Faculty of Science and Engineering

Department of Spatial Sciences

Assessing the impacts of agglomeration on train ridership under the spatial economic transport interaction framework

Siti Nurlaela

This thesis is presented for the degree of

Doctor of Philosophy

of

Curtin University

June 2016

This page intentionally left blank.

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: ………………………………………………

Date: ………30 June 2016………………...

This page intentionally left blank.

My Lord, cause me to enter a sound entrance and to exit a sound exit and grant me from Your-self a supporting authority (Translation of Al-Israa – The night Journey, 17:80).

This page intentionally left blank.

ABSTRACT

Transport-induced agglomeration is defined as the concentration of jobs or employed residents arising through transport infrastructure, and is measured in both density and proximity. While density is an important parameter in land use and proximity is a result of the transport provision, the origin framework for land use and transport integration has not yet incorporated the simultaneous effect of both density and proximity, known as "effective density". Effective density is one measurement of agglomeration that has been developed in the context of the wider impact assessment of transportation infrastructure. The land use-transport interaction (LUTI) framework is one of the major theoretical frameworks used in understanding the integration of transport and land use. However, there remains a lack of a spatial-economic dimension in LUTI in the literature.

The aim of this thesis is to model train ridership under the influence of public transport induced agglomeration. Thus, the spatial economic dimension is added to the LUTI framework, allowing research that incorporates spatial organisation and the economy of an area into transport and land use planning in a holistic manner. This thesis named this framework an extension of the LUTI framework that is part of what is commonly defined as a Spatial Economic Transport Interaction (SETI.). Specifically, the density concept in the LUTI framework was extended to an effective density, thus simultaneously represents scale and proximity.

Four reseach objectives were set up to reach research aim as follows:1) to investigate the link between transportation investment and agglomeration by providing evidence that support the existence of public transport induced agglomeration; (2) to understand the influence of public transport induced agglomeration on train ridership in terms of both effective job density and effective employed resident density; (3) to investigate the geographical extent of agglomeration relative to the distribution of stations that assumed as being the focal point (source) of agglomeration; and (4) to investigate the implications of agglomeration on job-housing balance from the wages/land rents and travel costs trade-off perspective and its further implication on train ridership.

The empirical evidences supported the existence of public transport induced agglomeration in the study area: the strong association of changes in travel time before-and-after the Perth-Mandurah railway extension with the changes in effective density and the changes in train ridership; the high rate of changes in train ridership and effective density for suburbs located adjacent to the Perth-Mandurah line, relative to other lines; the clustering of high effective density near train stations; and the exponential distance decay of effective density, whereby a distance increase of 1 km from a train station correlates to a decrease in effective job density of 11.2%, and in employed resident density of 9.5%.

The SETI model, utilising effective density variables, improved train ridership prediction capability by 2-5% compared to the LUTI model, while the level of errors in SETI was 64% of the LUTI model by validation or back testing analysis. Adding effective density variables has significantly improved the LUTI model, as shown by the partial F-test. The effective job density with the interaction term is statistically significant in both train trip production and train trip attraction, and the effective employed resident density is only statistically significant in train trip attraction.

Distance (station accessibility) has a negative influence on the relationship between effective density and train ridership attraction, but a positive one for train ridership production.

The geographic extents of the positive impact of effective job density on train ridership were found to be up to a distance of 2.4 - 4.1 km from the train station, and 5.32 – 8.84 km for the impacts of effective employed resident density. The retail sector has a smaller geographical extent of influence on train ridership than the construction and manufacturing sector, but at lower levels of multiplicative effects. Higher influences of effective density emerged when higher values of effective density were confined to smaller areas and this may be found in the construction sector, relative to the manufacturing and retail sectors. Job-housing balance analysis showed a strong effect of effective job density on train ridership attraction in job-rich suburbs.

The derived results may be useful in understanding the land use and socioeconomic impacts of transit and travel behaviour of railway users, in managing station facilities, and in developing land use policies and transit orientation development (TOD) strategies.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful.

I am very grateful and thank Allah SWT for His Kindness and His Mercy, He who made my long journey short and my difficult process easy. Along the way, there were many people who supported me, encouraged me, prayed for me and were patient with me, and who provided great help in the completion of this thesis.

First of all I would like to express the best gratitude to my supervisor, Associate Professor Jianhong (Cecilia) Xia, for her encouragement, guidance, immense support and supervision throughout this PhD study. I am indebted to her not only for scientific supervision, but also for her understanding, help, and friendship. I am also grateful to my co- supervisors Prof. Ross Taplin (the Curtin Bussiness School or CBS) and Prof. Bert Veenendaal particularly for their support, advices, and immense help especially in build my understanding on statistical modelling and spatial analysis issues. I also thank Assoc. Prof. Greg Costello from the Department of Economics and Property (the CBS) for providing access to the Landgate property value transaction dataset. My grateful also goes to Department of Planning of Western Australia, especially to Mr Gary McCarney for providing travel time data matrices and reviewing policy analysis chapter; The Western Australian Land Information Authority (Landgate), Department of Agriculture and Food of Western Australia, Department of Transport of Western Australia, and the Australian Bureau of Statistics for abundant data on land use, transportation, and socio-demographic/economic of Perth metropolitan region. I would like to thank all staff at the Department of Spatial Sciences, especially Ms Caroline Rockliff and Ms Lori Patterson, to all Curtin University staff especially those in the Graduate Research School (GRS), Curtin Library, Curtin Learning Centre, and the ISSU.

I acknowledge the Directorate General of Higher Education (DGHE) of Ministry of National Education of Indonesia for the funding of my PhD scholarship. I would like to thank also the Rector of Institute of Sepuluh Nopember, the Dean of Faculty of Civil Engineering and Planning, the Head of Department of Urban and Regional Planning, and all collegues and staffs of the Institute of Sepuluh Nopember especially in my department and the Laboratorium Perencanaan Wilayah for their patient, understanding, and great supports during my study leave.

I would like to thank Dr. Andrew Kirker for editing and proof read my thesis, Ferry Irfan Nurrahman and Al Huda for making some maps, Dinesh Tuladhar and Ting Lin for cooperation and team-work, and other PhD collegues in the Department of Spatial Science 207.310A.

Finally, I would like to thank my husband Haris Sri Purwoko and my son Faiz Farrosian Karim as "a young member" of the 207.310A workstation for their immense loves, supports and cherish, my son Alwan Fayyadh and Adnan Farrosi for their patient, loves, and prays; and Mbah Kakung Panidi's big family for their supports. I am truly indebted to my parents that had been passed away, M.A Sonjaya and Siti Hindasyah, my love and prays for them always, and my brothers and sisters especially Dewi Kristanti, Hidir Kurniawan, Imas Herlina, Aep Saefudin, Neni Nuraeni, and Nunung Nurhayati for keeping me strong and steadfast in this process. I thank to all friends and families I cannot named them one by one to have come across in a good relationship in my life and during my PhD journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ACRONYMS

ABS	Australian Bureau of Statistics
AHURI	Australian Housing and Urban Research Institute
ANZSIC	The Australian and New Zealand Standard Industrial Classification
BART	Bay Area Rapid Transit
BTRE	Bureau of Transport and Regional Economics
CBA	Cost-Benefit Analysis
CBD	Central Business District
CPI	Consumer Price Index
CSR	Complete Spatial Randomness
DC	Development Control
DOP	Department of Planning
DOT	Department of Transport
DZN	Destination Zone
ESRI	the Environmental Systems Research Institute
GDP	Gross Domestic Product
GHG	Green-house gases
GIS	Geographical Information System
INCP	Total Personal Income (Weekly)
ITS	Intelligent Transport System
LDV	Limited Dependent Variable
LFSP	Labour force status by persons
LNs	Liveable Neighbourhoods
LQ	Location Quotient
LU-T	Land use to Transport (Interaction)
LUTI	Land Use and Transport Interaction
MTWP	Method of travel to work by persons
NMT	Non-Motorized Transport
Pkm	Passenger kilometre
PMR	Perth Metropolitan Region
POR	Place of Residence
POW	Place of Work
SA2	Statistical Area Level 2
SETI	Spatial Economics Transport Interaction

SPSS	Statistical Product and Service Solutions
SSC	State Suburb Code
STEM	The Strategic Transport Evaluation Model
TCRP	Transit Cooperative Research Program
TDM	Travel Demand Management
TFP	Total Factor Productivity
The UK	The United Kingdom
The US	The United States
T-LU	Transport to Land Use (Interaction)
TOD	Transport Oriented Development
TSM	Transport System Management
VIF	Variance Inflation Factor
VMT	Vehicle Miles Travelled
WA	Western Australia
WEBs	Wider Economic Benefits

CHAPTER 1. INTRODUCTION

Although land use has influenced human life in many aspects (ecology, economic development, health sector and social equity), the influence of urban form on transport and mobility has been the most widely-explored from an empirical and theoretical point of view (Williams, 2005). The concept of integrated land use and transport (LUTI) had been studied for a long history (Cao, Mokhtarian, & Handy, 2009; Chen, Gong, & Paaswell, 2008). Many attempts at creating sustainable urban form and transportation have been presented as far back as 10,000 years ago since the development of ancient cities, such as Abu Hureya, Mureybat, Jericho (Soja, as cited in Hickman & Banister, 2005) The development of this theory within the urban planning context had been growing over the last 120 years (Hickman & Banister, 2005).

The traditional “predict and provide” approach to solving urban transportation problems is no longer an option: simply forecasting demand for travel and then meeting this demand with transport infrastructure solutions is no longer considered to be viable (Handy, 2008). Solving urban transport problem is now considered to require sustainable mobility solutions (Bertolini, Clercq, & Straatemeier, 2008). Moving toward a new way in which integration of land use and transport in cities are planned and developed is thought to be crucial (Tennøy, 2010). In this sense, it has been suggested that "it is necessary to analyse these two systems (land use and transport) in an integrated approach as they are very dependent on one another…" (Jonsson, 2008). Therefore, the impacts of transportation systems in terms of transportation infrastructure need to be evaluated not only on their aspect of land use but also on travel behaviour (Brons, Givoni, & Rietveld, 2009; Van Acker, Witlox, & Van Wee, 2007).

Further, other factors such as the spatial economic dimension may come into play when transportation system-impacted land use and travel behaviour has been examined in a holistic manner. The definition of the term 'spatial economic' is derived from Fujita and Thisse (2002) and Russo and Musolino (2012). Interactions between agents (firms or households) that involve commodities (good or services), including their locations are defined as "economy" (Arrow, K et al. as cited in Fujita and Thiesse, 2002, p. 20). The term 'spatial' refers to locations, landscapes, geographical scales, where the economic activity has occurred. The transport system creates spatial organisation by affecting activity locations, production levels, and trade patterns (Russo & Musolino, 2012). The spatial economic perspective provides an economic reason behind the concentration of production (firms) and consumption (households). Thus, there is a relationship between spatial attributes and economic theory, i.e. the distribution of economic activities (the concentration) is formed not by natural causes (such as rivers or ports) but due to the economic mechanism involving trade-offs between various forms of increasing returns and mobility costs (Fujita & Thisse, 2002). Furthermore, Fujita and Thiesse explained that the spatial configuration of economic activities emerged as an outcome of a process of two opposing forces, namely the centripetal forces (agglomeration or the pull factor) and the centrifugal forces (dispersion or the push factor). The spatial configurations of economic activity that are observed, then, are the results of the balance between these two forces that push and pull consumers and firms (Fujita & Thiesse, 2002, p. 5). This research, in line with the theoretical concept of agglomeration in Fujita and Thiesse, however, adapted a specific kind of agglomeration force, i.e. public transport induced agglomeration. That is, transportation infrastructure facilitates the clustering of employment. There are commuting cost reductions that form the grounds for firms' and households' location behaviour (Alpkokin, Cheung, Black, & Hayashi, 2008; Graham, 2007).

Nevertheless, the LUTI framework often treats spatial economic factors as independent from the influence of other factors in a land use-transportation interaction system (Southworth, 2001; Taylor, Miller, Iseki, & Fink, 2009; Wardman, 2006; Wardman, Lythgoe, & Whelan, 2007). In fact, spatial economic factors affect transport in a way that influences travel demand patterns, while the urban structure framed by the transportation system affects activity locations and production levels (Russo & Musolino, 2012). Russo & Musolino (2012) pointed out that the framework of Spatial Economic Transportation Interaction (SETI) is based on the concept that LUTI and the economic dimension are inherently part of SETI and are integrated in a two-way interaction between both.

On the other hand, most advances in the SETI field have been directed towards justifying substantial investment in public transportation spending. Transportation infrastructure impact assessment studies have been gaining popularity in this area, including the recent development of wider economic benefits (WEBs) assessments (Hensher, Truong, Mulley, & Ellison, 2012; T. Lin & Truong, 2012). Unfortunately, most of these studies refer to the economic impact of the transportation system on urban productivity as the focus or outcomes, and not to travel behaviour.

This thesis attempts to use the extension of the LUTI framework, referred to as the SETI framework, to derive factors affecting one form of transportation, train ridership. The spatial economic dimension, or specifically, in terms of public transport, "induced agglomeration" was found to be one of the predictor variables that influence train ridership, among others. This agglomeration is an outcome of improved accessibility gained from public transport infrastructure development. Under the SETI framework, the interrelationship between the changes in the transportation-system-influenced agglomeration and train ridership were examined.

1. 1 BACKGROUND

Cities all over the world have experienced common problems in urban transportation. Climate change, traffic congestion, air pollution, and traffic accidents have been major issues of concern for the past few decades. Banister (2008) pointed out a transport-led future has consisted of an increase in car dependence due to the increased suburbanisation; thus non-motorized transportation and local public transport have become less attractive due to the greater use of cars. Mohan (2008) examined most big cities have not been able to reduce their car use to the level less than 10% of mode shares or lower, despite the extensive public transport investments that have been put in place. Almost all major cities continue to face severe congestion on arterial roads. For example, car speeds during peak hours reduced to only 10-15 km/h in major cities like Paris, London, Tokyo, Tehran, Jakarta or Mexico City (Mohan, 2008). There was a contradictory situation when public transport investment was not followed by a substatial gain by a modal switch into public transport. This dilemma has created a challenge for sustainable urban transport development of a 21st Century goal. Providing an effective urban public transport development strategy that can match the demand for accessibility and mobility is still needed.

Australia's urban transportation has experienced the same challenges for sustainable transportation. Public transport share is only 10 percent out of the total metropolitan passenger kilometer (pkm). Currently, the share of private vehicles accounts for about 86 percent of the total passenger task over all Australian cities. Market shares of both modes have remained reasonably constant since the 1980s. Nevertheless, there has been some improvement in terms of public transport supply, with an increase from 8.9% (2005) to about 10.5% (2010) out of the total pkm in Australia (Bureau of Transport and Regional Economics [BTRE], 2007).

The data indicates a major challenge for the delivery of sustainable urban transport for Western Australia (WA), renowned for its low density housing and high dependence on the car (Curtis, 2012). The Australian Bureau of Statistics reported total kilometres travelled by motor vehicles in Western Australia (WA) increased by 23% in within period 2004 to 2010 - far above Australia's statistics which was a 14% increase within the same period (Australian Bureau of Statistics (ABS), 2010). In addition, the number of vehicles registered in WA increased by 27% compared to 19% in Australia between 2004 – 2010 (Australian Bureau of Statistics (ABS), 2010). On the other hand, Perth continues to focus on investment in public transport development, such as the first revitalization of suburban rail in the 1980s (Curtis, 2008) and the recent project being the extension of the Perth to Mandurah railway line in 2008.

The extension of the new Perth – Mandurah railway line in the Perth Metropolitan Region (PMR) has been assumed to have made a direct influence on land use and property development, or on economic development. Studies in public transport extension impacts have mentioned that public transport extension has prompted better access (Cervero & Kang, 2011) and the increased capitalization of property value (Cervero & Landis, 1993, 1997; K. Geurs, Zondag, de Jong, & de Bok, 2010; Rodríguez & Mojica, 2009), Other studies have mentioned the transport impact on land use developments (Cervero & Landis, 1997).

The question of the contribution of stations along the new line to increased train ridership has been considered for more than eight years. There was a need to examine the factors that can contribute to the success or failure of the transit system (the impact of railway line on train ridership) given the substantial investment in public transport.

