Assessing Environmental Inequality in Air Pollution Exposure: Evidence from Montréal

API6321 Final Research Paper

Ayaka Naganuma

Student ID: 300411721

December 17, 2025

1 Introduction

1.1 Background and Motivation

Socioeconomic inequality in environmental exposure is a central concern in research on environmental justice and urban policy. A large body of empirical work shows that exposure to ambient air pollution is unevenly distributed across social groups, with lower-income and racialized communities often facing higher concentrations of harmful pollutants. These disparities have important public health implications, given the well-established links between fine particulate matter ($PM_{2.5}$) and adverse respiratory and cardiovascular outcomes.

In urban settings, pollution exposure reflects the interaction of social and physical processes. Residential sorting by income and race intersects with the spatial distribution of pollution sources such as road networks, industrial facilities, and dense built environments. As a result, observed socioeconomic gradients in pollution exposure may reflect not only social disadvantage but also the spatial organization of cities. Distinguishing between these mechanisms is essential for understanding whether inequality arises from socioeconomic status itself or from the urban environments in which different groups reside.

This distinction raises a methodological challenge. Many studies estimate associations between socioeconomic indicators and pollution exposure using conventional regression models that implicitly assume spatial independence. Yet air pollution and its physical determinants exhibit strong spatial clustering. If spatial structure is not adequately addressed, estimated socioeconomic effects may conflate social inequality with unobserved spatial processes. This study is motivated by the need to explicitly incorporate spatial structure into the analysis of environmental inequality and to clarify how urban form conditions the relationship between socioeconomic characteristics and pollution exposure.

1.2 Literature Review and the Canadian Context

Research in environmental justice has consistently documented socioeconomic disparities in air pollution exposure. Studies in the United States show that racialized and lower-income communities are exposed to higher concentrations of $PM_{2.5}$ than more advantaged groups (Hajat et al., 2015; Liu et al., 2021). These disparities are especially pronounced in urban areas, highlighting the role of urban environments in shaping the distribution of pollution burdens. Importantly, evidence also shows that while average pollution levels have declined, relative disparities across social groups have persisted or widened (Jbaily et al., 2022).

At the same time, this literature emphasizes that estimated socioeconomic gradients are sensitive to spatial scale and modeling choices. Within-city analyses using fine spatial resolution often reveal patterns that are obscured in coarser comparisons, and failure to account for spatial dependence can bias inference about socioeconomic effects.

Evidence from Canada is more limited and mixed. Several studies find that socioeconomic disadvantage is associated with higher exposure to traffic-related pollutants in major cities, but also show that these associations weaken once spatial structure is accounted for (Pinault et al., 2016). Other work suggests that exposure gradients vary across cities and depend on local urban form and land-use patterns (Giang and Castellani, 2020). In some dense urban contexts, higher-income populations may reside closer to pollution sources, complicating simple interpretations of SES–pollution relationships.

Recent studies therefore emphasize improved treatment of both socioeconomic status and spatial processes. Spatially explicit analyses show that attenuation of socioeconomic coefficients after controlling for physical determinants does not imply the absence of inequality, but rather that inequality may be embedded in the spatial organization of cities. These insights motivate the present study's focus on distinguishing between mediation and interaction mechanisms and on examining how urbanization conditions socioeconomic exposure gradients within Montréal.

1.3 Research Questions

This paper addresses two research questions.

First, do socioeconomic disparities in $PM_{2.5}$ exposure persist once spatial dependence and key physical confounders are taken into account? Second, does urban structure modify the relationship between socioeconomic characteristics and pollution exposure, such that disparities are amplified or attenuated depending on urban form?

Together, these questions shift the focus from whether inequality exists to how it is produced and distributed within urban space.

1.4 Contributions

This paper makes three contributions. First, it explicitly addresses spatial dependence in air pollution data by comparing conventional Ordinary Least Squares models with Spatial Error Models. Second, it distinguishes between mediation and interaction by treating urban structural characteristics as both potential pathways and conditioning factors linking socioeconomic status to pollution exposure. Third, it separates average pollution levels from the social distribution of exposure, showing how urban structure shapes who bears pollution burdens even when overall pollution levels are largely explained by physical characteristics.

sectionConceptual Framework and Hypotheses

subsectionConceptual Framework

This study conceptualizes socioeconomic disparities in air pollution exposure as arising from the interaction between social characteristics and urban spatial structure, rather than as the direct effect of socioeconomic status (SES) alone. Socioeconomic status, measured by income and visible minority concentration, is expected to shape residential location within the city and, through this channel, influence exposure to ambient $PM_{2.5}$. At the same time, physical characteristics of the urban environment affect both neighborhood composition and pollution levels, complicating the interpretation of simple socioeconomic gradients.

Urban structure plays two roles in the framework. First, characteristics such as greenness and land-use composition may mediate the relationship between SES and pollution exposure by shaping the environments in which different social groups reside. Second, urban structure may condition the strength of this relationship, modifying how socioeconomic characteristics translate into pollution exposure across space.

Physical determinants such as population density, major road infrastructure, and industrial emissions are treated as confounders. These factors are closely linked to pollution generation and dispersion and are correlated with both SES and $PM_{2.5}$ concentrations. Controlling for them allows socioeconomic effects to be isolated without interpreting these variables as part of the causal pathway.

A key implication is that attenuation of SES coefficients after accounting for spatial structure does not imply the absence of inequality. Rather, it suggests that inequality may be embedded in the spatial organization of cities. This motivates the separate examination of mediation and interaction effects.

