Marginalised Place and Academic Attainment: A School-Level Comparison of British Columbia and Alberta

Theodore Young

Note that some graphs/tables/images have been removed in order to comply with copyright restrictions.

MSc Education (Research Design & Methodology), 2025-26

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Theodore Young
Date:	11 November 2026

Abstract

Rural students tend to score lower than their urban counterparts on standardised assessments across national contexts, yet Canadian evidence on this gap is thin, dated, and reliant on settlement size alone to measure rurality. This dissertation offers the first comparative provincial examination of the predictors of the rural–urban attainment gap in Canada, framing its subject as *marginalised place* and treating rurality as a multidimensional construct, operationalised here through settlement size and remoteness. Under a Most Similar Systems Design, hierarchical regression models were estimated on publicly available, school-level provincial assessment results in British Columbia (n = 863 schools, Grade 7) and Alberta (n = 986 schools, Grade 6), with census-linked measures of socioeconomic status, settlement size, remoteness, Indigenous share of the local population, and school type, cluster-robust inference, and replication in secondary cohorts (n = 1,092; n = 583). Parental education, not family income, was the operative socioeconomic predictor in both provinces. Place retained associations that SES did not account for, but differently by system. Remoteness remained significant net of all modelled factors in Alberta but attenuated to non-significance in BC once Indigenous population share and school type were modelled. Settlement size showed a conditional reversal in Alberta, as lower settlement size was associated with higher attainment once remoteness was held constant. Indigenous population share was modestly and robustly associated with lower attainment in both provinces. Attainment differences by school type were particularly pronounced in BC — independent schools scored 1.13 SD above public schools, largely unexplained by measured covariates — whereas Alberta's school-type advantages were smaller and structurally distinct. These findings motivate further work on the interplay of SES, place, and school type, ideally employing student-level measures of family circumstance as well as measures of Indigenous educational attainment developed with, and useful to, Indigenous communities themselves.

Table of Contents

Abstract..... 2

List of Tables.....7

List of Figures.....9

Chapter 1: Introduction.....10

Chapter 2: Literature Review.....13

2.1 Conceptualising Rurality: From Settlement Size to Social Space……..... 13

2.2 The SES Explanation: Evidence and Critique..... 17

2.3 Rural Effects on Attainment and Their Mechanisms: Settlement Size and Remoteness……. 18

2.4 Canada: An Underdeveloped Literature..... 20

2.5 Indigenous Discrimination and Maltreatment in Canada Is Inseparable From Place..... 21

2.6 School Type: Advantage Layered Onto Place……..... 24

2.7 The Second Gap: Provincial Place-Based Attainment Differences..... 25

2.8 Relevant Provincial Context: BC and Alberta..... 26

2.8.1 SES and Economic Context..... 26

2.8.2 Settlement Size and Remoteness..... 27

2.8.3 Indigenous Share: Geographical Distribution and Governance..... 28

2.8.4 School Type..... 28

2.9 Conclusion..... 29

2.10 Research Questions..... 30

Chapter 3: Methods..... 32

3.1 Study Design: Provincial Comparison via Secondary Data..... 33

3.1.1 Provincial Comparison..... 33
3.1.2 Secondary Data Analysis..... 36
3.2 Data Source Selection…….... 37
3.2.1 Attainment..... 37
3.2.2 Socioeconomic Status..... 38
3.2.3 Share of Indigenous Peoples..... 39
3.2.4 Settlement Size, Remoteness, and School Type..... 40
3.2.5 Summary of Data Sources..... 41
3.2.6 Data Sources Not Included..... 43
3.3 Analytic Samples..... 43
3.4 Measures..... 45
3.4.1 Attainment..... 45
3.4.2 Socioeconomic Status (SES)..... 46
3.4.3 Rurality..... 46
3.4.4 Indigenous Share..... 47
3.4.5 School Type..... 47
3.5 Hypotheses..... 47
3.6 Analytic Approach..... 49
3.6.1 Hierarchical Regression Modelling..... 49
3.6.2 Comparison Level..... 51
3.6.3 Assumptions..... 52
3.7 Ethical Considerations..... 53
Chapter 4: Findings..... 55
4.1 Descriptive Overview..... 55
4.2 British Columbia Grade 7: Primary Model Results..... 56
4.3 Alberta Grade 6: Primary Model Results..... 59

4.4 Cross-Provincial Comparison: Primary Pairing..... 62

4.5 The Patterns Hold in the Secondary Cohorts (BC Grade 4 / Alberta Grade 9).... 64

4.6 Do the Associations Vary by Context? (RQ6)..... 64

4.6.1 In Both Provinces, the Remoteness and Indigenous Share Associations Do Not Depend on Each Other (H6)..... 64

4.6.2 In BC, the Dramatic Independent-School Advantage Shrinks With Remoteness but Never Disappears…….…….65

4.6.3 In Alberta, the Separate-School Advantage Widens With Remoteness but Remains Modest……..…… 67

4.6.4 The Independent-School Advantage Does Not Detectably Vary With Parental Education (Exploratory, BC).....67

4.7 Summary of Hypothesis Outcomes..... 68

Chapter 5: Discussion..... 70

5.1 Indigenous Share and Remoteness..... 70

5.2 Settlement Size..... 72

5.3 School Type..... 73

5.4 Socioeconomic Status..... 74

Chapter 6: Conclusion..... 77

References..... 81

Appendix A Outcome Distributions..... 100

Appendix B Residual Diagnostics.....103

Appendix C Variance Inflation Factors..... 110

Appendix D Robustness Checks: Rank Transformation and Wild Cluster Bootstrap..... 113

D.1 Rank-Transformed Outcomes..... 113

D.2 Wild Cluster Bootstrap..... 115

Appendix E Correspondence Confirming Ethics Review Was Not Required..... 117

Appendix F Supplementary Descriptive Statistics..... 118

Appendix G Exurban Classification Sensitivity Analyses..... 120

Appendix H Supplementary Settlement-Size Analyses (Alberta)..... 123

H.1 Decomposition of the Rurality Block..... 123

H.2 Remoteness-Quartile-Restricted Regressions..... 123

H.3 School-Type Sector Profiles..... 124

Appendix I Alternative Block-Order Regressions (British Columbia and Alberta)..... 126

Appendix J Secondary Cohort Results (British Columbia Grade 4 and Alberta Grade 9)....129

Appendix K Supplementary School-Type Analyses..... 132

K.1 Independent-School Presence Across the Remoteness Distribution (British Columbia)..... 132

K.2 School Type × Parental Education (Exploratory, British Columbia)..... 134

K.3 Full Interaction Models (Both Provinces)..... 135

Appendix L R Script..... 138

List of Tables

Table 3.1 Datasets Utilised..... 42

Table 3.2 Sample Flow by Province..... 44

Table 4.1 Descriptive Statistics for the Primary Analytic Samples, Alberta Grade 6 and British Columbia Grade 7……...55

Table 4.2 Hierarchical Regression Predicting School Mean Foundation Skills Assessment (FSA) Score (Standardised Coefficients), British Columbia Grade 7..... 57

Table 4.3 Hierarchical Regression Predicting School Mean Provincial Achievement Test (PAT) Score (Standardised Coefficients), Alberta Grade 6..... 60

Table 4.4 Hypothesis Outcomes for the Primary Pairing, British Columbia Grade 7 and Alberta Grade 6……...69

Table A1 Distributional Statistics for the Composite Outcomes.....102

Table C1 Generalised Variance Inflation Factors for the Final Models, All Four Cohorts..111

Table C2 Generalised Variance Inflation Factors for the Interaction Models..... 112

Table D1 Final Models on Original Versus Rank-Transformed Outcomes..... 114

Table D2 Satterthwaite Versus Wild-Cluster-Bootstrap p-Values, Final Models..... 116

Table F1 Distribution of Schools and Family Income by Settlement Size (Popctr), British Columbia Grade 7 (N = 863).....118

Table F2 Distribution of Schools and Family Income by Settlement Size (Popctr), Alberta Grade 6 (N = 986).....118

Table F3 Family Income by Remoteness Quartile (IR), British Columbia Grade 7 (N = 863).....119

Table F4 Family Income by Remoteness (IR), Alberta Grade 6 (N = 986)..... 119

Table F5 School Type Composition of the Analytic Samples, n (%)..... 119

Table G1 Mean DA Income of Economic Family by Settlement Category Under the Original and Restricted Rural Classifications, British Columbia Grade 7..... 120

Table G2 Primary Model Coefficients With the 40 Exurban Schools Reclassified, British Columbia Grade 7..... 120

Table G3 Sensitivity of the Primary Model to Exurban Reclassification and Exclusion, Alberta Grade 6 121

Table G4 Primary Model Coefficients After Excluding Exurban Schools, Both Provinces……122

Table H1 Rurality Variables Net of SES: Alone Versus Joint Entry, Alberta Grades 6 and 9.. ……123

Table H2 Within-Quartile Regressions, Alberta Grade 6.....124

Table H3 Mean Socioeconomic and Rurality Characteristics by School Type, Alberta Grade 6 …….125

Table I1 Alternative Block-Order Hierarchical Regression, British Columbia Grade 7 (FSA).....127

Table I2 Alternative Block-Order Hierarchical Regression, Alberta Grade 6 (PAT).....128

Table J1 Hierarchical Regression Predicting School Mean FSA Score (Standardised Coefficients), British Columbia Grade 4……...130

Table J2 Hierarchical Regression Predicting School Mean PAT Score (Standardised Coefficients), Alberta Grade 9……..….131

Table K1 Independent-School Share by IR Quartile, British Columbia Grade 7.....133

Table K2 Regional Concentration of Independent Schools by Remoteness Quartile, British Columbia Grade 7……..……133

Table K3 Formal Tests of the Remoteness–School Type Association, British Columbia Grade 7 ……..……....134

Table K4 School Type × Parental Education Interaction Model, British Columbia Grade 7.....134

Table K5 Full Interaction Models (Standardised Coefficients), Both Provinces.....136

Table K6 Public–Separate Attainment Gap by Remoteness Quartile, Alberta Grade 6.....137

List of Figures

Figure 2.1 A Model of Rural Social Space..... 15

Figure 3.1 Map of Canadian Cities and Provinces.....35

Figure 4.1 Mean Foundation Skills Assessment (FSA) Score by School Type and Remoteness Quartile, British Columbia Grade 7……..... 66

Figure A1 Distributions of School-Level Composite Attainment Across the Four Analytic Cohorts……..…… 101

Figure B1 BC Grade 7 Final Model: Core Diagnostic Panel.....104

Figure B2 Alberta Grade 6 Final Model: Core Diagnostic Panel..... 105

Figure B3 BC Grade 7: Partial Residual Plots (DA income, share bachelor's+, popctr)..... 106

Figure B4 Alberta Grade 6: Partial Residual Plots (DA income, share bachelor's+, popctr).... 107

Figure B5 BC Grade 4 Final Model: Core Diagnostic Panel..... 108

Figure B6 Alberta Grade 9 Final Model: Core Diagnostic Panel..... 109

Chapter 1: Introduction

Where students live can be consequential to their educational experiences and outcomes. On the 2022 Programme for International Student Assessment (PISA), students in cities outperformed those in rural areas in 34 of 35 OECD countries (OECD, 2023c). Attainment differences are among the most measurable manifestations of the broader educational inequalities affecting rural communities (Sullivan et al., 2018), and studying their predictors can illuminate structural inequalities extending well beyond measured performance (Corbett, 2014).

This dissertation was prompted by a 2019 OECD finding that “On average across OECD countries, the academic performance gap between rural and urban students disappears completely after accounting for students’ and schools’ socio-economic profile” (Echazarra & Radinger, 2019, p. 20). This finding has significant potential to be influential: PISA is not merely the world's largest assessment programme but a policy instrument in its own right, with documented influence on national education reforms (Pons, 2017; Li et al., 2025). Countries may thus use this finding to attribute place-based educational differences to differences in wealth, and address rural attainment by giving schools and families more money, either directly or by creating more jobs in specific areas.

What is problematic about the influential potential of this finding is that it does not hold universally. Internationally, recent studies have found that rurality’s association with attainment persists independent of SES (Rojas Apaza et al., 2024; Borbely et al., 2024; Cao & Huo, 2026). A pointed example concerns Australia: in Echazarra and Radinger's (2019) analysis, Australia's rural–urban gap loses significance once SES is accounted for — yet recent Australian studies find that location's association with attainment persists net of SES and other controls (Dean et al., 2023; Perales et al., 2023).

Echazarra and Radinger’s (2019) finding also overstates what its evidence can support, given how narrowly rurality was conceptualised. PISA measures rurality solely through settlement size: the population of the settlement where a school sits. However, there is arguably something more ‘rural’ about an isolated town of 20,000 accessible only by seasonal ice road than one on the outskirts of a large metropolitan area; settlement size

measures fail to capture this difference (Corbett, 2014; Puryear & Kettler, 2017; Roberts et al., 2021). Perhaps rural areas are those which, because they are removed from urban centres, rely more on subsistence activities than more market-integrated areas do; the isolated town's subsistence ice fishing could then be considered a 'rural' activity. Or perhaps rurality lies in what education demands of a place's young people: in the isolated town, further education may structurally require leaving — schooling becomes preparation for departure (Corbett, 2007). In either case, the two towns differ not merely in location but in how life is organised and experienced there: the work available, the futures imaginable, and the institutions within reach (Green & Reid, 2021; Reid et al., 2010). To summarise, rural–urban differences cannot be reduced to differences in settlement size, because settlements of the same size can differ in nearly everything else that shapes education.

Aligned with these differences, Shucksmith (2012) posits that inequality exists between places, just as it does between social classes. ***Place*** shapes not merely the material resources available to students but the social practices, aspirational horizons, and institutional environments within which learning occurs (Green & Reid, 2021; Reid et al., 2010). Much like Dean et al. (2023), the present study is animated by Shucksmith's question: "How far is *place* central to the construction, maintenance and representation of social differences and divisions?" (p. 384). In this dissertation, ***marginalised place*** denotes places subject to marginalisation in one or both of two senses: material, in economic and infrastructural disadvantage, and epistemic, in being understudied and/or mismeasured. Rurality is an instance of marginalised place, epistemically, because rural places are increasingly understudied in education research (Cao & Huo, 2026; Corbett & Gereluk, 2020), rendering them what Thier et al. (2021) call 'research deserts'.

The present dissertation focuses on how marginalised place may relate to attainment in Canada, though it aligns itself with a larger body of education research that refuses to reduce the concept of ***place*** (Bæck, 2016; Corbett, 2014; Green & Reid, 2021; Roberts et al., 2021); it is offered, in part, as an invitation for similar work in other national contexts. As a Canadian study, it also focuses on the place-based character of Indigenous marginalisation: reserves and the sites of residential schooling constitute deliberately bounded, marginalised

places, and the historic maltreatment of Indigenous peoples remains inseparable from where it was enacted (Peters, 1998; TRC, 2015).

It is also prudent for a Canadian study concerned with schooling to operate provincially: each Canadian province manages its education system, so implications for how those systems work can only arise at the provincial level. In this dissertation, two provinces — British Columbia and Alberta — are compared, a pairing whose rationale the Literature Review and Methods develop in full. Within each, schooling runs from kindergarten to Grade 12 (roughly ages 5–18), delivered by local school ‘districts’ (in BC) or ‘divisions’ (in Alberta) under provincial education ministries. The grade cohorts whose assessments this study analyses — 4 and 7 in BC, 6 and 9 in Alberta — span students aged roughly 9 to 15.

In Canada especially, research on place and school-aged attainment is barely trodden ground (Corbett, 2014). Accordingly, this study does not pursue causal influence. It aims instead to describe and compare: to establish how each posed dimension of marginalised place is associated with school-level attainment net of SES, and whether place matters beyond what socioeconomic circumstance explains. Within each province, independent and conditional associations are sought. Research questions are developed at the close of Chapter 2 and, where warranted, formalised as directional hypotheses in Chapter 3. Patterns are then compared across the two provinces, candidate mechanisms suggested, and future research signalled.

Chapter 2: Literature Review

The Introduction established this study's subject — rural–urban attainment differences in Canada, compared across British Columbia and Alberta — and its broader concern, marginalised place, in its material and epistemic senses. This Literature Review shows that education research internationally has persistently misrepresented rurality, with empirical consequences: chief among them, the attribution of rural attainment gaps solely to socioeconomic status.

Two further points emerge from the Canadian literature examined in this review. First, a study interested in marginalised place and attainment in Canada must include an Indigenous dimension. Second, Canada's limited literature on place and attainment has rarely looked within provinces, yet claims about Canadian education systems can only be established at the provincial level. Accordingly, the review closes at that level, with contextual literature on this study’s two focus provinces.

2.1 Conceptualising Rurality: From Settlement Size to Social Space

Rurality has been repeatedly marginalised and misrepresented, such that its relationship with attainment has been obscured. Across policy and discourse relating to demography, international assessment, and popular culture (Corbett, 2014), there is a general tendency toward metrocentricity: the assumption that urban life is the norm against which rural life is measured (Bæck, 2015; Fulkerson & Thomas, 2016; Green & Reid, 2021). Metrocentricity is also often entangled with the deficit model of rurality, in which rural schools are deemed deficient relative to a standard designed by and for urban populations, tested through instruments (standardised assessments, credential frameworks) that reflect urban knowledge and aspirations (Corbett, 2014, 2025; Corbett & Gereluk, 2020). Reflecting this present entanglement, measures of rurality are sometimes even defined by what rural areas lack relative to the metropolis (Corbett, 2014; Corbett & Gereluk, 2020; Green & Reid, 2021).

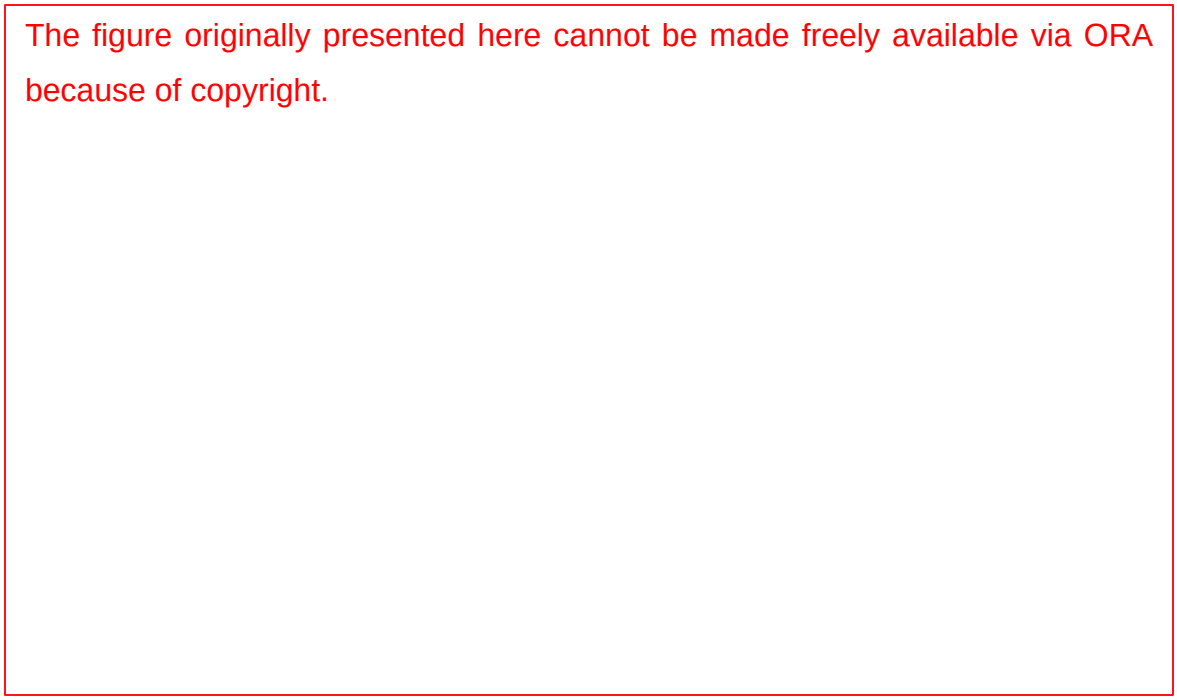
Studies beyond Echazarra and Radinger (2019) tend to equate rurality solely with settlement size, meaning any 'rural' association they report is, by construction, a settlement-size association (Basque & Bouchamma, 2013; Ma & Crocker, 2007; Sullivan et al., 2018); however, considerable literature argues against this practice. Rurality is also conceptualised through remoteness: the degree to which a community is distant from services, population centres, and markets, encompassing both physical distance and connectivity to economic and social infrastructure (Echazarra & Radinger, 2019; Knipe & Bottrell, 2023; UNCTAD, 2021). Measuring rurality solely by settlement size obscures precisely this distance-driven variation (Corbett, 2014; Roberts et al., 2021). This remoteness definition deliberately reads ‘distant from’ rather than ‘lacking’, resisting the definition-by-lack noted above; that said, concepts such as SES and remoteness ‘are never value neutral, especially if they are applied to people and communities’ (Dean et al., 2023, p. 98), a caution this study carries into its own interpretations.

The evident lack of substance in narrow, settlement-size rurality concepts extends further. Space is never neutral (Green & Reid, 2021): a large body of literature argues that rural places are ***complex*** and heterogeneous, encompassing distinct economic structures, geographies, and demographic compositions that vary dramatically across communities and cannot be captured by population alone (Corbett, 2014, 2025; Corbett & Gereluk, 2020; Dean et al., 2023; Green & Reid, 2021; Roberts et al., 2021). Measuring a ‘rural effect’ — the effect that living in a rural place has on attainment — therefore risks obscuring various place-based dynamics if reduced to a settlement-size effect alone.

In establishing the complexity of rurality, many of these studies (e.g., Corbett, 2014; Dean et al., 2023) draw on the Rural Social Space (RSS) model, developed by Reid et al. (2010) and elaborated by Green and Reid (2021) (Figure 2.1).

Figure 2.1

A Model of Rural Social Space



Note. From "Regenerating Rural Social Space? Teacher Education for Rural–Regional Sustainability," by J.-A. Reid, B. Green, M. Cooper, W. Hastings, G. Lock, and S. White, 2010, Australian Journal of Education, 54(3), p. 266 (https://doi.org/10.1177/000494411005400304). Copyright 2010 by SAGE Publications. Used under SAGE's thesis/dissertation reuse policy.

The RSS model provides a theoretical understanding of how different aspects of *place* interact to shape rural educational experience. It is complex in a positive sense: rural social space, like social life anywhere, cannot be reduced to one or two characteristics. The model focuses on the social practices that emerge from the interplay of three dimensions — economy, geography, and demography — in particular places over time. Economy sits at the apex of the model, since it shapes what work is available and what relationship young people understand education to have with their futures; however, any given economy implies particular geographies and demographies, and these interact. A forestry town that develops a tourism industry attracts different people, changes students' aspirational horizons, and transforms the lived experience of teaching there — none of which is visible if place is reduced to a population threshold.

The model also carries an implication for measurement: for any given measure of place, it is rarely clear where the spatial ends and the social begins. Place, on the RSS account, is simultaneously spatial and social: a settlement's size or remoteness is entangled with the labour markets, aspirations, and community histories that accompany it. Any measured association between place and attainment should therefore be expected to carry both spatial and social content, a point that recurs throughout this study.

Hawley et al. (2016), Roberts and Guenther (2021), and Siemens (2023) each emphasise that a rural education study must define what it means by ‘rural’. That said, the recent conceptual and theoretical literature in rural education posits that no simple definition is adequate. This is not a failure of the field but a reflection of the phenomenon of rurality itself: as the RSS model makes clear, rurality is irreducibly ***complex***. Correspondingly, authors stress that no single dimension — settlement size or remoteness — can carry the ‘rural effect’ on its own; capturing this effect requires encompassing several dimensions of rural social life at once (Corbett, 2014; Green & Reid, 2021). ***Distinctly*** measuring dimensions also matters (Roberts et al., 2021): a measure that combines, for example, settlement size and remoteness cannot show which of the two any observed association runs through.

Empirical work has demonstrated that narrowly operationalising rurality can have severe implications for analysis and interpretation. Puryear and Kettler (2017) found that rural fringe districts, ‘rural’ in settlement size but urban-adjacent, more closely resembled urban and suburban districts in educational outcomes than other rural districts — meaning aggregation into a single ‘rural’ category based on settlement size would obscure meaningful internal variation. Cao and Huo (2026) — who measured rurality through settlement-size categories — acknowledged that narrowing rurality in this way is empirically problematic: the lack of detailed geographic characteristics limited their exploration of nuanced rural–urban patterns and their “ability to identify hidden factors related to student achievement” (p. 201). Therefore, this review, and the study it grounds, treats rurality as a complex, multidimensional construct rather than simplifying it to settlement size.

2.2 The SES Explanation: Evidence and Critique

Socioeconomic status (SES) is the most prominent and widely cited explanation for why rural students underperform their urban peers (Marks & O'Connell, 2021). SES is commonly measured through family income, parental education, and parental occupation; the OECD combines the latter two with a home-possessions wealth proxy in its index of Economic, Social and Cultural Status (ESCS), which it measures during its Programme for International Student Assessment (PISA) (OECD, 2023c).

Some researchers have gone further, finding that SES more or less fully explains rural–urban attainment differences. Echazarra and Radinger (2019), in an OECD working paper, found the rural–urban PISA science gap disappeared almost entirely once individual and school SES were controlled. Nissinen et al. (2018) and Jónsson (2019) found capital-versus-regional PISA gaps in Finland and Iceland similarly eliminated by a handful of background controls. Finally, Byun and colleagues (2012) found the US rural–nonrural gap in bachelor's-degree attainment to be largely SES-explained.

Beyond this empirical evidence, this claim has logical force, resting on two well-established links. SES strongly predicts attainment: socioeconomically advantaged students scored on average 93 points higher on the 2022 PISA mathematics assessment, equivalent to about three years of schooling (OECD, 2023c). Rurality, in turn, predicts SES: tertiary-education shares are higher in cities across nearly all OECD countries (OECD, 2023b). Rural Canadians, specifically, earn less, aspire to less education, and experience more persistent low pay than their urban counterparts (Vera-Toscano et al., 2003; Zarifa et al., 2019). If rural places are poorer and poverty depresses attainment, an SES-driven gap is exactly what one would expect.

However, the full-explanation finding faces serious objections. Borbely et al. (2024), engaging Echazarra and Radinger's (2019) claim directly, found the opposite pattern in Scottish primary schools: rural schools showed lower educational mobility for disadvantaged pupils despite being less deprived on average across most domains — the reverse of what the SES explanation predicts. The measurement basis of Echazarra and Radinger's finding is also

concerning: their school-SES control is the mean ESCS of sampled students, and Avvisati (2020) found this aggregate's measurement error is largest precisely in the small rural schools where disadvantage concentrates. Moreover, Nissinen et al.'s models eliminated the gap only after also controlling for students' occupational expectations and motivation. Yet aspirations are precisely one of the channels through which place shapes attainment, both theoretically (Corbett, 2014; Reid et al., 2010) and empirically: net of SES, students' university intentions decline with distance from metropolitan areas (Cooper et al., 2018). The adjustment thus partly explains the gap away by its own mechanism. Finally, the evidence of Byun et al. (2012) concerns bachelor’s degree completion, which is largely achieved beyond where place-based constraints on school-aged learning — staffing shortages, narrow curricula, connectivity — bear most directly (Knipe & Bottrell, 2023).

A clear clue that more than SES is at play comes from comparing contexts: if SES fully explained rurality’s association with education, places with similar rural-urban SES profiles should show similar rural-urban attainment differences — yet they do not. Sullivan et al. (2018) found markedly poorer rural attainment in Australia than in Canada or New Zealand, despite rural communities sharing broadly similar average SES across all three countries. This suggests place-specific factors independently predict rural-urban attainment differences beyond SES. The following section examines this possibility through empirical evidence involving two dimensions of rural place theorised in the literature: settlement size and remoteness.

2.3 Rural Effects on Attainment and Their Mechanisms: Settlement Size and Remoteness

Internationally, considerable recent literature indicates that factors beyond socioeconomic background help explain rural attainment differences (Rojas Apaza et al., 2024; Drescher et al., 2022). Much of this evidence comes from Australia, which has perhaps the most developed national literature on rurality and attainment (Dean et al., 2023; Guenther, 2013; Knipe & Bottrell, 2023; Perales et al., 2023; Sullivan et al., 2018).

Two dimensions of rurality — settlement size and remoteness — have each been found to predict attainment independently of SES. For settlement size, such findings span the OECD (OECD, 2013), England (Cao & Huo, 2026; Midouhas & Flouri, 2015), China (Zhang et al., 2015), and Peru (Castro & Rolleston, 2015; Rojas Apaza et al., 2024), with mixed results in the United States (Drescher et al., 2022; Irvin et al., 2017). For remoteness, the limited evidence is mostly concentrated in Australia: remoteness retained a significant negative association with attainment net of SES, prior attainment, and Indigenous status (Dean et al., 2023), and rural, remote, and regional (RRR) students' attainment was not fully accounted for by socio-demographic composition (Perales et al., 2023).

While independent associations have been found for both dimensions, the proposed mechanisms differ, consistent with their conceptual distinctness. Settlement size operates chiefly through what smaller places can sustain: less-populated schools tend to have narrower curriculum offerings, fewer subject specialists, and higher rates of out-of-field teaching (Echazarra & Radinger, 2019). Larger centres, meanwhile, generate denser, more visible labour markets — making returns to education more observable and raising the social cost of not acquiring credentials (Boone & Van Houtte, 2016) — and higher peer SES and aspirations, both of which influence a student's own aspirations (Dickerson et al., 2018; Echazarra & Radinger, 2019).

Remoteness, by contrast, appears to operate primarily through supply-side mechanisms. The most prominent is teacher quality and retention: qualified teachers systematically avoid or leave remote postings and are hard to replace (Knipe & Bottrell, 2023). Long distances also make professional development prohibitively costly, deepening initial staffing deficits over time (Echazarra & Radinger, 2019). Access to services beyond the school worsens with remoteness (Alasia et al., 2017; Griffith, 1994): not only do university intentions decline with distance net of SES (Cooper et al., 2018), but weaker access to health services, connectivity, and specialist provision may limit attainment through diminished mental and physical supports.

Consistent with the RSS model, place-related disadvantage goes beyond these two rurality measures. In Australia, place-related educational disadvantage also appears to

substantially represent Indigenous disadvantage: the majority of students in very remote schools are Indigenous, and they face layered disadvantage through inadequate staffing, under-resourcing, and policy neglect (Guenther, 2013; Knipe & Bottrell, 2023). Here the spatial disadvantage of remote schooling is entangled with the social marginalisation of Indigenous communities; this interplay is precisely what the RSS model theorises, and provides further reason to measure place multidimensionally.

Whether the settlement-size and remoteness associations documented above — and the mechanisms proposed for each — operate similarly in Canada is difficult to establish from existing evidence: as the following section demonstrates, Canadian research on place-related attainment differences is comparatively scarce. What evidence does exist suggests that the entanglement of place and Indigeneity is, if anything, deeper in the Canadian context.

2.4 Canada: An Underdeveloped Literature

Where the Australian literature on place-based attainment differences is extensive — and Australia's dispersed rural populations, sizable Indigenous communities, and English-colonial history make it a natural reference point for Canada (Capano, 2015; Sullivan et al., 2018) — its Canadian counterpart is thin and offers little explanatory insight. More generally, Canadian work on the contextual and demographic predictors of attainment is underdeveloped: in the first-ever review of such factors, Mayor and Suarez (2019) extended the typical 10-year scoping review boundary to 17 years ‘due to the limited literature’ (p. 46).

Of the 54 articles Mayor and Suarez (2019) identified, only two assessed the rural–urban association with achievement. Seven years later, these remain the two most recent peer-reviewed journal articles this review identified on this topic. Basque and Bouchamma (2013), analysing the mathematics achievement of 2,436 Grade 8 students across 60 New Brunswick francophone schools, found in regression analyses that urban school attendance was a significant predictor, once prior achievement and teaching quality — but not SES — were accounted for. Ma and Crocker (2007), applying hierarchical linear

models to Canada's PISA 2000 sample (29,687 students in 1,117 schools across the ten provinces), with SES accounted for, found urban schools outperformed rural ones in all but three provinces (Alberta, Quebec, and Manitoba).

These articles are limited by more than their age. In both studies, rurality was not the primary focus: each pursued broader school-level questions, with rural/urban entering only as one covariate among many. Ma and Crocker's 100,000-population rural/urban cutoff is additionally problematic: many ‘rural’ schools were likely suburban, which the authors acknowledge may partly explain the rural advantage observed in some provinces.

Reflecting this scarcity of recent, dedicated work, Corbett (2014) argues that 'the field of rural education has not been significantly developed in Canada,' attributing this directly to 'the marginal status of the rural itself' in Canadian educational policy and discourse (p. 1). The present study — which examines how dimensions of rurality, and other aspects of place alongside them, are associated with school-aged attainment — is thus well positioned to contribute to this underdeveloped literature.

2.5 Indigenous Discrimination and Maltreatment in Canada Is Inseparable from Place

Indigenous peoples are arguably the most marginalised group in Canadian history, and their marginalisation is inseparable from place (Scott & Louie, 2020). Perhaps the foremost example of *marginalised place* in Canada is reserves, land set aside by the Canadian government for use by and occupancy of Indigenous peoples (The Canadian Encyclopedia, 2011). The reserve system has never been a neutral administrative category: the colonial administration established reserves as bounded, segregated spaces premised on assimilation; First Nations people were assigned to rural locations while urban areas were defined as hostile to their cultures (Peters, 1998). By 1940, almost all First Nations people lived on reserves, most geographically remote from urban economic opportunity.

The Indian Act further entrenches this spatial marginality: because on-reserve land cannot be mortgaged or transferred to non-band members, it cannot serve as collateral, effectively excluding First Nations communities from private capital markets and making housing and infrastructure investment dependent on federal transfers alone (Lyeo et al.,

2024). Housing conditions themselves have a persistent negative impact on health and well-being (Lyeo et al., 2024), and thereby plausibly on attainment. This legal-structural layer of economic exclusion means that Indigenous disadvantage cannot be collapsed into remoteness, even though reserves were created to be remote: the Indian Act's constraints are connected to reserve status rather than distance, and would operate even on a reserve adjacent to a city.

Empirical evidence further ties Indigenous dimensions of place to attainment. Looker and Bollman (2020) find that rural areas hold roughly double Canada's overall concentration of Indigenous-identifying people, and that dropout rates are substantially higher for Indigenous students — especially on-reserve — such that the rural dropout problem is substantially an Indigenous identity problem rather than a rurality problem per se. Yet Indigeneity is not merely a proxy for the rural or socioeconomic concepts already discussed: net of remoteness, income, and reserve, Treaty, and family characteristics, First Nations people's odds of completing high school remain 60% lower than non-Indigenous peers' (OR = 0.4; Layton, 2023). Nor is the association placeless: Indigenous attainment is higher on reserves near urban centres and declines steeply across remoteness classes (Bleakney & Melvin, 2022; Looker & Bollman, 2020). Moreover, Layton (2023) finds the remoteness–completion association is itself stronger for Indigenous than for non-Indigenous people. Rural and Indigenous dimensions of place are thus entangled but separable — each mattering at given levels of the other.

Indigenous educational disadvantage in Canada also operates through social transmission. First, the intergenerational legacy of residential schools is a deep historical wound. The Indian Residential School System, which ran until 1996, forcibly separated 150,000 Indigenous children from their families with the aim of ‘whitening’ the child (TRC, 2015). While residential school attendance marginally increased human capital accumulation in the first generation (due to forced schooling), it is associated with *lower* educational attainment among subsequent generations (Jones, 2024) — seemingly through disrupted parent-child transmission of educational aspiration. While the trauma of residential schools was location-borne, the associated intergenerational effects are socially transmitted.

Indigeneity is also associated with a range of present-day social disadvantages beyond intergenerational trauma — including experiences of racism and discrimination within educational settings, cultural stigmatisation, and dampened educational aspiration — that independently impede attainment regardless of geographic context (Chandler, 2010; Gordon & White, 2014; Melvin, 2023; Zinga & Gordon, 2016).

