (within 10 km of both systems) excluding schools in the City of Cape Town, with the harmonised in-R surfaces, full controls, fixed effects and CR2 division-clustered standard errors (in parentheses). The system-specific 10-km samples excluding Cape Town give the same qualitative pattern (docs/v23_noct_trajectories.csv). Sample sizes: Grade 3 mathematics 328; Grade 6 mathematics 316; Grade 9 mathematics 149. Powers, grades and samples were fixed before estimation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.13: Performance gradients per treatment standard deviation across the interpolation power

Term	Grade 3 mathematics	Grade 6 mathematics	Grade 9 mathematics
Public/poor-school performance			
Power 1	1.31**	1.32*	3.34***
Standard error	(0.50)	(0.59)	(0.77)
Power 2	0.58	0.67	1.96***
Standard error	(0.33)	(0.36)	(0.45)
Power 3	0.48	0.67*	1.72**
Standard error	(0.28)	(0.33)	(0.53)
Power 4	0.47	0.75*	1.70**
Standard error	(0.27)	(0.36)	(0.60)
Power 6	0.43	0.84*	1.66*
Standard error	(0.27)	(0.41)	(0.64)
Mission-school performance			
Power 1	-0.22	-0.13	0.18
Standard error	(0.27)	(0.36)	(0.59)
Power 2	0.20	0.40	0.77
Standard error	(0.25)	(0.32)	(0.58)
Power 3	0.31	0.51	0.63
Standard error	(0.22)	(0.31)	(0.57)
Power 4	0.38	0.55	0.57
Standard error	(0.20)	(0.31)	(0.63)
Power 6	0.46*	0.58	0.59
Standard error	(0.20)	(0.31)	(0.72)

Notes. Entries report the association between 2018 mathematics and a one-standard-deviation increase in the interpolated 1905 mean Standard, where the interpolation uses the named inverse-distance power; standard errors, scaled by the same treatment standard deviation, appear below. Lower powers average historical performance over a wider area. All specifications use the headline controls, fixed effects and division-clustered CR2 inference. All surfaces in this table are computed with the same in-sample interpolation engine; its power-4 surface correlates at 0.994 with the precomputed surface used in Table 4, so the power-4 rows differ marginally from the headline estimates. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Tables B.14, B.15 and B.16 provide the common-support and scale diagnostics used in Section 6. Table B.14 re-estimates the headline comparison for schools within 10 kilometres of both systems. It enters each surface separately and then jointly. Table B.15 reports endpoint tests for the two smoothing trajectories and their difference on the same sample. Table B.16 describes the spatial scale of recorded quality in 1905 within each system.

Table B.14: Both systems' 1905 quality surfaces on the common-support sample

Specification	Coefficient	SE	Per SD	(SE)
<i>Grade 3 mathematics</i> (n = 875)				
Public/poor, separate	0.68	(0.47)	0.49	(0.33)
Mission, separate	3.60**	(1.16)	0.78	(0.25)
Public/poor, joint	0.57	(0.47)	0.41	(0.34)
Mission, joint	3.45**	(1.15)	0.75	(0.25)
<i>Grade 6 mathematics</i> (n = 846)				
Public/poor, separate	1.16*	(0.56)	0.83	(0.40)
Mission, separate	3.82**	(1.38)	0.82	(0.30)
Public/poor, joint	1.05	(0.60)	0.75	(0.43)
Mission, joint	3.53*	(1.48)	0.76	(0.32)
<i>Grade 9 mathematics</i> (n = 399)				
Public/poor, separate	2.82**	(0.90)	2.04	(0.65)
Mission, separate	5.41	(3.93)	1.15	(0.83)
Public/poor, joint	2.72**	(0.86)	1.97	(0.62)
Mission, joint	4.88	(4.06)	1.03	(0.86)

Notes. Sample: contemporary schools within 10 km of a georeferenced 1905 school of each system; each grade uses the identical complete-case sample for all four rows. Treatments are the power-4 headline surfaces. Separate rows re-estimate the single-system specification on this common sample; joint rows enter both surfaces in one equation. Full controls, 1891-division and ethnic-group fixed effects, CR2 clustered by division. The two surfaces correlate at 0.26 on this sample (0.29 within the City of Cape Town, 0.17 outside). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.15: Endpoint tests of the smoothing trajectories on the common-support sample

Contrast	Estimate	SE	p	d.f.
<i>Grade 3 mathematics</i> (n = 875)				
Public/poor: power 4 minus power 1	-0.96*	(0.48)	0.083	7.4
Mission: power 4 minus power 1	0.80***	(0.25)	0.009	9.4
Difference-in-differences (public minus mission)	-1.77***	(0.41)	0.002	9.5
Cross-system contrast at power 4 (public minus mission)	-0.37	(0.34)	0.319	5.6
<i>Grade 6 mathematics</i> (n = 846)				
Public/poor: power 4 minus power 1	-0.75	(0.68)	0.305	7.2
Mission: power 4 minus power 1	0.81***	(0.23)	0.006	9.5
Difference-in-differences (public minus mission)	-1.56*	(0.73)	0.061	9.3
Cross-system contrast at power 4 (public minus mission)	-0.01	(0.55)	0.990	5.4
<i>Grade 9 mathematics</i> (n = 399)				
Public/poor: power 4 minus power 1	-1.93	(1.12)	0.136	5.8
Mission: power 4 minus power 1	1.29	(0.75)	0.113	10.6
Difference-in-differences (public minus mission)	-3.22*	(1.55)	0.065	9.4
Cross-system contrast at power 4 (public minus mission)	1.03	(0.64)	0.173	4.6