Provided public transport investment contributes to economic growth (urban productivity), how does this economic growth, in turn, influence the level of train usage? Understanding the relationship between agglomeration as the economic impact of transportation investment, and the level of train ridership as a consequence of transport investment, is the focus of this thesis. The research objective is to establish a model to understand the relationships between the agglomeration and train ridership. In order to reach this objective, there is a need to empirically investigate the influence of the agglomeration on train ridership following the construction of the Perth – Mandurah railway line extension. A conceptual approach is proposed based on the land use, the transport integration of urban travel demand and the economic impact of public transport investment. This approach highlights the importance of agglomeration to be incorporated in any model for forecasting travel demand. The agglomeration reflects changes caused by the extension of public transport facilities, in which firms, households and employment are moved closer to each other, creating concentrated clusters of these agents, leading to higher urban productivity such as higher wages, urban land rents, and employment rates.

Spatial clustering and distance decay principles are two aspects of agglomeration investigated by this study. The analyses consist of: identifying the potential and actual state of the agglomeration; investigating the location and the form of distribution of the agglomeration; and measuring the influence of agglomeration on the demand for journeys to work by train. This agglomeration was defined by two measurements: effective job density and effective employed resident density.

Four reseach objectives were set up to reach research aim as follows:1) to investigate the link between transportation investment and agglomeration by providing evidence that support the existence of public transport induced agglomeration; (2) to understand the influence of public transport induced agglomeration on train ridership in terms of both effective job density and effective employed resident density; (3) to investigate the geographical extent of agglomeration relative to the distribution of stations that assumed as being the focal point (source) of agglomeration; and (4) to investigate the implications of agglomeration on job-housing balance from the wages/land rents and travel costs trade-off perspective and its further implication on train ridership.

The incorporation of agglomeration aspect in this thesis is investigated based on some research hypotheses:

(i) There is a distance decay and concentration pattern (clustering effect) of agglomeration according to the distance from train stations. The closer a location is to a train station, the higher the concentration of agglomeration, and the further a location is from a train station, the lower the agglomeration. Improved accessibility relates to a lower value (less negative) of the distance decay parameter of agglomeration. A less-negative distance decay parameter will be associated with higher agglomeration values and a higher chance of train ridership, following Graham’s (2007) assertion that there is a direct relationship between the magnitude of potential of transit demand and the accessibility level of a location.

(ii) There is a strong influence of employed resident numbers and job numbers (the scale or size of activities) on train ridership. The higher the number of employed residents or jobs, such as measured by density, the higher the chance that train ridership will be generated or attracted in a location (the LUTI model). In terms of agglomeration (SETI-LUTI model), the number of employed residents and jobs are defined not only by the scale or size of activity but also in terms of accessibility or as effective employed resident density or as effective job density. Therefore, the higher the effective employed resident density and effective job density, the greater degree of train ridership will be generated from residential areas (train trip production model) or attracted into workplace areas (train trip attraction model).

(iii) Greater distance of employed residents from boarding stations and of jobs from alighting stations results in a lesser influence on agglomeration induced train ridership.

(iv) If the number of employed residents and jobs in a location is balanced, there is more chance that people will work and live in the same (suburb) area. The more balanced the ratio of job to employed residents is (for example, between 0.75 – 1.5), the lower the number of spatial interactions (trip numbers) are between residential and workplace suburbs. The higher the ratio is (jobs are much higher than employed residents), the greater is the level of trip attraction to workplace suburbs. The lower the ratio is (employed residents are much higher than jobs), the greater is the level of trip production from residential suburbs.

1. 2 RESEARCH SIGNIFICANCE

Two mainstream approaches have focused on the relationship between transportation system and land use interaction (LUTI), on one hand, and the relationship between transportation system and spatial economic interaction (SETI) on the other hand. There has been limited research on the relationship between transportation system, land use, and spatial economic interrelationship. The concept of transport-induced agglomeration in terms of effective density has been adopted in this thesis as part of an exercise in understanding the relationship between spatial economics of the transportation system and land use and travel behaviour.

This thesis was expected to shed light on any important interrelationship between transportation-land use and transportation-spatial economic model, for the case of train ridership. Understanding to what extent effective density is differentiated from density in their influence on train ridership would give different implications for land use and transportation policies. Modifying parameters from density to effective density could reveal a larger degree of spatial variation in model findings, thus allowing application to a wider variety of land use-transport policy options.

Understanding the geographical extent of agglomeration benefits can assist in the assessment of difference policies for taking benefits from agglomeration externalities (Graham & Melo, 2010). Effective density provides a more comprehensive measurement parameter by combining the effect of scale and proximities, and thus generates a meaningful geographical extent for assessing agglomeration benefits of train ridership. For example, the questions of where to add jobs and employed residents in which economy sector; where to develop new stations; and where to strengthen the transit facilities to enable greater ridership can be clarified based on understanding that the geographical scale of agglomeration is influenced by different agglomeration externalities among different locations, especially in relation to the position of train stations as the focal point.

The SETI model was an attempt to make improvements on LUTI parameters for train ridership. Thus, the SETI model was expected to increase predictive accuracy due to the addition of an agglomeration component to a train demand model under LUTI. 'Public Transport for Perth in 2031' reported that increases in train patronage since the rail extension may go beyond initial expectations. For example, Perth has experienced a rapid growth in train ridership that was greater than an average of national and international cities. Annual train patronage increased more than four times in two decades to where in 2010 it reached 54.7 million passengers (Department of Transport (DOT) of Western Australia, 2011). This significant change in mode share of public transport may not be revealed using the LUTI model, due to the omission of a spatial economic factor.

The findings of updated modelling may be able to inform stakeholders as to how current TOD strategies can be enhanced in more effective ways to support train ridership production and attraction. The derived results could be useful in understanding the land use and travel behaviour impact of transportation infrastructure in a more comprehensive manner.

1. 3 RESEARCH SCOPE

Hensher et al. (2012) suggested to differentiate agglomeration and agglomeration effects. Agglomeration effects are the costs/benefits of employment agglomeration or dis-agglomeration, while agglomeration (or effective density) is a "measure of the extent of employment agglomeration itself". Graham and Melo (2010) measured agglomeration at a disaggregate geographical level; hence agglomeration was defined as an aggregation of firms, workers, or population in the geographical neighbourhood of individual firms. The effective density was defined by the amount of agglomeration experienced by a firm located at a site *i* that was captured by two measurements: the quantity of employment in another location (*j*), and the connectedness or proximity of site *i* and site *j*. "Agglomeration economies" were positive externalities or the increasing return of urban productivity advantages derived from the spatial concentration of economic activities (agglomeration) (Graham & Melo, 2010; Mare & Graham, 2009). The term "effective density" is also known as "market potential", i.e. a type of accessibility measure of agglomeration economies as an "intermediate concept" to establish a link between transport and agglomeration economies, or transport induced agglomeration effects (Hensher et al., 2012; Melo, Graham, Levinson, & Aarabi, 2013).

The scope of this thesis concerns only the magnitude of agglomeration of employment, or jobs and employed residents (the magnitude of agglomeration itself) and does not involve agglomeration economies or the agglomeration effects of transportation externalities on urban productivity or a firm's productivity. However, this thesis considered variables such as land rent and wages as an exogenous factor influencing train ridership demand. These two variables are sometimes regarded as an indirect or partial productivity measurement.

Therefore, this thesis is differentiated from other research in agglomeration studies in that other studies often measure the magnitude of agglomeration, then quantify this magnitude in terms of urban productivity and wider economic benefits of transportation infrastructure. While travel behaviour was regarded as a short term response to transportation system changes, it was also assumed that higher public transportation usage will stimulate higher metropolitan density in the medium term, contributing to higher effective density which in turn translates into an increase in urban productivity (Hensher et al., 2012; Weisbrod & Reno, 2009). Instead of assuming that public transport usage facilitates higher density, this thesis assumes that it was higher density and effective density that will influence the success of public transport usage. Therefore, this thesis examines this transportation/land use/spatial economic interaction from a 'reverse' perspective, such that the transportation infrastructure will eventually change the effective density and these changes feedback to travel behaviour. Nevertheless, the interrelationship between transportation/land use/travel behaviour essentially runs both ways, i.e. "public transportation facilitates higher density, higher density requires more public transportation" (Weisbrod & Reno, 2009).

In terms of agglomeration studies, the influence of effective density on the level of commuter or train ridership was viewed via the mechanism of labor market pooling, in which transportation infrastructure increases accessibility of activities, reduces travel costs between activity locations, and creates higher employed resident and job supply accessibility. The clustering of jobs and employed residents also makes the search for jobs and skilled workers more efficient and reduces costs of commuting to workplaces.

Measurement of effective density was conducted at an aggregated geographical scale. The scale of density of employed residents and jobs were quantified based on the proximity between suburbs in the overall metropolitan area. The examination of the influence of effective density on train ridership was conducted across both place of residence (train ridership production) and place of work (train ridership attraction), and across different economic sectors such as construction, manufacturing, retail and the benchmark total sectors. In addition, "agglomeration has been shown to dissipate more rapidly with distance from the source" (Mare & Graham, 2009, p. 13). This thesis establishes a model that incorporates the distance decay effect of agglomeration from its source, assumed to be train stations, which facilitate clustering around them as the railway line is extended.

The derivation of parameters affecting train ridership was conducted both under the LUTI and SETI framework. These variables were represented as external factors of train ridership, and therefore this thesis omits the internal factors of train ridership, i.e. the factors under the control of the transit agencies, such as transit services and fare levels.

1. 4 A MODIFICATION OR EXTENSION OF DENSITY INTO AN EFFECTIVE DENSITY CONCEPT IN TRAIN RIDERSHIP PREDICTION

Density has been considered to be the main feature of land use among all other components that determine public transport usage. Density or densification is the expected core result of improved accessibility in a land use system when new transportation infrastructure intervenes in a system. Broad studies have provided evidence on density increases, such as changed residential use from single housing into multi-storey housing and condominiums/apartments (Cervero & Landis, 1993), or changes from brownfield areas into developed areas, or intensified development by land use infill (Curtis & Olaru, 2011). These changes take place in the medium to longer term with other gradual changes (Wagener, 2004), such as changes in employment and residential patterns (Hensher et al., 2012).

However, studies of a less dense city with a strong dependence on private cars, such as Perth, may face some challenges when mostly relying on density (the core of a LUTI strategy), such as in the case of transit oriented development or TOD studies. The extent to which higher residential densities affect travel by public transport had been questioned (Chatman, 2008; Curtis, 2008; Grant, 2009) . Employment density or density of work locations is sometimes a more influential factor than density at place of residence (Chen et al., 2008; Ory & Mokhtarian, 2009). As a critique to Newman and Kenworthy's work (1998 and 1999), Van de Coevering and Schwanen (2006) argued that a single measure such as density, while important, left out some important ingredients which reveal the nature of some relationships, such as the role of individual travellers and space-time context. They proposed other factors to complement spatial factors (density and centrality of activity) such as "socio-demographic structure", "housing structure" and "urban development history". Chatman (2008) suggested adding a spatial scale as a consideration when measuring density. Chatman has proposed the concept of 'activity density' which was assumed to be related to distance travelled, and that density measurement should be measured across different scales to adjust for different types of modes.

How density is distributed across urban areas is one feature of urban form (Bamford, 2009). For example, 'monocentric' urban form presents the city centre as the core of activity development, hence degree of density is considered to be the most important controlling factor and spatial interaction is oriented toward the city centre. The 'multiplied nuclei' is the urban form of more advanced stages of urban development, where interaction between spaces is no longer concentrated in single direction toward the city centre, but also between regional centres and their services areas. The TOD concept suggests the train station as a core of activity where development of jobs and/or residential areas will be intensified within the station precinct.

The simultaneous measurement of density or the scale (or density) of economic activity and the proximity between activities was proposed in Venables (2004). This concept is known as 'effective density'. Effective density has been widely recognized and used in agglomeration studies to assess the influence of public transport investment on wider economic impacts or urban productivity (Graham, 2007; Venables, 2007; Graham & Melo, 2010; Melo et al., 2013; Hensher et al., 2012). Effective density is a variable that represents urban agglomeration and also measures a market potential. Effective density is an accessibility-type measure of agglomeration, such that it can better describe the potential for opportunities that are available to firms, workers, or the city and the real accessibility than employment density or total population measures (Melo et al., 2013).

Travel behaviour is a result of the short-term impact of a ‘shock’ to the transport system (such as the development of new stations or a railway line extension). On the other hand, agglomeration is a medium to long term effect of transport investment that arises through the changes in the distribution of housing and employment (Hensher et al., 2012). Studies using the LUTI framework often consider that changes in travel behaviour would lead to the changes in medium term decision-making that affect the spatial distribution of housing and employment. These changes in distribution of activity patterns would further facilitate the agglomeration process and feedback to urban economic and land use such as land rents, wages, and urban productivity. How agglomeration, as defined in terms of effective density, may further influence travel behaviour such as train ridership is focus of this thesis. This thesis uses the reverse relationship between land use-transportation interactions, assuming that effective density is endogenously defined by travel time changes or accessibility changes that feedback to train ridership.

Improvements in the transport system can impact the ‘effective’ employment density even before or without any of the actual employment numbers changing, provided travel times are used to indicate the relative positions of these employment numbers with respect to the reference point (Hensher et al., 2012). For example, during a period of constant decay of agglomeration, if a proposed transport improvement reduces travel costs to the east of site by 20%, then the new effective density at that site will change in way that is equivalent to moving employment to the east 20% closer (Graham & Melo, 2010) without necessarily changing the physical density in the various locations.

There is an issue of non-constant spatial parameters, for example, due to a varying geographical context resulting in the variation of coefficients that indicate the degree of relationship between land use and the transportation system in a geographical space (Páez, 2006). The varying geographical context will give rise to the varying influence of land use variables on travel behaviour, such as in the case of the influence of density on travel behavior. The effective density concept embodies the quality of accessibility since either implicit or explicit travel time savings are incorporated into the parameter. This is the main difference between the concept of density and effective density.

1. 5 RESEARCH METHODS

Research methods in the agglomeration analysis were based on a comparison of before-and-after public transport extension; i.e. 2006 and 2011. Research methodology for modelling purposes was based on a cross-sectional approach. The derived model attempts to predict train ridership production and attraction based on 2011 data, and is then calibrated on 2006 data for train ridership production only.

The stages of research consist of:

1. Investigating the impact of public transport investment.

   This stage aims to investigate the impact of public transport investment on some key variables. This analysis compares the changes in effective density, train ridership, and travel time by park-and-ride between 2006 (before extension) and 2011 (after extension). Statistics descriptives are used to examine if the changes in all variables are larger in the area of the Perth-Mandurah railway lines than other lines.

   In this stage, the aim is to provide evidence towards testing if the effective density is related to public transport investment in the Perth-Mandurah railway line extension. The correlation between the changes in effective density before-and-after the extension, to the changes in travel time and the changes in train ridership was examined to determine whether the link between effective density and train ridership is a result of public transport investment or public transport induced agglomeration.

2. Investigating the degree of agglomeration.

   This stage aims to investigate the degree of agglomeration of jobs and employed residents. While the literature has proposed some methods for agglomeration measurement, such as the Gini index, Moran I statistics, and location quotient or LQ (Guillain & Le Gallo, 2007), this thesis adopts the measurement of agglomeration in terms of effective density, as used in

Venables (2004) and Graham (2007). The agglomeration in this context is related to public transport induced agglomeration. Further, the meaning of public transport induced agglomeration would be exchangeable with the term "effective density" in this thesis.

In order to measure the effective density, a decay parameter was first estimated, using the gravity model. This parameter was then used to determine the effective density for each suburb. The proximity between suburbs was measured based on a travel time matrix by park-and-ride produced by the STEM model of the Department of Planning of Western Australia.

3. Measuring spatial patterns such as clustering and decay patterns around train stations.

   This analysis aims to identify if agglomeration showing tendencies of decay in favour of proximity to train stations (distance decay effect) and if a concentration of activities (hot spots) emerge around station, such as in the context of transit oriented development (clustering effect). In this analysis, this thesis investigates whether or not the agglomeration and clustering (jobs and employed residents) exists around train stations and across different economic sectors. The method Getis-Ord G* was used to measure cluster phenomena; the probability density function and exponential distance decay curve was used to measure decay phenomena.

4. Establishing a model to understand the relationship between the agglomeration and train ridership.

   This thesis incorporates train trip production (place of residence) and train trip attraction (place of work) into the model for train ridership prediction. This is a process where factors affecting train ridership are derived under the LUTI framework and SETI framework.

5. Establishing the relationships between effective density and distance of suburbs to train stations to investigate the geographical extent of the influence of effective density on train ridership. If any agglomeration is influenced by the

proximity of suburbs to the station, the predictions for train ridership will be more accurate when the model incorporates this distance decay effect.

6. Analysing of the implication of models on land use policies. This thesis derives the beta coefficients of effective density with the interaction terms from various models. The interpretation of these beta coefficients with regard to policy scenarios is based on calculation of their multiplicative effects on the proportion of train ridership. The underlying hypothesis of agglomeration is used to assess the implications of job-housing balance for train ridership, by using wage equation models (Timothy & Wheaton, 2001). Further implications of the various statistical models on TOD strategies are discussed and the influence of effective density on train ridership is evaluated in comparison to other studies.

