

The Influence of Urban Forms on Transit Behavior in the Auckland Region

Dr. Mingyue Sheng
Energy Centre
The University of Auckland
Auckland, New Zealand
m.sheng@auckland.ac.nz

Prof. Basil Sharp
Department of Economics and Energy Centre
The University of Auckland
Auckland, New Zealand
b.sharp@auckland.ac.nz

Abstract — As well documented in the literature, urban form plays an essential role in determining transit ridership. However, among these studies, the majority of empirical work has not considered space as a relevant factor. Instead, most of the findings are based on a strong assumption that there is no spatial effect across the research area. This general negligence of spatial effects will, in turn, produce biased estimators if substantial geographical patterns exist. Given the observational heterogeneous distribution of transit patterns in the Auckland region, it is exceedingly doubtful whether the assumption of no spatial interdependence is valid. Based on cross-sectional data, mainly extracted from the New Zealand (NZ) 2006 census with additional geographical information compiled by ArcMap for the Auckland region, this paper contributes to the existing literature by offering insight into the spatial structure of the current public transport sector. The use of a spatial Durbin model provides a better understanding of the urban form factors that influence bus mode share by decomposing the total effect of one explanatory variable into direct and indirect effects. The results show that the total effects are comprised mostly of spatial spillover impacts. In addition to urban form variables, several other dimensions of potential bus mode share predictors are considered, including transit supply quality, accessibility to other modes of public transport, plus variables that describes household characteristics.

Keywords- Spatial dependence; spatial Durbin model; spillover effect; urban form; transit behavior

I. INTRODUCTION

A. Background Information

Throughout the world, as people's incomes rise, many shift to faster, more comfortable and more individually flexible means of transportation [1]. Not surprisingly, like most modern cities, the recent commuting pattern in Auckland, where one-third of NZ's population lives, is dominated by the automobile, with almost 88% of the share for the morning journey to work (JTW) attributed to private motor vehicles, while public transport accounted for only around 8% of the journeys [2]. In comparison to other competitor cities, data [3] from Auckland Regional Transport Agency (ARTA) confirm Auckland's position in terms of public transport use, with only 41 public transport trips made per capita per annum, while Wellington generated almost twice this number at 91 and Sydney had almost threefold (as

shown in Fig. 1).¹ Auckland is thus characterized by an elevated level of car-dependence and a low public transport patronage. Low public transport use may imply a relatively less developed public transport system, which could limit Auckland's potential to become more internationally competitive, attracting more international investment, events, and tourism.

With the aim of reducing automobile dependence and inducing non-automobile commuting, transport planners around the world are attempting to tackle the travel growth in travel demand by implementing transport planning projects that can promote forms of sustainable urban development [4]-[5]. In the case of Auckland, transport authorities have implemented several projects to facilitate the development of public transport, from small-scale projects such as expanding bus priority lanes to large-scale development such as bus and rail infrastructure initiatives. Therefore, from the perspective of local government and urban planners, it is crucial to have a solid understanding of how well the design and layout of urban areas do in terms of contributing to a reduction in automobile use and public transport travel promotion. In other words, what will be the likely impact of urban form on people's travel behavior?

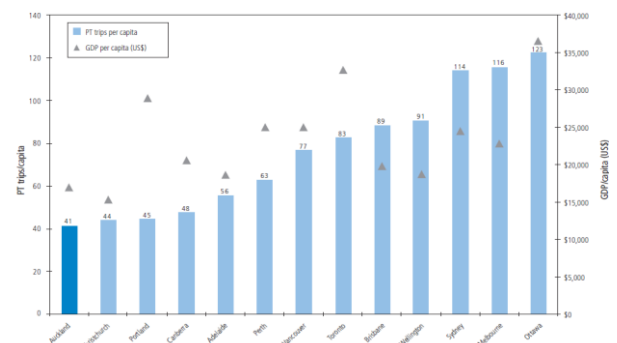


Figure 1 Public Transport Trips and GDP per capita (2007/08, US\$)

B. Objective, Motivation and Scope

ARTA reports that in Auckland, the total number of (unlinked) trips traveled by the public transport system in 2007 was 52.4 million, with buses, trains and ferries

¹ Auckland Regional Transport Agency, or ARTA, was replaced by Auckland Transport as part of the re-organization of local government on 1st November 2010. This paper still refers to ARTA since all of the data used here were compiled when ARTA was in existence.

contributed to 82%, 10%, and 8%, respectively [6]. Given its current dominant position in public transport usage, the analysis undertaken here focuses on the bus only.²

The motivation behind this paper is that to properly understand the relationship between urban form and transit ridership, it is necessary to consider the associated spatial structures more specifically. Some studies have attempted to identify the impact of urban form on different travel behaviors such as mode choices, travel demand, and travel patterns, over the past few decades [7]-[10]. A key problem with the above literature is the possibility that coefficient estimates of the impact of urban form might be attenuated by spatial dependence. Specifically, these analyses assume that observations are independent of one another in a geographical context. However, in reality, it seems unlikely that region i 's transport network regarding vehicles and public transport infrastructure is independent of that of its neighboring region j . Furthermore, from the econometric point of view, ignoring spatial characteristics between observations could, in turn, produce biased and inconsistent estimators [11].

This general limitation from past literature gives rise to the need for spatial estimation method, such as the spatial Durbin model (SDM) that had the advantage of separating the total effect of a particular variable on the transit ridership into own-region and neighborhood effects. To the best of our knowledge, [12] produced the only published paper which specifically takes spatial effects into account when analyzing the determinants of work trip bus ridership in the context of NZ, using the spatial error model (SEM) model. Their paper is also most relevant to this study. The authors use cross-sectional data involving 318 area units for the Auckland region in 2006. Once positive spatial autocorrelation is confirmed by a statistically significant Moran's I value, an SEM model is chosen because in the case of the spatial error test, its robust Lagrange Multiplier (LM) statistic is more significant than the alternative, spatial autoregressive (SAR) model. Regarding the choice of the spatial weights matrix, a Rook contiguity of order 2, including the lower order of 1, is applied based on the fact that it provides the best fit. The estimation method employed in their study is ML. After adjusting for spatial dependency, the SEM model provided more accurate parameter estimates and improving the overall predictive power of the empirical model.

However, there remains a potential weakness in interpreting [12]'s results. In addition to the spatial lag of the dependent variable included on the right-hand-side of the regression equation, it seems plausible that neighboring area unit's characteristics, for instance, population density and rush hour frequency could also play a significant role in explaining variations in a given area unit's bus ridership. This implies that further investigations of the impact of lagged explanatory variables on transit ridership are required. This study applies the SDM model, which has the ability to capture the characteristics of neighboring regions to account for any influence they may exert on their neighbor's transit ridership patterns.

² NZ census data such as ferry usage by commuters in Auckland do not exist.

The remainder of the paper is organized as follows: section two provides a review of spatial regression models, section three describes the dataset, outlines the variables used, and specifies the regression models employed. Section 4 presents some preliminary results from spatial econometric tests. Section 5 delivers the empirical results of the non-spatial ordinary least squares (OLS) model and the spatial models. The final section provides a conclusion by summarizing key findings, outlining limitation and suggesting future works of this study.