Notes. Each grade's block stacks two joint two-system equations, at inverse-distance powers 1 and 4, on the identical grade-specific complete-case sample of schools within 10 km of both systems. Treatments are the harmonised in-R surfaces, standardised on that sample, so all quantities are per treatment standard deviation. All controls and fixed effects are interacted with the stack; standard errors and tests use CR2 clustered by 1891 division (clubSandwich HTZ small-sample F with the reported denominator degrees of freedom). These are endpoint contrasts between powers 1 and 4; they do not test monotonicity across the full power grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.16: Spatial structure of recorded 1905 quality, by system

Distance band	Raw	District-demeaned	Pairs
<i>Public/poor schools</i>			
0–2 km	0.059	-0.180	537
2–5 km	0.101	-0.034	448
5–10 km	0.182	0.008	897
10–25 km	0.191	-0.034	2,989
25–50 km	0.181	-0.008	7,517
<i>Mission schools</i>			
0–2 km	0.469	0.132	397
2–5 km	0.222	-0.018	531
5–10 km	0.238	-0.067	1,360
10–25 km	0.191	-0.012	4,559
25–50 km	0.096	-0.021	8,886

Notes. Each entry is the mean standardised cross-product of 1905 mean Standard over unordered pairs of georeferenced schools of the named system whose distance falls in the band (a Moran-style band statistic). The district-demeaned column removes 1905 report-district means first. The statistic is descriptive: schools appear in many pairs, so no sampling inference is attached. Positive values indicate that nearby schools recorded similar quality; the comparison across systems and bands shows at what scale each system’s quality was spatially organised.

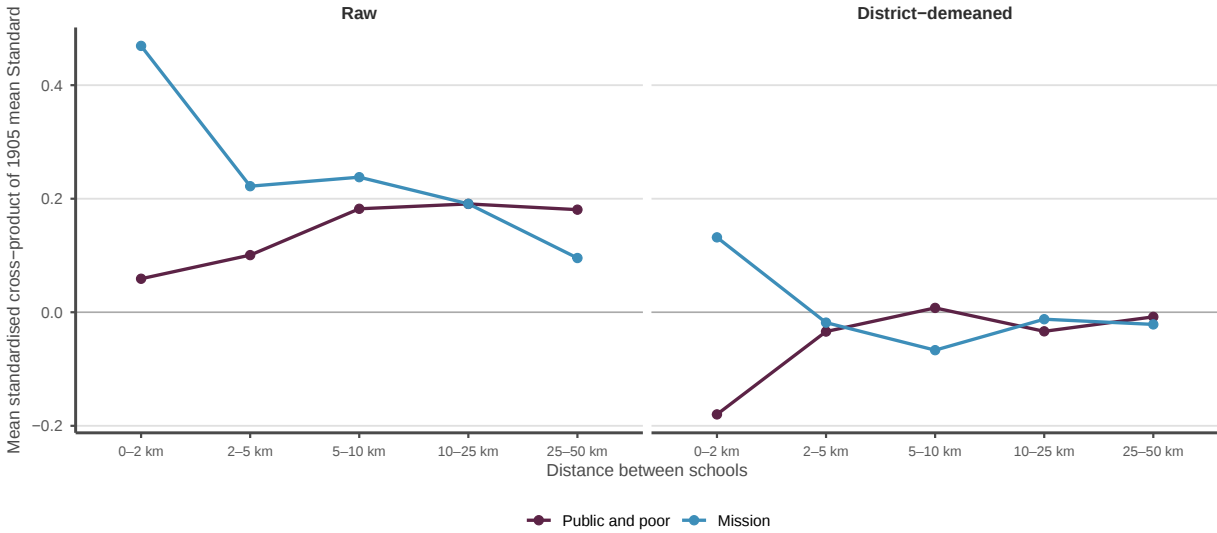


Figure B.1: The spatial scale of recorded quality in 1905, by system. Each point is the Moran-style band statistic reported in Table B.16: the mean standardised cross-product of 1905 mean Standard over unordered pairs of georeferenced schools in the named system whose separation falls within the distance band. The left panel uses the raw measure; the right removes 1905 report-district means. The statistic is descriptive, and the table reports the number of pairs. Mission quality is strongly correlated among schools within two kilometres, with the correlation declining over distance. Part of this fine-scale correlation remains within report districts. For public and poor schools, the correlation is weak at short distances, extends over a broader range, and disappears almost entirely after report-district means are removed. These are the spatial scales reflected in the smoothing trajectories in Figure 6.

Table B.17: Leave-one-out cross-validation of the IDW power parameter (mean Standard, 1905)

IDW power p	LOO-CV RMSE (Public/poor)	LOO-CV RMSE (Mission)
1	1.0072	0.4184
2	1.0358	0.4313
3	1.0694	0.4437
4	1.0972	0.4565
6	1.1301	0.4731

Notes. RMSE is measured in Cape Standard units. Each historical school is predicted from all other georeferenced schools in the named system.