1. 6 RESEARCH STRUCTURE

The research structure is represented as topics separated out into each chapter of this thesis. These are as follows:

1. Chapter 1 provides the introduction to this thesis, including the background to the research that initiated the current problems and research hypotheses; the research significance and scope; the main features of density and effective density and why it is important to differentiate the influence of these two measurements on train ridership; the stages of research, and the research structure.
2. Chapter 2 presents a literature review, mainly discussed factors affecting train ridership under the LUTI and SETI frameworks and the potential for ongoing research.
3. Chapter 3 explains the proposed research methodology, sources of data, definitions and measurements of each of the predictor variables and dependent variables, and the research method.

4. Chapter 4 describes the preliminary analysis that was required to prepare the data sets into an adaptable and useable set of variables for use in the analysis section.

5. Chapter 5 describes the first analysis. This chapter presents the measurement of agglomeration; the estimation of decay parameters for effective density; the empirical evidence for public transport induced agglomeration by linking the railway line extension to the association of changes in effective density to changes in travel time and changes in train ridership; then finally the analysis of spatial patterns of effective density comprising the clustering and distance decay phenomenon of agglomeration toward the nearest train station.

6. Chapter 6 presents the second analysis. This chapter discusses the result of train ridership prediction and compares the strength and magnitude of the relationships among predictor variables for train ridership. The results of the model in terms of factors affecting train ridership, and compares these to others in the literature.

7. Chapter 7 presents the third analysis. The geographical scope and the degree to which density and effective density influence train ridership is discussed. The comparison between density and effective density is focus of this chapter. Three aspects of the model results are discussed in this chapter: the overall model fit; the contribution of effective density to improving the prediction capability of the model by the measurement of coefficient of determination and partial determination in LUTI and SETI; and the model validation or back-testing based on the 2006 dataset.

8. Chapter 8 discusses the model findings in relation to the underlying research hypothesis concerning the trade-off concept between wages/land rents and transport costs. The application of the job-housing balance hypothesis is discussed with regard to some alternative policies for enhancing the effectiveness of Transit Oriented Development, or TOD. Another component of discussion involves the potential uses for incorporating effective density variable and its geographical extent into future modelling of land use

development and transportation policies, especially in the context of the Perth metropolitan study area. The significance of incorporating an effective density component into train ridership prediction is re-evaluated.

9. Chapter 9 provides the conclusions. Key findings are presented, and the limitations of this thesis are elaborated. Some policy implications are briefly stated. Critical points on how to improve the findings and methods are proposed for future research.

CHAPTER 2. LITERATURE REVIEW

2. 1 INTRODUCTION

This chapter builds a conceptual framework for train ridership prediction in which the factors involved in train ridership are derived based on the Land-Use Transport Interaction (LUTI) framework and its extension. In the context of travel behaviours, the key dimensions of land use and transport interaction are discussed. The main concept and the nature of the relationships in each study are mentioned. The discussion is divided into two parts: (1) LU-T studies in which land use influences travel behaviour; and (2) T-LU studies where transportation infrastructure shapes urban form or influences land use. Some gaps in the parameters used in LUTI are discussed and a framework for LUTI and its extension is proposed for deriving the appropriate factors for train ridership prediction.

2. 2 LU-T STUDIES

Mainstream studies in LUTI have assumed a causal relationship between land use and transportation. The term "land use transport feedback cycle" has been essential in developing the two-way relationship between trip and location and the way they co-determine each other (Wagener, 2004). Some minor but popular studies have concluded that the effect of self-selection has influenced this apparent causality. "Self-selection" refers to the possibility that individuals/households endogenously self-select themselves into environments that support their inclinations for certain transport modes. For instance, travel preferences of individuals/households means they select a location where they can behave in their preferred way and therefore, provided one's travel preference is to use train, she/he will move to a location where train is catered for (Nurlaela & Curtis, 2012). Self-selection therefore implies an over-estimation of the effect of the land use factor on travel behaviour. However, a non-spurious, causal relationship appears to be true based on the evidence for a statistically significant association between land use and transport after controlling for self-selection (Olaru, Smith, Xia, & Lin, 2014). Regardless, both the causal relationship and self-selection effects tend to promote the use of public transport and curbing the dominance of cars. Other studies report inconclusive results on the nature of causality but have emphasised its importance. The table 2.1 presents the above-mentioned references subdivided by their field of study and type of relationships.

Table 2.1 The division of field of study, type of inference of relationship, and references for land use influence transport studies

No.	***Type of study***	***Type of inference***	***References***
1.	LU - T Land use influences transport	Causality relationship	(Lindsey, Schofer, Durango-Cohen, & Gray, 2011), (Boustan, 2009), (Chen et al., 2008) (Ory & Mokhtarian, 2009) (Van de Coevering & Schwanen, 2006) (Limtanakool, Dijst, & Schwanen, 2006) Van Acker, V., F. Witlox, et al. (2007).

	Self-selection effect	(Levine & Frank, 2007) (Pinjari, Pendyala, Bhat, & Waddell, 2007) (Pinjari, Bhat, & Hensher, 2009) (Bhat & Guo, 2007) (Bhat & Eluru, 2009) Eluru, Bhat, Pendyala, and Konduri (2010); (Mokhtarian & Cao, 2008).
	Inconclusive relationship	(Næss, 2009) (Naess, 2006) (C. Kim, 2008) (Senbil, Kitamura, & Mohamad, 2009).

Density is the most common land use variable used in LU-T studies. Most findings agree on the conclusion that there exists a positive influence of density on public transport uses. Other factors besides density, included as part of a building environment factor, consist of design and diversity (the 3Ds) as well as distance to transit and destination access (the 5Ds) (Ewing & Cervero, 2001). Further development of these Ds components involved the addition of two more components, such as demand management as a sixth D and demographics as the seventh D (Ewing & Cervero, 2010).

Density and diversity variables are usually assumed to have some correlation with mode choice, travel distance, and trip rate. Van Acker, et al. (2007) explains that there has been a common belief among researchers that high density and diversity always correlated with the higher use of public transport and Non-Motorized Transport (NMT). In their study, travel distance negatively correlates with density. Characteristics of design such as dwellings, streets, and neighbourhoods may influence mode choice, but limited to trips for shopping, social, and recreational activities. Urban form and urban size were shown to be the main contributors to mode choices for commuting, especially for working trips.

Design aspects with more detailed measurements of urban form, i.e. urban form at the micro level, such as pedestrian-oriented design with small block sides, sidewalk systems and limited residential parking, are also often considered. However, these micro-level land use components are often too small-scale to define a clear relation between urban form and travel behaviour (Boarnet & Crane, 2001). Design aspects at the medium scale of urban form have been considered more appropriate to model. Examining a large area may dilute the effects of urban form, while a much smaller area may omit some important effects potentially. Therefore, recognising areas smaller than a zip code would likely have a better correspondence to the activity space in which non-car travel takes place (Vance & Hedel, 2007).

Falconer, Newman, and Giles-Corti (2010) criticized the "new urbanist" principles applied as part of the design code in developing 'liveable neighbourhoods' in Western Australia. Liveable Neighbourhoods (LNs) represents urban design in the form of straighter roads, provision of footpaths, shorter block lengths, and overall network connectivity. However, these design components did not deliver any local utilitarian destinations but were only anchored by public open spaces and water features; hence the purpose of a local walking trip would be undoubtedly for leisure. In fact, they found the average distance to key destinations was found to be significantly lower in conventional neighbourhoods when compared to LNs.

In addition to the density concept, another influence on travel behaviour was discussed in relation to residential location choice by Bhat and Guo (2007). They considered that land use factors have an influence on both residential location and mode of transport choice. Studies in this area investigate the joint-decision, the co-location, and the job-housing balance. However, the causal relationship in this field is often viewed as being confounded by the self-selection effect. Including the possibility of self-selection interfering with other causal effects, it has still been shown that land use moderately influences travel behaviour.

Research in the area of joint mobility for residential locations and workplaces supports the phenomenon of self-selection. Self-selection is defined as the effect of external bias in the decision-making process for residential location choice. Individuals/households choose the location or built environment that supports their preferences, lifestyles, personal traits, attitudes or socio-economic characteristics. Self-selection might happen due to a combination of intervening factors when residential location and travel behaviour variables are related to one another. Three sources of interrelated variables have been identified: socio-demographics (e.g. income), attitudes that comprise auto-disinclination or inclination, and travel preference. The three factors may be observable, such as in the case of household income; or unobserved such as attitudes, perceptions and environmental considerations (Pinjari et al., 2007). For example, individuals who value walking are likely to locate themselves in the neighbourhood with walking trail facilities (Guo & Bhat, 2007).

The co-location hypothesis was proposed in the research by C. Kim (2008). The hypothesis was stated as: “the decentralization of job centres or suburbanization would reduce the trip distance or commuting time due to relocation of residence or workplace as an adaptation process to the worsening congestion” (C. Kim, 2008). The results of this study found that the number of trips decreased with increases in commuting time and distance. However, there was not enough evidence to conclude if the decentralization of employment caused a reduction in the commuting time. The connection between the workplace and residential mobility was found to be weak.

Table 2.2 Key concepts in land use influence transport studies

No.	***Key concepts in land use influence transport***	***Variables***	***References***
1.	Density concept	Employment density, residential density, activity density, population density, proximity to infrastructure (accessibility), and land use diversity. Distance of residential area to the city centre. Vehicle mile travelled (VMT).	(Limtanakool et al., 2006) (Cervero, 2002) (Chatman, 2008) Naess (2006) Lindsey et al. (2011).

2.	Joint decision of residential location and travel choice, self-selection effect.	Decision maker attributes (household structure, vehicle ownership, etc.) Neighbourhood characteristics The preference of residence on neighbourhood type Public transport accessibility Land use density and mix (diversity) Residential location choice (e.g. between centre and sub-centre, between different type of neighbourhood) Mode choice or mode switching (e.g. between car and public transport).	(Bhat & Guo, 2007) (Levine & Frank, 2007) (Senbil et al., 2009) (Lee & Waddell, 2010) (Vega & Reynolds-Feighan, 2009) (Chang & Mackett, 2006) (Eluru et al., 2010).
3.	Co-location hypothesis	The number of trips Commuting time and distance Suburbanization employment centre (decentralization) Residential mobility Workplace mobility House price Wages	(Boustan, 2009) (C. Kim, 2008) (So, 2001) (Manaugh, Miranda-moreno, & El-geneidy, 2010).

2. 3 T-LU STUDIES

T-LU studies refer to studies in the LUTI field where an emphasis is put on the examination or understanding of how the transportation infrastructure influences land use or urban form. Transportation infrastructures are considered to influence land use in three areas:(1) the effects of transit on development; (2) how transit affects the quality of life; and (3) the available mechanisms to implement those effects (Catanese, 1988). In a US-based context study, three possible forms of the relationship were also stated: 1) how highways and mass transit contribute to decentralisation trends; (2) how transport affects the local balance between jobs and housing; and (3) how transport affects the pattern of commercial investment (Boarnet & Crane, 2001).

This thesis has identified two typologies of T-LU studies: the influence of transportation infrastructure on land use or urban form, and transportation evaluation or transport impact assessment, where studies relate to the environment or economic impact of transportation infrastructure. Recent developments in this area include extension to a wider impact assessment that adds the dimension of public transport-induced agglomeration.

Transportation infrastructure development, such as massive highways and electric railway development has shaped cities in the US into a sprawling form and created more car dependences due to leap-frog development (Cushman, 1988). Transport has created the segregation of development, created more distance between sub-centres and between the sub-centres and the CBD and created low-density suburbs, resulting in inefficiency of transport services. Conversely, the subway development in Toronto, Canada, encouraged fill-in development and strengthening of urban form. It created more dense population and jobs, and strengthening of some suburban centres, whereas the centre city became less purely a CBD. Suburbs became local employment centres where more people were attracted to live and work in traditional neighbourhoods (Greenberg, 1988).

Gallivan (2015) used a sophisticated statistical analysis tool which included a calculator that planners can use to predict the impacts of transportation on land use and travel behaviour (Gallivan, as cited in Department of Transport, 2012, p. 28). The analysis found that the effects of transit systems have created more compact and sustainable US cities. The tool simulated urban form without a transit system, which resulted in otherwise unexpected effects on the US urban population in aggregate, such as lower population density, lower activity density around transit nodes, increasing vehicle mile travels and fuel uses, which in turn led to increases in GHG emissions.

However, new transportation infrastructure was shown to induce growth through the multiplier effect of development, by increasing the accessibility and mobility of site users. This multiplier effect of new development from new highway corridor and interchanges applied not only to sites in the vicinity but also often to locations at quite a distance from the interchanges (Bezler and Autler, as cited in Evans, Pratt, Stryker, & Kuzmyak, 2007, pp. 17-14).

In the second typology of T-LU study, the empirical approach for ex-ante transport evaluation has been a standard cost-benefit analysis (CBA). The components of transport project appraisal have been mentioned in some of the literature. These consist of spending, travel improvement and access improvement impact. Impacts have been measured as direct, indirect and induced impacts. Spending impacts have been related to the calculation of benefits for transport industry and its supplied businesses. Spending money on public transportation infrastructure has been shown to create immediate jobs and income by supporting manufacturing, construction, and public transport operation activities. Travel improvement impacts have been related to more broad areas such as travel time saving, travel cost saving, reliability improvement and safety improvement. Access improvement impacts have resulted in both expanded public transportation and reduced traffic congestion, such as mobility and market access, and spatial agglomeration economies (Weisbrod & Reno, 2009).

In the context of wider impact assessments, the framework of evaluation has consisted of transport investment assessments under the environment of externality known as Wider Economic Benefits (WEBs). WEBs was an indirect and long term transport project, hence it was difficult to be measured using traditional cost and benefit analysis (CBA).(T. Lin & Truong, 2012) An agglomeration effect was the largest effect of Wider Economic Benefits. The agglomeration benefits accounted for more than 50% of the wider economic benefits produced by related transport projects (Rognlien, as cited in T. Lin & Truong, 2012, p. 1)

The assessment of wider economic benefits goes beyond the scope of the standard conventional calculation of cost benefit analysis (CBA). Applying the new form of appraisal to CrossRail in the UK, it was found that total user benefits based on CBA were valued at 12.832 million while the WEBs were 15,926 million. The CBA in this study consisted of business time-saving, commuting time-saving, and leisure time-saving while the WEBs consisted of agglomeration benefits measured with a Total Factor Productivity (TFP) parameter (Graham, 2007).

2. 4 LUTI FRAMEWORK APPLIED IN THE AREA OF TRAIN RIDERSHIP MODELLING

An integrated transportation-land use model was argued as the only approach to investigate the “true” impacts of rail transit system on a long run (Badoe & Miller, 2000). The goals of such a model are to predict the responses in travel behaviour due to changes in one or more components of land use and the transportation system. Models are also used in testing both land use and transportation policies and to assess their impacts (Wagener, 2004).

In this thesis, the LUTI framework is used to predict the aggregate response of the population to the level of train usage due to changes in land use and transportation. This thesis adopts the LUTI framework approach as proposed by Southworth (2001) and Broarnet and Crane (2001). Southworth (2001) mentions that the influence on travel demand can be understood from the interaction of exogenous demand related to people and the economy, developmental demand related to land use system, and the transportation supply in terms of transportation infrastructure development and investment. In addition, Boarnet and Crane (2001) use the price/costs-changes mechanism to understand how the supply relating to the transportation system and demand interact to create travel demand.

Southworth (2001) offered a framework to understand the origin of travel demand in response to supply-demand of land use and transportation system. On the demand side, the author mentioned transportation had grown rapidly due to the natural growing of population, employment and the economy. Those factors were named exogenous demands. The developmental demand referred to the land use system, consisting of spatial pattern, density and the level of land use mixing. The joint influence of exogenous and developmental demand combined to create travel demand. Southworth (2001) also explained the micro-economic approach through which demand and supply interacted. There was considered to be a causal loop between price elasticity and travel demand. As supply of transport increased (transport side), the cost decreased. Thus, a new highway (or other transportation infrastructure investment) led to lower travel costs and induced more travel. Over time, this influenced the value of urban land parcels leading to new building and new travel demands (Figure 2.1).