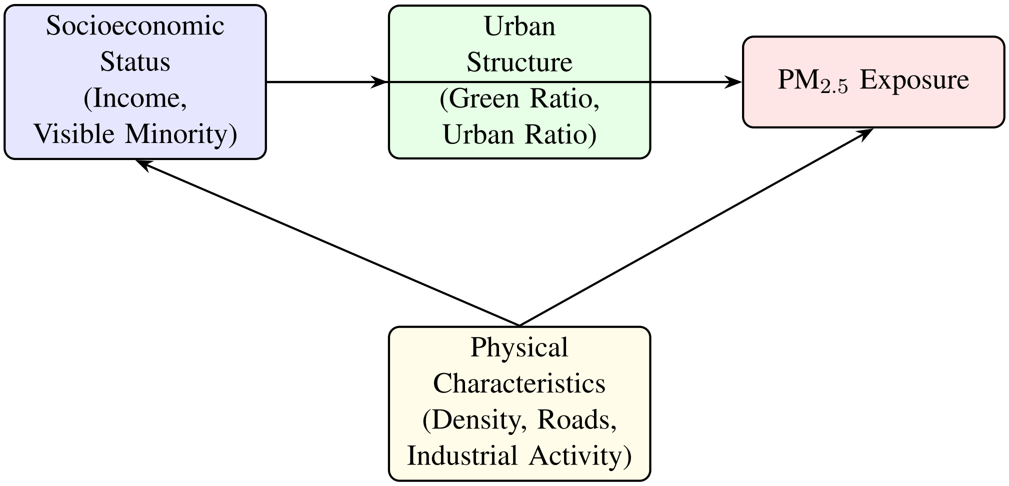


Figure 1: Conceptual framework linking socioeconomic status, urban structure, and $PM_{2.5}$ exposure.

1.5 Hypotheses

Based on this framework, the study tests four hypotheses.

H1 (Baseline association). Census tracts with lower socioeconomic status are associated with higher average $PM_{2.5}$ concentrations.

H2 (Attenuation under spatial controls). The association between socioeconomic status and $PM_{2.5}$ exposure attenuates once key physical determinants and spatial dependence are explicitly accounted for.

H3 (Mediation by urban structure). Urban structural characteristics, such as population density and land-use composition, partially mediate the relationship between socioeconomic status and $PM_{2.5}$ exposure.

H4 (Amplification through urbanization). Socioeconomic disparities in $PM_{2.5}$ exposure are amplified in more urbanized environments. An amplifying effect under H4 would be indicated by a positive interaction between minority share and urbanization. While H4 posits a potential amplification mechanism, the direction of moderation is ultimately an empirical question and may also reveal attenuation effects.

Together, these hypotheses allow the analysis to distinguish between average exposure differences, indirect pathways, and conditional mechanisms through which urban structure shapes environmental inequality.

2 Data and Variables

2.1 Data Sources

This study constructs a census-tract–level dataset for the Montréal metropolitan area by combining census, satellite-based, and geospatial data. The unit of analysis is the Census Tract (CT).

Socioeconomic characteristics are obtained from the 2016 Canadian Census Profile. Air pollution exposure is measured using satellite-based $PM_{2.5}$ concentration estimates for the period 2016–2019. Industrial pollution sources are captured using facility-level data from the National Pollutant Release Inventory (NPRI) for 2016–2018.

Urban structure variables are derived from the 2016 Road Network File and a national 2015 land cover classification raster. All datasets are harmonized using standardized census tract identifiers and a common coordinate reference system.

The initial dataset contains 970 census tracts in Montréal. After excluding observations with missing values, the final analytical sample consists of 478 census tracts.

2.2 Variable Construction

The outcome variable is ambient $PM_{2.5}$ exposure, measured as the census tract–level annual mean concentration. $PM_{2.5}$ values are averaged across multiple years to reduce short-term variability.

Socioeconomic status is summarized using measures of income, educational attainment, and visible minority share. Population density is included as a control for urban intensity. Industrial activity is captured using aggregated emissions and facility counts within a fixed spatial buffer.

Urban structure is characterized by major road length(km) and land cover composition, including green space ratio and urban land ratio. All variables are constructed at the census tract level and merged using consistent geographic identifiers. Variables with skewed distributions are transformed where appropriate in the empirical analysis.

Table 1: Variable Definitions and Data Sources

Variable	Definition	Data Source
$PM_{2.5}$	Annual mean $PM_{2.5}$ concentration averaged over 2016–2019 (μg/m^3)	NASA satellite-based estimates
Median income	Median total household income	Census 2016
Education rate	Share of population with a university degree (bachelor's level or above)	Census 2016
Visible minority share	Share of residents identifying as visible minorities	Census 2016
Population density	Total population divided by land area (persons/km^2)	Census 2016
Industrial emissions	Cumulative NPRI-reported emissions within 5 km of CT centroid, summed over 2016–2018	National Pollutant Release Inventory
Major road length(km)	Total length of major roads within census tract (km)	2016 Census Road Network File
Green space ratio	Proportion of census-tract land area classified as vegetated or natural land cover, constructed by averaging a binary green-space indicator across raster cells within each census tract	2015 Land Cover of Canada
Urban land ratio	Proportion of census-tract land area classified as highly built-up land cover, constructed analogously by averaging a binary urban-land indicator across raster cells within each census tract	2015 Land Cover of Canada

3 Exploratory Spatial Analysis

This section presents an exploratory spatial analysis of $PM_{2.5}$ exposure across census tracts in Montréal. The objective is to document the spatial distribution of air pollution and its bivariate relationship with socioeconomic characteristics, and to motivate the subsequent use of spatial econometric models. The analysis relies on visual inspection of spatial patterns and simple descriptive relationships rather than causal inference.

3.1 Spatial Distribution of $PM_{2.5}$

Figure 2 illustrates the spatial distribution of average $PM_{2.5}$ concentrations across Montréal census tracts. Pollution levels are highest in the central and eastern parts of the city, while lower concentrations are observed in peripheral and suburban areas. This pattern suggests a clear urban gradient in air pollution exposure. Importantly, $PM_{2.5}$ concentrations display strong spatial clustering. Adjacent census tracts tend to exhibit similar pollution levels, indicating that air pollution is not randomly distributed across space. Such spatial dependence is consistent with the diffusion of pollutants through atmospheric processes and with the spatial concentration of emission sources such as traffic corridors and industrial activity. This visual evidence already suggests that standard regression approaches assuming independent observations may be inappropriate in this context.