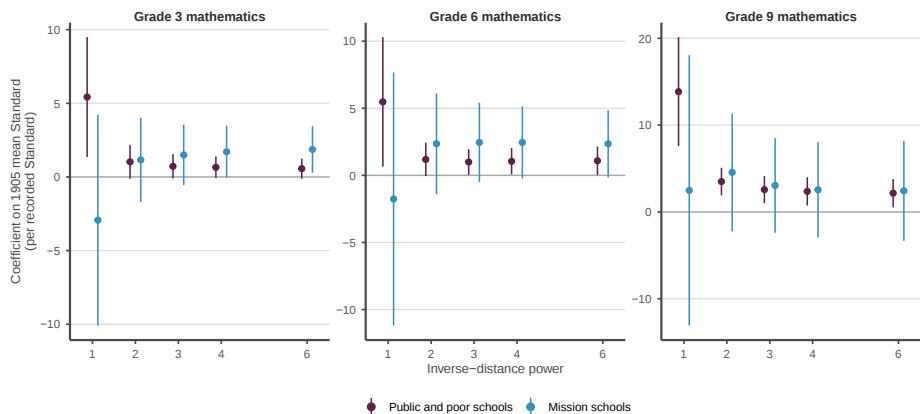


Figure B.2: Cape mean-Standard coefficients across inverse-distance powers. The figure reports the stability of the coefficients. Table B.17 reports leave-one-out prediction error for historical school quality. Neither exercise identifies the causal relevance of a modern exposure surface.

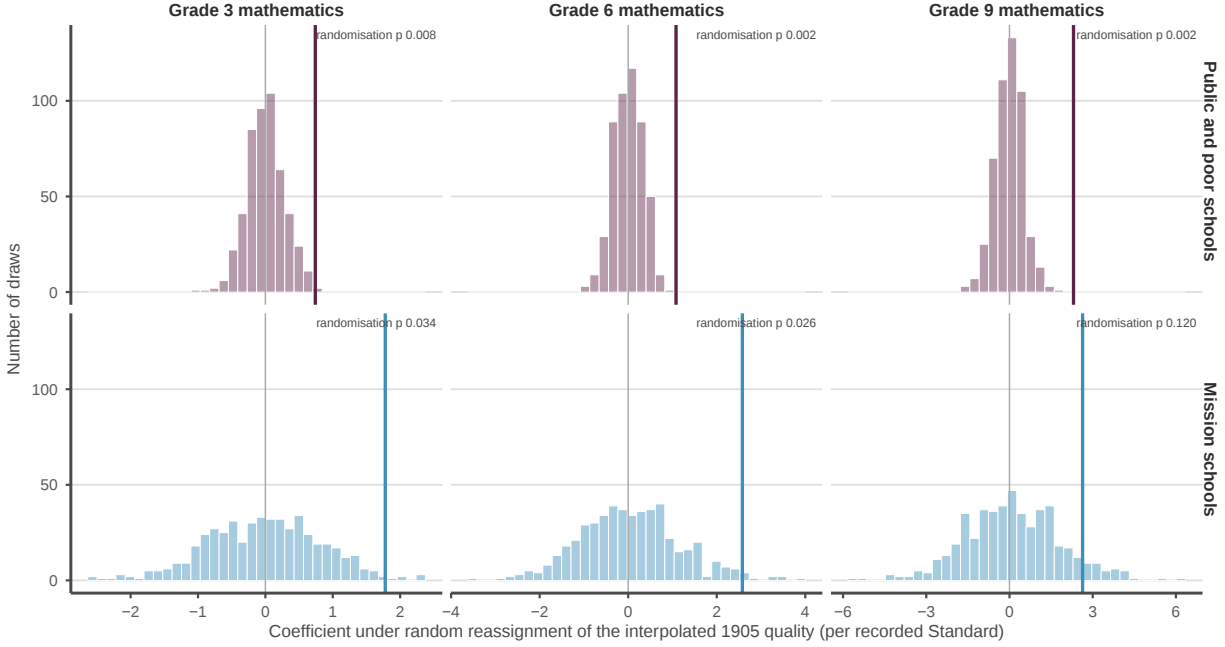


Figure B.3: Permutation placebo for both systems and all three grades. Each panel shows coefficients from 500 random reassignments of the named system’s interpolated 1905 quality across contemporary school locations, with the baseline specification re-estimated after each reassignment. The coloured vertical line marks the estimate under the observed assignment. Randomisation p -values count the observed assignment as one draw: one plus the number of random draws at least as large in absolute value, divided by 501. The smallest attainable value is therefore 0.002. The exercise rejects unrestricted reassignment, but the exchangeability assumption does not remove spatially structured omitted variables.

Figure B.3 and Table B.18 report the permutation placebo for both systems under the baseline specification. Its comparison with the division-clustered p -values in the headline table is informative but not reassuring.

Permuting quality across locations removes spatial structure. The resulting null distribution is therefore narrower than the sampling distribution represented by the clustered standard errors. The placebo shows that the observed assignment of quality to locations differs from an unrestricted reshuffling. Because that reshuffling also removes spatial dependence, the result is weaker than rejection of a substantive null of no relationship. It does not establish the precision of any individual estimate. We interpret the exercise no more strongly than its caption does.

Table B.18: Spatial permutation placebo, both systems

Outcome	Coef.	(SE)	CR2 p	Randomisation p	Draws
<i>Public and poor schools</i>					
Grade 3 mathematics	0.74	(0.40)	0.115	0.008	500
Grade 6 mathematics	1.08	(0.49)	0.072	0.002	500
Grade 9 mathematics	2.31	(0.89)	0.053	0.002	500
<i>Mission schools</i>					
Grade 3 mathematics	1.78	(0.94)	0.109	0.034	500
Grade 6 mathematics	2.58	(1.37)	0.112	0.026	500
Grade 9 mathematics	2.63	(3.02)	0.425	0.120	500

Notes. The interpolated 1905 quality of the named system is permuted across contemporary school locations and the baseline specification re-estimated. The randomisation p -value counts the true assignment as one draw: one plus the number of permutation draws at least as large in absolute value, divided by the number of draws plus one, so its minimum is $1/501$. Permutation removes spatial structure, so the resulting null distribution is narrower than the one the division-clustered standard errors describe; the exercise rejects unrestricted reassignment rather than establishing the precision of any individual estimate.