II. REVIEW OF SPATIAL REGRESSION MODELS

A. Spatial Lag Model

To address the issue of spatial autocorrelation mentioned above, prior spatial studies are mainly concerned with models that contain only one type of spatial interaction effect viz. the spatial lag model and the spatial error model. The former incorporates a spatially lagged dependent variable on the right-hand-side of a regression whereas the latter contains a spatial autoregressive process in the disturbance [13]. Following [14], the point of departure is a simple SAR model:

$$\begin{aligned} y &= \rho Wy + X\beta + u \\ u &\sim N(0, I_n\sigma^2) \end{aligned} \quad (1)$$

where y is an $n \times 1$ vector of observations on the dependent variable; X is an $n \times k$ matrix of observations on independent variables; and ρ is a spatial autocorrelation or spatial dependence parameter [15]. Additionally, ρ represents the intensity of the spatial dependence between neighboring locations, W is an $n \times n$ exogenous spatial weights matrix that specifies the assumed spatial structure and also describes the spatial arrangement of the spatial units in the sample. The element w_{ij} of W measures the nearness of area units i and j . Wy is thus the spatially lagged dependent variable which has the ability to account for various spatially related dependencies. Finally, β represents $k \times 1$ vector of estimators to be estimated, and u is an $n \times 1$ vector of independently and identically distributed (*i.i.d.*) random error terms. Equation (1) can be solved for y , and the reduced form is shown in (2):

$$y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} u \quad (2)$$

Moreover, [16] points out that in the SAR model, each observation y_i is a function of the spatial lag term Wy which represents an explanatory variable denoting the weighted average of spatially neighboring values, e.g., $y_2 = \rho(w_{12}y_1 + w_{13}y_3)$. Additionally, [17] notes that intuitively, the spatial lag term Wy , is correlated with the random error terms u , even when the latter are *i.i.d.* Consequently, it must be treated as an endogenous variable and the proper estimation technique such as Maximum Likelihood (ML), Spatial Two Stage Least Squares (S2SLS) and/or Generalised Method of Moments (GMM) should apply to account for this endogeneity problem, because OLS estimates for β are biased and inconsistent.

B. Spatial Error Model

If spatial dependence exists in disturbances, a SEM is usually applied in order to improve the precision of the estimated parameter because this kind of regression involves a

non-spherical error term. A general representation of the SEM model can be written as follows [17]:

$$y = X\beta + \varepsilon \quad (3)$$

and

$$\varepsilon = \theta W\varepsilon + u \quad (4)$$

Substituting (4) into (3) gives:

$$y = X\beta + (I_n - \theta W)^{-1}u \quad (5)$$

where θ is known as the spatial autoregressive parameter, which needs to be estimated jointly with the regression coefficients; W is the spatial weight matrix; and ε is a vector of *i.i.d* random error terms, which is assumed to be uncorrelated to u .

What is notable about the SEM model is that the spatial weights matrix W now relates to shocks in the unobserved variables (i.e. the error term ε) but not to the independent variables of the SEM model (i.e. variable X). In other words, (5) shows that the value of the dependent variable for each geographical location is influenced by the stochastic error terms at all other locations through the spatial multiplier $(I_n - \theta W)^{-1}$.

Consequently, in the context of public transport, the bus mode share at any spatial area is not only a function of the local characteristics but also of the unobserved variables at adjacent locations. In the case of spatially correlated disturbances, one should notice that even though OLS estimated-results are still unbiased, they are no longer efficient. In addition, according to [18], the classical estimators for standard errors are biased.

C. Spatial Durbin Model

In practice, one should realize that spatial dependence can have effects on both dependent and explanatory variables. Hence, according to [19], a “mixed” SDM introduced by [20], offers a more flexible alternative and might be more appropriate to apply by including the “inherent spatial autocorrelation” and the “induced spatial dependence” simultaneously. The SDM is specified as follows:

$$y = \rho W y + X\beta + W X \gamma + u \quad (6)$$

This model can be reduced to either (1) if $\gamma = 0$ or (3) if $\gamma = -\rho\beta$.

The reduced form of (6) is:

$$y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} W X \gamma + (I_n - \rho W)^{-1} u \quad (7)$$

Based on the above equations, an additional term $W X \gamma$ must be included in the model to capture the $k \times 1$ autoregression coefficient vector γ of the spatially lagged explanatory variables $W X$, which measures the marginal impact of the independent variables from adjacent observations on the dependent variable y [21].

Furthermore, [19] argues that this SDM could be developed from either an SEM [22] or from an SAR [23], and this “mixed” model can be viewed as an unrestricted model of either SEM or SAR. In other words, the SDM further nests the SAR and the SEM by involving spatial dependence in the error

term as well as in the dependent variable. Fig. 2 illustrates the theoretical relationship between SDM, SAR and SEM in a cross-sectional case.

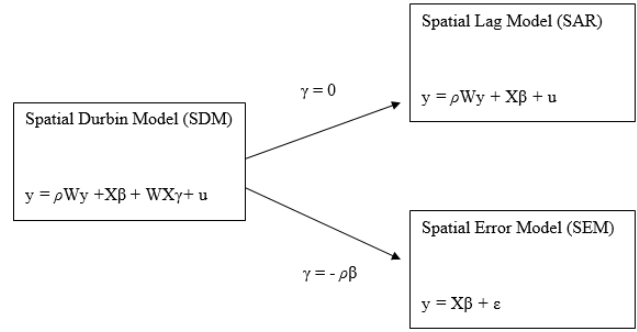


Figure 2 Relationship between SDM, SAR and SEM

According to [24], SDM is the only model that will produce unbiased estimates regardless of the true data-generation process (i.e. whether it is a spatial lag or a spatial error model). This is why the SDM is often viewed as the dominant spatial model among others. Although most transport data are geographically linked, past transport studies incorporating spatial effects are relatively scarce compared to their rich applications in other fields, such as agricultural and resource economics, housing and real estate. This empirical gap thus leads us to consider the use of spatial econometric models in the field of transport analysis.

III. DATA AND EMPIRICAL MODELS

A. Data

The major source of data for this study was the NZ Census, collected and compiled by the Statistics NZ on the census day, 6th March 2006. Additional data, such as distance to Auckland’s CBD, distance to the nearest rail or ferry terminals, and census area unit land areas, were calculated using ArcMap. Rush hour frequency, which combined the total number of buses passing through and stopping within each area unit, during both morning and afternoon peak periods, was compiled using the programs of ArcMap and Microsoft Excel. The data were geocoded at the centroid of each area unit.

