Population, Space and Place

WILEY

RESEARCH ARTICLE OPEN ACCESS

Mapping Educational Inequalities in Wales: Spatial and Socio-Economic Determinants of Pupils' Attainment

Alexandra Sandu[1,2,3] | Katy Huxley[4] | Jen Keating[2,4,5] | Tony Whiffen[6] | Rob French[2]

[1]School of Geography and Planning, Cardiff University, Cardiff, Wales, UK | [2]Administrative Research Data Wales, Cardiff University, Cardiff, Wales, UK | [3]Current affiliation: London School of Economics and Political Science, London, UK | [4]Social Science Research Park (SPARK), Cardiff University, Cardiff, Wales, UK | [5]School of Social Science, Cardiff University, Cardiff, Wales, UK | [6]Welsh Government, Cardiff, Wales, UK

Correspondence: Alexandra Sandu (a.sandu@lse.ac.uk)

Received: 8 July 2025 | **Revised:** 28 January 2026 | **Accepted:** 4 February 2026

Funding: Economic and Social Research Council, Grant/Award Number: ES/W012227/1

Keywords: administrative data linkage | educational inequalities | geographically weighted regression | socio-economic status | spatial analysis

ABSTRACT

Understanding spatial variations in educational outcomes is important for addressing educational inequalities. This study examines how socio-economic factors and household characteristics influence age 16 standardised attainment across Wales using linked administrative and census data. In terms of methodology, we employed logistic regression modelling at the individual level, while at the Lower Layer Super Output Area level, we used both Ordinary Least Squares and Geographically Weighted Regression. At the Individual level, results reveal strong associations between attainment and household characteristics, with household education level having positive effects, while socio-economic disadvantage is negatively associated with attainment. The spatial analysis highlights significant variations in how these factors impact attainment across Wales. Household education level shows consistently positive effects throughout the country, while eligibility for free school meals and special educational needs demonstrate varying negative associations across small geographies. Overall, this study provides novel insights into the complex relationship between place, socio-economic status, and educational outcomes in Wales. These findings suggest that one-size-fits-all educational policies may be insufficient and emphasise the need for geographically targeted interventions.

1 | Introduction

In recent years, educational inequalities have remained a persistent challenge in Wales, despite policy efforts to improve outcomes for all pupils. Both the 2018 and 2022 Programme for International Student Assessment (PISA) results highlighted that Wales lags behind other UK nations in the three key subjects tested—reading, mathematics, and science. The 2022 results were particularly concerning, marking the lowest scores ever recorded and falling below the OECD average. However, it is important to acknowledge that this decrease in scores reflects a broader global trend (Cadwallader et al. 2023a, 2023b; Ingram et al. 2023; Sibieta 2024; Sizmur et al. 2019).

This raises important questions about the underlying factors contributing to these disparities in educational performance. While socio-economic factors, including higher poverty levels and a diverse ethnic mix, are often cited as potential explanations, the evidence suggests that these elements alone do not account for low educational outcomes observed in Wales. The average pupil in Wales performed at a level comparable to the most disadvantaged children in England, and the gap in GCSE results between disadvantaged and other students in Wales was notably larger than that in England (Sibieta 2024).

Socio-economic factors are integral to understanding and improving academic performance, with a growing body of research

establishing strong links between household characteristics, socio-economic factors, and educational outcomes. Factors such as parental education, household income, family structure, and neighbourhood deprivation have been shown to influence a pupil's academic outcomes, although these influences vary across different social contexts and education systems (Bandyopadhyay et al. 2023; Broer et al. 2019; Imeraj et al. 2024; Perales et al. 2023).

Nevertheless, while studies have examined these relationships at a broad level, there is increasing recognition of the importance of spatial variations in educational outcomes (Banerjee et al. 2022; Gulson and Symes 2007; Smith et al. 2019; Zuccotti 2019). The impact of socio-economic factors on pupil attainment may vary significantly across different geographic areas so that a more nuanced, spatially-conscious approach to understanding and addressing educational inequalities may be needed (Ansong et al. 2015; Fotheringham et al. 2001).

Therefore, this study investigated how socio-economic factors, family type, and household characteristics relate to - and interact with – pupils' academic performance within Wales, with a focus on smaller geographic areas (specifically, Lower Layer Super Output Areas – LSOAs –, in which approximately 1,600 individuals reside). Specifically, the following research questions will be addressed:

1. How important are different measures of socio-economic status (SES) in determining educational outcomes?
2. What is the association between SES, household composition factors (family type) and pupil performance, and does geography moderate this relationship?

The paper is structured into four sections. The first provides an overview of research that has considered the association between household characteristics, socio-economic factors and pupil attainment. We highlight that while educational research has examined urban-rural differences in Wales, there has been limited spatial analysis beyond this simple divide when studying variations in pupil attainment across the country. The second section details the study's data sources and methodological approach utilising both logistic regression at the individual level and Geographically Weighted Regression (GWR) at the LSOA level. This dual approach allows us to capture both overall trends, and spatial variations in the relationships between socio-economic factors and educational outcomes across Wales. The third section presents the results, starting with descriptive statistics, followed by logistic regression results at individual level, OLS results at LSOA level, and finally, the GWR results at LSOA level, examining spatial variations in the relationships between household characteristics, socio-economic factors, and pupil attainment across LSOAs. The fourth section discusses the findings, comparing global and local (GWR) models, and explores the implications for educational disparities in Wales. Our analysis reveals significant spatial variations in how socio-economic factors influence educational outcomes, with particularly strong effects of household education level, special education needs (SEN) and free school meal eligibility (eFSM) varying across different regions of Wales. The conclusion summarises the key findings (while acknowledging limitations of the study) and emphasises the importance of considering spatial variations in socio-economic factors and their impact on educational outcomes in Wales.

Overall, while educational inequalities have been studied across the UK, particularly in England, few studies if any have undertaken a geospatial analysis of educational outcomes across Wales. This study offers a disaggregated spatial analysis of how socio-economic inequalities translate into educational attainment across the entire Welsh territory. By integrating linked administrative data with spatial regression methods, this paper contributes to a growing international literature using GWR to explore how socio-economic disadvantage operates unevenly across space (Chew et al. 2020; Ansong et al. 2015; Kopecna et al. 2021). Moreover, the application of GWR to educational data in a European context is still limited, and thus this study adds both methodological and theoretical insights to spatial inequality research.

2 | Theoretical Considerations

The influence of individual and household characteristics on educational outcomes has been a cornerstone of educational research for decades, and recent studies continue to refine the understanding of these relationships.

SES remains a key factor influencing pupils' educational attainment. A comprehensive meta-analysis by Sirin (2005) confirmed a strong correlation between SES and academic achievement. However, recent work by Chmielewski (2019) suggests that this relationship has strengthened over time in many countries, raising concerns about growing educational inequality. In the UK, researchers have revealed both persistent patterns, and emerging complexities in this relationship which brings into question the role of household characteristics, such as family income or social grade and their influence on educational performance. For instance, research indicates that family income has a causal relationship with educational attainment, as demonstrated by Blanden (2004) who found a strong correlation between family income and educational outcomes over time in the UK. Moreover, Gorard and Siddiqui (2019) conducted an analysis of the attainment gap in England, focusing on the impact of eligibility for free school meals (eFSM)—a common proxy for low SES in UK educational research (Ilie et al. 2017; Taylor 2018). While they confirmed the persistent nature of the attainment gap, their longitudinal analysis revealed that the gap has narrowed slightly over time, particularly in primary education which could suggest that targeted interventions may be having some positive effects. Furthermore, in Wales, Herbert and Thomas (1998) analysis revealed that schools in more affluent areas often outperformed those in deprived areas, with factors such as social class, unemployment rates, and housing conditions strongly associated with GCSE results. The complexity of the SES-attainment relationship is further highlighted by Strand (2014), who examined the association between SES, ethnicity, and gender in educational achievement, revealing that while low SES negatively impacts attainment across all ethnic groups, the magnitude of this effect varies significantly. These dynamics are also inherently intersectional. Gender, ethnicity, and class do not operate as independent predictors but combine to shape pupils' experiences and outcomes. Strand's findings highlight this, and in the Welsh context, intersectional disadvantage may be spatially concentrated due to the geography of migration, labour markets, and

community structures. Recognising these intersecting dimensions strengthens the case for a spatial approach that can capture where such compounded inequalities are most acute.

Besides SES, the role of parental education has also been analysed with its positive influence being well-established across the literature (Davis-Kean 2005; Imeraj et al. 2024; Ludeke et al. 2021). In Northern Ireland, Early et al. (2023) found that a multidimensional approach to SES, including parental education and eFSM, significantly impacted GCSE attainment. However, these patterns are not simply about 'SES' in a narrow statistical sense: they are manifestations of broader class-based processes. Parental education, occupational status, and household resources function as key dimensions of class position, shaping children's exposure to cultural capital, expectations, and institutional know-how. Indeed, the research shows that the effects of parental education and household resources are mediated by institutional environments, shaping how socio-economic disadvantage translates into educational outcomes rather than operating as isolated characteristics(Keating et al. 2025). Furthermore, Erola et al. (2016) demonstrate that educational transmission are more complex than previously thought, involving both direct and indirect pathways, reflecting the reproduction of class advantage rather than isolated socio-economic traits This is consistent with theories of socio-cultural capital, which posits that families with higher educational backgrounds not only provide academic support but also foster a rich socio-cultural environment (Davis-Kean et al. 2021; Harding et al. 2015).

Socio-economic disadvantage is never abstract or placeless because it is produced, reproduced, and experienced through specific spatial configurations rather than operating as an individual attribute. Classic spatial theory stresses that socio-economic processes are always territorially embedded: they arise from, and are sustained by, the uneven geography of labour markets, housing systems, institutional infrastructures, and settlement patterns (Galster 2012; Massey 1994; Pacione 1997). This means disadvantage accumulates differently across places, shaped by local histories, path-dependent economic restructuring, and the spatial clustering of resources and constraints. Empirically, studies of educational inequality repeatedly show that socio-economic conditions exert stronger or weaker effects depending on local opportunity structures and neighbourhood contexts, reinforcing the idea that disadvantage cannot be detached from its spatial setting (Ballas et al. 2012). Therefore, the same nominal indicator of deprivation (eFSM, social grade, household deprivation) can operate differently depending on the spatial configuration in which it is rooted in, reflecting historical legacies linked to socio-economic opportunities or constraints. A spatial lens is therefore key, not because it adds geography as an extra variable, but because it captures the territorially differentiated ways in which classed inequalities materialise. Demographic shifts, labout market shifts and selective migration reshape the socio-demographic composition of urban and rural areas, and in turn condition school resources, peer effects and neighbourhood infrastructures. Spatial variation is therefore not noise, but a constitutive element of how socio-economic disadvantage is reproduced.