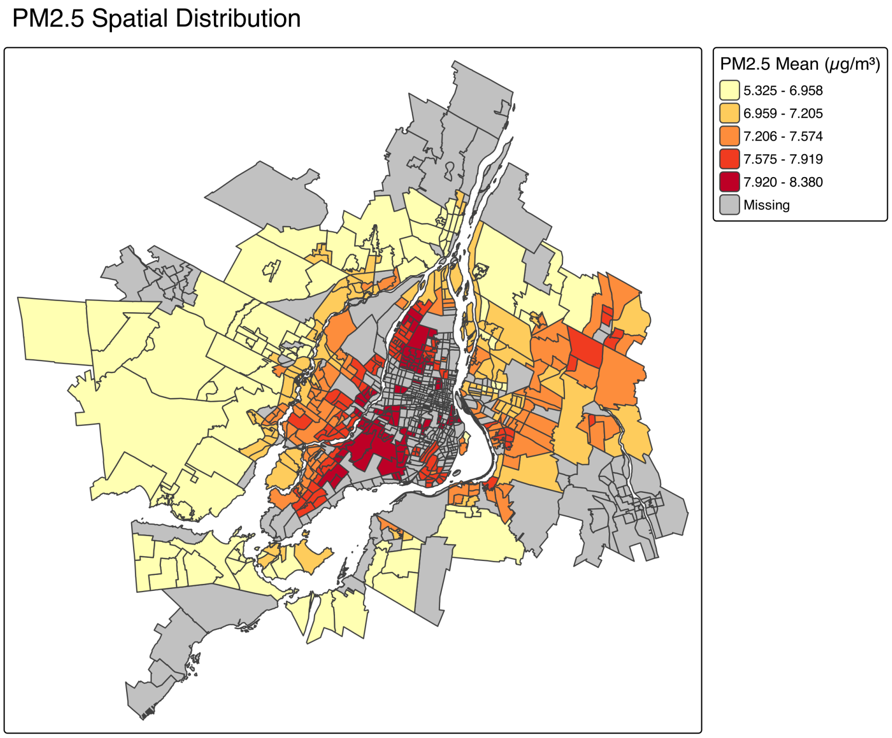
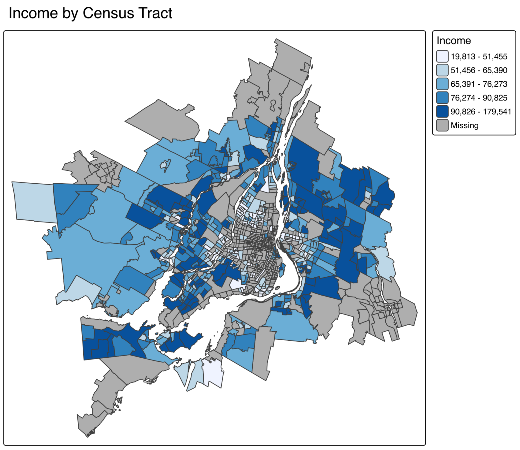


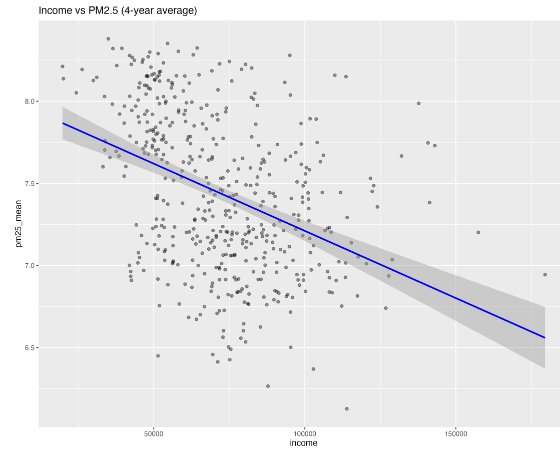
Figure 2: Spatial distribution of average $PM_{2.5}$ concentrations

3.2 Income, Education, and $PM_{2.5}$

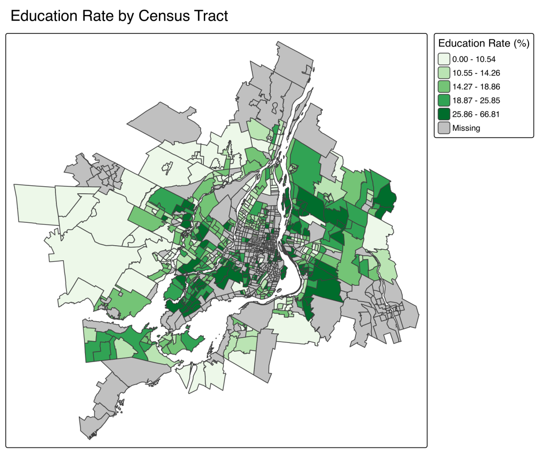
Figure 3b presents the spatial distribution of median household income alongside a scatter plot of income and $PM_{2.5}$. Census tracts with lower median income tend to experience higher $PM_{2.5}$ exposure. The negative bivariate association is consistent with the hypothesis that lower-income communities face disproportionate environmental burdens. In contrast, the relationship between education and $PM_{2.5}$ differs from that of income. As shown in Figure 3d, areas with higher educational attainment are often associated with higher pollution levels. This pattern reflects the concentration of highly educated residents in the urban core, where pollution is also elevated. In this setting, education appears to capture residential location within dense central neighborhoods rather than socioeconomic disadvantage per se. This distinction is important for interpreting education coefficients in multivariate models.



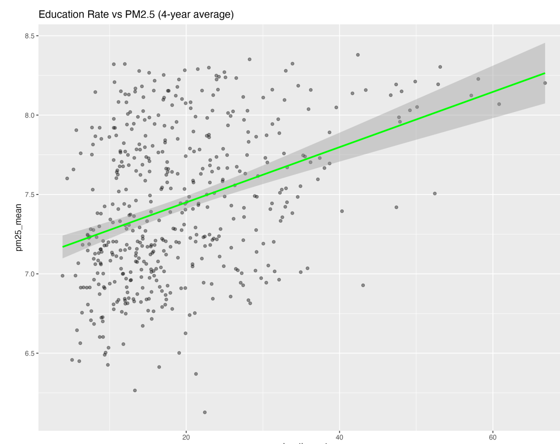
(a) Income by census tract



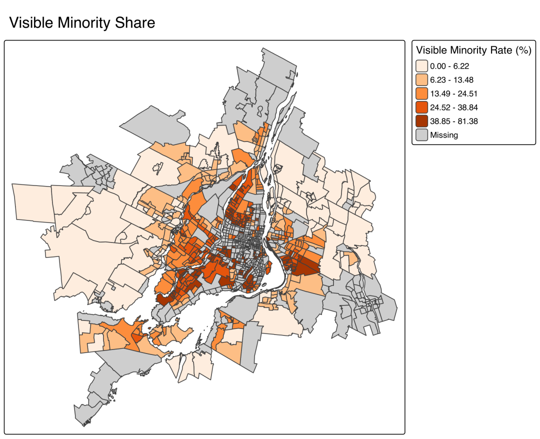
(b) Income vs. $PM_{2.5}$



(c) Education rate by census tract



(d) Education rate vs. $PM_{2.5}$



(e) Visible minority share by census tract

(f) Visible minority share vs. $PM_{2.5}$

Figure 3: Spatial distributions and bivariate associations between socioeconomic characteristics and $PM_{2.5}$ exposure. Maps illustrate residential sorting across Montréal, while scatter plots show simple bivariate relationships. These descriptive patterns motivate the multivariate and spatial econometric analysis that follows.