The census area unit is the second smallest geographical unit defined by Statistics NZ. Area units are aggregations of meshblocks and they are non-administrative areas that are in between meshblocks and territorial authorities in size [25]. All data used in this study were compiled at this geographical level. In line with [26], smaller units such as the meshblocks would render too much variation, and consequently, increase analytical instability, while larger units such as territory authorities would aggregate data too much and are thus incapable of providing useful results.³

B. Variables

The selection of variables is mainly inspired by previous bus patronage studies. The dependent variable Bus_i is the percentage of workers in area unit i who take bus as their main

³ There are more than ten thousand mesh blocks and only seven territorial authorities in the Auckland region.

transport to work, self-reported on the census day. It was obtained by dividing the total number of bus passengers by the total number of JTW commuters in the i^{th} area unit. The percent mode share to bus offers an overall measure of the prominence of bus transport in the Auckland region.

Fig. 3 presents the spatial distribution of bus mode share in the Auckland region based on 2006 census data. From this figure, it is evident that the bus mode share is not evenly distributed across area units. More specifically, the observations do not seem to be randomly distributed over space. Area units which have a high level of bus mode share, represented by the darker color zones, tend to be closely concentrated in the center, while the area units which have a relatively low bus ridership, shown in the lighter color parts, are scattered around the boundaries.

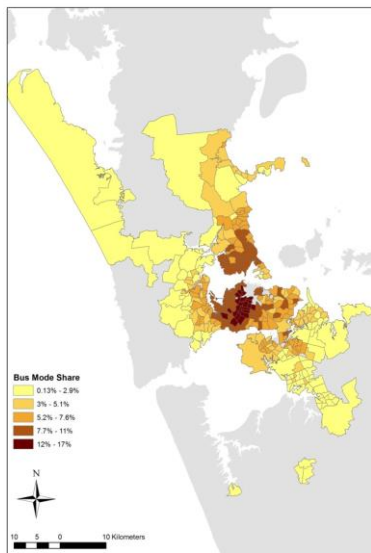


Figure 3 Spatial distribution of bus mode share in the Auckland region

Small clusterings of high values are also detected on the northeast and southeast corners of the map, which further indicates the spatially heterogeneous nature of the distribution of bus mode share. Therefore, spatial autocorrelation is apparently observed, because undoubtedly the probability of a specific value of the bus mode share variable in one specific location (area unit) depends on its value in neighboring locations.

Potential bus mode share predictors are divided into three categories: urban form, transit service, and household characteristics. The final dataset includes eight independent variables, where:

1. Urban form variables:

- *Population Density*: gross population density in the i^{th} area unit in the Auckland region, measured by the total number of inhabitants per square kilometer;
- *Employment Density*: employment density, measured by the total number of full-time and part-time employees per capita in the i^{th} area unit in the Auckland region;

- *Dwelling*: total number of private owner occupied dwellings in the i^{th} area unit in the Auckland region; used as an indicator of land use patterns;

- *CBD*: distance to CBD from the centroid of the i^{th} area unit in the Auckland region, in kilometers;

2. Transit service variables:

- *Station*: distance to the nearest public transport terminal/stop other than bus (either train or ferry) from the centroid of the i^{th} area unit in the Auckland region, measured in kilometres;
- *Frequency*: frequency of bus service within the i^{th} area unit in the Auckland region;

3. Household characteristic variables:

- *Income*: median household income measured in thousands of NZ Dollars (NZD) within the i^{th} area unit in the Auckland region;
- *Car*: mean number of motor vehicles per household within the i^{th} area unit in the Auckland region;

Reference [27] pointed out that several urban form variables such as road network type and neighborhood type, also have some influence on the demand for public transport. In addition, according to [28], a few transit service variables which describe the quality of transit service, such as the in-vehicle time and an indicator of the waiting environment, will have some effects on the demand for public transport as well. Unfortunately, these data are not available. A summary of key descriptive statistics of the variables used in this analysis are presented in Table I. As can be seen from this table, the bus share for JTW trips in the Auckland region is fairly low; the average figure for all 317 area units is only 5.65%, ranging from a low of 0.13% to a high of 17.43%.

TABLE I. AREA UNIT LEVEL DESCRIPTIVE STATISTICS OF VARIABLES FOR AUCKLAND REGION

Variables	Mean	Std. Dev.	Min	Max
Dependent variable				
<i>Bus (%)</i>	5.65	3.44	0.13	17.43
Urban Form variables				
<i>Population Density (per km²)</i>	833.98	405.64	1.47	1726.74
<i>Employment Density (per capita)</i>	0.48	0.08	0.27	0.66
<i>Dwelling</i>	1241.25	503.51	114	3270
<i>CBD (km²)</i>	16.68	8.36	2.23	43.29
Transit Service variables				
<i>Station (km²)</i>	3.67	4.17	0.14	35.53
<i>Frequency</i>	130.03	94.48	2	476
Household characteristic variables				
<i>Income (in 000 of NZD)</i>	27.11	6.32	14.4	48.4
<i>Car</i>	1.71	0.2	1.18	2.32

C. Empirical Bus Mode Share Models

A logarithmic transformation is applied to both dependent and explanatory variables with the intention of capturing the *a priori* belief that *ceteris paribus*, the impact of each explanatory variable on bus mode share is diminishing [29]-[30].

Therefore, firstly, the non-spatial bus mode share model in log-log form is specified as below:

$$\ln Bus = X\beta_{OLS} + \varepsilon_{OLS} \quad (8)$$

Equation (8) posits that the variation in the natural logarithm of the bus mode share (*lnBus*) in area unit *i* is explained by the variables in matrix *X*, which comprises a constant term, the natural logarithm of urban form, transit service, and the household characteristic variables. Since (8) is estimated by ordinary least squares, it is labelled as the OLS model and hence the estimated results from this model serve as a benchmark against the following spatial model estimations.

Secondly, the following SAR model is:

$$\ln Bus = \rho W \ln Bus + X\beta_{SAR} + u_{SAR} \quad (9)$$

Similarly, the SEM is:

$$\ln Bus = X\beta_{SEM} + \varepsilon \quad (10)$$

$$\text{where } \varepsilon = \theta W\varepsilon + u_{SEM}$$

Lastly, the SDM is given as:

$$\ln Bus = \rho W \ln Bus + X\beta + WX\gamma + u_{SDM} \quad (11)$$

IV. PRELIMINARY TESTS

A. Spatial Weights Matrix

In empirical spatial econometric models, the selection of a spatial weights matrix, normally denoted as *W*, plays an important role. As [31] outlined that there are many possible means to quantify the structure of spatial dependence between observations. Typical approaches include distance decay [32], structure of a social network [33], economic distance [34] and *k* nearest neighbors [35]. However, as [36] illustrated, one major challenge facing spatial econometric models is that the spatial weights matrix *W* cannot be directly estimated and needs to be explicitly specified a priori. Current economic theory provides no formal guidance for this. Although a wide range of literature, echoed by [37], proposed several approaches to create the spatial weights matrix; there barely exists a formal guidance on how to select the “optimal” spatial weights as existing specifications all seem somewhat arbitrary.

Practically, in spite of their lesser theoretical appeal, geographically derived weights are among the most widely applied specification in the spatial econometric analysis [18]. In addition, as [38] argued, popularity of geographically derived weights is due to the fact that the structure of *W* is constrained so that the weights are truly exogenous to the model, thus avoiding identification problems. There are two types of geographically derived weights based on proximities, namely, a binary measure of continuity (when two areas share common borders) and a continuous measure of distance. Following a majority of empirical studies [39]-[42], we use a two-dimensional Cartesian coordinate system with the ordered pair (*x*, *y*) coordinates to create a spatial weights matrix *W*, given the distance decay specification and its eigenvalues matrix *E*.

