PUBLIC TRANSPORT ACCESS AND ITS IMPLICATIONS FOR HOUSEHOLD CAR OWNERSHIP: A CASE STUDY OF THE SYDNEY GREATER METROPOLITAN AREA

Bernard Trendle
Project Officer, Statistics and Analytics, Transport Performance and Analytics, Freight, Strategy and Planning, Transport for NSW, Chippendale, NSW, 2008, Australia. Email: bernard.trendle@transport.nsw.gov.au.

ABSTRACT: A comparison of data from census 2011 and census 2016 indicates that a slightly higher proportion of commuters within the Sydney Greater Metropolitan Area use public transport, with the figure increasing from 17.3 per cent in 2011 to 19.8 per cent in 2016. The same data also indicates that, not only has the number of cars increased, but so too has the number of cars per household. Furthermore, car ownership by household is found to vary across the Greater Metropolitan Area. Other work has reported that Australians in the bottom income quintile are more likely to experience transport difficulties. Households in this category are more likely to live on the urban fringe, where there is lower access to public transport. This change in accessibility will likely impact on the relationship between household income and car ownership. The aim of this paper is to estimate the strength of this relationship and see if it varies across Sydney's Greater Metropolitan Area. This is done by introducing spatial regimes into the modelling. The novelty of the work done here, is that rather than specifying these regimes, they are created by letting the data speak for itself. Natural breaks in the variable measuring of the proportion of employed persons journeying to work by public transport are identified and used to define the regimes.

KEY WORDS: Regional; Urban; Land Use; Transportation.

1. INTRODUCTION

Sydney's Greater Metropolitan Area (GMA) is a planning area used by Transport for New South Wales (TfNSW). It extends from Port Stephens in the North, to beyond Ulladalla in the south and is bordered in the west by the Great Dividing Range. At the time of the 2016 Census, there were 5.8 million people residing in the region, with around 2.7 million of them employed. Of these employed people, the same data source indicates that of those who indicated a mode of transport, 20.1 per cent indicated they used public transport.

Census data also indicates that there were, on average, 1.7 cars per household in the GMA at census time 2016. This figure is up slightly from the 1.6 cars per household recorded in the 2011 census. The same data source indicates that the number of cars owned per household varies systematically across the Sydney GMA, with the incidence of car ownership lowest in some of the SA2s with the highest income. Access to public transport also varies across the GMA, and census data indicates that higher proportions of employed people journey to work using some form of public transport from SA2s within the Inner West and Eastern Suburbs of Sydney. These SA2s also tend to have higher average incomes. Additionally, the census data indicates that many low income families, living on the urban fringe or rural areas, rely heavily on private motor vehicles.

Rosier and Macdonald (2011) note that, while the proportion of Australians who feel they often cannot travel to destinations they need to visit is small, it is Australians in the bottom income quintiles who are more likely to experience transport difficulties. These authors also note that transport disadvantage is common in specific geographical locations, including outer-urban and rural areas. Rosier and Macdonald (2011) suggest that in outer-urban areas, transport disadvantage is the result of a range of factors, including the fact that the population in these areas comprises a higher proportion of low income households, and that public transport infrastructure is poorer in these areas.

One of the consequences of differences in access to public transport, is that the relationship between household income and car ownership may vary systematically across geographic space. A direct implication of Rosier and Macdonald (2011), is that households on the urban fringe, where access to public transport is lower, will rely more heavily on private motor vehicles. The aim of this paper is to estimate the strength of the relationship between household income and car ownership levels and, in particular, explore how this relationship varies across Sydney's Greater Metropolitan Area (GMA), where income levels, demographic characteristics and urban density change (see, for example, O'Sullivan 2014).

The paper proceeds as follows, section 2 provides a literature review, focusing on determinants of car ownership. This is followed, in section 3, with an outline of the data used. This section also provides a descriptive analysis of the relationship between income and car ownership. Section 4 presents a number of models of the variation in household car ownership.

The section starts with a simple Ordinary Least Squares (OLS) model, before introducing a model incorporating spatial regimes. A brief conclusion is provided in section 5.

2. MODELLING HOUSEHOLD CAR OWNERSHIP - A LITERATURE REVIEW

Reviews of international literature by de Jong *et al*. (2004) and Anowar, *et al*. (2014) provide a comprehensive description of approaches used to model car ownership. These papers are complemented by the work of Goodwin *et al*. (2004) and Graham and Glaister (2004). These latter two papers set out to establish the consensus on the impact of various household, demographic and geographic factors on the level of car ownership. This consensus is established by comparing elasticities from the studies reviewed in these publications.

A common theme in the work cited in Goodwin *et al*. (2004) and Graham and Glaister (2004) is that income is an important factor in determining car ownership, with car ownership levels increasing as incomes rise. The strength of this relationship is measured as the elasticity of car ownership with respect to income. Depending on the nature of the data used in the reported study, this elasticity typically lies between a low of 0.5 and an upper limit of just above 1.0 (see, for example, Graham and Glaister, 2004; Clark, 2009).

The modelling strategy adopted in this work follows Clark (2007; 2009), using data derived from regional or zonal units. In the current study, SA2 level data from the 403 SA2s which make up Sydney's GMA are used to explore the factors determining variation in local car ownership. Anowar *et al*. (2014) note that these aggregate type studies are seldom used for policy work as they fail to capture the underlying behavioural mechanisms that guide the household decision process. The use of data of this type limits the analysis to a focus on a static point in time. Complex feedback loops between the household location decision, mode of travel to work, and the household car ownership decision cannot be included in this type of work. However, the main focus of the current study is the possibility that the relationship between income and car ownership varies across the GMA. In this case models using aggregate zonal data are likely to prove adequate, as demonstrated by Clark (2007; 2009).

While the primary objective of this paper is to explore the relationship between income and car ownership, there are also likely to be other variables that influence household car ownership levels. Demographic variables such as gender, age, education and the employment status of the population are often cited as significant in transportation research. For age, studies have suggested that older people are more likely to be tied to a motor vehicle through habit and preference (Clark, 2009). Household size and composition (i.e. presence of children) positively influence the level of car ownership (see Clarke, 2007; 2009), while higher levels of education have been found to increase environmental concern and change people's attitudes toward vehicle ownership (Flamm, 2009).

Regional land use characteristics are likely to play an important role in car ownership. Variables in this category include the level of urbanization of the area and the area's access to transport infrastructure (Potoglou and Kanaroglou, 2008). Households in more urbanised areas are less likely to own a car than households in thinly populated areas, while access to public transport infrastructure has an effect on car ownership levels among households (Clark, 2009).

Finally, transportation factors have also been explored in previous research in an attempt to understand their influence on household car ownership levels (see Maltha *et al*., 2017). These factors include availability of substitutes to motor vehicles, such as public transport, the accessibility of destination regions, or the distances that are travelled.