3.3 Visible Minority Concentration and $PM_{2.5}$

Figure 3f shows the spatial distribution of visible minority share and its relationship with $PM_{2.5}$. Census tracts with a higher share of visible minority residents systematically exhibit higher pollution exposure. Both the spatial map and the bivariate scatter indicate a positive association between minority concentration and $PM_{2.5}$ levels. This pattern suggests the presence of racialized exposure disparities within Montréal. However, as with income, these disparities may partly reflect residential sorting into dense and infrastructure-heavy areas rather than direct discrimination in pollution siting. Distinguishing between these mechanisms requires controlling for urban structural factors.

3.4 Structural Variables and Spatial Confounding

Figure 4 presents the spatial distribution of key structural variables: population density, road density, and industrial emissions. These variables are strongly concentrated in the central city and closely mirror the spatial pattern of $PM_{2.5}$. Their alignment with pollution hotspots indicates that physical and infrastructural characteristics are major determinants of air pollution exposure.

At the same time, these structural variables are correlated with socioeconomic characteristics. High-density areas tend to host both higher-income, highly educated residents and larger visible minority populations. As a result, urban structure may confound observed associations between socioeconomic status and pollution exposure.

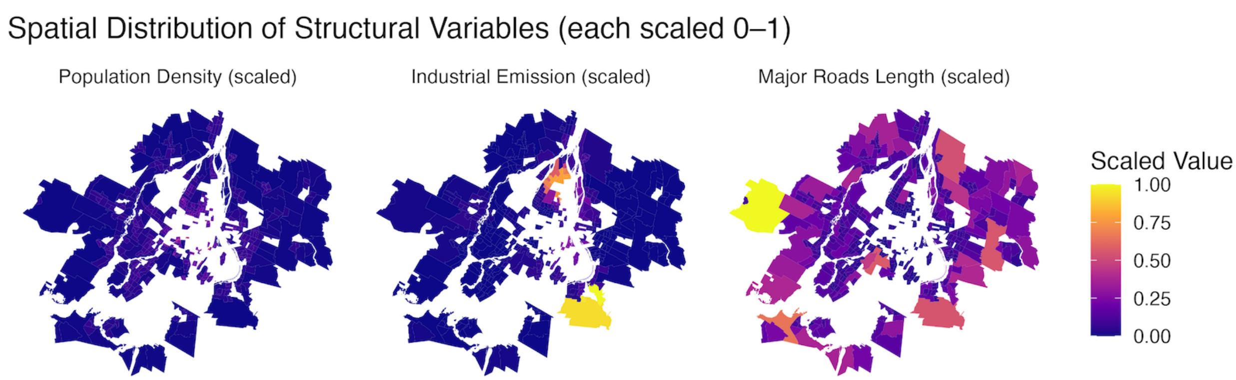


Figure 4: Spatial distribution of key structural variables (each scaled to 0–1): population density, industrial emissions, and major road length. Darker colors indicate higher relative values within the Montréal metropolitan area.

3.5 Implications for Empirical Strategy

The exploratory spatial analysis yields three main insights. First, $PM_{2.5}$ exposure in Montréal exhibits strong spatial clustering. Second, income and visible minority concentration are associated with higher pollution exposure, while education reflects urban centrality rather than disadvantage. Third, structural characteristics such as density, roads, and industrial activity are closely linked to both pollution and socioeconomic sorting.

These findings motivate the econometric strategy adopted in the next section. The presence of spatial dependence calls for spatial regression models, while the overlap between socioeconomic and structural patterns highlights the importance of explicitly controlling for physical determinants of pollution.

4 Empirical Strategy

This section outlines the empirical strategy used to examine socioeconomic inequality in $PM_{2.5}$ exposure. The analysis proceeds in three steps. First, baseline Ordinary Least Squares (OLS) models are estimated to assess the association between socioeconomic status (SES) and air pollution. Second, a series of diagnostic tests are conducted to evaluate whether the assumptions underlying OLS are satisfied. Finally, a Spatial Error Model (SEM) is introduced to explicitly account for spatial dependence in the data.

4.1 Baseline OLS Models

4.1.1 SES-Only Specification

The empirical analysis begins with a baseline Ordinary Least Squares (OLS) specification that relates $PM_{2.5}$ exposure to socioeconomic characteristics alone:

$$PM_{2.5,i} = \beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Education}_i + \beta_3 \text{Minority}_i + \varepsilon_i, \tag{1}$$

where $PM_{2.5,i}$ denotes the average $PM_{2.5}$ concentration in census tract i. Income represents median household income, Education is the share of residents with a university degree, and Minority denotes the share of visible minority residents.

This specification corresponds to the standard environmental inequality framework, in which socioeconomic disadvantage is expected to be associated with higher pollution exposure. At this stage, the model is intentionally parsimonious and abstracts from physical and infrastructural determinants of pollution in order to establish baseline socioeconomic gradients.

4.1.2 Extended Model with Physical Confounders

The baseline model is then extended to account for key physical determinants of air pollution:

$$PM_{2.5,i} = \beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Education}_i + \beta_3 \text{Minority}_i + \beta_4 \text{Industry}_i + \beta_5 \text{Roads}_i + \beta_6 \text{Density}_i + \varepsilon_i. \quad (2)$$

Population density proxies overall urban intensity, road infrastructure captures traffic-related emissions, and industrial emissions represent point-source pollution. Including these controls allows socioeconomic gradients to be evaluated conditional on urban structure, clarifying whether observed SES–pollution associations reflect social disadvantage per se or differences in the physical environments where households reside.