By convention, the weights matrix *W* has been row-standardized such that every row of the matrix sums to one (i.e. $\sum_j w_{ij} = 1$). Each element of *W* is therefore defined as:

$$w_{ij} = 0 \text{ if } i = j$$

$$w_{ij} = \frac{1}{d_{ij}} \text{ if } d_{ij} \leq d^* \text{ and;}$$

$$w_{ij} = 0 \text{ if } d_{ij} > d^* \text{ if observation } i \neq j$$

where *d_{ij}* is the spherical distance between the centroids of area units *i* and *j*, and *d** is the critical cut-off distance. This inverse Euclidean distance, *d_{ij}*, contains a maximum threshold band of 24.14 kilometers to guarantee connections between all area units, that is, each spatial unit must have at least one neighbor.⁴ This indicates that two area units are considered neighbors when the distance between their centroids is less than 24.14 kilometers and not neighbors if their centroids lie 24.14 or more kilometers apart.

B. Moran's I Test

A univariate Moran's *I* test for residuals is the most commonly employed first-step specification test for spatial autocorrelation [43]. The test does not specify an explicit alternative spatial model (i.e. either SAR or SEM models) but has power against both [44].

The Moran's *I* test for residuals in matrix notation is captured by:

$$I = (N / S_0)(e'We / e'e)$$

where *e* denotes a vector of OLS residuals, and $S_0 = \sum_j \sum_j w_{ij}$, a standardization factor that refers to the sum of the weights for the non-zero cross products [45].

According to [46], the interpretation of Moran's *I* should be parallel to a correlation coefficient; however the major distinction is that its value is not bounded by the closed, (-1, +1) interval. A positive value signals positive spatial autocorrelation, measuring the occurrence of similar levels of a variable being found over contiguous or nearby spaces. By contrast, a negative value signals negative spatial autocorrelation, measuring the joint occurrence of high and low attribute values in adjoining locations.

The Moran's *I* statistic shows a positive value of 18.733 with a p-value that is less than 0.0001. As expected, this result indicates that the null hypothesis of no spatial dependence should be rejected. Furthermore, the test statistic indicates that positive spatial autocorrelation exists, and in order to obtain unbiased and consistent estimators, spatial models should be adapted instead of the non-spatial OLS estimations.

C. The Lagrange Multiplier Test

By applying the Lagrange Multiplier (LM) test, we select between a spatial lag and a spatial error alternative. Basically there are two major forms of the LM test. The LM_{lag} statistic tests the null hypothesis of no spatial autocorrelation in the dependent variable; the LM_{error} statistic, on the other hand, tests the null hypothesis of no significant spatial autocorrelation in the error terms.

The LM test against a spatial lag alternative (LM_{lag}) is demonstrated in [20] and took the following form:

⁴ The default unit for cut-off length is in miles in *Stata 11*, by conversion, 15 miles are approximately equal to 24.14 kilometers.

$$LM_{lag} = [e'Wy / (e'e/N)]^2 / D$$

where $D = [(WX \beta)'(In - X(X'X)^{-1}X'(WX \beta)/\sigma^2)] + tr(W^2 + W'W)$.⁵

By contrast, the LM test against a spatial error alternative (LM_{error}) which is initially outlined in [47], takes the form of:

$$LM_{error} = [e'We / (e'e/N)]^2 / [tr(W^2 + W'W)]$$

Aside from a scaling factor, this statistic corresponds to the squared value of Moran's I .

If both hypotheses can be rejected, one should consider constructing robust forms of these LM tests which have the ability to correct for the presence of local misspecification of the other form [48]-[49]. The test procedures of LM_{lag}^r and LM_{error}^r are identical to the one described above.⁶ Both the classic and the robust LM tests are based on the residuals of the OLS model and are asymptotically distributed as $\chi^2(1)$.

Table II presents the diagnostics for spatial dependence. Under the classic LM test, both hypotheses of no spatially lagged dependent variable and of no spatially autocorrelated disturbances can be rejected at a 1% significance level. Robust LM tests consistently show the same results, with rejection of both hypotheses at a 1% significance level. This implies that OLS is rejected for both SAR and SEM models.

TABLE II. OLS DIAGNOSTICS FOR SPATIAL DEPENDENCE

Measure	Statistic	Significance
SEM		
LM_{error}	123.601	***
LM_{error}^r	30.395	***
SAR		
LM_{lag}	128.174	***
LM_{lag}^r	34.968	***

Unlike what holds for the SAR's counterpart, the Autoregressive (AR) model in time series analysis, the OLS estimation in the presence of spatial dependence will be inconsistent, simply because of the endogeneity issue discussed before. Therefore, in this study, the SAR and SEM models are estimated using ML estimation [50].

V. ESTIMATION RESULTS

A. Non-spatial and Spatial Models

The results from the non-spatial OLS is reported in Table III, alongside with the estimated results under the SDM. Several distinctions are evident. Firstly, the estimated coefficients of the urban form variables are significant and of expected signs. However, against expectations, the variable $ln(Dwelling)$ is not significant. Additionally, compared with the SDM, most estimated coefficients from the OLS are larger in magnitude,

⁵ "tr" denotes the trace of the matrix W.

⁶ The subscript "r" denotes "robust".

implying that without the inclusion of potential spatial autocorrelation between dependent, independent variables and the error terms, the OLS results simply ignore this spatial variation and produce biased estimates.

TABLE III. NON-SPATIAL OLS, SPATIAL AUTOREGRESSIVE MODEL AND SPATIAL DURBIN MODEL (DEPENDENT VARIABLE: $LNBUS$)

Explanatory Variable	OLS Estimates	SDM Estimates
Intercept	2.749 ***	3.522
$ln(\text{Population Density})$	0.138 ***	0.141 ***
$ln(\text{Employment Density})$	1.646 ***	-0.285
$ln(\text{Dwelling})$	-0.218	-0.166 *
$ln(\text{CBD})$	-0.97 ***	-0.511 ***
$ln(\text{Station})$	0.26 ***	0.12 ***
$ln(\text{Frequency})$	0.158 ***	0.143 ***
$ln(\text{Income})$	-1.053 ***	-0.579 ***
$ln(\text{Car})$	-0.865 ***	-0.732 ***
Lag $ln(\text{Population Density})$		-0.379 **
Lag $ln(\text{Employment Density})$		3.528
Lag $ln(\text{Dwelling})$		-0.066
Lag $ln(\text{CBD})$		0.48 **
Lag $ln(\text{Station})$		0.098 *
Lag $ln(\text{Frequency})$		0.281
Lag $ln(\text{Income})$		0.41
Lag $ln(\text{Car})$		-3.613 ***
ρ		0.823 ***
Squared Correlation	0.73	0.852
Variance Ratio		0.779
Log likelihood		-59.929

The value of R-squared (R^2) is 0.730, indicating a reasonable model fit.⁷ However, as the results from the Moran's I statistic and model diagnostic tests in Table 3.2 show, estimates using the OLS method suffers from a major problem: there is evidence of a positive spatial autocorrelation, and the LM test statistic suggests the lag specification as the appropriate alternative. Thus, the above OLS estimates should be interpreted with caution.