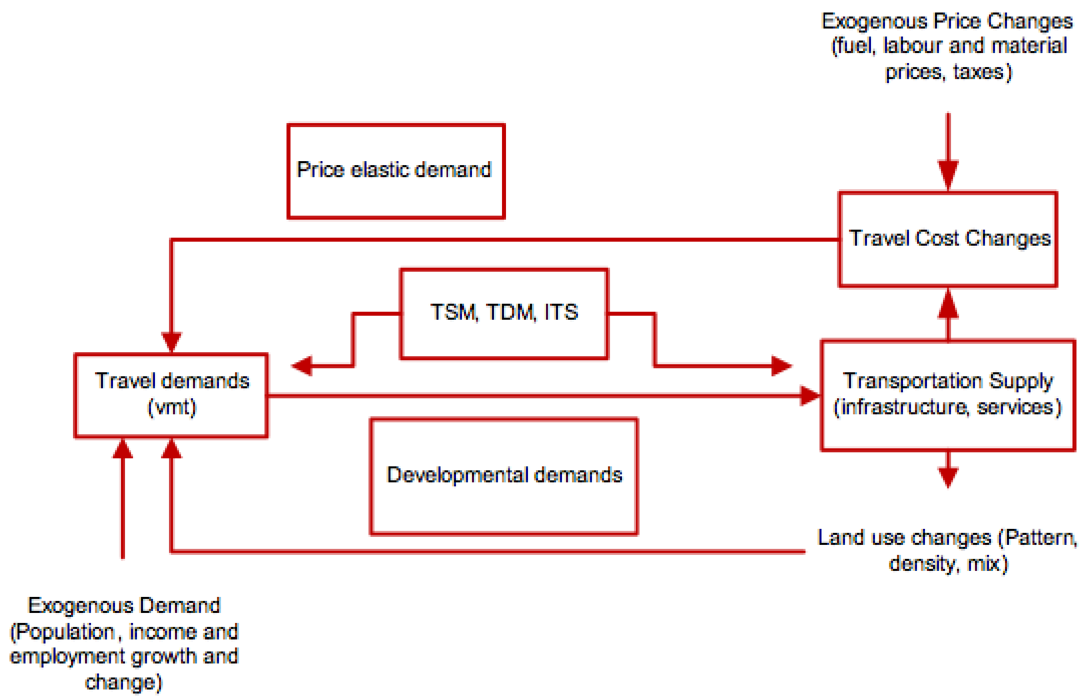


Figure 2.1 Sources of VMT growth and their interaction (Southworth, 2001, p. 1273)
In order to solve the problem of increasing travel demand, Southworth suggested practical, short term management. This consisted of Transport System Management (TSM), Travel Demand Management (TDM), and ITS (Intelligent Transport System) measures.

The economic dimension was incorporated into the framework of supply-side land use-transport interaction through the implicit prices/costs changes (Boarnet & Crane, 2001). The development of an efficient public transit network and massive highway or freeway construction reduced the travel costs between origin and destination, especially between homes and work places. On the demand side of land use and transport interaction, Boarnet & Crane (2001) used the price/costs approach in understanding and explaining the mechanism of how urban form changes impacted transit or travel behaviour. Land use and design would influence travel behaviour by changing the price of travel, i.e. through linking the neighbourhood design components to price variables. They proposed urbanism, a design principle that adjusts the land use and transport system to minimize travel costs and directs the 'rational choice' of the traveller into more sustainable transport by non-motorized means (bicycle or walking) and transit. An example of this is the grid street pattern, which lowers trip costs for cars and walking in comparison to the conventional design.

This thesis adopts the framework of LUTI mentioned in Southworth (2001) and applies it in deriving the important factors affecting train ridership. In short, the LUTI framework offers a component of demand side (land use and the exogenous factors; such as population and economy), supply side (the transportation system, and transportation infrastructure), and the interaction of supply and demand due to price/costs changes.

Parameters affecting transit use are generally comprised of both internal and external factors. Internal factors are those factors under control of the transit agencies (Taylor et al, 2009) including transit services such as fares, frequency of service, service design, and marketing (Fleishman et al., 2007). External factors are those factors beyond the control of the transit agencies. These comprise the local or regional economy, land use, population, transit policy, and attributes of competing modes (Fleishman et al., 2007). This thesis focuses on the external factors of train ridership under the LUTI framework, thus omitted the internal factors. Although the transit agencies cannot control the external factors, they can assess the impact of external factors on travel demand and take action to respond in a way that benefits them, or avoid risk by mitigating negative impacts (Fleishman et al., 2007).

2.4.1 Parameters for train ridership: land use factor

The parameters contributing to train ridership that were part of the demand side of LUTI are related to land use. Land use generates travel demand from various activities according to each land use type. Land use consists of (a) the amount, quality and spatial distribution of opportunities supplied at each destination (jobs, shops, health, social and recreational facilities), (b) the demand for these opportunities at origin locations (such as households), (c) the interaction of supply and demand for opportunities, which may result in competition for activities with restricted capacity, such as job and school vacancies (K. T. Geurs & van Wee, 2004, p. 128).

Density, land use mix, and the TOD design components have been the most popular land use variables used as predictors for transit ridership (Evans et al., 2007). Density and the distance from the CBD has been shown to correlate with a consistent increase in public transport share, independent of other factors (Rickwood & Glazebrook, 2009). Density-related built-environment impacts have often been found in research to be the influencing factors for both housing choice and mode choice. The density factor is usually complicated by many other factors such as self-selection issues, generalized travel costs, accessibility and access to transit stations.

Chen et al. (2008) examined the importance of density on mode choice decision. The case study used was the New York Metropolitan Region. Density was examined in more detail with a division between employment density in work places and residential density in housing areas, and these two sections were compared to identify how much they impacted mode choice. The policy implied in the research was similar to the ABC policy applied in the Netherlands, in which there was an aim to channel new employment into nodes that were well served by the public transport network in order to restrain the use of the private car (Chen et al., 2008). Therefore, the areas with the highest employment densities, such as the CBD, usually have a concentration of transit hubs and a moderate level of population density due to high rents excluding many residents from the CBD. The policy to create mixed-use neighbourhoods in residential areas only (trip origin) was not considered to be enough, due to the importance of work place neighbourhoods outweighing that of residential neighbourhoods (Ory & Mokhtarian, 2009).

On the other hand, other research has found the effect of job density in the CBD can be less than the population density (Van de Coevering & Schwanen, 2006). The influence of density on travel behaviour was seen to be variable and to depend on the focus of the study area. For example, total distance by public transport only correlated to the employment density in the inner area. The percentage of jobs in the central area was found to be correlated negatively with the average commuting distance. This reflected the larger distances travelled by public transport in metropolitan areas with more jobs per hectare in the core (Van de Coevering & Schwanen, 2006).

Provided there were enough people to be carried on (mass) transit, the more people could be transported per vehicle, per hour, or per mile of service, the more effective and productive the transit would be (Cushman, 1988). Therefore, the more densely-concentrated people are, whether in home areas, shopping areas, or schools, the more people the transit system was able to carry per unit of service offered, the more revenue would be made, and the more productive the transit system would be.

The connection between density and public transport use may be explained by the indirect effect of density. Higher density creates activity locations closer to each other and allows for shorter access and egress times to public transport facilities. This indirect effect is captured by an indicator called proximity to infrastructure networks or distance to transit. Research by Cervero (2002), and Limtanakool et al. (2006), found that travellers with good highway accessibility are likely to commute by private car while travellers with good access or residing close to railway stations are likely to use public transport.

Furthermore, there has been a need to use a more specific density indicator to measure the impact on transport with precision and to reduce bias - an example is the use of the parameter “activity density”. Activity density is defined as “the number of local desirable non-work activities, whereas higher activity density implied lower average distance to activities over all activities from home” (Chatman, 2008) . Accessibility of activities has been assumed to relate to the reduction in travel distance, and the number of trips, especially the linking trips of multipurpose trips (Van Acker et al., 2007).

Density of development, zoning controls, and relative locations of major employer and residential areas; have been found to be statistically significant influences on mode choices and the need for transit (Fleishman et al., 2007). Based on extensive surveys of the 1992 California TOD stations, the effect of heavy residential TOD areas was shown to be different to the effect of retail-based or office-based TOD areas. The residential TOD areas contributed to lower ridership than the retail- and office-based TOD areas. Transit-oriented office centres created compact access due to building to building walking. Among 1,400 employees at 18 worksites within 1.2 mile of each other, the surveys found that workers near rail stations were 2.7 times more likely to commute by rail than the average for cities as a whole. It is worth noting that employees working within Transit-oriented retail centres generally had lower vehicle ownership and walking was the predominant travel mode besides rail (Evans et al., 2007). Similarly, rail commuting was stronger in stations that served employment centres than those which do not (Blainey, 2010). Distance decay weighted regression revealed different sensitivities according to the location of new developments within the station catchment areas. For example, adding 100 new commercial jobs located within the 0-100 m band from a station would create more journeys than adding new jobs in 700 – 800 m band at a 6 times larger rate (Gutiérrez, Cardozo, & García-Palomares, 2011).

2.4.2 Population and economic component

Mode choice studies have emphasised the underlying factors for why people choose certain modes of transport over others, for example, why people choose trains compared to cars or the other way around. The underlying mechanism that determines choice by individuals or households has been discussed using several approaches, such as microeconomics, sociology, and psychological analysis and market segmentation (McFadden, 2007; Ohnmacht, Götz, & Schad, 2009; Van Acker & Witlox, 2009):

1. The micro-economic approach: the microeconomic approach views travel behaviour with regard to the perception of transport or travel as a commodity (Ferguson, 2000). Hence, travel demand is considered to be a derived demand, based on activities and goods that require travel (Boarnet & Crane, 2001). In line with Ferguson (2000), Liu (2011) applied the perspective of "commodity" for travel using the assumption that people will choose to travel (travel choice, vehicle choice) based on their preferences for different options (goods/commodities/choices) on the costs of these goods (such as monetary and time costs), and the budget constraints, i.e. the availability of income and time in order to maximize their utility.

   The economic approach as a framework to analyse travel behaviour often values rational-behaviour-based decision making (McFadden, 2007). People are assumed to give consistent responses to certain situation, with the main objective being to maximize utility. Put another way, people would optimize the benefits they received and minimize all costs likely to be incurred. One simple example, stated in McFadden (2007), examined intermodal transportation, which provided good service for people to utilize non-motorized combination public transport (like buses equipped with bike racks). These would be valued as giving more positive benefits to users, and would increase people's inclination to use buses and bikes together.

The micro-economic approach discussed in depth in McFadden (2007) concerns the utilization of a random utility model of discrete choice, that was originally used by Domencich and McFadden (1975). It modelled the urban travel demand (mode choice) assuming a utility structure that would give a multinomial logit and nested logit model for mode choice. This approach also potentially applies to both very detailed analyses of household activities, all the way up to broad questions of transportation system investment and operation. The premise of the random utility model was that individuals or households would obtain utility from the choices they make in order to maximize that utility value. Since different people would have different tastes, these differences would be captured by making the utility random. The model used the observed discrete choices to estimate the distribution of utility and its relationship with transportation variables (independent variables) of defined research interest. This approach is a highly disaggregated form of modelling, unlike the spatial approach of travel demand modelling, which is related to aggregate modelling.

2. The sociology approach (vertical and horizontal segregation approach): the terminology of vertical and horizontal segregation may be found in Ohnmacht et al. (2009). **C**haracteristics such as age, gender, household size, household income, level of education, employment status, mobility constraints, are commonly used as variables. These variables sometimes are used in combination with each other to form categorisations of individuals or of households according to their similarity of characteristics. This creates segregation vertically, also known as 'bundle of individual or household' or as 'household type'.

3. The psychology/personality approach: McFadden (2007) refers this approach to cognitive psychology. McFadden (2007) proposes a combination of sociology, anthropology, and psychology. This approach focuses on the consideration of human tastes, personal traits, life style, and preferences as the main predictors for travel behaviour. Using this point of view, people are no longer segregated by the usual socio-economic or demographic types, but rather by internal preferences (horizontal segregation). In this way, people from the same household type possibly have different travel behaviours and people from a different socio-economic category possibly have similar travel behaviour.

   Most studies in this area attach a higher significance to the effect of personality variables on travel behaviour, compared to social economic and land use variables**.** However, lacks of data and difficulties in collecting data have been major obstacles for these studies. In fact, research in this area has mostly focused on leisure trips. According to Ohnmacht et al. (2009), the personality research approach has aimed to reveal more complex travel behaviour including need to take into account the social realities of individualization and pluralism. For that reason, leisure trip studies were appropriate for examining, and able to successfully confirm, the 'in group homogeneity of lifestyle' often controlled for in lifestyle trips due to the similarities in interests or preferences.

4. Market segmentation has been defined as: "a customer classification where subsets of people shared some common characteristics (socio-economic and personality traits) that were identified within a larger sample" (Diana & Mokhtarian, 2009, p. 456). In addition, Shiftan, Outwater, and Zhou (2008) used **s**egmentation by attitudes which were aimed to understand motivational and affective factors. It was hypothetically stated that segmentation by attitudes was important in order to identifying potential mode switchers since socio demographic factors had little bearing on travel profiles of the segment.

Table 2.3 The approaches in mode choice studies

No	***The Approach in Mode choice Studies***	***References***
1.	The microeconomic theory: Using utility theory, economic instruments, econometric model	(Domencich & McFadden, 1975) (Ferguson, 2000) (Liu, 2011) (Hensher & Puckett, 2007) (Hess, Rose, & Hensher, 2008) (S. Kim & Ulfarsson, 2008) (Eluru et al., 2010)) (Dargay, 2007).
2.	The sociology approach (vertical and horizontal segregation, market segmentation)	(Nolan, 2010) (Commins & Nolan, 2011) (Diana & Mokhtarian, 2009) (Shiftan et al., 2008).
3.	The personality/psychological analysis on human preference/taste/lifestyle (affective dimension, perception, attitudes)	(Ohnmacht et al., 2009) (Ory & Mokhtarian, 2009) (Lois & López-Sáez, 2009) (Holden, 2007) (Choocharukul, Van, & Fujii, 2008).

It is worth pointing out that these approaches are differentiated by the context of their research approach. Aggregate studies define population variables as a bundle of socio-economic or demographic characteristics (Hensher & Ton, 2002), hence such studies will tend to categorise individuals or households within the same socio-economic and demographic level into the same category without differentiating their level of taste, preference or individual inclination. In fact, the socio-economic and demographic level is assumed to dictate their preference of travel mode, by assigning the same level of utility for the set of behavioural choices into the same category. Therefore, vertical segregation is the most common approach in determining population component measurements used in aggregate studies. This includes defining the characteristics of population groups based on income level, household structure, household size, educational background, employment structure, and ethnic background.

Train ridership studies have broadly used socio-demographic and socio-economic variables as determinants of ridership. Taylor et al., (2009) used these variables as external factors, given that these factors were outside of the control of the public transit system. They found that the percent population in college, percent population of recent immigrants, and percent carless households all positively influenced the total urbanized area transit ridership. However, these influences were far less than the influence of the predicted service supply, i.e. an internal factor.

Car ownership levels and car usage were shown to be important not only in relation to the socio-economic group but also as a response to the urban form that eventually dictated travel behaviour. Wagener (2004) in his famous 'land-use transport feedback cycle' identified car ownership to be a result of the inter-relationship between transport and land use. Sub-urbanization development and the movement outward to the suburbs of town shopping centres, for example, need to be served by an increased level of car ownership in the population (Hall, 1969). Regulation to restrict automobile ownership, such as through tax on car acquisitions, has been demonstrated to have a positive impact on transit patronage (de Grange, Troncoso, & González, 2012). Car ownership also has been shown to have a strong association with the choice to use cars or trains, especially for those who live in the immediate surroundings of a rail station.

The amount of income is strongly related to the degree of car ownership level, where both influence ridership. Income has been assumed to have a negative correlation with ridership. Chan and Miranda-Moreno (2013) examined average household income at the 1000 m buffer zone around stations for ridership prediction. Chan and Miranda-Moreno found that increasing 10% of income would decrease ridership 8.2% for the Metro railway in Quebec, confirming that people with higher incomes were less attracted to transit. The effect of increasing income on demand for transit ridership was found to counter much of the increase in train ridership resulting from increases in employment (McLeod et al. (1991) as cited in Chen, Varley, & Chen, 2011). Cervero (1994) showed that 42.3 percent of trips came from no-vehicle households and 3.5 percent was from households with three or more vehicles. Apparently, there was no indication of a relationship between rail usage and household income. Nevertheless, a study by Wardman (2007) produced contradictory results to the common studies on income. Income has an expected positive effect on rail demand, with elasticity 0.52. However, in this study, car ownership, which often has positive association with income, led to fewer rail trips, with elasticity -0.79. With the inclusion of two demographic variables, the level of car ownership and the number of jobs located within station catchment, Blainey (2010) reported that both variables were statistically significant, but that they only added minor improvement to the model fit (absolute-changes 1.5% of R-square or 4% or relative changes).