3. THE DATA

Table 1 provides summary statistics for a number of variables relating to the SA2's of the Sydney GMA. The data in this table indicates that, on average, there were just under 1.75 cars per household in the Sydney GMA at the time of the 2016 census, with this statistic ranging from a high of 3 (Horsley Park - Kemps Creek), to a low of 0.56 (Sydney - Haymarket - The Rocks). Similarly, on average, households comprised 2.76 persons across the GMA, and range from a high of 3.70 (Cecil Hills in Sydney's west - adjacent to Kemps creek), to a low of 1.58 persons (Potts Point - Woolloomooloo).

In table 1, *Cars_Hhld* is defined as the average number of cars per household by SA2 at census 2016. This variable is derived by dividing the number of cars for each SA2, by the number of households. *Ave_inc* is the average weekly income of each SA2 and in this study has been derived by multiplying the mid-point of each income band by the number of persons in the band, summing all bands, and dividing by the number of persons who responded to the income question. For the final, 'upper' income band, the approach followed Needleman (1978), creating a midpoint by multiplying the minimum income of the upper band by 1.5.

Table 1. Summary Statistics.

Variable name	Variable description	Average	Std Dev.	Max	Min
Cars_Hhld	Average no. Of cars per household in each SA2	1.75	0.37	3.00	0.55
Ave_inc	Average household weekly income	1 998.67	502.87	3 408.63	1 026.74
P_Ptransport	Per cent of employed persons travelling by public transport	0.20	0.13	0.51	0.00
Hhld_size	Average no. of persons per household for the SA2	2.75	0.38	3.70	1.58
Pop_Density	No. Of persons per square kilometre	2 338.84	2 272.54	1 4529.50	1.21
P_Aged65plus	Per cent of the population aged 65 or older	0.15	0.05	0.45	0.04
P_NESB	Per cent of the population that is non-English speaking	0.14	0.05	0.42	0.02
P_Bachelor	Per cent of persons aged 20 and over with a bachelor degree	0.26	0.13	0.57	0.05

Source: the Author.

P_Ptransport is the proportion of employed persons who indicated they travelled to work by public transport. This variable has been derived by dividing the number of persons who indicated that they journeyed to work by public transport, for each SA2, by the total number of employed persons who responded to the mode question in Census 2016 for the SA2 in question. Two additional variables capturing regional or landuse characteristics are included, being Hhld_size and Pop_Density. Hhld_size is the average number of persons per household, and is derived by dividing the population of each SA2 in the study area, by the number of households in the corresponding SA2. Pop_Density is the population density, and is derived by dividing each SA2's area in square kilometres by its population.

Three demographic variables are also included in Table 1, being *P_Aged65plus, P_NESB* and *P_Bachelor. P_Aged65plus* is the proportion of the population of the SA2 aged 65 or older at census time 2016. This variable is derived by dividing the population aged 65 or older in each SA2 by its total population. P_Bachelor is defined as the proportion of each SA2s population with a Bachelor or higher level qualification and is derived as the number of persons with a Bachelor or higher level qualification, divided by the population aged 20 or older. P_Nesb, the proportion who don't speak English well has been derived by dividing the sum of respondents who indicated that they didn't speak English well (Not well or Not at all in the published categories), by all persons who responded to this question (i.e. Very well, Well, Not Well and Not at all).

The distribution of car ownership across the GMA is provided in figure 1. This figure provides a quintile map, showing cars per household across the SA2s of the GMA in five bands. In figure 1, the lightest shaded areas have less than 1.47 cars per household. As the shading darkens the average number of cars per household for the SA2 in question rises, with the highest category being more than 2.04 cars per household. The map, together with the inset, indicates that areas with relatively few cars per household can be found in the Inner West and Eastern Suburbs, a hardly surprising result, given the population density, associated traffic congestion and high accessibility to public transport in these areas.

Figure 2 provides details of the geographic distribution of household income, again a quintile map is provided. Here the data indicates that high income areas tend to be clustered in Sydney's Inner West and Eastern Suburbs, the North Shore and along the Northern Beaches of Sydney. However, a careful comparison of figures 1 and 2 suggests that the relationship between the number of cars owned by household and the income of households might not be as straightforward as expected. For example, areas on the southern side of Broken Bay are in the highest income category of the map and also fell into the highest band for car ownership in figure 1. In contrast, in the Inner West there are many SA2s which fall into the highest income quintile, but are in the lowest quintile for car ownership.

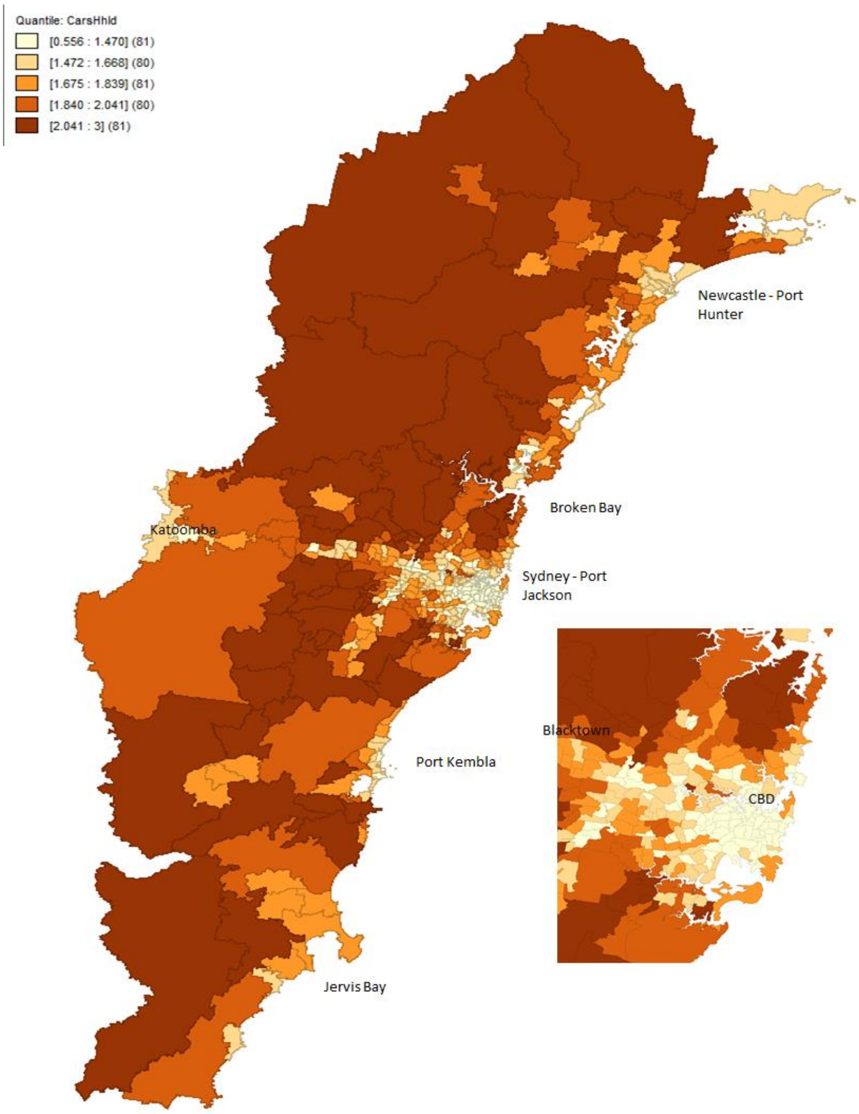

Figure 1. Quintile Map: Cars Per Household. Source: the Author.