Overall, the consensus in the literature on educational inequalities is that they are perpetuated or exacerbated by a multitude of factors. Spatial variation also plays a role, although the primary focus remains on urban-rural divides (Kimosop et al. 2015; Midouhas and Flouri 2015; Sajjad et al. 2022). Nevertheless, the growing use of spatial analysis in educational research has provided new insights into the geography of educational attainment, allowing for more context-specific analyses that move beyond simple urban-rural dichotomies. Lubienski and Lee (2017) argue for the value of geospatial analysis in understanding and uncovering educational inequalities patterns that might be obscured in traditional, non-spatial analyses. To this end, GWR has become an increasingly popular geospatial tool for exploring spatial variations in educational outcomes, as it can offer insights that traditional regression analysis often overlooks. Studies have used GWR to analyse factors influencing academic achievement at the district, school, and individual levels, with common predictors including SES, race/ethnicity, and school resources across different grade levels and subjects. Studies across diverse geographical contexts, from Asia (He and Huang 2021) to Africa (Ansong et al. 2015; Naidoo et al. 2014) or Europe (Kopecna et al. 2021; Sacco and Falzetti 2021) and the United States (Chew et al. 2020; Fiduccia 2017; Qiu and Wu 2019; Thorne Wallington 2014; Thorne-Wallington 2016; Wei et al. 2018), have demonstrated GWR′s capacity to reveal how relationships between SES factors and educational outcomes vary across space. Spatial approaches reveal patterns that remain hidden in global models, particularly where broad classifications such as the urban–rural divide mask substantial local differences in access to resources and opportunity structures. By estimating relationships locally, GWR highlights that socio-economic influences on attainment are context-dependent rather than uniform across geographical areas. This is especially relevant in settings like Wales, where the urban–rural distinction captures some broad contrasts but also obscures important within-area variation linked to different local histories and settlement patterns. Recognising these place-specific contexts provides the theoretical rationale for employing GWR in this study. However, the application of GWR in educational research is not without limitations. Fotheringham et al. (2001), in their study of primary school performance in northern England, noted that GWR results can be sensitive to the choice of spatial scale. This highlights the need for careful consideration of methodological choices when applying GWR to educational data. Different geographical scales capture different social realities: school catchments, neighbourhood boundaries, and small-area socio-economic contexts rarely align, and each contains distinct patterns of interaction, resource access, and opportunity structures relevant to pupils' educational experiences. Relying solely on administrative units such as school districts or LSOAs may therefore obscure fine-grained variations that operate at the level of streets, estates, or community networks These methodological challenges underscore the importance of combining GWR with other analytical approaches to provide a more comprehensive understanding of educational inequalities.

Finally, while educational inequalities have been a major subject of interest in the literature, gaps remain, particularly regarding the use of geospatial analysis to better understand spatial patterns in educational performance. Despite growing acknowledgment of spatial variations in educational outcomes,

as is the case for Wales, there is a notable lack of studies examining these dynamics, especially given its unique socio-economic context and educational landscape. Furthermore, existing research often treats individual and household socio-economic characteristics separately, failing to fully capture the complex relationship between them. More studies are needed that consider individual pupil attributes, household dynamics, and broader socio-economic contexts within a spatial framework, as such an integrated approach could provide key insights into how these multi-level factors together shape educational outcomes across different geographical contexts. Indeed, spatial analysis, particularly GWR, allows for the identification of geographically uneven effects that remain hidden in global models. It shows where the coupling between deprivation and attainment intensifies or weakens, rather than assuming parameter homogeneity. This is key in a context like Wales, where the socio-historical structuring of space and place remains central to the production of educational inequalities. Therefore, this integrated perspective allows us to identify geographically uneven effects that remain hidden in global models, showing how population-level inequalities are structured through place-specific socio-economic configurations rather than operating as uniform individual attributes.

3 | Data and Methodology

This study employs a bi-scalar approach to investigate the socio-spatial determinants of educational attainment in Wales for one cohort (pupils sitting their General Certificate of Secondary Education (GCSE) exams in 2011), employing a comprehensive dataset that integrates administrative education data and census data.

As such, the analysis draws upon three primary data sources, which provide different information:

- **Administrative Examination Data**: The dependent variable is derived from GCSE results for the 2011 cohort and has a binary outcome of achieving the core subject indicator (CSI), or not achieving the CSI. CSI indicates that a pupil achieved GCSE grade A*-C in Welsh or English first language, mathematics and science. This standardised measure of academic achievement serves as our primary indicator of educational attainment. We adopt this binary outcome for clearer interpretation of our econometric models and to maintain policy relevance, following its established use in Wales and UK attainment studies (ap et al. 2017; Tseliou et al. 2024)
- **Administrative Pupil Data**: Individual-level variables include eFSM, a widely-used proxy for SES in educational research (Hobbs and Vignoles 2007; Ilie et al. 2017; Taylor 2018), as well as gender, SEN status, and ethnicity, operationalized as the percentage of pupils from ethnic and other minorities backgrounds. While eFSM is widely employed, its limitations as an indicator of disadvantage have been noted, particularly the risk of misclassification highlighted in debates on contextual indicators (Boliver et al. 2020, 2022). Here, eFSM is therefore treated as a useful but imperfect marker of socio-economic disadvantage.
- **Census data**: Census-based household deprivation indicators (housing, health, employment, education and composite) provide a multidimensional view of disadvantage at the household level. Household-level socio-economic indicators are extracted from the 2011 Census, encompassing family type, highest social grade (a standardised socio-economic classification ranging from AB (higher managerial/professional) to DE (semi-skilled/unskilled), serving as a proxy for family income), highest qualification, number of adults, and number of adults employed. These variables provide important context for understanding the influence of household characteristics on educational outcomes, being often identified as pertinent and important in the literature (Burger 2019; Davis-Kean 2005; Early et al. 2023; Pacione 1997; Schulz 2006).

Overall, the final linked dataset consisted of data on 31,295 pupils, and 1,625 LSOAs. Two-hundred and eighty-three, or 15%, of LSOAs were not considered for the analysis due to low sample sizes that increased the risk of identifying any individuals, as per Higher Education Statistics Agency (HESA) and Office for National Statistics (ONS) guidelines which govern the use of data in SAIL for the purposes of this study. Additionally, as the spatial analysis is undertaken at the LSOA level using aggregated pupil–household data, school-level characteristics (e.g., language medium, school type) which are recorded at a different granularity were not included into these small-area measures. Importantly, while this study draws upon census data, our analysis focuses on 31,295 pupils who sat GCSE exams in 2011. However, the exclusion of 283 LSOAs due to small sample sizes may particularly affect representation of rural and sparsely populated areas, which should be considered when interpreting results. This exclusion introduces a mild spatial imbalance, as rural LSOAs—especially in Mid and West Wales—are more likely to fall below disclosure thresholds, potentially reducing the visibility of some rural patterns in the local analyses. Therefore, although this represents the complete observed population rather than a random sample, we employ statistical significance testing within a superpopulation framework, which provides theoretical justification for statistical inference even when analysing complete populations. This approach conceptualises observed data as one realisation of an underlying stochastic process that could generate different outcomes under similar conditions (Little 2003; Valliant et al. 2000). In educational contexts, this perspective is particularly valuable because cohorts represent temporal snapshots of ongoing social processes that operate across different years and settings (Gibbs et al. 2017).

In terms of the methodological approach, this study integrates individual-level analysis, LSOA-level analysis and spatial analysis with the aim of capturing individual effects and local geographical variations in educational attainment. First, logistic regression is employed to model the probability of pupils' achievement of CSI as a function of individual socio-economic characteristics and household characteristics. To account for potential heterogeneity in error variance across observations, robust standard errors were employed. This allows for the estimation of odds ratios for each predictor variable, providing insights into the relative importance of each factor in determining educational outcomes, while ensuring valid statistical inference. The model is specified as:

$$logit(KS4_{CSI}) = \alpha_0 + \alpha_1(Gender) + \alpha_2(Rural - Urban\ Flag) + \alpha_3(Highest\ Social\ Grade) + \alpha_4(Adults\ working) + \alpha_5(Highest\ Qualificaton) + \alpha_6(eFSM) + \alpha_7(SEN) + \alpha_8(Ethnicity) + \alpha_9(Housing\ Deprivation) + \alpha_{10}(Health\ Deprivation) + \alpha_{11}(Employment\ Deprivation) + \alpha_{12}(Education\ Deprivation) + \alpha_{13}(Family\ Type)$$

Where:

$logit(KS4_{CSI})$ - the log odds of passing the CSI at GCSE;

α_0 - the intercept;

α_1 to α_{13} - the regression coefficients for each predictor variable;

Second, to get insights about the relationships between predictor variables and educational outcomes at the LSOA level, two complementary regression methods were used. The Ordinary Least Squares (OLS) regression served as a global model, providing an overall assessment of the relationships across Wales, with robust standard errors employed to account for non-constant error variance. The model is as follows:

$$Y = \beta_0 + \beta_1(Gender) + \beta_2(Rural - Urban\ Flag) + \beta_3(Highest\ Social\ Grade) + \beta_4(Adults\ working) + \beta_5(Highest\ Qualificaton) + \beta_6(eFSM) + \beta_1(SEN) + \beta_7(Ethnicity) + \beta_8(Household\ Multiple\ Deprivation) + \beta_9(Family\ Type) + \varepsilon$$

Where:

Y - the proportion of pupils passing the CSI at GCSE within each LSOA

β_0 - the intercept

β_1 to β_9 - the regression coefficients for each predictor variable

ε - the error term

Prior to analysis, diagnostic tests were performed to assess model assumptions. The Breusch-Pagan test revealed heteroskedasticity and therefore to address this while maintaining coefficient estimates, robust standard errors were calculated. This approach provides more reliable inference by adjusting standard error estimates to account for heteroskedasticity, without altering the coefficient values themselves.