The 1992 California transit surveys found that, in both high-housing-cost residential TOD areas and low income residential TOD areas, ridership decreased compared to the average 1990 Census data. The suggested explanation was that the increased rates of walking in the former areas and the excellent pedestrian facilities accompanied by the short drives to areas of employment in the latter areas (Evans et al., 2007).

The 1990 US Census and 1990 Nationwide Personal Transportation Study reported that transit users who were: workers age 17-29; those with no car; those with no education background; black people; women; those with immigrant status and living in US less than 5 years; those with an income level ranging from $5k to 10k per year, and those having mobility limitation, were more likely than average to use transit as their principal mode for commuting to work in U.S metropolitan areas in 1990 (Fleishman et al., 2007). Wardman (2006) reported, based on business trip model, that men make 60% more business trips than women and those aged over 50 make fewer trips. All household types make fewer business trips by rail than single person households. In general, residential characteristics that influenced rail share the most consisted of African-American ethnicity, small size of household, middle age, and working in clerical or sales positions (Cervero, 1994).

Gutiérrez et al. (2011) predicted the monthly station ridership by using population and employment variables as predictor variables to describe service area characteristics. The population was segmented into groups of workers, foreigners, population under 20 years old, population over 60 years old and non-car owning households. Similarly, employment was divided into groups by sector, such as commercial, administration, education, health and industrial sector. Variables of employment in the commercial and education sectors were highly significant in affecting train ridership. Variables for workers and foreign population also positively influenced ridership. Other socio demographic variables were not statistically significant, including the non-car households, people under 20 years and people over 60 years old, and employment in the industrial, health, and administration sectors.

Wardman (2006) created a workers' structure that was comprised of percent professional/managerial workers, percent non-manual workers, and percent skilled workers. Percent professional/managerial workers and percent skilled workers were shown to have a negative contribution to ridership with elasticity -0.68 and -0.55 respectively, while percent non-manual workers had a positive contribution with elasticity 1.27.

With regard to the effect of the regional economy, Chiang, Russell, and Urban (2011) investigated the impact of volatility in gasoline prices and the general economic downturn on the monthly ridership of the Metropolitan Tulsa Transit Authority. The impact of gas price was seen to be only marginal. While it was suggested by other research that a declining economy (reflected by the issue of food stamps) may explain reduced ridership, this did not happen in Tulsa. A simple explanation was given that unemployed people who tend to receive food stamps have less need to travel to work.

Wardman (2006) hypothesised that the high levels of the growth in railway demand in the 1990s in Great Britain may due to GDP. He found that while GDP was important, other variables such as variations in car times, fuel costs, car ownership, population and a post-privatisation time trend were found to have statistically significant contributions. GDP at a growth of 2.1% had a high elasticity on ridership at 1.95. While prices fell by 2.9%, income had an elasticity of 0.08. The critical importance of GDP to rail demand growth was clear. Car time and fuel costs also produced elasticities similar to the other variables, but lower than the effect of population growth.

2.4.3 The supply side of the system: transportation component

The third key factor affecting ridership, as defined in this thesis is the "transportation component." The influence of public transport infrastructure adds to the attractiveness of public transport, and can be measured by variables relating to land use factors, as explained in Van de Coevering and Schwanen (2006). They have shown that the ratio of public transport to road supply and rail density, and the lower public parking supply in the CBD correlate positively with the total travel distance by public transport. Further, as emphasised by Nolan (2010) and also in Commins and Nolan (2011), the reality that a household's decision to use a car is not only affected by internal factors such as car types and their socio economic/demographic profiles, but also affected by external factors in terms of supply side variables such as public transport provision and the presence of parking restrictions. In line with these, Commins and Nolan (2011) also attempted to analyse the supply side determinants of mode choice, in addition to an analysis of the demographic, and socio-economic aspects of households. Comparing locations between good and poor rail availability, the authors found that those who lived and worked in areas with good rail facilities are likely to use rail to get to work, combined with walking and cycling. It did not matter whether car ownership and work location had been taken into account.

Other studies agreed upon the importance of railway stations, the proximity of living environments to station precincts, and the spatial accessibility of location to the increased use of trains for commuters, for example Cervero (1994), J. Lin and Long (2008), and Limtanakool et al. (2006). Some studies found that these factors seemed to be more important than that of vehicle ownership and household income (J. Lin and Long, 2008) and others found that the elasticity of vehicle miles travelled is higher for these factors than for land use factors such as density and land use mix (Liu, 2011). Cervero (1994) also showed that rail usage share declined by 0.85 percent for every 100 foot increase in walking distance to the rail station. Nevertheless, the relationship between walking distance and train ridership was weak. This may be explained by the fact that the train station in the study area mostly functions as part of a greater commuter system. Stations provide ample park and ride facilities, resulting in a statistically significant share of rail users beyond walking distance.

The policy of expanding metro or train networks stimulates the use of public transit (de Grange et al., 2012). In the case of Metro extensions in the Netherlands, Limtanakool et al. (2006) examined the impact of this extension on the propensity of people to move house or switch mode in the case of medium distance travel to work. Medium and longer distance travel in the study was defined as home-based trips as long as 50 km and over for a one-way trip. They found that, as households had a stronger resistance to move than to switch mode, households would choose to travel longer than to move house, if one of two earners in the family changed job location or if they have different job locations. Train extensions that connected the origin and destination at medium and long distance would therefore increase the share of train use. Therefore, travel distance impacted on the trade-off between the residential choice and mode choice decision.

The effect of parking restrictions is also important for those who work in the city centre, as they prefer to use public transport and are significantly more likely to walk and cycle. Based on a model for forecasting the total number of trips made from local rail stations in England and Wales over a one year period, Blainey (2010) identified a strong association between car park size at stations and ridership, by showing that the number of parking spaces added to the set of predictor variables, increasing the R-square of 5% in absolute-value or 13% in relative value. Rail usage was as high as nine out of ten work trips to the downtown compared to one out of twenty trips for most other destinations or to regional sub-centres (Cervero, 1994).

Gutiérrez et al. (2011) pointed out the importance of feeder buses to station ridership and considered two variables to describe this: the number of urban and suburban bus lines with stops within 200 metres of the station catchment area, and a dummy variable of park and ride spaces. Regression revealed that both variables were statistically significant. However, the value of the beta coefficient was too high and therefore the data was difficult to interpret. Cervero (1994) explained that access to and from rail stations played a role in train usage. Nine out of ten residents who reached the station did so by walking; nearly 10 percent used park and ride within a quarter to half mile distance from their home. At the destination station, almost 75% of workers also walked to their destination from the station, and bus travel was used by more than 20% of workers.

2. 5 PUBLIC TRANSPORT INDUCED AGGLOMERATION

Since the 1920s, agglomeration has been recognised by economists as a general trend for industries to reduce production costs by “physical spillovers” and to increase productivity by "knowledge spillovers" (Marshall, 1920). Agglomeration is defined as the clustering of economic activities that may result from production externalities (or "spill-overs"), transportation externalities, communication externalities, technological externalities and pecuniary externalities (Brinkman, 2011; Fujita & Thisse, 2002) Agglomeration arises due an economic mechanism based on the trade-off between the various forms of increasing returns and mobility costs. Agglomeration may be seen as the underlying mechanism for different types and scales of spatial phenomena in regions and cities. Some examples of these phenomena are the core-periphery structure between the Northern and Southern Hemisphere, regional disparities, the stability of the urban hierarchy within most countries, and commercial clustering, ranging from district to neighbourhood scale. Different types of agglomeration economies may interact at different distance scales, with a particular requirement for spatial proximity (Fujita & Thisse, 2002).

Agglomeration of production externalities explains why commerce may be found to be clustered in global regions, metropolitan areas, or specifically at the neighbourhood or district level (Brinkman, 2011). Technological externalities have been differentiated from pecuniary externalities depending on whether or not the agglomeration was a product of market interaction (Fujita & Thiesse, 2002). Pecuniary externalities result from market interaction when companies, consumers, or workers are involved in trade, mediated by the price mechanism or the market. This often emerges in the form of a "snowball effect" when the number of agents continue to grow and benefit more from larger diversity or either specialization. This cumulative process is a combination of markets achieving increasing returns from monopolistic situation.

Communication externalities may be considered as part of technological externalities. Their geographical cluster has been based on a competitive paradigm for highly specialized, creative and scientific or industrial districts such as garment districts, financial districts, jewellery and advertising districts. Technological externalities emerge as a result of non-market interaction whereas the process affects the utility of agents involved or the production function of a firm (Fujita & Thiesse, 2002).

Despite these different types of agglomeration, agglomeration has common features and principles behind it. First, agglomeration may involve any of these four processes:

mass production; the availability of specialized input services; the formation of highly specialized labour forces and the production of new ideas; and the existence of modern infrastructure. Second, the agglomeration process often results from an increasing return to scale, indivisibilities and imperfect competition, under the competitive paradigm. Third, there is a balance between the “centripetal force” and the “centrifugal force” in an urban system that encourages agglomeration effects (Alonso 1964 and Krugman 1995, as cited in Fujita & Thisse, 2002).

Transport development influences patterns of urban development and location choices for households and firms. These major changes in land use patterns influence the number of trips, destinations and mode choices (Waddell, Ulfarsson, Franklin, & Lobb, 2007). Public transport infrastructure creates the effect of agglomeration through transportation externalities. Urban economic agglomeration arise from the transportation infrastructure since this service provides a facility for sharing and reducing the cost per unit of production through economies of scale (Rawnsley, Davies, Szafraniec, & Ratnam, 2014). Public transport-induced agglomeration economies, which can also be considered to be part of technological externalities, may be defined as the concentration of economic activities and the clustering of offices, shops, entertainment centres, and other land uses that emerge around public transportation stops. The benefits from this clustering are increased efficiency through lower infrastructure costs and reduced labour costs, while more opportunities are created for greater access to specialized labour (Weisbrod & Reno, 2009).

There are two types of transport infrastructures and their corresponding influence on land use: namely, the “strategic” and the “structural” infrastructure. The strategic infrastructure are the major infrastructure projects that are intended to boost urban productivity, due to the change in overall relative accessibility and in location decisions made by households and firms. Structural infrastructure provides the skeletal shape of an urban area and underpins the economic adaptability of a region, for example, by the development of major hospitals, university campuses or arterial roads (Rawnsley et al., 2014).

Agglomeration economies created by transportation infrastructure are affected by three features of cities: the total number of jobs; the density of jobs; and the transport connection that existed between those jobs. Weisbrod & Reno (2009) stated that the most important issue in economic impact analysis of transportation is to understand how investment leads to real changes in the economy, either for agents (firms, employee, households) or for areas.

The conceptual benefits from transportation infrastructure translate into a form of urban agglomeration economy through four mechanism (Rawnsley et al., 2014):

- Reduction in the cost of transport that leads to lower production costs of firms through lower supply chain costs and higher profits and income.
- Increased productivity of firms allowed them to penetrate to larger markets.
- Higher incomes and lower prices lead to higher demands for these goods/services.
- Improved accessibility changes demand for land in some particular areas and leads to high-rise building (commercial and residential developments).

There is a geographical extent of agglomeration that has been seen as important for policy decisions that attempt to include the effect of agglomeration for facilitating their goals. In agglomeration studies, the effect of agglomeration of employment on urban wages is dictated by the distance to workers' place of work. For example, wages increase by 1.5% to 2.14% for an additional 100,000 full-time workers within 5 miles of their place of work, but fall sharply thereafter (Rosenthal & Strange in Graham & Melo, 2010, p. 13). Di Addario and Patacchini (2008) followed Rosenthal and Strange's approach and estimated that an increase of 100,000 inhabitants within 4 kilometres should raise wages by 0.1-0.2%, but the increase falls sharply thereafter. The impact of urban size on wages is found to be statistically significant only up to 12 kilometres, whereas the average radii of Italian local labour market is 14.7 kilometres (Graham & Melo, 2010, p. 13).

In the case of effective density, as measured by real network travel time, the geographical scope of productivity-employment accessibility effects can extend up to 6o minutes travel time, but that the magnitude of the productivity-agglomeration effect decays very rapidly with time and is very strong within 20 minutes driving time. Doubling the number of jobs accessible within 20 minutes of driving time leads to an increase in real average wages of 6.5%, while the impact for a similar increase within 20 to 30 minutes is as small as 0.5% (Melo et al., 2013).

By a similar principle, the public transport-induced agglomeration effect can explain how public transport improvement creates a more dense and intensive development of activities (land use) around public transport facilities. This implies that the spatial proximity and distance decay principles may be attributable to the characteristics of agglomeration. That is to say, the magnitude of one measurement in an area will influence the magnitude of the same measurement in nearby areas, and that the effect of agglomeration economies decreases with distance from the transport facilities.

2. 6 LITERATURE GAPS

Major LUTI research extensively discusses the influence of land use on travel behaviour but there is a lack of research in the reverse relationship, i.e. how transportation systems influence land use (T-LU) (Van Acker et al., 2007). Among T-LU literature, some studies emphasize the impact of transportation infrastructure development on land use developments (Cervero & Landis, 1993, 1997) as part of the need to justify investment in the transportation system. Furthermore, a more specific aim of these studies has been to examine how public transport extension relates property values and housing appraisal (Greenberg, 1988; Watkins, 1999). Property valuation is often used in the calculation of value capture mechanisms in the form of taxes (McIntosh, Newman, Crane, & Mourtiz, 2011). Other studies fall in the category of accessibility studies to measure the accessibility impact of transportation infrastructure development. The accessibility variable influences house prices in terms of location advantages (Buxton & Taylor, 2010; Cho, Poudyal, & Lambert, 2008; Du & Mulley, 2007; Kryvobokov, 2007).

A close link has been demonstrated between planning directives (regardless the land use) and the economy (Gonçalves, Portugal, & Nassi, 2009). In fact, the two-way interrelationship between the economy and the transportation system (Hall, 1969; Russo & Musolino, 2012) defines how the development of transportation infrastructure will contribute to the economic system, yet the growth of economy activity requires additional supply via the transportation system. Lakshmanan (2011) pointed out three approaches in which transportation infrastructure contributed to the economy, which had previously been discussed as part of T-LU studies. These approaches were: a traditional cost-benefit assessment; the measurement of externalities of transport infrastructure investment in terms of economy-wide cost reductions and output expansions; and the analysis of the full effects of transport infrastructure on the growth of the total factor productivity (TFP) of the economy. This TFP often measured the effects of transport in terms of urban activity clusters resulting from spatial agglomeration and the innovation/commercialization of new knowledge. Nonetheless, current literature has not paid sufficient attention to understanding the impact of transportation, land use, and economic integration on travel behaviour.

The missing component that enables the merging of the two-way relationship of transportation and land use, and the two-way relationship between transportation and economy, is the agglomeration effect of transportation infrastructure development, which is also assumed to eventually influence travel behaviour. While urban activity clusters resulting from spatial agglomeration have been commonly related to the density concept of land use (resulting from the improvement of accessibility), how agglomeration or effective density and density by itself are related or differentiated from each other or how they impact on travel behaviour has not yet been explored.

In the context of train ridership prediction, there is little or no knowledge about the application of effective density in comparison to, or complementary to, the density concept. Variable effective density that incorporates accessibility measurements has been assumed to carry a higher degree of spatial variation than the variable of density alone. Changing analysis from density to effective density may reveal a larger degree of spatial variation in model findings, and would thus provide a more flexible option for a greater variety of land use-transport policy options.

The geographical extent of agglomeration has been an important aspect in the literature, and related knowledge has been useful in assessing the ability of different policy scenarios to make use the benefit of agglomeration (Melo et al., 2013). Similarly, in LUTI studies, train ridership has often been assessed based on the fact that ridership decays with distance from train stations. Concentration of development near stations such as TOD produces a statistically significant increase in ridership

(Cervero, 2006; Cervero & Duncan, 2006). Studies consistently show that transit usage decays exponentially with distance from a station (Bernick and Cervero, as cited in Cervero, 2006, p. 286). Aggregate studies in this thesis call for an adaptable distance band suiting the commuting pattern between suburbs (i.e. medium distance travel) and park-and-ride ridership to accommodate the effective density phenomena. Instead of using fixed distance bands typical of other studies, or arbitrarily-assigned distance bands, the interaction between suburb distance to train station and the effective density have been applied here. This method is relevant, especially with regard to assigning distance decay from the agglomeration source (the train station) so that this may be modelled as a continuous rather than discrete entity.