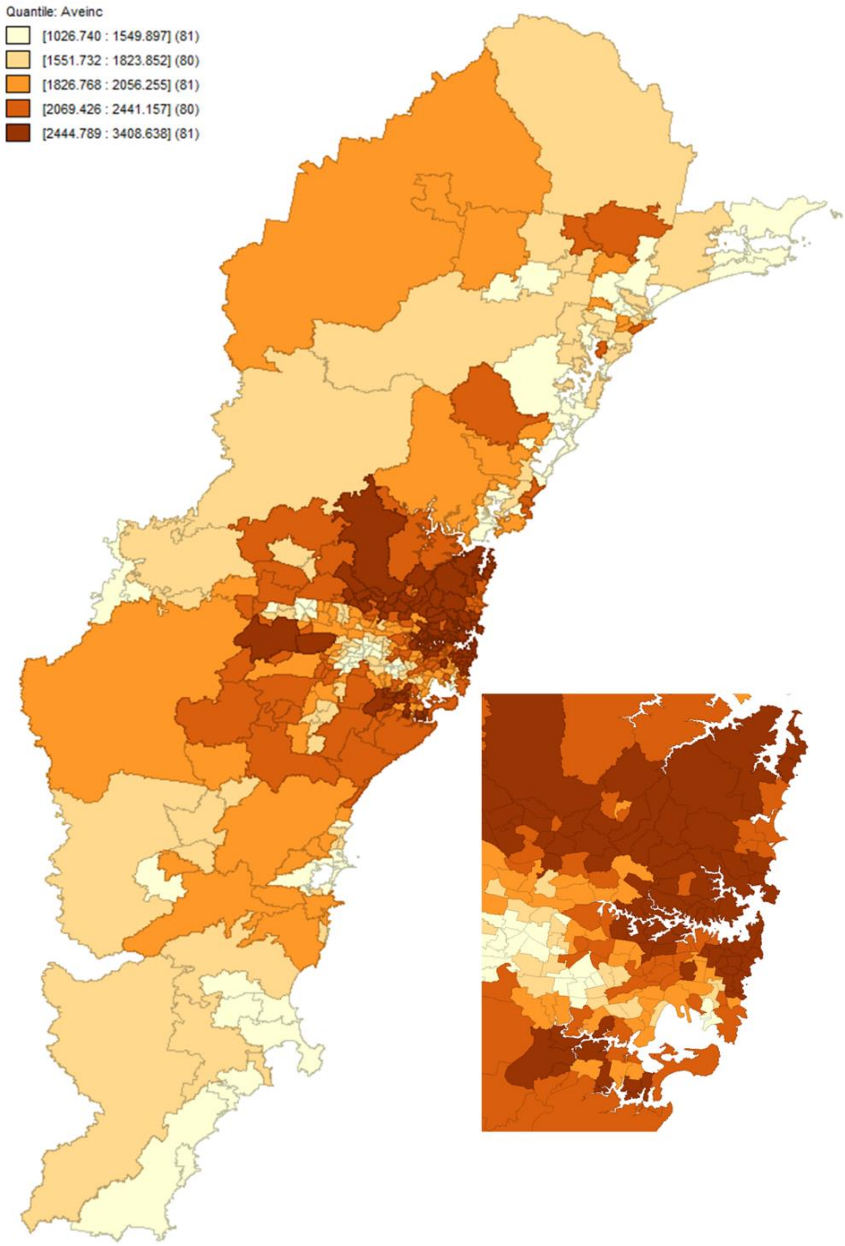

Figure 2. Quantile Map: Average Income. Source: the Author.

The complexity of the relationship between the number of cars owned per household and income is highlighted in figure 3. Each of the nine scatter diagrams in this figure shows average weekly household income by SA2 along the horizontal axis and the number of cars per household along the vertical axis. Before presenting the data, the SA2s in the sample are subsetted according to where they sit in terms of the proportion traveling to work by public transport and the population density of the SA2s. Specifically, the SA2s are first divided into three equally sized groups in terms of the proportion travelling to work by public transport (lowest third, middle third and highest third). Each of these three groups are then divided into three equally sized groups according to the population density of the SA2 (again, lowest, middle and highest third). The result are nine charts showing the relationship between the number of cars owned per household and average weekly income for groups of SA2s, depending on where they sit according to both the proportion traveling to work by public transport and the population density.

The scatter plot in the bottom left hand corner of Figure 3 shows the relationship between income and the number of cars per household for households in SA2s with the lowest population densities, which also use the lowest levels of public transport to journey to work. Note that the trend line through the points on this plot is steeper than any of the other trend lines, indicating that, for this group, changes in the number of cars owned per household is more sensitive to income changes than for any other group subsetted in this figure.

In contrast, the scatter plot in the top right hand corner shows the relationship between household income and the number of cars owned per household in the most densely settled SA2s of metropolitan Sydney. The SA2s in the scatter diagram also have the highest proportion of residents commuting to work via pubic transport and are in the top income group identified in the figure. In contrast to all other scatter diagrams in figure 3, the relationship highlighted in the scatter diagram demonstrates a negative (but insignificant) relationship between household income and the number of cars per household.

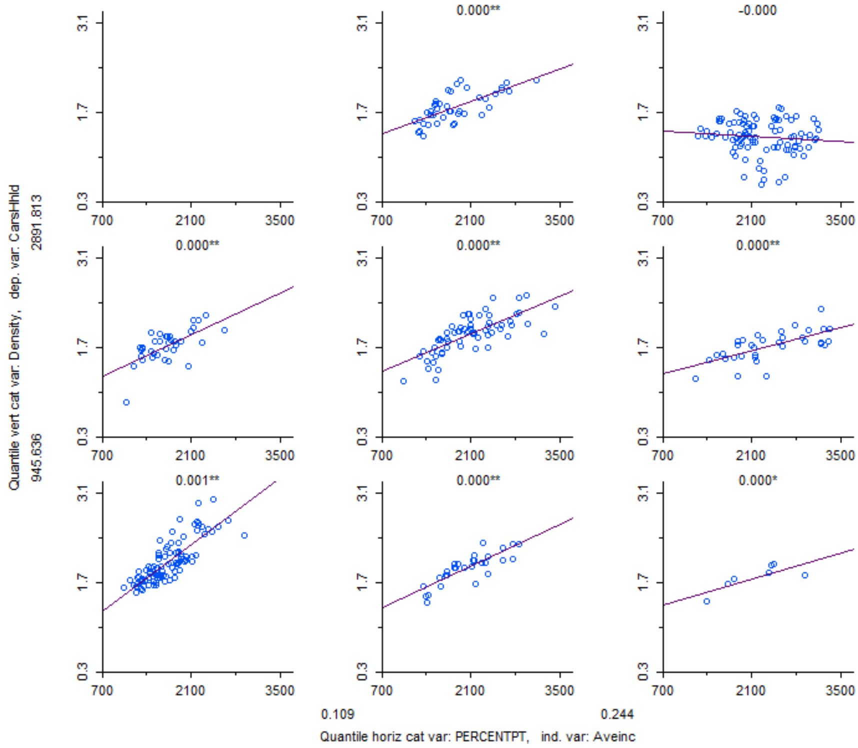

Figure 3: Cars Per Household and Average Income Sub-Setted by Density and per cent of Working Age Persons Using Public Transport. Source: the Author.