Indigeneity is, of course, an identity rather than a location; prejudice attaches to Indigenous persons wherever they live. The claim here is narrower: in Canada, Indigenous marginalisation has been deliberately spatialised. The reserve system converted a people into a geography, confining First Nations communities to bounded and mostly remote spaces. Even the socially transmitted legacies of residential schooling are not spatially neutral, as they travel through families and communities that remain geographically clustered. These legacies remain anchored in place: the schools were physical sites, with 44% located on reserve lands (Indigenous Services Canada, 2024). The removals residential schools brought about were also from particular home places (Truth and Reconciliation Commission of Canada [TRC], 2015), and the resulting trauma is carried collectively in the communities from which children were taken (Bombay et al., 2014). High-Indigenous-concentration communities are therefore marginalised places twice over. First, they are social spaces in which trauma is communally carried. Second, they are physical sites of harm: the home places from which children were taken and, in many cases, the lands on which the schools stood. This study's Indigenous population share (‘Indigenous share’) measure identifies schools in precisely such communities (see Methods).

Recent work measuring Indigenous educational outcomes in Canada exists: British Columbia reports assessment results by Indigenous student share (Education Analytics Office, 2025), while Statistics Canada (Layton, 2023) and Looker and Bollman (2020) compare graduation and dropout rates. However, peer-reviewed literature specifically measuring Indigenous examination attainment is thin and dated (Friesen & Krauth, 2010; Richards, Hove & Afolabi, 2008).

This lack of studies is concerning: addressing such disparities is now explicit national policy, with the Truth and Reconciliation Commission's (TRC, 2015) Calls to Action

demanding the elimination of educational achievement and funding gaps between Indigenous and non-Indigenous Canadians. Establishing these gaps also matters politically: opposition to redistribution toward Indigenous peoples is driven partly by the belief that Indigenous disadvantage reflects individual or cultural failings rather than structural forces (Beauvais, 2022). Finding that Indigenous population share predicts attainment independently of SES, settlement size, and remoteness directly challenges this narrative, and provides empirical grounding for the policy responses the TRC calls for.

2.6 School Type: Advantage Layered onto Place

In this study, ‘school type’ refers to a school's funding and governance category — whether attendance is fee-paying or free, and under whose authority the school operates — rather than to its pedagogical model or the stage of schooling it covers (e.g., elementary, secondary). The present study is interested in marginalised place: specifically, both rural and Indigenous dimensions of place in Canada. School type is not itself a form of marginalised place, but it is of substantial research interest for two reasons. First, it may capture socioeconomic advantage that measured SES variables do not. Second, both the availability of school types and the advantage they confer may vary across places.

Uptake of school choice, where it exists, is not distributed evenly across SES grades. In Canada, Yoon et al. (2018) found that Vancouver's low-income families participate in school choice at lower rates than every higher income bracket, and that when they do choose, only 20% enrol in selective programmes, against 59–68% of every income group above them. Even under a formally universal choice policy, then, uptake of the more academically differentiated school types may skew toward higher-income families. However, the selection at work operates through more than income: educationally engaged families transmit advantage through aspirations, involvement, and cultural capital, which travel with the child into whichever school is chosen (Lareau, 2011).

Selection — not what schools contribute themselves — appears to capture much of school-type advantage. Controlling for student intake and composition in multilevel regression, Dronkers and Robert (2008) found that most of the raw private-school advantage

in reading across the OECD disappeared. Dronkers and Avram (2010) pressed further: arguing that regression controls under-correct for parental self-selection, they applied the stricter test of propensity score matching, under which the private advantage disappeared in most countries. These studies indicate that the better selection into private schools is accounted for, the less advantage remains for the schools to claim. Notably, however, Canada was among the countries where the advantage survived the stricter test, suggesting that school differences may explain some of school-type advantage in the national context of the present study.

Spatially, different school types are rarely distributed evenly across the urban–rural continuum. Alternative school types — private, specialised — tend to cluster disproportionately in larger urban centres that alone can sustain their enrolment, and the highest-performing and most selectively marketed among them are concentrated in particular urban neighbourhoods (Adamson et al., 2016; Ball, 2003; Forsey et al., 2008).

2.7 The Second Gap: Provincial Place-Based Attainment Differences

Because each provincial government in Canada is responsible for its own education system, Canada-wide research struggles to draw system-level insights about what are, institutionally, entirely distinct systems. However, the last peer-reviewed journal article this review identified that measured place-based attainment differences within provinces — aside from New Brunswick (Basque & Bouchamma, 2013) — was Ma and Crocker (2007), nearly 20 years ago. A British Columbia School Trustees Association (BCSTA, 2024) publication does measure rural schooling outcomes within a single province. Yet its aggregated rurality measure obscures which dimension is specifically at work, and its findings directly concern the school districts that fund the organisation; reporting may therefore tilt in whatever direction suits those districts.

Here, a second gap emerges: not only is evidence on place and attainment limited Canada-wide, but there is no recent peer-reviewed research within individual provinces. This gap arguably matters more than the first, as it is province-level evidence that can enable system-level change.

2.8 Relevant Provincial Context: BC and Alberta

To address this provincial gap, the present study compares the socioeconomic and place-based predictors of school-aged attainment within British Columbia (BC) and Alberta, using population-wide samples; the full design rationale is detailed in the Methods section. The following section reviews evidence on how each construct of interest — SES, settlement size, remoteness, Indigenous share, and school type — is situated in each province's economic, geographic, and institutional context.

2.8.1 SES and Economic Context

The channels linking income to attainment depend on what local labour markets reward (Boone & Van Houtte, 2016) — and here the provinces diverge sharply. Alberta's economy is centred on resource extraction. Its Athabasca Oil Sands generate strong demand for both skilled and unskilled labour, sometimes drawing young people out of school and into high-wage, low-skill work (Corbett, 2014): the 2001–2008 oil boom measurably reduced Albertan young men's university enrolment (Morissette et al., 2015; Emery et al., 2012). Such work is both increasingly available — post-pandemic job postings for workers holding only a high school diploma nearly tripled (OECD, 2023a) — and well paid enough to draw fly-in, fly-out labourers from Atlantic Canada (Ferguson, 2011). BC does not mirror this pattern: services account for roughly 75% of BC's GDP but only around 53% of Alberta's, where mining and oil and gas extraction alone contributes 27% in 2022 (BC Stats, 2023; Government of Alberta, 2023); correspondingly, BC projects that roughly three-quarters of jobs in the coming decade will require education beyond high school (Province of British Columbia, 2023). One consequence of this market structure is that income and parental education may be less tightly coupled in Alberta, where resource-sector wages depend less on credentials than service-sector salaries do.

Parental education is a strong predictor of achievement (Tan, 2017), including in Canada (Areepattamannil & Lee, 2014; Christofides et al., 2012; Kaushal, 2014), and it carries dimensions of advantage that income does not — aspirations, cultural capital,

familiarity with academic institutions (Lareau, 2011). If income and education are indeed less tightly coupled in Alberta's labour market, these dimensions may emerge more visibly in Alberta than BC.

No identified study compares income and parental educational capital as distinct predictors of attainment across the two provinces — a further gap this study addresses. Recent Canadian SES–attainment evidence is national rather than provincial (Bacic & Zheng, 2022; Caro et al., 2009), and the nearest provincial studies are single-dimensional: Gill et al. (2025) linked poverty severity to attainment in BC but did not examine parental education.

2.8.2 Settlement Size and Remoteness

Settlement size sits differently in the two provinces' economic geographies. Alberta's urban centres are growing hubs of high-skilled employment while, as discussed above, the Oil Sands offer well-paid low-skill work away from the major cities (Corbett, 2014); these conditions may sharpen the difference between what large and small settlements reward educationally. In BC, by contrast, housing-cost-driven out-migration from Vancouver has dispersed relatively affluent, educated households into smaller communities (McQuillan, 2024). To the extent these households carry not only income but the aspirations and educational investment that accompany it, the advantage gradient across settlement sizes may be flattened, thereby muting what settlement size can add to attainment differences in BC.

The provinces' physical geographies also differ in ways that shape what a given distance means. BC is dissected by mountain ranges and a fjord-cut coastline, with some island and inlet communities reachable only by ferry or air; Alberta is predominantly prairie and parkland, served by a relatively direct road network. The same nominal distance to services can therefore impose greater time and cost in BC, bearing on the access and staffing mechanisms discussed in Section 2.3. Remoteness may thus be more isolating, and more consequential for schooling, in BC than in Alberta.

2.8.3 Indigenous Share: Geographical Distribution and Governance

The provinces' reserve geographies and histories of Indigenous governance differ. British Columbia's reserve system comprises numerous small, scattered, and mostly non-treaty reserves, whereas Alberta's numbered treaties established far fewer, larger ones (Harris, 2002); therefore, Indigenous population share and remoteness may be interwoven differently in the two provinces. BC is also where First Nations have most extensively pursued self-governance, including band-level control over education, health, and child welfare; in Alberta, these domains remain largely administered under the federal Indian Act. This self-determination appears educationally consequential: Chandler (2010) found that in BC bands demonstrating all six measured cultural continuity markers — such as self-government and band-level control over education — timely high school completion reached 48.3%, against 24.3% in bands with none. This advantage is visible in direct interprovincial comparison as well: Richards (2013) found BC's Indigenous K–12 outcomes markedly exceed Alberta's, attributing the difference partly to greater First Nations institutional capacity and more comprehensive performance monitoring.

2.8.4 School Type

The types of schools available and the frequency with which students attend schools of different types vary across BC and Alberta. BC awards full funding only to its secular public system; independent schools — including private and religious schools — may receive only partial provincial subsidy, so are school-fee-charging (Asadolahi et al., 2022; Taylor, 2001). Fees vary dramatically within British Columbia's independent school sector, which accounts for 15% of students: tuition ranges from roughly $1,000 at some faith schools to $20,000 or more at top-ranked preparatory schools (Azimi et al., 2023).

Alberta's system is more complex: Asadolahi et al. (2022) rank Alberta the most school-choice-expansive province. Catholic separate schools are constitutionally protected under the Alberta Act (1905), fully publicly funded, and run by their own elected boards, enrolling roughly a quarter of Alberta students (Hunt & DeJong VanHof, 2024). Charter

schools — first introduced in Canada by Alberta (Taylor, 2001) — are likewise publicly funded but pedagogically autonomous. Private schools make up only 5.3% of enrolment, and are majority faith-based (58%): schools of non-Catholic faiths — Islamic, Jewish, other Christian — are only partially funded and sit outside the separate system (Hunt & DeJong VanHof, 2024). After faith-based schools, private schools are most frequently special emphasis (sports, arts, STEM; 30.6%), followed by top-tier ‘elite’ schools (8.9%).

Evidence that these structural differences translate into attainment is unclear. In BC, private schools outperform public schools by 0.12–0.15 SD net of lagged scores and neighbourhood SES, with the advantage concentrated in Catholic and preparatory schools (Azimi et al., 2023; see also Boerema, 2009); this gap is plausibly inflated by self-selection, since fee-paying signals both means and family investment in schooling. In Alberta, the small charter sector carries variable results (Bosetti & Butterfield, 2016), and the separate and private sectors' outcomes are untested.

2.9 Conclusion

This review has traced a gap in the Canadian literature on place and attainment that becomes more troubling at the provincial level, where school-system decision-making is concentrated. While international evidence shows that place-based attainment differences persist independently of SES, the Canadian literature is thin, dated, and methodologically limited on this topic; here, the marginalisation of rural and Indigenous place is furthered in its ‘research desert’ (Thier et al., 2021). No peer-reviewed study in the past decade has measured a rural association with school-aged attainment in Canada, and rurality is rarely operationalised beyond settlement size. Moreover, the only peer-reviewed studies measuring Indigenous attainment are BC-specific and over 15 years old, and none model it jointly with SES, rurality, and school type. In both cases, these associations have gone unexamined within individual provinces for over a decade, and uncompared across provinces since Ma and Crocker (2007).

The evidence reviewed here, together with each province's context, motivates the research questions below. These ask whether each discussed dimension of marginalised place

— settlement size, remoteness, and Indigenous population share — predicts attainment net of SES; whether school type predicts attainment and layers further advantage onto these dimensions of place; whether any of these associations are conditional on one another; and, throughout, whether they differ between British Columbia and Alberta.

2.10 Research Questions

Primary RQ: To what extent do socioeconomic status, rurality, Indigenous population share, and school type predict school-level academic attainment within each of British Columbia and Alberta, and how does the relative contribution of each factor compare across the two provinces?

Unless stated otherwise, each question below is asked within each province, and — at pattern level — across the two.

RQ1 (SES): To what extent do family income and parental education predict school-level attainment?

RQ2 (Settlement size): Beyond socioeconomic status, to what extent does settlement size predict attainment?

RQ3 (Remoteness): Beyond socioeconomic status and settlement size, to what extent does remoteness predict attainment?

RQ4 (Indigenous share): Beyond socioeconomic status and rurality, to what extent does Indigenous population share predict attainment?

RQ5 (School type): Beyond socioeconomic status, rurality, and Indigenous share, to what extent does school type predict attainment[1], and how does this association compare across British Columbia's two-category and Alberta's four-category systems?

RQ6 (Conditional associations): Do the associations identified under RQ1–5 depend on one another? That is, does the strength of any association vary with the level of another dimension?

[1] Further justification on why school type's association is of interest specifically independent of these factors is provided in the Methods.

Chapter 3: Methods

The present study is best characterised as a quantitative comparative case study (Gerring, 2017), in which the cases are two provincial education systems: British Columbia and Alberta. Within each province, regression models estimate how socioeconomic status, rurality, Indigenous share, and school type predict school-level attainment. The comparison between provinces then operates at the level of these estimated relationships rather than of absolute scores, a restriction required by the non-equivalence of the provinces' assessment instruments (see Analytic Approach). Additionally, conditional associations are examined — whether the association of one dimension with attainment varies across levels of another — extending the comparison from whether each dimension matters to where it matters most.

The study's contribution is one of hypothesis-guided descriptive comparison: while directional hypotheses structure the analysis, what is produced are detailed patterns of association — the strength of each dimension, independent and relative, and how it interacts with others — rather than causal effects or mechanisms. In a literature as thin as rural Canadian education, such description is a contribution in its own right: it documents socially important, previously unrecorded patterns, and points later causal work toward the mechanisms worth investigating (Loeb et al., 2017).

Within this descriptive frame, inferential statistics are used to strengthen observation rather than to test causal claims. Because the analytic samples approach the full population of eligible schools, significance testing does not serve sample-to-population generalisation. What any single cohort's results do reflect are transient influences: which students happened to write the assessment, the cohort's composition, and measurement error (Kane & Staiger, 2002; Leckie & Goldstein, 2017). P-values accordingly address whether an association could plausibly be an accident of the particular cohort observed, or whether a similar estimate should be expected were the exercise repeated (Field, 2018). Statistical significance is only one source of confidence in this study; variance accounted for and whether coefficients survive robustness checks (see Analytic Approach) also inform certainty about observed

associations. But such confidence only ever establishes that an association merits attention — never that one variable causally affects another.

3.1 Study Design: Provincial Comparison via Secondary Data

3.1.1 Provincial Comparison

The comparison is structured according to the Most Similar Systems Design (MSSD) framework (Przeworski & Teune, 1970), widely employed in comparative social science (Anckar, 2008, 2020; Della Porta, 2008; Seawright & Gerring, 2008). This design posits that when compared cases possess similar background conditions, differences in outcomes can be attributed to distinct explanatory variables rather than to structural factors. Similarity thus functions as a statistical control; while unaccounted-for contextual differences may persist, equivalent background variables are held approximately constant through purposeful case selection.

BC and Alberta satisfy the MSSD criteria: they diverge in ways that carry theoretically meaningful implications for rural education, while remaining similar enough that much of the surrounding context can be held approximately constant. The divergences — in economy, governance, and geography — were established as provincial context in the Literature Review; the directional expectations they inform are formalised below as hypotheses (Section 3.5). This section concentrates on the similarity side of the design: the shared context that allows differences in the estimated relationships to be linked to the identified divergences rather than to confounding context.

First, both provinces operate under a shared federal framework, exposing them to identical monetary policy, equalisation formulas, and social insurance programmes, and are bound by the Canadian Charter of Rights and Freedoms (1982). The two also draw from the same national teacher recruitment pool. Having these shared national conditions reduces the risk of omitted variable bias by holding numerous potential confounders constant (Gerring, 2017; Snyder, 2001).

BC and Alberta also have geographic profiles that are highly comparable in scale and settlement structure (Figure 3.1) — though not, as the Literature Review established, in their topography.

Figure 3.1

Map of Canadian Cities and Provinces



Note. British Columbia and Alberta (outlined in teal) are the focus of this study; the hatched area marks the Athabasca oil sands region. Population figures represent 2021 Census populations for Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs; Statistics Canada, 2021).

Both provinces are territorially vast, with major urban centres forming discrete nodes amid large rural expanses and rural population shares of 10–15% (Statistics Canada, 2021). Other provinces differ on exactly these dimensions: the Atlantic provinces are smaller, with short rural–urban distances and roughly 45% rural populations, while Ontario and Quebec's Quebec City–Windsor Corridor concentrates about half of Canada's population in a continuous urbanised belt absent in BC and Alberta (Statistics Canada, 2021). 'Rural' therefore means something similar in the two provinces — long distances from discrete urban nodes — in a way it would not in, say, a BC–Nova Scotia comparison.

Furthermore, both provinces have comparable Indigenous population shares — 5.9% in BC and 6.8% in Alberta, against a range of 2.2% to 18.1% across all provinces (Statistics Canada, 2021) — and both are subject to the same constitutional duty to consult Indigenous peoples in matters affecting their rights or title (Papillon, 2024). They also operate predominantly English-language K-12 public education systems with similar diploma structures and graduation requirements, and census-based provincial assessment programmes (Mulcahy, 2009).

No MSSD comparison achieves perfect control. Because only two cases are compared and they diverge in several ways at once, MSSD cannot isolate which divergence — theorised or not — drives any cross-provincial difference in the estimated relationships. Instead, interpretation relies on the hypotheses, each of which ties a predicted difference to a specific provincial mechanism in advance. Additionally, an ideal MSSD would hold the outcome measure constant, whereas the two assessment instruments differ in design and ages tested; the direction and pattern-level comparison adopted in the Analytic Approach mitigates but does not eliminate this limitation.

3.1.2 Secondary Data Analysis

Secondary data analysis provided the most viable approach for this study, most prominently because the project's nine-month timeframe and its being conducted from the United Kingdom made primary data collection unrealistic. This approach carries two principal limitations identified in the literature that apply to the current study and are

discussed as they arise. First, secondary datasets may fit the research question poorly: variables are often operationalised for institutional rather than academic purposes, and administrative data contain systemic errors that secondary analysts cannot trace or correct (Boo & Froelicher, 2012; Cohen et al., 2017; Smith, 2008, 2009). Second, official data are social and political products, and may reflect the agendas of originating institutions rather than the phenomena they purport to measure (Cohen et al., 2017; Smith, 2008).

3.2 Data Source Selection

The next task was identifying datasets able to operationalise attainment and the predictors established in the Literature Review: SES, rurality, Indigenous share, and school type. The same time and location constraints (Section 3.1.2) meant only publicly available data were considered; restricted Canadian microdata (e.g., Statistics Canada's Research Data Centres) can require in-country presence and long approval timelines. This section justifies the choices made among the available options, beginning with attainment.

3.2.1 Attainment

Standardised assessment scores were preferred from the outset as this study's attainment measure. They are more commonly employed than high school completion rates in the Canadian rural education literature (Cartwright & Allen, 2002; Ma & Crocker, 2007; cf. Layton, 2023). Completion is binary: it indicates graduation status but obscures variation in academic proficiency — the underlying construct of interest — whereas assessment scores capture the full spectrum of performance. Finally, contemporary Canadian research remains sceptical of graduation rates as a proxy for academic competence (Brown & Parekh, 2022; Pulchny et al., 2025).

School-level attainment data were necessary: Canadian school districts can span nearly 200,000 square kilometres (School District 91, Nechako Lakes), so district-level data could make a school in a regional hub indistinguishable from one in a remote hamlet. This study’s unit of analysis is therefore the school.

Alberta's Provincial Achievement Tests (PATs) and British Columbia's Foundation Skills Assessment (FSA) were chosen as this study’s attainment data sources because they

are available at the school level. While PISA is the attainment measure most commonly employed in comparable studies (Dean et al., 2023; Sullivan et al., 2018), school-level analysis of Canadian PISA data is prohibited under national governance policy, and its small sampled-school design could not be linked to the school-level covariates this study requires; the same school-level limit holds for the Pan-Canadian Assessment Program (PCAP; CMEC, n.d.).

One caveat applies to the chosen data: FSA and PAT results feed into school rankings (Cowley et al., 2025; Shang & Cowley, 2024) that inform parental choice. Schools may respond to these stakes to create bias — through unequal test preparation or lenient administration — but those that do so cannot be distinguished from those that do not; any resulting score inflation is acknowledged as a limitation.

3.2.2 Socioeconomic Status

Turning to SES — conventionally measured through income, parental education, and occupation — no school-level measure of any component was identified: PISA's ESCS exists only for its own small sample of schools, unrelated to the FSA and PAT records used here. Occupational data were unavailable at any suitable geographic level, so SES was operationalised through its income and education components, for which only one viable source was identified.

The Canadian Census of Population includes both family income and parental education data. Both measures are available from the federal level down to the Dissemination Area (DA) level — Canada's smallest published geographic unit, at approximately 400–700 residents. The next level up in Statistics Canada's geographic hierarchy, Census Subdivisions (CSDs), can span up to 24,000 square kilometres and 2.7 million residents. DA-level assignment was therefore employed throughout, as any coarser unit would risk the same distortion already noted for Nechako Lakes.

To employ these Census measures at the school level, each school is assigned the values of the DA in which it is located. This approach assumes that the socioeconomic character of a school's surrounding area reflects the circumstances of the students attending it — an assumption that requires justification. Davis et al. (2023) found that only 17% of

Canadians were assigned to the same income quintile as their DA, with agreement weakest in rural areas (weighted $\kappa = 0.11$); however, these findings concern whether a DA reflects a given individual's circumstances. This study's unit of analysis is the school, so the relevant question is population-level: whether schools in higher-income DAs tend to serve higher-income communities. Direct empirical support for this tendency is lacking, but comparable constraints led others to the same solution: Smyth and Darmody (2026) measured neighbourhood composition at Ireland's broadly analogous Electoral Division level because individual- and school-level data were unavailable.

The assumption should be strongest where this study's focus lies: rural schools typically serve a single, bounded community (Corbett, 2014), so the surrounding DA should approximate actual intake well. However, school choice — an overwhelmingly urban phenomenon (Adamson et al., 2016; Ball, 2003) — may let school populations diverge from their DA's residents (Yoon et al., 2018). Independent schools are the clearest case: enrolment is not residence-based and families often travel far to attend, so DA values may be especially misaligned with actual intake. The misalignment is also plausibly directional: fee-paying intakes may be considerably wealthier than the school's surrounding DA. Because all schools contribute to the SES estimate — and most are urban — urban mismeasurement weakens this estimate overall, not merely for the urban subsample. Finally, two caveats apply everywhere: catchments can exceed a single DA, and equal-population DAs can differ in their share of school-aged residents.

The assumption therefore remains a significant limitation, and associations involving SES are interpreted cautiously. Given SES's centrality in the rural education literature, the measure was nonetheless worth employing.

3.2.3 Share of Indigenous Peoples

A school-level Indigenous share measure exists in BC — the education ministry reports student counts by Indigenous identity — but Alberta has no equivalent. Using different instruments in each province would risk conflating instrument differences with

genuine cross-provincial variation, so Indigenous share was instead measured using the same Census approach as SES: the share of Indigenous peoples in a school's DA.

Unlike the SES case, this area-level approach arguably makes the coefficient's interpretation more aligned with this study’s interests. The Literature Review established that Indigenous marginalisation is inseparable from place — through reserves, the siting of residential schools, and their intergenerational legacies. A school-Indigenous-count measure, however, would also sweep in students schooled far from these geographies: Indigenous children are vastly overrepresented in foster care — 7.7% of the child population but 53.7% of foster children in 2021 (Hahmann et al., 2024) — and are disproportionately placed in urban, low-Indigenous-concentration areas.

The DA measure employed instead captures the association with being educated in a high-Indigenous-concentration area specifically, rather than with being Indigenous anywhere — a closer match to these place-associated dimensions of Indigenous disadvantage. And because residential-school trauma is transmitted through the families and communities who endured it (Jones, 2024), a community-level measure is well placed to register its intergenerational imprint. The measure will, however, capture less of the disadvantage characteristic of urban, low-concentration schools, such as racism and discrimination from classmates and teachers (Angus Reid Institute, 2021). Both of these processes — trauma transmission and in-school discrimination — are individual-level phenomena, better captured through individual-level data than through any area-level proxy; the claims here concern only which broad patterns the DA measure is better and worse positioned to reflect (see Section 3.2.6).

3.2.4 Settlement Size, Remoteness, and School Type

Settlement size was available through the ‘popctr’ classification in the 2021 Census Geographic Attribute File (Statistics Canada, 2022), which assigns every dissemination block to a population centre size class or to rural (see Measures).

Statistics Canada developed the Index of Remoteness (IR) in consultation with Aboriginal Affairs and Northern Development Canada and Health Canada, situating it within

a research tradition concerned with health accessibility and population disadvantage (Alasia et al., 2017); its use in the most recent Canadian research involving place, Indigeneity, and education (Layton, 2023) further supports its applicability here.

School type classification was extracted directly from the BC FSA database; for Alberta, where the raw PAT data omitted this information, school types were cross-referenced using Alberta Education's Authority and School Listing.

3.2.5 Summary of Data Sources

All datasets discussed in this section are summarised in Table 3.1 below.

Table 3.1

Datasets Utilised

Dataset	Provider	Level	Access	Period
Foundation Skills Assessment (FSA) results	BC Ministry of Education and Child Care	School	Public	2024/25
Provincial Achievement Test (PAT) results	Alberta Education	School	Public	2024/25
Census of Population 2021 — family income, parental education (25–64), Indigenous identity share	Statistics Canada	Dissemination area (DA)	Public	2021
Population centre classification (popctr)	Statistics Canada (2021 Census geography)	Dissemination block (DB)	Public	2021
Index of Remoteness (IR)	Statistics Canada (Alasia et al., 2017)	Census subdivision (CSD)	Public	2021
School type — BC (public/independent)	BC Ministry of Education (within FSA school file)	School	Public	2024/25
School type — Alberta (authority type)	Alberta Education (authority and school listing)	School	Public	2024/25

3.2.6 Data Sources Not Included

While primary data collection was not feasible (Section 3.1.2), it would have addressed several of this study's shortcomings. A closed-ended school questionnaire could have directly measured student-level SES and reached Indigenous students' individual experiences of schooling. Primary data collection also could have captured supply-side mechanisms through which remoteness is theorised to affect attainment, such as teacher retention and specialist staffing (Echazarra & Radinger, 2019; Knipe & Bottrell, 2023). No secondary measure of teacher staffing adequate for cross-provincial comparison was identified: only BC publishes school-level teacher headcounts, class-size data are operationalised differently across the two systems, and no teaching-specific labour-market data exist for either province. The omission of teacher-related data is acknowledged as a limitation; while the remoteness association is measured, the staffing dynamics that may constitute it cannot be examined with these data.

3.3 Analytic Samples

Because the FSA is administered in Grades 4 and 7 (BC) and the PAT in Grades 6 and 9 (Alberta), statistical models were estimated separately for each grade cohort. The cross-provincial comparison between BC Grade 7 and Alberta Grade 6 is the most developmentally proximate pairing and carries primary analytical weight, as elaborated in the Measures section. The primary analytic samples were therefore made up of schools with unsuppressed Grade 7 FSA or Grade 6 PAT results for the 2024/25 academic year; the Grade 4 (BC) and Grade 9 (Alberta) cohorts are subsidiary, used only for within-province robustness checks (Section 4.5).

Schools were matched to their corresponding DA using coordinate data where available; for schools without recorded coordinates, addresses were geocoded using R's tidygeocoder package. Schools were included in the analytic samples only where complete data were available across all variables. The school serves as the primary unit of analysis, with each observation representing a school-year composite attainment mean.

Assessment performance results are suppressed for schools with fewer than 10 ‘writers' — students who sat the assessment — in BC, or fewer than 6 in Alberta (Government of British Columbia, n.d.; Government of Alberta, n.d.-b). Suppressed institutions are by construction the smallest; the implications of this pattern are taken up below.

The primary analytic samples comprised 986 schools in Alberta and 863 in British Columbia; the full sample flow from reporting institutions is displayed in Table 3.2.

Table 3.2

Sample Flow by Province

Province	Reporting institutions	Analytic sample	% retained	Lost	% lost
Alberta (Grade 6)	1,619	986	60.9	633	39.1
British Columbia (Grade 7)	1,110	863	77.7	247	22.3

Alberta's writer-count field classifies all 470 score-related losses precisely (168 did not offer the relevant grade; 302 fell below the suppression minimum), whereas BC suppresses the writer count itself when small, so its 217 score-related losses cannot be attributed to a specific cause. The asymmetry reflects disclosure practice, not inconsistency in method.

A further 49 Alberta schools (15 DL/outreach, 34 Francophone) and 16 BC schools (DL/outreach) fell outside the study's scope. Another 114 Alberta and 14 BC schools were lost at the covariate-joining stage, where Census measures are suppressed in DAs with too few tax filers — a pattern concentrated in reserve and remote communities. Whether triggered by writer or tax-filer counts, suppression biases the sample in the same direction: the smallest schools, typically serving the most remote communities, are the least likely to appear in the analytic sample. These communities are precisely where the associations under study should register most strongly; this exclusion is noted as a limitation of the present analysis.

Participation, not suppression, drives a further absence: none of Alberta's 82 band-operated 'Federal First Nations' schools appear in any 2024/25 PAT dataset. Alberta recognises them as eligible reporting authorities but cannot require participation, and none currently participate. Many would likely fall below the six-writer threshold regardless, but their absence limits what this study can say about reserve-operated schooling.

This does not mean reserve-connected students are absent from the analytic sample. In Alberta, many on-reserve students access provincial testing through Education Services Agreements, attending neighbouring public or charter schools (Government of Alberta, n.d.-a); consistent with this, among the 15 retained Alberta schools with the highest DA-level Indigenous shares, half sit within 10 km of a reserve and all but one within roughly 50 km. Given the scale of both provinces' rural geography, and that many reserves have no school of their own, students from these communities plausibly attend these proximate schools even at distances other countries would not call 'nearby'. Reserve-connected students are present in BC's sample too, and more directly: seven of its equivalent 15 schools sit on reserve-type CSDs. In both cases, though, the DA-level measurement employed cannot confirm the share of enrolment any school draws from a reserve.

3.4 Measures

Having defined the case design, data sources, and analytic samples, this section specifies how each core variable is operationalised: its construction, its scale, and its linkage to schools. Descriptive statistics for all measures are reported in the Findings (Table 4.1).

3.4.1 Attainment

The dependent variable is a composite school-level attainment score. In British Columbia, this is the average of the school's Literacy and Numeracy FSA results; in Alberta, of its English Language Arts and Mathematics PAT results. Because the two provinces use different scoring scales — BC's FSA scale score (~200–800) versus Alberta's PAT percentage (0–100) — the composite is reported separately by province rather than pooled.

3.4.2 Socioeconomic Status (SES)

SES is operationalised using two Census variables. Income was measured through average after-tax economic-family income (continuous, in 2020 Canadian dollars), consistent with the wider literature (Duncan et al., 2014; Schulz, 2005) and the household-investment mechanism theorised in the Literature Review (Schneider et al., 2018). Parental education was measured through the share of the DA population aged 25–64 holding a bachelor's degree or higher (a proportion from 0 to 1). Each school is assigned the values of the DA in which it is located.

3.4.3 Rurality

Rurality is operationalised as two distinct measures. Settlement size uses the Census 'popctr' classification from the 2021 Geographic Attribute File, which places each area in one of four ordered classes, from rural (1) to large city (4). Popctr is published at the dissemination-block level — a physical geographic unit roughly equivalent to a city block in urban areas. It was instead assigned at the DA level, aligning its unit with the SES and Indigenous share measures, as the most common class among each DA's blocks. Little rests on this choice: in 87% of sample DAs, all blocks share one class, and a population-weighted alternative reclassifies under 2% of DAs.

The Index of Remoteness is a continuous 0–1 measure (higher = more remote) combining road-network distance to service points with population-weighted connectivity to economic and social infrastructure. It is measured at the CSD level, leaving less opportunity for variation than DA measurement permits. The largest CSDs are, however, overwhelmingly the most remote (Alasia et al., 2017), so a school's precise location within one is unlikely to alter its proximity to services materially; the unavailable within-CSD variation is therefore unlikely to be consequential.

3.4.4 Indigenous Share

Indigenous share is operationalised as the proportion of Indigenous peoples in a school's DA, drawn from the 2021 Census of Population and expressed on a 0–1 scale.

3.4.5 School Type

School type is operationalised to reflect each province's institutional structure. BC schools are coded public or independent — a binary that, per the funding structure established earlier (Section 2.8.4), accurately represents the provincial landscape.

Alberta requires four categories: public, Catholic separate, private (accredited), and charter. Collapsing constitutionally protected, fully funded separate schools with fee-charging private schools would be institutionally inaccurate, and charter schools — publicly funded but pedagogically autonomous — are retained as their own category to avoid misattributing their performance to either group. Francophone schools are excluded from both samples: the mechanism distinguishing them is linguistic rather than the religious or pedagogical factors presently of interest.

Because 'school type' names genuinely different institutional realities in each province, the variable is not comparable across provinces even in principle; each typology is interpreted within its own system.

3.5 Hypotheses

With the comparative provincial context established, data sources selected, and measures specified, the hypotheses can now be stated in operational terms. Consistent with this study’s ethical commitments (see Section 3.7), the directional hypotheses below predict associations within current educational structures; negative predictions do not imply that people or places are deficient. The hypotheses are as follows:

H1 (RQ1) — SES: Higher family income and parental education will each independently predict higher attainment in both provinces. The income association is expected to be weaker

in Alberta, where resource-sector employment weakens the education–income link (Corbett, 2014), with parental education's association, net of income, expected to be correspondingly larger.

H2 (RQ2) — Settlement Size: Smaller settlement size will predict lower attainment beyond SES, with the association expected to be larger in Alberta than in BC. The Oil Sands create demand for low-skill labour concentrated in small rural communities, actively discouraging educational attainment in ways not captured by family income alone (Corbett, 2014); no comparable mechanism operates in BC's more service-oriented economy.

H3 (RQ3) — Remoteness: Greater remoteness will predict lower attainment beyond SES and settlement size, with the association expected to be stronger in BC than in Alberta. BC's mountainous, fjord-cut terrain makes a given level of remoteness more consequential than in Alberta (Section 2.8.2), and, with the settlement-size association expected weaker in BC (H2), more residual variation remains for remoteness to capture.

H4 (RQ4) — Indigenous Share: Higher DA Indigenous share will predict lower attainment in both provinces beyond SES and rurality, reflecting intergenerational trauma, structural underfunding, and the Eurocentric content of provincial assessments. The association is expected to be less negative in BC, where First Nations self-governance and cultural continuity markers moderate Indigenous disadvantage (Chandler, 2010).

H5 (RQ5) — School Type: In BC, independent schools are expected to predict higher attainment beyond SES, reflecting socioeconomic selection into fee-paying schools. In Alberta, the separate-school association is less directionally certain: public funding removes the fee barrier, so any association may reflect pedagogical or values-based differentiation more than compositional advantage. No expectations are stated for Alberta's small private and charter sectors (Section 2.8.4, Table F5).

H6 (RQ6) — Remoteness × Indigenous share: Based on Layton's (2023) finding (Section 2.5), the negative association between remoteness and attainment is expected to be stronger where the Indigenous share of a school's surrounding area is higher; equivalently, the Indigenous share association is expected to be stronger in more remote areas. Statistically, this corresponds to a negative remoteness × Indigenous share interaction term in each province, specified in Section 3.6.

3.6 Analytic Approach

3.6.1 Hierarchical Regression Modelling

Hierarchical regression modelling was chosen to estimate the hypothesised predictors' associations with attainment because it allows a theoretical ordering of blocks (Field, 2018): each block's association with attainment is estimated net of all blocks preceding it.

