Multilevel Modelling & MAIHDA: From Educational Segregation to Intersectional Analysis

Three Research Applications Demonstrating MAIHDA's Transformative Potential

Dr Yiyang Gao

13/06/2025

Research Profile

Dr Yiyang Gao
PhD in Advanced Quantitative Methods, Q-Step Centre, Exeter
Current Postdoc, Evidence Centre for Education, Durham University

Research Positioning: At the intersection of Quantitative Geography (Space), Demography/Sociology (Population), Education, and Health

Featured Studies in this Presentation:

1. **Ethnicity & Segregation**
    - Multilevel Analysis of Ethnic Clustering Across Local Schools
    - British Education Research Journal (R&R)
2. **Workforce Inequality**
    - Understanding Teacher Retention in England: A Multi-Cohort Survival Analysis
    - Research Papers in Education (Under Review)
3. **Accessibility, Simulation, Health**
    - Grant: Educational Equity - Transport & School Segregation
    - With Department of Sociology, Durham

What is MAIHDA?

Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy

The “Gold Standard” for Quantitative Intersectional Analysis

Treats intersectional identities as random effects in multilevel models, revealing how multiple characteristics combine to create unique experiences beyond simple additive effects.

Conventional Approach vs MAIHDA Approach:

Conventional	MAIHDA
Examines factors separately	Creates intersectional strata
Assumes additive effects	Reveals multiplicative effects
Misses critical intersections	Handles small cell sizes via shrinkage
Generic policy recommendations	Enables targeted interventions

Three Extended MAIHDA Applications:

- **Spatial MAIHDA** → School Segregation Study
- **Longitudinal MAIHDA** → Teacher Retention Study
- **Policy Evaluation MAIHDA** → Transport & Accessibility Study

Study 1: Educational Segregation - Current Approach

Research Question

Which students experience educational segregation, and how does this vary across space and time?

Current Aggregate Approach:

- **Calculates:** Segregation indices by ethnicity and area
- **Shows:** Overall levels of ethnic concentration in schools
- **Misses:** Which specific students attend segregated schools
- **Cannot reveal:** Intersectional vulnerabilities

The Segregation Paradox

Aggregate data shows decreasing segregation, but which students still experience it?

Who Attends Segregated Schools?
Current methods can't answer: - Are Pakistani boys more affected than Pakistani girls? - Does social class intersect with ethnicity? - Where do vulnerable groups cluster?

Definition: Segregated School = Schools where >70% of students are from a single ethnic group

Study 1: Spatial MAIHDA Enhancement

Individual-Level Spatial MAIHDA Innovation

Method Enhancement: - Unit: Individual students (N = 3.2 million) - Outcome: Attends segregated school (binary) - 24 intersectional strata: Ethnicity × SES × Gender - Cross-classified with spatial units (LSOAs) - Monte Carlo: 10,000 simulations

Model: `logit(Y_ijk) =  + u j[strata] + v k[LSOA] + W*spatial_effects`

Probability of Attending Segregated School:

Intersectional Group	Probability	Odds Ratio
Pakistani × Low SES × Male	**71.2%**	3.57 [3.18-3.99]
Pakistani × Low SES × Female	**68.4%**	3.24 [2.89-3.61]
Black × Low SES × Male	52.3%	2.48 [2.21-2.78]
White × High SES × Female	21.1%	Reference

Discriminatory Accuracy: ICC = 0.15
85% of variation is within intersectional groups - context matters more than identity alone

Spatial Clustering

- **Bradford:** 72% risk
- **Birmingham:** 69% risk
- **East London:** 74% risk
- **Moran's I = 0.82** (strong spatial clustering)

Critical Finding: Monte Carlo reveals Pakistani low-SES boys have 71.2% [95% CI: 68.4%-73.8%] probability of attending segregated schools, with strong spatial clustering in specific urban areas.

Study 2: Teacher Retention - Current Approach

Tracking the 2011 Teacher Cohort: Traditional Survival Analysis

Current Method Limitations: - Cox model with additive effects - Assumes: ethnicity + gender + ITT = risk - Misses intersectional experiences - Cannot identify time-specific vulnerabilities

Hidden Variation:

- **Female: 16.54%** - but ranges from 8% to 24% by intersectional group
- **No ITT: 9.75%** - but some groups retain at 18%, others at 3%

11-Year Retention Rates:

- Female: 16.54%
- Male: 9.75%
- With ITT: 16.85%
- No ITT: 9.75%

Female teacher retention ranges from 8% to 24% across intersectional groups!