Figure 4 shows the location of these last mentioned SA2s. It can be seen in this figure that for the most part, they are clustered around Sydney harbour, in Sydney's Eastern or Inner Western Suburbs, or along major train lines (i.e. Hornsby is on the Northern line, Parramatta on the Western line, while Bankstown and Revesby are on the Bankstown and Illawarra lines respectively.

Table 2 provides details of the average values of selected variables within the region highlighted by figure 4 (under the heading 'Regional subset'), along with the average for the GMA. The data in this table indicates that the regions highlighted in figure 4 have a higher average weekly income ($2 485, compared to $1 999 for the GMA), while also having, on average, less cars per household (1.63) than the average for the GMA (1.75). Household size is almost at the GMA average at 2.74 persons per household (compared to 2.75 for the GMA), while 30 per cent of employed persons in these SA2s indicated that they used public transport to journey to work on census day, compared to an average for the GMA of over 20 per cent.

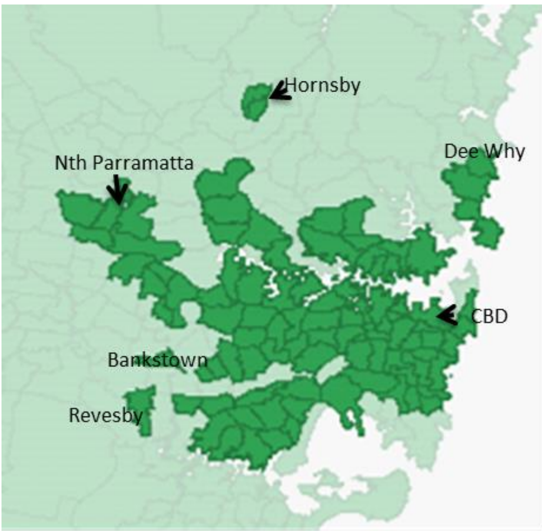

Figure 4. SA2s in Sydney Where There Appears to be a Negative Relationship Between Average Income and the Number of Cars Per Household. Source: the Author?

Table 2. Characteristics of Subregion.

	Unit	Regional Subset	GMA average
Cars_Hhld	No.	1.63	1.75
Hhld_size	No.	2.74	2.75
Pop_Density	Persons sqkm.	2 452.53	2 338.85
Ave_inc	$	2 485.38	1 998.68
P_Ptransport	%	0.30	0.20
P_NESB	%	0.15	0.14
P_Bachelor	%	0.40	0.26
P_Aged65plus	%	0.15	0.16

Source: the Author.

4. MODELLING SPATIAL VARIATION IN HOUSEHOLD CAR OWNERSHIP

This section commences with a global model of household car ownership within the Sydney GMA. In the current situation, global model refers to the assumption that the relationship between the dependent variable (no. of cars owned per household) and the explanatory variables, is invariant across the GMA. Initial modelling uses OLS and the results of this modelling are presented in the first panel of table 3. The variables in levels, i.e. *Cars_Hhld, Hhold_size, Pop_Density* and *Ave_inc* all enter this relationship in natural log form. In this formulation, the coefficient estimates are interpreted as elasticities.

The per cent travelling to work by public transport *(P_Ptransport)* is not included in the modelling here[1]. It is highly likely that decisions about car ownership and the mode of travel to work are determined simultaneously. In this situation incorporating *P_Ptransport* may result in a problem of endogeneity, which is known to produce biased and inconsistent parameter estimates in regression models. Instead of incorporating *P_Ptransport* directly, the next stage of modelling uses this variable to create spatial regimes, and work focusses on examining whether the explanatory variables vary systematically across the regimes.

The results presented in table 3 indicate that all variables are significant (low p - values), while the model diagnostics in the lower left panel suggest that the model explains a significant amount of the observed variation in the number of cars owned per household, i.e. the F - statistic is significant, while the R^2 is 0.88. These diagnostics also indicate that heteroscedasticity may be a problem, with the Breusch-Pagan (BP) test statistic highly significant.

In contrast, the tests for spatial autocorrelation in the panel below the OLS regression suggest this problem is present in the residuals of the OLS version of the model. The tests of spatial autocorrelation favour a spatial error specification (SEM). In this application, SEM is also favoured on theoretical grounds. This is because it is difficult to see a situation where

---

[1] An additional consideration is that *P_Ptransport* is only one of a number of measures of accessibility that could be used. The natural log of the average distance of commutes, or an index of job accessibility are others. All these variables attempt to measure the same thing, accessibility, and for the sake of simplicity, it was decided to use *P_Ptransport* in the current study. Additionally, a measure of service level was also considered, however this measure is still under development and does not cover the entire GMA.

one of the explanatory variables included in the model could affect the number of cars owned per household in a neighbouring region, i.e. it does not appear likely that there are spillover effects from spatially lagged explanatory variables. Likewise, there seems no theoretical grounds to justify spillover effects from a lagged dependent variable, i.e. a spatial autoregressive (SAR) model. For this reason, in this application, spatial autocorrelation is treated as a nuisance parameter, also making the SEM the most appropriate specification. The SEM specification used here also provides estimates robust to heteroscedaticity, which is likely to be a problem in the current application, as indicated by the BP test results of the OLS version of the model.

Table 3. Global Models of No. Cars Per Household Sydney GMA.

	OLS regression			*SEM (KP HET)*		
Variable	Coefficient	t-stat	P-value	Coefficient	z-stat	P-value
Intercept	-2.99	-11.70	0.00	-3.88	-15.37	0.00
P_Nesb	-0.62	-6.10	0.00	-0.51	-4.48	0.00
P_Bachelor	-0.84	-12.23	0.00	-1.06	-13.79	0.00
P_ Aged65plus	1.09	11.84	0.00	0.57	6.34	0.00
Ln(Pop_density)	-0.04	-11.95	0.00	-0.03	-10.98	0.00
Ln(Hhld_size)	1.02	23.85	0.00	0.72	10.24	0.00
Ln(Ave_inc)	0.39	10.54	0.00	0.55	13.62	0.00
lambda	-	-	-	0.81	18.38	0.00
Mean dep var	0.53	AIC	-839.30	S.D. dep var	0.40	
R^2	0.88	Adjusted R^2	0.88	Pseudo R^2	0.66	
SSRl	2.84	F-statistic	477.45	No. of iter	1	
Sigma^2	0.01	Prob(F-stat)	0.00	No. vars	7	
BP test	231.67	p-val	0.00	DF	396	
Tests for spatial autocorrelation	Robust LM (lag) = 8.8, p-val = 0.00 Robust LM (error) = 115.4, p-val = 0.00					

Source: the Author.