2. 7 CHAPTER SUMMARY

The contribution of this thesis is to improve the LUTI framework using a modified SETI-LUTI framework. This framework considers the effect of public transport-induced agglomeration to influence the two-way LUTI relationship. The interaction between land use, transport, and spatial-economic factors may be explained with three interrelated concepts:

a. Transportation infrastructure influences land use: transport investment, such as train station construction, influences accessibility by public transport and the level of density and effective density.

b. Transport-induced agglomeration effects: the transport-land use relationship influences urban productivity such as land rents or property values, the income of workers, and the wages of employees (employed non-resident).

c. Land use influences travel behaviour: the relationship between transport-induced agglomeration effects may be viewed with respect to the role of agglomeration in increased train ridership.

This thesis attempts to fill the research gap in the literature by understanding how public transport-induced agglomeration influences the level of train ridership. It is assumed that there is an interaction between the two sides: the public transport facilities and the economy. Increases in urban productivity by transport facilities can reinforce transport usage, such as train ridership. This study will not measure the wider economic impacts of agglomeration economies for businesses, but will be limited to its impact on employees (employed residents and job numbers) in terms of effective density. These effective density influences are assumed to facilitate the effective use of trains as a mode of travel to work and will be modelled endogenously. Chapter 3 provides detailed explanations of the proposed research framework where a train ridership model will incorporate the parameters of public transport induced agglomeration.

CHAPTER 3. RESEARCH METHODOLOGY

3. 1 INTRODUCTION

Spatial economic impacts of the transportation system have been discussed in other research fields such as in agglomeration studies, within the domain of urban economics. However, there has not been much research dedicated to studying the influence of agglomeration on travel behaviour responses, such as train ridership. This thesis proposes a model to understand the impact of public transportation development, such as railway line extension, on train ridership responses, through induced agglomeration in terms of effective density.

The hypothetical reasoning for this thesis is discussed in section 3.2. The attributes of the proposed study framework are then discussed in section 3.3. Methods to test these hypotheses are developed in section 3.4. Furthermore, causality relationships are proposed as a function of sets of independent variables (all component of land use, transportation system or accessibility, spatial economic, and socio demographic/economic) on the dependent variable (train ridership) in sections 3.5 to 3.7. This thesis uses two data structures: the first is based on an aggregate dataset defined at the suburb administrative level; and the second is the derivation of a fishnet data structure for a more detailed measurement of some variables, which is explained in section 3.8. Descriptions of the study area, the determination of industry sectors, data, sources and software requirements are described in sections 3.9 to 3.12.

3. 2 RESEARCH HYPOTHESISES

Train ridership prediction is a specific task in travel behaviour modelling. Standard urban travel demand modelling has shown that socio-demographic factors and land use factors largely influence demand (Ferguson, 2000). For example, the densification and intensification of land use and population, measured by various indicators, have been shown to influence travel demand (Chatman, 2008).

The literature gaps identified in chapter 2 suggest that there are spatial economic effects central to the densification concept that has been rarely accounted for in the modelling of travel behaviour, including the case of train ridership prediction. A study performed by Graham and Melo (2010) is one rare example of study in this area. This spatial economic effect has been termed “effective densities induced through transport investment” (Graham, 2007, p. 337). Improvement of transport infrastructure is believed to induce more development around this infrastructure (Blainey, 2010). These densification-agglomeration snowball effects need to be considered in the modelling of travel behaviour, whereas current literature on agglomeration studies have been more focuses on their impact on urban economics or urban productivity. Ignoring this phenomenon potentially leads to a bias toward underestimation of train ridership, since these agglomeration effects potentially act as exogenous factors that may contribute to the magnitude of errors (unobserved factors) in the modelling. This thesis therefore attempts to propose a modified SETI-LUTI framework to model the rail demand. The aim of the proposed model has been to accommodate the dimension of public transport induced agglomeration in train ridership prediction and to determine the extent that this factor influences the level of train ridership.

The first research hypothesis has been derived to propose a link between transportation infrastructure development and agglomeration, referred to as public transport induced agglomeration. That is, “the intermediate stage which attempts to understand the relationship between transport accessibility and sources of agglomeration” (Graham & Melo, 2010, p. 19). Thus, the research hypothesis proposes an underlying reasoning for the evidence of public transport induced agglomeration by assuming a distance decay and concentration pattern of agglomeration according to the distance from train stations.

The public transport induced agglomeration concept was initiated by Venables (2004) and further developed by Graham (2007). The study of Venables (2004) highlighted the relationship between dense spatial units and the clustering of economic activity with an increase in productivity. This concept was explained in the following figure 3.1. Venables' original concept explained the relationship between wages, travel costs, and land rent or housing costs and its impacts on city size. The Y axis plotted unit costs or unit benefits. Unit costs and benefits could be represented as wage level, commuting costs, and/or land rents. The X axis plotted the number of workers and also represented the distance from the zero point (the CBD or centre) to other locations at increasing distance from the CBD.

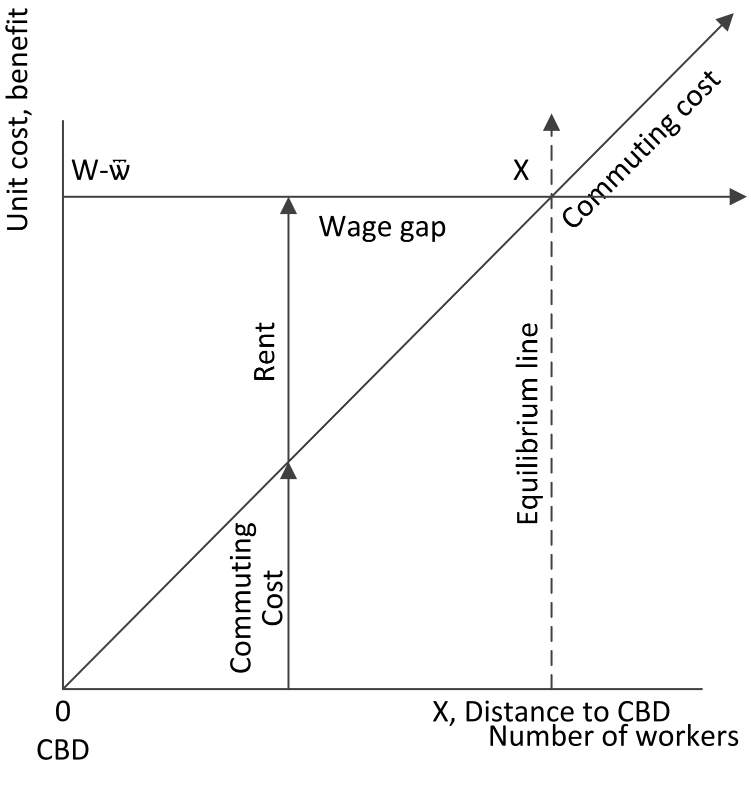


Figure 3.1 Urban equilibrium: agglomeration economies from transport infrastructure (Venables, 2004, p. 19)

Figure 3.1 represented the trade-off between land rent or housing costs and commuting costs as: "*workers located closer to the CBD face lower commuting costs but higher rents, as given by the distance between the horizontal line (W-$\bar{w}$) and the commuting cost curve*" (Venables, 2004, p. 9). There was a balance in the trade-off between land rents and commuting costs, where the increase in commuting costs by moving along the X axis was off-set exactly by the same amount of the decrease in land rent. Thus, there was a rent gradient, with a converse commuting cost gradients. "*The size of the city was determined at the point X, where the wage gap was equal to the travel costs of the most distant city worker. At point X no further workers want to be employed in the city*" (Venables, 2004, p. 9). Apparently, to be able to work at the CBD, workers lived at point *X* need to gain a premium wage which was not in evidence in reality. Nevertheless, the equilibrium condition required the wage gap between locations to be almost invariant, and thus workers were indifferent between living in the city or non-city locations.

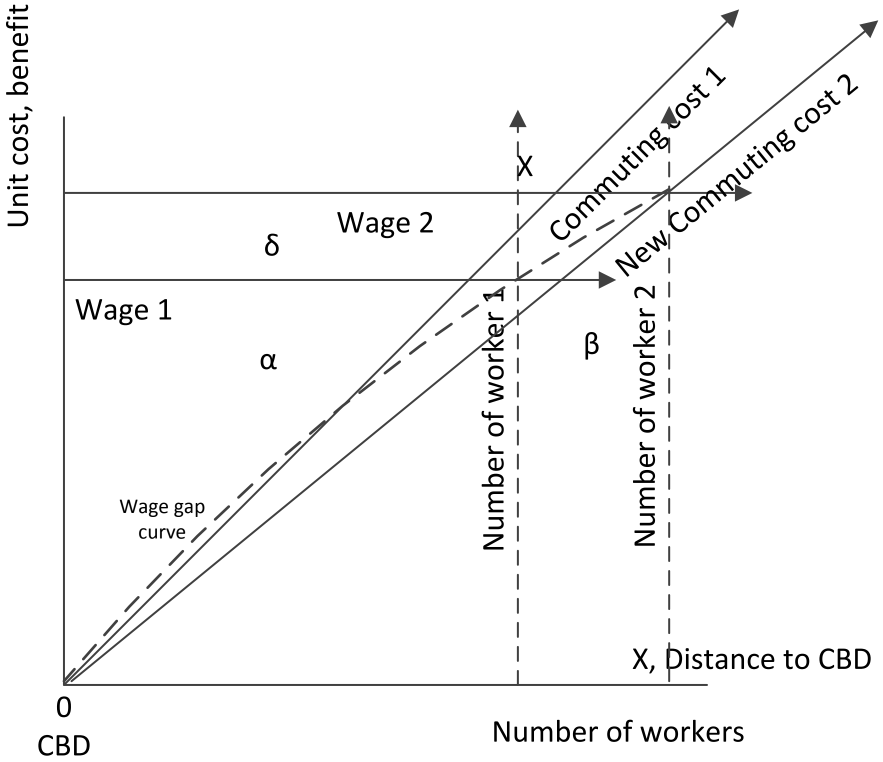


Figure 3.2 Net gain from transport improvement with endogenous productivity (Venables, 2004, p. 19)

Further, figure 3.2 described how an endogenous productivity effect could be added to the framework. In Venables' model formulation, transport improvement was modelled as a reduction in commuting costs, which derived a further effect on the real income. As workers were able to save more money (a conversion from travel costs savings) workers could afford to live closer to the centre. Transportation externalities created therefore increased urban productivity, which in turn increased the city size. The agglomeration benefits, represented as the 'convex wage-gap curve' instead of 'constant curve' in figure 3.1, would bring two consequences: "*First, the change in city size from the transport improvement is larger. Second, the higher productivity of existing city workers raises real income by area* δ *or an induced productivity gain. The real income gain now is* α + β + δ" (Venables, 2004, p. 10).

The indirect productivity in terms of the incomes of workers and land rents was derived from the "urban equilibrium" concept. The classical urban economics theory, as stated by Alonso 1964 (re-explored in Fujita and Thiesse 2002, as cited in Venables (2004)) stated that urban equilibrium was achieved if: *"the city expands up to the point at which these are high enough that a worker is indifferent between locating at the edge of the city and commuting to the CBD, or living (and working) in a non-city location*"(Venables, 2004, pp. 3-4).

This thesis adapts the Venables model by following Graham's interpretation of the model. Graham (2007) used Venables' framework to further calculate the agglomeration magnitude of employment in terms of accessibility or opportunity. The income gain from transport improvement (δ) was regarded by Graham as a type of measure of the elasticity of urban productivity with respect to the city size. Therefore, Graham interpreted agglomeration as defined in Venables, to be quantified from elasticities of urban productivity with respect to some measures of urban density. Graham assumed that transport investment may or may not increase densities or increase the city size. The measure of density argued by Graham incorporated an implicit transport dimension in terms of travel distance. Graham, after Venables, modelled agglomeration economies using a measure that incorporates both proximity (accessibility) and the scale of economic activity (the size of employment), what so called as the effective density (Graham, 2007; Venables, 2004). The total effective density of employment defined in Graham is a measurement of agglomeration. In this thesis, adjustment of the transport dimension was made by changes to or replacement of the travel distance with the travel time (further shown in equation 5.1 of chapter 5), which is a more explicit measurement of accessibility (Graham & Melo, 2010). The formula in Graham (2007, p. 327) was specified as follows:

$$ED_{im} = \frac{E_i}{\sqrt{(A_i/\pi)}} + \sum_j^{i \neq j} \frac{E_j}{td_{ij}^{\alpha}} \quad \text{Equation 3.1}$$

Where:

ED_{im} = the employment density of ward *i* for industry in sector *m*.

E_i = the number of employment of ward *i*

A_i = the land area of ward *i*

E_j = the number of employment of ward *j*

td_{ij}^{α} = the travel distance between ward *i* and ward *j*, weighted by the distance decay parameter α.

The important concept in Venables and Graham that is adapted in this thesis was the introduction of effective density concept. As pointed out by Venables (2004), there were two mechanisms by which transport infrastructure may affect productivity. *First*, transport infrastructure improved links between firms that in turn raised the effective density of the cluster (jobs or employed residents). *Second*, transport infrastructure relaxed constraints on access to the centre, where this increased the overall city employment. The agglomeration externalities were related to these two mechanisms. Failures to include these effects in a transport appraisal were considered to undervalue the benefits of transport.

Hensher et al. (2012) pointed out that the effective employment density may be defined not only in the terms of the actual (physical) number of employment in various locations, but also in terms of their relative positions with respect to a particular reference point (for example the central city or CBD). Even without any changes in the employment numbers, improvement in a transport system can impact the effective employment density once travel times are used to indicate the relative positions of these employment numbers with respect to the reference point. This thesis assumed that when public transport agglomeration is a function of an improvement in travel time due to railway line extension, a train station can be viewed as a reference point. A reference point in Venables' concept is regarded as a CBD, while in the context of public transport agglomeration applied in this thesis, it is a station. This is because overall travel times change due to the presence of train stations in the system following the railway line extension and the development of new stations.

Figure 3.2 attempts to explain the difference in wage levels ('a convex wage gap') as a cause for the distance decay pattern of agglomeration from a location closer to the reference point such as the CBD.

Figure 3.3 explains an adaptation by this thesis of the Venables and Graham framework on the transport induced agglomeration. First of all, this thesis makes an adaptation by assuming the centre point location 0 in figure 3.2 is not a CBD, but a train station. The edge of the city (point *X*) is then interpreted as the boundary of the catchment of the train station (see section 4.1.2.15 in chapter 4 for determination of the edge or the size of a catchment area). The *Y* axis then also represents the magnitude of agglomeration that also represented the potential of travel demand, where its value would peak at the zero location (station precinct). Note that the location *a* is added to the graph in order to better illustrate the distance decay pattern. New (improved) travel times that reduced travel costs would be translated into real income (Venables, 2004), thus increasing the affordability of longer commuting, increasing the distance from the edge to the centre, and increasing the number of workers or level of employment, which therefore may expand the catchment radius. In addition, improved services and infrastructure of public transport would increase the attractiveness of public transport. More workers may be willing to travel longer distances to use train services. Therefore, the catchment area of the train station would be larger.

This thesis attempts to apply the trade-off concept, specifically in terms of how the trade-off mechanism explains rail demand in terms of the size of potential demand or effective density. In this case, the trade-off is between housing costs and transport costs as a function of distance to the train station, instead of distance to the city centre or CBD. It is assumed that the railway extension would improve accessibility levels, leading to greater travel time savings. Jobs would be concentrated near stations and the value captured would increase land rents in this area. The premium on wages, land rents, and effective density from travel time saving would follow a distance decay, with similar patterns among each other. When jobs are assumed to be located in the centre (near stations), through the optimization process, the density distribution for housing around the centre would be expected to decline, in turn affecting the total amount of commuting (Cervero, Round, Goldman, & Wu, 1995). This trade-off process would eventually lead to the distribution pattern of job numbers and residential numbers. The travel time saving due to transport cost reduction would be converted directly to income. People living near the edge of the catchment would need to receive higher wages in order to afford commuting to the centre. As they are able to save enough money to afford longer commuting trips, the catchment area would be extended, the employed resident and job numbers increased, thus creating higher potential travel demand for train ridership (figure 3.3).