B.6 Language outcomes

Table B.19: Historical recorded performance and 2018 language performance

Term	Grade 3 language	Grade 6 language	Grade 9 language
Public/poor-school performance			
1905 mean Standard	0.68	0.86	1.74
Standard error	(0.66)	(0.84)	(1.06)
Observations	981	948	427
R^2	0.209	0.212	0.240
Mission-school performance			
1905 mean Standard	1.07	0.34	-1.43
Standard error	(1.32)	(1.93)	(3.39)
Observations	905	874	411
R^2	0.199	0.198	0.237

Notes. Specifications match Table 4 but use the corresponding 2018 language score as the outcome. Standard errors use CR2 clustered by 1891 census division with Satterthwaite degrees of freedom. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.7 Inference and selection diagnostics

Table B.20: Cape headline estimates: inference sensitivity

System	Outcome	Estimate	Conley 25 km SE	Conley 50 km SE	Conley 100 km SE	Division CR2 SE [p; df]	HC1 SE [p]
Public/poor	Grade 3 mathematics	0.74	0.31 [†]	0.10 [†]	0.15 [†]	0.40 [0.115; 5.9]	0.36 [0.041]
Public/poor	Grade 6 mathematics	1.08	0.33 [†]	0.24 [†]	0.16 [†]	0.49 [0.072; 5.9]	0.47 [0.021]
Public/poor	Grade 9 mathematics	2.31	0.52 [†]	0.54 [†]	0.27 [†]	0.89 [0.053; 4.5]	0.85 [0.007]
Mission	Grade 3 mathematics	1.78	1.44 [†]	1.37 [†]	0.86 [†]	0.94 [0.109; 5.7]	1.04 [0.087]
Mission	Grade 6 mathematics	2.58	1.55 [†]	1.63 [†]	1.01 [†]	1.37 [0.112; 5.6]	1.32 [0.051]
Mission	Grade 9 mathematics	2.63	3.42 [†]	— [†]	1.84 [†]	3.02 [0.425; 4.7]	2.42 [0.278]

Notes. Full-controls mean-standard specification and 10 km sample. Conley entries report standard errors from the un-repaired covariance matrix; [†] marks a covariance matrix with a negative eigenvalue, and — marks a negative treatment variance. CR2 clusters by 1891 division; brackets report the p-value and Satterthwaite degrees of freedom. HC1 is heteroskedasticity robust. The main table uses CR2.

Table B.21: Oster selection diagnostics for the Cape performance regressions

System	Outcome	Short coefficient	Full coefficient	Full R^2	R_{max}	Oster δ	$\beta^*(\delta = 1)$	Interpretation
Public/poor	Grade 3 mathematics	0.99	0.74	0.20	0.26	3.70	0.43	
Public/poor	Grade 6 mathematics	1.49	1.08	0.19	0.25	3.86	0.50	
Public/poor	Grade 9 mathematics	3.07	2.31	0.21	0.27	1.41	1.96	
Mission	Grade 3 mathematics	2.70	1.78	0.20	0.26	1.73	0.95	
Mission	Grade 6 mathematics	3.08	2.58	0.17	0.22	4.91	2.10	
Mission	Grade 9 mathematics	3.85	2.63	0.20	0.26	0.77	2.20	

Notes. The short regression includes historical-division fixed effects. The full regression adds ethnic-group fixed effects, the geographic and historical controls. $R_{max} = 1.3R^2$ capped at one. A negative δ is not evidence of robustness: it occurs when the included controls move the coefficient away from zero, contrary to the proportional-selection pattern needed for an interpretable Oster bound. Short and full regressions are estimated on the common complete-case sample of the full specification. The mission rows use the full 10-km mission sample.

The Oster calculations are descriptive sensitivity diagnostics (Oster, 2019). Under the baseline controls, every bias-adjusted $\beta^*(\delta = 1)$ has the same sign as the corresponding estimate. Moreover, δ exceeds one in every cell except mission Grade 9, where it is 0.77. In every other cell, selection on unobservables would therefore need to be at least as strong as selection on the included controls to reduce the coefficients to zero. The mission rows use the full mission sample within 10 kilometres.

C Linked-school construction

Candidate links are generated from normalised historical and modern school names, suburb and town fields, and geodesic distance. Generic school terms are removed before tokens are matched. A candidate must share a distinctive token, lie within two kilometres of the historical school, and map uniquely to a modern school. Manual review removes six false positives based on generic place names.

Confidence tier A requires a distance of no more than one kilometre and a near-exact occurrence of the shared token in the modern name or suburb. The procedure retains 60 links. Seven lack either historical mean Standard or a usable test outcome in 2018, leaving 53 observations. The replication archive contains the crosswalk and the code used to construct the leave-own-out neighbourhood surface.

D The society-orientation exercise and its coding artefact

This appendix reports the national test of Proposition 4 and the diagnostic showing why it does not identify an operator effect.