Results from the SEM version of the model are broadly in line with results from the OLS version of the model. The largest difference is in the magnitude of the estimated coefficient of the proportion of the population aged 65 or greater (P_aged65plus), with the OLS model returning a coefficient of 1.09, compared to an estimate of 0.57 in the SEM specification. The sensitivity of the number of cars owned per household to changes in household size (Ln(Hhld_size)) is also lower in the SEM version of the model, declining from 1.02 to 0.72. In contrast, moving from the OLS to SEM, sees the estimated coefficient of (the log of average weekly income) (Ln(Ave_inc)) increase from 0.39 to 0.55, indicating that in the SEM specification, household car numbers are more sensitive to income changes than in the OLS version of the model.

The global models presented in Table 3 assume that the relationship between household car ownership and the explanatory variables is constant across the entirety of the GMA. However, what the data presented in figures 3 and 4 suggest, is that there may be differences in this relationship, that is, that this relationship may vary, with estimated coefficients taking different values at different points in space. There are a number of ways to test this hypothesis, i.e. through the estimation of spatial Bayesian models (Finley *et al*., 2007), Geographically Weighted Regressors (Brunsdon *et al*., 1998), or through the creation of spatial regimes (Anselin, 1990, Lauridsen 1996, Flores and Rodriguez-Oreggia, 2014) and this latter approach is adopted in the current study.

Spatial Regimes

In empirical spatial research, it is often assumed that the relationships under consideration are stable over the spatial structure being considered in the analysis. In the current study, for example, the model results presented in table 3 assume that the estimated coefficients apply to all SA2s of the Sydney GMA. However, the data presented in figure 3 suggests that the relationship between the average household income of an SA2 and the number of cars owned per household, may vary systematically with changes in the proportion of workers who chose to commute by public transport.

The creation of spatial regimes provides a way of exploring this variation in the relationship between the number of cars owned per household and the explanatory variables of our model. Spatial regimes are a way of introducing and testing for spatial heterogeneity, which arises when structural changes related to location exist in the data. In situations where spatial regimes are present, they are characterised by differing parameter values or functional forms. Here, the assumption of a fixed relationship between dependent and independent variables that holds over the complete data set is formally investigated.

A formal approach to test the structural stability of the regression coefficients across spatial subsets is possible through the spatial Chow test (Anselin, 1990). A spatial switching regression, or spatial regimes model, applies spatial Chow tests to diagnose structural instability in parameters across regimes. A significant coefficient variable suggests a 'level' shift in the effect of the explanatory variables on the dependent variable across specific areas of study.

A standard regime model takes the form:

$$\left[\frac{y_i}{y_j}\right] = \left(\frac{X_i, 0}{0, X_j}\right) = \left[\frac{\beta_i}{\beta_j}\right] + \left[\frac{\varepsilon_i}{\varepsilon_j}\right] \tag{1}$$

where *i* and *j* index discrete spatial subsets or regimes of the data and a test of the null hypothesis consists of $\beta_i = \beta_j$, where the β are estimated in the above equation. The standard Chow test distributed as an *F* with (*K*, *N*-*2K*) degrees of freedom is given by:

$$C = \left[\left(\varepsilon_r^{'}\varepsilon_r - \varepsilon_u^{'}\varepsilon_u\right)/\,K\right]\!\left[\left(\varepsilon_u^{'}\varepsilon_u\right)/\left(N-2K\right)\right] \sim F_{(k,N-2K)} \tag{2}$$

where ε_R and ε_U are the OLS residuals from a restricted model and from an unrestricted model, respectively; *N* is the number of observations and *K* is the number of regressors. However, when the error terms are spatially autocorrelated, the above expression is no longer valid. A corrected version of the test is referred to as a spatial Chow test (see Anselin 1998, 1990) is instead, given by:

$$c_s = \left\lfloor \varepsilon_R^{'}\left(I-\theta W\right)^{'}\left(I-\theta W\right)\varepsilon_R^{'}\varepsilon_U^{'}\left(I-\theta W\right)^{'}\left(I-\theta W\right)\varepsilon_U \right\rfloor /\, \sigma^2 \chi_K^2 \tag{3}$$

Where θ represents the ML estimate for the spatial parameter and σ2 the estimate for the error variance, while I is an identity matrix of dimension n x n, where n is the number of observations, or regions).

There are a number of ways to define regimes, for example, Lauridsen (1996) adopts three approaches to defining regimes, comprising regimes based on natural (political or administrative) boundaries, regimes based on univariate sorting (regimes defined by breaks in a single variable), and regimes based on multivariate sorting (regimes defined by breaks in a composite variable, which is itself created from a number of variables). In contrast, Flores and Rodreguez-Oreggia (2014) use exploratory spatial data analysis, in particular, local indicators of spatial association to define regimes. In the current work, the regimes are defined by the per cent of workers of SA2s who indicated that they journeyed to work by public transport on the day of the 2016 census (equivalent to the univariate sorting approach of Lauridsen 1996). Given that there are just over 400 SA2s (and thus observations) for modelling, initial experimentation looked at the creation of either 3 or 4 spatial regimes to maintain sufficient observations for model estimation.

The software used in this analysis (Geoda), provides an algorithm to identify natural breaks[2] in the data, and these natural breaks are here used to define the regimes. Both regimes based on 3 or 4 natural breaks resulted in the creation of a number of SA2s which were isolated, and after some experimentation it was realised there would be far less problems with a three regime specification which required a more limited reclassification of SA2s to create contiguous regimes and avoid the problem of islands among the regimes (i.e. where say an SA2 assigned to regime 1 is surrounded by SA2s defined as regime 3). The resulting regimes, after this respecification are shown in figure 5.

The regimes specified in figure 5 are closely related to the data provided in the subsetted scatter diagrams of figure 3. In figure 3, the subsets presented along the horizontal axis are created by dividing P_Ptransport (the per cent reporting that the mode of journey to work was public transport) into three equally sized groups. The three regimes used in the modelling are thus closely related to the summation of the SA2s captured within each of the three columns of figure 3. For example, there are about 135 SA2s in the three scatter diagrams that makes up a column in figure 3. Regime 3 of figure 5 below is a subset of 116 members of the SA2s with the highest proportion of public transport users (i.e. the final column of Figure 3).

---

[2] To find natural breaks, a nonlinear algorithm developed by Jenks (1977) which groups observations such that the within-group homogeneity is maximized is used. In essence, this is a one dimension clustering algorithm which determines the break points that yield groups with the largest internal similarity.