Finally, to account for spatial heterogeneity in educational outcomes and their determinants, Geographically Weighted Regression (GWR) was used. Prior to implementing GWR, we tested for spatial autocorrelation using Global Moran's I to determine whether educational outcomes exhibited significant spatial clustering. This preliminary analysis informed our decision to employ spatial regression techniques. This method extends the global OLS model by allowing coefficients to vary spatially, allowing for the exploration of local variations in relationships between predictors and educational outcomes across LSOAs in Wales. The model was optimised and only included five key independent variables: Highest Household Qualification, eFSM, SEN, Household Deprivation Index, and Family Type. This was necessary in order to address multicollinearity and spatial non-stationarity issues (Wheeler and Tiefelsdorf 2005), due to strong correlations (over 0.7) between socio-economic indicators, particularly between various household deprivation dimensions (housing, health, employment, and education) and other measures of disadvantage like social grade and employment status. The variable selection process was further guided by statistical diagnostics from ArcGIS Pro, which revealed model instability with the full set of variables. To address this, we employed a robust bandwidth optimisation technique using the Golden Search selection method. This approach systematically determines the optimal spatial weighting configuration by minimising the Akaike Information Criterion (AIC), balancing model complexity with goodness of fit. Given the dataset has1,625 Lower Layer Super Output Areas (LSOAs), the neighbourhood selection process explored a range from 81 to 812 neighbours. The minimum threshold of 81 neighbours (5% of total LSOAs) provides sufficient local data points, while the maximum of 812 neighbours (half of total LSOAs) prevents over-smoothing of spatial relationships. These thresholds reflect a conceptual balance between capturing meaningful local variation in educational contexts and avoiding bandwidths so narrow or so broad that they obscure the spatial structure of socio-economic disadvantage across Wales. We employed the golden search algorithm in ArcGIS Pro to automatically identify the optimal adaptive bandwidth within these boundaries. This method systematically narrows the search range using the golden ratio to find the bandwidth that minimises the AICc criterion, efficiently balancing local specificity with statistical reliability. This approach aligns with best practice in spatial econometrics, where model parsimony is balanced against explanatory power to ensure robust results (Fotheringham et al. 2002; Li et al. 2019). The reduced model mitigates the risk of overfitting and improves the reliability of local parameter estimates, which is important in the context of geographically weighted regression analyses.

$$Z(X_i, Y_i) = \gamma_0(X_i, Y_i) + \gamma_1(X_i, Y_i)(Highest\ Qualificaton) + \gamma_2(X_i, Y_i)(eFSM) + \gamma_3(X_i, Y_i)(Household\ Multiple\ Deprivation) + \gamma_4(X_i, Y_i)(Family\ Type) + \gamma_5(X_i, Y_i)(SEN) + \varepsilon_i'$$

Where:

$Z(X_i, Y_i)$ - the proportion of students passing the CSI at GCSE at location i

(X_i, Y_i) - the projected x-y coordinates of each LSOA

$\gamma_1(X_i, Y_i)$ *to* $\gamma_5(X_i, Y_i)$ - the varying regression coefficients for each predictor variable

at location i

ε'_i - the error term at location i

Prior to analysis, we conducted several diagnostic tests to ensure model validity. Variance Inflation Factor (VIF) analysis was performed to assess multicollinearity among predictor variables. All VIF values were below 5, indicating acceptable levels of collinearity. Moreover, to ensure the validity and reliability of the GWR model, diagnostic tests were conducted to assess multicollinearity and observation influence. Local collinearity, defined as the correlation between predictor variables within each local regression rather than across the entire dataset, was assessed using condition numbers, which remained below 30, suggesting stable local parameter estimates(Wheeler and Tiefelsdorf 2005). Additionally, Cook's Distance was examined to identify potential outliers affecting local estimates and the values were consistently very low across all observations (maximum Cook's distance was less than 0.01), reflecting that no single data point had a strong influence on the estimated coefficients. The absence of influential points reinforces the stability and reliability of the local parameter estimates, ensuring that the regression model provides a consistent and accurate representation of the relationships under investigation (Brunsdon et al. 2012).

The *p*-values in this study are interpreted within a superpopulation framework(Little 2003), which provides theoretical justification for statistical inference even when analysing complete populations. While we analyse complete data from the 2011 pupil cohort, significance testing remains appropriate as it helps assess whether observed relationships reflect stable structural patterns. As Little (2003) argues, this framework allows researchers to make inferences about generative mechanisms that extend beyond the specific dataset at hand.

Finally, to visualise the variation in results across LSOAs, ArcGIS Pro was used to create a series of maps showing magnitude, direction, and statistical significance of the final parameter estimates across Wales. Additionally, two maps displaying the R-squared values and regression residuals were generated to assess both the model's explanatory power and sufficiency.

4 | Results

The characteristics of our sample are shown in Table 1. Half of the pupils achieved their CSI at GCSE level. The data reveals important patterns of disadvantage that extend beyond traditional economic indicators. While 15% of pupils were eFSM, higher proportions experienced various forms of deprivation: housing (18%), employment (27%), education (17%), and particularly health deprivation (37%). This highlights the multidimensional nature of disadvantage affecting educational outcomes. The high standard deviations across these deprivation measures (approximately 7.98–7.99) indicate considerable variation in socioeconomic conditions across the LSOAs, suggesting the importance of examining spatial patterns in educational inequality.

4.1 | Factors Associated With Pupils' Attainment—Individual Level Data

The logistic regression model examines the factors associated with the likelihood of achieving the GCSE CSI using individual level data. This comprehensive model includes control variables, household socio-economic factors, individual pupil characteristics, household deprivation measures, and family structure variables (see Table 2).

The presence of a degree-level qualification in a household was associated with higher odds of achieving the CSI by 140%. Females had 14% higher odds of achieving the CSI than males. SEN showed the strongest negative effect, reducing the odds by 85%, whilst eFSM was also strongly negatively associated with achieving the CSI, with odds being 41% lower. Households with the lowest social grades (D or E) were associated with 32% lower odds of achieving the CSI, Pupils residing in urban areas were associated with 15% lower odds of achieving the CSI compared to their rural counterparts. The percentage of working adults in a household showed a slight positive correlation

TABLE 1 | Sample characteristics.[a]

Variable	Percentage	St. Dev	*N*
CSI Pass	50%	0.5	31, 295
Gender: Female	49%	0.5	31, 295
Gender: Male	51%	0.5	31, 295
Rural-Urban Flag (urban)	71%	0.46	31, 295
Highest social grade in the household (DE)	4%	0.19	31, 295
Adults working in the Household	74%	36.18	31, 295
Highest qualification in the household (degree)	36%	0.48	31, 295
eFSM (eligible)	15%	0.35	31, 295
SEN	19%	0.39	31, 295
Ethnicity (minorities and other ethnicities, beside WBRI)	8%	0.28	31, 295
Housing deprivation	18%	7.98	31, 295
Health deprivation	37%	7.99	31, 295
Employment deprivation	27%	7.99	31, 295
Education deprivation	17%	7.98	31, 295
Family type (married couples)	68%	0.47	31, 295

[a]The characteristics in Table 1 closely reflect those of the full 2011 GCSE cohort in Wales. Minor differences arise due to the exclusion of LSOAs with small pupil counts for disclosure control, which predominantly affects sparsely populated rural areas. No major demographic groups are systematically excluded.

TABLE 2 | Logistic regression results—individual level Key Stage 4 Core Subject Indicator Achievement.

Variables	CSI Pass (Cofficient, *p*-value and standard error)	CSI Pass (Odds ratio)
Gender (female)	0.132*** (0.0257)	1.141***
Rural-urban flag (urban)	−0.155*** (0.0282)	0.856***
Highest Social Grade in the household (DE)	−0.376*** (0.0907)	0.687***
Adults working in the household (%)	0.00238*** (0.000551)	1.002***
Highest qualification in the household (degree)	0.871*** (0.0280)	2.389***
eFSM (eligible)	−0.519*** (0.0488)	0.595***
SEN (additional needs)	−1.863*** (0.0394)	0.155***
Ethnicity (minorities and other ethnicities, beside WBRI)	0.0350 (0.0472)	1.036
Housing deprivation	−0.525*** (0.0412)	0.592***
Health deprivation	−0.0771** (0.0310)	0.926**
Employment deprivation	−0.293*** (0.0427)	0.746***
Education deprivation	−0.556*** (0.0464)	0.573***
Family type (married couples)	0.339*** (0.0288)	1.404***
Constant	−0.0858 (0.0619)	
Observations	31,295	
Pseudo R-squared	0.171	

Note: Standard errors in parentheses; *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

(0.2%), while ethnicity demonstrated no statistically significant relationship.

All four forms of household deprivation as measured by the Census dataset were negatively associated with pupils' achieving the CSI. Housing deprivation was associated with 41% lower odds of achieving the CSI, education deprivation with43% lower odds, health deprivation with 8% lower odds, and employment deprivation with 26% lower odds.

Finally, being from a married couple family was positively associated with achieving the CSI, with odds being 37% higher.

4.2 | Factors Associated With Pupils' Attainment—LSOA-Level Aggregated Data

Table 3 presents the OLS regression results using LSOA-level aggregated data. The results indicate that having a higher proportion of females within an LSOA corresponds to higher achievement rates (e.g., having 10% more females is associated with a 0.72% increase in achievement rates). Urban LSOAs, on average, show a 1.56% lower proportion of pupils reaching the GCSE CSI achievement threshold compared to their rural counterparts. A 10% higher level of households with degrees is associated with a 3.86% higher CSI achievement rate. Conversely, economic disadvantage, as measured by eFSM, has a significant negative effect: 10% higher eFSM corresponds to CSI achievement rates being 2.96% lower. A 10% higher percentage of pupils with SEN results is related to achievement rates being 2.93% lower. Ethnicity also plays a role, with a 10% higher proportion of ethnic minorities or other ethnicities (besides White British) being associated with achievement rates being 0.81% lower. Household multiple deprivation, measured by households experiencing one or more dimensions of deprivation, is negatively linked to educational attainment, with 10% higher deprivation corresponding to achievement rates being 1.94% lower. LSOAs with higher proportions of married couples show higher achievement rates, with a 10% higher proportion of married couple families in an LSOA linked to 0.83% higher CSI achievement rates.