The employed theoretical ordering follows from the Literature Review. If SES alone fully accounts for rural underperformance (Echazarra & Radinger, 2019), then, using the *complex* understanding of rurality discussed, it must wholly explain the social and spatial (dis)advantage that exists within the RSS. Entering SES first, before settlement size and remoteness, therefore tests whether these dimensions retain an independent association with attainment once SES is held constant. Positioning the Indigenous share block after these yields its association net of SES and rurality. This ordering matters twice over. First, Indigenous share falls within the RSS model's demography dimension, meaning its association would also need to be accounted for if SES were to account for rural underperformance. Second, an association persisting independently of SES and rurality would strengthen the case for greater Indigenous-directed support. School type enters last because it is the most policy-manipulable of the four factors: a government can authorise new schools or adjust funding formulas directly, but cannot quickly alter a community's income, settlement pattern, or trauma response. This positioning also serves the study of marginalised place: because school type can confer advantage on places (see Literature Review), it is a potential lever for redistributing place-based advantage; measuring its independent association is a first step in assessing that lever.

This study performs a hierarchical regression on attainment with the following blocks:

Block 1: SES (family income, parental education) — RQ1

Block 2: Rurality (settlement size, remoteness) — RQ2, RQ3

Block 3: Indigenous share (share of Indigenous peoples) — RQ4

Block 4: School type (4-type typology in Alberta, 2-type in BC) — RQ5

Following the main block structure, a remoteness by Indigenous share interaction term is entered as a supplementary test. RQ6 encompasses other conditional patterns — examined as exploratory analyses — but this term alone is assigned a formal hypothesis (H6) for three reasons. First, it produces a school-level analogue of a central finding of Layton (2023): that remoteness was more negatively associated with high-school completion for Indigenous than for non-Indigenous people. Layton's is arguably this study's closest Canadian counterpart in timing and content, sharing even the Index of Remoteness. Second, the term sits at the centre of this study's concern with marginalised place: the Indigenous dimension of place is arguably the most marginalised, so formally testing its intersection with remoteness — a less-measured instance of marginalised place (Section 2.1) — is not incidental to the study's purpose but an expression of it. Third, the provinces give reason to expect the interaction itself to differ: Alberta's fewer, larger treaty reserves tie Indigenous concentration to remoteness relatively uniformly, whereas BC's numerous small, scattered reserves span a wider remoteness range (Harris, 2002).

Within each provincial model, two statistics are of primary interest. First, standardised regression coefficients (β) indicate the direction and magnitude of each predictor's association with attainment, net of preceding blocks at its point of entry and of all predictors in the final model. Because the outcome and continuous predictors are z-scored, each β represents the SD-unit change in attainment per one-SD increase in the predictor within the province's own distributions. Standardisation serves within-province interpretation only — not cross-instrument magnitude comparison (see Comparison Level) — and each β is reported alongside the real-world value of one predictor SD (e.g., a dollar amount) to aid substantive reading. School type is the exception: it enters as an unstandardised dummy

(reference: Public), since standardising a binary predictor yields a coefficient that depends on category balance and corresponds to no meaningful change (Fox, 2016); its coefficient is a group difference in SD units of attainment, not comparable in magnitude to that of the continuous predictors.

Second, the incremental variance explained (ΔR^2) at each block indicates the additional attainment variation that block's variables account for beyond preceding blocks — directly testing whether rurality, Indigenous share, and school type each contribute independently once SES is held constant. Because predictors are correlated, a block's ΔR^2 depends on entry order: whichever block enters first claims the variance it shares with later blocks. Each ΔR^2 is therefore read as a block's contribution given the theoretical ordering, not as an order-free measure of importance, and models entering school type first are reported as robustness checks on this. Where the analysis ranks predictors or blocks within a province, ranking is based on variance explained rather than coefficient size, whose mixed units preclude it. ΔR^2 is likewise reported within each province only: the share of attainment variance that is explainable at all differs between instruments for reasons unrelated to the strength of the modelled relationships (see Comparison Level).

3.6.2 Comparison Level

Cross-province comparisons are drawn only at the level of direction and relative pattern — whether a predictor is significant in both provinces, and how it ranks within each provincial model — never as claims about the size of a difference between two coefficients. This reflects a genuine measurement concern: FSA and PAT differ in scale and scoring method, creating real comparability problems when scores are compared directly (Fung & Chu, 2015). Moreover, Ho (2009) shows that effect-size statistics — the class to which standardised coefficients belong — are not invariant under non-linear rescaling. An apparent cross-province difference in coefficient size therefore cannot be distinguished from a by-product of how the two tests' scales are constructed. Conversely, direction, significance, and rank-order claims are substantially more robust to such rescaling (Ho, 2009); thus,

hypotheses phrased as relative-magnitude expectations (e.g., a larger settlement-size association in Alberta) are assessed through these converging pattern-level indicators.

Neither composite exhibits the marked skew or boundary censoring under which Ho's (2009) concerns bind most strongly (Appendix A). As a further check, all four models were re-estimated on rank-transformed outcomes (Appendix D); because ranks are invariant to any monotone rescaling, stability under this specification indicates results are not artefacts of scale shape. In both primary models (BC Grade 7, Alberta Grade 6), every predictor's sign and significance remained unchanged.

3.6.3 Assumptions

Several standard regression assumptions warrant consideration. First, linearity was safely assumed for the continuous predictors — income, parental education, remoteness (IR), and settlement size (popctr), the last treated as continuous given its four ordered categories (Field, 2018) — as residual plots revealed no systematic curvature (Appendix B). Second, normality of residuals and homoscedasticity were assessed through standard diagnostic plots: heteroscedasticity was mild and confined to the lower tail of the attainment distribution, inconsequential for inference at these sample sizes (Field, 2018; Appendix B). Third, multicollinearity warranted checking given the established intercorrelations among the predictors; all VIFs fell well below conventional thresholds, with the highest being 3.14 (Appendix C). The strong settlement size–remoteness correlation ($r \approx -0.71$ to -0.73) still calls for interpretive caution — their coefficients are partial, each conditional on the other — but both remained stable as each entered, indicating correlation rather than destabilising collinearity.

Fourth, full independence of observations cannot be assumed: schools within a CSD share remoteness scores by construction, and schools within a district share administrative policies and resources. Standard errors are therefore clustered at the CSD level, using the CR2 small-sample adjustment with Satterthwaite degrees of freedom (Pustejovsky & Tipton, 2018; R script in Appendix L). As a further sensitivity check, key coefficients were re-tested with a wild cluster bootstrap (Cameron, Gelbach, & Miller, 2008), using Webb weights given

the unbalanced cluster sizes (MacKinnon & Webb, 2017) and 9,999 replications; bootstrap and Satterthwaite p-values are reported together in Appendix D (Table D2). All checks held equally for the secondary pairing (BC Grade 4 / Alberta Grade 9).

3.7 Ethical Considerations

All data analysed in this study are publicly available aggregates: no human participants were recruited, and no identifiable individual data were accessed, so institutional ethics review was not required (see Appendix E). The provinces' suppression thresholds (Section 3.3) also offer privacy protection: results for cohorts small enough to risk identifying individual students are never published in the first place.

By associating dimensions of marginalised place with lower measured attainment, this study risks implying that people or places are 'deficient' (Corbett, 2014) — an implication it rejects. Both the FSA and PATs are designed by provincial governments based in urban centres, and Canadian curricula have been shown to marginalise Indigenous perspectives, with Indigenous incorporation substantially improving Indigenous student outcomes (Spencer, 2025). These instruments may therefore understate the abilities of the very populations this study centres; its findings describe attainment within current educational structures, not the capacities of communities.

A related consideration concerns Indigenous data governance. This study analyses state-collected data about Indigenous communities without those communities' involvement — the configuration that Indigenous data-sovereignty principles (OCAP®: ownership, control, access, and possession; FNIGC, n.d.) were articulated against. However, public school-level aggregates fall outside OCAP's formal scope, and analysing provincial assessment data for Indigenous inequities aligns with a documented First Nations stance: BC's First Nations Leadership Council has publicly defended provincewide assessment as 'necessary for addressing inequities' in K–12 education, against campaigns for its cancellation (Ungerleider, 2022). Still, research conducted with Indigenous communities remains the ideal (see Conclusion). In its absence, this study honours OCAP's underlying principle — care in the use and representation of data concerning Indigenous peoples — by

reporting findings only as coefficients, categories, and quartiles, naming no school serving an identifiable Indigenous community.

Finally, the author is neither Indigenous nor a rural dweller. This work is intended to serve rather than speak for these populations, but writing about communities one does not belong to carries a risk of mischaracterisation; the descriptive register adopted throughout, and the consultation-based future research recommended in the Conclusion, are, in part, responses to that risk.

Chapter 4: Findings

This chapter reports the study's results: the descriptive overview (4.1), each province's primary model (4.2–4.3), the cross-provincial comparison (4.4), robustness in the secondary cohorts (4.5), and conditional associations (4.6), closing with the hypothesis outcomes (4.7). Throughout, results are reported and patterns characterised. Interpretations related to stated hypotheses and to exploratory findings are reserved for the Discussion.

4.1 Descriptive Overview

The Methods established BC and Alberta's overall similarity as the basis for the MSSD design; descriptive statistics presented in Table 4.1 further bear this similarity out.

Table 4.1

Descriptive Statistics for the Primary Analytic Samples, Alberta Grade 6 and British Columbia Grade 7

			SES		Rurality		
Province	*N* schools	Attainment	DA income ($k)	Share bachelor's+	Settlement size (1 = rural, 4 = large city)	Remoteness	Indigenous share
Alberta	986	53.80 (8.80)	109.50 (34.80)	0.28 (0.19)	3.09 (1.14)	0.18 (0.10)	0.07 (0.07)
British Columbia	863	466.00 (58.90)	110.60 (26.50)	0.32 (0.16)	3.18 (1.11)	0.18 (0.13)	0.06 (0.10)

Note. Values are *M* (*SD*). Attainment is reported on each province's native scale (see Methods).

The high mean settlement sizes in Table 4.1 reflect a top-heavy distribution: large-urban schools comprise the majority of each analytic sample (55.0% in Alberta, 59.9% in BC; see Appendix F). Income, remoteness, and Indigenous share have near-identical means across the provinces, with parental education only slightly differing (M = 0.32 and 0.28).

The two rurality measures relate to income differently — a contrast relevant to interpreting their associations in Sections 4.2 and 4.3. Income is not monotonic with settlement size in either province. In BC, small towns — not rural DAs — record the lowest

mean income (~$96k versus ~$103k; Tables F1–F2). This likely reflects both the exclusion of income-suppressed remote DAs (see Analytic Samples) and the classification of affluent exurban DAs as 'rural' by settlement size: reassigning these DAs (IR < 0.20) restores a monotone income gradient (Table G1). In Alberta, the same ordering appears, but the gap is negligible (~$100k versus ~$101k). Conversely, remoteness orders income monotonically in both provinces, though more steeply in BC (a ~$25k decline to the most remote quartile, versus ~$12k in Alberta).

Because census suppression concentrates school losses in remote and reserve communities (Section 3.3), excluded DAs are likely lower-income and higher in Indigenous share than retained ones; the income and parental education means in Table 4.1 thus likely overstate the population means. The same exclusions should shape the regression estimates in the opposite direction: narrowing each predictor's range precisely where its association with attainment would register most strongly attenuates estimated associations. The remoteness and Indigenous share associations reported here are therefore conservative; the true scale of remoteness- and Indigenous-share-related inequality is likely greater than estimated.

Finally, school-type composition is reported in Appendix F (Table F5). The independent-school share is considerably greater in BC (19.4%, n = 167) than in Alberta (5.4%, n = 53), giving BC's H5 contrast the more substantial estimation base. BC may appear to have the larger public sector (80.6% vs 67.3%), but Alberta's separate (25.9%) and charter (1.4%) schools are also fully publicly funded and non-fee-paying.

4.2 British Columbia Grade 7: Primary Model Results

Table 4.2 presents the BC Grade 7 hierarchical regression results. Coefficients are standardised throughout, and p-values and variance explained are read as strengthening confidence in an observed association rather than establishing it (see Methods).

Table 4.2

Hierarchical Regression Predicting School Mean Foundation Skills Assessment (FSA) Score (Standardised Coefficients), British Columbia Grade 7

Predictor	*SD*	Model 1: SES	Model 2: + Rurality	Model 3: + Indigenous share	Model 4: + School type
Intercept	—	0.00 (0.05)	0.00 (0.06)	0.00 (0.06)	−0.91*** (0.09)
Block 1: SES					
DA income	$26.5k	0.08 (0.04)	0.06 (0.04)	0.05 (0.04)	0.05 (0.03)
Share bachelor's+	0.16	0.46*** (0.05)	0.40*** (0.06)	0.37*** (0.07)	0.35*** (0.05)
Block 2: Rurality					
Settlement size	1.11	—	−0.06 (0.05)	−0.07 (0.06)	−0.05 (0.05)
Remoteness	0.13	—	−0.16*** (0.05)	−0.11* (0.04)	−0.08 (0.05)
Block 3: Indigenous share					
Share Indigenous	0.10	—	—	−0.16*** (0.03)	−0.19*** (0.03)
Block 4: School type					
Independent (ref: Public)	—	—	—	—	1.13*** (0.13)
Model fit					
R^2		.25	.26	.28	.48
ΔR^2		—	+.01	+.02	+.20

Note. N = 863 schools. β with cluster-robust (CR2) standard errors in parentheses; outcome and continuous predictors z-scored (see Methods). *SD* = predictor's standard deviation in raw units. Independent is a 0/1 dummy (reference: Public). *p < .05. **p < .01. ***p < .001.

When only SES measures were accounted for, parental education emerged as the dominant SES predictor. A one-SD increase in parental education — approximately 16 percentage points of bachelor's-plus share — was associated with a nearly half-SD increase in attainment (β = 0.457, p < .001), equivalent to a median school rising to about the 68th percentile of the provincial distribution. This result is consistent with H1's prediction that parental education would independently predict attainment in BC. Family income, by contrast, showed no significant association with attainment (β = 0.077, p > .05), against H1's prediction.

When settlement size and remoteness entered at the next stage, remoteness was the more predictive of the two: a one-SD increase (≈0.13 on the 0–1 Index of Remoteness) was associated with a 0.165-SD decline in attainment (β = −0.165, p < .001) — equivalent to a median school falling to about the 43rd percentile. This supports H3: remoteness negatively predicted attainment, independently of SES measures already accounted for. Conversely,

settlement size showed no significant association with attainment on entry ($\beta = -0.061$, $p > .05$), a null result that does not support H2 in BC. The settlement-size null is robust to measurement: upon reclassifying the 40 schools coded ‘rural’ (popctr = 1) but sitting on city outskirts (IR < 0.20), settlement size remained non-significant ($\beta = -0.073$; Table G2). Finally, the SES coefficients attenuated modestly upon the rurality block’s entrance (income from 0.077 to 0.058; parental education from 0.457 to 0.404) with no changes in significance, consistent with some of their predictive power overlapping with rurality.

Upon its entry into the model, a one-SD increase in Indigenous share (≈9.8 percentage points) was associated with 0.159-SD lower attainment ($\beta = -0.159$, $p < .001$), supporting H4. The sample spans roughly 190 CSD clusters. However, the estimation of this coefficient behaves as if drawing on about thirteen independent places (Satterthwaite df ≈ 13.4), consistent with the right-skewed distribution of Indigenous share: in a province whose population is roughly 6% Indigenous (Table 4.1), high-share schools are rare. This association should therefore be interpreted cautiously, though it survives the wild cluster bootstrap (Appendix D).

Indigenous population share explained a further 1.8% of the variation in attainment ($\Delta R^2 = .018$) — modest in absolute terms, yet comparable to the rurality block's combined contribution ($\Delta R^2 = .012$), despite being a single predictor. While most coefficients were essentially unchanged upon its inclusion (shifting ≤ ~0.03 SD, with no changes in significance), remoteness attenuated from $\beta = -0.165$ ($p < .001$) to -0.107 ($p = .013$) — a shift consistent with the geographic overlap between remoteness and Indigenous population share (in-sample r = .51; Section 4.6.1).

Arguably the most notable result of the BC regressions, H5 was supported to a dramatic extent, as independent schools scored on average 1.134 SD above public schools (SE = 0.126, $p < .001$). This difference persists after accounting for SES, rurality, and Indigenous share; moreover, entered first, school type's coefficient is 1.165 against 1.134 in the full model (Table I1) — the association overlaps little with the measured covariates. BC’s public–independent gap is comparable to the difference in rank between a school at the 50th and 87th percentile of the score distribution. In outcome terms: for parental education — the

largest standardised coefficient among the continuous predictors — to be associated with an equivalent attainment difference, a school's DA bachelor's share would need to reach roughly 86%, 3.25 standard deviations above the sample mean. Only one school in the sample reaches that level; school type's contrast applies to every school.

While ΔR^2 is order-dependent (Section 3.6.1) and school type enters last, its addition still increased R^2 by .20 — nearly seven times the combined contribution of the rurality and Indigenous share blocks. The remoteness coefficient shrank further upon school type's entry into the full model, from $\beta = -0.107$ to -0.081, and was no longer significant. This reflects neither collinearity (all VIFs well within limits; Appendix C) nor school type standing in for location: independent schools make up a broadly similar share of every remoteness quartile, and the two variables are near-uncorrelated ($r \approx -0.03$; Section 4.6.2).

4.3 Alberta Grade 6: Primary Model Results

Table 4.3 presents the Alberta Grade 6 regression results — the second half of the primary analytic pairing. Each school-type coefficient represents that sector's contrast with public schools in provincial standard-deviation units.

Table 4.3

Hierarchical Regression Predicting School Mean Provincial Achievement Test (PAT) Score (Standardised Coefficients), Alberta Grade 6

Predictor	SD	Model 1: SES	Model 2: + Rurality	Model 3: + Indigenous share	Model 4: + School type
Intercept	—	0.00 (0.04)	0.00 (0.04)	0.00 (0.05)	−0.09 (0.06)
Block 1: SES					
DA income	$34.1k	0.07 (0.04)	0.06 (0.03)	0.06 (0.03)	0.05 (0.04)
Share bachelor's+	0.19	0.35** (0.05)	0.39* (0.09)	0.35* (0.10)	0.34* (0.10)
Block 2: Rurality					
Settlement size	1.14	—	−0.20** (0.06)	−0.17* (0.06)	−0.18** (0.06)
Remoteness	0.10	—	−0.16** (0.05)	−0.11* (0.05)	−0.11* (0.05)
Block 3: Indigenous share					
Share Indigenous	0.07	—	—	−0.13* (0.05)	−0.12* (0.05)
Block 4: School type					
Separate	—	—	—	—	0.15* (0.05)
Charter	—	—	—	—	1.22* (0.22)
Private	—	—	—	—	0.62* (0.14)
Model fit					
R^2		.16	.18	.19	.23
ΔR^2		—	+.02	+.01	+.04

Note. As Table 4.2; N = 986 schools. School type entered as 0/1 dummies against a Public reference (Separate n = 255; Charter n = 14; Private n = 53). *p < .05. **p < .01. ***p < .001.

When only SES measures were accounted for, income was again non-significant (β = 0.071, p > .05), while parental education emerged as the stronger of the two. A one-SD increase in a school's bachelor's-plus share — approximately 18.5 percentage points — was associated with a 0.352-SD increase in attainment (p = .003), equivalent to a median school rising to about the 64th percentile of the provincial distribution. This coefficient warrants care, however: in the final model it is significant under the primary inference (p = .032) but not under a wild cluster bootstrap check (p = .129), consistent with this variable's very low effective degrees of freedom (df ≈ 4; see Methods). It is therefore treated as suggestive rather than established in Alberta. That said, the parental education association is substantial and positive in every specification, recurs at Grade 9, survives rank transformation, and parallels

BC's robustly estimated finding. Moreover, the SES block's substantial variance contribution ($\Delta R^2 = .16$) persists largely intact when school type enters first (Table I2).

Both rurality measures were significant upon entry. A one-SD increase in settlement size (≈1.14 categories on the four-point scale) was associated with a 0.196-SD decline in attainment ($p = .004$), and a one-SD increase in remoteness score (≈0.10 IR) with a 0.163-SD decline ($p = .003$). The remoteness result supports H3 within Alberta: remoteness predicted attainment independently of SES and settlement size. The settlement size result, by contrast, runs contrary to H2's prediction: rather than smaller settlements underperforming beyond SES, attainment declined as settlement size increased. Supplementary decomposition (Table H1) clarifies the structure of this result: entered alone, neither settlement size ($p = .120$) nor remoteness ($p = .503$) was significantly associated with attainment; entered together, both settlement size and remoteness were, indicating that each predicts attainment only once the other is held constant.

The negative settlement-size coefficient should not be interpreted as a simple urban penalty gradient. Of the remoteness quartiles, Q3 is the only band in which all four settlement-size categories co-occur (Table H2). There, rural-classified schools (popctr = 1) outperformed small-town schools net of SES (medium-city schools were also directionally lower, $p = .090$), with large-urban schools indistinguishable from rural ones. This pattern is evidence that attainment does not move steadily across the four categories as the linear specification assumes: the negative coefficient summarises a non-monotone pattern across internally heterogeneous categories. The association nonetheless survived reclassification of the 33 exurban schools (popctr = 1, IR < 0.20): $\beta = -0.178$ to -0.169, $p = .008$ to .020 (Table G3).

Indigenous population share was significantly associated with attainment upon entry: a one-SD increase (≈7.2 percentage points) predicted a 0.129-SD decline ($p = .024$), supporting H4 in Alberta. Its entry modestly attenuated both rurality coefficients (settlement size from −0.196 to −0.170; remoteness from −0.163 to −0.112), with both remaining significant ($p = .012$; $p = .025$), and added $\Delta R^2 = .01$. As in BC, while the sample spans roughly 190 clusters, the estimation of this coefficient behaves as though it draws on about

nine independent places (Satterthwaite df ≈ 9.0); Alberta's Indigenous share estimate is thus treated with the same caution.

The final Alberta Grade 6 model included school type as three dummy variables against a public reference. All three non-public sectors carried positive, significant coefficients: separate (Catholic) schools scored 0.150 SD above public schools (p = .021), private schools 0.619 SD above (p = .010), and charter schools 1.222 SD above (p = .018). The private and especially charter estimates warrant caution: the 53 private schools fall within just 18 CSD clusters, the 14 charter schools within only 6 (Table F5), so the cluster-robust inference for these terms rests on few effective clusters (Satterthwaite df = 4.1 and 2.6, respectively). This limits not only the significance tests but the estimates themselves: a contrast estimated from so few schools is imprecise — the charter coefficient's 95% CI spans [0.43, 2.01] — so its magnitude should not be treated as a reliable sector gap, nor ranked against other coefficients. School type explained a further 3.9% of the variation in attainment ($\Delta R^2 = .04$). This places school type second in variance explained, behind SES ($\Delta R^2 = .16$), but it remains more proximate to the concepts that explain less variance: rurality ($\Delta R^2 = .02$) and Indigenous share ($\Delta R^2 = .01$).

In the school-type-first ordering (Table I2), the private coefficient attenuated most upon SES entry (0.875 to 0.662, ≈24%) — directionally consistent with selection into fee-charging schools — though the small sample (n = 53) limits any firm reading. The separate coefficient, by contrast, was little moved by SES (0.155 to 0.106) and recovered once rurality entered (0.150 in the final model), consistent with the sector's more urban but SES-similar profile (Table H3). In this ordering, school type's variance contribution rises to nearly half the SES block's; however, school type in Alberta is the only block in any tested model to span three predictors, and blocks with more parameters tend to capture more variance regardless of substantive importance.

4.4 Cross-Provincial Comparison: Primary Pairing

This section compares the two provinces' final models, drawing on the last columns of Tables 4.2 and 4.3. Per the Analytic Approach, comparison operates at the level of

direction, significance, and predictor ranking within each model; the size of any cross-provincial gap is not itself interpreted.

Indigenous share is negative and significant in both provinces' final models, and robust to inference procedure in both (see Methods); it thus behaves the most uniformly of any predictor across the two systems, supporting H4 in each. The SES pattern converges too, though on less even ground. In both provinces, parental education is the stronger SES predictor and DA income is significant in no final model. Still, where BC's parental education estimate is robust across inference procedures, Alberta's is significant only under the primary inference, not the bootstrap check. H1 is therefore supported for parental education in BC and suggestively so in Alberta, and unsupported for income in both systems — though Alberta's null tests are underpowered at the cluster level, so the income null is more firmly established in BC.

The rurality blocks diverge sharply in two respects. First, remoteness follows the same initial pattern in both provinces — significant on entry, attenuating once Indigenous share enters — but reaches different endpoints. In BC, remoteness attenuates to non-significance ($p < .001$ at entry to $p = .085$ in the final model), whereas Alberta's remains significant throughout ($p = .003$ to $.037$). The BC estimate is borderline rather than clearly null, however, so the contrast is better read as a difference of degree than a categorical divergence. Second, settlement size is null in every BC specification but significant in Alberta's final model, in the direction opposite to H2's prediction and only once remoteness is jointly modelled (Section 4.3). H2 is therefore unsupported in both provinces, for informatively different reasons: a simple null in BC, a reversed conditional association in Alberta.

School type predicts attainment beyond SES, rurality, and Indigenous share in both provinces, supporting H5, but its structure and associations differ. BC's single contrast — independent versus public — produces a large advantage that is stable across entry orders (Table 4.2; Table I1). In Alberta, separate, private, and charter schools score above public schools — separate least, private more, charter most (Table 4.3) — though the private and especially charter estimates rest on few schools in few clusters. School type sits second in

variance explained, behind SES, in both provinces. However, Alberta's contribution (ΔR^2 = .04) sits nearer to rurality (.02) and Indigenous share (.01) than SES (.16), despite spanning three predictors, while BC's approaches SES itself (.20 vs .25); the school-type-first model orderings echo these proximities (Appendix I). With the variables measured here, pattern-level differences signal that school type may matter more in BC than in Alberta.

Finally, the four blocks collectively account for far more attainment variance in BC (R^2 = .48) than in Alberta (.23) — a gap reported descriptively rather than interpreted, since the instruments' differing scale construction means it may reflect measurement error as much as substance (see Methods). Whatever its source, this gap undoubtedly contributes to Alberta's coefficients being less significant throughout. Explained variance shrinks the residual noise around estimates; with less of it, Alberta's standard errors stay inflated, which weakens significance for any given coefficient size. Correspondingly, no Alberta specification produces a predictor significant at $p < .001$, where BC's final model has three.

4.5 The Patterns Hold in the Secondary Cohorts (BC Grade 4 / Alberta Grade 9)

Robustness checks on the secondary cohorts reproduce the primary patterns within each province (Tables J1–J2), with only minor deviations. At BC Grade 4, DA income turns weakly significant ($p \approx .03$) and school type's contribution grows stronger still (ΔR^2 = .21, moving above SES's .17). In Alberta, every substantive Grade 6 pattern recurs at Grade 9, though charter and private estimates, while near their Grade 6 values in magnitude, fall to borderline non-significance on even smaller cluster bases.

4.6 Do the Associations Vary by Context? (RQ6)

4.6.1 In Both Provinces, the Remoteness and Indigenous Share Associations Do Not Depend on Each Other (H6).

H6 predicted that the negative association between remoteness and attainment would strengthen as Indigenous population share rises; equivalently, that the Indigenous share

association would strengthen with remoteness. This is a school-level analogue of Layton's (2023) finding that remoteness matters more for Indigenous than for non-Indigenous completion. However, the interaction term was non-significant in both provinces ($\beta = 0.011$, $p = .716$ in BC; $\beta = -0.023$, $p = .445$ in Alberta), with model R^2 essentially unchanged in each case and no collinearity concerns among the interacting terms (GVIF^(1/(2df)) < 1.65). H6 is therefore not supported: neither dimension's association with attainment detectably depends on the level of the other. This null concerns only their conditionality, not their presence; both dimensions' independent associations remain as reported in the primary models.

In both provinces, remoteness and Indigenous share are themselves moderately correlated ($r = 0.51$ in BC; $r = 0.39$ in Alberta). They are intertwined enough that the most remote quartile carries roughly three times the Indigenous share of the least remote in Alberta, and roughly ten times in BC. Yet they have sufficient independent variation for the interaction to be precisely estimated: the estimates are near zero with small standard errors in both cases (SE = 0.028 and 0.026). The nulls are therefore interpretable in both provinces: the two dimensions travel together, but any conditionality between their associations is at most very small.

4.6.2 In BC, the Dramatic Independent-School Advantage Shrinks With Remoteness but Never Disappears

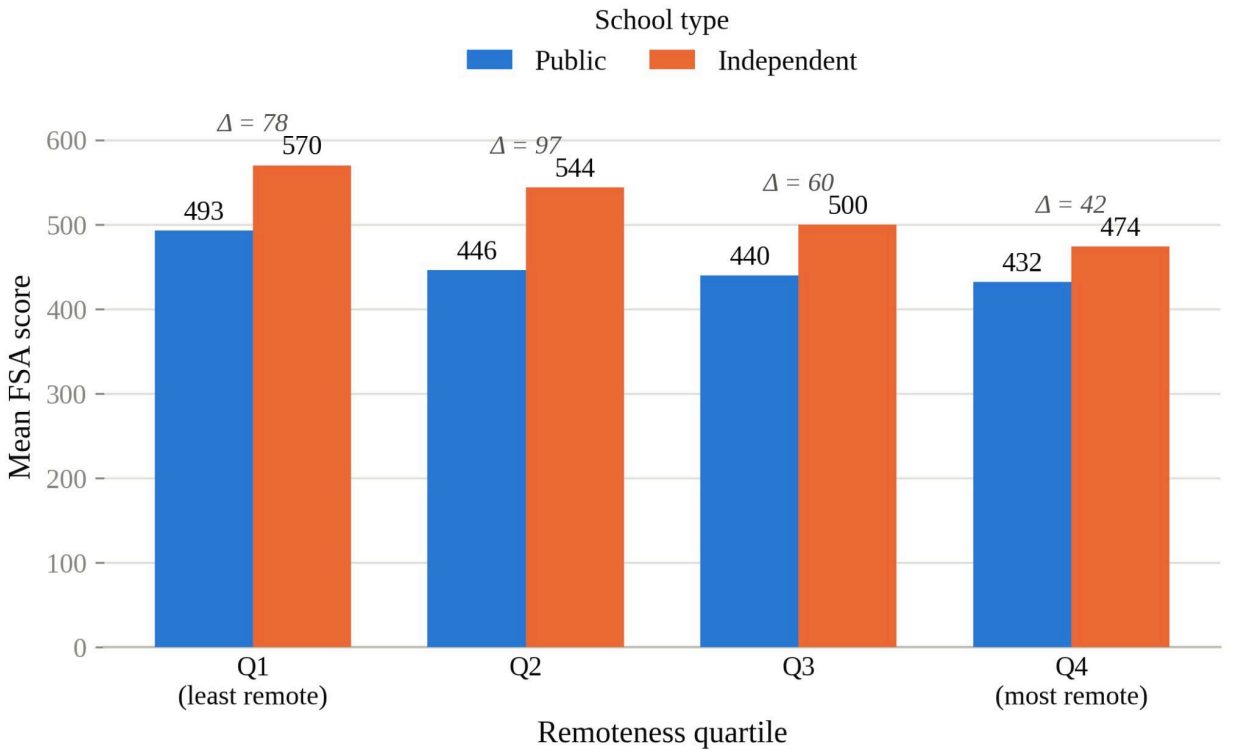
A second conditional pattern was examined post hoc: whether the school-type attainment gap varies with remoteness. Prompted by the size of the BC school-type coefficient and not hypothesised, it is reported as exploratory. Remoteness was used as the conditioning dimension rather than settlement size, which showed no significant association with attainment in BC in any specification.

Two distinct questions were separated. First, whether access to independent schooling varies by remoteness quartile: independent schools are present across the full remoteness distribution, comprising 19.4%, 15.7%, 26.4%, and 15.8% of schools from the least to most remote quartile, with no clear gradient (formal tests agree; Tables K1–K3). Second, whether

the independent-school advantage varies with remoteness; here, the school type by remoteness interaction was significant (β = −0.312, p = .003; Table K5): the advantage is larger among less remote schools (Figure 4.1).

Figure 4.1

Mean Foundation Skills Assessment (FSA) Score by School Type and Remoteness Quartile, British Columbia Grade 7



Note. Values are raw school-level mean FSA scores with no covariate adjustment (*N* = 863). Δ = the Independent − Public difference at each quartile, computed before rounding.

While both public and independent schools score lower with greater remoteness, the independent–public gap is actually larger at remoteness Q2 than Q1 — and both gaps are sizeable: 97.3 points at Q2 (Cohen's d = 2.31) and 77.5 at Q1 (d = 1.75). In other terms, the Q2 gap is comparable to the difference between a school at the 50th and the 99th percentile of the quartile's distribution; at Q1, between the 50th and the 96th. The gap does shrink to

42.5 points at Q4, but there it remains comparable to the difference between the 50th and 80th percentile, and is strongly significant ($p < .001$) with a large effect size ($d = 0.83$).

While access to independent schooling may not vary drastically by remoteness quartile, that does not mean independent schools are evenly distributed across BC (Table K2). In fact, 118 of the sample's 167 independent schools (71%) sit in the Lower Mainland (including Vancouver and its surrounding communities), Greater Victoria, or Kelowna, and schools in these areas make up 100% of Q1 and Q2 and 74% of Q3. This leaves only 49 (29%) spread across the rest of the province; independent-school advantage exists at every remoteness level, but the schools carrying it overwhelmingly inhabit a few urban regions.

4.6.3 In Alberta, the Separate-School Advantage Widens With Remoteness but Remains Modest

The equivalent tests in Alberta — motivated by the BC result and likewise exploratory — produced a partial parallel running in the opposite direction. The separate by remoteness interaction was positive and significant ($\beta = 0.121$, $p = .027$; full model in Table K5): the separate-school advantage is larger among more remote schools. The interaction appears anchored in the observed data at the remote end (Table K6): in Q4, separate schools' raw mean exceeds public schools' by 2.6 percentage points ($p = .011$) — a median separate school sitting at roughly the 64th percentile of the public-school distribution — while the gap is small and non-significant in the three less remote quartiles. This Q4 difference remains only small-to-medium in effect size ($d = 0.35$).

4.6.4 The Independent-School Advantage Does Not Detectably Vary With Parental Education (Exploratory, BC)

A further exploratory test — whether the independent-school advantage varies with DA-level parental education — was null ($\beta = -0.04$, $p = .563$; full model in Table K4). However, the DA measure is weak precisely for choice-driven schools, whose students often travel from outside the surrounding school catchment (Section 3.2.2).

4.7 Summary of Hypothesis Outcomes

Table 4.4 summarises the outcomes of H1–H6 for the primary pairing; the final column records the verdict on each hypothesis's cross-provincial prediction, drawing on the comparisons in Section 4.4. Verdicts follow the comparison level established in the Methods, and 'bootstrap sensitivity' refers to the wild cluster bootstrap check (Appendix D, Table D2).

Table 4.4

Hypothesis Outcomes for the Primary Pairing, British Columbia Grade 7 and Alberta Grade 6

Hypothesis	Prediction (abbreviated)	BC verdict	Alberta verdict	Cross-provincial prediction
H1 (income)	Income independently predicts attainment; weaker in AB	Not supported (*ns* throughout)	Not supported (*ns* throughout; power limited)	No clear evidence: null in both provinces
H1 (parental education)	Parental education independently predicts attainment; larger independent contribution in AB	Supported (significant throughout; robust to bootstrap sensitivity)	Suggestive: significant under primary inference; fails bootstrap check	Evidence against: AB's association is the less firmly established, not the larger
H2 (settlement size)	Smaller settlements predict lower attainment beyond SES; larger association in AB	Not supported (*ns* throughout)	Not supported as specified; reversed conditional association (Section 4.3)	Evidence against: AB showed no larger association in the predicted direction
H3 (remoteness)	Remoteness predicts lower attainment beyond SES and settlement size; larger association in BC	Supported at entry ($p < .001$); not sustained in the final model ($p = .085$)	Supported: significant throughout	Evidence against: the pattern ran opposite (persists in AB, attenuates in BC), though the difference is one of degree (p = .037 vs .085)
H4 (Indigenous share)	Higher Indigenous share predicts lower attainment beyond SES and rurality; less negative in BC	Supported ($p < .001$, final model; bootstrap-robust)	Supported ($p = .024$, final model; bootstrap-robust)	No clear evidence: identical patterns; no difference discernible at pattern level
H5 (school type)	BC: independent-school advantage beyond SES. AB: no directional prediction	Supported (significant throughout; robust to entry order and bootstrap sensitivity)	No direction predicted; a modest separate-school advantage emerged	n/a — none made
H6 (remoteness × Indigenous share)	Remoteness association strengthens with Indigenous share (negative interaction)	Not supported (*ns*)	Not supported (*ns*)	n/a — none made

Note. ns = not statistically significant. For cross-provincial predictions, 'evidence against' indicates pattern-level evidence ran contrary to the predicted difference; 'no clear evidence' indicates the observed patterns could not discriminate for or against it.