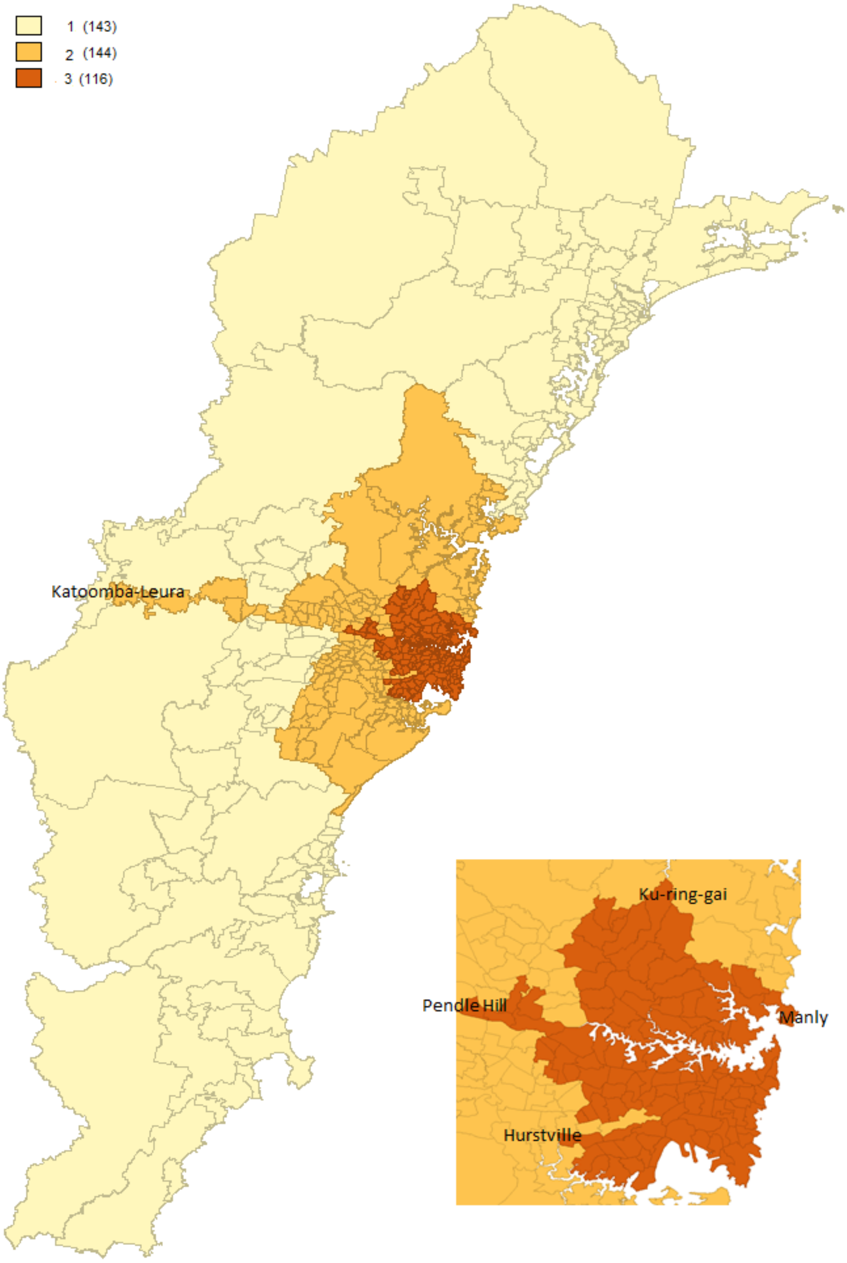

Figure 5. Spatial Regimes Defined by Per Cent Travelling to Work by Public Transport. Source: the Author.

The regimes run from 1 (SA2s with residents who proportionally were the lowest users of public transport) to regime 3, the highest users of public transport. Table 4 provides details of the differences in the magnitude of the variables used in this analysis across these three regimes. The data in this table indicates that while households in regime 1 have the highest number of cars per household (an average of 1.9), they also have the lowest average weekly household income ($1 732). In contrast, in regime 3, which is entirely composed of SA2s within metropolitan Sydney, average weekly income is the highest ($2 322), while the number of cars owned per household is lowest (1.4) in this regime.

Table 4. Summary of Regional Differences by Select Variables.

Variable	Unit	Regime 1	Regime 2	Regime 3
Cars_Hhld	No.	1.94	1.86	1.39
Hhld_size	No.	2.64	2.98	2.60
Pop_density	No.	717.22	2 096.32	4 638.99
Ave inc	No.	1 731.89	2 003.13	2 322.03
P_Ptransport	%	0.05	0.19	0.37
P_Nesb	%	0.12	0.15	0.16
P_Bachelor	%	0.17	0.24	0.41
P_aged65	%	0.19	0.15	0.16

Source: the Author.

The SEM versions of the model incorporating the three regimes are presented in table 5 and as with the global SEM model, the model has been estimated with standard errors robust to the presence of heteroscedasticity. An inspection of the results indicates that the estimated coefficients, and their significance, varies across the three regimes. For example, the results indicate that □, the coefficient of the spatial error term, is not significant in regime 1. The residual tests for spatial autocorrelation from the OLS model are also included in the lower panel of this table and for regime 1, also suggest that spatial autocorrelation is not an issue in this regime.

The spatial chow test presented in the final three columns of the table, indicates that the proportion of the population that is non-English speaking *(P_Nesb)*, the proportion with a Bachelor or higher level qualification *(P_Bachelor)* and the natural log of population density *(Ln(Pop_density))* have the same effect across all spatial regimes included in the model, i.e.

that the impact of these variables does not vary across the study area. In contrast, there are significant differences in the estimated coefficients of the proportion of the SA2s population aged 65 or older (*P_age65plus*), ranging from 0.35 and 0.31 in regimes 1 and 2 respectively, to 1.45 in regime 3, where access to public transport is highest. The effect of household size (*Ln(Hhld_size)*) also varies across the 3 regimes, with the effect similar in regimes 1 and 2 (at 0.70 and 0.61 respectively), while markedly higher in regime 3, where the estimated coefficient is 0.99.

The coefficient estimates of *Ln(ave_inc)* also vary across the three regimes. In regime 3, where average incomes are highest at $2 322 per week, a 1 per cent increase in income is found to increase the number of cars owned per household by 0.32 per cent. This regime largely comprises the SA2s in the three scatter diagrams in the final columns of figure 3, including those SA2s in the inner city, where the scatter diagram indicates a marginal negative relationship between income and the number of cars owned per household (as highlighted in figure 4). In contrast, the responsiveness of household car ownership to income changes is highest in regime 2, which comprises the outer Sydney region and SA2s along the Blue Mountains line. In this regime, a 1 per cent increase in weekly household income is estimated to result in a 0.54 per cent increase in the number of cars owned per household. Census data indicates that commuters in this region are less reliant on public transport than are commuters in regime 3 (i.e. 19 per cent of residents of regime 2 indicated that they used public transport to travel to work, compared to 37 per cent in regime 3). This lower reliance on public transport appears to create greater reliance on car ownership. In this situation, income changes see relatively larger changes in car ownership even though average incomes remain lower than in regime 3. Finally, in regime 1, the periphery of the GMA, where incomes are lowest (i.e. $1 732), the modelling indicates that a 1 per cent increase in household incomes leads to a 0.38 per cent increase in car ownership.