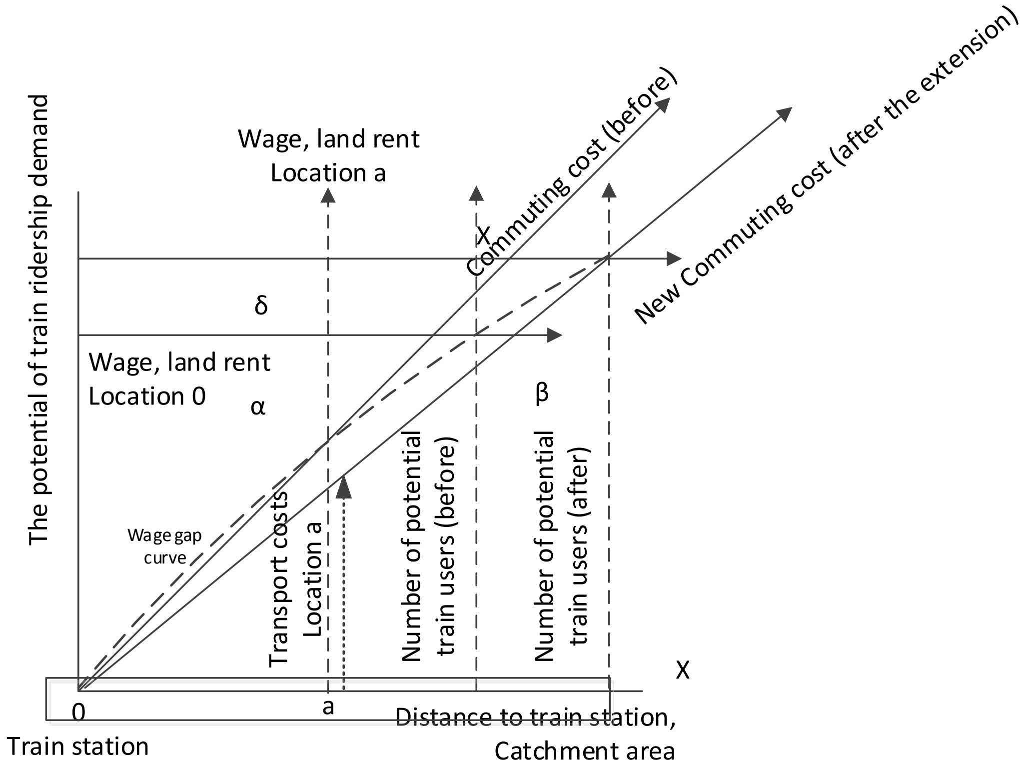


Figure 3.3 Adaptation of the Venables's framework (2004) to public transport induced agglomeration.

Evidence on the presence of the trade-off mechanism are provided in the literature, such as how the trade-off between housing costs and transport costs may affect urban sizes. Workers are willing to commute only to the extent that its costs are offset by lower housing costs. So (2001) explained that, holding housing prices fixed, increasing commuting time required wage increases in the metropolitan market of between 7.7% to 13%. Rosenthal and Strange (2004) found that a 10% increase in commuting time lowered the non-metropolitan population by 1.1%. An increase in 100,000 residences within 4 kilometres of employment centres also raised wages by 0.1-0.2%.

The adaptation of Venables (2004) and Graham (2007) made in this thesis, such as this thesis assumes these reference points to be train stations and that the decay pattern would emerge between stations and locations farther away. This new assumption leads to **the first research hypothesis** for the basis of understanding **public transport induced agglomeration**: *The first hypothesis states that there is a distance decay and concentration pattern (clustering effect) of agglomeration according to the distance from train stations. The closer a location is to a train station, the higher the concentration of agglomeration, and the further a location is from a train station, the lower the agglomeration. Improved accessibility relates to a lower value (less negative) of the distance decay parameter of agglomeration. A less-negative distance decay parameter will be associated with higher agglomeration values and a higher chance of train ridership, following Graham's (2007) assertion that there is a direct association between the level of location accessibility of and the level of potential of transit demand.*

Public transport induced agglomeration in terms of effective density will only be demonstrated if the spatial interaction or travel demand increases or decreases in response to changes in travel times. Graham and Melo (2010) provided evidence on the travel time decay of business trips and commuting trips to infer potential effects of new transport infrastructure (such as an extension of public transport facilities) and to ultimately determine the size of agglomeration benefits. The change in the effective density resulting from a reduction in travel time (through an increase in the average speed after a transport investment was incorporated via a decreased distance decay gradient. Thus, the improvement in travel times reduced the spatial impedance and other obstacles to travel. The level of spatial interaction that was determined by this decay parameter was also assumed to apply in relation to public transport induced agglomeration, i.e. the decay parameter was revised after a transport intervention.

The second research hypothesis developed in this thesis relates to the concept of **agglomeration-induced train ridership**. This thesis seeks to establish reasoning that links transportation system development to agglomeration and travel behaviour (train ridership) following from the first hypothesis.

Graham and Melo (2010) suggested that investigation of the changes in trip patterns may indicate the existence of public transport induced agglomeration. New development of transport infrastructure was proposed to increase spatial interactions (such as railway trips) and these spatial interactions would become more efficient due to travel time savings following the development, thus increasing urban productivity. Therefore, Graham and Melo (2010) estimated the time decay gradient of commuting trips, segregated by trip purpose. These time-decay gradients that connect firms and workers may be used to infer the potential effects of transport investment. Using empirical estimates, they provide indications of the potential magnitude of agglomeration benefits of a transport project, such as high speed rail. In this thesis, a similar approach to the estimation of the time decay of agglomeration from spatial interactions is used. However, the resultant effect of the estimated decay gradient on effective density was used as a predictor of train trip production and train trip attraction. Thus, this thesis assumes that public transport induced agglomeration would facilitate greater ridership, where after a few years following railway expansion, it is assumed that intensive economic growth around train stations would induce a higher demand for train ridership.

In the context of this thesis (the railway line extension), effective density as a form of public transport induced agglomeration is assumed to influence train ridership. The agglomeration mechanism implies that locations of clusters of employment are more attractive since they offer better accessibility to other jobs or employed residents once the transportation system has changed or improved (AHURI, 2006). In this sense, the spatial concentration of economic activity creates positive externalities via agglomeration economies. Transport improvement induces or results in agglomeration or spatial concentration of economic activity.

This thesis therefore derives a second hypothesis based on the assumption that higher agglomeration is associated with spatial concentration of economic activity, which means higher density and scale of activity and geographical proximity to activity. Both factors (scale and accessibility/proximity) are assumed to directly influence travel behaviour. Transportation development (such as infrastructure) influences urban patterns and locations (Waddell et al., 2007), and urban patterns affect the available choice of transportation in turn (Hall, 1969). Therefore, agglomeration or effective density or employment accessibility may be regarded as one determinant of train ridership. Based on equation 3.1, lower travel impedance (by public transport) between places associated with higher zonal accessibility (by public transport) may be translated into higher willingness to travel (by public transport). The original concept of Venables (2004) and Graham (2007) on the trade-off between housing costs and transport costs, based on the level of wages, is used to explain the size of urban areas, hence indirectly referring to the size of the population and the potential travel demand.

This thesis models train ridership prediction based on **the second research hypothesis named as the agglomeration induced train ridership**: *There is a strong influence of employed resident numbers and job numbers (the scale or size of activities) on train ridership. The higher the number of employed residents or jobs, such as measured by density, the higher the chance that train ridership will be generated or attracted in a location (the LUTI model). In terms of agglomeration (SETI-LUTI model), the number of employed residents and jobs are defined not only by the scale or size of activity but also in terms of accessibility or as effective employed resident density or as effective job density. Therefore, the higher the effective employed resident density and effective job density, the greater degree of train ridership will be generated from residential areas (place of residence model) or attracted into workplace areas (place of work mode).*

This thesis assumes that public transport induced agglomeration determines the spatial concentration of employed residents and the spatial concentration of jobs. Different locations around a station should experience different levels of spatial concentration due to a distance decay agglomeration pattern measured away from stations (hypothesis 1). **The third research hypothesis** relates to the influence of distance of a location from train stations, referred to as the "**geographical extent of agglomeration**". The distance from train stations should influence the strength of influence of agglomeration on train ridership in different ways. Following the second research hypothesis, the influence of agglomeration is tested in two different models, the train trip attraction (place of work) and the train trip production model (place of residence). Thus, the distance of employed residents to boarding stations and the distance of jobs to alighting stations have been developed as separate variables that are assumed to influence agglomeration induced ridership. **The third research hypothesis** states that *a greater distance of employed residents from boarding stations and of jobs from alighting stations results in a lesser influence on agglomeration induced train ridership.*

This thesis assumes that stations at places of work (alighting stations) have different characteristics to stations at places of residence (boarding stations), for example, in the difference in their potential to generate agglomeration and train ridership level. Alighting stations are mostly located at the city centre, where the catalyst function of the station to integrate with adjacent activities may be explained by the centrality of its location (Gonçalves et al., 2009). This thesis tests the idea that suburbs near alighting stations would have higher numbers of jobs and public facilities. Indicators which compare the level of jobs to the level of residential activities are a measurement of self-containment/sufficiency characteristics. Self-containment/sufficiency represents the tendency of people to live or work in the same area or different areas, and thus determines the level of overall travel demand (Cervero & Duncan, 2006). The self-containment/sufficiency concept may also be measured using indicators of the job-housing balance, which is part of the spatial structure measurement.

Thus, the **fourth research hypothesis,** which is based on the concept of **job-housing balance**, states that: *If the number of employed residents and jobs in a location is balanced, there is more chance that people will work and live in the same (suburb) area. The more balanced the ratio of job to employed residents is (for example,* *between 0.75 – 1.5), the lower the number of spatial interactions (trip numbers) are between residential and workplace suburbs. The higher the ratio is (jobs are much higher than employed residents), the greater is the level of trip attraction to workplace suburbs. The lower the ratio is (employed residents are much higher than jobs), the greater is the level of trip production from residential suburbs.*

3. 3 THE PROPOSED RESEARCH FRAMEWORK

This thesis develops a theoretical framework to understand the underlying relationship between land use and transport when the spatial economic dimensions are incorporated into the LUTI framework. Many efforts towards this integration process have been made in the last decades (Russo & Musolino, 2012). Mainstream studies developing the SETI framework have focused on the economic impact of new or improved transportation supply systems. These were motivated by the continuous pressure to reduce public expenditure on public transport infrastructure - significant spending on public transport investment needs to be justified (Miyagi, as cited in Russo & Musolino, 2012, p. 190). Studies on the value capture of transportation, agglomeration economies and the transportation “spatial spill-over” were incorporated into one component of the SETI system in order to assess the wider economic impact of transportation (Russo & Musolino, 2012). However, many of these studies examine the economic effects of transportation as feedback to the financial system of public transport expenditure, not train ridership itself.

The broader idea of an integrated framework between the two-way relationship between spatial economic and transport systems at both the national and urban scales has been developed by Russo and Musolino (2012), who proposed the unifying of ‘SETI’ or ‘Spatial Economic Transport Interaction’ (figure 3.4). The attempts to understand the link between transport and economic growth had been made possible by identifying the multiple causal mechanisms between the two. These consist of market expansion, gains from trade, technological shifts, processes of spatial agglomeration and processes of innovation and commercialization of new knowledge in urban clusters that were encouraged by transport improvement (Lakshmanan, 2011). This thesis adds agglomeration into the LUTI framework and terms this the modified SETI-LUTI model.

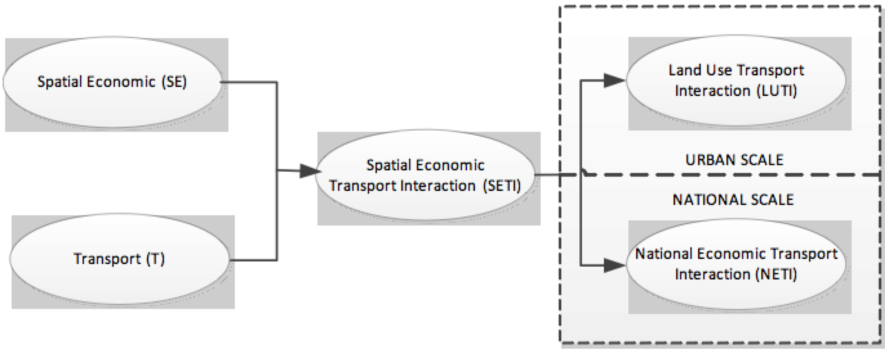


Figure 3.4 SETI models: urban and national specifications (Russo & Musolino, 2012, p. 191)

The link between transport and urban growth is two-way: transportation at any point in time will affect subsequent urban growth, and the urban pattern affects the available choice of transportation in turn (Hall, 1969). The addition of the spatial economic dimension is proposed in this thesis as a tool for bridging the connectivity between the impacts of transportation development on land use and land use on travel behaviour.

Specifically, this thesis follows the concept of travel time saving to explain the interrelationship between transportation-land use-spatial economic and travel behaviour. The process of how travel time savings emerge following public transport extensions has been explained by two mechanisms (Cervero & Kang, 2011). *First*, improved access increases the service quality of public transport and reduces travel time. There would be a form of travel time saving due to this time reduction. This then triggers land use changes, i.e. the increased demand for residential housing and other activities along the area of improved public transport services. These increasing demands lead to more intensive activities. Then, these intensifications are translated into higher real estate prices by market mechanisms. *Second*, there would be value captured in the form of benefits assessments; the improved access would create benefits for commercial sectors around the transport facilities. Therefore, the benefits may be calculated as the amount of taxes paid by commercial sectors. These tax income increases would create pressures on the market to build with a higher density along the service corridors, creating job opportunities.

In relation to the train ridership model, density will be compared/combined with effective density. A direct association has been shown between accessibility levels of a location and the magnitude of potential for transit demand (Graham, 2007). Therefore, if the economic benefits of transport are to translate into effective density, one can examine the effect of job supply or the number of employed residents in an area on the potential for train ridership.

The emphasis of this thesis is to examine in more detail one of sub-components within the SETI framework (i.e. the LUTI) as proposed by Russo and Musolino (2012). Figure 3.6 shows the specific contribution of public induced agglomeration in train ridership modelling based on the SETI-LUTI model. Comparing figure 3.5 to figure 3.6 demonstrates the modification that the SETI-LUTI applies to the LUTI model. There are several factors relevant to agglomeration that may be used to extend the LUTI model (Figure 3.6). These consist of (1) the component of public transport induced agglomeration, such as the existence of distance decay patterns from stations, measured by the clustering effects and the decay effects. Distance decay and clustering are then assumed to be relevant spatial factors when agglomeration emerges as part of railway line and station development. (2) the component of agglomeration induced train ridership, where agglomeration is defined as effective density consisting of a scale and proximity sufficient to influence the level of travel demand. Land use density (based on the LUTI concept) is expressed as the effective density in the SETI-LUTI extension model. Transportation development and the accessibility concept (based on LUTI) is expressed as a distance proximity in the SETI-LUTI extension. Also included are components of urban *productivity* (indirect or partial) in terms of land rent and wages, and the agglomeration magnitude itself, in terms of effective employed resident density and effective job density. These are modelled assuming the trade-off mechanism (as discussed in section 3.2).

The LUTI model is a subset of the framework of the SETI model as pointed out by Russo and Musolino (2012). All components that are involved in the framework the LUTI model can also be included in the framework of the SETI-LUTI model.

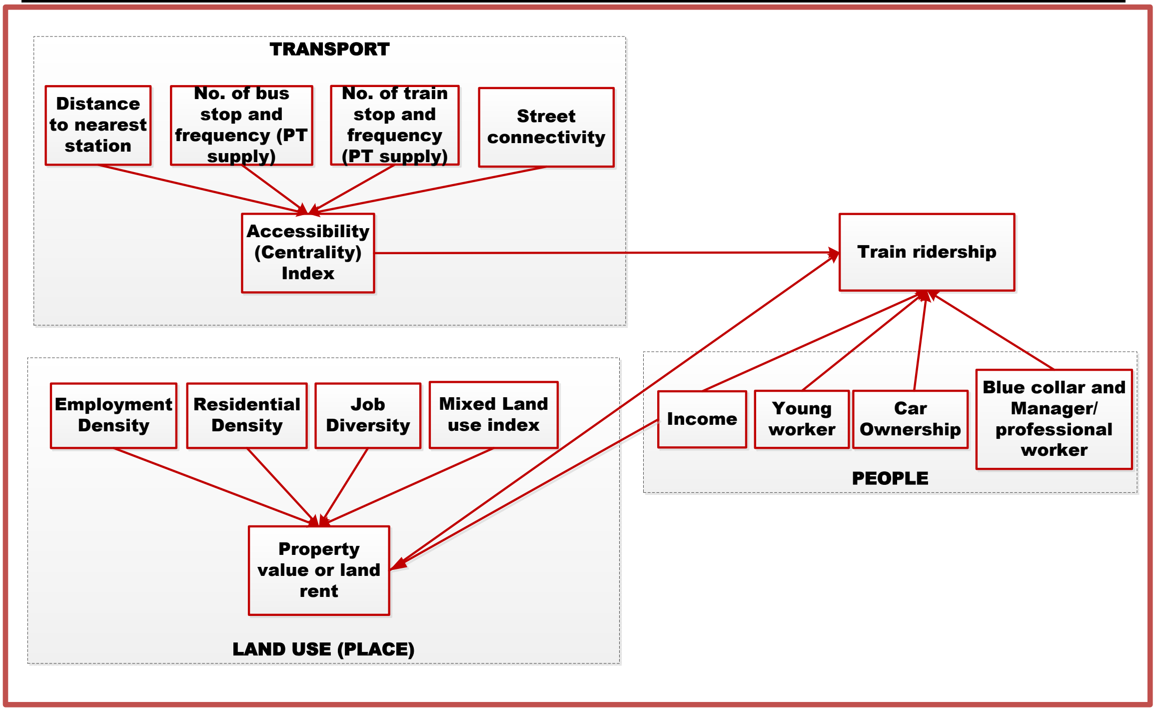


Figure 3.5 LUTI framework in modelling train ridership

The variables derived from this proposed SETI-LUTI framework which have been regressed on the model of train ridership consist of:

(a) Effective employed resident density, as an extended measurement to employed resident density in the LUTI model.