Chapter 5: Discussion

This chapter interprets the findings in relation to the literature and provincial contexts developed earlier, with marginalised place as its organising concern. Measures capturing the posed dimensions of marginalised place — Indigenous share and remoteness (5.1), and settlement size (5.2) — are thus discussed first, followed by school type (5.3) and socioeconomic status (5.4). Throughout, consistent with the commitments of Section 3.7, negative associations describe attainment within prevailing educational structures, not the capacities of the communities this study centres (Corbett, 2014; Spencer, 2025).

5.1 Indigenous Share and Remoteness

Indigenous share behaved as hypothesised in both provinces — negatively and significantly associated with attainment net of SES and rurality — and both associations survived the wild cluster bootstrap, making this the study's clearest cross-provincial convergence. The apparent difference in significance ($p < .001$ in BC against $p = .024$ in Alberta) should not be over-read: Alberta's model leaves more variance unexplained and rests on a more concentrated cluster structure, so its estimates are less precise across the board. More notable is what did not appear: H4's expectation that self-governance and cultural continuity would visibly soften the association in BC (Chandler, 2010; Richards, 2013) finds no support at pattern level; if anything, BC's association is more firmly established.

In variance-explained terms, the contribution is modest — $\Delta R^2 \approx .01–.02$ in each province — with Indigenous share adding little variance beyond what SES and rurality already carry; however, that modesty deserves scrutiny rather than acceptance, because the association is almost certainly understated. Suppression based on the number of writers — and, in Alberta, non-participation by band-operated schools — excludes communities where Indigenous share is highest (Section 3.3). Within the retained sample, the upper range of Indigenous share is remarkably thin: only four and two sample DAs in BC and Alberta, respectively, sit near or above 90%. Implications at that extreme therefore rest on thin

ground: a 90% share sits some nine to twelve standard deviations above the provincial means, so the model's predictions there amount to linear extrapolation through a handful of observations, and are unreliable as estimates.

The modest figures are therefore best read as floors, resting on a sample that under-represents the highest-Indigenous-share communities and on estimation dominated by the low end of the share distribution, the range where the association is plausibly weakest. Moreover, the association’s statistical modesty is partly a product of the very marginalisation it measures: the communities the measure most concerns are those the data infrastructure renders least visible. Finally, the DA measure employed adds some error of its own: a school's DA Indigenous share need not mirror its actual enrolment, and the direction of any resulting bias is unknown; however, there was no available alternative for comparative provincial research.

Next, only one of H3's predictions held: while remoteness did retain a significant association through Alberta's final model, its coefficient attenuated to non-significance in BC. Thus, at pattern level, remoteness was not the stronger presence in BC as H3 predicted. However, it would be a mistake to read this as remoteness persisting in one province and vanishing in the other: Alberta's final-model association is modest ($\beta = -0.11$, $p = .037$), BC's sits at $\beta = -0.08$ ($p = .085$), and neither rurality block contributes substantial variance on entry ($\Delta R^2 = .01$ in BC; .02 in Alberta). Regardless, one possibility for the Alberta coefficient — if its small association is not merely noise — is that the remoteness term partly absorbs the labour-market mechanism H2 attached to settlement size, given the settlement-size sign reversal discussed shortly.

In BC, remoteness entered strongly ($\beta = -0.16$, $p < .001$) and attenuated once Indigenous share arrived ($\beta = -0.11$, $p = .013$) — and that attenuation is itself informative. These two dimensions overlap more tightly in BC ($r = .51$, against .39 in Alberta; Section 4.6.1), which the provincial context anticipated: BC's numerous small reserves scatter across the distance spectrum, with several of its highest-Indigenous-share communities at the remoteness extreme. Much of what BC's remoteness association captures may therefore be Indigenous-community disadvantage travelling under the name ‘remoteness’ — the RSS

model's entanglement of the spatial and the social, made empirical. This proposed entanglement is correlational, not conditional: the dimensions travel together, but neither strengthens the other's association, which is precisely what H6 tested.

H6 — a school-level analogue of Layton's (2023) finding that remoteness matters more for Indigenous than non-Indigenous high-school completion — was not supported in either province, and the nulls are precise rather than underpowered: the estimates are near zero with small standard errors ($\beta = 0.01$, SE = 0.03, $p = .716$ in BC; $\beta = -0.02$, SE = 0.03, $p = .445$ in Alberta). This does not contradict Layton, however. Her finding was established at the individual level, measured attainment through completion, and drew on national data; the present study differs on all three points, and divergence could have arisen through any of them. Nothing here invalidates Layton's result; if anything, hers remains the more reliable of the two: her Census individual-level data carry far fewer omissions, and they genuinely describe the Indigenous students concerned — where each lives, whether each graduated — rather than just characterising the areas around their schools.

5.2 Settlement Size

Settlement size produced the study's clearest hypothesis failures. In BC, settlement size was null in every model specification. In Alberta, no association appeared in the form H2 specified (beyond SES); instead, once remoteness was jointly modelled, a significant association emerged in the opposite direction — larger settlements predicting lower attainment, net of SES and remoteness — a reversal that replicates at Grade 9. However, the reversal of settlement size in Alberta invites caution before interpretation: settlement size and remoteness are strongly correlated ($r = -.71$ to $-.73$), and each is null when entered alone (Appendix H) — a structure in which jointly estimated coefficients can turn significant for statistical rather than substantive reasons.

If the reversal reflects a real pattern rather than the mechanics of the split, one candidate explanation concerns settlement classification. Many affluent DAs just outside cities are coded 'rural' (popctr = 1) (Puryear & Kettler, 2017); in this way, they may not capture the mechanisms the Literature Review attached to settlement size: narrower

curricula, specialist staffing shortages, less visible labour-market returns to education. However, robustness checks revealed that even after reclassifying these exurbs, the pattern in both provinces held (Appendix G). For Alberta, another explanation is substantive: the labour-market mechanism H2 attached to settlement size may instead belong to remoteness, with Alberta's persistent — though modest — remoteness association (5.1) carrying what H2 expected settlement size to show. This explains the missing association, but not the reversal: relocating the mechanism should leave settlement size null, and it is not.

Neither province's rurality block was practically large — even Alberta's, the larger, stayed small under any entry order (Section 4.3; Appendices H and I). Nonetheless, the results align with Corbett's (2014) argument that rural places, considered in their circumstances, are not underperformers: BC's settlement-size association is null, and Alberta's runs in rural places' favour. Corbett's argument defending rural places was anchored in Cartwright and Allen's (2002) finding that rural–urban attainment differences disappeared in every province once community variables related to adult educational attainment were controlled for, with a single exception: Alberta. In the present data, this historical exception has inverted: conditional on SES and remoteness, Alberta's settlement-size association now runs in rural places' favour. Finally, these results show that multidimensional rurality measurement reveals more. BC's null and Alberta's conditional reversal are different findings with different (albeit limited) implications — structure that a single 'rural' measure would have collapsed (Roberts et al., 2021).

5.3 School Type

School type explained a substantial share of BC's attainment variation ($\Delta R^2 = .20$) — second only to SES (.25), and nearly its equal under the alternative entry ordering (Appendix I), despite comprising a single binary contrast. The final-model gap is correspondingly large: independent schools scored 1.13 SD above public schools ($p < .001$).

Conditional evidence indicates BC's school-type advantage may be, at least in part, urban advantage. 118 of the sample's 167 independent schools (71%) lie in the Lower Mainland (including Vancouver and its surrounding communities), Greater Victoria, or

Kelowna, with only 49 (29%) spread across the rest of the province (Table K2). Moreover, as the school type by remoteness interaction ($\beta = -0.312$, $p = .003$) indicates, that advantage grows as remoteness falls: the gap is equivalent to a median independent school sitting at roughly the 80th percentile of the public-school distribution in the most remote quartile, and at roughly the 96th in the least remote (Section 4.6.2). While elite preparatory schools are unaffordable for most families wherever they are located, faith-based independent schools are often considerably cheaper (Azimi et al., 2023) but may remain inaccessible for geographic reasons: in a province as vast as BC, families outside urban centres may have no independent school within reach. Without one, the option carrying the model's largest advantage does not exist.

These differences are particularly consequential for equity if they reflect what schools do rather than whom they enrol. While that question cannot be resolved here, wider evidence speaks to the present estimate. The selection account draws its strongest support from Azimi et al. (2023): across three designs — lagged Grade 4 scores, comparisons within micro-neighbourhoods of roughly nineteen households, and children who switched school types — the school-added component in BC's own FSA data was modest (0.12–0.15 student-level SDs). The international pattern points the same way: whether by statistical controls (Dronkers & Robert, 2008) or by comparing only similar students (Dronkers & Avram, 2010), accounting for who attends private schools made the school-type advantage smaller. In this second study, however, Canada was among the few countries where a significant advantage persisted even between similar students. Follow-up work linking student-level intake to school-level gaps, ideally capturing instructional differences directly, is needed. If instruction genuinely differs between school types, independent schools become an immediate equity concern: they offer a real advantage while limiting who can attend by cost and location.

5.4 Socioeconomic Status

In both provinces' final models, parental education was the dominant continuous predictor, while family income fell short of significance at even the most lenient

conventional threshold. H1's within-province expectations are therefore supported for education and unsupported for income — though only suggestively so for education in Alberta, for data-structure reasons discussed shortly. H1's cross-provincial element — a weaker income association, and correspondingly larger education association, in Alberta — cannot be validated: at pattern level the two provinces simply appear similar, with income null in both and education near-identical ($\beta = 0.34$, $p = .032$ in Alberta; $\beta = 0.35$, $p < .001$ in BC).

The income null may stem from issues with the DA income measure itself. As discussed in the Methods, DA income agrees poorly with the incomes of the individuals living there (17% same-quintile assignment, weakest in rural areas, $\kappa = 0.11$; Davis et al., 2023); the school-level unit softens this — the study never claimed DA income matches any household — but does not necessarily rescue the measure. Moreover, another result speaks against a true income null. With a school-type association as large as BC's, produced by schools that draw fee-paying families from well beyond their DAs, it is hard to claim that family income appears non-predictive.

The parental education result carries informative implications: it may indicate that the dimensions of advantage theorised as unique to parental education — aspirations, cultural capital, familiarity with academic institutions (Lareau, 2011) — are at work in both provinces. Unfortunately, while Alberta's SES block explains the most variance in its model by a wide margin, its parental education association failed the wild cluster bootstrap ($p = .129$, against .032 under primary inference). This reflects data structure rather than substance: with half the sample in two metropolitan clusters, the association is identified from very few effective community-level comparisons. It is unfortunate that, in the province whose estimates are already the less certain (Section 4.4), the association that resists firm establishment is the model's largest contributor.

Future work involving student-level SES linkage could escape both these problems at once: it would avoid the present clustering issue and genuinely approximate students' SES. Student-level linkage is especially necessary in BC, where fee-paying school choice is both prominent and, in these results, powerfully associated with attainment. Here, arguably the

study's largest contribution and most consequential measurement limitation converge on the same unmeasured quantity: the household resources of the students actually in each school.

Chapter 6: Conclusion

This dissertation examined how socioeconomic status, each posed dimension of marginalised place — settlement size, remoteness, and Indigenous population share — and school type are associated with school-level attainment in British Columbia and Alberta, with a particular interest in what place contributes once SES is accounted for. It further asked whether these associations depend on one another, and whether they differ at pattern level between the two provinces.

The answers produced, by research question, are as follows. For SES (RQ1): parental education, not family income, was the operative socioeconomic predictor in both provinces. That said, Alberta's parental-education association, failing the wild cluster bootstrap, is established suggestively rather than firmly. Settlement size (RQ2) showed no hypothesis-aligned association, but a small Albertan association emerged indicating larger settlements predicted lower attainment; this association was robust to exurban reclassification and replicated at Grade 9. Remoteness (RQ3) had a modest but persistent association independent of all modelled factors in Alberta, but attenuated to borderline non-significance upon other variables' introduction in BC. Indigenous share (RQ4) was modestly but robustly associated with lower attainment in both provinces, though sampling limitations mean the true association likely differs from what was estimated here. School type (RQ5) produced the largest attainment gap in BC's final model — independent schools scored 1.13 SD above public schools — and contributed substantially to explained variance ($\Delta R^2 = .20$); Alberta's associations were smaller (separate schools) or rested on few effective clusters (private, charter). The hypothesised remoteness by Indigenous share interaction (RQ6, H6) was null in both provinces. Exploratory tests (RQ6) further revealed that the school-type advantage itself was remoteness-dependent in both systems, and did not vary detectably with parental education in BC.

In the Introduction, the stated aim of this study was 'to establish how each posed dimension of marginalised place is associated with school-level attainment net of SES, and whether place matters beyond what socioeconomic circumstance explains'. At the specified

associational level, this aim was largely met. Place clearly has some association with attainment beyond SES: SES offered real but never full explanatory power in either province, and the place dimensions continued to yield significant associations after SES had been accounted for. Evidence also supported the empirical value of the multidimensional rurality operationalisation the Rural Social Space model motivates: in Alberta, measuring settlement size and remoteness jointly yielded two rural associations where the conventional settlement-size-only approach would have found none. Conditional associations indicate something further: place matters not only directly but by shaping how other advantages operate. For example, in BC, access to an independent school was strongly associated with attainment net of measured DA-level SES, but that access is unevenly distributed across the province, and the advantage itself varied with where a school sits.

However, the verdict that place demonstrably had an association with attainment independent of SES is held back by limitations, which largely stem from the census measures employed. In Alberta, the concentrated cluster structure — half the sample's schools fall in just two metropolitan Census Subdivisions (CSDs), the unit at which standard errors are clustered — left the parental-education association, though large in coefficient, short of firm statistical support. Since that association carries nearly all of the SES block's explained variance, its uncertainty means the extent to which the place associations are ‘independent of SES’ is itself less certain. Additionally, the SES measures are all at the DA level, which describes choice-driven schools least well, since their intake is least likely to come from the surrounding DA. In this way, the substantial school-type differences in BC may reflect high-SES selection into independent schools, a possibility DA-level measurement struggles to capture. Lastly, the Indigenous-share estimate, while small, was robustly significant beyond SES and rurality. However, the high-Indigenous-concentration communities the measure most concerns were hardly present in the sample, so the estimate is not necessarily reflective of the true association between high-Indigenous-concentration areas and attainment.

Broader limitations add to these. The 2021 Census precedes the 2024/25 attainment data by three to four years, meaning areas may have changed in the intervening years in ways

the models cannot see. The two assessment instruments also differ in construction and measurement error, which may affect even the pattern-level contrasts drawn between provinces. Both tests suppress results for schools with the fewest writers, a practice that protects privacy but likely removes many schools in marginalised places from view. The work was also not conducted in consultation with Indigenous peoples; this is permissible for public aggregate data (Section 3.7), but it continues a long and unfortunate history of research about, rather than with, these communities.

This study's contribution is descriptive: it maps associations without establishing causality or mechanism. While its practical implications are limited, systematic description of socially significant patterns is a contribution in its own right, especially where phenomena are poorly documented (Loeb et al., 2017). The Literature Review established precisely that: peer-reviewed Canadian rural and Indigenous educational attainment work is meaningfully thin.

Practical implication thus awaits future work building on these patterns. That work should pursue student-level SES linkage, allowing SES's contribution to be estimated cleanly before place's independent share. Indigenous educational research should move beyond the DA measure, toward school-level enrolment data where they exist, and categorical indicators of reserve-adjacent and high-share communities; these measures and designs should be developed in consultation with Indigenous peoples. What this work could eventually accomplish is of substantial scale: a literature able to establish where, and through what structures, place shapes educational opportunity in Canada — and what would change it. Such a literature has yet to emerge, but this study offers a foundation from which it can begin.

Finally, it is important to restate that the attainment differences documented in this dissertation are not the failings of marginalised communities. Even if causal accounts of these differences were eventually established, they could not honestly be framed as deficiencies: the assessment programmes, schools, and data infrastructures through which attainment is defined and measured were built by and for urban Canada (Corbett, 2014; Green & Reid, 2021). However, because these structures were built, they are not inevitable.

What was made by choice can be remade by it. This dissertation is offered as a small contribution to that remaking.

References

Adamson, F., Åstrand, B., & Darling-Hammond, L. (Eds). (2016). *Global Education Reform: How Privatization and Public Investment Influence Education Outcomes (1st edn).* Routledge. https://doi.org/10.4324/9781315680361

Alasia, A., Bédard, F., Bélanger, J., Guimond, E., & Penney, C. (2017). *Measuring remoteness and accessibility: A set of indices for Canadian communities (Catalogue no. 18-001-X).* Statistics Canada. https://www150.statcan.gc.ca/n1/pub/18-001-x/18-001-x2017002-eng.htm

Alberta Act, S.C. 1905, c. 3.

Ananian, S., & Dellaferrera, G. (2024). *Employment and wage disparities between rural and urban areas (ILO Working Paper No. 107).* International Labour Organization. https://doi.org/10.54374/LNZT9066

Anckar, C. (2008). On the Applicability of the Most Similar Systems Design and the Most Different Systems Design in Comparative Research. *International Journal of Social Research Methodology, 11*(5), 389–401. https://doi.org/10.1080/13645570701401552

Anckar, C. (2020). *The most-similar and most-different systems design in comparative policy analysis.* In B. G. Peters & G. Fontaine (Eds), Handbook of Research Methods and Applications in Comparative Policy Analysis. Edward Elgar Publishing. https://doi.org/10.4337/9781788111195.00008

Andres, L., & Looker, E. D. (2001). Rurality and Capital: Educational Expectations and Attainments of Rural, Urban/Rural and Metropolitan youth. *Canadian Journal of Higher Education, 31*(2), 1–45. https://doi.org/10.47678/cjhe.v31i2.183387

Areepattamannil, S., & Lee, D. H. L. (2014). Linking Immigrant Parents' Educational Expectations and Aspirations to Their Children's School Performance. *The Journal of Genetic Psychology, 175*(1), 51–57. https://doi.org/10.1080/00221325.2013.799061

Asadolahi, S., Farney, J., Triadafilopoulos, T., & White, L. A. (2022). Charting the Rise of School Choice across Canadian Provinces: A Policy Index. *Canadian Journal of Political Science, 55*(1), 188–207. https://doi.org/10.1017/S0008423921000901

Avvisati, F. (2020). The measure of socio-economic status in PISA: A review and some suggested improvements. *Large-Scale Assessments in Education, 8*(1), 8. https://doi.org/10.1186/s40536-020-00086-x

Azimi, E., Friesen, J., & Woodcock, S. (2023). Private Schools and Student Achievement. *Education Finance and Policy, 18*(4), 623–653. https://doi.org/10.1162/edfp_a_00405

Bacic, R., & Zheng, A. (2022). *Income-achievement gaps in Canada* (Department of Economics Working Paper No. 2022-04). McMaster University. http://socialsciences.mcmaster.ca/econ/rsrch/papers/archive/2022-04.pdf

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy, 18*(2), 215–228. https://doi.org/10.1080/0268093022000043065

Basque, M., & Bouchamma, Y. (2013). Academic Achievement in Effective Schools. *Alberta Journal of Educational Research, 59*(3), 503–519. https://doi.org/10.55016/ojs/ajer.v59i3.55790

BC Stats. (2023). *BC's gross domestic product (GDP) by industry: 2023*. https://www2.gov.bc.ca/assets/gov/data/statistics/economy/economic-other/gdp_by_industry_2023.pdf

Bæck, U.-D. K. (2016). Rural Location and Academic Success—Remarks on Research, Contextualisation and Methodology. *Scandinavian Journal of Educational Research, 60*(4), 435–448. https://doi.org/10.1080/00313831.2015.1024163

Beauvais, E. (2022). The Political Consequences of Indigenous Resentment. *Journal of Race, Ethnicity, and Politics, 7*(1), 37–64. https://doi.org/10.1017/rep.2020.25

Beauvais, E., & Stolle, D. (2022). The Politics of White Identity and Settlers' Indigenous Resentment in Canada. *Canadian Journal of Political Science, 55*(1), 59–83. https://doi.org/10.1017/S0008423921000986

Bleakney, A., & Melvin, A. (2022). *Indigenous women and girls: Socioeconomic conditions in remote communities compared with more accessible areas* (Insights on Canadian Society, Catalogue no. 75-006-X). Statistics Canada.

Boerema, A. J. (2009). Does Mission Matter? An Analysis of Private School Achievement Differences. *Journal of School Choice, 3*(2), 112–137. https://doi.org/10.1080/15582150902996708

Bombay, A., Matheson, K., & Anisman, H. (2014). The intergenerational effects of Indian Residential Schools: Implications for the concept of historical trauma. *Transcultural Psychiatry, 51*(3), 320–338. https://doi.org/10.1177/1363461513503380

Boo, S., & Froelicher, E. S. (2013). Secondary analysis of national survey datasets. *Japan Journal of Nursing Science, 10*(1), 130–135. https://doi.org/10.1111/j.1742-7924.2012.00213.x

Boone, S., & Van Houtte, M. (2016). More Ambitious Educational Choices in Urban Areas: A Matter of Local Labor Market Characteristics? *Urban Education, 51*(8), 940–963. https://doi.org/10.1177/0042085914549363

Borbely, D., Gehrsitz, M., McIntyre, S., Rossi, G., & Roy, G. (2024). Rurality, socio-economic disadvantage and educational mobility: A Scottish case study. *British Educational Research Journal, 50*(1), 162–182. https://doi.org/10.1002/berj.3917

Bosetti, L., & Butterfield, P. (2016). The politics of educational reform: The Alberta charter school experiment 20 years later. *Global Education Review, 3*(2), 26–46.

British Columbia School Trustees Association. (2024). *Interim report from Rural and Remote Network*. BCSTA.

Brown, R. S., & Parekh, G. (2022). *The secondary to postsecondary transition process in Canadian public education: Trends and new shifts.* Proceedings of the Canada International Conference on Education (CICE-2022) and World Congress on Education (WCE-2022). https://doi.org/10.20533/cice.2022.0026

Brown, R., & Parekh, G. (2022, February 9). *Creating a university 'graduation predictor'.* Higher Education Quality Council of Ontario. https://heqco.ca/creating-a-university-graduation-predictor/

Bukodi, E., & Goldthorpe, J. H. (2013). Decomposing ‘Social Origins’: The Effects of Parents’ Class, Status, and Education on the Educational Attainment of Their Children. *European Sociological Review, 29*(5), 1024–1039. https://doi.org/10.1093/esr/jcs079

Byun, S., Meece, J. L., & Irvin, M. J. (2012). Rural-Nonrural Disparities in Postsecondary Educational Attainment Revisited. *American Educational Research Journal, 49*(3), 412–437. https://doi.org/10.3102/0002831211416344

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics, 90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Cao, G., & Huo, G. (2026). Reassessing urban-rural education disparities: Evidence from England. *Educational Studies, 52*(2), 189–208. https://doi.org/10.1080/03055698.2025.2478831

Capano, G. (2015). Federal Dynamics of Changing Governance Arrangements in Education: A Comparative Perspective on Australia, Canada and Germany. *Journal of Comparative Policy Analysis: Research and Practice, 17*(4), 322–341. https://doi.org/10.1080/13876988.2014.952530

Caro, D., McDonald, J., & Willms, J. (2009). Socio-economic Status and Academic Achievement Trajectories from Childhood to Adolescence. *Canadian Journal of Education/Revue Canadienne de l'éducation, 32*(3), 558-590-558-590. https://cje-rce.ca/index.php/cje-rce/article/view/3052

Cartwright, F., & Allen, M. K. (2002). *Understanding the rural–urban reading gap (Catalogue no. 81-595-MIE2002001; Education, Skills and Learning Research Papers).* Statistics Canada.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A

meta-analysis. *Educational Research Review, 14*, 33–46. https://doi.org/10.1016/j.edurev.2015.01.002

Chandler, M. J. (2010). Social determinants of educational outcomes in Indigenous learners. *Education Canada, 50*(2), 46–50.

Chmielewski, A. K. (2019). The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015. *American Sociological Review, 84*(3), 517–544. https://doi.org/10.1177/0003122419847165

Christofides, L. N., Hoy, M., Milla, J., & Stengos, T. (2012). The Implication of Peer and Parental Influences on University Attendance: A Gender Comparison. *University of Guelph, Department of Economics and Finance Working Paper Series, 2012*(01), 1–44. https://uoguelph.ca

Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education (8th edn).* Routledge. https://doi.org/10.4324/9781315456539

Constitution Act, 1982, Pub. L. No. C. 11, Being Schedule B to the Canada Act 1982 (UK) (1982). https://justice.gc.ca

Cooper, G., Strathdee, R., & Baglin, J. (2018). Examining geography as a predictor of students' university intentions: A logistic regression analysis. *Rural Society, 27*(2), 83–93. https://doi.org/10.1080/10371656.2018.1472909

Corbett, M. (2007). *Learning to leave: The irony of schooling in a coastal community.* Fernwood Publishing.

Corbett, M. (2014). Toward a Geography of Rural Education in Canada. *Canadian Journal of Education, 37*(3). https://eric.ed.gov/?id=EJ1057952

Corbett, M. (2025). *A reflection on methodological considerations in my own rural education research.* In J. D. Johnson & H. L. Harmon (Eds), Handbook on Rural and Remote Education (pp. 492–507). Edward Elgar Publishing. https://doi.org/10.4337/9781035307722.00042

Corbett, M., & Gereluk, D. (2020). Introduction: Why rural matters in Canadian education. In M. Corbett & D. Gereluk (Eds.), *Rural teacher education: Connecting land and people* (pp. 3–17). Springer Singapore. https://doi.org/10.1007/978-981-15-2560-5_1

Council of Ministers of Education, Canada. (n.d.). *Pan-Canadian Assessment Program (PCAP).* https://www.cmec.ca/240/Pan-Canadian_Assessment_Program_(PCAP).html

Cowley, P., Emes, J., & Shang, M. (2025). *Report card on British Columbia's elementary schools 2025.* Fraser Institute. https://www.fraserinstitute.org/studies/report-card-british-columbias-elementary-schools-2025

Davis, L. E., Mahar, A. L., & Strumpf, E. C. (2023). Agreement between individual and neighborhood income measures in patients with colorectal cancer in Canada. *JNCI: Journal of the National Cancer Institute, 115*(5), 514–522. https://doi.org/10.1093/jnci/djad017

Dean, J., Roberts, P., Downes, N., & Goldsmith, A. (2023). The spatial implications of academic achievement in Year 12: Rethinking discourses of disadvantage in rural locations. *Australian Journal of Education, 67*(1), 94–114. https://doi.org/10.1177/00049441231155708

della Porta, D. (2008). Comparative analysis: Case-oriented versus variable-oriented research. In D. della Porta & M. Keating (Eds.), *Approaches and methodologies in the social sciences: A pluralist perspective* (pp. 198–222). Cambridge University Press. https://doi.org/10.1017/CBO9780511801938.012

Dickerson, A., Maragkou, K., & McIntosh, S. (2018). *The causal effect of secondary school peers on educational aspirations (CVER Discussion Paper No. 017).* Centre for Vocational Education Research, London School of Economics. https://cver.lse.ac.uk/textonly/cver/pubs/cverdp017.pdf

Drescher, J., Podolsky, A., Reardon, S. F., & Torrance, G. (2022). The Geography of Rural Educational Opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 8*(3), 123–149. https://doi.org/10.7758/RSF.2022.8.3.05

Dronkers, J., & Avram, S. (2010). A cross-national analysis of the relations of school choice and effectiveness differences between private-dependent and public schools. *Educational Research and Evaluation, 16*(2), 151–175. https://doi.org/10.1080/13803611.2010.484977

Dronkers, J., & Robert, P. (2008). Differences in Scholastic Achievement of Public, Private Government-Dependent, and Private Independent Schools: A Cross-National Analysis. *Educational Policy, 22*(4), 541–577. https://doi.org/10.1177/0895904807307065

Duman, E. (2021). The impacts of family background on children's educational attainment and earnings. *Journal of Politics Economy and Management, 4*(1), 25–37. https://dergipark.org.tr/en/pub/jopem/article/908901

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2014). Boosting Family Income to Promote Child Development. *The Future of Children, 24*(1), 99–120. https://doi.org/10.1353/foc.2014.0008

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2015). *Children and socioeconomic status.* In R. M. Lerner (Ed.), Handbook of Child Psychology and Developmental Science (7th ed., Vol. 4). Wiley. https://doi.org/10.1002/9781118963418.childpsy414

Echazarra, A., & Radinger, T. (2019). *Learning in rural schools: Insights from PISA, TALIS and the literature* (OECD Education Working Papers No. 196; OECD Education Working Papers, Vol. 196). https://doi.org/10.1787/8b1a5cb9-en

Edgerton, J., Peter, T., & Roberts, L. (2008). Back to the Basics: Socio-Economic, Gender, and Regional Disparities in Canada's Educational System. *Canadian Journal of Education/Revue Canadienne de l'éducation, 31*(4), 861–888. https://cje-rce.ca/index.php/cje-rce/article/view/3025

Education Analytics Office. (2025). *Aboriginal report – How are we doing?* British Columbia Ministry of Education and Child Care. https://www2.gov.bc.ca/gov/content/education-training/k-12/support/aboriginal-students-hawd

Emery, J. C. H., Ferrer, A., & Green, D. (2012). Long-Term Consequences of Natural Resource Booms for Human Capital Accumulation. *ILR Review, 65*(3), 708–734. https://doi.org/10.1177/001979391206500310

Faíña, J. A., & López-Rodríguez, J. (2006). Market access and human capital accumulation: The European Union case. *Applied Economics Letters, 13*(9), 563–567. https://doi.org/10.1080/13504850500401684

Ferede, E. (2018a). Alberta's Fiscal Responses to Fluctuations in Non-Renewable-Resource Revenue. *SPP Briefing Papers, 11*(23). https://ideas.repec.org//a/clh/briefi/v11y2018i23.html

Ferede, E. (2018b). *Alberta's fiscal responses to fluctuations in non-renewable-resource revenue.* University of Calgary School of Public Policy. https://www.policyschool.ca/wp-content/uploads/2018/09/NRR-Ferede.pdf

Ferguson, N. (2011). *From Coal Pits to Tar Sands: Labour Migration Between an Atlantic Canadian Region and the Athabasca Oil Sands.* Just Labour. https://doi.org/10.25071/1705-1436.35

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics (5th edn).* SAGE Publications Inc.

First Nations Leadership Council. (2021, November 18). *Re: Foundation Skills Assessment* [Letter to Hon. Jennifer Whiteside, Minister of Education]. https://www.sd61.bc.ca/wp-content/uploads/sites/91/2021/12/2021-11-17_FNLC_LTR_to_Whiteside_RE_FSA-1.pdf

Fischer, A., & Roodman, D. (2021). *fwildclusterboot: Fast wild cluster bootstrap inference for linear models[Computer software].* Comprehensive R Archive Network (CRAN). https://cran.r-project.org/package=fwildclusterboot

Forsey, M., Davies, S., & Walford, G. (2008). *The Globalisation of School Choice?* Symposium Books.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models (3rd edn).* Sage.

Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association, 87*(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190

Friesen, J., & Krauth, B. (2010). Sorting, peers, and achievement of Aboriginal students in British Columbia. *Canadian Journal of Economics/Revue Canadienne d'économique, 43*(4), 1273–1301. https://doi.org/10.1111/j.1540-5982.2010.01614.x

Fulkerson, G. M., & Thomas, A. R. (2016). *Reimagining Rural: Urbanormative Portrayals of Rural Life.* Lexington Books.

Fung, K., & Chu, M.-W. (2015). Fairness of standardized assessments: Discrepancy between provincial and territorial results. *Journal of Contemporary Issues in Education, 10*(1), 2–24. https://doi.org/10.20355/C5KG6P

Gaskell, J., & Levin, B. (2011). *The Challenges of Poverty and Urban Education in Canada: Lessons from Two School Boards.* In C. Raffo, A. Dyson, H. Gunter, D. Hall, L. Jones, & A. Kalambouka (Eds), Education and Poverty in Affluent Countries (pp. 148–160). Routledge.

Gayle, G.-L., Golan, L., & Soytas, M. A. (2018). Intergenerational mobility and the effects of parental education, time investment, and income on children's educational attainment. Federal Reserve Bank of St. *Louis Review, 100*(3), 281–295. https://doi.org/10.20955/r.100.281-95

Gerring, J. (2017). *Case Study Research: Principles and Practices.* Cambridge University Press.

Gill, R., Karim, M. E., Puyat, J. H., Petteni, M. G., Guhn, M., Janus, M., Forer, B., & Gadermann, A. (2025). Poverty type, immigration background, and secondary school academic outcomes for children in British Columbia. *Social Psychology of Education, 28*(1), 167. https://doi.org/10.1007/s11218-025-10093-x

Giraudy, A., Moncada, E., & Snyder, R. (Eds). (2019). *Inside Countries: Subnational Research in Comparative Politics(1st edn).* Cambridge University Press. https://doi.org/10.1017/9781108678384

Glenlyon Norfolk School. (n.d.). *Tuition & other costs.* https://www.mygns.ca/admissions/tuition-and-financial-support/tuition-and-other-costs/

Gordon, C. E., & White, J. P. (2014). Indigenous Educational Attainment in Canada. *International Indigenous Policy Journal, 5*(3). https://doi.org/10.18584/iipj.2014.5.3.6

Government of Alberta. (2023). *Economic spotlight: 2022 GDP by industry*. Treasury Board and Finance. https://open.alberta.ca/dataset/13ab3f73-6e4e-4aac-b56b-bff38800aa65/resource/44df3fc1-b4f4-49c5-8aa9-1ffb62a49060/download/tbf-economic-spotlight-2023-07.pdf

Government of Alberta. (2024). *Alberta mining and oil and gas extraction industry profile, 2023*. Jobs, Economy and Trade. https://open.alberta.ca/dataset/f4f39b9e-48cb-4f6a-b491-25ee6f9c281e/resource/ebdc

160e-9a8a-459b-94dc-b3b2fedf882e/download/jet-alberta-mining-and-oil-and-gas-extraction-industry-profile-2023.pdf

Government of Alberta. (n.d.-a). *Education services agreement standards*. https://www.alberta.ca/education-services-agreement-standards

Government of Alberta. (n.d.-b). *Provincial achievement tests*. https://www.alberta.ca/provincial-achievement-tests

Government of British Columbia. (n.d.). *Foundation Skills Assessment (FSA)*. https://www2.gov.bc.ca/gov/content/education-training/k-12/administration/program-management/assessment/foundation-skills-assessment

Green, B., & Reid, J.-A. (2021). *Rural Social Space: A Conceptual-Analytical Framework for Rural (Teacher) Education and the Rural Human Services.* In P. Roberts & M. Fuqua (Eds), Ruraling Education Research (pp. 29–46). Springer Singapore. https://doi.org/10.1007/978-981-16-0131-6_3

Griffith, D. A. (1994). *A Northern Territory approach to quantifying 'access disadvantage' to educational services in remote and rural Australia (D.* McSwan & M. McShane, Eds; pp. 311–314). James Cook University.

Gristy, C. (2019). Journeys to school in rural places: Engaging with the troubles through assemblages. *Journal of Rural Studies, 72*, 286–292. https://doi.org/10.1016/j.jrurstud.2019.10.016

Guenther, J. (2013). *Towards educational advantage in very remote Australia: An analysis of 2012 NAPLAN data [Paper presentation].* Australian Association for Research in Education Annual Conference, Adelaide, Australia.

Hahmann, T., Lee, H., & Godin, S. (2024). *Indigenous foster children living in private households: Rates and sociodemographic characteristics of foster children and their households.* Statistics Canada. https://www150.statcan.gc.ca/n1/pub/41-20-0002/412000022024001-eng.htm

Harris, C. (2002). *Making Native Space: Colonialism, Resistance, and Reserves in British Columbia.* UBC Press.