Table 5. SEM Model with Three Spatial Regimes, Sydney GMA.

	Regime 1		*Regime 2*		*Regime 3*		*Regime diagnostic- Spatial Chow test*		
Variable	Coefficient	Probability	Coefficient	Probability	Coefficient	Probability	df	Value	Prob
Intercept	-2.66	0.00	-3.79	0.00	-2.24	0.00	2	8.70	0.01
P_Nesb	-0.24	0.03	-0.24	0.05	-0.93	0.02	2	2.97	0.23
P_Bachelor	-0.48	0.00	-0.67	0.00	-0.67	0.00	2	2.41	0.30
P_aged65plus	0.35	0.01	0.31	0.05	1.45	0.00	2	6.26	0.04
Ln*(Pop_density)*	-0.03	0.00	-0.03	0.00	-0.08	0.02	2	2.95	0.23
Ln*(Hhld_size)*	0.70	0.00	0.61	0.00	0.99	0.00	2	11.31	0.00
Ln*(ave_inc)*	0.38	0.00	0.54	0.00	0.32	0.00	2	7.90	0.02
Lambda (λ)	0.13	0.19	0.44	0.00	0.63	0.00	2	14.65	0.00
Global test							16	220.09	0.00
	Mean dep var = 0.65	Df = 136	Mean dep var = 0.61	Df = 137	Mean dep var = 0.29	Df = 109			
	S.D. dep var =0.16	Pseudo R^2 = 0.93	S.D. dep var =0.15	Pseudo R^2 = 0.91	S.D. dep var =0.25	Pseudo R^2 = 0.85			
	No. obs = 143		No. obs = 144		No. obs = 116				
	LM lag = (2.21,0.14)		LM lag = (3.53,0.06)		LM lag = (22.16,0.00)				
	Robust LM = (2.41,0.12)		Robust LM = (31.20,0.00)		Robust LM = (1.46,0.23)				
	LM error = (0.04, 0.85)		LM error = (32.84,0.00)		LM error = (56.44,0.00)				
	Rob LMerr = (0.24,0.63)		Rob LMerr = (34.7,0.00)		Rob LMerr = (35.80,0.00)				

Source: the Author.

5. CONCLUSION/DISCUSSION

This paper has explored the factors impacting on the number of cars owned by households. A range of variables capturing SA2 demographic composition, along with regional and transport characteristics were included in the modelling. The findings are not unlike those presented in previous research.

The current work is most like Clarke (2007; 2009), who also used aggregate zonal data. Unlike Clark, whose focus is the relationship between household car ownership and income, and so estimates a bivariate model, this study has incorporated several variables to capture the effect of demographic and regional characteristics on household car ownership. Because there is a risk of simultaneity between transport characteristics and household car ownership, transport characteristics have not entered the model directly. Instead, transport characteristics, in the form of the proportion of each SA2's workers who commute to work using public transport, is used to define spatial regimes. The novelty of the approach adopted here is that rather than specifying these regimes directly, they have been created by letting the data speak for itself. Natural breaks in the variable *P_Ptransport* (the proportion of people using public transport to commute to work) were identified and used to define the regimes.

Like Clarke (2007; 2009), the modelling here has found that the impact of average household income on the number of cars owned per household varies by regime. Using data from England and Wales, Clark (2007) derived an estimate of the number of cars owned per household with respect to income of 0.92 in a global version of his model. This is far above the estimates derived in the analysis presented here (0.55 in the global model incorporating a spatial error term). This difference may not be surprising. Australian cities tend to have a lower population density than many of their European counterparts. This lower level of population density, and corresponding greater sprawl, means that public transport accessibility is lower for much of the population. This lower accessibility will increase resident's reliance on privately owned motor vehicles. As a result, car ownership is higher per household and less sensitive to changes in household income.

While studies like Clarke (2007; 2009) and the work presented here cannot hope to provide the level of detail of studies using unit record data from specialised household travel surveys, the focus on identifying spatial regimes, and exploring the way space impacts the relationship between the explanatory variables and the number of cars owned per household, has revealed relationships which may be missed when using unit record files.

In the model including spatial regimes, the sensitivity of household car ownership to changes in average weekly income is highest in regime 2 (with an estimated elasticity of 0.54). This regime broadly aligns to the periphery of metropolitan Sydney. In contrast, the sensitivity of household car ownership to income is lowest in regime 3, a regime which includes the urban core of metropolitan Sydney and where average incomes are highest. Public transport access is also higher in this regime, with more residents indicating they use public transport to journey to work. In contrast, persons on the urban fringe and in the regions beyond (i.e. the Hunter and Illawarra regions) tend to have lower incomes while also having less access to public transport.

The relationship between average household income and the number of cars owned per household, and the fact that this relationship varies across the regimes may have policy implications. Options like congestion levies or environmental taxes and fuel excise, levied directly on the car or fuel will, according to the modelling undertaken here, likely have a larger impact on lower income households on Sydney's urban fringe (regime 2 in this analysis). Additionally, it is people living in these areas that Rosier and Macdonald (2011) identify as being much more likely to experience transport difficulties.

Troy (2004) notes that, in regards to structure, Australian cities like Sydney have a similar structure to US and Canadian cities. Population density is relatively low and structure is dominated by sprawl, while average commutes are longer than in their European counterparts. This outcome is attributed to the take up of private motor vehicles as a form of mass transit. This use of motor vehicles has contributed greatly to urban sprawl and Troy (2004) argues that, for Australia at least, this phenomenon has had a long lasting, if not permanent impact on the structure of Australian cities. This impact includes low population densities, long commutes and difficulty in providing cost effective public transport.

To reduce the reliance on automobiles, Soltani and Somenahalli (2012) conclude that policies should be directed at reducing the reliance on private vehicle travelling. Such policies should focus on creating more accessible land use patterns and more liveable communities, while at the same time reducing public transport costs. These authors note that some aspects of urban structure are important in influencing car ownership. With regard to density, Soltani and Somenahalli (2012) note that early studies concluded that the low density areas generate private motor vehicle dependence due to a combination of factors including greater distances to travel and little option to walk or use public transport. Similarly, Pushkarev and Zupan (1982) in a study of American cities suggests that the higher density is associated with the lower ownership and the lower use of private motor vehicles, while households living in high density areas own fewer cars than those in the low density areas, a finding supported by the work here.