Table D.1: Missionary-society educational orientation: a proximity design and its coding artefact

Term	Matric pass Province fixed effects	Matric pass District fixed effects	Bachelor pass Province fixed effects	Bachelor pass District fixed effects
Panel A: continuous education-orientation score				
Maximum society score nearby	1.14***	1.23***	2.32***	1.59**
Standard error	(0.26)	(0.45)	(0.45)	(0.73)
Controlling for log mission density	0.34	0.28	0.74**	-0.14
Standard error	(0.24)	(0.29)	(0.32)	(0.43)
Mean society score nearby (for comparison)	0.47	-0.07	0.85	-0.55
Standard error	(0.43)	(0.53)	(0.61)	(0.70)
Observations	3,054	3,052	3,054	3,052
R^2	0.123	0.165	0.118	0.149
Panel B: mutually exclusive society categories				
Near high-education-focus society	3.58***	3.61***	6.95***	6.25***
Standard error	(0.60)	(0.63)	(0.88)	(0.76)
Near low-education-focus society	0.73	0.09	0.68	0.13
Standard error	(0.84)	(0.75)	(0.90)	(0.96)
High-focus minus low-focus	2.85***	3.53***	6.27***	6.11***
Standard error	(0.76)	(0.80)	(1.06)	(1.16)
Near unscored-society mission only	1.18	0.58	2.27*	1.36
Standard error	(1.02)	(0.86)	(1.33)	(1.22)
Observations	6,777	6,777	6,777	6,777
R^2	0.148	0.193	0.143	0.176
Panel C: the contrast and mission density				
High-focus minus low-focus, society-clustered	2.85***	3.53***	6.27***	6.11***
Standard error	(0.88)	(0.72)	(1.19)	(0.97)
Contrast controlling for log mission density	1.13	1.80***	2.67**	2.90***
Standard error	(0.76)	(0.69)	(1.07)	(0.91)
Contrast under nearest-mission assignment (corrected)	0.79	1.46*	2.17*	2.65**
Standard error	(0.87)	(0.83)	(1.30)	(1.19)
Contrast, exactly one scored mission (post hoc)	1.84*	1.70	2.99**	2.58
Standard error	(0.97)	(1.27)	(1.25)	(1.70)
Urban indicator	Yes	Yes	Yes	Yes
Fixed effects	Province	Education district	Province	Education district
Contemporary quintile control	No	No	No	No
Inference	Conley, 50 km; row 1 society-clustered	Conley, 50 km; row 1 society-clustered	Conley, 50 km; row 1 society-clustered	Conley, 50 km; row 1 society-clustered

Notes. The Brown–Martin society score is the first principal component of society-level mean investment in five categories recorded in the worldwide mission data of Beach (1903): primary schools, secondary schools, colleges, kindergartens and industrial schools. The component is standardised over the worldwide distribution of missionary societies, so one unit is one standard deviation of that distribution, not of the South African sample. High-focus societies are those above the median of the South African score distribution weighted by the number of missions each society operated, so that the high- and low-focus groups hold roughly equal numbers of missions (343 and 329 of the 672 missions operated by scored societies). Panel A restricts the sample to contemporary schools within 10 km of a scored historical mission. The primary measure is the maximum society score among those missions, an upper-envelope aggregation in which a low-orientation operator nearby does not dilute a high-orientation one. The maximum-score association shrinks by two-thirds or more after controlling for log mission density and remains significant only for bachelor passes within provinces. Because density may itself transmit orientation, the adjusted rows diagnose sensitivity of the exposure coding rather than estimate a direct or total effect of orientation; the permutation and station-level comparisons below support the coding-artefact interpretation. The mean score, reported for comparison, is insignificant throughout. Panel B uses the full national sample and assigns mutually exclusive high-focus, scored low-focus and unscored-only categories; no mission within 10 km is omitted.

Present-day schools are not classified as mission schools. The high-minus-low row is a linear contrast within the same regression. Panel C reports the same contrast under four departures and the diagnostic that motivates them. The first re-estimates Panel B unchanged but clusters on the society of the nearest scored mission, with schools having no scored mission within 10 km treated as their own clusters, and takes critical values from $t(G - 1)$ on the 32 society clusters rather than the normal. The second adds the log number of all missions within 10 km. The third replaces the high-over-low precedence rule with a symmetric assignment that classifies each school by the orientation of its nearest scored mission; this is the corrected estimate. The fourth restricts to schools with exactly one scored mission within 10 km, where the operator is unambiguous; it is post hoc, discards half the exposed sample, and leaves groups that differ in province composition, distance to the mission and contemporary quintile, so it is a diagnostic rather than a preferred specification. The imbalance that motivates these departures is direct. Because the high-focus category indicates any high-focus mission within 10 km, it equals one more often where missions are dense: 6.15 missions within 10 km against 1.71 for the low-focus category. Assigning the orientation labels at random across societies reproduces that gap (ratio 3.75 against 3.61 observed), and the symmetric nearest-mission rule removes it (0.99). The gap is therefore a property of the coding rule, not of where high-orientation societies operated; at station level they are in sparser neighbourhoods than low-orientation societies. Standard errors allow spatial correlation within 50 km except where stated. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: Balance of observables by mission-orientation category

Term	Near high-education-focus society	Near low-education-focus society
Schools	2,051	1,003
Missions within 10 km	6.15	1.71
Distance to nearest scored mission (km)	4.20	5.68
Urban share	0.27	0.30
Share in Eastern Cape	0.25	0.09
Share in KwaZulu-Natal	0.29	0.34
Share in Limpopo	0.04	0.24
Mean matric pass rate	82.2	78.6

Notes. Each column describes the contemporary schools the precedence rule of Table D.1 assigns to the high- or low-education-focus category (a school is high-focus if any high-focus mission lies within 10 km, low-focus if a scored low-focus but no high-focus mission does). The two groups are not comparable places: the high-focus category sits amid more than three times as many missions, closer to the nearest one, and in a different provincial mix. Because the place characteristics missions selected on do not difference out across the groups, the high-minus-low outcome contrast is not identified as an operator-quality effect. This balance table takes the place of a formal proposition on operator orientation.