4.2 Model Diagnostics

While OLS provides a useful benchmark, several diagnostic tests indicate that its assumptions are violated in this setting. Variance inflation factor diagnostics for the full OLS specification without interaction terms (Model 2) indicate that multicollinearity is not a major concern. All VIF values in this specification remain within conventional thresholds, suggesting that the estimated coefficients are not driven by linear dependence among regressors. In contrast, models that include interaction terms mechanically exhibit substantial VIF values for the interacted variables. This behavior reflects the known limitations of standard VIF diagnostics in the presence of interactions rather than substantive multicollinearity. Accordingly, interaction terms are evaluated based on coefficient stability and statistical significance, and key variables are mean-centred in subsequent specifications. The Ramsey RESET test strongly rejects the correct functional form, and Moran's I reveals pronounced spatial autocorrelation in $\text{PM}_{2.5}$ levels and OLS residuals. Together, these diagnostics suggest that linear OLS models fail to capture important nonlinear and spatial processes, motivating the use of spatial econometric methods.

Taken together, these diagnostics motivate the use of spatial econometric models and the adoption of log-transformed covariates in subsequent specifications to address scale differences and potential nonlinearities in key physical and socioeconomic variables.

4.3 Spatial Error Model

4.3.1 Model Specification

To explicitly account for spatial dependence, the analysis adopts a Spatial Error Model (SEM):

$$PM_{2.5} = X\beta + u, \tag{3}$$

$$u = \lambda W u + \varepsilon, \tag{4}$$

where X is the matrix of observed covariates, W is a spatial weights matrix, λ captures spatial error dependence, and ε is an i.i.d. error term.

In addition to main effects, the SEM specifications allow for interaction terms between socioeconomic characteristics and urban structural variables, mirroring the interaction structures estimated in the OLS models. This enables the analysis to examine whether the association between socioeconomic composition and pollution exposure is conditioned by urban form even after accounting for spatially correlated unobservables.

4.3.2 Intuition

The SEM can be interpreted as correcting for omitted spatially correlated influences, such as prevailing wind patterns, historical land-use decisions, or unobserved infrastructural features. If left unmodeled, these factors induce spatial correlation in the regression residuals and can bias or inflate socioeconomic coefficients in conventional OLS models.

By explicitly modeling this correlation structure, the SEM absorbs spatial clustering into the error term, yielding more reliable estimates of socioeconomic and urban-structural associations with pollution exposure.

4.3.3 Spatial Weights Matrix

The spatial weights matrix W is constructed using a k-nearest neighbors approach with $k = 4$, such that each census tract is connected to its four closest neighbors. The matrix is row-standardized.

This specification ensures a uniform number of neighbors across tracts and is well suited to urban settings with irregular census tract sizes. It reflects the assumption that unobserved spatial dependence in pollution exposure is driven primarily by local, rather than long-distance, spatial processes.

5 Results

5.1 Main Results: OLS versus SEM

Table 2 contrasts conventional OLS specifications with spatially explicit estimates. In the SES-only OLS model (Model 1), all socioeconomic variables are strongly associated with $PM_{2.5}$ exposure. Median income is negatively related to $PM_{2.5}$ ($\hat{\beta} = -4.678 \times 10^{-6}$, $p < 0.001$), while both education rate ($\hat{\beta} = 0.0152$, $p < 0.001$) and visible minority share ($\hat{\beta} = 0.0142$, $p < 0.001$) are positively associated with $PM_{2.5}$ levels. These bivariate SES gradients account for a substantial share of variation ($R^2 = 0.562$).

Adding physical and infrastructural controls in the full linear OLS specification (Model 2) does not eliminate the SES gradients: income remains negative and statistically significant, and both education rate and visible minority share remain positive and statistically significant. Major road length enters with a negative and statistically significant coefficient ($\hat{\beta} = -0.003$, $p < 0.001$), while industrial emissions and population density are not statistically distinguishable from zero in this linear-scale specification. Variance inflation factor diagnostics for Model 2 indicate that multicollinearity is not a major concern.

Functional-form diagnostics indicate that the linear level specification may be misspecified (RESET for Model 2: $p = 2.3 \times 10^{-9}$). To assess the robustness of the estimated socioeconomic gradients to potential nonlinearities and scale effects, the analysis therefore also reports a log-transformed OLS model with an interaction between log income and visible minority share (Model 4). In this model, log income is negative and statistically significant ($\hat{\beta} = -0.334$, $p < 0.001$), log population density is positive and statistically significant ($\hat{\beta} = 0.054$, $p = 0.015$), and log major-road length is negative and statistically significant ($\hat{\beta} = -0.082$, $p = 0.027$). The visible-minority main term is marginal ($\hat{\beta} = -0.051$, $p = 0.057$), while the interaction term is positive and statistically significant ($\hat{\beta} = 0.00574$, $p = 0.019$), implying that the visible minority–$PM_{2.5}$ association increases with income and should be interpreted through marginal effects rather than the main term alone.

The Spatial Error Model (SEM) estimated on the same log-transformed specification (Model 4) shows very strong spatial error dependence ($\hat{\lambda} = 0.9206$, $p < 2.2 \times 10^{-16}$) and substantially improved fit relative to OLS (log X) (AIC improves from 241.26 to -409.41). Under the SEM, the log-income coefficient attenuates markedly and becomes only marginal ($\hat{\beta} = -0.0919$, $p = 0.056$). The visible-minority main term is also marginal ($\hat{\beta} = -0.0253$, $p = 0.075$), and the log-income-by-minority interaction is likewise marginal ($\hat{\beta} = 0.00246$, $p = 0.059$). In contrast, log population density remains large and highly significant ($\hat{\beta} = 0.0619$, $p < 0.001$), and education remains positive and significant ($\hat{\beta} = 0.00471$, $p < 0.001$). Road and industrial terms are not statistically significant in the SEM.

Spatial diagnostics confirm the importance of the spatial specification. $PM_{2.5}$ is extremely clustered across census tracts (Moran's I $= 0.893$, $p < 2.2 \times 10^{-16}$), while residual spatial autocorrelation is greatly reduced under the SEM (Moran's I $= 0.0766$, $p = 0.00418$). Overall, Hypothesis 1 is supported in non-spatial OLS models, while Hypothesis 2 is supported insofar as SES coefficients attenuate substantially once spatial dependence is modeled explicitly.