Therefore, we should concentrate on the estimated results from the SDM. First of all, upward bias is found in most of the least-squares estimates, suggesting over-estimation of the sensitivity of bus mode share to the urban form, transit supply, and socio-economic and demographic characteristics when spatial dependence is disregarded. Secondly, the spatial lags on $ln(\text{Population Density})$, $ln(\text{CBD})$ and $ln(\text{Station})$ are all significant in the SDM, implying possible omitted variable issue if we do not include them in the non-spatial OLS model. Thirdly, by taking the spatial lag into account, the fit of the model has improved dramatically. The R^2 statistic for the SDM model is 0.852, which has higher value compared to the

⁷ Adjusted R-squared for OLS estimation is 0.725.

one in OLS. Therefore, after adjusting for spatial dependence, the overall fitness of the model has been improved.

B. Choosing between Alternative Spatial Dependence Models

As [13] described, if the OLS model is rejected in favor of both SAR and SEM models, then the SDM should be estimated. Therefore, a likelihood ratio (LR) test, also known as the score test, can subsequently be used to test two separate hypotheses that $H_0: \gamma = 0$ and $H_0: \rho\beta + \gamma = 0$.

Recall that the SDM model is reduced to the SAR model if $\gamma = 0$. Reference [19] proposed that when there is evidence of maintaining the SAR or SEM model, the SDM model specified by (4) and the following log-likelihood tests may be useful regarding determining the “true” spatial process. Thus, for the SAR model, one can determine the dominant model by testing the null hypothesis $\gamma = 0$. Rejecting the null hypothesis implies rejecting the SAR. Similarly, a common factor constraint: $\gamma = -\rho\beta$ should be tested to determine the best model between the SDM and its SEM. Likewise, if the null is rejected, this indicates statistical evidence for the SDM. With the aid of the LR test, one can decide the better model between the SDM and its restricted versions.

The likelihood ratio (λ) is defined as:

$$\lambda = 2[\ln(L_U) - \ln(L_R)] \sim \chi^2(m)$$

where L_U is the likelihood function of the unrestricted model (i.e. $L_U = L_{SDM}$) whereas L_R is the likelihood function of the restricted model (i.e. $L_R = L_{SAR}$ or L_{SEM}), and m is the number of restrictions imposed. The idea is that if the restrictions are valid, the log-likelihood functions should appear to be similar in values, and accordingly, λ should equal zero.

The following results are obtained: $L_{SDM} = -59.929$, $L_{SAR} = -86.097$ and $L_{SEM} = -75.867$.

With 8 degrees of freedom, the critical values at 1%, 5% and 10% significance are 1.646, 2.733 and 3.490, respectively. The test statistics exceed the critical values for all cases, therefore we can reject the null hypothesis that the underlying spatial process is SAR or SEM at a 1% significance level. In other words, the restriction on parameter γ associated WX and the common factor constraint are invalid. As a result, the unrestricted SDM should be employed to represent the data-generation process of the spatial dependence. This result further implies that the spatial lags of both the dependent and explanatory variables should be included in the model. In fact, the inclusion of the spatial lags of independent variables makes reasonable sense as area units located near each other should have some degree of similarity regarding urban form, transit supply and household characteristics variables, because economic activities tend to interact largely across space.

The estimation results for the SDM model are summarized in Table III, alongside the OLS estimates. Overall, the SDM explains over 85% of the variation in the bus mode share.

C. Decomposing Total Effect into Direct and Indirect Effects

Interpretation of the SDM model differs from that of its non-spatial regression counterpart, the ordinary least squares,

as the k^{th} parameter vector β is no longer a partial derivative of y with respect to change in the k^{th} independent variable from the $n \times k$ matrix of X [51]. Essentially, the spatial dependence components in the SDM expand the spatial information set to include additional information from neighboring area units. To see the impact of this additional spatial information, consider the partial derivative of the SDM in (5) with respect to a particular explanatory variable x_k :

$$M = \frac{\partial y}{\partial x_k} = \underbrace{(I - \rho W)^{-1}}_{n \times n} \underbrace{[\beta_k]}_{n \times n} + \underbrace{W\gamma_k}_{n \times n} \tag{12}$$

The partial derivative results in an $n \times n$ matrix M representing marginal effects, which is shown in (12). The impact on the dependent variable from a change in a coefficient can be decomposed into three ways, namely, direct, indirect and total effects. The direct effect is defined as the average of the diagonal elements of matrix M by [24], it provides a summary measure that represents an average of the impacts on bus mode share arising from own-region changes in variable x_k . The indirect effect is defined as the average of the off-diagonal elements of matrix M ; this effect is also known as the spatial spillover effect as it measures the impact on bus mode share in area unit i arising from changes in variable x_k from all other area units. The total effect is calculated as the average row sums of matrix M ; it includes both direct plus indirect effects. The total effect measures the average cumulative impact on each observation from changing the k^{th} explanatory variable by one unit across all observations.

Average direct, indirect and total effects estimated are reported in Table IV, along with inferential statistics (i.e. the figures in parenthesis are bootstrapped standard errors) calculated using a bootstrap method with 1,000 draws. Because all of the variables are expressed in natural logs, the coefficient estimates can also be interpreted as elasticities.

For the total effects, all estimated parameter values have the expected signs, with one exception for $\ln(\text{Population Density})$. The total effects of $\ln(\text{Station})$ and $\ln(\text{Frequency})$ on transit ridership are all positive and significant; while the total effects of $\ln(\text{Dwelling})$, $\ln(\text{CBD})$, $\ln(\text{Income})$ and $\ln(\text{Car})$ and $\ln(\text{Population Density})$ are negative and significant. Separating the total effect of a regressor into direct and indirect effects yields further insights.

For the two transit service variables, first of all, the total effect, which comprises the direct and indirect effects of $\ln(\text{Station})$, is positive and significant, implying that across the Auckland region, as the distance to train station and/or ferry terminal increases, commuters will prefer to choose buses as their transport mode. Next, both the direct and indirect effects of rush hour frequency show a significant positive effect on the bus mode share in a given area. This result provides insights to transport planning *viz.* that by increasing the number of buses during morning and peak hours, the effect will not only be reflected in a rise in the percentage of commuters who choose to take bus to work in their own district, but also an additional spillover benefit which can be reflected in nearby areas. The elasticity of total effect of this

variable is about 2.4, which indicates that increasing the transit frequency in area unit i by one percent, the average bus mode share across all area units will rise by 2.4%, holding other variables constant.