Table D.2 shows the lack of comparability between the orientation groups. Schools assigned to the high-focus category have, on average, 6.15 missions within ten kilometres, compared with 1.71 for the low-focus category. They are closer to the nearest mission, at 4.2 rather than 5.7 kilometres. Their provincial composition also differs: one quarter are in the Eastern Cape, compared with less than one tenth of the low-focus group, while one twenty-fifth are in Limpopo, compared with one quarter.

The place characteristics on which missions selected therefore do not cancel across the two groups. This violates the condition required by Proposition 4, so the high-minus-low contrast does not identify an operator-orientation effect.

Table D.1 reports the full exercise. Panel A regresses outcomes on the continuous society score. The preferred aggregation is the maximum score among scored missions within 10 kilometres rather than the average. This upper-envelope measure prevents a nearby low-orientation operator that built little from diluting a high-orientation operator. It also makes the measure mechanically increasing in the number of nearby missions, which is the problem isolated by the density diagnostic below.

Using the maximum, the association is large and precise. A one-standard-deviation increase in the worldwide score is associated with 1.14 percentage points in the matric pass rate and 2.32 percentage points in the bachelor pass rate. Both estimates are significant at one per cent and remain significant when education-district fixed effects replace province fixed effects. An operator-quality channel would produce this result.

The density diagnostic attenuates it. Adding the log number of nearby missions reduces the matric coefficient by more than two thirds, to 0.34, and removes its significance in every specifi-

cation except the bachelor-pass regression with province fixed effects. Because the maximum over a larger number of missions is mechanically higher where missions are dense, the raw association is mechanically sensitive to density. In principle, local mission density could itself be a channel through which an education-focused society affected later schooling, so the density-adjusted coefficient should not be read as the total effect of orientation. We use it here as a diagnostic of the exposure coding: random relabelling reproduces the density gap, and high-orientation societies did not operate in denser station neighbourhoods. The average score is reported for comparison and is insignificant throughout. With province fixed effects, its matric-pass coefficient is 0.47 with a standard error of 0.43.

The continuous measure therefore leads to the same conclusion as the categorical measure. A strong apparent orientation signal reflects where missions cluster. Panel B uses mutually exclusive categories in the full national sample, with no mission within 10 kilometres as the omitted group. Relative to that group, proximity to a high-focus society predicts a matric pass rate 3.58 percentage points higher and a bachelor pass rate 6.95 points higher. The corresponding low-focus estimates are 0.73 and 0.68 and are not statistically distinguishable from zero. The high-minus-low contrasts are 2.85 and 6.27 percentage points.

Panel C shows why these estimates do not answer the operator-orientation question. The category rule assigns a school to the high-focus group when *any* mission within 10 kilometres belongs to a high-focus society. A dense mission area provides more draws, making the indicator more likely to equal one. The high-focus category consequently has an average of 6.15 missions within 10 kilometres, compared with 1.71 in the low-focus category.

A label permutation tests this mechanism directly. We randomly assign high-focus status across the thirty-nine scored societies while holding the number of high-focus societies at ten. Across 4,999 draws, the mean density ratio is 3.75, with a standard deviation of 0.56 and a central 95 per cent interval from 2.63 to 4.72. The observed ratio is 3.61. Random labels thus reproduce the imbalance, which is a property of the coding rule.

The imbalance is not a feature of missionary placement. At station level, high-orientation societies operated in sparser neighbourhoods than low-orientation societies. They had an average of 4.89 other scored stations within 25 kilometres, compared with 6.53, and the score does not predict station clustering. The score does predict educational activity: high-orientation societies operated 4.93 atlas primary schools per station, compared with 1.24 for low-orientation societies. This is the activity the measure is designed to capture.

The same permutation applied to the coefficient rather than the density gives the same diagnosis. Under the baseline coding, random society labels produce a high-minus-low matric contrast with a mean of 2.25 percentage points, a standard deviation of 0.71, and a central 95 per cent interval from 0.73 to 3.59 across 4,999 draws. The observed contrast of 2.85 is at the 81st percentile of this distribution and lies well within it.

Under symmetric nearest-mission assignment, the permutation distribution has a mean of 0.02 and a standard deviation of 0.93. The observed contrast is 0.79 with a standard error of 0.87, again

within the distribution generated by random labels. The replication archive contains the draws, seeds and summary statistics.

The remaining rows in Panel C report three additional specifications. Clustering by the society of the nearest scored mission and using critical values from $t(G-1)$ for the thirty-two society clusters increases the baseline standard error from 0.76 to 0.88 without changing the point estimate. The level of inference is therefore not the source of the result.

Controlling for the log number of nearby missions reduces the matrix contrast to 1.13. Restricting the sample to the 1,406 schools with exactly one scored mission within 10 kilometres, for which the operator is unambiguous, gives a contrast of 1.84 with a standard error of 0.97. We report this estimate but do not rely on it. The restriction is post hoc and removes half of the exposed sample. The remaining groups also differ in provincial composition, distance to the mission, and contemporary poverty quintile. It therefore replaces one imbalance with others.

The national data do not identify the effect of operator orientation. We retain the exercise because its diagnostic applies more broadly. When a treatment indicator equals one if any unit of a particular type lies within a radius, and one type takes precedence over another, the indicator also measures local unit density. Permuting type labels across units shows how much of an estimated contrast the coding rule generates by itself.