Hawley, L. R., Koziol, N. A., Bovaird, J. A., McCormick, C. M., Welch, G. W., Arthur, A. M., & Bash, K. (2016). Defining and describing rural: Implications for rural special education research and policy. *Rural Special Education Quarterly, 35*(3), 3–11. https://doi.org/10.1177/875687051603500302

Ho, A. D. (2009). A Nonparametric Framework for Comparing Trends and Gaps Across Tests. *Journal of Educational and Behavioral Statistics, 34*(2), 201–228. https://doi.org/10.3102/1076998609332755

Hung, J. (2022). Hukou System Influencing the Structural, Institutional Inequalities in China: The Multifaceted Disadvantages Rural Hukou Holders Face. *Social Sciences, 11*(5), 194. https://doi.org/10.3390/socsci11050194

Hunt, D., & DeJong VanHof, J. (2024). *Exploring Alberta's independent school landscape: Diversity, growth, and trends.* Cardus. https://www.cardus.ca/research/education/reports/exploring-albertas-independent-school-landscape/

Indigenous Services Canada. (2024). *Former Indian residential schools – Environmental scan: Status of sites and buildings.* Government of Canada. https://sac-isc.gc.ca/eng/1719411519382/1719411537769

Irvin, M., Byun, S., Smiley, W. S., & Hutchins, B. C. (2017). Relation of Opportunity to Learn Advanced Math to the Educational Attainment of Rural Youth. *American Journal of Education, 123*(3), 475–510. https://doi.org/10.1086/691231

Island Catholic Schools. (n.d.). *Admissions*. https://cisdv.bc.ca/admissions

Jones, M. E. C. (2024). The Intergenerational Legacy of Indian Residential Schools. *Demography, 61*(6), 1871–1895. https://doi.org/10.1215/00703370-11679677

Jónsson, Þ. A. (2019). Er búsetumunur á námsárangri þegar ólík þjóðfélagsstaða er tekin með í reikninginn? *Tímarit Um Uppeldi Og Menntun, 28*(1). https://doi.org/10.24270/tuuom.2019.28.4

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives, 16*(4), 91–114. https://doi.org/10.1257/089533002320950993

Kaushal, N. (2014). Intergenerational Payoffs of Education. *The Future of Children, 24*(1), 61–78. https://doi.org/10.1353/foc.2014.0005

Kneebone, R. D., & Gres, M. (2013). *A recovery program for Alberta: A 10-year plan to end the addiction to resource revenues.* University of Calgary School of Public Policy. https://ideas.repec.org/a/clh/resear/v6y2013i11.html

Kneebone, R. D., & Wilkins, M. (2018). *50 years of Government of Alberta budgeting.* University of Calgary School of Public Policy.

https://www.policyschool.ca/wp-content/uploads/2018/10/50-Years-AB-Budget-Kneebone-Wilkins.pdf

Knipe, S., & Bottrell, C. (2023). Staffing remote schools: Perennial failure. *Journal of Global Education and Research, 7*(2), 183–198. https://doi.org/10.5038/2577-509X.7.2.1197

Korzinski, D. (2021, October 19). *Kids, Schools and Racism in Canada.* Angus Reid Institute. https://angusreid.org/canada-school-kids-racism-diversity/

Lareau, A. (2011). *Unequal Childhoods: Class, Race, and Family Life (2nd ed.).* University of California Press.

Layton, J. (2023). *Distance as a factor for First Nations, Métis, and Inuit high school completion* (Education, Learning and Training: Research Paper Series, Catalogue no. 81-595-M). Statistics Canada.

Lazatin, E. (2017). *Hide FSA results so that outside groups don't rank schools: B.C. school trustees*. https://globalnews.ca/news/3796174/b-c-school-trustees-urge-education-minister-to-hide-standardized-test-results/

Leckie, G., & Goldstein, H. (2017). The evolution of school league tables in England 1992–2016: 'Contextual value-added', 'expected progress' and 'progress 8'. *British Educational Research Journal, 43*(2), 193–212. https://doi.org/10.1002/berj.3264

Li, J., Xue, E., & Guo, S. (2025). The effects of PISA on global basic education reform: A systematic literature review. *Humanities and Social Sciences Communications, 12,* Article 106. https://doi.org/10.1057/s41599-025-04403-z

Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). *Descriptive analysis in education: A guide for researchers* (NCEE 2017–4023). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/20174023/

Looker, E. D., & Bollman, R. D. (2020). *Setting the Stage: Overview of Data on Teachers and Students in Rural and Urban Canada.* In M. Corbett & D. Gereluk (Eds), Rural Teacher Education (pp. 21–73). Springer Singapore. https://doi.org/10.1007/978-981-15-2560-5_2

Ludeke, S. G., Gensowski, M., Junge, S. Y., Kirkpatrick, R. M., John, O. P., & Andersen, S. C. (2021). Does parental education influence child educational outcomes? A developmental analysis in a full-population sample and adoptee design. *Journal of Personality and Social Psychology, 120*(4), 1074–1090. https://doi.org/10.1037/pspp0000314

Lyeo, J. S., Wong, M. D., Clyke, N., Big Canoe, B., Kinnear, P., Stopps, H., Spence, N. D., & Haines, S. R. (2024). Ten questions concerning First Nations on-reserve housing in Canada. *Building and Environment, 257*, 111544. https://doi.org/10.1016/j.buildenv.2024.111544

Ma, X., & Crocker, R. (2007). Provincial Effects on Reading Achievement. *Alberta Journal of Educational Research, 53*(1). https://doi.org/10.55016/ojs/ajer.v53i1.55201

MacKinnon, J. G., & Webb, M. D. (2017). Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics, 32*(2), 233–254. https://doi.org/10.1002/jae.2508

Marks, G. N., & O'Connell, M. (2021). Inadequacies in the SES–Achievement model: Evidence from PISA and other studies. *Review of Education, 9*(3), e3293. https://doi.org/10.1002/rev3.3293

Martinez, N. T., Xerxa, Y., Law, J., Serdarevic, F., Jansen, P. W., & Tiemeier, H. (2022). Double advantage of parental education for child educational achievement: The role of parenting and child intelligence. *European Journal of Public Health, 32*(5), 690–695. https://doi.org/10.1093/eurpub/ckac044

Mayor, C., & Suarez, E. B. (2019). A scoping review of the demographic and contextual factors in Canada's educational opportunity gaps. *Canadian Journal of Education, 42*(1), 42–87.

McQuillan, K. (2024). *Leaving the big city: New patterns of migration in Canada.* University of Calgary School of Public Policy. https://journalhosting.ucalgary.ca/index.php/sppp/article/view/78322

Melvin, A. (2023). *Postsecondary educational attainment and labour market outcomes among Indigenous peoples in Canada, findings from the 2021 Census.* Statistics Canada. https://www150.statcan.gc.ca/n1/pub/75-006-x/2023001/article/00012-eng.htm

Michael, D., & Kyriakides, L. (2023). Mediating effects of motivation and socioeconomic status on reading achievement: A secondary analysis of PISA 2018. *Large-Scale Assessments in Education, 11*(1), 31. https://doi.org/10.1186/s40536-023-00181-9

Midouhas, E., & Flouri, E. (2015). Rural/Urban Area Differences in the Cognitive Abilities of Primary School Children in England: Rurality/Urbanity and Child Ability in England. *Population, Space and Place, 21*(2), 157–170. https://doi.org/10.1002/psp.1815

Mitra, S., Dangwal, R., & Thadani, L. (2008). Effects of remoteness on the quality of education: A case study from North Indian schools. *Australasian Journal of Educational Technology, 24*(2). https://doi.org/10.14742/ajet.1219

Morissette, R., Chan, P. C. W., & Lu, Y. (2015). Wages, Youth Employment, and School Enrollment: Recent Evidence from Increases in World Oil Prices. *Journal of Human Resources, 50*(1), 222–253. https://doi.org/10.3368/jhr.50.1.222

Mulcahy, D. M. (2009). Developing government policies for successful rural education in Canada. In T. Lyons, J.-Y. Choi, & G. McPhan (Eds.), *Innovation for equity in rural education: Symposium proceedings* (pp. 23–32). University of New England.

Nissinen, K., Ólafsson, R. F., Rautopuro, J., Halldórsson, A. M., & Vettenranta, J. (2018). *The urban advantage in education?* Science achievement differences between metropolitan and other areas in Finland and Iceland in PISA 2015. In Northern lights on TIMSS and PISA 2018 (pp. 183–212). Nordic Council of Ministers. https://doi.org/10.6027/TN2018-524

OECD. (2013). *What makes urban schools different?* (PISA in Focus No. 28). OECD Publishing. https://doi.org/10.1787/5k46l8w342jc-en

OECD. (2023a). *Labour and skills demand in Alberta: Insights using big data intelligence.* OECD Publishing. https://doi.org/10.1787/659ce346-en

OECD. (2023b). *OECD Regional Outlook 2023: The Longstanding Geography of Inequalities.* OECD Publishing. https://doi.org/10.1787/92cd40a0-en

OECD. (2023c). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education.* OECD Publishing. https://doi.org/10.1787/53f23881-en

Papillon, M. (2024). Comparing Policy Feedback Effects in Federal Systems: The Case of Provincial Indigenous Consultation Policies in Canada. *Journal of Comparative Policy Analysis: Research and Practice*, *26*(6), 686–705. https://doi.org/10.1080/13876988.2024.2386363

Perales, F., Johnstone, M., Xiang, N., & Tomaszewski, W. (2023). Explaining achievement gaps between students from regional and metropolitan areas: Accounting for socio-demographic and school climate factors. *Australian Journal of Education*, *67*(1), 76–93. https://doi.org/10.1177/00049441231152943

Perry, L. B., Saatcioglu, A., & Mickelson, R. A. (2022). Does school SES matter less for high-performing students than for their lower-performing peers? A quantile regression analysis of PISA 2018 Australia. *Large-Scale Assessments in Education*, *10*(1), 17. https://doi.org/10.1186/s40536-022-00137-5

Peters, E. J. (1998). Subversive Spaces: First Nations Women and the City. *Environment and Planning D: Society and Space*, *16*(6), 665–685. https://doi.org/10.1068/d160665

Pons, X. (2017). Fifteen years of research on PISA effects on education governance: A critical review. *European Journal of Education, 52*(2), 131–144. https://doi.org/10.1111/ejed.12213

Province of British Columbia. (2023). *British Columbia labour market outlook: 2023 edition.* WorkBC. https://www.workbc.ca/research-labour-market/bc-labour-market-outlook

Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. Wiley-Interscience.

Pulchny, A., Robson, K., & Brown, R. S. (2025). Who transitions into post-secondary education and why? A systematic review of the Canadian landscape. *Social Sciences & Humanities Open, 11*, 101289. https://doi.org/10.1016/j.ssaho.2025.101289

Puryear, J. S., & Kettler, T. (2017). Rural Gifted Education and the Effect of Proximity. *Gifted Child Quarterly, 61*(2), 143–152. https://doi.org/10.1177/0016986217690229

Pustejovsky, J. E., & Tipton, E. (2018). Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models. *Journal of Business & Economic Statistics, 36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004

Reid, J.-A., Green, B., Cooper, M., Hastings, W., Lock, G., & White, S. (2010). Regenerating Rural Social Space? Teacher Education for Rural—Regional Sustainability. *Australian Journal of Education, 54*(3), 262–276. https://doi.org/10.1177/000494411005400304

Reimer, D., Sortkear, B., Oskarsson, M., Nilsen, T., Rasmusson, M., & Nissinen, K. (2018). *Northern Lights on TIMSS and PISA 2018 (2018:524).* Nordic Council of Ministers. https://doi.org/10.6027/TN2018-524

Richards, J. (2013). *Why is BC best?* The role of provincial and reserve school systems in explaining Aboriginal student performance (Commentary No. 390). C.D. Howe Institute. https://cdhowe.org/publication/why-bc-best-role-provincial-and-reserve-school-systems-explaining-aboriginal-student-performance/

Richards, J., Hove, J., & Afolabi, K. (2008). *Understanding the Aboriginal/non-Aboriginal gap in student performance: Lessons from British Columbia (Commentary No. 276).* C.D. Howe Institute. https://www.cdhowe.org/pdf/commentary_276.pdf

Roberts, P., & Fuqua, M. (Eds). (2021). *Ruraling Education Research: Connections Between Rurality and the Disciplines of Educational Research.* Springer Singapore. https://doi.org/10.1007/978-981-16-0131-6

Roberts, P., & Guenther, J. (2021). Framing rural and remote: Key issues, debates, definitions, and positions in constructing rural and remote disadvantage. In P. Roberts & M. Fuqua (Eds.), *Ruraling education research: Connections between rurality and the disciplines of educational research* (pp. 13–27). Springer Singapore. https://doi.org/10.1007/978-981-16-0131-6_2

Roberts, P., Thier, M., & Beach, P. (2021). *Erasing Rurality: On the Need to Disaggregate Statistical Data.* In P. Roberts & M. Fuqua (Eds), Ruraling Education Research (pp. 107–127). Springer Singapore. https://doi.org/10.1007/978-981-16-0131-6_8

Rojas Apaza, R., Paredes, R. P., Arpi, R., Quispe Lino, C. N., & Chura-Zea, E. (2024). Urban-rural gap in education performance in Peruvian public institutions during

2018: An analysis using the Oaxaca-Blinder decomposition. *Frontiers in Education, 9*, 1394938. https://doi.org/10.3389/feduc.2024.1394938

Schneider, B., Kim, S., & Klager, C. (2017). Co-Development of Education Aspirations and Postsecondary Enrollment Especially Among Students Who Are Low Income and Minority. *Research in Human Development, 14*(2), 143–160. https://doi.org/10.1080/15427609.2017.1305811

Schneider, D., Hastings, O. P., & LaBriola, J. (2018). Income Inequality and Class Divides in Parental Investments. *American Sociological Review, 83*(3), 475–507. https://doi.org/10.1177/0003122418772034

Schulz, W. (2005, April). *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003* [Paper presentation]. Annual Meeting of the American Educational Research Association, San Francisco, CA, United States. https://files.eric.ed.gov/fulltext/ED493510.pdf

Scott, D., & Louie, D. (2020). Reconsidering rural education in the light of Canada's Indigenous reality. In M. Corbett & D. Gereluk (Eds.), *Rural teacher education: Connecting land and people* (pp. 113–135). Springer. https://doi.org/10.1007/978-981-15-2560-5_6

Seawright, J., & Gerring, J. (2008). Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options. *Political Research Quarterly, 61*(2), 294–308. https://doi.org/10.1177/1065912907313077

Shang, M., & Cowley, P. (2024). *Report card on Alberta's elementary schools 2024.* Fraser Institute. https://www.fraserinstitute.org/studies/report-card-on-albertas-elementary-schools-20

Shucksmith, M. (2012). Class, Power and Inequality in Rural Areas: Beyond Social Exclusion? *Sociologia Ruralis*, *52*(4), 377–397. https://doi.org/10.1111/j.1467-9523.2012.00570.x

Siemens, E. (2023). Multi-Graded, Rural Alberta Schools: Quality Education or Not? *Alberta Journal of Educational Research, 69*(2), 167–181. https://doi.org/10.55016/ojs/ajer.v69i2.74854

Sjøberg, S. (2015). PISA and Global Educational Governance – A Critique of the Project, its Uses and Implications. *EURASIA Journal of Mathematics, Science and Technology Education, 11*(1). https://doi.org/10.12973/eurasia.2015.1310a

Smith, C., Parr, N., & Muhidin, S. (2019). Mapping schools' NAPLAN results: A spatial inequality of school outcomes in Australia. *Geographical Research, 57*(2), 133–150. https://doi.org/10.1111/1745-5871.12317

Smith, E. (2008). *Using Secondary Data in Educational and Social Research.* Open University Press.

Smith, E. (2009). What can secondary data analysis tell us about school exclusions in England? *International Journal of Research & Method in Education, 32*(1), 89–101. https://doi.org/10.1080/17437270902749304

Smyth, E., & Darmody, M. (2026). The impact of school social composition and neighborhood social mix on upper secondary exam performance in Ireland.

International Journal of Comparative Sociology, 67(1), 119–136. https://doi.org/10.1177/00207152231223725

Snyder, R. (2001). Scaling Down: The Subnational Comparative Method. *Studies in Comparative International Development, 36*(1), 93–110. https://doi.org/10.1007/BF02687586

Spencer, N. (2025). *Neighbourhood effects and education policy in Indigenous communities.* SSRN. https://doi.org/10.2139/ssrn.5574931

Statistics Canada. (2021). *Census of population, 2021* [Data set]. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/index-eng.cfm

Statistics Canada. (2022). *Geographic attribute file, 2021 Census* (Catalogue no. 92-151-X) [Data set]. https://www150.statcan.gc.ca/n1/en/catalogue/92-151-X

Sullivan, K., McConney, A., & Perry, L. B. (2018). A Comparison of Rural Educational Disadvantage in Australia, Canada, and New Zealand Using OECD's PISA. *Sage Open, 8*(4), 2158244018805791. https://doi.org/10.1177/2158244018805791

Tan, C. Y. (2017). Examining Cultural Capital and Student Achievement: Results of A Meta-Analytic Review. *Alberta Journal of Educational Research, 63*(2), 139–159. https://doi.org/10.55016/ojs/ajer.v63i2.56285

Taylor, A. (2001). *The Politics of Educational Reform in Alberta.* University of Toronto Press.

The Canadian Encyclopedia. (2011). *Reserves in Canada.* Historica Canada. https://www.thecanadianencyclopedia.ca/en/article/aboriginal-reserves

Thier, M., Longhurst, J. M., Grant, P. D., & Hocking, J. E. (2021). *Research Deserts: A Systematic Mapping Review of U.S.* Rural Education Definitions and Geographies. Journal of Research in Rural Education, Volume 37. https://doi.org/10.26209/JRRE3702

Truth and Reconciliation Commission of Canada. (2015). *Honouring the truth, reconciling for the future: Summary of the final report of the Truth and Reconciliation Commission of Canada*. https://nctr.ca/records/reports/

Ungerleider, C. (2022, November 29). *First Nations leaders are right to back student assessments.* The Tyee. https://thetyee.ca/Opinion/2022/11/29/First-Nations-Right-Back-Student-Assessments/

United Nations Conference on Trade and Development. (2021). *State of Commodity Dependence 2021*. United Nations. https://doi.org/10.18356/9789210057790

Vera-Toscano, E., Weersink, A., & Phimister, E. (2003). Earnings Mobility of Rural versus Urban Workers in Canada. *Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie, 51*(2), 241–258. https://doi.org/10.1111/j.1744-7976.2003.tb00175.x

Weber, M. (1978). *Economy and society: An outline of interpretive sociology (G.* Roth & C. Wittich, Eds.). University of California Press. (Original work published 1922)

Western Canadian Protocol for Collaboration in Basic Education, Kindergarten to Grade 12. (1993). [Interprovincial agreement].

https://www.edu.gov.mb.ca/k12/framework/publications/math/framework_k-8/docs/background.pdf

Winton, S., & Staples, S. (2022). *Shifting meanings: The struggle over public funding of private schools in Alberta, Canada.* Education Policy Analysis Archives, 30. https://doi.org/10.14507/epaa.30.7002

Xiang, L., & Stillwell, J. (2023). Rural–Urban Educational Inequalities and Their Spatial Variations in China. *Applied Spatial Analysis and Policy, 16*(2), 873–896. https://doi.org/10.1007/s12061-023-09506-1

Yoon, E.-S., Lubienski, C., & Lee, J. (2018). The geography of school choice in a city with growing inequality: The case of Vancouver. *Journal of Education Policy, 33*(2), 279–298. https://doi.org/10.1080/02680939.2017.1346203

Zarifa, D., Seward, B., & Milian, R. P. (2019). Location, location, location: Examining the rural-urban skills gap in Canada. *Journal of Rural Studies, 72*, 252–263. https://doi.org/10.1016/j.jrurstud.2019.10.032

Zhang, D., Li, X., & Xue, J. (2015). Education Inequality between Rural and Urban Areas of the People's Republic of China, Migrants' Children Education, and Some Implications. *Asian Development Review, 32*(1), 196–224. https://doi.org/10.1162/ADEV_a_00042

Zhao, Y. (2020). Two decades of havoc: A synthesis of criticism against PISA. *Journal of Educational Change, 21*(2), 245–266. https://doi.org/10.1007/s10833-019-09367-x

Zinga, D. M., & Gordon, M. K. (2016). ‘Racism under the radar’: Student perceptions of school experiences in a multicultural context. *Race Ethnicity and Education, 19*(5), 1088–1116. https://doi.org/10.1080/13613324.2014.885429

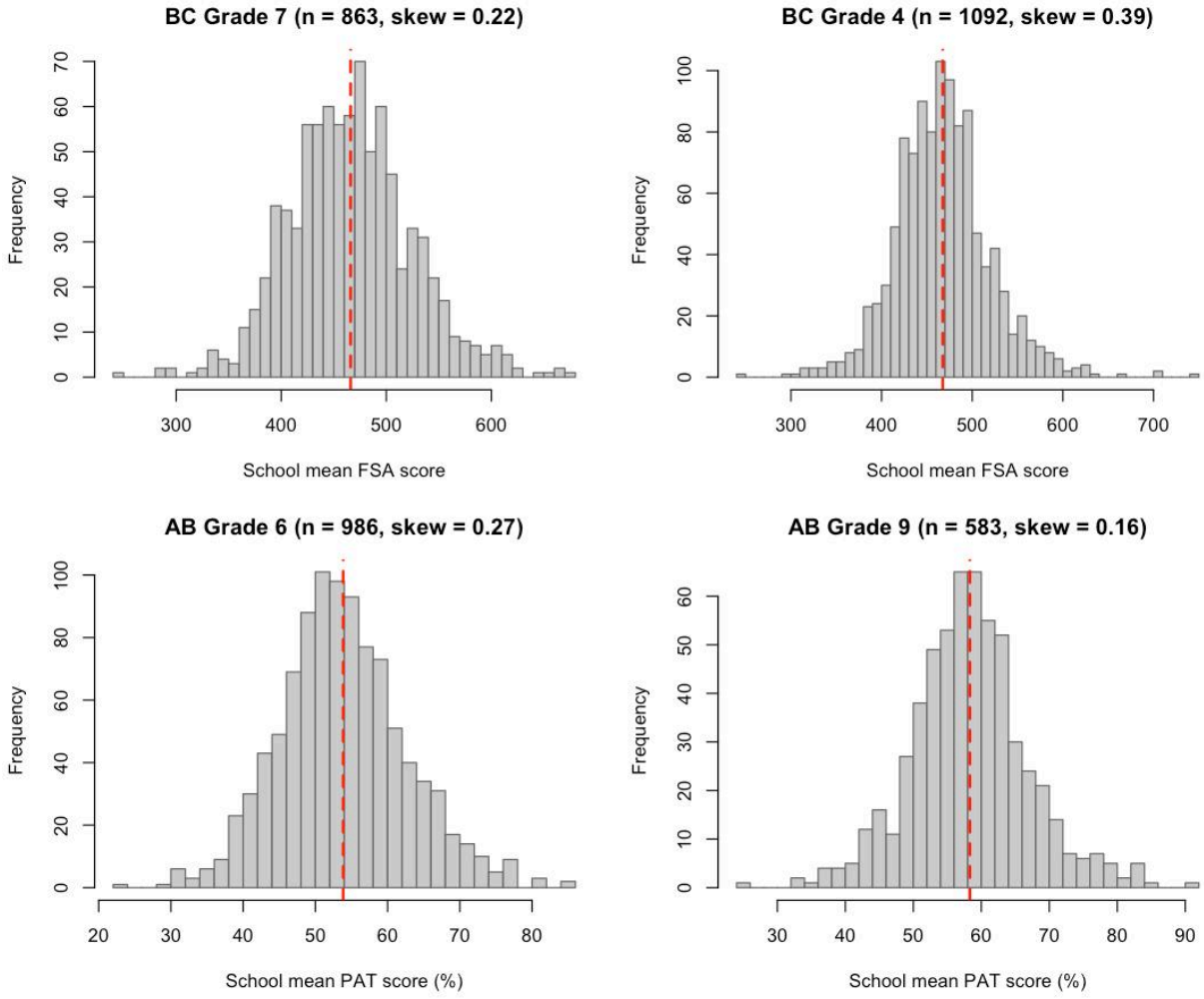
Appendix A

Outcome Distributions

Figure A1 shows the distribution of the composite outcome (school mean of literacy and numeracy) for each analytic cohort: the BC FSA scale-score composites (Grade 7 primary, Grade 4 secondary) and the Alberta PAT percentage composites (Grade 6 primary, Grade 9 secondary). Table A1 reports the corresponding distributional statistics. No cohort shows marked skew (all |skewness| ≤ 0.39), and boundary concentration is negligible: as the Floor and Ceiling columns of Table A1 show, only 0.1–0.3% of schools fall within the extreme 2% of either end of the observed scale range in any cohort. These figures indicate no meaningful ceiling or floor compression of the kind under which scale-shape concerns (Ho, 2009) would bind most strongly. Kurtosis is somewhat elevated in the BC Grade 4 cohort (5.02, against a normal-distribution value of 3), indicating heavier tails than in the other cohorts. Because rank transformation is invariant to such distributional features, the stability of results under the rank-transformed re-estimation (Appendix D) indicates that these tails do not drive them.

Figure A1

Distributions of School-Level Composite Attainment Across the Four Analytic Cohorts



Note. Dashed lines indicate cohort means. BC panels show Foundation Skills Assessment (FSA) scale-score composites; Alberta panels show Provincial Achievement Test (PAT) percentage composites.

Table A1

Distributional Statistics for the Composite Outcomes

Cohort	*N*	*M*	*SD*	Min	Max	Skewness	Kurtosis	Floor (%)	Ceiling (%)
BC Grade 7	863	466.0	59.0	244.6	674.6	0.22	3.62	0.1	0.3
BC Grade 4	1,092	467.4	52.8	247.4	749.5	0.39	5.02	0.1	0.1
AB Grade 6	986	53.8	8.8	22.1	85.9	0.27	3.38	0.1	0.1
AB Grade 9	583	58.3	8.7	25.5	91.2	0.16	3.98	0.2	0.2

Note. BC outcomes are FSA scale scores; Alberta outcomes are PAT school mean percentages. Floor and Ceiling give the percentage of schools within the bottom and top 2% of the observed scale range, respectively.

Appendix B

Residual Diagnostics

Figures B1–B4 present diagnostic panels for the final models of the two primary cohorts, estimated by OLS on the analytic samples. Panels comprise residuals versus fitted values (with loess), normal Q–Q, scale–location, and component-plus-residual (partial residual) plots, for each continuous predictor. Residual plots reveal no systematic curvature, supporting the linearity assumption for income, parental education, and remoteness. Heteroscedasticity is mild and confined to the lower tail of the attainment distribution, and the clustered (CR2) standard errors are robust to it by construction (Pustejovsky & Tipton, 2018). Figures B5–B6 repeat the core panel for the secondary cohorts, which show the same structure.

Figure B1

BC Grade 7 Final Model: Core Diagnostic Panel

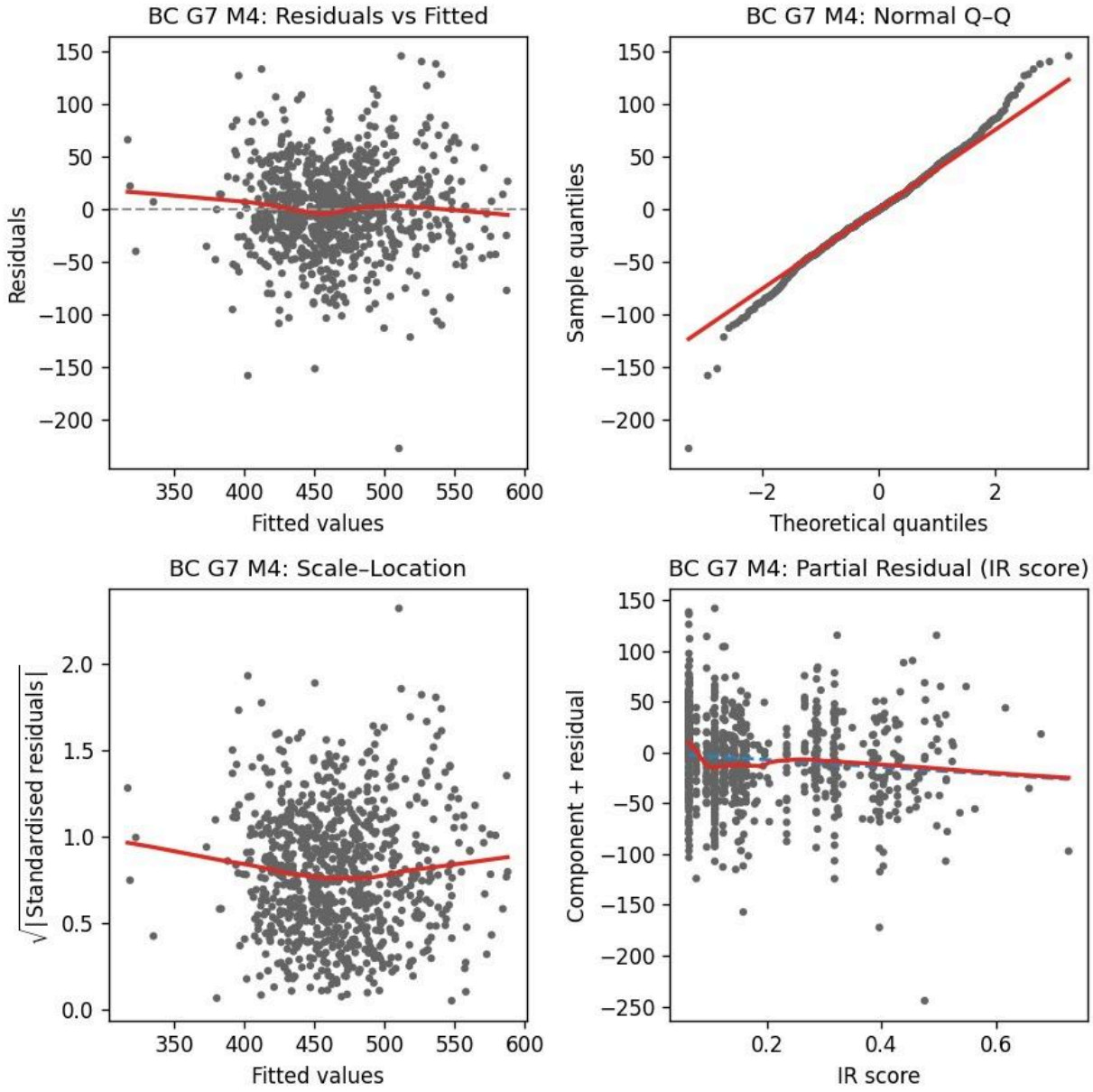


Figure B2

Alberta Grade 6 Final Model: Core Diagnostic Panel

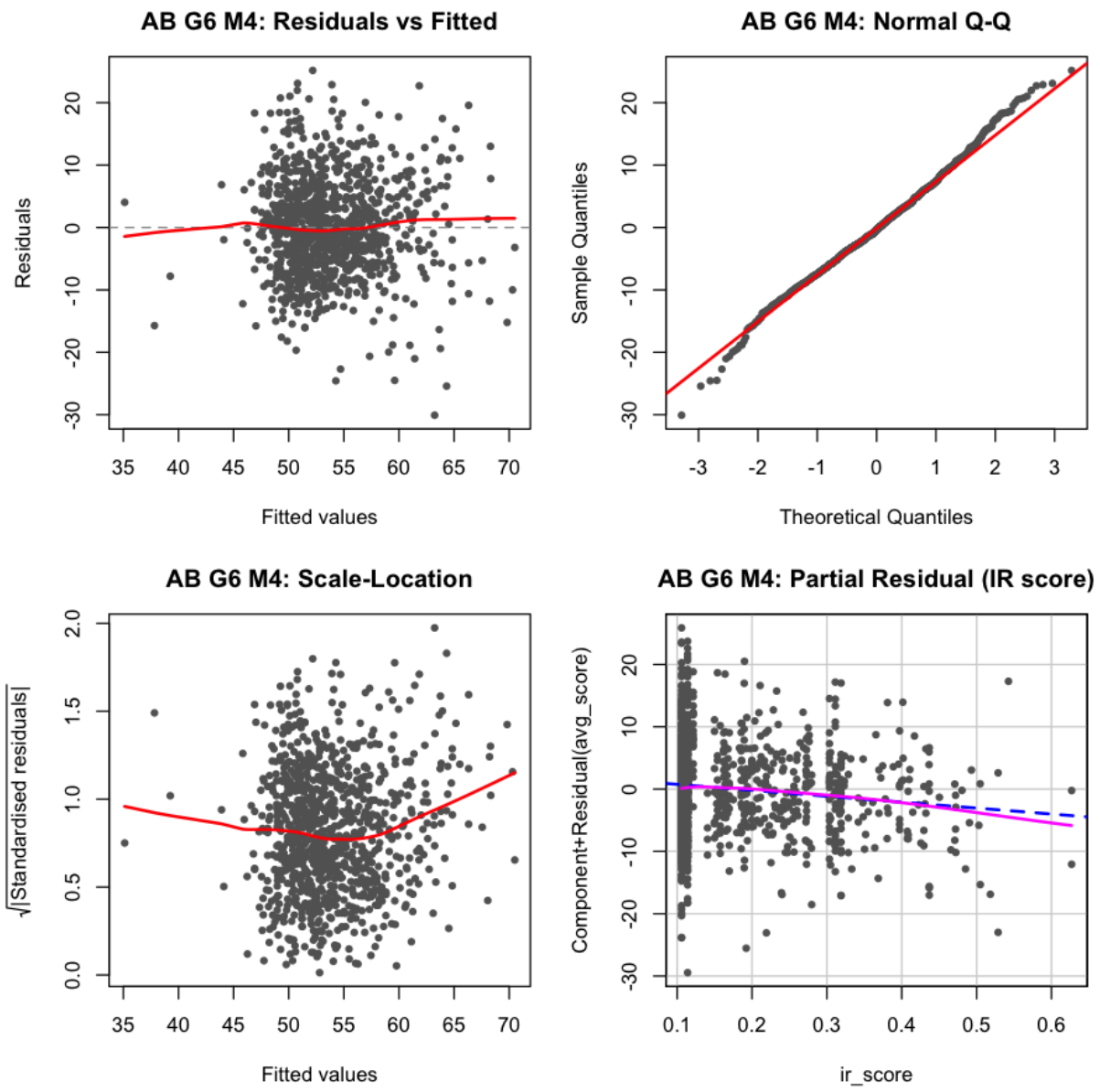


Figure B3

BC Grade 7: Partial Residual Plots (DA income, share bachelor's+, popctr)

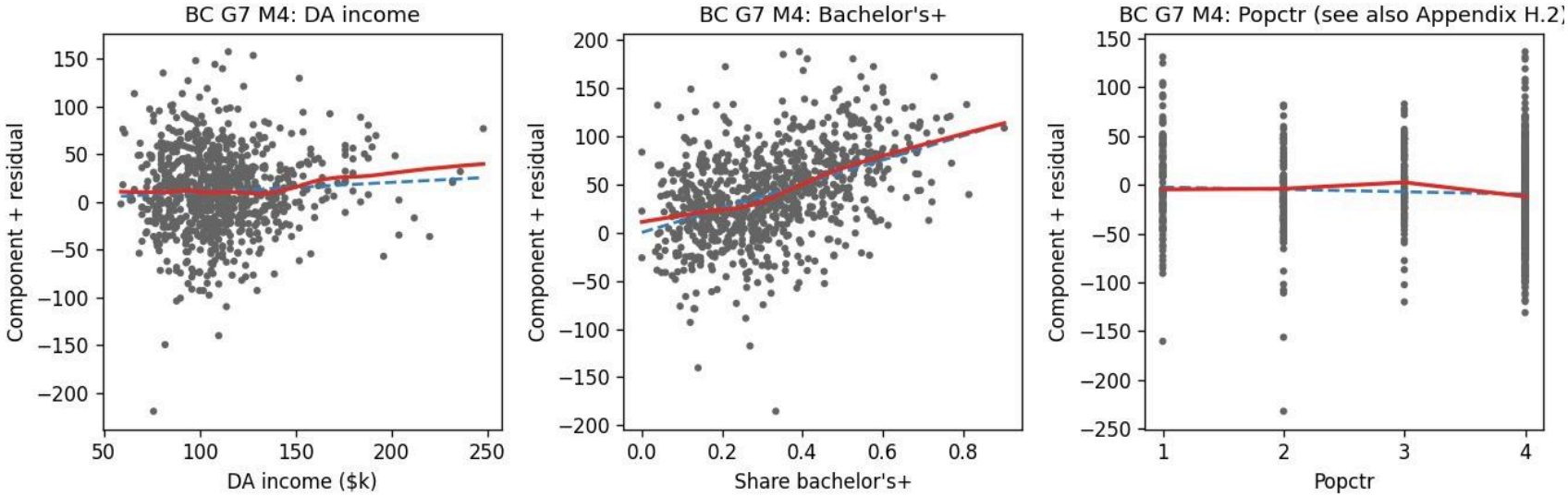


Figure B4

Alberta Grade 6: Partial Residual Plots (DA income, share bachelor's+, popctr)