TABLE IV. DIRECT, INDIRECT AND TOTAL EFFECTS OF THE SPATIAL DURBIN MODEL

Variables	Direct Effect	Indirect Effect	Total Effect
$\ln(\text{Population Density})$	0.136 ***	-1.48 ***	-1.345 ***
	-4.70E-05	-4.70E-05	-5.80E-08
$\ln(\text{Dwelling})$	-0.17 ***	-1.141 ***	-1.311 ***
	-3.50E-05	-3.50E-05	-3.50E-08
$\ln(\text{CBD})$	-0.51 ***	0.335 ***	-0.175 ***
	-1.00E-05	-1.00E-05	-1.30E-08
$\ln(\text{Station})$	0.124 ***	1.108 ***	1.232 ***
	-3.40E-05	-3.30E-05	-5.10E-08
$\ln(\text{Frequency})$	0.151 ***	2.245 ***	2.395 ***
	-7.00E-05	-7.10E-05	-1.20E-07
$\ln(\text{Income})$	-0.58 ***	-0.374 ***	-0.955 ***
	-1.10E-05	-1.20E-05	-7.00E-09
$\ln(\text{Car})$	-0.816 ***	-23.732 ***	-24.548 ***
	-0.00075	-0.00069	-8.40E-07

Secondly, for the group of variables that have negative impact to bus mode shares, the parameter estimate on one of the urban form variables, $\ln(\text{Dwelling})$, indicates that the larger the share of private owner occupied dwellings within an area unit, the lower the share of commuters who take bus to work, which seems intuitively plausible. The estimated coefficient on the total effect of another urban form variable, $\ln(\text{CBD})$, is negative and significant, suggesting that the propensity to take a bus decreases as the area unit is farther away from the CBD in the Auckland region.

For the two household characteristic variables, both the direct and indirect effects of income level exert a significant negative impact on the bus mode share, reflecting the idea that bus transport is an inferior good: as the commuters become wealthier, they will make fewer bus patronages for their JTW trips. Moreover, the direct effect of $\ln(\text{Car})$ show that there is an inverse relationship between the number of cars owned per household and the bus usage rate, which is in line with the findings on car ownership variable found in [52] and [53]. The indirect effect of cars exhibits the same tendency, suggesting that with a one percent rise in the number of cars owned in adjacent area unit j , the average bus mode share in any given area unit i tends to decline by approximately 23.7%. A possible explanation for this phenomenon is that increasing car use and access in neighboring areas results in feedback forces and a "follow the neighbor" philosophy.

Thirdly, the estimated coefficient on the total effect of $\ln(\text{Population Density})$ is negative and significant at the 1% level. This implies that the population density from all observed area units affects the percentage of workers who take bus as their main transport to work negatively, which runs

counter to our original hypothesis that high population density leads to high transit ridership.

As discussed earlier, total effect can be unraveled into direct (own-region) and indirect (spatial spillover) effect. Some notable findings were revealed by our results: two urban form variables: $\ln(\text{Population Density})$ and $\ln(\text{CBD})$ have opposite signs for direct and indirect effect parameters; while the rest stays consistent. The estimated coefficient on the direct effect of $\ln(\text{Population Density})$ is positive and significant at the 1% level. The result is consistent with the assumptions made by previous studies without consideration of spatial effects [54]-[55], where people living in high-density sectors prefer to use more public transport or walk more frequently, but will make fewer and shorter trips by private vehicles. However, the indirect effect is negative and also significant, suggesting that once the population density in nearby regions increases, the bus mode share in area unit i will decline. This outcome may be because commuters in area unit i interpret the rise in population density in their neighboring regions as a sign of potential congestion issues and dissatisfaction of the transit service, since buses might not be running on time, in such cases, taking private vehicles will be a better alternative than using public transport. Because the indirect effect is larger in magnitude, the total effect of $\ln(\text{Population Density})$ is negative.

For the next urban form variable $\ln(\text{CBD})$, the direct effect is negative and significant at the 1% level, suggesting that commuters are less willing to take the bus to work if they live farther away from Auckland city center. While the spatial spillover effect is positive and significant, suggesting that the bus mode share in region i will tend to rise if commuters in nearby regions live further away from the CBD. The parameter estimated for the direct effect of distance to CBD is negative and this may reflect the less attractiveness of using public transit for commuters who live far away from the city center, due to the fact there may be less public transport options available in their region. On the other hand, the positive indirect impact indicates that as a result of the spillover effect, local commuters may find riding buses is a better option for longer trips, especially when the workplace is far from the commuter's residential address. Moreover, the variable $\ln(\text{CBD})$ may also serve as a proxy for the cost of public transport, therefore, those living further away may find it more cost effective to use their private vehicles to go to work. The sign of the total effect of this variable is negative because the magnitude of the negative direct effect outweighs the positive indirect effect.

Another significant finding from the SDM is that except for $\ln(\text{CBD})$ and $\ln(\text{Income})$, the total effects are comprised mostly of the spatial spillover impacts, while only a relatively small portion is attributed to the direct effects on bus mode share that arose from own-region changes in variable x_k . For instance, the indirect effect of $\ln(\text{Car})$ constitutes nearly 97% of the total impact of number of cars on bus mode share. Therefore, for the case of spatial dependence considered in the SDM model, least-squares regressions that ignore this spatial spillover effect and only produce the coefficients that

represent the summary impact measures, result in biased and inconsistent estimates. The result also reveals that spatial spillovers dominate in transit behavior analysis and greater attention should be paid from transport and urban planners to both own-region effect and the impact from neighborhood when evaluating new projects, or making transport investment decisions.

VI. CONCLUSION, LIMITATION AND FUTURE RESEARCH

This paper estimated how urban form variables are related to bus mode share and how these effects vary across the Auckland region. Overall, based on area unit level data, the analysis highlighted the complexity and importance of spatial structure in determining the factors that influence bus mode share.

OLS as used in many previous transport-related studies assumes that the observations/regions are independent of one another in a geographical context. OLS thus looks for similarities in different spatial areas and provides an 'average' figure to cover the whole space. However, this is not plausible when using spatially-defined data because they are likely to exhibit positive spatial autocorrelation. That is, correlation of a variable with itself through space. Ignoring the spatial characteristics between observations/regions can, in turn, produce biased and inconsistent estimators.

By conducting an in-depth case study using Auckland data, urban form, coupled with other factors that affect the bus mode share are explored and these, in turn, are related to a spatial context. Moran's *I* test shows that there is statistically significant evidence of the presence of positive spatial autocorrelation. Therefore, by taking spatial dependence into account, spatial regression models are selected over the non-spatial OLS model to obtain unbiased and consistent estimators. The empirical results show that bus mode share in one area unit exhibits a positive relationship with the share in neighboring area units.

However, the interpretation of these findings based on SAR and SEM models is confounded by the strong spatial autocorrelation of the urban form and other transit characteristics such as transit supply, and socioeconomic/demographic differences across area units. By applying the likelihood ratio tests, this paper confirms the existence of spatial dependency in the lags of both dependent and independent variables. This dominating spatial issue has been addressed by the use of the spatial Durbin model. Estimated results from SDM show that the total effects comprised mostly of spatial spillover impacts, and only a relatively small percentage is attributed to the direct effects on bus mode share that arose from own-region changes in any given explanatory variable. For planners and developers, the SDM model is not only technically superior, but also preferable for evaluating new projects and making investment decisions [56]. As unlike traditionally estimated coefficient interpretations, one can easily unravel total effect into own-region and spatial spillover effect. The results presented in this paper indicate that knowledge of a specific spatial lag may

provide clues about the importance of future land use patterns on transit ridership.