TABLE 3 | OLS regression results—LSOA-level aggregated data.

Variables	CSI Pass (GCSE)
Gender (female)	0.0716*** (0.0268)
Rural-urban flag (urban)	−0.0156** (0.00731)
Highest Social Grade in the household (DE)	−0.0615 (0.0708)
Highest Qualification in the household (degree)	0.386*** (0.0234)
eFSM (eligible)	−0.296*** (0.0392)
SEN (additional needs)	−0.293*** (0.0291)
Ethnicity (minorities and other ethnicities, besides WBRI)	−0.0807** (0.0332)
Household deprivation dimensions (1 or more dimensions)	−0.194*** (0.0273)
Family type (married couples)	0.0833*** (0.0282)
Constant	0.552*** (0.0455)
Observations	1,625
R-squared	0.559

Note: Standard errors in parentheses; *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

To explore the spatial distribution of educational outcomes, we first mapped CSI pass rates across Welsh LSOAs and conducted spatial autocorrelation analysis – the extent to which results are clustered geographically (Figure 1). The significant Global

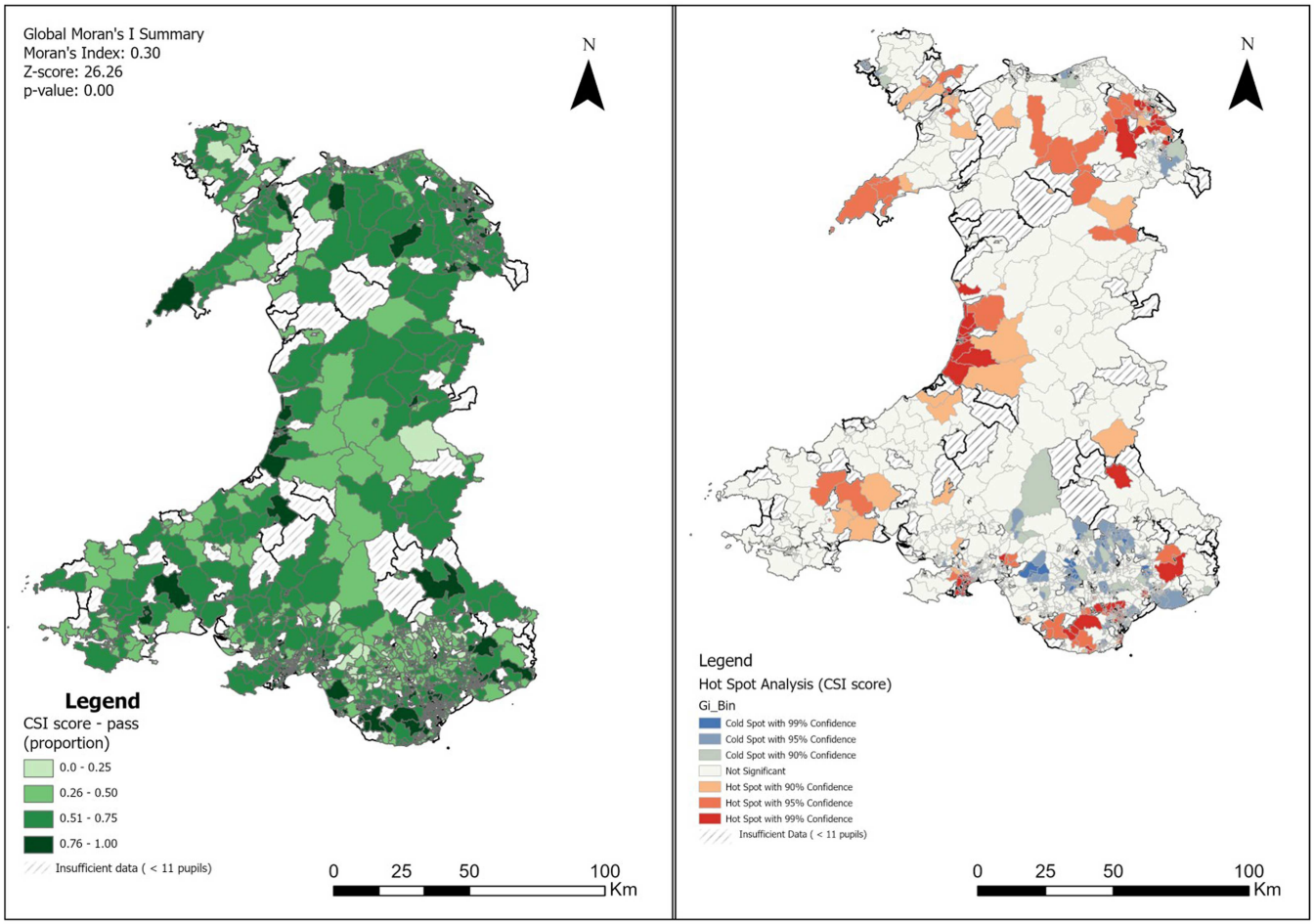

FIGURE 1 | Spatial variations of GCSE results across LSOAs in Wales (left); Hot-Spot Analysis Results for GCSE results (right).

Moran's I value (0.30, $p < 0.001$) confirmed spatial clustering of educational outcomes. Figure 1 presents the spatial distribution of GCSE attainment across Wales. The choropleth map (on the left) displays the proportion of pupils within LSOAs reaching the achievement threshold for CSI for GCSE, while the 'hot-spot' map (on the right) shows the results of the hot-spot analysis for CSI achievement rates within LSOAs.

The choropleth map reveals distinct spatial patterns in GCSE attainment. Lower proportions of pupils achieving the pass threshold for CSI are predominantly observed in the South-East and mid-Wales (indicated by the lighter green shading). In contrast, higher proportions are more frequently found in central and northern areas of Wales (darker green shading), although there are a few LSOAs with very high pass rates in South-East Wales too.

The hot-spot analysis further confirms these spatial patterns. Statistically significant clusters of low proportions (cold spots shown in blue) and high proportions (hot spots shown in red) are evident across the country. A notable cold spot cluster in South-East Wales aligns with the low proportion observed in that region on the choropleth map. North Wales exhibits several hot spots, while North-East Wales presents a mix of both cold- and hot spots.

This moderate positive spatial autocorrelation (Moran's $I = 0.30$) indicates that similar GCSE attainment levels tend to cluster geographically at the LSOA level. The statistically significant z-score of 26.26 ($p < 0.001$) allows us to reject the null hypothesis of random spatial distribution, confirming that the geographic clustering of educational outcomes is unlikely to have occurred by chance.

Following the spatial autocorrelation analysis that confirmed geographic clustering of educational outcomes, we examined how socioeconomic factors relate to these patterns across Wales. Table 4 presents the local GWR results, where we account for geographical variation at the LSOA level. Compared to the global results from OLS where we do not account for variation across LSOAs, we find support for employing GWR methods. The GWR model ($R^2 = 0.57$) slightly outperforms the OLS aggregate model ($R^2 = 0.56$). Moreover, the Akaike Information Criterion (AIC) values provide stronger evidence of an improved model fit, favouring the GWR approach. This is important in a spatial context, as AIC penalises model complexity and therefore provides a more appropriate criterion than R^2 for comparing global and local models with different degrees of freedom. Comparing coefficients between the two approaches further reveals that it is important to consider spatial variation in how factors are associated with educational outcomes. Accounting for spatial variation significantly improves model performance beyond what is captured by R^2 alone. The GWR model reveals important spatial variation in how factors are associated with educational outcomes. Households where at least one adult has a degree-level qualification shows consistently positive effects across Wales,

TABLE 4 | Local(GWR) vs Global Regression (OLS) results.

Variable	Local results (GWR)					Global results (OLS)
	Mean of βs	SD of βs	Min	Max	Median	β (SE)
Intercept	0.48	0.04	0.36	0.66	0.48	0.50*** (0.03)
Highest qualification in the household (degree)	0.38	0.06	0.23	0.52	0.36	0.39*** (0.02)
Household deprivation dimensions (1 or more dimensions)	−0.16	0.09	−0.35	0.05	−0.16	−0.20*** (0.03)
eFSM (eligible)	−0.28	0.13	−0.55	0.01	−0.27	−0.30*** (0.04)
SEN (additional needs)	−0.36	0.12	−0.67	−0.16	−0.33	−0.29*** (0.03)
Family type (married couples)	0.1	0.07	−0.09	0.27	0.11	0.09*** (0.03)
R-squared			0.57			0.55
AIC			−3255.6461			−1971.2368

Note: Standard errors in parentheses; $^{**}p < 0.05$; $^{*}p < 0.1$; $^{***}p < 0.01$.