Table 2: Socioeconomic Gradients in $PM_{2.5}$ Exposure: Level OLS and Log Specifications

	Panel A: Level specification (OLS)		**Panel B: Log specification**	
	(1) SES OLS	(2) Full OLS	(3) OLS (log X)	(4) SEM (log X)
Socioeconomic variables				
Income (level)	$-0.000***$	$-0.000***$		
Education rate	$0.015***$	$0.014***$	$0.013***$	$0.005***$
Visible minority share	$0.014***$	$0.013***$	-0.051	-0.025
Physical / infrastructural controls				
Total industrial emissions		0.000		
Major road length (km)		$-0.003***$		
Population density		-0.000		
Log covariates and interaction				
Log income			$-0.334***$	-0.092
Log major road length			$-0.082*$	0.019
Log population density			$0.054*$	$0.062***$
Log industrial emissions			-0.007	0.006
Log income $\times$ Visible minority share			$0.006*$	0.002
λ (spatial error)				$0.921***$
N	478	478	478	478
R^2	0.562	0.583	0.592	
Adj. R^2	0.560	0.577	0.586	
AIC	266.480	249.819	241.256	-409.406

Notes: Panel A reports level-based OLS models. Column (2) controls for population density, major road length, and industrial emissions. Panel B reports log-transformed specifications; Column (4) estimates a spatial error model including log income, log population density, log major road length, and log industrial emissions, as well as a log income $\times$ visible minority interaction. All models use the same sample of census tracts.

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

5.2 Mediation Results

Table 3 reports SEM estimates incorporating urban-structure variables as potential mediators or correlates of $PM_{2.5}$ exposure. In the baseline SEM including centered green ratio and centered urban ratio, both variables are positively and statistically significantly associated with $PM_{2.5}$ (green ratio: $\hat{\beta} = 1.454$, $p = 0.0049$; urban ratio: $\hat{\beta} = 0.217, p = 0.0197$). In this specification, centered log income is not statistically significant, whereas centered visible minority share remains positive and significant ($\hat{\beta} = 0.00176$, $p = 0.0149$). Log population density continues to exhibit a strong positive association with $PM_{2.5}$.

Allowing for an interaction between income and green ratio provides no evidence of moderation: the income-by-green interaction is statistically insignificant, and the green-ratio main effect becomes only marginal. In the full SEM reported in Table 3, which includes the income-by-green interaction, green ratio remains positive and statistically significant, while the income-by-green interaction itself is not statistically significant and the visible minority coefficient remains positive and significant.

Because these models rely on cross-sectional data and do not exploit exogenous variation in urban form, the results are interpreted as associational evidence consistent with an urban-structure pathway rather than as causal evidence of mediation.

Table 3: SEM Estimates with Urban Structure Variables

	SEM-base	SEM-green	SEM-full
Centered log income	-0.028	-0.032	-0.025
	(0.036)	(0.037)	(0.036)
Centered visible minority	$0.00176*$	$0.00151*$	$0.00209**$
	(0.00072)	(0.00073)	(0.00073)
Green ratio	$1.454**$	0.770	$1.500**$
	(0.517)	(0.430)	(0.512)
Urban ratio	$0.217*$		0.164
	(0.093)		(0.095)
Income $\times$ green		0.148	1.515
		(1.311)	(1.340)
λ	$0.922***$	$0.924***$	$0.923***$
N	478	478	478
AIC	-412.5	-407.1	-418.1

Standard errors in parentheses.

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

5.3 Amplification and Interaction Effects

Table 4 then introduces a separate SEM specification that includes an interaction between visible minority share and urban land ratio, allowing urbanization to act as a moderator of minority exposure. Although H4 hypothesized amplification, the estimated interaction reveals attenuation, underscoring the importance of empirically testing the direction of urban moderation. Table 4 evaluates whether urban form modifies the association between visible minority concentration and $PM_{2.5}$ exposure. In the SEM with centered visible minority share interacted with centered urban ratio, the interaction term is negative and statistically significant ($\hat{\beta} = -0.0112, p = 0.0053$). In the full SEM that includes both the minority-by-urban interaction and the income-by-green interaction, the minority-by-urban interaction remains negative and statistically significant ($\hat{\beta} = -0.0126$, $p = 0.0021$), while the centered minority main effect at mean urban ratio is positive and statistically significant ($\hat{\beta} \approx 0.00209$, $p = 0.0041$).

With centering, these estimates imply that the marginal association between visible minority share and $PM_{2.5}$ is strongest in lower-urban-ratio tracts and attenuates as the built-up share increases. Under the paper's definition, where amplification requires the marginal visible-minority effect to increase with urban intensity, these results therefore indicate attenuation rather than amplification, and Hypothesis 4 is not supported.

Table 4: Urban Form as a Moderator of Minority Exposure

	SEM-Urban	SEM-Full
Centered visible minority	$0.00183*$	$0.00209**$
	(0.00073)	(0.00073)
Centered urban ratio	0.00480	0.164
	(0.080)	(0.095)
Minority × urban	$-0.0112**$	$-0.0126**$
	(0.00403)	(0.00410)
λ	$0.922***$	$0.923***$
N	478	478
AIC	-412.4	-418.1

Standard errors in parentheses.

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

6 Discussion

6.1 Spatial Structure and the Attenuation of Socioeconomic Gradients

A central result of the analysis is that socioeconomic gradients in $PM_{2.5}$ exposure that appear strong in non-spatial OLS specifications weaken materially once spatial dependence is explicitly modeled. While the SES-only OLS model yields large and highly significant coefficients for income and visible minority concentration, the preferred Spatial Error Model (SEM) produces substantially smaller and only marginal SES effects, even as measures of structural intensity remain robust.

Spatial diagnostics clarify why this attenuation occurs. $PM_{2.5}$ exposure exhibits extremely strong spatial autocorrelation (Moran's I $= 0.893$), indicating that nearby census tracts share common, unobserved determinants of pollution exposure. By explicitly modeling this spatial process, the SEM absorbs a large share of the residual spatial clustering, reducing residual autocorrelation to 0.0766. Consistent with this pattern, the estimated spatial error parameter is very large ($\hat{\lambda} \approx 0.921$), suggesting that unobserved, spatially clustered factors, such as infrastructural configurations or environmental conditions, constitute a major component of the data-generating process. In non-spatial models, these factors can be inadvertently captured by socioeconomic regressors, inflating apparent SES gradients.