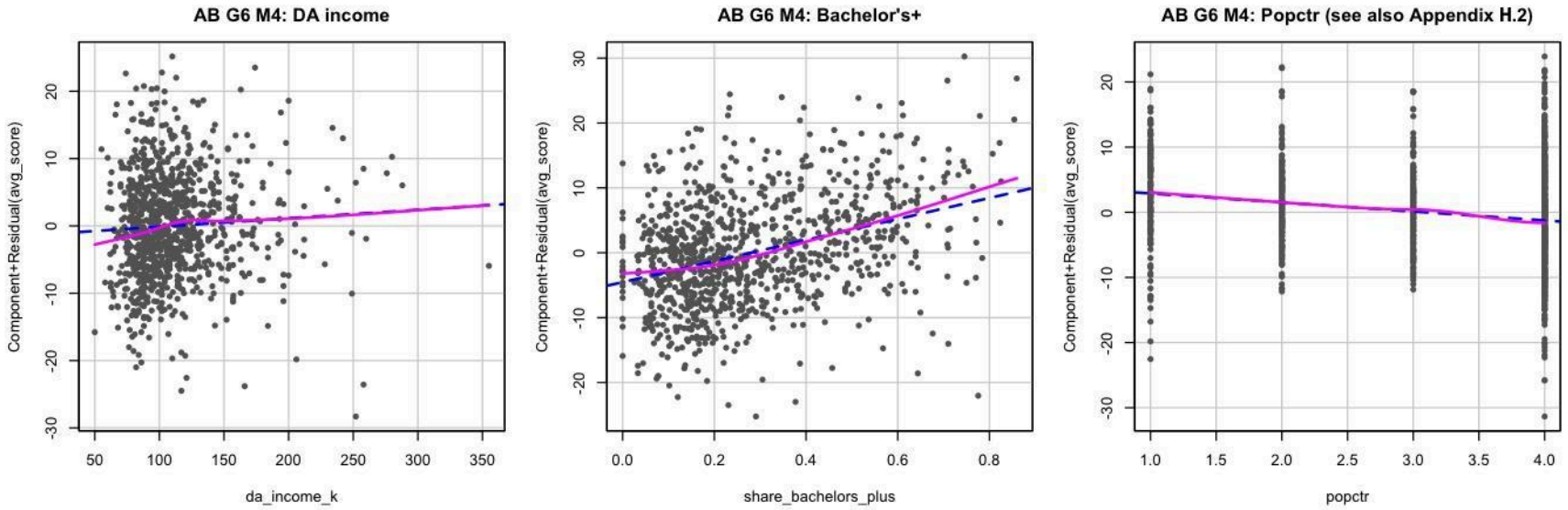


Figure B5

BC Grade 4 Final Model: Core Diagnostic Panel

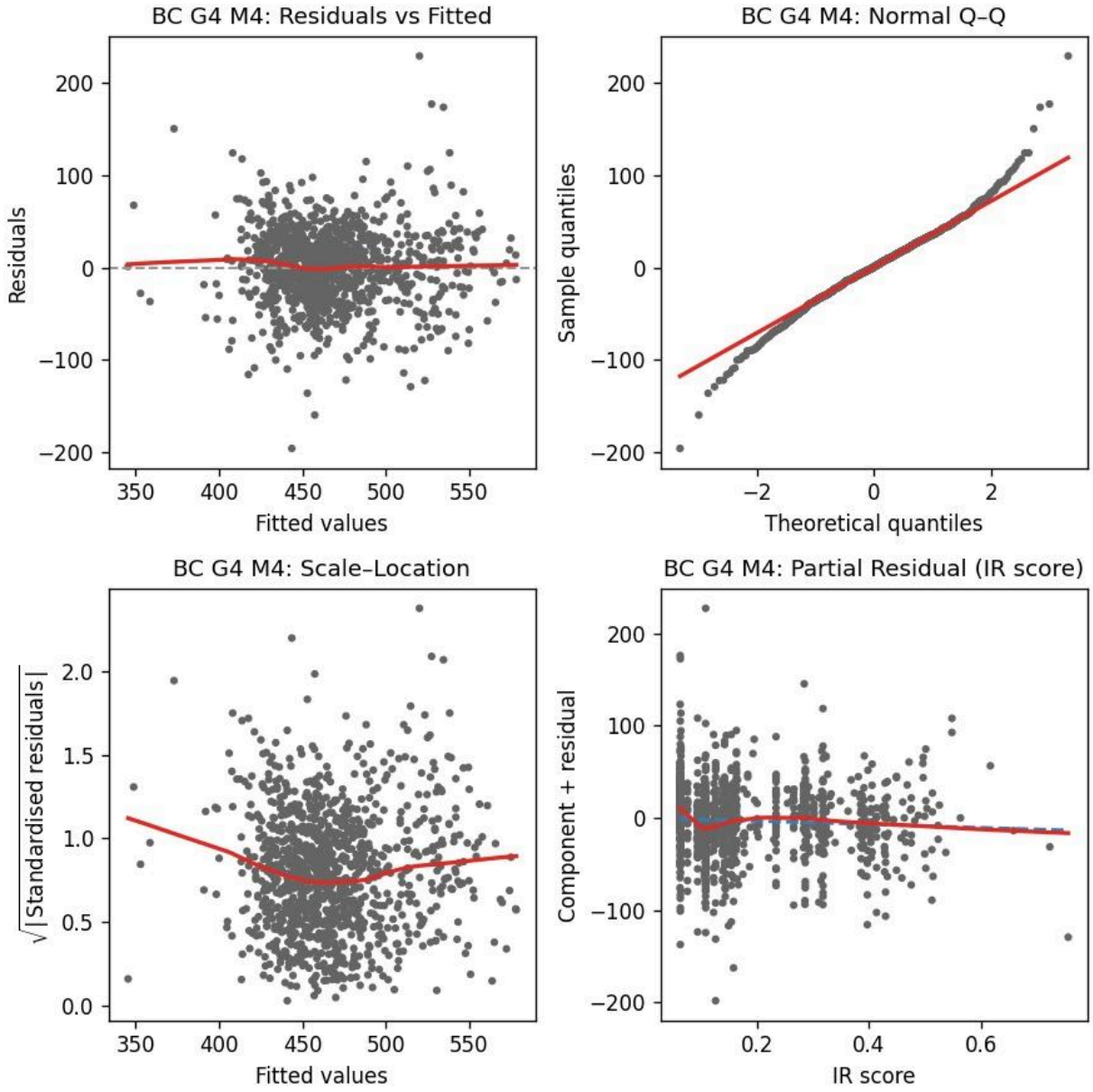
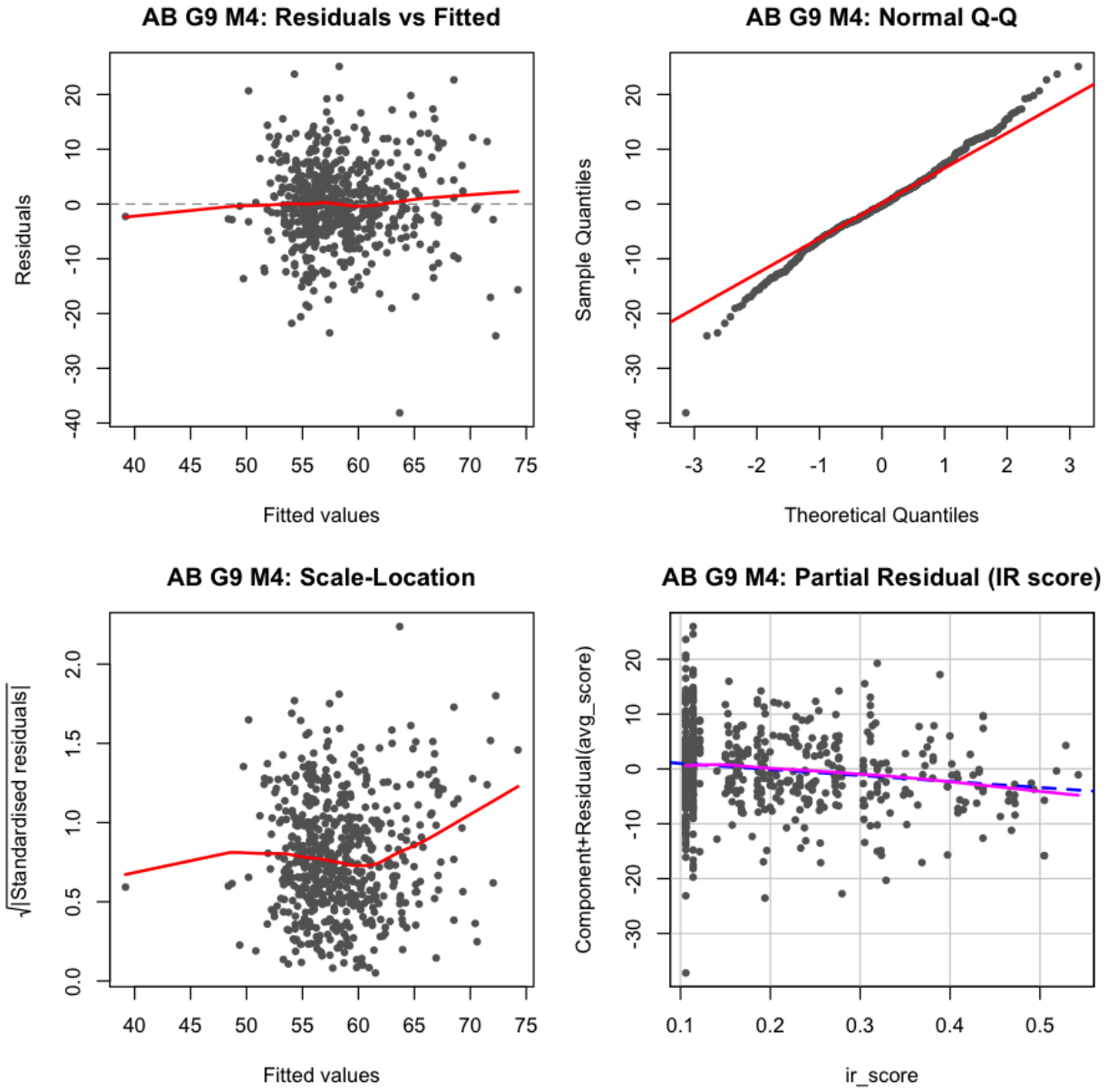


Figure B6

Alberta Grade 9 Final Model: Core Diagnostic Panel



Note. The Alberta popctr partial residual plots should be read alongside Section H.2, which shows the settlement-size association is not strictly monotone within the mid-remoteness band; popctr is nonetheless retained as continuous in the main models (see Section 3.6.3).

Appendix C

Variance Inflation Factors

Table C1 reports variance inflation factors for the final model of each cohort. Because the Alberta models contain a multi-category school type predictor, generalised VIFs are reported throughout, with GVIF^(1/(2·Df)) as the dimension-adjusted statistic; the VIF-equivalent column squares this to offer comparability with conventional thresholds. All values fall well below the conventional threshold of 10 (Field, 2018): the maxima are 2.61 (remoteness, BC Grade 7), 2.66 (parental education, Alberta Grade 6) and 3.14 (parental education, Alberta Grade 9). The strong settlement size–remoteness correlation ($r \approx -0.71$ to -0.73) therefore manifests as interpretive partialling rather than destabilising collinearity. Table C2 reports the interaction models. Interaction-constituent VIFs are parameterisation-dependent; under the public-reference coding reported in Table K5, all terms fall below 2.9.

Table C1

Generalised Variance Inflation Factors for the Final Models, All Four Cohorts

Term	GVIF	*df*	$GVIF^{1/(2df)}$	VIF-equivalent
British Columbia Grade 7				
DA income	1.38	1	1.17	1.38
Share bachelor's+	1.87	1	1.37	1.87
Popctr	2.20	1	1.48	2.20
IR score	2.61	1	1.61	2.61
Share Indigenous	1.43	1	1.20	1.43
School type	1.01	1	1.00	1.01
British Columbia Grade 4				
DA income	1.38	1	1.18	1.38
Share bachelor's+	1.87	1	1.37	1.87
Popctr	2.13	1	1.46	2.13
IR score	2.57	1	1.60	2.57
Share Indigenous	1.51	1	1.23	1.51
School type	1.01	1	1.01	1.01
Alberta Grade 6				
DA income	1.76	1	1.33	1.76
Share bachelor's+	2.66	1	1.63	2.66
Popctr	2.49	1	1.58	2.49
IR score	2.32	1	1.52	2.32
Share Indigenous	1.27	1	1.13	1.27
School type	1.07	3	1.01	1.02
Alberta Grade 9				
DA income	1.79	1	1.34	1.79
Share bachelor's+	3.14	1	1.77	3.14
Popctr	3.07	1	1.75	3.07
IR score	2.34	1	1.53	2.34
Share Indigenous	1.24	1	1.11	1.24
School type	1.15	3	1.02	1.05

Note. GVIF = generalised variance inflation factor (Fox & Monette, 1992); values are from each cohort's final model (Model 4). School type has $df = 3$ in Alberta (three dummies against a Public reference) and $df = 1$ in BC.

Table C2

Generalised Variance Inflation Factors for the Interaction Models

Term	GVIF	*df*	$GVIF^{1/(2df)}$	VIF-equivalent
British Columbia: School type × IR				
DA income	1.38	1	1.17	1.38
Share bachelor's+	1.87	1	1.37	1.87
Popctr	2.20	1	1.48	2.20
Share Indigenous	1.44	1	1.20	1.44
School type	1.01	1	1.00	1.01
IR score	2.82	1	1.68	2.82
School type × IR	1.22	1	1.10	1.22
British Columbia: IR × Indigenous share				
DA income	1.39	1	1.18	1.39
Share bachelor's+	1.97	1	1.40	1.97
Popctr	2.22	1	1.49	2.22
School type	1.01	1	1.00	1.01
Share Indigenous	2.56	1	1.60	2.56
IR score	2.67	1	1.64	2.67
IR × Share Indigenous	2.05	1	1.43	2.05
Alberta: IR × Indigenous share				
DA income	1.77	1	1.33	1.77
Share bachelor's+	2.85	1	1.69	2.85
Popctr	2.50	1	1.58	2.50
Share Indigenous	2.24	1	1.50	2.24
School type	1.07	3	1.01	1.02
IR score	2.44	1	1.56	2.44
IR × Share Indigenous	2.05	1	1.43	2.05
Alberta: School type × IR				
DA income	1.78	1	1.33	1.78
Share bachelor's+	2.67	1	1.63	2.67
Popctr	2.51	1	1.58	2.51
Share Indigenous	1.27	1	1.13	1.27
School type	1.81	3	1.10	1.22
IR score	2.80	1	1.67	2.80
School type × IR	2.44	3	1.16	1.35

Note. GVIF = generalised variance inflation factor (Fox & Monette, 1992); IR = Index of Remoteness. Interaction-constituent GVIFs are parameterisation-dependent; under the public-reference coding reported in Table K5, no term exceeds 2.85.

Appendix D

Robustness Checks: Rank Transformation and Wild Cluster Bootstrap

D.1 Rank-Transformed Outcomes

All four final models were re-estimated with the outcome replaced by its within-cohort rank (standardised). Because ranks are invariant to any monotone rescaling of the underlying test scale, stability under this specification indicates results are not artefacts of scale shape — the concern raised by the FSA and PAT scales' differing constructions (Section 3.6.2). In both primary models (BC Grade 7, Alberta Grade 6), every predictor's sign and significance level is unchanged. In the secondary cohorts there is one borderline change, flagged in the final column of Table D1: the Alberta Grade 9 private coefficient crosses into significance ($p = .059 \rightarrow .043$), a shift consistent with the sector's small sample ($n = 49$ schools, Satterthwaite $df \approx 5$) rather than any substantive difference.

Table D1

Final Models on Original Versus Rank-Transformed Outcomes

	Original outcome		Rank-transformed outcome		
Term	*β* (*SE*)	*p*	*β* (*SE*)	*p*	Significance change
British Columbia Grade 7					
DA income	0.046 (0.028)	.129	0.048 (0.031)	.141	
Share bachelor's+	0.349 (0.050)	< .001	0.401 (0.059)	< .001	
Popctr	−0.045 (0.048)	.354	−0.062 (0.053)	.249	
IR score	−0.081 (0.046)	.085	−0.078 (0.050)	.122	
Share Indigenous	−0.189 (0.032)	< .001	−0.137 (0.029)	< .001	
Independent school (ref: Public)	1.134 (0.126)	< .001	1.074 (0.116)	< .001	
British Columbia Grade 4					
DA income	0.065 (0.027)	.027	0.071 (0.027)	.018	
Share bachelor's+	0.291 (0.052)	< .001	0.333 (0.071)	< .001	
Popctr	−0.071 (0.054)	.197	−0.083 (0.070)	.249	
IR score	−0.042 (0.053)	.428	−0.052 (0.063)	.410	
Share Indigenous	−0.205 (0.042)	< .001	−0.139 (0.036)	< .001	
Independent school (ref: Public)	1.253 (0.158)	< .001	1.107 (0.114)	< .001	
Alberta Grade 6					
DA income	0.048 (0.040)	.292	0.044 (0.041)	.335	
Share bachelor's+	0.339 (0.103)	.032	0.364 (0.104)	.027	
Popctr	−0.178 (0.061)	.008	−0.216 (0.063)	.002	
IR score	−0.108 (0.050)	.037	−0.124 (0.047)	.012	
Share Indigenous	−0.124 (0.046)	.024	−0.082 (0.029)	.018	
Separate (ref: Public)	0.150 (0.045)	.021	0.171 (0.050)	.020	
Charter	1.222 (0.224)	.018	1.047 (0.189)	.017	
Private	0.619 (0.136)	.010	0.500 (0.095)	.006	
Alberta Grade 9					
DA income	−0.022 (0.037)	.579	−0.006 (0.034)	.867	
Share bachelor's+	0.355 (0.067)	.003	0.351 (0.078)	.006	
Popctr	−0.214 (0.059)	.001	−0.242 (0.067)	.001	
IR score	−0.132 (0.054)	.018	−0.172 (0.055)	.003	
Share Indigenous	−0.129 (0.034)	.003	−0.099 (0.030)	.008	
Separate (ref: Public)	0.219 (0.059)	.008	0.251 (0.070)	.010	
Charter	1.140 (0.294)	.061	0.962 (0.363)	.118	
Private	0.537 (0.220)	.059	0.386 (0.144)	.043	Yes

Note. Cluster-robust (CR2) standard errors, clustered at the census subdivision level, in parentheses. Exact *p*-values are reported; values below .001 are shown as < .001. Significance change flags coefficients whose status at α = .05 differs between outcomes.

D.2 Wild Cluster Bootstrap

Key coefficients were re-tested using a wild cluster bootstrap (Cameron, Gelbach, & Miller, 2008), implemented in the ‘fwildclusterboot’ package in R (Fischer & Roodman, 2021). Webb weights were used given the unbalanced cluster sizes (MacKinnon & Webb, 2017), with 9,999 replications and the null imposed. Table D2 reports bootstrap p-values alongside the CR2/Satterthwaite p-values from the main models. Agreement at $\alpha = .05$ holds for 25 of 28 coefficients. The three exceptions are Alberta Grade 6 share bachelor's+ (.032 → .129), Alberta Grade 6 Charter (.018 → .051) and Alberta Grade 9 Charter (.061 → .048); however, the Charter cases rest on very small sector samples (n = 14 and 16 schools respectively), consistent with the caution already applied to Charter estimates in the text.

Table D2

Satterthwaite Versus Wild-Cluster-Bootstrap p-Values, Final Models

Term	*β*	Satterthwaite *p*	Bootstrap *p*	Agreement at .05
British Columbia Grade 7				
DA income	0.046	.129	.098	Yes
Share bachelor's+	0.349	< .001	< .001	Yes
Popctr	−0.045	.354	.391	Yes
IR score	−0.081	.085	.088	Yes
Share Indigenous	−0.189	< .001	< .001	Yes
Independent school (ref: Public)	1.134	< .001	< .001	Yes
British Columbia Grade 4				
DA income	0.065	.027	.019	Yes
Share bachelor's+	0.291	< .001	< .001	Yes
Popctr	−0.071	.197	.214	Yes
IR score	−0.042	.428	.433	Yes
Share Indigenous	−0.205	< .001	< .001	Yes
Independent school (ref: Public)	1.253	< .001	< .001	Yes
Alberta Grade 6				
DA income	0.048	.292	.246	Yes
Share bachelor's+	0.339	.032	.129	No
Popctr	−0.178	.008	.037	Yes
IR score	−0.108	.037	.032	Yes
Share Indigenous	−0.124	.024	.029	Yes
Separate (ref: Public)	0.150	.021	.027	Yes
Charter	1.222	.018	.051	No
Private	0.619	.010	.018	Yes
Alberta Grade 9				
DA income	−0.022	.579	.573	Yes
Share bachelor's+	0.355	.003	.001	Yes
Popctr	−0.214	.001	.007	Yes
IR score	−0.132	.018	.017	Yes
Share Indigenous	−0.129	.003	.003	Yes
Separate (ref: Public)	0.219	.008	.018	Yes
Charter	1.140	.061	.048	No
Private	0.537	.059	.053	Yes

Note. *β* = standardised coefficients from the final models (Model 4); both tests cluster at the census subdivision level. Wild cluster bootstrap: Webb weights, 9,999 replications, null imposed. Values below .001 are shown as < .001.

Appendix E

Correspondence Confirming Ethics Review Was Not Required

RM

Research Management <research.man...

Monday, January 5, 2026 at 6:26 AM

To: [redacted] Research Management

Dear [redacted]

Thank you for your submission to the research ethics internal checklist.

We have reviewed your submission and, based on the details provided, can confirm that you do not need to apply for ethics review.

Should your methodology need substantially to change during the project, you are welcome to check with us again via the checklist or proceed straight to ethics review if clearly required.

All the best,

Niall

Niall Summers *(he/him)*
Research Support Officer
Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY

DEPARTMENT OF EDUCATION

UNIVERSITY OF OXFORD

Appendix F

Supplementary Descriptive Statistics

This appendix reports the descriptive tables referenced in the Descriptive Overview (Section 4.1): income by settlement size (Tables F1–F2) and by remoteness (Tables F3–F4), and the school-type composition of the analytic samples (Table F5).

Table F1

Distribution of Schools and Family Income by Settlement Size (Popctr), British Columbia Grade 7 (N = 863)

Popctr category	*n* schools	%	*M* income ($k)	*SD*
1 (Rural)	107	12.4	102.63	20.32
2 (Small population centre)	148	17.1	96.05	18.65
3 (Medium population centre)	91	10.5	100.55	17.91
4 (Large population centre)	517	59.9	118.13	27.95

Note. Popctr is coded 1 = rural to 4 = large city (Census 2021 dissemination area modal value). % is the share of the provincial analytic sample in each popctr category. Income is average after-tax income of economic family in the school's DA, in thousands of 2020 CAD. Population centre size classes follow Statistics Canada definitions; the main text refers to these informally as small towns, medium cities, and large cities.

Table F2

Distribution of Schools and Family Income by Settlement Size (Popctr), Alberta Grade 6 (N = 986)

Popctr category	*n* schools	%	*M* income ($k)	*SD*
1 (Rural)	143	14.5	101.39	33.68
2 (Small population centre)	168	17.0	99.57	17.93
3 (Medium population centre)	133	13.5	113.37	31.64
4 (Large population centre)	542	55.0	114.99	37.26

Note. Popctr is coded 1 = rural to 4 = large city (Census 2021 dissemination area modal value). % is the share of the provincial analytic sample in each popctr category. Income is average after-tax income of economic family, in thousands of 2020 CAD. *N* = 986 matches the regression-analytic sample used throughout Chapter 4 (see Analytic Samples).

Table F3

Family Income by Remoteness Quartile (IR), British Columbia Grade 7 (N = 863)

IR quartile	IR range	*n* schools	*M* income ($k)	*SD* ($k)
Q1 (least remote)	0.06–0.07	216	121.57	33.42
Q2	0.07–0.12	216	120.60	22.64
Q3	0.12–0.29	216	103.46	20.17
Q4 (most remote)	0.29–0.73	215	96.59	17.43

Note. IR = Index of Remoteness (Alasia et al., 2017), assigned at the census subdivision level. Quartiles rank the sample by IR score into four equal-n groups; on the allocation of tied IR values, see the Table K1 note.

Table F4

Family Income by Remoteness (IR), Alberta Grade 6 (N = 986)

IR band	IR range	n schools	M income ($k)	SD ($k)
Q1–Q2 (metropolitan)	0.106–0.114	494	114.53	37.74
Q3	0.114–0.231	246	108.80	29.59
Q4 (most remote)	0.233–0.627	246	102.79	28.72

Note. IR = Index of Remoteness (Alasia et al., 2017), assigned at the census subdivision level. Because Calgary and Edmonton share near-minimum IR values, the two least remote quartiles contain only large-urban schools and are reported as a single metropolitan band. Income is average after-tax income of economic family, in thousands of 2020 CAD. *n*s sum to 986, matching the regression-analytic sample in Table 4.1.

Table F5

School Type Composition of the Analytic Samples, n (%)

Province	Public	Separate	Charter	Private (AB) /Independent (BC)
Alberta	664 (67.3)	255 (25.9)	14 (1.4)	53 (5.4)
British Columbia	696 (80.6)	—	—	167 (19.4)

Note. Percentages are of each province's analytic sample (Alberta *n* = 986; British Columbia *n* = 863). Dashes indicate school types that do not exist in the province.

Appendix G

Exurban Classification Sensitivity Analyses

This appendix reports the sensitivity checks for the classification of exurban schools — those coded rural on settlement size (popctr = 1) but located on the outskirts of cities (IR < 0.20) — referenced in Sections 4.1, 4.2, 4.3, and the Discussion. In BC, 40 of 107 rural-classified schools (37.4%) were exurban, whereas in Alberta it was 33 of 143 (23.1%). Table G4 reports the exclusion specification for both provinces.

Table G1

Mean DA Income of Economic Family by Settlement Category Under the Original and Restricted Rural Classifications, British Columbia Grade 7

	Original classification		Rural restricted to IR ≥ 0.20	
Settlement category	*n*	*M* income ($000s)	*n*	*M* income ($000s)
Rural	107	102.63	67	94.21
Small population centre	148	96.05	188	100.45
Medium population centre	91	100.55	91	100.55
Large population centre	517	118.13	517	118.13

Note. The restricted classification limits the rural category to DAs with IR ≥ 0.20, reassigning the 40 exurban schools to the small-population-centre category; under it, mean income orders monotonically with settlement size. IR = Index of Remoteness.

Table G2

Primary Model Coefficients With the 40 Exurban Schools Reclassified, British Columbia Grade 7

	Original model		Exurban schools reclassified	
Term	*β* (*SE*)	*p*	*β* (*SE*)	*p*
DA income	0.046 (0.028)	.129	0.047 (0.028)	.127
Share bachelor's+	0.349 (0.050)	< .001	0.351 (0.049)	< .001
Popctr	−0.045 (0.048)	.354	−0.073 (0.057)	.209
IR score	−0.081 (0.046)	.085	−0.104 (0.050)	.044
Share Indigenous	−0.189 (0.032)	< .001	−0.189 (0.032)	< .001
Independent school (ref: Public)	1.134 (0.126)	< .001	1.135 (0.126)	< .001

Note. Standardised coefficients with cluster-robust (CR2) standard errors, clustered at the census subdivision level, in parentheses; Satterthwaite *p*-values. *N* = 863. Exurban schools (popctr = 1, IR < 0.20) reassigned to the small-population-centre category.

Table G3

Sensitivity of the Primary Model to Exurban Reclassification and Exclusion, Alberta Grade 6

	Original model		Exurbs reclassified		Exurbs excluded	
Term	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*
DA income	0.048 (0.040)	.292	0.056 (0.041)	.240	0.048 (0.030)	.206
Share bachelor's+	0.339 (0.103)	.032	0.330 (0.106)	.038	0.345 (0.100)	.032
Popctr	−0.178 (0.061)	.008	−0.169 (0.069)	.020	−0.140 (0.072)	.061
IR score	−0.108 (0.050)	.037	−0.112 (0.051)	.034	−0.082 (0.052)	.123
Share Indigenous	−0.124 (0.046)	.024	−0.125 (0.046)	.024	−0.125 (0.047)	.026
Separate (ref: Public)	0.150 (0.045)	.021	0.145 (0.045)	.024	0.150 (0.046)	.023
Charter	1.222 (0.224)	.018	1.216 (0.223)	.018	1.225 (0.223)	.018
Private	0.619 (0.136)	.010	0.626 (0.136)	.010	0.629 (0.130)	.014

Note. Standardised coefficients with cluster-robust (CR2) standard errors, clustered at the census subdivision level, in parentheses; Satterthwaite *p*-values. Reclassified: exurban schools (popctr = 1, IR < 0.20) reassigned to the small-population-centre category (popctr = 2). Excluded: exurban schools dropped, with all standardised variables re-standardised on the reduced sample (N = 953).

Table G4

Primary Model Coefficients After Excluding Exurban Schools, Both Provinces

	Full sample		Exurbs excluded	
Term	*β* (*SE*)	*p*	*β* (*SE*)	*p*
British Columbia Grade 7				
DA income	0.046 (0.028)	.129	0.055 (0.029)	.084
Share bachelor's+	0.349 (0.050)	< .001	0.342 (0.052)	< .001
Popctr	−0.045 (0.048)	.354	−0.119 (0.067)	.085
IR score	−0.081 (0.046)	.085	−0.142 (0.061)	.027
Share Indigenous	−0.189 (0.032)	< .001	−0.192 (0.034)	< .001
Independent school (ref: Public)	1.134 (0.126)	< .001	1.155 (0.132)	< .001
Alberta Grade 6				
DA income	0.048 (0.040)	.292	0.048 (0.030)	.206
Share bachelor's+	0.339 (0.103)	.032	0.345 (0.100)	.032
Popctr	−0.178 (0.061)	.008	−0.140 (0.072)	.061
IR score	−0.108 (0.050)	.037	−0.082 (0.052)	.123
Share Indigenous	−0.124 (0.046)	.024	−0.125 (0.047)	.026
Separate (ref: Public)	0.150 (0.045)	.021	0.150 (0.046)	.023
Charter	1.222 (0.224)	.018	1.225 (0.223)	.018
Private	0.619 (0.136)	.010	0.629 (0.130)	.014

Note. Standardised coefficients with cluster-robust (CR2) standard errors, clustered at the census subdivision level, in parentheses; Satterthwaite *p*-values. BC: $N = 863$ full sample, 823 excluding the 40 exurban schools. Alberta: $N = 986$ full sample, 953 excluding exurbs.

Appendix H

Supplementary Settlement-Size Analyses (Alberta)

This appendix reports the three supplementary analyses referenced in Sections 4.3 and 5.2.

H.1 Decomposition of the Rurality Block

Because settlement size (popctr) and remoteness (IR) are strongly correlated (r = −0.71 in Alberta), each was estimated net of SES alone, then jointly. As Table H1 shows, neither predicts attainment net of SES alone; entered jointly, both are significant, meaning each predicts attainment only once the other is held constant. The pattern replicates across both Alberta cohorts.

Table H1

Rurality Variables Net of SES: Alone Versus Joint Entry, Alberta Grades 6 and 9

		Grade 6		Grade 9	
Specification	Term	β (*SE*)	*p*	β (*SE*)	*p*
Settlement size + SES (no IR)	Popctr	−0.097 (0.057)	.120	−0.101 (0.061)	.130
Remoteness + SES (no popctr)	IR score	−0.055 (0.081)	.503	−0.079 (0.075)	.298
Both + SES	Popctr	−0.196 (0.061)	.004	−0.230 (0.070)	.003
Both + SES	IR score	−0.163 (0.051)	.003	−0.193 (0.056)	.001

Note. The dependent variable and continuous predictors are z-scored, as in the main models; cluster-robust (CR2) standard errors, clustered at the census subdivision level, in parentheses.

H.2 Remoteness-Quartile-Restricted Regressions

The Grade 6 regression (attainment on SES and settlement-size category) was re-estimated within remoteness-quartile subsamples, with schools ranked by IR. The first two quartiles contain only large-urban schools (Calgary and Edmonton share near-minimum IR values); results are therefore reported for the third quartile — the only band in which all four categories co-occur — and the fourth. IR ranges were 0.114–0.231 (Q3) and 0.233–0.627 (Q4). Per-cell n in Q3: rural = 62, small town = 82, medium city = 54, large urban = 48 (total

246); in Q4: rural = 81, small town = 86, medium city = 79 (total 246; no large-urban schools). Within the mid-remoteness band, rural-classified DAs outperform small-town schools net of SES; no settlement-size category differs significantly within the most remote quartile (Table H2).

Table H2

Within-Quartile Regressions, Alberta Grade 6

Term	*β*	*SE*	*p*
Quartile 3 (mid-remoteness)			
DA income	−0.021	0.106	.847
Share bachelor's+	0.095	0.109	.402
Popctr 2 (small town)	−0.402	0.154	.013
Popctr 3 (medium city)	−0.326	0.179	.090
Popctr 4 (large urban)	0.232	0.207	.305
Quartile 4 (most remote)			
DA income	−0.031	0.053	.594
Share bachelor's+	0.146	0.112	.216
Popctr 2 (small town)	0.115	0.145	.428
Popctr 3 (medium city)	0.163	0.168	.356

Note. Reference category: rural (popctr = 1). The dependent variable is z-scored; standard errors are cluster-robust (CR2).

H.3 School-Type Sector Profiles

Table H3 reports each school type's mean socioeconomic and rurality characteristics, contextualising the sector attenuation patterns discussed in Section 4.3. Private schools show the highest sector income (≈$17k above public). Separate schools barely differ from public schools in SES but skew more urban and connected (higher popctr, lower IR), consistent with the separate sector's attenuation once rurality controls enter the model.

Table H3

Mean Socioeconomic and Rurality Characteristics by School Type, Alberta Grade 6

School type	*n*	*M* income ($k)	*M* bachelor's+	*M* popctr	*M* IR
Public	664	109.0	0.266	2.95	0.186
Separate	255	109.3	0.292	3.41	0.170
Charter	14	119.6	0.368	3.64	0.136
Private	53	126.3	0.367	3.13	0.154

Note. Income is the mean after-tax income of economic family in the schools' dissemination areas, in thousands of 2020 CAD; bachelor's+ is the mean share of adults (25–64) holding a bachelor's degree or higher. School types follow Section 3.4.5: separate schools are Catholic and fully publicly funded; charter schools are publicly funded but pedagogically autonomous; private schools are accredited and fee-charging; public is the provincial default system.

Appendix I

Alternative Block-Order Regressions (British Columbia and Alberta)

This appendix re-estimates both provincial models with school type entered first — a robustness check on the hierarchical ordering, referenced in Sections 4.2 and 4.3. Reversing the entry order shows the ΔR^2 attributable to school type before any other predictors are controlled, and each subsequent block's ΔR^2 net of school type. Specifications are otherwise identical to the primary models (Tables 4.2-4.3): entries are standardised coefficients (β), with cluster-robust (CR2) standard errors (clustered at the census subdivision level) in parentheses, and the outcome and continuous predictors are z-scored. Because both orderings converge on the same specification, Model 4 in each table is identical to the corresponding primary final model.