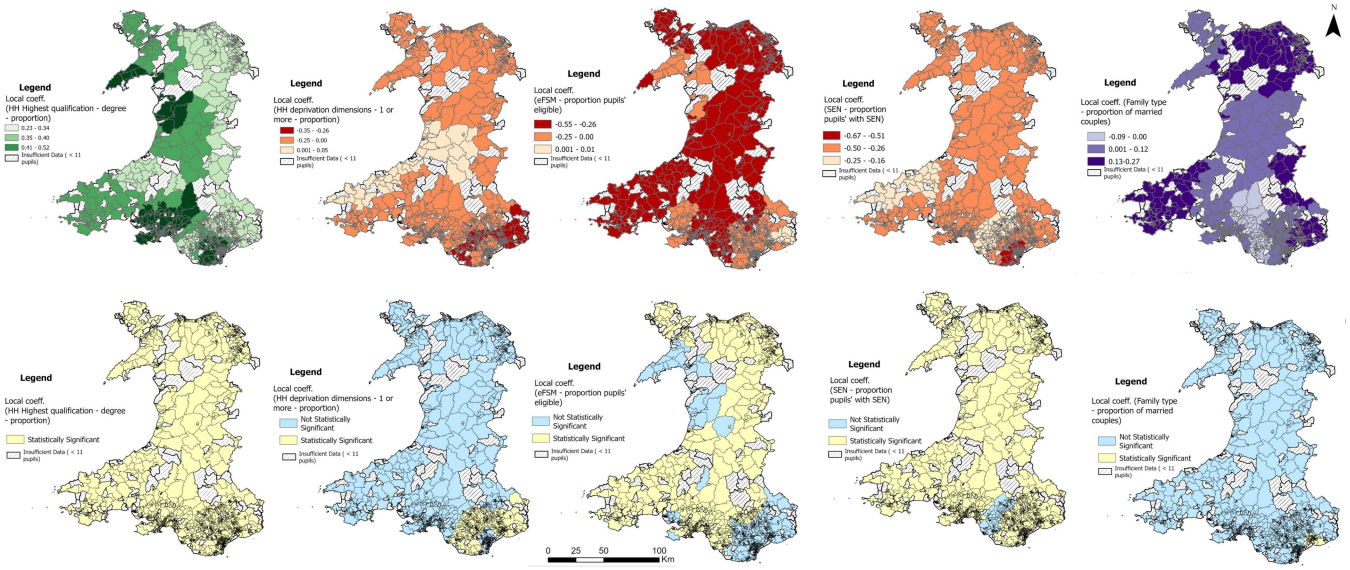

FIGURE 2 | GWR results - local coefficients & statistical significance.

with GWR coefficients ranging from 0.23 to 0.52, compared to the OLS estimated of 0.39. Household multidimensional deprivation exhibits more variable effects (GWR: −0.35 to 0.05; OLS: −0.20), with some areas showing no significant associations. Similarly, eFSM demonstrates substantial spatial variation (GWR: −0.55 to 0.01; OLS: −0.30), with the unexpected positive coefficients likely reflecting, as in the case of the household multidimensional deprivation, local data noise in areas with fewer observations, as discussed earlier. SEN status shows the widest range of local effects (GWR: −0.67 to −0.06; OLS: −0.29), indicating a consistently negative association across Wales, but with clear spatial differentiation. The strongest effects appear across the post-industrial LSOAs of the South Wales Valleys. In contrast, the weakest effects are found in parts of the south-West coastal peninsula, while most other regions of Wales exhibit moderate, mid-range effects. This variation matters because it highlights that the extent to which SEN status impacts attainment is shaped by local educational provision and support capacity, with some areas better able to buffer its effects than others. Finally, in relation to family type, the model shows a positive but variable association (GWR: −0.09 to 0.27; OLS: 0.09), suggesting that while coming from a married couple family is generally associated with better educational outcomes, this effect varies considerably across different areas of Wales.

To visualise these spatial variations identified in our GWR model, Figure 2 maps the local coefficient values and their statistical significance across Wales. The top five maps indicate the spatial variation for different variables of interest, whilst the bottom five maps show corresponding significance across LSOAs for each of the variables of interest.

First, the association between presence of household members with degrees and GCSE attainment varies substantially across Wales. A 10% higher proportion of households with degrees is associated with the proportion of pupils achieving CSI being 2.3% to 5.2% higher, depending on the LSOA. This relationship is statistically significant throughout Wales, underscoring the consistent importance of higher education in household environments.

Second, household deprivation, which is often negatively associated with attainment, varied in its' association with attainment across Wales. A 10% higher level of household deprivation is associated with effects ranging from being 3.5% lower to 0.5%

higher GCSE attainment, depending on the LSOA. Whilst, this relationship is not statistically significant in most areas, the strongest negative effects are concentrated in parts of South-East Wales, particularly across the post-industrial LSOAs of the South Wales Valleys (including Merthyr Tydfil and Blaenau Gwent) and the Cardiff–Newport urban fringe. This matters because these areas combine long-standing labour-market restructuring, higher concentrations of multidimensional disadvantage, and more constrained local resources - place-specific conditions that intensify the influence of household deprivation on educational outcomes.

Thirdly, the association between eFSM and GCSE attainment varies widely across Wales, with a 10% lower proportion of eFSM in an LSOA being associated with differences ranging from 5.5% lower to 0.1% higher GCSE attainment across different LSOAs. This relationship is statistically significant in most areas, except in South-East Wales and some LSOAs in West Wales, where positive local coefficients are observed. The variation in findings between household deprivation and FSM indicate that multiple and non-economic deprivation have differing impacts, particularly in South-East Wales.

For the association between SEN and GCSE attainment, a 10% higher proportion of pupils with SEN is associated with 6.7% to 1.6% lower achievement of CSI, depending on the LSOA. This relationship is statistically significant in most areas, with some exceptions in South-East Wales, around Maesteg and Merthyr.

Finally, the association between family structure and achievement of CSI shows that a 10% lower proportion of households with married couples is associated with differences ranging from 0.9% lower to 2.7% higher GCSE attainment, depending on the LSOA. However, this relationship is not statistically significant in most LSOAs.

A small number of counterintuitive coefficients emerge in the GWR outputs for both household deprivation and eFSM. For household deprivation, these take the form of a band running through Mid-Wales and extending into the central West Wales coastline, covering LSOAs from northern Pembrokeshire through coastal Ceredigion and into south-central Gwynedd, with additional neighbouring LSOAs in central Powys. These areas display small positive coefficients (0.001–0.05). For eFSM, only a very limited number of LSOAs in the South-East exhibit similarly weak positive coefficients, and these do not form any wider spatial grouping. None of these coefficients are statistically significant. Conditions. These may be attributed to local data noise in the GWR model. In areas with fewer individuals, the model becomes more susceptible to random fluctuations, potentially leading to unstable coefficient estimates (Kiani et al. 2024; Leung et al. 2000; Yu et al. 2020)

Finally, to assess where our model most effectively captures the factors driving educational outcomes, Figure 3 maps the local R^2 values and model residuals across the LSOAs, providing information about the model fit. The model demonstrates its

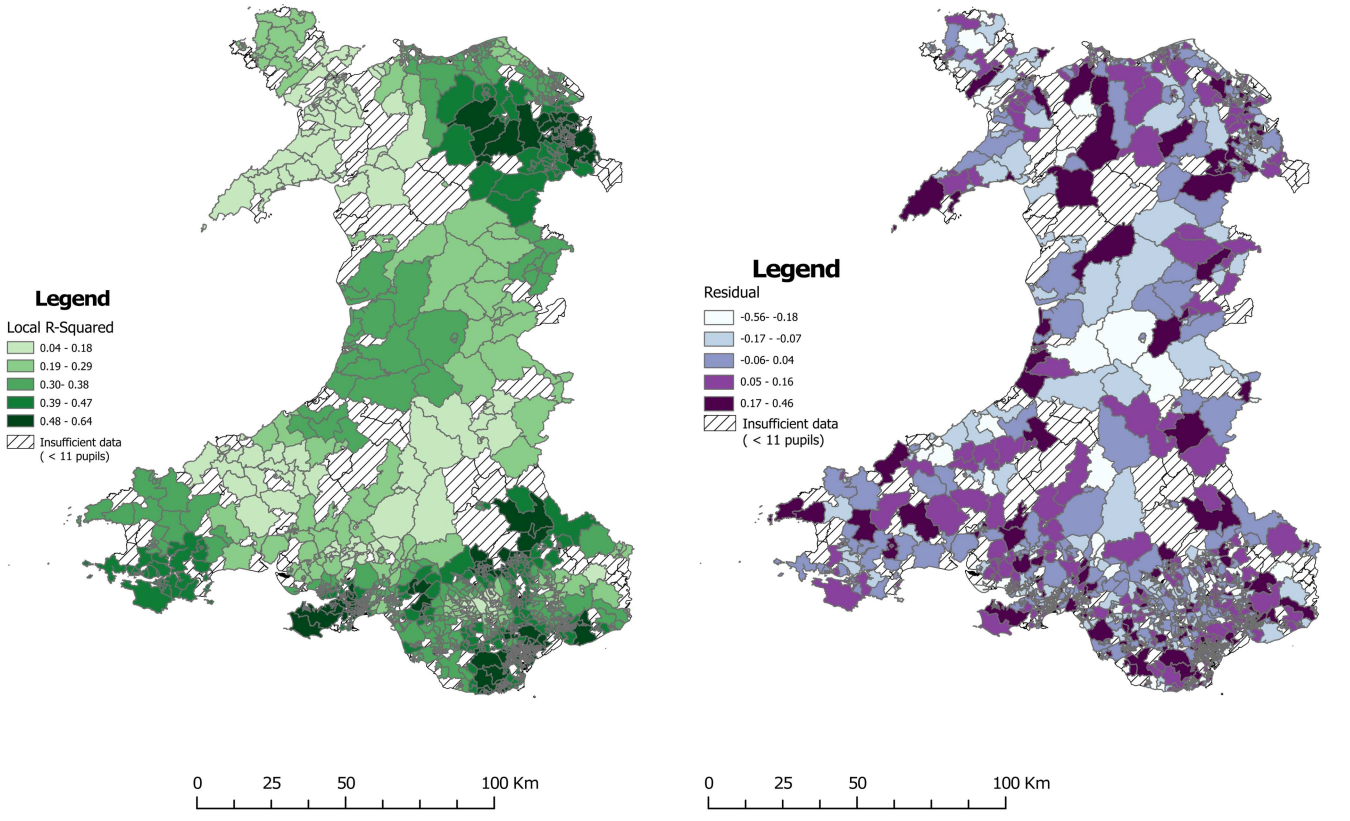

FIGURE 3 | GWR results - Local R2 (left) and Residuals (right).

strongest explanatory power in South-East Wales, where it accounts for up to 75% of the variation in educational outcomes (Local R^2 values ranging from 0.68 to 0.75). A notable trend emerges, showing better model fit in more urbanised areas, particularly in the south of the country, where Local R^2 values consistently exceed 0.60. This pattern suggests that the socio-economic and demographic factors included in our model are particularly relevant in explaining educational attainment in urban contexts. However, the residual map highlights that while the model performs well overall, there are local variations it doesn't fully capture. The spread of residuals across Wales, ranging from −0.41 to 0.44, indicates the presence of other factors associated with educational outcomes that warrant further investigation. Areas with higher residuals (both positive and negative) need further investigation. These could represent LSOAs where local factors are influencing educational outcomes beyond the variables included in the current model, such as school quality and resources, teacher experience and retention rates, access to additional educational support services or community-level social capital.