At the same time, the interaction results indicate that environmental inequality is context-dependent rather than governed by a single, constant socioeconomic gradient. In the SEM interaction specifications, the interaction between visible minority share and urban land ratio is negative and statistically significant. This implies that the marginal association between visible minority concentration and $PM_{2.5}$ exposure declines as urban intensity increases. At the mean level of urban land ratio, the visible minority main effect remains positive and statistically significant, but this association becomes progressively smaller in more built-up tracts. The results therefore do not support amplification of minority exposure disparities in highly urbanized areas; instead, they point to attenuation of the minority–$PM_{2.5}$ relationship as urban intensity rises.

6.2 Policy Implications

The findings point to policy levers that operate primarily through the built environment rather than through socioeconomic targeting alone. Because population density and spatial clustering emerge as central predictors of $PM_{2.5}$ exposure, interventions aimed at reducing emissions and exposure in dense neighborhoods, such as traffic management, low-emission corridors, and land-use buffers near major emitters, may yield substantial health benefits even without explicitly targeting socioeconomic groups.

The results also highlight the importance of how green space is planned and measured. Greenness and urban-form indicators behave more like mediators than simple confounders, and their estimated associations are sensitive to spatial scale and model specification. This suggests that policy should prioritize functional green infrastructure, including vegetation barriers and street trees along high-traffic corridors, as well as equitable access to green space, rather than assuming that higher tract-level greenness necessarily translates into lower pollution exposure.

More broadly, the findings call for a rethinking of environmental justice in urban contexts such as Montréal. Rather than being defined by a stable association between socioeconomic status and exposure, environmental inequality appears to be shaped by the interaction between social composition, urban structure, and spatial processes. Justice-oriented policy should therefore incorporate spatial diagnostics and neighborhood-scale urban form as core components of both monitoring and intervention design.

7 Limitations and Future Research

This study intentionally focuses on a single metropolitan area (Montréal) and a cross-sectional snapshot in order to isolate the role of model specification, spatial dependence, and urban structure in shaping estimated SES gradients. This scope choice strengthens internal coherence (clear identification of spatial clustering and model misspecification) but limits external generalizability.

First, as a single-city analysis, the results should not be extrapolated to Canada as a whole. Montréal's historical development, road networks, and industrial geography may differ from other cities, potentially producing different SES–pollution relationships.

Second, the cross-sectional design limits causal interpretation. The analysis estimates conditional associations rather than causal effects, and it cannot distinguish between sorting, policy-driven siting, and historical land-use dynamics. For the same reason, the mediation results are presented as *suggestive* rather than causal: mediators and outcomes are measured contemporaneously, and mediator inclusion may absorb multiple mechanisms (including confounding). Relatedly, unobserved housing market mechanisms, such as land prices, rents, and residential choice, may jointly influence socioeconomic composition and pollution exposure, complicating interpretation of SES gradients in cross-sectional data.

Third, the spatial weights construction (4-nearest neighbors) implies local dependence by design, but the neighborhood graph contains multiple disconnected subgraphs. Future work should evaluate robustness across alternative k values and contiguity-based weights to ensure that inference is not sensitive to the chosen connectivity structure.

Finally, several data-related limitations should be noted. The analysis focuses on $PM_{2.5}$ exposure and does not incorporate other pollutants such as NO_2 or ozone, which may follow different spatial patterns and socioeconomic gradients. In addition, a subset of census tracts was excluded due to missing data, potentially introducing spatial selection bias if data availability is systematically related to neighborhood characteristics. Moreover, industrial emissions are aggregated at the census-tract level, which may underrepresent fine-grained variation from point sources and attenuate localized exposure effects. These limitations suggest caution in interpreting estimated coefficients as comprehensive measures of air pollution exposure.

Future research can extend this framework in three directions: (i) multi-city comparative analysis to assess heterogeneity across Canadian urban contexts, (ii) time-series or panel extensions to study changes in exposure and sorting over time, and (iii) quasi-experimental designs (e.g., policy changes, infrastructure interventions) to strengthen causal claims about mediation and interaction mechanisms.

8 Conclusion

This paper examined environmental inequality in $PM_{2.5}$ exposure across 478 Montréal census tracts using both OLS and spatial econometric models. The data exhibit extremely strong spatial clustering, and incorporating spatial dependence via a Spatial Error Model yields a very large spatial error parameter and greatly reduced residual spatial autocorrelation. Substantively, SES gradients that appear strong in conventional OLS attenuate substantially in the preferred SEM specification, while urban intensity remains robustly related to $PM_{2.5}$. Interaction evidence further indicates that the visible-minority–$PM_{2.5}$ association is not constant across urban contexts, but attenuates as urban land ratio increases. Methodologically, the central lesson is that spatial diagnostics and model choice materially affect inference about environmental inequality: much of the apparent SES gradient in non-spatial models reflects spatially structured processes and urban intensity rather than a uniform socioeconomic effect.

References

[1] Giang, A., & Castellani, K. (2020). Cumulative air pollution indicators highlight unique patterns of injustice in urban Canada. *Environmental Research Letters*, 15(12), 124063.

[2] Hajat, A., Hsia, C., & O'Neill, M. S. (2015). Socioeconomic disparities and air pollution exposure: A global review. *Current Environmental Health Reports*, 2(4), 440–450.

[3] Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, 601(7892), 228–233.

[4] Liu, J., Clark, L. P., Bechle, M. J., Hajat, A., Kim, S. Y., Robinson, A. L., … Marshall, J. D. (2021). Disparities in air pollution exposure in the United States by race/ethnicity and income, 1990–2010. *Environmental Health Perspectives*, 129(12), 127005.

[5] Pinault, L., Crouse, D., Jerrett, M., Brauer, M., & Tjepkema, M. (2016). Spatial associations between socioeconomic groups and NO_2 air pollution exposure within three large Canadian cities. *Environmental Research*, 147, 373–382.