Table I1

Alternative Block-Order Hierarchical Regression, British Columbia Grade 7 (FSA)

Predictor	Model 1: School type	Model 2: + SES	Model 3: + Rurality	Model 4: + Indigenous share
Intercept	−0.225* (0.103)	−0.216** (0.057)	−0.215** (0.063)	−0.219** (0.068)
Block 1: School type				
Independent (ref: Public)	1.165*** (0.143)	1.115*** (0.126)	1.111*** (0.127)	1.134*** (0.126)
Block 2: SES				
DA income	—	0.069* (0.031)	0.051 (0.028)	0.046 (0.028)
Share bachelor's+	—	0.444*** (0.040)	0.389*** (0.045)	0.349*** (0.050)
Block 3: Rurality				
Popctr	—	—	−0.040 (0.046)	−0.045 (0.048)
IR score	—	—	−0.149** (0.048)	−0.081 (0.046)
Block 4: Indigenous Share				
Indigenous share	—	—	—	−0.189*** (0.032)
Model fit				
R^2	.212	.444	.455	.480
ΔR^2	—	+.232	+.011	+.025

Note. N = 863 schools. School type is Independent entered against a Public reference. Dashes indicate predictors not entered in that model. ΔR^2 = incremental R^2 from block entry.

*p < .05. **p < .01. ***p < .001.

Table I2

Alternative Block-Order Hierarchical Regression, Alberta Grade 6 (PAT)

Predictor	Model 1: School type	Model 2: + SES	Model 3: + Rurality	Model 4: + Indigenous share
Intercept	−0.107 (0.064)	−0.080 (0.051)	−0.092 (0.049)	−0.089 (0.055)
Block 1: School type				
Separate (ref: Public)	0.155* (0.058)	0.106 (0.043)	0.158* (0.046)	0.150* (0.045)
Charter	1.409* (0.218)	1.207* (0.220)	1.232* (0.238)	1.222* (0.224)
Private	0.875* (0.202)	0.662** (0.135)	0.627** (0.138)	0.619* (0.136)
Block 2: SES				
DA income	—	0.064 (0.041)	0.051 (0.038)	0.048 (0.040)
Share bachelor’s+	—	0.329** (0.061)	0.373* (0.101)	0.339* (0.103)
Block 3: Rurality				
Popctr	—	—	−0.203** (0.060)	−0.178** (0.061)
IR score	—	—	−0.157** (0.053)	−0.108* (0.050)
Block 4: Indigenous Share				
Indigenous share	—	—	—	−0.124* (0.046)
Model fit				
R^2	.064	.200	.218	.230
ΔR^2	—	+.136	+.018	+.012

Note. N = 986 schools. School type is entered as binary dummies against a Public reference. Dashes indicate predictors not entered in that model. ΔR^2 = incremental R^2 from block entry.

*p < .05. **p < .01. ***p < .001.

Appendix J

Secondary Cohort Results (British Columbia Grade 4 and Alberta Grade 9)

This appendix reports full hierarchical regression results for the secondary cohorts, estimated identically to the primary models: same blocks, same exclusions, standardised coefficients (β) with cluster-robust (CR2) standard errors (clustered at the census subdivision level) in parentheses, the outcome and continuous predictors z-scored, and the SD column giving one predictor standard deviation in raw units. BC Grade 4 reproduces the Grade 7 structure with two minor deviations (Section 4.5): DA income turns weakly significant, and school type's contribution to explained variance grows stronger still ($\Delta R^2 = .212$, moving above the SES block's .174); the Indigenous-share coefficient is robust to rank transformation (Table D1). The Alberta Grade 9 estimates mirror Grade 6 throughout, yet while charter and private coefficients retain their magnitudes, they fall short of significance under an even smaller cluster base.

Table J1

Hierarchical Regression Predicting School Mean FSA Score (Standardised Coefficients), British Columbia Grade 4

Predictor	*SD*	Model 1: SES	Model 2: + Rurality	Model 3: + Indigenous share	Model 4: + School type
Intercept	—	0.000 (0.052)	0.000 (0.053)	0.000 (0.056)	−0.204** (0.069)
Block 1: SES					
DA income	$25.6k	0.084* (0.034)	0.068* (0.032)	0.061 (0.032)	0.065* (0.027)
Share bachelor's+	0.16	0.368*** (0.038)	0.345*** (0.052)	0.322*** (0.055)	0.291*** (0.052)
Block 2: Rurality					
Popctr	1.13	—	−0.077 (0.060)	−0.073 (0.060)	−0.071 (0.054)
IR score	0.13	—	−0.129* (0.058)	−0.058 (0.056)	−0.042 (0.053)
Block 3: Indigenous share					
Indigenous share	0.11	—	—	−0.151*** (0.039)	−0.205*** (0.043)
Block 4: School type					
Independent school (ref: Public)	—	—	—	—	1.253*** (0.158)
Model fit					
R^2		.174	.181	.197	.409
ΔR^2		—	+.007	+.015	+.212

Note. N = 1,092 schools (Grade 4 FSA, 2024/25). Dashes indicate predictors not entered in that model. ΔR^2 = incremental R^2 from block entry. *SD* = one standard deviation of the predictor in raw units. *p < .05. **p < .01. ***p < .001.

Table J2

Hierarchical Regression Predicting School Mean PAT Score (Standardised Coefficients), Alberta Grade 9

Predictor	*SD*	Model 1: SES	Model 2: + Rurality	Model 3: + Indigenous share	Model 4: + School type
Intercept	—	0.000 (0.041)	0.000 (0.044)	0.000 (0.049)	−0.129 (0.072)
Block 1: SES					
DA income	$35.5k	0.010 (0.037)	−0.025 (0.036)	−0.023 (0.035)	−0.022 (0.037)
Share bachelor's+	0.19	0.366*** (0.035)	0.437** (0.065)	0.401** (0.068)	0.355** (0.067)
Block 2: Rurality					
Popctr	1.20	—	−0.230** (0.070)	−0.200** (0.072)	−0.214** (0.059)
IR score	0.10	—	−0.193** (0.056)	−0.140** (0.052)	−0.132* (0.054)
Block 3: Indigenous share					
Indigenous share	0.07	—	—	−0.127** (0.037)	−0.129** (0.034)
Block 4: School type					
Separate (ref: Public)	—	—	—	—	0.219** (0.059)
Charter	—	—	—	—	1.140 (0.294)
Private	—	—	—	—	0.537 (0.220)
Model fit					
R^2		.139	.161	.175	.226
ΔR^2		—	+.023	+.013	+.051

Note. N = 583 schools (Grade 9 PAT: English Language Arts 9 and Mathematics 9, 2024/25). School type is entered as binary dummies against a Public reference. Dashes indicate predictors not entered in that model. ΔR^2 = incremental R^2 from block entry. *p < .05. **p < .01. ***p < .001.

Appendix K

Supplementary School-Type Analyses

Model-based analyses in this appendix follow the primary specifications: the outcome and continuous predictors are z-scored, school type is entered as binary dummies, and cluster-robust (CR2) standard errors are clustered at the census subdivision level. Exact *p*-values are reported throughout, with values below .001 shown as < .001.

K.1 Independent-School Presence Across the Remoteness Distribution (British Columbia)

Independent schools comprise 19.4%, 15.7%, 26.4%, and 15.8% of schools from the least to the most remote IR quartile (Table K1). The one feature that survives every tie-breaking reorder is the peak at Q3: across 2,000 random permutations of tied IR values, Q3's lowest attainable share (25.0%) still exceeds Q1's highest (22.2%), so the third-quartile peak — unlike the exact shares themselves — is a robust pattern. Beyond that peak, the ordering shows no clear gradient. Table K2 details the regional concentration underlying these shares: three metropolitan regions account for every independent school in Q1–Q2 but none of Q4's 34 (see Section 4.6.2).

Formal tests agree (Table K3): the association between remoteness and independent status is null on every metric. The Wilcoxon rank-sum test is the appropriate robustness check given IR's right skew; its non-significance, alongside a negligible effect size ($|d| = 0.07$, well under Cohen's 0.2 benchmark), indicates no substantively meaningful remoteness gradient in independent-school presence.

Table K1

Independent-School Share by IR Quartile, British Columbia Grade 7

IR quartile	*n* schools	*n* independent	% independent
Q1 (least remote)	216	42	19.4
Q2	216	34	15.7
Q3	216	57	26.4
Q4 (most remote)	215	34	15.8

Note. IR = Index of Remoteness. Quartiles rank the analytic sample (n = 863) by IR score into four equal-n groups (216, 216, 216, 215) rather than fixed percentage cutoffs, keeping group sizes — not threshold placement — constant across quartiles. Because IR is measured at the CSD level, many schools share tied values (87% of BC G7 schools; all three quartile boundaries fall within tie blocks); allocation of ties is arbitrary, and Q1–Q3 shares vary by roughly ±1–2 points under alternative allocations without affecting the Q3 pattern.

Table K2

Regional Concentration of Independent Schools by Remoteness Quartile, British Columbia Grade 7

		Lower Mainland, Greater Victoria, or Kelowna		
IR quartile	*n* independent	*n*	%	Elsewhere in BC (*n*)
Q1 (least remote)	42	42	100.0	0
Q2	34	34	100.0	0
Q3	57	42	73.7	15
Q4 (most remote)	34	0	0.0	34
Total	167	118	70.7	49

Note. IR = Index of Remoteness. Percentages are of each quartile's independent schools. Of the 118 independent schools in the three regions, 98 are in the Lower Mainland, 13 in Greater Victoria, and 7 in the Kelowna area.

Table K3

Formal Tests of the Remoteness–School Type Association, British Columbia Grade 7

Test	Statistic	*p*
Point-biserial correlation	$r = -.028$	.405
Welch *t*-test (mean IR, independent vs. public)	$t(267.6) = -0.88$	.382
Cohen's *d*	$d = -0.07$	—
Wilcoxon rank-sum	$W = 57{,}473$	.824
Logistic regression (OR per 0.1 IR)	OR = 0.95	.405

Note. Point-biserial *r* = linear association between IR score and independent status; OR = odds ratio. Negative statistics indicate independent schools sit at marginally lower IR (less remote) locations.

K.2 School Type × Parental Education (Exploratory, British Columbia)

The exploratory test of whether the independent-school advantage varies with DA-level parental education is null: interaction β = −0.041 (SE = 0.068), p = .563 (Table K4). The DA-level measure is weak for choice-driven schools whose students travel from outside the catchment; see Discussion.

Table K4

School Type × Parental Education Interaction Model, British Columbia Grade 7

Term	β (*SE*)	*p*
Intercept	−0.219 (0.067)	.004
DA income	0.046 (0.029)	.135
Popctr	−0.044 (0.048)	.361
IR score	−0.078 (0.046)	.095
Share Indigenous	−0.190 (0.032)	< .001
Independent school (ref: Public)	1.137 (0.131)	< .001
Share bachelor's+	0.357 (0.061)	< .001
Independent school × Share bachelor's+	−0.041 (0.068)	.563

Note. N = 863 schools. Standardised coefficients; CR2 standard errors clustered at the CSD level.

K.3 Full Interaction Models (Both Provinces)

Table K5 reports full coefficient tables for the conditional-association models of Section 4.6. The school type × IR interactions run in opposite directions: in BC, the independent-school advantage narrows with remoteness (independent × IR $\beta = -0.312$, $p = .003$), while in Alberta the separate-school advantage widens with it (separate × IR $\beta = 0.121$, $p = .027$). The remoteness × Indigenous share interactions testing H6 are non-significant in both provinces.

Table K5

Full Interaction Models (Standardised Coefficients), Both Provinces

Term	β (*SE*)	*p*
British Columbia Grade 7: School type × IR		
Intercept	−0.220 (0.066)	.003
DA income	0.048 (0.027)	.102
Share bachelor's+	0.351 (0.048)	< .001
Popctr	−0.038 (0.046)	.412
Share Indigenous	−0.177 (0.031)	< .001
Independent school (ref: Public)	1.118 (0.103)	< .001
IR score	−0.027 (0.050)	.594
Independent school × IR	−0.312 (0.095)	.003
British Columbia Grade 7: IR × Indigenous share		
Intercept	−0.225 (0.063)	.001
DA income	0.045 (0.029)	.146
Share bachelor's+	0.344 (0.054)	< .001
Popctr	−0.047 (0.048)	.330
Independent school (ref: Public)	1.134 (0.126)	< .001
Share Indigenous	−0.203 (0.047)	.003
IR score	−0.084 (0.050)	.099
IR × Share Indigenous	0.011 (0.028)	.716
Alberta Grade 6: IR × Indigenous share		
Intercept	−0.081 (0.048)	.153
DA income	0.050 (0.039)	.272
Share bachelor's+	0.355 (0.106)	.031
Popctr	−0.176 (0.059)	.007
Share Indigenous	−0.086 (0.065)	.234
Separate (ref: Public)	0.150 (0.045)	.022
Charter	1.224 (0.227)	.019
Private	0.623 (0.139)	.011
IR score	−0.095 (0.055)	.094
IR × Share Indigenous	−0.023 (0.026)	.445
Alberta Grade 6: School type × IR		
Intercept	−0.089 (0.057)	.191
DA income	0.042 (0.039)	.348
Share bachelor's+	0.341 (0.102)	.031
Popctr	−0.186 (0.064)	.008
Share Indigenous	−0.122 (0.045)	.024
Separate (ref: Public)	0.161 (0.052)	.030
Charter	1.100 (0.547)	.181
Private	0.559 (0.125)	.004
IR score	−0.136 (0.054)	.016
Separate × IR	0.121 (0.050)	.027
Charter × IR	−0.265 (0.805)	.781
Private × IR	−0.221 (0.131)	.170

Note. BC $N = 863$; Alberta $N = 986$. Corresponding variance inflation factors are reported in Appendix C, Table C2.

Table K6

Public–Separate Attainment Gap by Remoteness Quartile, Alberta Grade 6

IR quartile	*n Public*	*n Separate*	Gap (points)	*d*	Percentile equivalent	*p*
Q1 (least remote)	154	78	1.4	0.15	50th → 56th	.247
Q2	151	75	−0.3	−0.03	50th → 49th	.793
Q3	180	43	1.6	0.23	50th → 59th	.118
Q4 (most remote)	179	59	2.6	0.35	50th → 64th	.011

Note. Percentile equivalent = the percentile of the public-school distribution at which the median separate school sits, under a normal approximation. p-values are from unadjusted within-quartile Welch's t-tests, which do not assume equal variances (d, by convention, uses the pooled SD, so d and p do not share the same variance assumption). Because IR contains extensive ties, Q1–Q3 boundaries are arbitrary among tied schools — only Q4's membership is stable across tie-breaking order, and Q4 is the quartile 4.6.3 interprets. Q1 and Q2 span only one to two CSDs each, so quartile-level comparisons there cannot be CSD-clustered.