Soltani and Somenahalli (2012) also note that land use diversity is also important. Mixed land uses have been found to provide more transport options. A more accessible land use pattern means that less mobility (physical travel) is needed to reach goods, services and activities. The clustering of different land uses, such as shopping, offices and retailing has been found to encourage trips done by walking and public transportation. Cervero (1996) concluded that such development is associated with a decline in car ownership levels in large American cities.

Based on evidence like this, government land use policies in Australia and overseas have focused their attention on the density, mix and location of urban development to public transport. Soltani and Somenahalli (2012) note that planning policy based on access to public transport has been embraced in the UK to minimise additional car travel, reduce trip lengths and encourage use of other, more sustainable means of travel. In the UK, policy has moved towards integrating public transport accessibility levels with the social needs of the population, reducing their reliance on private motor vehicles and providing more accessible opportunities to socially excluded populations. The same authors note that in the Netherlands, policy is focussed on locating each business in a location with an accessibility profile in accordance with its mobility characteristics. The aim here is to reduce the number of trips made by car.

There is a great deal of uncertainty about future levels of car ownership, stemming from changes in technology and environmental concerns. Future private transport may be driven by battery, or hydrogen fuel cells. There may be a move from private ownership to car sharing. This aside, in Sydney, policy responses like Future Transport (TfNSW, 2016) which aims to make public transport more accessible across the GMA and the Greater Sydney Commissions (GSC) strategy 'Metropolis of three cities', (see GSC 2016) may well act to reduce the number of cars owned per household, and subsequently, the proportion of trips made by private motor vehicles on the urban fringe of Sydney. The Metropolis of three cities' strategy of GSC aims at encouraging both residential and employment developments away from the central business district, focusing instead on Parramatta, and beyond, in a third city centered around the second airport at Badgerys Creek. This strategy, if successful, will see population densities in the urban fringe increase, providing a greater return to increased investment in public transport, while also reducing average commute times and the reliance on private motor vehicles.

REFERENCES

Anowar, S., Eluru, N. and Miranda-Moreno, L. (2014). Alternative Modelling Approaches Used for Examining Automobile Ownership: A Comprehensive Review. *Transport Reviews*, 34(4), pp. 441-473.

Anselin, L. (1990). Dependence and Structural Instability in Applied Regression Analysis. *Journal of Regional Science*, 30, pp. 185-207.

Brunsdon, C, Fotheringham, S. and Charlton, M. (1998). Geographical Weighted Regression—Modelling Spatial Non-Stationarity. *The Statistician*, 47(3), pp. 431-443.

Cervero, R., (1996). Mixed Land-Uses and Commuting: Evidence from the American Housing Survey. *Transportation Research A*, 30(5), pp. 361-377.

Clark (2009). Spatial Modelling of Car Ownership Data: A Case Study from the United Kingdom. *Applied Spatial Analysis*, 3, pp. 45-65.

Clark, S. D. (2007). Estimating Local Car Ownership Models. *Journal of Transport Geography*, 15(3), pp. 184-197.

de Jong, G., Fox, J., Pieters, M., Vonk, L. and Daly, A. (2004). Comparison of Car Ownership Models. *Transport Reviews*, 24(4), pp. 379-408.

Finley, A. O., Banerjee, S. and Carlin, B. P. (2007). spBayes: A Program for Multivariate Point Referenced Spatial Modeling. *Journal of Statistical Software*, 19(4).

Flamm, B. (2009). The Impacts of Environmental Knowledge and Attitudes on Vehicle Ownership and Use. *Transportation Research Part D: Transport and Environment*, 14(4), 2009, pp. 272–279. https://doi.org/10.1016/j.trd.2009.02.003.

Flores, M. and Rodriguez-Oreggia, E. (2014). Spillover Effects of Homicides Across Mexican Municipalities: A Spatial Regime Model Approach. Online version accessed from Munich Personal RePEc Archive, https://mpra.ub.uni-muenchen.de/56507/1/MPRA_paper_56507.pdf

Goodwin, P., Dargay, J. and Hanly, M. (2004). Elasticities of Road Traffic and Fuel Consumption with Respect to Price and Income: A Review. *Transport Reviews*, 24(3), pp. 275-292.

Graham, J. and Glaister, S. (2004). Road Traffic Demand Elasticity Estimates: A Review. *Transport Reviews*, 24,(3), 261-274.

Greater Sydney Comission (GSC) (2016). Greater Sydney Commission, Metropolis of Three Cities. Website accessed from https://www.greater.sydney/metropolis-of-three-cities.

Jenks, G. (1977). Optimal Data Classification for Choropleth Maps. Kansas City, KS: Department of Geography, University of Kansas, Occasional Paper 2. Cited in Anselin, L. (2014). GeoDa: An Introduction to Spatial Data Analysis, doi: https://spatial.uchicago.edu/geoda

Lauridsen J. (1996). Dependence and Regimes in Applied Spatial Regression Analysis. *Paper presented at the 36th European Regional Science Association Congress, Zurich*, https://pdfs.semanticscholar.org/dd89/284b8814e57b6e0e299b8f86b3435ec93bb8.pdf.

Maltha, Y., Kroesen, M., Van Wee, B. and van Daalen, E. (2017). Changing Influence of Factors Explaining Car Ownership Levels in the Netherlands. *Transportation Research Record: Journal of Transportation Research Board*, 2666, pp. 103-111.

Needleman, L. (1978). On the Approximation of the Gini Coefficient of Concentration. *The Manchester School*, 46, pp. 105-122.

O'Sullivan, M. (2014). Two or Three-Car Families New Norm in Sydney Despite Public Transport Increase. *Sydney Morning Herald*, 14 August 2016, doi: https://www.smh.com.au/national/nsw/two-or-threecar-families-new-norm-in-sydney-despite-public-transport-increase-20160805-gqm4qw.html.

Potoglou, D. and P. S. Kanaroglou. (2008). Modelling Car Ownership in Urban Areas: A Case Study of Hamilton, Canada. *Journal of Transport Geography*, 16(1), pp. 42-54. https://doi.org/10.1016/j.jtrangeo.2007.01.006.

Pushkarve, B. and Zupan, J. (1982). Where Transit Works: Urban Densities for Public Transportation. In H. S. Levinson and R. A. Weant (Eds) *Urban Transportation: Perspectives and Prospects*, Eno Foundation for Transportation, Westport, Connecticut.

Rosier, K. and Macdonald, M. (2011). The Relationship Between Transport and Disadvantage in Australia. *CAFCA Resource Sheet*. doi: mhttps://aifs.gov.au/cfca/publications/ Relationship -between -transport-and-disadvantage-austr

Soltani, A. and Somenahalli, S. (2012). *Household Vehicle Ownership: Does Urban Structure Matter?* Proceedings of the 28th Australasian Transport Research Forum, Sydney, Australia.

Transport for New South Wales (TfNSW) (2016). Transport for New South Wales, Future Transport. https://future.transport.nsw.gov.au/designing-future/program

Troy, P. (2004). The Structure and Form of the Australian City: Prospects for Improved Urban Planning. *Urban Policy Program Issues Paper No. 1*., Griffith University, Brisbane.