One limitation is that there is commonly an endogeneity issue with service frequency as an explanatory variable in a regression model with patronage as the dependent variable. Public transport providers often base the frequency they provide on patronage within an area unit; therefore, frequency may depend on patronage, since public transport providers gear service levels according to patronage. The variable *Frequency* in this SDM is thus suspected to be an endogenous variable to the bus mode share. One possible way to investigate the potential endogeneity issue here is to apply the Durbin-Wu-Hausman (DWH) endogeneity test [57]. The DWH endogeneity test requires the use of a valid instrumental variable (IV) to *Frequency*. Ideally, this IV is assumed to be correlated with *Frequency* but uncorrelated with ε_{OLS} . But, within this dataset, a valid IV variable is impossible to obtain. Another limitation is the lack of enough data to include other public transport modes (rail and ferry) in this study, unfortunately current dataset (NZ Census 2006) is not comprehensive enough to take consideration of alternative transport modes other than buses, and our results might be sensitive to such inclusions.

This empirical analysis suggests several directions for future empirical research. First of all, this cross-sectional approach will not be sufficient to show the impact of variables that do not vary across area units. With the presence of panel data, which has the ability to capture time trends, it would be necessary to investigate explicitly the dynamics of these other variables such as the fuel prices.⁸ Secondly, current models would benefit from a more comprehensive dataset which comprises more transport supply side variables such as seat capacity and labor/capital cost. Lastly, regarding model methodology, another approach, Bayesian estimation, which has the advantage of allowing comparison of various weight matrices based on Bayesian posterior model probability, could be applied in future research and compared with the maximum likelihood estimation.

REFERENCES

- [1] A. Downs, *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*. Brookings Institution Press, Washington, D.C, 2003.
- [2] MoT, "Travel patterns: household travel," Wellington: Ministry of Transport, 2013. <http://www.transport.govt.nz/ourwork/tmif/travelpatterns/tp006/>
- [3] ARTA, "2009/10 – 2011/12 Auckland regional land transport programme," Auckland: Auckland Regional Transport Authority.
- [4] D. Banister and S. Marshall, *Encouraging Transport Alternatives: Good Practice in Reducing Travel*. London: The Stationery Office, 2000.
- [5] H. Barton, G. Davies, and R. Guise, *Sustainable Settlements: A Guide for Planners, Designers and Developers*. Bristol: University of the West of England, Severnside Research and Consultancy Unit, 1995.

⁸ Reference [58] conducted a recent study that attempts to determine the factors that influence the demand for public transport in both short-term and long-run for three cities in NZ. Using time series data from the period 1996 to 2008, this paper highlights that considerable fluctuation in fuel price exhibits a positive effect on public transport patronage in all three cities.