Appendix L

R Script

The complete analysis script follows. It first constructs the four analytic samples from the source files — assessment results, the 2021 Census, the Index of Remoteness, and school locations — then estimates every model reported in Chapters 3-4 (descriptive statistics, the hierarchical models, the school-type × remoteness comparisons, interactions, and diagnostics), and finally generates the computations for all relevant appendices.

```
setwd(path.expand("~/Diss Coding/Main"))

library(data.table)
library(dplyr)
library(tidyr)
library(stringr)
library(janitor)
library(readxl)
library(readr)
library(sf)
library(estimatr)
library(car)
library(moments)
library(fwildclusterboot)

sf::sf_use_s2(FALSE)
options(scipen = 999)
out <- "replication_out"
dir.create(out, showWarnings = FALSE)

# Index of Remoteness
ir_data <- read.csv("~/Downloads/172600012020001-eng/2021IR_DATABASE.csv")
ir_csd <- ir_data %>%
  transmute(csd_id = as.integer(CSDuid),
            ir = as.numeric(Index_of_remoteness)) %>%
  filter(!is.na(csd_id), !is.na(ir))

# DB-level census spines: DB -> DA -> CSD nesting plus population-centre
# class (popctr)
bc_census <- fread("~/Diss Coding/bc_census_data.csv", data.table = FALSE) %>%
  rename(da_id = DAUID_ADIDU, csd_id = CSDUID_SDRIDU)
ab_census <- fread("~/Diss Coding/alberta_census.csv", data.table = FALSE) %>%
  rename(csd_id = CSDUID_SDRIDU)

# DA boundary polygons
da_sf <- sf::st_read("~/Downloads/lda_000a21a_e/lda_000a21a_e.shp", quiet = TRUE) %>%
```

```
  clean_names() %>%
  sf::st_make_valid() %>%
  mutate(dauid8 = str_pad(as.character(dauid), 8, pad = "0")) %>%
  select(dauid8)

# ODEF school locations
odef <- read_csv("~/Diss Coding/odef_v3_0_1.csv", show_col_types = FALSE) %>%
  clean_names()

# Census, restricted to the rows relevant to the BC side of the present study
bc_census_extract <- fread("~/Diss Coding/bc_census_extract.csv")

# DA income
bc_da_income <- bc_census_extract %>%
  filter(CHARACTERISTIC_ID == 322) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         da_income = as.numeric(C1_COUNT_TOTAL)) %>%
  filter(str_starts(dauid8, "59")) %>%
  select(dauid8, da_income)

# Indigenous population share
indig_wide_bc <- bc_census_extract %>%
  filter(CHARACTERISTIC_ID %in% c(1402, 1403)) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         value = as.numeric(C1_COUNT_TOTAL),
         key = ifelse(CHARACTERISTIC_ID == 1402, "pop_total", "n_indigenous")) %>%
  select(dauid8, key, value) %>%
  pivot_wider(names_from = key, values_from = value) %>%
  mutate(share_indigenous = n_indigenous / pop_total) %>%
  filter(str_starts(dauid8, "59"))

# Parental education (DA bachelors+ share)
parental_ed_bc <- bc_census_extract %>%
  filter(CHARACTERISTIC_ID %in% c(2014, 2024)) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         value = as.numeric(C1_COUNT_TOTAL),
         key = ifelse(CHARACTERISTIC_ID == 2014, "pop_25_64", "n_bachelors_plus")) %>%
  select(dauid8, key, value) %>%
  pivot_wider(names_from = key, values_from = value) %>%
  mutate(share_bachelors_plus = n_bachelors_plus / pop_25_64) %>%
  filter(str_starts(dauid8, "59"))

# School locations -> DA spatial join
bc_odef <- odef %>%
  filter(str_detect(tolower(provider), "british columbia|province of british columbia"))
geom_txt <- bc_odef$geometry
is_point_wkt <- !is.na(geom_txt) &
  grepl("^\\s*POINT(\\s+Z)?\\s*\\(", geom_txt, ignore.case = TRUE)
bc_sf <- sf::st_sf(bc_odef[is_point_wkt, , drop = FALSE],
                   geometry = sf::st_as_sfc(geom_txt[is_point_wkt], crs = 4326))
```

```
da_sf_bc <- da_sf %>% filter(str_starts(dauid8, "59"))
if (st_crs(bc_sf) != st_crs(da_sf_bc)) bc_sf <- st_transform(bc_sf, st_crs(da_sf_bc))
bc_schools_with_da <- st_join(bc_sf, da_sf_bc, join = st_within, left = TRUE)

# DA-level popctr: modal popctr across DBs within each DA
da_popctr_bc <- bc_census %>%
  filter(!is.na(popctr), !is.na(da_id)) %>%
  mutate(dauid8 = str_pad(as.character(da_id), 8, pad = "0")) %>%
  group_by(dauid8) %>%
  summarise(popctr = as.integer(names(sort(table(popctr), decreasing = TRUE))[1]),
            .groups = "drop")

# DA -> CSD lookup (for clustering and the IR join)
da_to_csd_bc <- bc_census %>%
  distinct(da_id, csd_id) %>%
  mutate(dauid8 = str_pad(as.character(da_id), 8, pad = "0"),
         csd_uid = as.character(csd_id))

# Building corrected BC samples
NEW_FSA <- path.expand(
  "~/Downloads/foundational_skills_assessment_2021-22_to_2025-26_residents_only (2).xlsx")
# This is called NEW_FSA for clarity as there were some schools added
# to the FSA database mid-way through the project.

fsa_school_new <- read_excel(NEW_FSA) %>%
  filter(DATA_LEVEL == "School Level") %>%
  mutate(AVG_SCORE = as.numeric(AVG_SCORE),
         NUMBER_WRITERS = as.numeric(NUMBER_WRITERS))

type_lookup <- fsa_school_new %>%
  mutate(school_id = as.character(as.integer(SCHOOL_NUMBER))) %>%
  distinct(school_id, PUBLIC_OR_INDEPENDENT)

# Sample-flow inputs BC
g7_raw <- fsa_school_new %>%
  filter(SUB_POPULATION == "All Students", SCHOOL_YEAR == "2024/2025",
         as.integer(GRADE) == 7)
cat("BC G7 schools reporting:", n_distinct(as.integer(g7_raw$SCHOOL_NUMBER)), "\n")
g7_valid <- g7_raw %>% filter(is.finite(AVG_SCORE), AVG_SCORE > 0)
cat("BC G7 with valid scores:", n_distinct(as.integer(g7_valid$SCHOOL_NUMBER)), "\n")

# BC School type
bc_school_type <- fsa_school_new %>%
  filter(AVG_SCORE > 0) %>%
  transmute(school_id = as.character(as.integer(SCHOOL_NUMBER)),
            school_type = PUBLIC_OR_INDEPENDENT) %>%
  distinct(school_id, .keep_all = TRUE)

# BC school-level covariate frame
bc_schools_with_ir <- bc_schools_with_da %>%
```

```
  mutate(school_id = as.character(school_id)) %>%
  left_join(bc_da_income, by = "dauid8") %>%
  left_join(indig_wide_bc %>% select(dauid8, share_indigenous), by = "dauid8") %>%
  left_join(parental_ed_bc %>% select(dauid8, share_bachelors_plus), by = "dauid8") %>%
  left_join(bc_school_type, by = "school_id") %>%
  left_join(da_popctr_bc, by = "dauid8") %>%
  sf::st_drop_geometry() %>%
  left_join(da_to_csd_bc %>% select(dauid8, csd_uid), by = "dauid8") %>%
  left_join(ir_csd %>% transmute(csd_uid = as.character(csd_id), ir_score = ir),
            by = "csd_uid")

# Distance-learning schools were excluded as they don't reflect their DA;
# identified by name filter on the FSA file ('distance', 'virtual' etc)
# the list is fixed here so the sample is stable across FSA data releases/runs.

dl_school_ids <- c("2396738","2399159","3499101","3896843","3996721","3999214",
                   "4099136","5796841","5798005","5898002","6196745","6396767",
                   "6398007","7198008","7396838","7999145","8396713","8396714",
                   "9199085")

bc_sch_df_new <- fsa_school_new %>%
  filter(SUB_POPULATION == "All Students", as.integer(GRADE) %in% c(4, 7)) %>%
  mutate(school_id = as.character(as.integer(SCHOOL_NUMBER)),
         avg_score = AVG_SCORE, grade = as.integer(GRADE),
         skill = tolower(FSA_SKILL_CODE)) %>%
  filter(is.finite(avg_score), avg_score > 0) %>%
  left_join(
    bc_schools_with_ir %>%
      mutate(school_id = as.character(school_id)) %>%
      select(school_id, dauid8, da_income, share_bachelors_plus,
             popctr, ir_score, share_indigenous, school_type, csd_uid),
    by = "school_id") %>%
  mutate(da_income_k = da_income / 1000)

# BC Grade 7 analytic sample
bc_g7_df <- bc_sch_df_new %>%
  filter(SCHOOL_YEAR == "2024/2025", grade == 7,
         !school_id %in% dl_school_ids) %>%
  select(school_id, skill, avg_score, dauid8, da_income_k,
         share_bachelors_plus, popctr, ir_score, share_indigenous,
         school_type, csd_uid) %>%
  pivot_wider(names_from = skill, values_from = avg_score,
              names_prefix = "score_") %>%
  mutate(avg_score = rowMeans(cbind(score_literacy, score_numeracy),
                              na.rm = TRUE)) %>%
  filter(is.finite(avg_score), is.finite(da_income_k),
         is.finite(share_bachelors_plus), is.finite(popctr),
         is.finite(ir_score), is.finite(share_indigenous)) %>%
  left_join(type_lookup %>% rename(fsa_type = PUBLIC_OR_INDEPENDENT),
            by = "school_id") %>%
```

```
  mutate(school_type = coalesce(school_type, fsa_type)) %>%
  select(-fsa_type) %>%
  mutate(school_type = factor(school_type,
                              levels = c("Public School", "Independent School")))

# Grade 4 sample (used in robustness checks)
bc_g4_df <- bc_sch_df_new %>%
  filter(SCHOOL_YEAR == "2024/2025", grade == 4,
         !school_id %in% dl_school_ids) %>%
  select(school_id, skill, avg_score, dauid8, da_income_k,
         share_bachelors_plus, popctr, ir_score, share_indigenous,
         school_type, csd_uid) %>%
  pivot_wider(names_from = skill, values_from = avg_score,
              names_prefix = "score_") %>%
  mutate(avg_score = rowMeans(cbind(score_literacy, score_numeracy),
                              na.rm = TRUE)) %>%
  filter(is.finite(avg_score), is.finite(da_income_k),
         is.finite(share_bachelors_plus), is.finite(popctr),
         is.finite(ir_score), is.finite(share_indigenous)) %>%
  left_join(type_lookup %>% rename(fsa_type = PUBLIC_OR_INDEPENDENT),
            by = "school_id") %>%
  mutate(school_type = coalesce(school_type, fsa_type)) %>%
  select(-fsa_type) %>%
  mutate(school_type = factor(school_type,
                              levels = c("Public School", "Independent School")))

cat("BC G7 n =", nrow(bc_g7_df), "| BC G4 n =", nrow(bc_g4_df), "\n")
print(table(bc_g7_df$school_type))

# z-scored versions of both
bc_g7_std <- bc_g7_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))

bc_g4_std <- bc_g4_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))

# Building Alberta samples

# PAT school-level scores
pat2 <- read.csv("~/Diss Coding/pat2.csv") %>% clean_names()
```

```
# 2021 Census -- rows that are relevant to Albertan side of the
# present study

ab_census_extract <- fread("~/Diss Coding/ab_census_extract.csv")

alberta_da_income <- ab_census_extract %>%
  filter(CHARACTERISTIC_ID == 322) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         da_income = as.numeric(C1_COUNT_TOTAL)) %>%
  filter(str_starts(dauid8, "48")) %>%
  select(dauid8, da_income)

indig_wide_ab <- ab_census_extract %>%
  filter(CHARACTERISTIC_ID %in% c(1402, 1403)) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         value = as.numeric(C1_COUNT_TOTAL),
         key = ifelse(CHARACTERISTIC_ID == 1402, "pop_total", "n_indigenous")) %>%
  select(dauid8, key, value) %>%
  pivot_wider(names_from = key, values_from = value) %>%
  mutate(share_indigenous = n_indigenous / pop_total) %>%
  filter(str_starts(dauid8, "48"))

parental_ed_ab <- ab_census_extract %>%
  filter(CHARACTERISTIC_ID %in% c(2014, 2024)) %>%
  mutate(dauid8 = str_pad(as.character(ALT_GEO_CODE), 8, pad = "0"),
         value = as.numeric(C1_COUNT_TOTAL),
         key = ifelse(CHARACTERISTIC_ID == 2014, "pop_25_64", "n_bachelors_plus")) %>%
  select(dauid8, key, value) %>%
  pivot_wider(names_from = key, values_from = value) %>%
  mutate(share_bachelors_plus = n_bachelors_plus / pop_25_64) %>%
  filter(str_starts(dauid8, "48"))

# School locations -> DA spatial join. Some schools had no coordinates and
# were geocoded once from their addresses (R's tidygeocoder package).
# Results were stored in ab_geocoded.csv and read back here. Geocoding is not
# re-done here as it can vary slightly between runs; having one fixed geocoding
# file ensures approximated locations of schools are held constant across runs.

ab_odef <- odef %>% filter(provider == "Province of Alberta")
geom_txt_ab <- ab_odef$geometry
is_point_wkt_ab <- !is.na(geom_txt_ab) &
  grepl("^\\s*POINT(\\s+Z)?\\s*\\(", geom_txt_ab, ignore.case = TRUE)
ab_sf <- sf::st_sf(ab_odef[is_point_wkt_ab, , drop = FALSE],
                   geometry = sf::st_as_sfc(geom_txt_ab[is_point_wkt_ab], crs = 4326))
ab_geocoded_sf <- read.csv("~/Diss Coding/ab_geocoded.csv", stringsAsFactors = FALSE) %>%
  filter(!is.na(lat), !is.na(lon)) %>%
  st_as_sf(coords = c("lon", "lat"), crs = 4326)

da_sf_ab <- da_sf %>% filter(str_starts(dauid8, "48"))
```

```
if (st_crs(ab_sf) != st_crs(da_sf_ab)) ab_sf <- st_transform(ab_sf, st_crs(da_sf_ab))
if (st_crs(ab_geocoded_sf) != st_crs(da_sf_ab)) ab_geocoded_sf <- st_transform(ab_geocoded_sf,
st_crs(da_sf_ab))

coerce_non_geom_to_char <- function(sf_obj) {
  sf_obj %>% mutate(across(where(\(col) !inherits(col, "sfc")), as.character))
}
ab_schools_with_da <- bind_rows(
  coerce_non_geom_to_char(st_join(ab_sf, da_sf_ab, join = st_within, left = TRUE)), coerce_non_geom_to_char(st_join
  (ab_geocoded_sf, da_sf_ab, join = st_within, left = TRUE))) %>% sf::st_drop_geometry()

# DA-level popctr
da_popctr_ab <- ab_census %>%
  filter(!is.na(popctr), !is.na(da_id)) %>%
  mutate(dauid8 = str_pad(as.character(da_id), 8, pad = "0")) %>%
  group_by(dauid8) %>%
  summarise(popctr = as.integer(names(sort(table(popctr), decreasing = TRUE))[1]),
            .groups = "drop")

da_to_csd_ab <- ab_census %>%
  distinct(da_id, csd_id) %>%
  mutate(dauid8 = str_pad(as.character(da_id), 8, pad = "0"),
         csd_uid = as.character(csd_id))

# School type
ab_school_type <- pat2 %>%
  distinct(school_code_original, authority_type) %>%
  mutate(school_id = as.character(as.integer(school_code_original))) %>%
  select(school_id, authority_type) %>%
  filter(!is.na(school_id), !is.na(authority_type))

# AB covariate frame
ab_schools_with_ir <- ab_schools_with_da %>%
  left_join(alberta_da_income, by = "dauid8") %>%
  left_join(indig_wide_ab %>% select(dauid8, share_indigenous), by = "dauid8") %>%
  left_join(parental_ed_ab %>% select(dauid8, share_bachelors_plus), by = "dauid8") %>%
  left_join(ab_school_type, by = "school_id") %>%
  left_join(da_popctr_ab, by = "dauid8") %>%
  left_join(da_to_csd_ab %>% select(dauid8, csd_uid), by = "dauid8") %>%
  left_join(ir_csd %>% transmute(csd_uid = as.character(csd_id), ir_score = ir),
            by = "csd_uid") %>%
  mutate(school_id = as.character(school_id))

ab_covariates <- ab_schools_with_ir %>%
  mutate(school_id = as.character(as.integer(school_id))) %>%
  select(school_id, dauid8, da_income, share_bachelors_plus,
         popctr, ir_score, share_indigenous, authority_type)

# Distance-learning/outreach exclusion — name filter was run externally on
```

```
# PAT school names, as was done for bc
ab_dl_school_codes <- c("319","1233","1546","1724","1903","1996","2317",
                        "2344","2349","2353","2379","2389","2419")

# PAT long frame
pat2_long <- pat2 %>%
  mutate(
    school_code = as.character(as.integer(school_code_original)),
    skill = case_when(
      str_detect(course_name, regex("english lang arts|english language arts", ignore_case = TRUE)) ~ "literacy",
      str_detect(course_name, regex("^mathematics", ignore_case = TRUE)) ~ "numeracy",
      TRUE ~ NA_character_),
    grade = case_when(
      str_detect(course_name, " 6") ~ 6L,
      str_detect(course_name, " 9") ~ 9L,
      TRUE ~ NA_integer_)) %>%
  filter(!is.na(skill), !is.na(grade), language == "Written in English",
         !school_code %in% ab_dl_school_codes)

make_ab_df <- function(grade_val, year_col) {
  pat2_long %>%
    filter(grade == grade_val) %>%
    mutate(mean_pct = suppressWarnings(as.numeric(.data[[year_col]]))) %>%
    filter(is.finite(mean_pct)) %>%
    select(school_code, skill, mean_pct) %>%
    pivot_wider(names_from = skill, values_from = mean_pct,
                names_prefix = "mean_pct_", values_fn = mean) %>%
    left_join(ab_covariates, by = c("school_code" = "school_id")) %>%
    mutate(da_income_k = da_income / 1000) %>%
    filter(is.finite(da_income_k), is.finite(share_bachelors_plus),
           is.finite(popctr), is.finite(ir_score), is.finite(share_indigenous))
}

ab_g6_2025 <- make_ab_df(6, "x2025sch_mean")
ab_g9_2025 <- make_ab_df(9, "x2025sch_mean")

# AB Grade 6 analytic sample
ab_g6_df <- ab_g6_2025 %>%
  left_join(ab_schools_with_ir %>%
              mutate(school_id = as.character(as.integer(school_id))) %>%
              select(school_id, csd_uid),
            by = c("school_code" = "school_id")) %>%
  mutate(
    da_income_k = da_income / 1000,
    school_type = case_when(
      authority_type == "Public"   ~ "Public",
      authority_type == "Separate" ~ "Separate",
      authority_type == "Charter"  ~ "Charter",
      authority_type == "Private"  ~ "Private",
      TRUE                         ~ "Other"),
```

```
    school_type = factor(school_type,
                         levels = c("Public", "Separate", "Charter", "Private", "Other")),
    avg_score = rowMeans(cbind(mean_pct_literacy, mean_pct_numeracy), na.rm = TRUE)) %>%
  filter(authority_type != "Francophone",
         is.finite(avg_score), is.finite(da_income_k),
         is.finite(share_bachelors_plus), is.finite(popctr),
         is.finite(ir_score), is.finite(share_indigenous),
         !is.na(csd_uid))

# Grade 9 sample (robustness checks)
ab_g9_df <- ab_g9_2025 %>%
  left_join(ab_schools_with_ir %>%
              mutate(school_id = as.character(as.integer(school_id))) %>%
              select(school_id, csd_uid),
            by = c("school_code" = "school_id")) %>%
  mutate(
    da_income_k = da_income / 1000,
    school_type = case_when(
      authority_type == "Public"   ~ "Public",
      authority_type == "Separate" ~ "Separate",
      authority_type == "Charter"  ~ "Charter",
      authority_type == "Private"  ~ "Private",
      TRUE                         ~ "Other"),
    school_type = factor(school_type,
                         levels = c("Public", "Separate", "Charter", "Private", "Other")),
    avg_score = rowMeans(cbind(mean_pct_literacy, mean_pct_numeracy), na.rm = TRUE)) %>%
  filter(authority_type != "Francophone",
         is.finite(avg_score), is.finite(da_income_k),
         is.finite(share_bachelors_plus), is.finite(popctr),
         is.finite(ir_score), is.finite(share_indigenous),
         !is.na(csd_uid))

cat("AB G6 n =", nrow(ab_g6_df), "| AB G9 n =", nrow(ab_g9_df), "\n")
print(table(ab_g6_df$school_type))

# z-scored versions
ab_g6_std <- ab_g6_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))

ab_g9_std <- ab_g9_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
```

```
        share_indigenous_z = as.numeric(scale(share_indigenous)))

# Descriptives
t41 <- rbind(
  bc_g7_df %>% summarise(Cohort = "BC G7", n = n(),
    score_m = round(mean(avg_score), 2), score_sd = round(sd(avg_score), 2),
    income_m = round(mean(da_income_k), 2), income_sd = round(sd(da_income_k), 2),
    bach_m = round(mean(share_bachelors_plus), 2), bach_sd = round(sd(share_bachelors_plus), 2),
    popctr_m = round(mean(popctr), 2), popctr_sd = round(sd(popctr), 2),
    ir_m = round(mean(ir_score), 2), ir_sd = round(sd(ir_score), 2),
    indig_m = round(mean(share_indigenous), 2), indig_sd = round(sd(share_indigenous), 2)),
  bc_g4_df %>% summarise(Cohort = "BC G4", n = n(),
    score_m = round(mean(avg_score), 2), score_sd = round(sd(avg_score), 2),
    income_m = round(mean(da_income_k), 2), income_sd = round(sd(da_income_k), 2),
    bach_m = round(mean(share_bachelors_plus), 2), bach_sd = round(sd(share_bachelors_plus), 2),
    popctr_m = round(mean(popctr), 2), popctr_sd = round(sd(popctr), 2),
    ir_m = round(mean(ir_score), 2), ir_sd = round(sd(ir_score), 2),
    indig_m = round(mean(share_indigenous), 2), indig_sd = round(sd(share_indigenous), 2)),
  ab_g6_df %>% summarise(Cohort = "AB G6", n = n(),
    score_m = round(mean(avg_score), 2), score_sd = round(sd(avg_score), 2),
    income_m = round(mean(da_income_k), 2), income_sd = round(sd(da_income_k), 2),
    bach_m = round(mean(share_bachelors_plus), 2), bach_sd = round(sd(share_bachelors_plus), 2),
    popctr_m = round(mean(popctr), 2), popctr_sd = round(sd(popctr), 2),
    ir_m = round(mean(ir_score), 2), ir_sd = round(sd(ir_score), 2),
    indig_m = round(mean(share_indigenous), 2), indig_sd = round(sd(share_indigenous), 2)),
  ab_g9_df %>% summarise(Cohort = "AB G9", n = n(),
    score_m = round(mean(avg_score), 2), score_sd = round(sd(avg_score), 2),
    income_m = round(mean(da_income_k), 2), income_sd = round(sd(da_income_k), 2),
    bach_m = round(mean(share_bachelors_plus), 2), bach_sd = round(sd(share_bachelors_plus), 2),
    popctr_m = round(mean(popctr), 2), popctr_sd = round(sd(popctr), 2),
    ir_m = round(mean(ir_score), 2), ir_sd = round(sd(ir_score), 2),
    indig_m = round(mean(share_indigenous), 2), indig_sd = round(sd(share_indigenous), 2)))
print(t41)

#Hierarchical models

# BC Grade 7
bc_g7_m1 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z,
                      data = bc_g7_std, clusters = csd_uid)
bc_g7_m2 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z,
                      data = bc_g7_std, clusters = csd_uid)
bc_g7_m3 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z,
                      data = bc_g7_std, clusters = csd_uid)
bc_g7_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z + school_type,
                      data = bc_g7_std, clusters = csd_uid)
print(summary(bc_g7_m1)); print(summary(bc_g7_m2))
print(summary(bc_g7_m3)); print(summary(bc_g7_m4))
```

```
# BC Grade 4
bc_g4_m1 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z,
                      data = bc_g4_std, clusters = csd_uid)
bc_g4_m2 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z,
                      data = bc_g4_std, clusters = csd_uid)
bc_g4_m3 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z,
                      data = bc_g4_std, clusters = csd_uid)
bc_g4_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z + school_type,
                      data = bc_g4_std, clusters = csd_uid)
print(summary(bc_g4_m1)); print(summary(bc_g4_m2))
print(summary(bc_g4_m3)); print(summary(bc_g4_m4))

# Alberta Grade 6
ab_g6_m1 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z,
                      data = ab_g6_std, clusters = csd_uid)
ab_g6_m2 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z,
                      data = ab_g6_std, clusters = csd_uid)
ab_g6_m3 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z,
                      data = ab_g6_std, clusters = csd_uid)
ab_g6_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z + school_type,
                      data = ab_g6_std, clusters = csd_uid)
print(summary(ab_g6_m1)); print(summary(ab_g6_m2))
print(summary(ab_g6_m3)); print(summary(ab_g6_m4))

# Alberta Grade 9
ab_g9_m1 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z,
                      data = ab_g9_std, clusters = csd_uid)
ab_g9_m2 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z,
                      data = ab_g9_std, clusters = csd_uid)
ab_g9_m3 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z,
                      data = ab_g9_std, clusters = csd_uid)
ab_g9_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                        popctr_z + ir_score_z + share_indigenous_z + school_type,
                      data = ab_g9_std, clusters = csd_uid)
print(summary(ab_g9_m1)); print(summary(ab_g9_m2))
print(summary(ab_g9_m3)); print(summary(ab_g9_m4))

# R^2 by block
bc_g7_l1 <- lm(avg_score ~ da_income_k + share_bachelors_plus, data = bc_g7_df)
bc_g7_l2 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score, data = bc_g7_df)
```

```
bc_g7_l3 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous, data
= bc_g7_df)
bc_g7_l4 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous +
school_type, data = bc_g7_df)
r2_bc_g7 <- c(summary(bc_g7_l1)$r.squared, summary(bc_g7_l2)$r.squared,
              summary(bc_g7_l3)$r.squared, summary(bc_g7_l4)$r.squared)
cat("BC G7 R2:", round(r2_bc_g7, 4), "| increments:", round(diff(r2_bc_g7), 4), "\n")

bc_g4_l1 <- lm(avg_score ~ da_income_k + share_bachelors_plus, data = bc_g4_df)
bc_g4_l2 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score, data = bc_g4_df)
bc_g4_l3 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous, data
= bc_g4_df)
bc_g4_l4 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous +
school_type, data = bc_g4_df)
r2_bc_g4 <- c(summary(bc_g4_l1)$r.squared, summary(bc_g4_l2)$r.squared,
              summary(bc_g4_l3)$r.squared, summary(bc_g4_l4)$r.squared)
cat("BC G4 R2:", round(r2_bc_g4, 4), "| increments:", round(diff(r2_bc_g4), 4), "\n")

ab_g6_l1 <- lm(avg_score ~ da_income_k + share_bachelors_plus, data = ab_g6_df)
ab_g6_l2 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score, data = ab_g6_df)
ab_g6_l3 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous, data
= ab_g6_df)
ab_g6_l4 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous +
school_type, data = ab_g6_df)
r2_ab_g6 <- c(summary(ab_g6_l1)$r.squared, summary(ab_g6_l2)$r.squared,
              summary(ab_g6_l3)$r.squared, summary(ab_g6_l4)$r.squared)
cat("AB G6 R2:", round(r2_ab_g6, 4), "| increments:", round(diff(r2_ab_g6), 4), "\n")

ab_g9_l1 <- lm(avg_score ~ da_income_k + share_bachelors_plus, data = ab_g9_df)
ab_g9_l2 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score, data = ab_g9_df)
ab_g9_l3 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous, data
= ab_g9_df)
ab_g9_l4 <- lm(avg_score ~ da_income_k + share_bachelors_plus + popctr + ir_score + share_indigenous +
school_type, data = ab_g9_df)
r2_ab_g9 <- c(summary(ab_g9_l1)$r.squared, summary(ab_g9_l2)$r.squared,
              summary(ab_g9_l3)$r.squared, summary(ab_g9_l4)$r.squared)
cat("AB G9 R2:", round(r2_ab_g9, 4), "| increments:", round(diff(r2_ab_g9), 4), "\n")

# VIFs and key correlations
print(round(vif(bc_g7_l4), 3))
cat("BC popctr-IR r:", round(cor(bc_g7_df$popctr, bc_g7_df$ir_score), 3), "\n")
cat("BC IR-Indigenous r:", round(cor(bc_g7_df$ir_score, bc_g7_df$share_indigenous), 3), "\n")
cat("AB G6 popctr-IR r:", round(cor(ab_g6_df$popctr, ab_g6_df$ir_score), 3), "\n")
cat("AB G9 popctr-IR r:", round(cor(ab_g9_df$popctr, ab_g9_df$ir_score), 3), "\n")

# Table 4.5 + regional concentration
m_type_only <- lm_robust(avg_score_z ~ school_type, data = bc_g7_std,
                         clusters = csd_uid)
print(summary(m_type_only))
```

```
dq <- bc_g7_df %>% mutate(q = ntile(ir_score, 4))
t44 <- dq %>% group_by(q, school_type) %>%
  summarise(m = round(mean(avg_score), 1), n = n(), .groups = "drop")
print(as.data.frame(t44))

for (qq in 1:4) {
  pub <- dq$avg_score[dq$q == qq & dq$school_type == "Public School"]
  ind <- dq$avg_score[dq$q == qq & dq$school_type == "Independent School"]
  pooled_sd <- sqrt(((length(ind) - 1) * var(ind) + (length(pub) - 1) * var(pub)) /
                    (length(ind) + length(pub) - 2))
  d <- (mean(ind) - mean(pub)) / pooled_sd
  cat("Q", qq, ": gap =", round(mean(ind) - mean(pub), 1),
      " d =", round(d, 2),
      " welch p =", signif(t.test(ind, pub)$p.value, 3), "\n")
}

ir_names <- ir_data %>% transmute(csd_uid = as.character(CSDuid), csd_name = CSDname)
lm_pat  <- paste0("Vancouver|Surrey|Burnaby|Richmond|Coquitlam|^Delta|",
                  "New Westminster|Langley|Maple Ridge|Pitt Meadows|Port Moody|",
                  "White Rock|Abbotsford|Chilliwack|Anmore|Belcarra|Lions Bay|",
                  "Bowen Island|Mission")
vic_pat <- "Victoria|Saanich|Oak Bay|Esquimalt|Colwood|Langford|Metchosin|Sooke|Sidney|View
Royal|Highlands"
kel_pat <- "Kelowna|Lake Country"
region_tab <- dq %>% filter(school_type == "Independent School") %>%
  mutate(csd_uid = as.character(csd_uid)) %>%
  left_join(ir_names, by = "csd_uid") %>%
  mutate(region = case_when(
    grepl(lm_pat, csd_name)  ~ "Lower Mainland",
    grepl(vic_pat, csd_name) ~ "Greater Victoria",
    grepl(kel_pat, csd_name) ~ "Kelowna area",
    TRUE ~ "Other"))
print(table(region_tab$region))
cat("Big-three share:", sum(region_tab$region != "Other"), "of", nrow(region_tab), "\n")
print(region_tab %>% group_by(q) %>%
        summarise(n = n(), pct_big3 = round(100 * mean(region != "Other"))))

#Interactions
bc_int_indig <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + ir_score_z * share_indigenous_z + school_type,
                        data = bc_g7_std, clusters = csd_uid)
print(summary(bc_int_indig))

bc_int_type <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                           popctr_z + ir_score_z * school_type + share_indigenous_z,
                         data = bc_g7_std, clusters = csd_uid)
print(summary(bc_int_type))

bc_int_ed <- lm_robust(avg_score_z ~ da_income_k_z + popctr_z + ir_score_z +
                       share_indigenous_z + school_type * share_bachelors_plus_z,
```

```
                    data = bc_g7_std, clusters = csd_uid)
print(summary(bc_int_ed))

ab_int_indig <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + share_indigenous_z + school_type +
                          ir_score_z * share_indigenous_z,
                        data = ab_g6_std, clusters = csd_uid)
print(summary(ab_int_indig))

ab_int_type <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                         popctr_z + share_indigenous_z + school_type * ir_score_z,
                       data = ab_g6_std, clusters = csd_uid)
print(summary(ab_int_type))

#Outcome distributions
x <- bc_g7_df$avg_score
span <- max(x) - min(x)
a_bc_g7 <- data.frame(Cohort = "BC Grade 7", N = length(x),
  Mean = round(mean(x), 1), SD = round(sd(x), 1),
  Min = round(min(x), 1), Max = round(max(x), 1),
  Skewness = round(skewness(x), 2), Kurtosis = round(kurtosis(x), 2),
  pct_bottom2 = round(100 * mean(x <= min(x) + 0.02 * span), 1),
  pct_top2 = round(100 * mean(x >= max(x) - 0.02 * span), 1))

x <- bc_g4_df$avg_score
span <- max(x) - min(x)
a_bc_g4 <- data.frame(Cohort = "BC Grade 4", N = length(x),
  Mean = round(mean(x), 1), SD = round(sd(x), 1),
  Min = round(min(x), 1), Max = round(max(x), 1),
  Skewness = round(skewness(x), 2), Kurtosis = round(kurtosis(x), 2),
  pct_bottom2 = round(100 * mean(x <= min(x) + 0.02 * span), 1),
  pct_top2 = round(100 * mean(x >= max(x) - 0.02 * span), 1))

x <- ab_g6_df$avg_score
span <- max(x) - min(x)
a_ab_g6 <- data.frame(Cohort = "AB Grade 6", N = length(x),
  Mean = round(mean(x), 1), SD = round(sd(x), 1),
  Min = round(min(x), 1), Max = round(max(x), 1),
  Skewness = round(skewness(x), 2), Kurtosis = round(kurtosis(x), 2),
  pct_bottom2 = round(100 * mean(x <= min(x) + 0.02 * span), 1),
  pct_top2 = round(100 * mean(x >= max(x) - 0.02 * span), 1))

x <- ab_g9_df$avg_score
span <- max(x) - min(x)
a_ab_g9 <- data.frame(Cohort = "AB Grade 9", N = length(x),
  Mean = round(mean(x), 1), SD = round(sd(x), 1),
  Min = round(min(x), 1), Max = round(max(x), 1),
  Skewness = round(skewness(x), 2), Kurtosis = round(kurtosis(x), 2),
  pct_bottom2 = round(100 * mean(x <= min(x) + 0.02 * span), 1),
  pct_top2 = round(100 * mean(x >= max(x) - 0.02 * span), 1))
```

```
a_stats <- rbind(a_bc_g7, a_bc_g4, a_ab_g6, a_ab_g9)
print(a_stats)

png(file.path(out, "figA1_outcome_distributions.png"), width = 1100, height = 900, res = 120)
par(mfrow = c(2, 2), mar = c(4.5, 4.5, 3, 1))
hist(bc_g7_df$avg_score, breaks = 40, col = "grey80", border = "grey40",
    xlab = "School mean FSA score", main = "BC Grade 7")
abline(v = mean(bc_g7_df$avg_score), col = "red", lwd = 2, lty = 2)
hist(bc_g4_df$avg_score, breaks = 40, col = "grey80", border = "grey40",
    xlab = "School mean FSA score", main = "BC Grade 4")
abline(v = mean(bc_g4_df$avg_score), col = "red", lwd = 2, lty = 2)
hist(ab_g6_df$avg_score, breaks = 40, col = "grey80", border = "grey40",
    xlab = "School mean PAT score (%)", main = "AB Grade 6")
abline(v = mean(ab_g6_df$avg_score), col = "red", lwd = 2, lty = 2)
hist(ab_g9_df$avg_score, breaks = 40, col = "grey80", border = "grey40",
    xlab = "School mean PAT score (%)", main = "AB Grade 9")
abline(v = mean(ab_g9_df$avg_score), col = "red", lwd = 2, lty = 2)
dev.off()

# Diagnostic panels
png(file.path(out, "figB1_bc_g7_m4_diagnostics.png"), width = 900, height = 900, res = 120)
par(mfrow = c(2, 2), mar = c(4.5, 4.5, 3, 1))
plot(fitted(bc_g7_l4), resid(bc_g7_l4), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "Residuals", main = "BC G7 M4: Residuals vs Fitted")
abline(h = 0, lty = 2, col = "grey60")
lines(lowess(fitted(bc_g7_l4), resid(bc_g7_l4)), col = "red", lwd = 2)
qqnorm(resid(bc_g7_l4), pch = 20, col = "grey40", main = "BC G7 M4: Normal Q-Q")
qqline(resid(bc_g7_l4), col = "red", lwd = 2)
plot(fitted(bc_g7_l4), sqrt(abs(rstandard(bc_g7_l4))), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "sqrt(|standardised residuals|)",
    main = "BC G7 M4: Scale-Location")
lines(lowess(fitted(bc_g7_l4), sqrt(abs(rstandard(bc_g7_l4)))), col = "red", lwd = 2)
crPlot(bc_g7_l4, variable = "ir_score", pch = 20, col = "grey40",
      main = "BC G7 M4: Partial Residual (IR score)")
dev.off()

png(file.path(out, "figB5_bc_g4_m4_diagnostics.png"), width = 900, height = 900, res = 120)
par(mfrow = c(2, 2), mar = c(4.5, 4.5, 3, 1))
plot(fitted(bc_g4_l4), resid(bc_g4_l4), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "Residuals", main = "BC G4 M4: Residuals vs Fitted")
abline(h = 0, lty = 2, col = "grey60")
lines(lowess(fitted(bc_g4_l4), resid(bc_g4_l4)), col = "red", lwd = 2)
qqnorm(resid(bc_g4_l4), pch = 20, col = "grey40", main = "BC G4 M4: Normal Q-Q")
qqline(resid(bc_g4_l4), col = "red", lwd = 2)
plot(fitted(bc_g4_l4), sqrt(abs(rstandard(bc_g4_l4))), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "sqrt(|standardised residuals|)",
    main = "BC G4 M4: Scale-Location")
lines(lowess(fitted(bc_g4_l4), sqrt(abs(rstandard(bc_g4_l4)))), col = "red", lwd = 2)
crPlot(bc_g4_l4, variable = "ir_score", pch = 20, col = "grey40",
```

```
      main = "BC G4 M4: Partial Residual (IR score)")
dev.off()

png(file.path(out, "figB3_ab_g6_m4_diagnostics.png"), width = 900, height = 900, res = 120)
par(mfrow = c(2, 2), mar = c(4.5, 4.5, 3, 1))
plot(fitted(ab_g6_l4), resid(ab_g6_l4), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "Residuals", main = "AB G6 M4: Residuals vs Fitted")
abline(h = 0, lty = 2, col = "grey60")
lines(lowess(fitted(ab_g6_l4), resid(ab_g6_l4)), col = "red", lwd = 2)
qqnorm(resid(ab_g6_l4), pch = 20, col = "grey40", main = "AB G6 M4: Normal Q-Q")
qqline(resid(ab_g6_l4), col = "red", lwd = 2)
plot(fitted(ab_g6_l4), sqrt(abs(rstandard(ab_g6_l4))), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "sqrt(|standardised residuals|)",
    main = "AB G6 M4: Scale-Location")
lines(lowess(fitted(ab_g6_l4), sqrt(abs(rstandard(ab_g6_l4)))), col = "red", lwd = 2)
crPlot(ab_g6_l4, variable = "ir_score", pch = 20, col = "grey40",
      main = "AB G6 M4: Partial Residual (IR score)")
dev.off()

png(file.path(out, "figB6_ab_g9_m4_diagnostics.png"), width = 900, height = 900, res = 120)
par(mfrow = c(2, 2), mar = c(4.5, 4.5, 3, 1))
plot(fitted(ab_g9_l4), resid(ab_g9_l4), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "Residuals", main = "AB G9 M4: Residuals vs Fitted")
abline(h = 0, lty = 2, col = "grey60")
lines(lowess(fitted(ab_g9_l4), resid(ab_g9_l4)), col = "red", lwd = 2)
qqnorm(resid(ab_g9_l4), pch = 20, col = "grey40", main = "AB G9 M4: Normal Q-Q")
qqline(resid(ab_g9_l4), col = "red", lwd = 2)
plot(fitted(ab_g9_l4), sqrt(abs(rstandard(ab_g9_l4))), pch = 20, col = "grey40",
    xlab = "Fitted values", ylab = "sqrt(|standardised residuals|)",
    main = "AB G9 M4: Scale-Location")
lines(lowess(fitted(ab_g9_l4), sqrt(abs(rstandard(ab_g9_l4)))), col = "red", lwd = 2)
crPlot(ab_g9_l4, variable = "ir_score", pch = 20, col = "grey40",
      main = "AB G9 M4: Partial Residual (IR score)")
dev.off()

# VIF Tables
cat("BC G7 M4 VIF:\n"); print(round(vif(bc_g7_l4), 2))
cat("BC G4 M4 VIF:\n"); print(round(vif(bc_g4_l4), 2))
cat("AB G6 M4 VIF:\n"); print(round(vif(ab_g6_l4), 2))
cat("AB G9 M4 VIF:\n"); print(round(vif(ab_g9_l4), 2))

bc_int_type_lm <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                     popctr_z + ir_score_z * school_type + share_indigenous_z,
                   data = bc_g7_std)
cat("BC type x IR VIF:\n"); print(round(vif(bc_int_type_lm), 2))
bc_int_indig_lm <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                      popctr_z + ir_score_z * share_indigenous_z + school_type,
                    data = bc_g7_std)
cat("BC IR x Indigenous VIF:\n"); print(round(vif(bc_int_indig_lm), 2))
ab_int_indig_lm <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
```

```
                    popctr_z + share_indigenous_z + school_type +
                    ir_score_z * share_indigenous_z,
                  data = ab_g6_std)
cat("AB IR x Indigenous VIF:\n"); print(round(vif(ab_int_indig_lm), 2))
ab_int_type_lm <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                     popctr_z + share_indigenous_z + school_type * ir_score_z,
                   data = ab_g6_std)
cat("AB type x IR VIF:\n"); print(round(vif(ab_int_type_lm), 2))

# Rank-transformed models
bc_g7_rank <- bc_g7_std %>% mutate(rank_z = as.numeric(scale(rank(avg_score))))
bc_g7_m4_rank <- lm_robust(rank_z ~ da_income_k_z + share_bachelors_plus_z +
                           popctr_z + ir_score_z + share_indigenous_z + school_type,
                         data = bc_g7_rank, clusters = csd_uid)
print(summary(bc_g7_m4_rank))

bc_g4_rank <- bc_g4_std %>% mutate(rank_z = as.numeric(scale(rank(avg_score))))
bc_g4_m4_rank <- lm_robust(rank_z ~ da_income_k_z + share_bachelors_plus_z +
                           popctr_z + ir_score_z + share_indigenous_z + school_type,
                         data = bc_g4_rank, clusters = csd_uid)
print(summary(bc_g4_m4_rank))

ab_g6_rank <- ab_g6_std %>% mutate(rank_z = as.numeric(scale(rank(avg_score))))
ab_g6_m4_rank <- lm_robust(rank_z ~ da_income_k_z + share_bachelors_plus_z +
                           popctr_z + ir_score_z + share_indigenous_z + school_type,
                         data = ab_g6_rank, clusters = csd_uid)
print(summary(ab_g6_m4_rank))

ab_g9_rank <- ab_g9_std %>% mutate(rank_z = as.numeric(scale(rank(avg_score))))
ab_g9_m4_rank <- lm_robust(rank_z ~ da_income_k_z + share_bachelors_plus_z +
                           popctr_z + ir_score_z + share_indigenous_z + school_type,
                         data = ab_g9_rank, clusters = csd_uid)
print(summary(ab_g9_m4_rank))

#Wild cluster bootstrap
set.seed(42)
dqrng::dqset.seed(42)

bc_g7_ols <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                  popctr_z + ir_score_z + share_indigenous_z + school_type,
                data = bc_g7_std)
cat("\nBootstrap BC Grade 7:\n")
for (term in names(coef(bc_g7_ols))[-1]) {
  bt <- boottest(bc_g7_ols, clustid = "csd_uid", param = term, B = 9999,
                 type = "webb", impose_null = TRUE)
  cat(term, ": boot p =", round(bt$p_val, 4), "\n")
}

bc_g4_ols <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                  popctr_z + ir_score_z + share_indigenous_z + school_type,
```

```
              data = bc_g4_std)
cat("\nBootstrap BC Grade 4:\n")
for (term in names(coef(bc_g4_ols))[-1]) {
  bt <- boottest(bc_g4_ols, clustid = "csd_uid", param = term, B = 9999,
                 type = "webb", impose_null = TRUE)
  cat(term, ": boot p =", round(bt$p_val, 4), "\n")
}

set.seed(42)
dqrng::dqset.seed(42)

ab_g6_ols <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                  popctr_z + ir_score_z + share_indigenous_z + school_type,
                data = ab_g6_std)
cat("\nBootstrap AB Grade 6:\n")
for (term in names(coef(ab_g6_ols))[-1]) {
  bt <- boottest(ab_g6_ols, clustid = "csd_uid", param = term, B = 9999,
                 type = "webb", impose_null = TRUE)
  cat(term, ": boot p =", round(bt$p_val, 4), "\n")
}

ab_g9_ols <- lm(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                  popctr_z + ir_score_z + share_indigenous_z + school_type,
                data = ab_g9_std)
cat("\nBootstrap AB Grade 9:\n")
for (term in names(coef(ab_g9_ols))[-1]) {
  bt <- boottest(ab_g9_ols, clustid = "csd_uid", param = term, B = 9999,
                 type = "webb", impose_null = TRUE)
  cat(term, ": boot p =", round(bt$p_val, 4), "\n")
}


#Supplementary descriptives
f1 <- bc_g7_df %>% group_by(popctr) %>%
  summarise(n = n(), M = round(mean(da_income_k), 2), SD = round(sd(da_income_k), 2)) %>%
  mutate(pct = round(100 * n / sum(n), 1))
cat("F1:\n"); print(as.data.frame(f1))

f2 <- ab_g6_df %>% group_by(popctr) %>%
  summarise(n = n(), M = round(mean(da_income_k), 2), SD = round(sd(da_income_k), 2)) %>%
  mutate(pct = round(100 * n / sum(n), 1))
cat("F2:\n"); print(as.data.frame(f2))

f3 <- bc_g7_df %>% mutate(q = ntile(ir_score, 4)) %>% group_by(q) %>%
  summarise(ir_min = round(min(ir_score), 3), ir_max = round(max(ir_score), 3),
            n = n(), M = round(mean(da_income_k), 2), SD = round(sd(da_income_k), 2))
cat("F3:\n"); print(as.data.frame(f3))

f4 <- ab_g6_df %>% mutate(q = ntile(ir_score, 4)) %>%
  mutate(q = ifelse(q <= 2, 2, q)) %>%    # Q1+Q2 merged (metropolitan band)
```

```
  group_by(q) %>%
  summarise(ir_min = round(min(ir_score), 3), ir_max = round(max(ir_score), 3),
            n = n(), M = round(mean(da_income_k), 2), SD = round(sd(da_income_k), 2))
cat("F4:\n"); print(as.data.frame(f4))

f5_ab <- ab_g6_df %>% count(school_type) %>%
  mutate(Province = "Alberta", pct = round(100 * n / sum(n), 1))
f5_bc <- bc_g7_df %>% count(school_type) %>%
  mutate(Province = "British Columbia", pct = round(100 * n / sum(n), 1))
f5 <- rbind(f5_ab, f5_bc)
cat("F5:\n"); print(as.data.frame(f5))

#Exurban sensitivity checks
bc_exurb <- bc_g7_df$popctr == 1 & bc_g7_df$ir_score < 0.20
ab_exurb <- ab_g6_df$popctr == 1 & ab_g6_df$ir_score < 0.20
cat("BC exurbs:", sum(bc_exurb), "of", sum(bc_g7_df$popctr == 1), "rural\n")
cat("AB exurbs:", sum(ab_exurb), "of", sum(ab_g6_df$popctr == 1), "rural\n")

g1 <- bc_g7_df %>%
  mutate(popctr2 = ifelse(popctr == 1 & ir_score < 0.20, 2, popctr)) %>%
  group_by(popctr2) %>%
  summarise(n = n(), M = round(mean(da_income_k), 2))
cat("G1:\n"); print(as.data.frame(g1))

# BC, exurbs reclassified to popctr 2
bc_recl_df <- bc_g7_df
bc_recl_df$popctr[bc_exurb] <- 2
bc_recl_std <- bc_recl_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))
bc_recl_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + ir_score_z + share_indigenous_z + school_type,
                        data = bc_recl_std, clusters = csd_uid)
print(summary(bc_recl_m4))

# BC, exurbs excluded (re-standardised on the reduced sample)
bc_excl_df <- bc_g7_df[!bc_exurb, ]
bc_excl_std <- bc_excl_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))
bc_excl_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + ir_score_z + share_indigenous_z + school_type,
```

```
                    data = bc_excl_std, clusters = csd_uid)
print(summary(bc_excl_m4))

# Alberta, exurbs reclassified
ab_recl_df <- ab_g6_df
ab_recl_df$popctr[ab_exurb] <- 2
ab_recl_std <- ab_recl_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))
ab_recl_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + ir_score_z + share_indigenous_z + school_type,
                        data = ab_recl_std, clusters = csd_uid)
print(summary(ab_recl_m4))

# Alberta, exurbs excluded
ab_excl_df <- ab_g6_df[!ab_exurb, ]
ab_excl_std <- ab_excl_df %>%
  mutate(avg_score_z = as.numeric(scale(avg_score)),
         da_income_k_z = as.numeric(scale(da_income_k)),
         share_bachelors_plus_z = as.numeric(scale(share_bachelors_plus)),
         popctr_z = as.numeric(scale(popctr)),
         ir_score_z = as.numeric(scale(ir_score)),
         share_indigenous_z = as.numeric(scale(share_indigenous)))
ab_excl_m4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                          popctr_z + ir_score_z + share_indigenous_z + school_type,
                        data = ab_excl_std, clusters = csd_uid)
print(summary(ab_excl_m4))

# Supplementary settlement size analyses, Alberta
h_g6_popctr <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + popctr_z,
                         data = ab_g6_std, clusters = csd_uid)
print(summary(h_g6_popctr))
h_g6_ir <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + ir_score_z,
                     data = ab_g6_std, clusters = csd_uid)
print(summary(h_g6_ir))
h_g6_both <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                         popctr_z + ir_score_z,
                       data = ab_g6_std, clusters = csd_uid)
print(summary(h_g6_both))

h_g9_popctr <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + popctr_z,
                         data = ab_g9_std, clusters = csd_uid)
print(summary(h_g9_popctr))
h_g9_ir <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + ir_score_z,
                     data = ab_g9_std, clusters = csd_uid)
print(summary(h_g9_ir))
```

```
h_g9_both <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z +
                       popctr_z + ir_score_z,
                     data = ab_g9_std, clusters = csd_uid)
print(summary(h_g9_both))

# Within-IR-quartile regressions, AB Grade 6
ab_g6_q <- ab_g6_std %>% mutate(ir_q = ntile(ir_score, 4),
                                popctr_f = factor(popctr, levels = 1:4))
print(ab_g6_q %>% filter(ir_q >= 3) %>% count(ir_q, popctr_f) %>% as.data.frame())

q3_data <- ab_g6_q %>% filter(ir_q == 3) %>% mutate(popctr_f = droplevels(popctr_f))
h2_q3 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + popctr_f,
                   data = q3_data, clusters = csd_uid)
print(summary(h2_q3))

q4_data <- ab_g6_q %>% filter(ir_q == 4) %>% mutate(popctr_f = droplevels(popctr_f))
h2_q4 <- lm_robust(avg_score_z ~ da_income_k_z + share_bachelors_plus_z + popctr_f,
                   data = q4_data, clusters = csd_uid)
print(summary(h2_q4))

# School type profiles, AB Grade 6
h3 <- ab_g6_df %>% group_by(school_type) %>%
  summarise(n = n(), income = round(mean(da_income_k), 1),
            bachelors = round(mean(share_bachelors_plus), 3),
            popctr = round(mean(popctr), 2),
            ir = round(mean(ir_score), 3))
cat("H3:\n"); print(as.data.frame(h3))

# Alternative block orders
bc_alt_m1 <- lm_robust(avg_score_z ~ school_type,
                       data = bc_g7_std, clusters = csd_uid)
bc_alt_m2 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z,
                       data = bc_g7_std, clusters = csd_uid)
bc_alt_m3 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z +
                         popctr_z + ir_score_z,
                       data = bc_g7_std, clusters = csd_uid)
bc_alt_m4 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z +
                         popctr_z + ir_score_z + share_indigenous_z,
                       data = bc_g7_std, clusters = csd_uid)
print(summary(bc_alt_m1)); print(summary(bc_alt_m2))
print(summary(bc_alt_m3)); print(summary(bc_alt_m4))

bc_alt_l1 <- lm(avg_score ~ school_type, data = bc_g7_df)
bc_alt_l2 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus, data = bc_g7_df)
bc_alt_l3 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus + popctr + ir_score, data =
bc_g7_df)
bc_alt_l4 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus + popctr + ir_score +
share_indigenous, data = bc_g7_df)
r2_bc_alt <- c(summary(bc_alt_l1)$r.squared, summary(bc_alt_l2)$r.squared,
               summary(bc_alt_l3)$r.squared, summary(bc_alt_l4)$r.squared)
```

```
cat("BC alt-order R2:", round(r2_bc_alt, 4), "| increments:", round(diff(r2_bc_alt), 4), "\n")

ab_alt_m1 <- lm_robust(avg_score_z ~ school_type,
                       data = ab_g6_std, clusters = csd_uid)
ab_alt_m2 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z,
                       data = ab_g6_std, clusters = csd_uid)
ab_alt_m3 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z +
                         popctr_z + ir_score_z,
                       data = ab_g6_std, clusters = csd_uid)
ab_alt_m4 <- lm_robust(avg_score_z ~ school_type + da_income_k_z + share_bachelors_plus_z +
                         popctr_z + ir_score_z + share_indigenous_z,
                       data = ab_g6_std, clusters = csd_uid)
print(summary(ab_alt_m1)); print(summary(ab_alt_m2))
print(summary(ab_alt_m3)); print(summary(ab_alt_m4))

ab_alt_l1 <- lm(avg_score ~ school_type, data = ab_g6_df)
ab_alt_l2 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus, data = ab_g6_df)
ab_alt_l3 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus + popctr + ir_score, data =
ab_g6_df)
ab_alt_l4 <- lm(avg_score ~ school_type + da_income_k + share_bachelors_plus + popctr + ir_score +
share_indigenous, data = ab_g6_df)
r2_ab_alt <- c(summary(ab_alt_l1)$r.squared, summary(ab_alt_l2)$r.squared,
               summary(ab_alt_l3)$r.squared, summary(ab_alt_l4)$r.squared)
cat("AB alt-order R2:", round(r2_ab_alt, 4), "| increments:", round(diff(r2_ab_alt), 4), "\n")

#Remoteness x School type
k1 <- bc_g7_df %>% mutate(ir_q = ntile(ir_score, 4)) %>%
  group_by(ir_q) %>%
  summarise(n = n(), n_indep = sum(school_type == "Independent School"),
            pct_indep = round(100 * mean(school_type == "Independent School"), 1))
cat("K1:\n"); print(as.data.frame(k1))

q25 <- quantile(bc_g7_df$ir_score, 0.25)
q50 <- quantile(bc_g7_df$ir_score, 0.50)
q75 <- quantile(bc_g7_df$ir_score, 0.75)
k1_cut <- bc_g7_df %>%
  mutate(ir_q = cut(ir_score, c(-Inf, q25, q50, q75, Inf), labels = 1:4)) %>%
  group_by(ir_q) %>%
  summarise(n = n(), pct_indep = round(100 * mean(school_type == "Independent School"), 1))
cat("K1 (tie-safe cut):\n"); print(as.data.frame(k1_cut))

# Independent school remoteness gradient BC
ir_ind <- bc_g7_df$ir_score[bc_g7_df$school_type == "Independent School"]
ir_pub <- bc_g7_df$ir_score[bc_g7_df$school_type == "Public School"]
n_all <- nrow(bc_g7_df)
is_ind <- as.integer(bc_g7_df$school_type == "Independent School")

r_pb <- cor(bc_g7_df$ir_score, is_ind)
t_pb <- r_pb * sqrt(n_all - 2) / sqrt(1 - r_pb^2)
p_pb <- 2 * pt(-abs(t_pb), df = n_all - 2)
```

```
cat("Point-biserial r =", round(r_pb, 3), ", p =", round(p_pb, 3), "\n")

welch <- t.test(ir_ind, ir_pub)
cat("Welch t(", round(welch$parameter, 1), ") =", round(welch$statistic, 2),
    ", p =", round(welch$p.value, 3), "\n")

pooled_sd <- sqrt(((length(ir_ind) - 1) * var(ir_ind) +
                   (length(ir_pub) - 1) * var(ir_pub)) / (n_all - 2))
cohens_d <- (mean(ir_ind) - mean(ir_pub)) / pooled_sd
cat("Cohen's d =", round(cohens_d, 2), "\n")

wilcox <- wilcox.test(ir_ind, ir_pub, exact = FALSE)
cat("Wilcoxon W =", wilcox$statistic, ", p =", round(wilcox$p.value, 3), "\n")

logit <- glm(is_ind ~ ir_score, data = bc_g7_df, family = binomial)
or_per_01 <- exp(coef(logit)["ir_score"] * 0.1)
logit_p <- summary(logit)$coefficients["ir_score", "Pr(>|z|)"]
cat("OR per 0.1 IR =", round(or_per_01, 2), ", p =", round(logit_p, 3), "\n")
```