Critical Limitation: These aggregate curves hide intersectional vulnerabilities. A Black male teacher without ITT in London has a fundamentally different trajectory than a White female teacher with ITT in rural areas.

Study 2: Longitudinal MAIHDA Enhancement

Revealing Intersectional Career Trajectories

Longitudinal MAIHDA Innovation: - 200 intersectional strata: Ethnicity × Gender × ITT × Region - Time-varying hazards: Captures non-linear career patterns - Individual trajectories: Random slopes for each stratum - Critical periods: Statistically identified vulnerability windows

Model: `logit(leaving) ~ poly(year,2) + (poly(year,2)|strata) + (1|school) + (1|teacher)`

Key Finding: 78% of variation is within intersectional groups - individual context dominates

Retention Statistics by Intersectional Group:

Group	Year 1 Risk	Year 2 Risk	Critical Period	11-Year Retention
Black × Male × Non-ITT × London	18%	**31%**	Year 2 Crisis	3.2%
Pakistani × Female × Non-ITT × North	**22%**	28%	Year 1 Crisis	4.8%
White × Female × ITT × North	8%	10%	Gradual	23.4%

Transformative Insight: Longitudinal MAIHDA reveals that “teacher retention” isn’t one problem but many distinct challenges. Black male non-ITT teachers in London face a Year 2 crisis (HR = 2.8, 95% CrI: [2.1-3.7]).

Study 3: London's Free School Transport Policy

Evaluating Complex Eligibility Criteria: Who Really Benefits?

London's Free Transport Policy:

Standard Eligibility: - Under 8: >2 miles to school - 8+: >3 miles to school - No safe walking route - Special educational needs

Low-Income Eligibility: - 8-11: >2 miles (nearest school) - 11-16: 2-6 miles (3 nearest schools) - 11-16: 2-15 miles (faith schools)

The Same Policy, Different Impacts:

Example: 11-year-old, low-income, 4 miles from school - **White British:** Likely uses local school (free transport) - **Pakistani Muslim:** May prefer faith school 8 miles away (not covered) - **Black Caribbean:** May avoid local "failing" school (complex choice)

Research Gap: Current models simulate transport use but can't explain WHY different groups make different choices under the same policy

Study 3: MAIHDA-Enhanced Policy Evaluation

Revealing Who Benefits and Who's Left Behind

MAIHDA + ABM Integration: - MAIHDA: Identifies differential policy impacts by intersectional group - ABM: Simulates spatial dynamics and school competition - 60 strata: Ethnicity × Income × Car × Distance - Monte Carlo: 10,000 policy scenarios

Eligibility ≠ Access:

Intersectional Group	Eligible	Take-up	Impact
Pakistani × Low Income × 4mi	Yes	42%	+1.2 schools
White × Low Income × 4mi	Yes	78%	+3.1 schools
Black × Low Income × 2.5mi	No	-	**0 schools**
Pakistani × Low Income × Faith	Yes*	68%	+4.8 schools

*Eligible for faith school provision (2-15 miles)

Monte Carlo Policy Refinements:

- **Scenario A:** Reduce threshold to 2.5 miles → 18,000 more eligible, reduces gap by 31%
- **Scenario B:** Cultural liaison for take-up → Pakistani take-up: 42%→71%, most cost-effective
- **Scenario C:** Expand faith provision → Addresses preferences, reduces segregation 24%

Transformative Finding: Pakistani families are eligible but face cultural/information barriers (42% take-up), whilst Black families in the 2.5-3 mile gap receive no support. Targeted outreach would be 3x more cost-effective than expanding eligibility.

MAIHDA: Transforming Understanding of Educational Inequalities

Evolution Across Three Studies

Study	Current Method	MAIHDA Extension	Key Innovation
School Segregation	Aggregate Indices	Spatial MAIHDA	Reveals which intersectional groups experience segregation & where they cluster
Teacher Retention	Survival Analysis	Longitudinal MAIHDA	Identifies career crisis points for specific groups
Transport Access	Agent-Based Model	Policy Evaluation MAIHDA	Quantifies differential policy impacts

Contributing to Sheffield's MAIHDA Leadership

Methodological Innovation: - Extending MAIHDA to spatial, longitudinal, and policy domains - Monte Carlo uncertainty quantification - Large dataset expertise (NPD, SWC)

Why MAIHDA Matters: - Identifies WHO needs help (intersections) - Shows WHERE to intervene (spatial) - Reveals WHEN to act (temporal) - Quantifies WHAT WORKS (policy)

Vision

Advancing intersectional quantitative methods to reveal not just that inequalities exist, but precisely **where**, **for whom**, **when**, and **why** – transforming our ability to design interventions that actually work