5 | Discussion

Overall, Wales is marked by persistent territorial inequalities that influence how deprivation translates into educational outcomes. South-East Wales, especially the former coalfield areas of the Valleys, combines high population densities with long-term labour-market restructuring, limited local employment diversity and concentrated multidimensional deprivation. These dynamics contrast with Mid and West Wales, where low-density rural areas face ageing demographics, limited public transport connectivity and sectoral dependence on agriculture and tourism, producing different opportunity structures despite lower deprivation intensity. Parts of North-East Wales form a cross-border economic corridor with North-West England, characterised by stronger labour-market links, higher out-commuting rates and more mixed socio-economic profiles. These territorially specific configurations - industrial legacies, labour-market reach, accessibility and demographic structure - provide the contextual conditions within which the spatial variation identified in our analysis emerges(Beatty and Fothergill 2018; Henley 2024).

Indeed, the findings of this study highlight that geography plays a central role in shaping the variability of factors associated with educational attainment across Wales. The GWR analysis revealed that relationships between key socio-economic indicators and GCSE attainment are not uniform, but rather exhibit considerable local variation, except for parental education. This spatial heterogeneity is particularly relevant in a geographically diverse context such as Wales, where socio-economic conditions, settlement patterns and educational challenges differ both between and within regions (Davies et al. 2011; Equality and Social Justice Committee 2023). Recognising this variation is essential, as it demonstrates that socio-economic disadvantage is embedded in place-specific contexts rather than operating consistently across space. The results therefore reinforce arguments in the wider literature that educational inequalities cannot be fully understood or effectively addressed without accounting for the spatial contexts in which they occur (Ballas et al. 2012; Frouillou 2022; Mishra 2023).

The mean CSI attainment rate of 0.50 indicates that while half of the pupils achieved the CSI at GCSE the underlying factors contributing to this outcome reveal a more complex narrative where both individual and household socio-economic characteristics, as well as geography converge and create opportunities but also challenges. Moreover, the observed spatial clustering of educational outcomes raises key questions about the underlying processes driving these geographical patterns implying that neighbouring areas are more likely to have similar educational outcomes, pointing to the influence of both broader regional/local factors and personal and household characteristics on pupils' educational performance. While our individual-level analysis identifies significant associations between socio-economic factors and educational attainment, the spatial heterogeneity in these relationships as highlighted by the GWR analysis, suggests a complex relationship between place and education than previously accounted for.

We shall now consider the differences in results between our individual, and LSOA models.

5.1 | Gender

The relationship between gender and educational attainment, with girls showing a higher likelihood of achieving CSI at GCSE level compared to boys, is consistent with previous research, suggesting that gender disparities in educational performance persist, often correlated with variations in motivation (Buchmann et al. 2008; Connolly 2006; Strand 2014; Voyer and Voyer 2014). However, one should also keep in mind that the narrative of girls outperforming boys is more complex and context-dependent, varying across school subjects, as several researchers point out (Connolly 2006; Watson et al. 2010).

5.2 | Urban-Rural Divide

An interesting finding from our analysis was the persistent negative association between urban residence and GCSE attainment. Across all models, pupils residing in urban areas showed a lower likelihood of achieving CSI compared to their rural counterparts. This urban disadvantage contrasts with several international studies that have found urban advantages in educational outcomes (Kopecna et al. 2021; Kryst et al. 2015; Rodriguez-Gómez et al. 2024; Welch et al. 2007), suggesting that the relationship between urban residency and educational attainment may be context-specific in Wales, or in the UK. As Graham (2024) also highlighted that in England pupils from rural areas sometimes outperform pupils from urban areas in exams. However, Echazarra and Radinger (2019) show that when SES is accounted for, the performance gap between rural and urban pupils often diminishes and thus the reasons for this urban-rural divide in Wales warrant further investigation. Apart from socio-economic characteristics, potential factors could include differences in school quality, community resources, or specific urban challenges not captured in our models.

5.3 | Socio-Economic Background: Household Social Grade, eFSM, Household Deprivation

It is widely accepted in the literature that household socio-economic factors play an important role in shaping pupils'

academic outcomes, providing insights into disparities in educational attainment. Our findings in the Welsh context align with this perspective, showing that pupils from households classified within the lowest social grades (D or E) experience a worrying decrease in their likelihood of passing GCSE exams. Furthermore, the negative associations between GCSE results and other socio-economic factors such as eFSM and various dimensions of household deprivation, highlight the persistent educational disadvantages faced by pupils from less privileged backgrounds, which in turn hinder their educational success. These results echo findings from numerous studies on the impact of SES on educational outcomes (Broer et al. 2019; Burger 2019; Early et al. 2023; Pacione 1997; Salma and Chaudhry 2024; Schulz 2006; Sirin 2005).

The spatial variations in the effects of eFSM and Household Deprivation on attainment underline the importance of local context in shaping educational outcomes. These differences matter because they reflect underlying place-specific conditions: socio-economic disadvantage is more concentrated and persistent in parts of South-East Wales, where post-industrial legacies and higher levels of multidimensional deprivation intensify its association with attainment. In other areas of Wales, where deprivation is less spatially clustered, these relationships are more moderate. This contrast helps explain why national policies designed around uniform assumptions may have limited effectiveness, as the scale and form of disadvantage differ between places. Efforts to mitigate socio-economic inequalities may therefore require greater intensity and targeting in areas where disadvantage exerts stronger penalties. The observed divergence between eFSM and household deprivation across regions further illustrates that socio-economic influences operate through different local contexts rather than functioning uniformly across the country. This highlights the need for geographically sensitive approaches to both research and policy aimed at reducing educational inequalities in Wales.

5.4 | Household/Parental Educational Attainment

In contrast, the presence of at least one degree-holder within a household correlates with a very high increase in the odds of a pupil passing their GCSE exam. This not only emphasises the impact of parental educational attainment on children′s academic performance but also raises questions about the cycle of inherited educational advantage and disadvantage across generations in Wales, aligning with the findings about intergenerational transmission of educational advantage highlighted by Chevalier et al.(2013) both in the USA and UK. This reinforces the importance of considering family educational level when aiming to understand pupils' educational attainment (Davis-Kean 2005; Desforges 2003; Ule 2015). Moreover, the varying effect of level of education in a household across Wales, while highlighting the benefits of parental education as evidenced by the literature, also raises questions about what local factors might be moderating this relationship and how it translates into children's academic success. Future research could employ mixed-methods approaches, combining quantitative spatial analysis with qualitative case studies to better understand the local dynamics observed, looking into community-level social capital or regional socio-economic structures.

5.5 | Family Structure

Finally, the analysis of the relationship between family structure and educational attainment shows that while the global regression model indicated a statistically significant association across Wales, when local variations were accounted for, family structure emerged as consistently non-significant across small geographies. Therefore, while there might be an overall effect at a Wales-level, potentially reflecting broader socio-economic patterns associated with family structures (Albright and Conley 2004; Chung 2015; Von Stumm et al. 2022), local non-significant results indicates that this effect is not consistently observable at smaller geographical scales. Other local factors may instead be correlated with pupils' educational outcomes. Thus, these findings caution against any assumption that trends observed at a national level translate to local contexts, underscoring the need for multi-scale analysis in understanding the geography of educational attainment. A more nuanced approach should consider how various family structures interact with socio-economic and community factors to influence educational outcomes, rather than focusing solely on family structure. This approach would avoid perpetuating stereotypes or unfairly stigmatising certain types of family.

5.6 | Special Education Needs

At the individual level, pupils with SEN show significantly lower odds of achieving CSI at GCSE level. The spatial analysis further nuances this finding, demonstrating that while the negative association is present across Wales, its magnitude varies substantially across LSOAs. Most notably, some areas, both in South-West and South-East Wales, show non-significant associations, suggesting the presence of potentially effective local strategies. These spatial variations raise questions about differences in local support systems, resource allocation, and educational practices. Areas where SEN has a weaker association with attainment could offer valuable insights into effective support mechanisms, whether these are formal educational interventions, specific school strategies, or community support networks. However, interpreting these patterns requires careful consideration of local context and local strategies/services.

6 | Conclusions

This study aimed to examine the spatial disparities in educational attainment across Wales, focusing on the multidimensional relationship between individual and household characteristics associated with GCSE exam results. By employing both global and local spatial analysis techniques, the study underscores key patterns that provide insights into the geography of educational outcomes, with implications extending beyond the Welsh context. Thus, the novelty of the research lies in its comprehensive integration of individual and household data within a spatial framework. The use of GWR has allowed the recognition of local variations that would have been obscured by traditional global models and this multi-level approach makes it possible to observe patterns of educational disadvantage that go beyond urban-rural or affluent-deprived dichotomies.

Beyond the methodological contribution, the analysis highlights how socio-economic disadvantage is spatially organised and

experienced unevenly across Wales. These findings offer a basis for qualitative and mixed-methods research to examine the local mechanisms underlying the observed spatial variation, particularly in areas where disadvantage appears to exert weaker or stronger effects than expected. The study therefore provides both a diagnostic mapping of inequalities and a platform for deeper place-based enquiry.

Indeed, the analysis revealed that different measures of SES are statistically associated with educational outcomes, with their relationship varying across spatial contexts. SEN, eFSM, and multidimensional household deprivation emerged as key predictors, although their effects differ notably between South-East Wales and other regions, particularly in the case of the latter two. This spatial heterogeneity underscores the complex, but also locally dependent mechanisms linking SES to educational outcomes. Overall, geography appears to play a key role in moderating the relationships between SES, household composition, and educational outcomes. It also suggests that one-size-fits-all educational policies may be less effective in addressing the diverse challenges across Wales. Thus, stakeholders and policymakers should consider developing context-sensitive interventions that account not just for overall educational gaps, but also consider the broader socio-economic conditions, tailored to local contexts. Areas where factors that usually have negative effects (such as multidimensional household deprivation or SEN) have less impact on educational outcomes could provide insights into effective local practices that could be scaled up or adapted for other areas.