Appendix

A.1 Correlation Structure of Variables

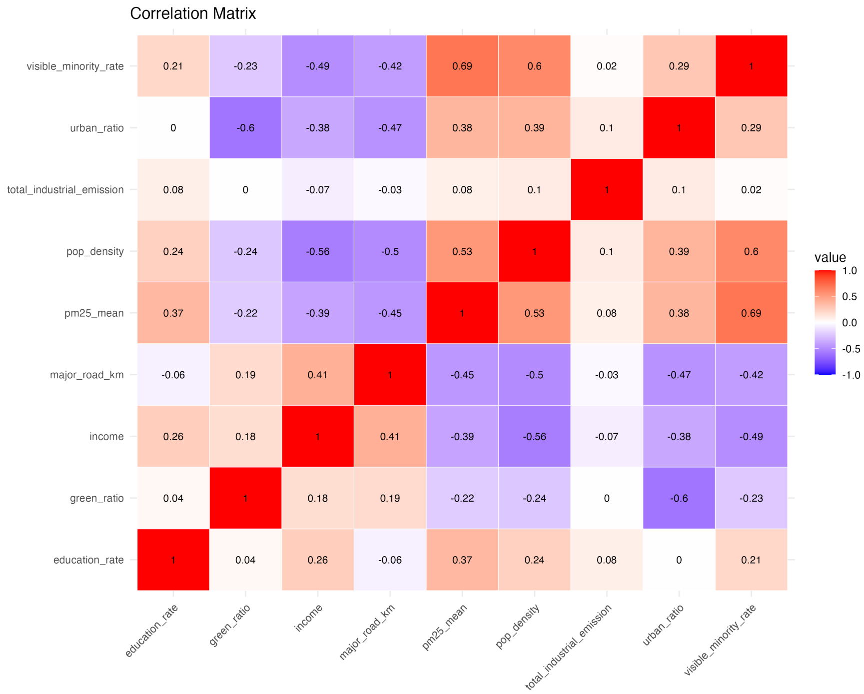


Figure 5: Pairwise correlation matrix among socioeconomic, structural, and pollution variables.

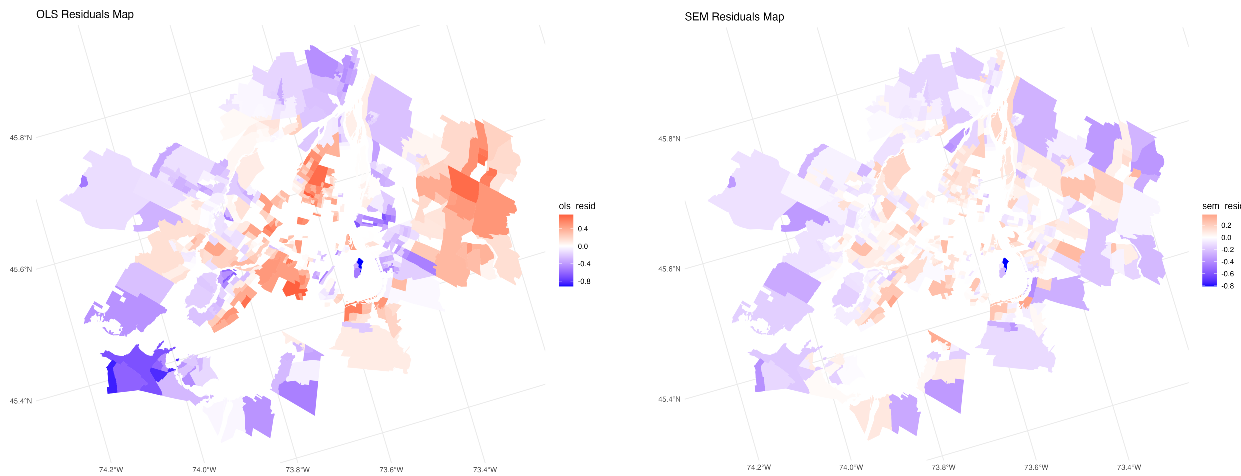
A.2 Full Regression Results

Table 5: Full OLS and Spatial Error Model Estimates (Appendix)

	SES OLS	Full OLS	Interaction OLS	OLS (log X)	SEM (log X)
(Intercept)	7.153***	7.200***	7.234***	10.480***	7.715***
	(0.068)	(0.077)	(0.090)	(1.092)	(0.555)
Income	$-0.000***$	$-0.000***$	$-0.000***$		
	(0.000)	(0.000)	(0.000)		
Education rate	$0.015***$	$0.014***$	$0.014***$	$0.013***$	$0.005***$
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Visible minority rate	$0.014***$	$0.013***$	$0.011***$	-0.051	-0.025
	(0.001)	(0.001)	(0.003)	(0.027)	(0.014)
Total industrial emissions		0.000	0.000		
		(0.000)	(0.000)		
Major road length (km)		$-0.003***$	$-0.003***$		
		(0.001)	(0.001)		
Population density		-0.000	0.000		
		(0.000)	(0.000)		
Income $\times$ minority			0.000		
			(0.000)		
Log income				$-0.334***$	-0.092
				(0.094)	(0.048)
Log major road length				$-0.082*$	0.019
				(0.037)	(0.016)
Log population density				$0.054*$	$0.062***$
				(0.022)	(0.009)
Log industrial emissions				-0.007	0.006
				(0.006)	(0.006)
Log income $\times$ minority				$0.006*$	0.002
				(0.002)	(0.001)
λ (spatial error)					$0.921***$
					(0.011)
N	478	478	478	478	478
R^2	0.562	0.583	0.583	0.592	
Adj. R^2	0.560	0.577	0.577	0.586	
AIC	266.480	249.819	251.285	241.256	-409.406
Log Likelihood	-128.240	-116.909	-116.642	-111.628	214.703

Notes: Standard errors in parentheses. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. R^2 is not reported for the spatial error model, as it is not directly comparable to OLS goodness-of-fit measures. All models use the same sample of census tracts.

A.3 Residual Spatial Patterns



(a) OLS residuals

(b) SEM residuals

Figure 6: Spatial distribution of regression residuals. The OLS residuals exhibit pronounced spatial clustering, whereas residual spatial structure is substantially reduced under the Spatial Error Model (SEM).

A.4 Spatial Autocorrelation Tests

Table 6: Moran's I Tests for Spatial Autocorrelation

	Moran's I	Expected I	Variance	p-value
$PM_{2.5}$ (levels)	0.893	-0.002	0.00090	$< 2.2 \times 10^{-16}$
SEM residuals	0.0766	-0.002	0.00089	0.00418

Notes: Moran's I statistics are computed under randomisation using a k-nearest neighbors spatial weights matrix ($k = 4$). Expected values correspond to the null hypothesis of spatial randomness.