We do not claim that every proximity design in the literature on historical institutions has this problem. We show that this design did, that the diagnostic requires only relabelling, and that we could not find it reported elsewhere.

E National staffing-continuity diagnostic

Table E.1: Society educational orientation and foreign-staff continuity to 1922

Term	Matric pass Province fixed effects	Matric pass District fixed effects	Bachelor pass Province fixed effects
Panel A: nearest staffed-1900 mission category			
Persisted, low focus	2.37***	2.13**	4.06**
Standard error	(0.92)	(0.87)	(1.73)
Persisted, high focus	2.11***	2.46***	5.09***
Standard error	(0.74)	(0.91)	(1.18)
Faded, low focus	3.35***	2.80***	4.54***
Standard error	(1.08)	(0.95)	(1.36)
Faded, high focus	2.68**	2.79**	5.40***
Standard error	(1.09)	(1.19)	(1.97)
Panel B: continuous score and staffing continuity			
Society education score	0.16	0.36	0.95
Standard error	(0.76)	(0.52)	(0.84)
Recorded with foreign staff in 1922	-1.12	-0.26	-0.90
Standard error	(1.00)	(0.82)	(1.55)
Education score $\times$ recorded in 1922	0.15	0.05	0.10
Standard error	(0.58)	(0.32)	(0.90)
Observations (Panel B)	2,022	2,018	2,022
Urban indicator	Yes	Yes	Yes
Fixed effects	Province	Education district	Province
Contemporary quintile control	No	No	No
Inference	Conley, 50 km	Conley, 50 km	Conley, 50 km

Notes. This is a measurement diagnostic, not an identified closure design. The sample universe is Beach missions recorded with foreign staff in 1900 and carrying a Brown–Martin education score. Panel A classifies each contemporary school by the nearest such mission within 10 km; no scored staffed-1900 mission within 10 km is omitted. Persisted means foreign staff are again recorded in 1922; faded means they are not. Panel A coefficients and stars compare each category with the no-nearby-mission baseline; the continuity diagnostic is the ordering across categories, not whether individual categories differ from zero. Panel B restricts to schools within 10 km and interacts the continuous education score with an indicator that the mission is again recorded with foreign staff in 1922. Absence of foreign staff in 1922 may reflect atlas coverage, closure, or transfer to African staff and is therefore not a direct measure of state destruction. Standard errors allow spatial correlation within 50 km. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The society education score used in this paper is stored in a versioned derived workbook distributed with the replication files. The workbook contains one row for each of the sixty societies operating Beach stations in South Africa. It records a score and high-focus indicator for the thirty-nine scored societies; both fields are missing for the remaining twenty-one. Together with station locations, contemporary school outcomes, controls and geographic identifiers, the workbook reproduces every merge, classification and regression reported in the paper.

The worldwide station-level inputs used by Brown and Martin (2026) to construct the principal component remain with that project and are not redistributed here. The component therefore cannot be reconstructed from raw inputs within this archive. At the data owner's request, the released South African mission records exclude variables not used in this paper, including the health fields.

References

- Abad, Leticia Arroyo, and Noel Maurer.** 2021. “History Never Really Says Goodbye: A Critical Review of the Persistence Literature.” *Journal of Historical Political Economy*, 1: 31–68.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson.** 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review*, 91(5): 1369–1401.
- Alesina, Alberto F., Marlon Seror, David Y. Yang, Yang You, and Weihong Zeng.** 2020. “Persistence Despite Revolutions.” NBER Working Paper 27053.
- Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou.** 2021. “Intergenerational Mobility in Africa.” *Econometrica*, 89(1): 1–35.
- Austin, Gareth.** 2008. “The “Reversal of Fortune” Thesis and the Compression of History: Perspectives from African and Comparative Economic History.” *Journal of International Development*, 20(8): 996–1027.
- Beach, Harlan P.** 1903. *A Geography and Atlas of Protestant Missions*. New York:Student Volunteer Movement for Foreign Missions.
- Beach, Harlan P., and Burton St. John.** 1916. *World Statistics of Christian Missions*. New York:Committee of Reference and Counsel of the Foreign Missions Conference of North America.
- Beach, Harlan P., and Charles H. Fahs.** 1925. *World Missionary Atlas*. New York:Institute of Social and Religious Research.
- Becker, Bastian, and Felix Meier zu Selhausen.** 2023. “Women on a Mission: Protestant Legacies of Gender Equality in Africa?” African Economic History Network Working Paper 72.
- Bell, Robert M., and Daniel F. McCaffrey.** 2002. “Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples.” *Survey Methodology*, 28(2): 169–181.
- Brown, Gabriel.** 2025. “The Persistent Effects of Bible Translations in Africa.” SSRN Working Paper 4695488, revised November 2025.
- Brown, Gabriel, and Lisa-Cheree Martin.** 2026. “No Church in the Wild? Mission Stations and Education in Africa.” Work in progress.
- Cape Colony.** 1906. *Report of the Superintendent-General of Education for the Year 1906*. Government Printer, Cape Town.
- Carrillo, Bladimir, Carlos Charris, and Wilman Iglesias.** 2023. “Moved to Poverty? A Legacy of the Apartheid Experiment in South Africa.” *American Economic Journal: Economic Policy*, 15(4): 183–221.