This study is not without limitations. It uses cross-sectional data, providing only a snapshot of educational attainment patterns and taking no account of the impact of subsequent policy changes. As the residuals from the GWR analysis suggest, there is potential for omitted variables, since educational outcomes are shaped by a multitude of factors—not only individual and household characteristics but also contextual and place-specific influences. Future research should explore temporal dynamics through longitudinal data and consider a mixed-methods approach to build a more complete picture of the factors shaping pupils' educational outcomes in Wales. Although the household-level data were derived from the 2011 Census, this remains the most recent small-area socio-economic data linked to pupil records. The release of the 2021 Census will allow for an update of this analysis, enabling temporal comparison and potential extension of the methodology to examine change over time. In addition, the spatial patterns and variations observed across Wales point to the need for more in-depth spatial analyses, incorporating additional geographic variables and employing advanced spatial modelling techniques to further understand the complex relationship between place and educational performance.

The patterns identified here reflect the socio-economic landscape captured by the 2011 cohort, but the intervening period has been marked by substantial shocks that may have amplified spatial inequalities. A decade of austerity, the Covid-19 pandemic, and the ongoing cost-of-living crisis have all affected households unevenly across Wales, with early evidence suggesting intensified pressures in already disadvantaged areas. Evidence from the Covid-19 period indicates that unequal access to digital learning resources and home learning environments further reinforced pre-existing socio-spatial inequalities in educational participation and attainment (Sandu and Taylor 2024). These shifts underscore the importance of revisiting the analysis once 2021 Census-linked data become available, not only to assess temporal change but also to examine how new vulnerabilities and support structures have altered the geography of educational inequality.

Nevertheless, despite these limitations, the current study provides a robust foundation to inform the development of more targeted, geographically sensitive policies to address educational inequalities in Wales. Overall, it foregrounds place as a moderator of socio-economic disadvantage, contributing to the emerging international literature on spatialised educational inequality (Lubienski and Lee 2017; Burger 2019). While the Welsh case is nationally specific, the broader lesson—that educational disadvantage is spatially uneven even in small jurisdictions—has relevance across devolved, federal, or regionally diverse education systems in Europe and beyond.

Acknowledgements

This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) system, which is part of the national e-health records research infrastructure for Wales. We would like to acknowledge all the data providers who make anonymised data available for research. The data analysis was carried out in the Wales Institute of Social and Economic Research and Data (WISERD) Education Data Lab project 'Participation in and Progression Through Education in Wales' within the SAIL databank, funded under the Administrative Data Research (ADR) Wales Education theme ES/W012227/1 and WISERD ES/S012435/1. This work uses data provided by patients and collected by the NHS as part of their care and support. ADR Wales brings together data science experts at Swansea University Medical School, staff from the WISERD at Cardiff University and specialist teams within the Welsh Government to develop new evidence which supports the Programme for Government by using the SAIL Databank at Swansea University, to link and analyse anonymised data. The ADRW Programme of Work 2022–2026 outlines the ten thematic areas that the ADR Wales team will focus their research on to help government address the most pressing issues facing society. ADR Wales is part of ADR UK and funded by the Economic and Social Research Council (part of UK Research and Innovation). This work was conducted while Alexandra Sandu and Jen Keating were at Cardiff University - Administrative Research Data Wales, UK.

Ethics Statement

The National Research Ethics Service has previously agreed that research carried out within SAIL does not require ethical review due to the anonymization process applied to the data.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Albright, K., and D. Conley. 2004. "After the Bell. Family Background." In *Public Policy and Educational Success*. Routledge.

Ansong, D., E. K. Ansong, A. O. Ampomah, and B. K. Adjabeng. 2015. "Factors Contributing to Spatial Inequality in Academic Achievement in Ghana: Analysis of District-Level Factors Using Geographically Weighted Regression." *Applied Geography* 62: 136–146.

ap, G., L. Spencer, J. Payne, et al. 2017. *Re-thinking Educational Attainment and Poverty (REAP) - in Rural Wales*. Bangor University.

Ballas, D., R. Lupton, D. Kavroudakis, et al. 2012. *Mind the Gap: Education Inequality across EU Regions*. Brussels: European Commission.

Bandyopadhyay, A., T. Whiffen, R. Fry, et al. 2023. "How Does the Local Area Deprivation Influence Life Chances for Children in Poverty in Wales: A Record Linkage Cohort Study." *SSM - Population Health* 22: 101370.

Banerjee, S., G. M. Szirony, N. McCune, W. S. Davis, S. Subocz, and B. Ragsdale. 2022. "Transforming Social Determinants to Educational Outcomes: Geospatial Considerations." *Healthcare* 10, no. 10: 1974.

Beatty, C., and S. Fothergill. 2018. "The contemporary labour market." In *Britain's Older Industrial Towns|Sheffield Hallam University*. Sheffield Hallam University.

Blanden, J. 2004. "Family Income and Educational Attainment: A Review of Approaches and Evidence for Britain." *Oxford Review of Economic Policy* 20, no. 2: 245–263.

Boliver, V., S. Gorard, M. Powell, and T. Moreira. 2020. "The Use of Access Thresholds to Widen Participation at Scottish Universities." *Scottish Affairs* 29, no. 1: 82–97.

Boliver, V., S. Gorard, and N. Siddiqui. 2022. "Who Counts as Socio-economically Disadvantaged for the Purposes of Widening Access to Higher Education?" *British Journal of Sociology of Education* 43, no. 3: 349–374.

Broer, M., Y. Bai, and F. Fonseca. 2019. "A Review of the Literature on Socioeconomic Status and Educational Achievement." In Socio-economic Inequality and Educational Outcomes. *IEA Research for Education*, 7–17. Cham: Springer International Publishing.

Brunsdon, C., M. Charlton, and P. Harris. 2012. Living with Collinearity in Local Regression Models.

Buchmann, C., T. A. DiPrete, and A. McDaniel. 2008. "Gender Inequalities in Education." *Annual Review of Sociology* 34, no. 1: 319–337.

Burger, K. 2019. "The Socio-Spatial Dimension of Educational Inequality: A Comparative European Analysis." *Studies in Educational Evaluation* 62: 171–186.

Cadwallader, S., G. Lee, and H. Kayton. 2023a. *PISA 2022: National Report for England*. Department for Education.

Cadwallader, S., G. Lee, and H. Kayton. 2023b. *PISA 2022: National Report for Wales*. Welsh Government.

Chevalier, A., C. Harmon, V. O' Sullivan, and I. Walker. 2013. "The Impact of Parental Income and Education on the Schooling of Their Children." *IZA Journal of Labor Economics* 2, no. 1: 8.

Chew, B., A. Satpathy, and E. Wong. 2020. "Geospatial Analyses to Determine Academic Success Factors in California's K-12 Education." *Annals of GIS* 26, no. 2: 81–100.

Chmielewski, A. K. 2019. "The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015." *American Sociological Review* 84, no. 3: 517–544.

Chung, K. K. H. 2015. "Socioeconomic Status and Academic Achievement." In *International Encyclopedia of the Social & Behavioral Sciences*, 924–930. Elsevier.

Connolly, P. 2006. "The Effects of Social Class and Ethnicity on Gender Differences in GCSE Attainment: A Secondary Analysis of the Youth Cohort Study of England and Wales 1997–2001." *British Educational Research Journal* 32, no. 1: 3–21.

Davies, R., S. Drinkwater, C. Joll, et al. 2011. *An Anatomy Of Economic Inequality In Wales*. Cardiff: Commission in Wales - Equality and Human Rights Commission.

Davis-Kean, P. E. 2005. "The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment." *Journal of Family Psychology* 19, no. 2: 294–304.

Davis-Kean, P. E., L. A. Tighe, and N. E. Waters. 2021. "The Role of Parent Educational Attainment in Parenting and Children's Development." *Current Directions in Psychological Science* 30, no. 2: 186–192.

Desforges, P. C. 2003. "The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievement and Adjustment: A Literature Review." *Department for Education and Skills* 1: 1.

Early, E., S. Miller, L. Dunne, and J. Moriarty. 2023. "The Influence of Socio-Demographics and School Factors on GCSE Attainment: Results From the First Record Linkage Data in Northern Ireland." *Oxford Review of Education* 49, no. 2: 171–189.

Echazarra, A., and T. Radinger. 2019. "Learning in rural schools: Insights From PISA, TALIS and the literature." OECD Education Working Papers (196). OECD Education Working Papers.

Equality and Social Justice Committee. 2023. Calling Time on Child Poverty: How Wales Can Do Better. Welsh Parliament - Equality and Social Justice Committee.

Erola, J., S. Jalonen, and H. Lehti. 2016. "Parental Education, Class and Income Over Early Life Course and Children's Achievement." *Research in Social Stratification and Mobility* 44: 33–43.

Fiduccia, P. 2017. School Performance and Community Development in New York State: A Spatial Statistics Perspective.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, England ; Hoboken, NJ, USA: Wiley.

Fotheringham, A. S., M. E. Charlton, and C. Brunsdon. 2001. "Spatial Variations in School Performance: A Local Analysis Using Geographically Weighted Regression." *Geographical and Environmental Modelling* 5, no. 1: 43–66.

Frouillou, L. 2022. "The Spatial Dimension of Educational Inequalities." In *Inequalities in Geographical Space*, 1–25. John Wiley & Sons, Ltd.

Galster, G. C. 2012. "The Mechanism(s) of Neighbourhood Effects: Theory, Evidence, and Policy Implications." In Neighbourhood Effects Research: New Perspectives. *Dordrecht*, edited by M. van Ham, D. Manley, N. Bailey, et al., 23–56. Netherlands: Springer.

Gibbs, B. G., K. Shafer, and A. Miles. 2017. "Inferential Statistics and the Use of Administrative Data in US Educational Research." *International Journal of Research & Method in Education* 40, no. 2: 214–220.

Gorard, S., and N. Siddiqui. 2019. "How Trajectories of Disadvantage Help Explain School Attainment." *Sage Open* 9, no. 1: 2158244018825171.

Graham, L. 2024. "The Grass Ceiling: Hidden Educational Barriers in Rural England." *Education Sciences* 14, no. 2: 165.

Gulson, K. N., and C. Symes. 2007. "Spatial Theories of Education: Policy and Geography Matters." In *Routledge Research in Education* (9). London New York: Routledge, Taylor & Francis Group.

Harding, J. F., P. A. Morris, and D. Hughes. 2015. "The Relationship Between Maternal Education and Children's Academic Outcomes: A Theoretical Framework." *Journal of Marriage and Family* 77, no. 1: 60–76.