- [6] ARTA, "Annual report 2006/07," Auckland: Auckland Regional Transport Authority, 2007.
- [7] P. Gordon, A. Kumar, and H.W. Richardson, "Congestion, changing metropolitan structure and city size in the United States," *International Regional Science Review*, vol. 12, pp. 45-56, 1989.
- [8] P. Headicar, and C. Curtis, "Residential development and car-based travel: does location make a difference?" 22nd PTRC European Transport Forum, Proceedings of Seminar C: Environmental issues, University of Warwick, 12-16 September 1994.
- [9] R. Kitamura, P. Mokhtarian, and L. Laidet, "A micro-analysis of land use and travel in five neighbourhoods in the San Francisco Bay area," *Transportation*, vol. 24, pp. 125-158, 1997.
- [10] P. Næss, P.G. Røe, and S. Larsen, "Travelling distances, modal split and transportation energy in thirty residential areas in Oslo," *Journal of Environmental Planning and Management*, vol. 38, pp. 349-370, 1995.
- [11] J.P. LeSage and R.K. Pace, "Spatial econometric models," in *Handbook Of Applied Spatial Analysis: Software Tools, Methods And Applications*, M.M. Fischer and A. Getis, Eds, Berlin: Springer, 2010, pp. 355-376.
- [12] M. Greer and B. van Campen, "Influences on public transport utilization: the case of Auckland," *Journal of Public Transportation*, vol. 14, pp. 51-68, 2011.
- [13] J.P. Elhorst, "Applied spatial econometrics: raising the bar," *Spatial Economic Analysis*, vol. 5, pp. 10-28, 2010.
- [14] L. Anselin, "Spatial data analysis with GIS: an introduction to application in the social sciences," Technical Report, Santa Barbara: University of California, National Center for Geographic Information and Analysis, 1992.
- [15] M.M. Wall, "A close look at the spatial structure implied by the CAR and SAR models," *Journal of Statistical Planning and Inference*, vol. 121, pp. 311-324, 2004.
- [16] J.P. LeSage, "Lecture 1: maximum likelihood estimation of spatial regression models," Unpublished lecture notes, University of Toledo, 2004.
- [17] L. Anselin, "Spatial externalities, spatial multipliers and spatial econometrics," *International Regional Science Review*, vol. 26, pp. 153-166, 2003.
- [18] L. Anselin, "Spatial econometrics," in *A companion to theoretical econometrics*, chapter 14, B.H. Baltagi, Ed. Massachusetts: Blackwell Publishers, 2001, pp. 310-330.
- [19] L. Osland, "An application of spatial econometrics in relation to hedonic house price modelling," *Journal of Real Estate Society*, vol. 32, pp. 289-320, 2010.
- [20] L. Anselin, *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers, 1988.
- [21] W.D. Kissling and G. Carl, "Spatial autocorrelation and the selection of simultaneous autoregressive models," *Global Ecology and Biogeography*, vol. 17, pp. 59-71, 2008.
- [22] L. Anselin, "Spatial econometrics," in *Palgrave handbook of econometrics: vol. I, econometric theory*. T.C. Mills and K. Patterson, Eds. Basingstoke: Palgrave Macmillan, 2006, pp. 901-969.
- [23] R. Bivand, "Regression modelling with spatial dependence: an application of some class selection and estimation methods," *Geographical Analysis*, vol. 16, pp. 25-37, 1984.
- [24] J.P. LeSage and R.K. Pace, *Introduction to Spatial Econometrics*, Boca Raton: Chapman and Hall/CRC, Taylor and Francis Group, 2009.
- [25] Statistics New Zealand. "Area unit," Wellington: Statistics New Zealand. http://www.stats.govt.nz/surveys_and_methods/methods/classifications-andstandards/classification-related-stats-standards/territorial-authority/definition.aspx
- [26] D. Yu, N.A. Peterson, M.A. Sheffer, R.J. Reid, and J.E. Schnieder, "Tobacco outlet density and demographics: analysing the relationships with a spatial regression approach," *Public Health*, vol. 124, pp. 412-416, 2010.
- [27] TRB, "TCRP Report 16, Transit and urban form", vol. I, Part I, Transit, urban form, and the built environment: a summary of knowledge. Washington, D.C: National Research Council, Transportation Research Board, 1996.
- [28] N. Paulley, R. Balcombe, R. Mackett, H. Titheridge, J. Preston, M. Wardman, J. Shires, and P. White, P, "The demand for public transport: the effects of fares, quality of service, income and car ownership," *Transport Policy*, vol. 13, pp. 295-306, 2006.
- [29] G. Bresson, J. Dargay, J-L. Madre, and A. Pirotte, "Economic and structural determinants of the demand for public transport: an analysis on a panel of French urban areas using shrinkage estimators," *Transportation Research Part A*, vol. 38, pp. 269-285, 2004.
- [30] J. Gomez-Ibanez, "Big city transit ridership, deficits and politics: avoiding reality in Boston," *Journal of the American Planning Association*, vol. 62, pp. 30-50, 1996.
- [31] J.P. LeSage, "Ph.D. Topics Lectures: spatial econometrics," Unpublished, University of Toledo, 2002.
- [32] L. Anselin, "Estimation methods for spatial autoregressive structures," *Regional Science Dissertation and Monograph Series No.8*. New York: Cornell University, Program in Urban and Regional Science, 1980.
- [33] P. Doreian, "Linear models with spatially distributed data, spatial disturbances or spatial effects," *Sociological Methods and Research*, vol. 9, pp. 29-60, 1980.
- [34] A. Case, H.S. Rosen, and J.R. Hines, "Budget spillovers and fiscal policy interdependence: evidence from the States," *Journal of Public Economics*, vol. 52, pp. 285-307, 1993.
- [35] J. Pinkse and M.E. Slade, "Contracting in space: an application of spatial statistics to discrete-choice models," *Journal of Econometrics*, vol. 85, pp. 125-154, 1998.
- [36] R. Leenders, "Modeling social influence through network autocorrelation: constructing the weight matrix," *Social Networks*, vol. 24, pp. 21-47, 2002.
- [37] L. Anselin, "Under the hood: issues in the specification and interpretation of spatial regression models," *Agricultural Economics*, vol. 27, pp. 247-267, 2002.
- [38] C.F. Manski, "Identification of endogenous social effects: the reflection problem," *Review of Economic Studies*, vol. 60, pp. 531-542, 1993.
- [39] B. Fingleton, "Estimates of time to economic convergence: an analysis of regions of the European Union," *International Regional Science Review*, vol. 22, pp. 5-34, 1999.
- [40] B. Fingleton, "Spatial econometrics economic geography, dynamics and equilibrium: a third way?" *Environment and Planning A*, vol. 32, pp. 1481-1498, 2000.
- [41] J. Le Gallo, C. Ertur, and C. Baumont, "A spatial econometric analysis of convergence across European regions, 1980-1995," in *European regional growth*, B. Fingleton, Ed. Berlin: Springer, pp. 99-129, 2003.
- [42] S.J. Rey, and M.G. Boarnet, "A taxonomy of spatial econometric models for simultaneous equations systems," in *Advances in spatial econometrics: Methodology, tools and applications*, L. Anselin L, R.J. Florax and S.J. Rey, Eds. New York: Springer, pp. 99-119, 2004.
- [43] P.A.P. Moran, "The interpretation of statistical maps," *Journal of the Royal Statistical Society B*, vol. 10, pp. 243-251, 1948.
- [44] L. Anselin and S. Rey, "Properties of tests for spatial dependence in linear regression models," *Geographical Analysis*, vol. 23, pp. 112-131, 1991.
- [45] L. Anselin, "Interactive techniques and explanatory spatial data analysis," in *Geographical information systems: Principles, techniques, management and applications*, P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind Eds. New York: John Wiley and Sons, pp. 251-264, 1999.
- [46] R.J. Florax and P. Nijkamp, "Misspecification in linear spatial regression models," *Discussion Papers no. TI 2003-081/3*, Amsterdam: Tinbergen Institute, 2003.
- [47] P. Burridge, "On the Cliff-Ord test for spatial autocorrelation," *Journal of the Royal Statistical Society B*, vol. 42, pp. 107-108, 1980.
- [48] L. Anselin, A. Bera, R.J., and M. Yoon, "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics*, vol. 26, pp. 77-104, 1996.
- [49] A. Bera and M.J. Yoon, "Specification testing with locally misspecified alternative," *Econometric Theory*, vol. 9, pp. 649-658, 1993.

- [50] J.K. Ord, "Estimation methods for models of spatial interaction," *Journal of American Statistical Association*, vol. 70, pp. 120-126, 1975.
- [51] J.P. LeSage and M.M. Fischer, "Spatial growth regressions: model specification, estimation and interpretation," *Spatial Economic Analysis*, vol. 3, pp. 275-340, 2008.
- [52] F. Zhao, L.F. Chow, M.T. Li, and X. Liu, "Transit ridership model based on geographically weighted regression," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1972, pp. 105-114, 2006.
- [53] C. Vance and R. Hedel, "The impact of urban form on car travel: disentangling causation from correlation," *Transportation*, vol. 34, pp. 575-588, 2007.
- [54] K. Maat, B. van Wee, and D. Stead, "Land use and travel behaviour: expected effects from the perspective of utility theory and activity based theories," *Environment and Planning B*, vol. 32, pp. 33-46, 2005.
- [55] R.L. Steiner, "Residential density and travel patterns: review of the literature," *Transportation Research Record*, vol. 1466, pp. 37-43, 1994.
- [56] F. Goetzke and P.M. Andrade, "Walkability as a summary measure in a spatially autoregressive mode choice model: an instrumental variable approach," in *Spatial Analysis, Advances in Spatial Science*, A. Páez et al., Eds. Berlin Heidelberg: Springer-Verlag, 2010.
- [57] R. Davidson J.G. MacKinnon, *Estimation and Inference in Econometrics*. Oxford University Press, New York: 1993.
- [58] J. Wang, "Appraisal of factors influencing public transport patronage," NZTA research report no.434. Wellington: New Zealand Transport Agency, 2011.

Author Profile:

Dr. Mingyue Sheng is a research fellow in the Energy Centre at The University of Auckland, New Zealand. She holds a PhD in Economics from the University of Auckland. She comes with expertise in applied economics/econometrics, and advanced data analysis especially in the field of transport economics. Before joining the Energy Centre, she was one of the lecturing team members from the Department of Civil and Environmental Engineering at the University of Auckland. Her current research focuses on intelligent transport systems, road safety issues and spatial econometrics.

Prof. Basil Sharp was appointed Director of the Energy Centre at the University of Auckland, New Zealand in 2008. His research interests span the field of environmental economics and law and economics. He has a specific interest in developing microeconomic foundations for analyzing different mechanisms for allocating natural resources, and the performance of rights-based systems of governance in water resources, nitrogen pollution, climate change and fisheries. He has applied (spatial) econometric analysis and non-market valuation methods to a range of environmental problems in New Zealand.