- Conley, Timothy G., and Morgan Kelly.** 2025. "The Standard Errors of Persistence." *Journal of International Economics*, 153: 104027.
- Dennis, James S., Harlan P. Beach, and Charles H. Fahs.** 1911. *World Atlas of Christian Missions*. New York: Student Volunteer Movement for Foreign Missions.
- Fedderke, Johannes W., Raphael de Kadt, and John M. Luiz.** 2000. "Uneducating South Africa: The Failure to Address the 1910–1993 Legacy." *International Review of Education*, 46(3–4): 257–281.
- Fenske, James, Bishnupriya Gupta, and Anwesh Mukhopadhyay.** 2025. "Colonial Persistence." CAGE Working Paper 752, University of Warwick. Forthcoming in *The Routledge Handbook of the Economic History of Colonialism*.
- Fiske, Edward B., and Helen F. Ladd.** 2004. *Elusive Equity: Education Reform in Post-Apartheid South Africa*. Washington, DC: Brookings Institution Press.
- Gallego, Francisco A., and Robert D. Woodberry.** 2010. "Christian Missionaries and Education in Former African Colonies: How Competition Mattered." *Journal of African Economies*, 19(3): 294–329.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales.** 2016. "Long-Term Persistence." *Journal of the European Economic Association*, 14(6): 1401–1436.
- Horrell, Muriel.** 1964. *A Decade of Bantu Education*. South African Institute of Race Relations.
- Hyslop, Jonathan.** 1999. *The Classroom Struggle: Policy and Resistance in South Africa, 1940–1990*. Pietermaritzburg: University of Natal Press.
- Imbens, Guido W., and Michal Kolesár.** 2016. "Robust Standard Errors in Small Samples: Some Practical Advice." *Review of Economics and Statistics*, 98(4): 701–712.
- Jedwab, Remi, Felix Meier Zu Selhausen, and Alexander Moradi.** 2022. "The Economics of Missionary Expansion: Evidence from Africa and Implications for Development." *Journal of Economic Growth*, 27(2): 149–192.
- Klein Goldewijk, Kees, Arthur Beusen, Jonathan Doelman, and Elke Stehfest.** 2017. "Anthropogenic Land Use Estimates for the Holocene – HYDE 3.2." *Earth System Science Data*, 9(2): 927–953.
- Loram, Charles Templeman.** 1917. *The Education of the South African Native*. Longmans, Green, and Co.
- MacKinnon, James G., and Halbert White.** 1985. "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics*, 29(3): 305–325.

- Malherbe, Ernst Gideon.** 1925. *Education in South Africa (1652-1922)*. Juta & Co., Ltd.
- Molteno, Frank.** 1987. “The Historical Foundations of the Schooling of Black South Africans.” In *Apartheid and Education: The Education of Black South Africa*, ed. Peter Kallaway. Ravan Press.
- Nunn, Nathan.** 2020. “The Historical Roots of Economic Development.” *Science*, 367(6485): eaaz9986.
- Okoye, Dozie, and Roland Pongou.** 2024. “Missions, Fertility Transition, and the Reversal of Fortunes: Evidence from Border Discontinuities in the Emirates of Nigeria.” *Journal of Economic Growth*, 29: 333–379.
- Oster, Emily.** 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics*, 37(2): 187–204.
- Spaull, Nicholas.** 2013. “Poverty & Privilege: Primary School Inequality in South Africa.” *International Journal of Educational Development*, 33(5): 436–447.
- Spaull, Nicholas.** 2019. “Equity: A Price Too High to Pay?” In *South African Schooling: The Enigma of Inequality. A Study of the Present Situation and Future Possibilities*, ed. Nicholas Spaull and Jonathan D. Jansen. Springer.
- Spolaore, Enrico, and Romain Wacziarg.** 2013. “How Deep Are the Roots of Economic Development?” *Journal of Economic Literature*, 51(2): 325–369.
- Streit, Karl.** 1913. *Atlas Hierarchicus: Descriptio Geographica et Statistica S. Romanae Ecclesiae*. Paderborn:Sumptibus Typographiae Bonifacianae.
- Streit, Karl.** 1929. *Atlas Hierarchicus: Descriptio Geographica et Statistica Sanctae Romanae Ecclesiae*. . 2 ed., Paderborn:Sumptibus Typographiae Bonifacianae.
- Tabellini, Guido.** 2010. “Culture and Institutions: Economic Development in the Regions of Europe.” *Journal of the European Economic Association*, 8(4): 677–716.
- Valencia Caicedo, Felipe.** 2019. “The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America.” *The Quarterly Journal of Economics*, 134(1): 507–556.
- Van der Berg, Servaas.** 2007. “Apartheid’s Enduring Legacy: Inequalities in Education.” *Journal of African Economies*, 16(5): 849–880.
- Von Fintel, Dieter, and Johan Fourie.** 2019. “The Great Divergence in South Africa: Population and Wealth Dynamics over Two Centuries.” *Journal of Comparative Economics*, 47: 759–773.
- Voth, Hans-Joachim.** 2020. “Persistence — Myth and Mystery.” CEPR Discussion Paper / SSRN Working Paper 3729382.

- Wallerger, Markus, Daniel Archambault, David Auber, Martin Nöllenburg, and Jaakko Peltonen.** 2022. “Edge-Path Bundling: A Less Ambiguous Edge Bundling Approach.” *IEEE Transactions on Visualization and Computer Graphics*, 28(1): 313–323.
- Wantchekon, Leonard, Marko Klačnja, and Natalija Novta.** 2015. “Education and Human Capital Externalities: Evidence from Colonial Benin.” *Quarterly Journal of Economics*, 130(2): 703–757.