He, G., and Q. Huang. 2021. "G." In *eospatial Analysis and Research on Social and Spatial Inequality of Compulsory Education: A Case Study of Hangzhou*, edited by J.-H. Pan, 2021, 1–14. China: Complexity.

Henley, A. 2024. "Welsh Productivity Performance: Lost Cause or Still Waiting for a Miracle?" *Welsh Economic Review* 29: 1–16.

Herbert, D. T., and C. J. Thomas. 1998. "School Performance, League Tables and Social Geography." *Applied Geography* 18, no. 3: 199–223.

Hobbs, G., and A. Vignoles. 2007. *Is Free School Meal Status a Valid Proxy for Socio-Economic Status (in Schools Research)?* 1 June. London School of Economics: Centre for the Economics of Education.

Ilie, S., A. Sutherland, and A. Vignoles. 2017. "Revisiting Free School Meal Eligibility as a Proxy for Pupil Socio-Economic Deprivation." *British Educational Research Journal* 43, no. 2: 253–274.

Imeraj, L., H. A. G. de Valk, and S. Gadeyne. 2024. "Parental Education and Neighbourhood-Effect Heterogeneity in Educational Attainments

of Native and Minority Youth in Belgian Metropolitan Cities." *Population, Space and Place* 30, no. 5: e2756.

Ingram, J., J. Stiff, S. Cadwallader, et al. 2023. *PISA 2022: National Report for Northern Ireland*. Department for Education.

Keating, J., C. Knight, A. Sandu, and R. French. 2025. "What Individual, Family, and School Factors Influence the Identification of Special Educational Needs in Wales?" *British Journal of Educational Psychology* 95, no. 2: 530–550.

Kiani, B., B. Sartorius, C. L. Lau, and R. Bergquist. 2024. "Mastering Geographically Weighted Regression: Key Considerations for Building a Robust Model." *Geospatial Health* 19, no. 1: 1271.

Kimosop, P. K., K. M. Otiso, and X. Ye. 2015. "Spatial and Gender Inequality in the Kenya Certificate of Primary Education Examination Results." *Applied Geography* 62: 44–61.

Kopecna, J., C. Bagnarol, S. Donno, et al. 2021. "Geographical Differences in Italian Students' English Test Per- Formance: A Role of Individual And Local Characteristics." In *invalsi data to Investigate the Characteristics of Students, School, and Society: IV Seminar 'INVALSI Data: A Research and Educational Teaching Tool'*, edited by P. Falzetti, 9–29. Milano, Italy: FrancoAngeli INVALSI - Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione.

Kryst, E. L., S. Kotok, and K. Bodovski. 2015. "Rural/Urban Disparities in Science Achievement in Post-Socialist Countries: The Evolving Influence of Socioeconomic Status." *Global Education Review* 2, no. 4: 4.

Leung, Y., C.-L. Mei, and W.-X. Zhang. 2000. "Testing for Spatial Autocorrelation Among the Residuals of the Geographically Weighted Regression." *Environment and Planning A: Economy and Space* 32, no. 5: 871–890.

Li, U., U. Fotheringham, U. Li, and Oshan. 2019. "Fast Geographically Weighted Regression (FastGWR): A Scalable Algorithm to Investigate Spatial Process Heterogeneity in Millions of Observations." *International Journal of Geographical Information Science* 33, no. 1: 155–175.

Little, R. 2003. To Model or Not to Model? Competing Modes of Inference for Finite Population Sampling." The University of Michigan Department of Biostatistics Working Paper Series. Epub ahead of print 12 November 2003.

Lubienski, C., and J. Lee. 2017. "Geo-Spatial Analyses in Education Research: The Critical Challenge and Methodological Possibilities." *Geographical Research* 55, no. 1: 89–99.

Ludeke, S. G., M. Gensowski, S. Y. Junge, R. M. Kirkpatrick, O. P. John, and S. C. Andersen. 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, no. 4: 1074–1090.

Massey, D. B. 1994. *Space, Place, and Gender*. Minneapolis: University of Minnesota Press.

Midouhas, E., and E. Flouri. 2015. "Rural/Urban Area Differences in the Cognitive Abilities of Primary School Children in England." *Population, Space and Place* 21, no. 2: 157–170.

Mishra, A. 2023. "Spatial Inequality and Education: Unraveling the Geographical Dimensions of Educational Disparities." *Technolearn An International Journal of Educational Technology* 13, no. 1: 1.

Naidoo, A. V., A. van Eeden, and Z. Munch. 2014. "Spatial Variation in School Performance, a Local Analysis of Socio-Economic Factors in Cape Town." *South African Journal of Geomatics* 3, no. 1: 78–94.

Pacione, M. 1997. "The Geography of Educational Disadvantage in Glasgow." *Applied Geography* 17, no. 3: 169–192.

Perales, R. D., S. D. M. Familara, R. T. D. Pena, et al. 2023. "Direct Influence of Socio-Economic Towards Education Quality: Growth of Standardization." *Buletin Edukasi Indonesia* 2, no. 3: 121–127.

Qiu, X., and S. Wu. 2019. "Contextual Variables of Student Math Proficiency and Their Geographic Variations in Missouri." *Applied Geography* 109: 102040.

Rodriguez-Gómez, W. F., E. Y. Rodríguez Gómez, and J. E. Gomez-Gonzalez. 2024. "The Rural-Urban Student Performance Gap in Colombia." *Desarrollo y Sociedad* 1, no. 97: 59–75.

Sacco, C., and P. Falzetti. 2021. "Spatial Variations of School-Level Determinants of Reading Achievement in Italy." *Large-scale Assessments in Education* 9, no. 1: 12.

Sajjad, M., H. Munir, S. Kanwal, and S. A. Asad Naqvi. 2022. "Spatial Inequalities in Education Status and Its Determinants in Pakistan: A District-Level Modelling in the Context of Sustainable Development Goal-4." *Applied Geography* 140: 102665.

Salma, U.-, and M. O. Chaudhry. 2024. "Impact of Socio-Economic Factors on the Educational Attainment of the Students in Pakistan: A Case Study of Bahawalpur District." *Pakistan Journal of Humanities and Social Sciences* 12, no. 1: 2141.

Sandu, A., and C. Taylor. 2024. "Digital Learning Technologies Usage During Covid-19 Lockdowns." *British Educational Research Journal* 51, no. 2: 607–628.

Schulz, W. 2006. *Measuring the Socio-Economic Background of Students and Its Effect on Achievement in PISA 2000 and PISA 2003* (2006). Annual Meetings of the American Educational Research Association (AERA).

Sibieta, L. 2024. *Sliding Education Results and High Inequalities Should Prompt Big Rethink in Welsh Education Policy*. Institute for Fiscal Studies.

Sirin, S. R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75, no. 3: 417–453.

Sizmur, J., R. Ager, J. Bradshaw, et al. 2019. *Achievement of 15-Year-Old Pupils in England: PISA 2018 National Report*. National Foundation For Educational Research.

Smith, C., N. Parr, and S. Muhidin. 2019. "Mapping Schools' NAPLAN Results: A Spatial Inequality of School Outcomes in Australia." *Geographical Research* 57, no. 2: 133–150.

Strand, S. 2014. "School Effects and Ethnic, Gender and Socio-Economic Gaps in Educational Achievement at Age 11." *Oxford Review of Education* 40, no. 2: 223–245.

Von Stumm, S., S. N. Cave, and P. Wakeling. 2022. "Persistent Association Between Family Socioeconomic Status and Primary School Performance in Britain Over 95 Years." *NPJ Science of Learning* 7, no. 1: 4.

Taylor, C. 2018. "The Reliability of Free School Meal Eligibility as a Measure of Socio-Economic Disadvantage: Evidence From the Millennium Cohort Study in Wales." *British Journal of Educational Studies* 66, no. 1: 29–51.

Thorne Wallington, E. 2014. Thinking Geospatially: How Variable Relationships With Reaching Achievement Test Scores in the State of Missouri Vary*by Geospatial Location.

Thorne-Wallington, E. 2016. "The Geospatial Analysis of L2 Reading Achievement: Challenges and Opportunities." *Reading in a Foreign Language* 28, no. 2: 294–298.

Tseliou, F., C. Taylor, and S. Power. 2024. "Formal School Exclusions Over the eDucational Lifecourse in Wales." *Oxford Review of Education* 50, no. 6: 854–875.

Ule, M. 2015. "The Role of Parents in cHildren′s Educational Trajectories in Slovenia." *Sodobna Pedagogika* 66, no. 1: 10–45A.

Valliant, R., A. H. Dorfman, and R. M. Royall. 2000. "Finite Population Sampling and Inference: A Prediction Approach." In *Wiley Series in Probability and Statistics Survey Methodology Section*. New York: Wiley.

Voyer, D., and S. D. Voyer. 2014. "Gender Differences in Scholastic Achievement: A Meta-Analysis." *Psychological Bulletin* 140, no. 4: 1174–1204.

Watson, A., M. Kehler, and W. Martino. 2010. "The Problem of Boys' Literacy Underachievement: Raising Some Questions." *Journal of Adolescent & Adult Literacy* 53, no. 5: 356–361.

Wei, Y. D., W. Xiao, C. A. Simon, B. Liu, and Y. Ni. 2018. "Neighborhood, Race and Educational Inequality." *Cities* 73: 1–13.

Welch, A., S. Helme, and S. Lamb. 2007. "Rurality and Inequality in Education." In International Studies in Educational Inequality, Theory and Policy. *Dordrecht*, edited by R. Teese, S. Lamb, M. Duru-Bellat, et al., 602–624. Netherlands: Springer.

Wheeler, D., and M. Tiefelsdorf. 2005. "Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression." *Journal of Geographical Systems* 7, no. 2: 161–187.

Yu, H., A. S. Fotheringham, Z. Li, T. Oshan, and L. J. Wolf. 2020. "On the Measurement of Bias in Geographically Weighted Regression Models." *Spatial Statistics* 38: 100453.

Zuccotti, C. V. 2019. "Ethnicity and Neighbourhood Attainment in England and Wales: A Study of Second Generations' Spatial Integration." *Population, Space and Place* 25, no. 7: e2252.