(b) Effective job density, as an extended measurement to job density in the LUTI model.

(c) Clustering as a spatial attribute of agglomeration in the SETI-LUTI model.

(d) Distance decay pattern of agglomeration, as a spatial attribute of agglomeration in the SETI-LUTI model.

(e) Job-housing balance measures the self-containment/sufficiency to indicate (1) if a suburb is more of a residential or workplace area; (2) if a station corresponding to that suburb is more of an origin (boarding) station or a destination (alighting) station (3) whether the level of spatial interaction from/to that suburbs will be high or low.

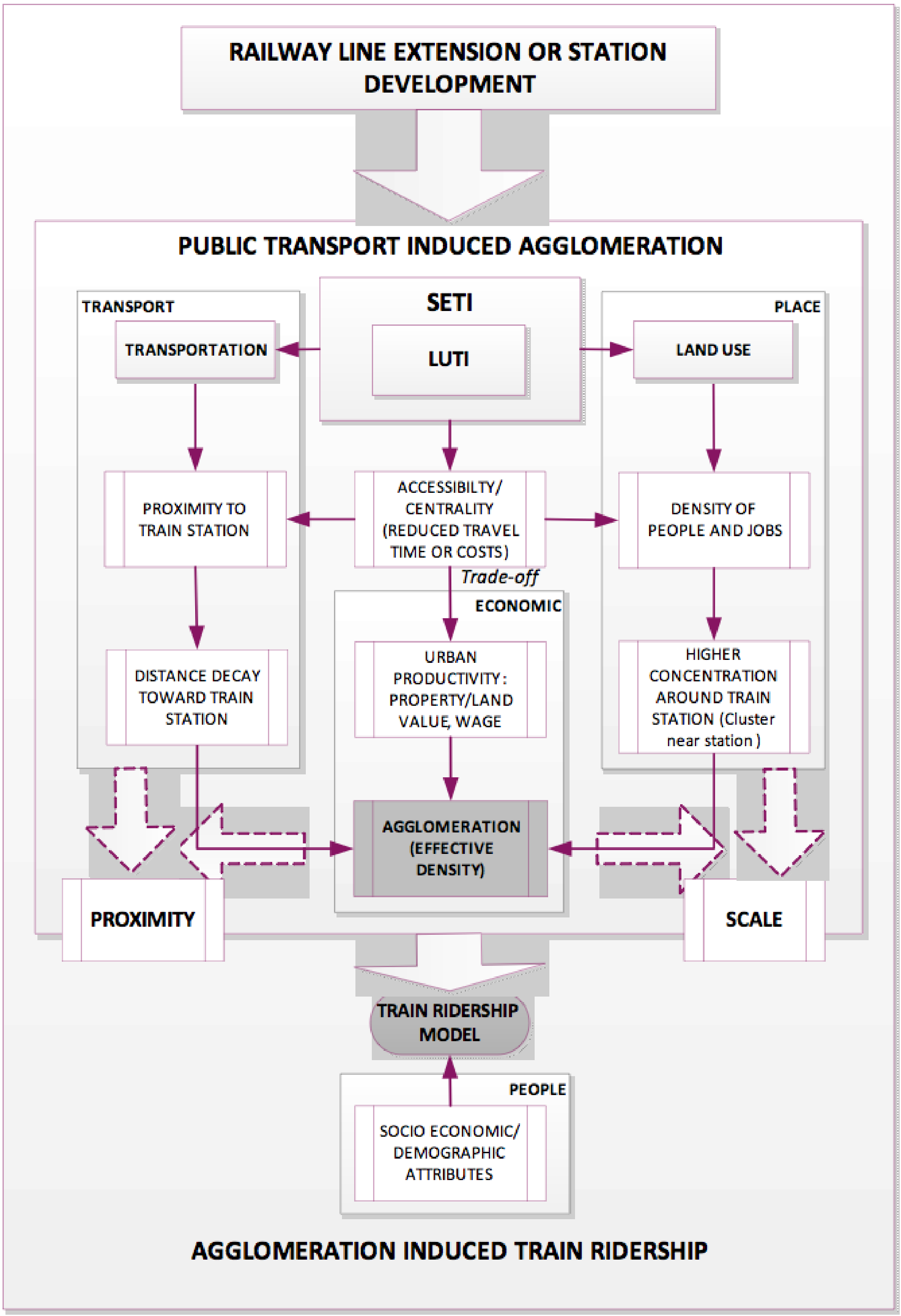


Figure 3.6 The LUTI-SETI framework in modelling train ridership

Knowing this information, such as the dominant activities (residential or non-residential), in what sector, at what scale and where, would be expected to inform the most important factors controlling the pattern of transportation within a city region (Hall, 1969).

3. 4 METHOD OF ANALYSIS AND RESEARCH STAGES

The previous sections described the characteristics of the proposed SETI-LUTI framework and research hypotheses based on the available theoretical concepts about public transport induced agglomeration and its modification for this thesis. This section discusses the method of analysis and research stages that have been conducted in order to examine each of the research hypotheses. The overall research methodology framework is presented in figure 3.7.

The following table summarises the methods of analysis.

Table 3.1 The methods of analysis and research stages based on research hypotheses

No.	*Research hypothesis*	*Research Stages and Method*
1.	Hypothesis 1: *There is a distance decay and concentration pattern (clustering effect) of agglomeration according to distance from train stations. The closer a location is to a train station, the greater is the concentration of agglomeration, and the further a location is from a train station the lesser the agglomeration, hence creating a distance decay pattern of agglomeration. Improved accessibility relates to a lower value (less negative) of the distance decay parameter of agglomeration. A less-negative (low value) distance decay parameter will be associated with higher agglomeration values and a higher chance of train ridership*.	Provides analysis to determine if public transport induced agglomeration exists and varies among different sectors. Stages: Measure the association between the changes in travel time before-and-after the Perth-Mandurah railway line extension. Measure the changes in effective density and the level of train ridership production using correlation analysis and compare them between different railway lines using descriptive statistics. Measure any clustering phenomena using Getis-Ord G* Model the spatial distance decay pattern of agglomeration by building a fishnet dataset (grid cells) and measure the average value of effective density based on a 1 km ring buffer from the fishnet dataset. The curve fitting by exponential distance decay model calculated based on probability density function to show the rate of decay.

2.	Hypothesis 2: *There is a strong influence of employed resident numbers and job numbers (the scale or size of activities) on train ridership. The higher the number of employed residents or jobs, as measured by density, the higher the chance that train ridership will be generated or attracted in a location (the LUTI model). In terms of agglomeration, (SETI-LUTI model), the number of employed residents and jobs are defined not only by the scale or size of activity but also in terms of accessibility or as effective employed resident density or effective job density. Therefore, the hypothesis states that the higher the effective employed resident density and effective job density, the higher level of train ridership will be generated from residential areas (place of residence model) or attracted into workplace areas (place of work model)*	Multiple regression of train ridership prediction by comparing the model with and without agglomeration. Stages: Provide analysis to determine if public transport induced agglomeration exists and varies among different sectors (Stage 1). Specify the distance decay parameter of agglomeration by estimating the travel impedance parameter in the gravity model. Measure the magnitude of effective density. Model train ridership with effective density as one of the explanatory variables in the regression model. Compare the models developed, based on both the LUTI and the SETI-LUTI framework.
3.	Hypothesis 3: *Greater distance of employed residents from boarding stations and of jobs from alighting stations results in a lesser influence on agglomeration induced train ridership.* .	Multiple regression of train ridership prediction by accommodating the interaction variable between the magnitude of effective density and distance between suburb and train station. Stages: Construct grid cells dataset in a form of a fishnet by dividing suburbs into a uniform size of *n* number of grid cells (fishnets) of 1 km times 1 km size. Measure the Euclidian distance from the centroid of each fishnet to the nearest train station's centroid. Calculate the average distance of each suburb by averaging distance across all *n* fishnets for each suburb based on two types of distances: from the residential suburbs to the nearest boarding stations; and from the workplace suburbs to the nearest alighting stations. Establish a continuous variable of the interaction term between the average distance of suburbs from their nearest train station and the magnitude of effective density. Regress the train ridership variable on this variable in a model of train ridership prediction. Determine from the model the threshold distance from stations beyond which the change in effective density will no longer influence the level of train ridership

4.	Hypothesis 4: *If the number of employed residents and jobs in a location is balanced, there is more chance that people will work and live in the same (suburb) area. The more balanced the ratio of job to employed residents is (for example, between 0.75 – 1.5, the lower the number of spatial interactions (trip numbers) are between residential and workplace suburbs. The higher the ratio is (jobs are much higher than employed residents), the greater is the level of trip attraction to workplace suburbs. The lower the ratio is (employed residents are much higher than jobs), the greater is the level of trip production from residential suburbs.*	Comparing the model of train ridership production and train ridership attraction. Investigate the policy implications from model results. Stages: Examine the differences in factors influencing ridership based on train ridership production model and train ridership attraction model. Apply the *job-housing balance* principle to understand the extent of each of the factors (especially socio-demographic and land use variables) in influencing train ridership. Apply the job-housing balance principle to understand the *trade-off* between wages and housing costs/travel costs. Build an equation to explain the *trade-off* mechanism, calculate a travel time saving, and explain why effective density may eventually influence train ridership in three suburbs categories (job-housing balance, job-rich suburbs, employed resident-rich suburbs) based on these trade-off perspectives.

These methods of analysis have been implemented in sequential research stages. The following research stages were conducted:

1. Data preparation (refer to Chapter 4).

   This stage involved three processes: *First,* data cleaning and data transformation to fit the data into the model. *Second,* data aggregation and conversion in relation to preparing the complete data set so that all data have consistent geographical boundary levels between difference data sources. *Third,* identification of the economic sector that contributes most to the economy, thus, may generate the highest potential demand for train ridership.
2. Analysis I: Investigating public transport induced agglomeration (refer to Chapter 5).

   This stage aims to determine if public transport agglomeration exists and if it varies by industry sector. There are three types of analysis conducted to support the hypothesis of public transport induced agglomeration: (1). Estimate the spatial decay (based on travel time) of agglomeration based on each sector by modelling spatial interaction (the gravity model) where the rate in the flow of people between locations is assumed to influence the rate in which the scope of agglomeration decreases with the increase of travel time between suburbs.

(2). Calculate the magnitude of agglomeration of the effective job and employed resident density in each sector.

(3). Investigate the association between the changes in travel time before and after the Perth-Mandurah railway line extension and the changes in effective density and in train ridership production. Compare these changes between different railway lines.

(4). Investigate any clustering phenomena of agglomeration by Getis-Ord G*, in particular to determine if agglomeration hot spots may be found near train stations.

(5). Investigate the distance decay pattern of agglomeration by curve fitting of exponential distance decay to the average effective density values from all fishnets based on probability density function within each 1 km ring distance, up to 16 km rings (as defined for the overall catchment area-determination of the size of 16 km radius of catchment may refer to section 4.1.2.15 in Chapter 4).

3. Analysis II: Use regression analysis to understand how agglomeration varies by sector and their influence on train ridership, based on train ridership production and train ridership attraction. Establish the regressions for both the LUTI and the SETI-LUTI framework (refer to Chapter 6). Regression analysis has been used to examine the strength of relationship between the explanatory variables for train ridership. The agglomeration influence on ridership may be discussed by comparing models including and without agglomeration, and by comparing models based on the elasticity of agglomeration between sectors on train ridership.
4. Analysis III: Establish a regression analysis that accommodates the interaction term between distance of suburbs from stations and the magnitude of effective density, in order to understand the influence of station as the reference point of agglomeration on train ridership. Regression including the interaction term includes the effect of distance within the influence of effective density on train

ridership, thus providing information on the effects of geographical extent of agglomeration on train ridership (Chapter 7).

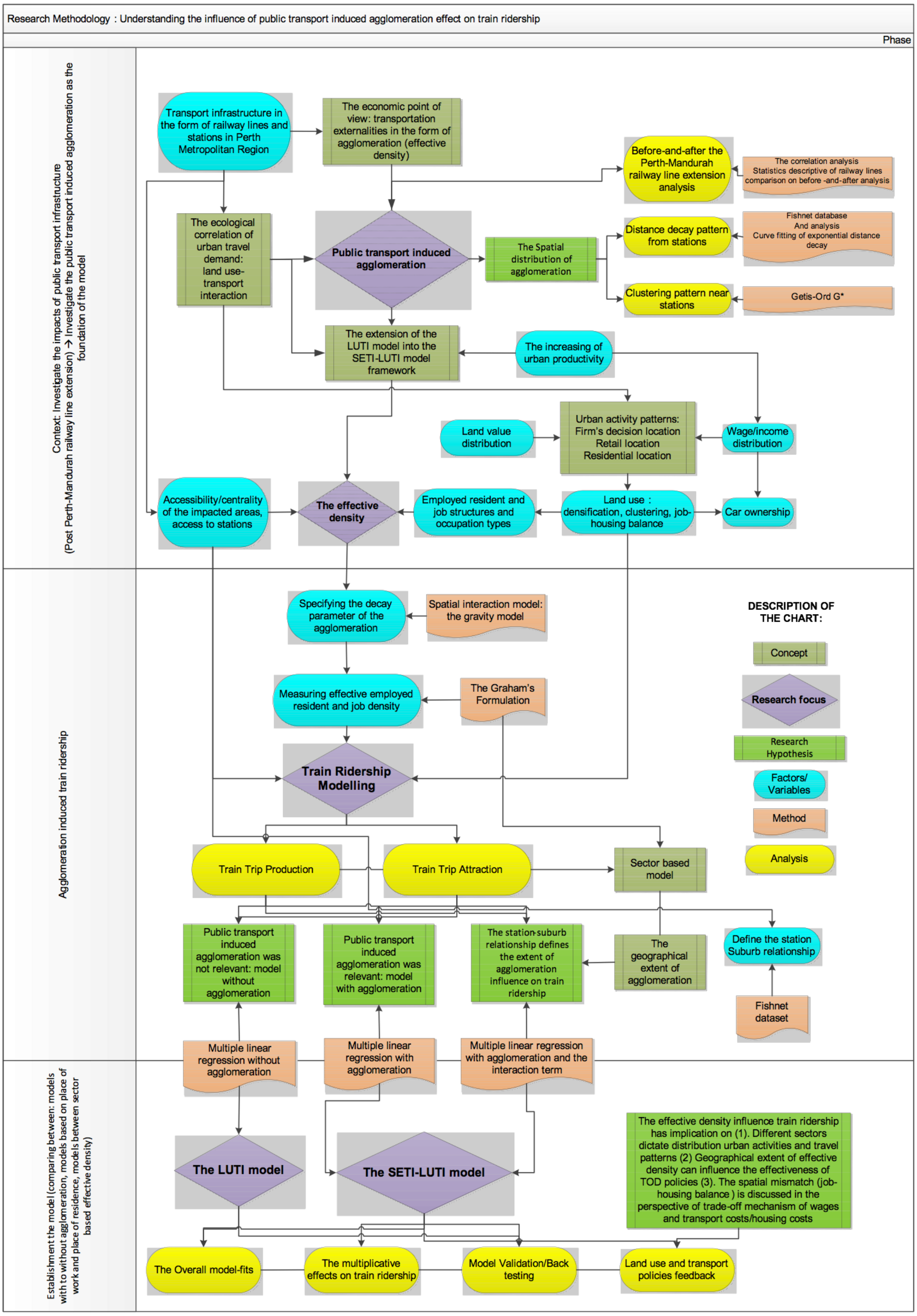


Figure 3.7 Research methodology framework

5. Investigate policy implications from the model results (refer to chapter 8). First, time decay parameters that may explain the effect of agglomeration on train ridership from different sectors have been interpreted. Second, application of the geographical extent of agglomeration has been interpreted, especially as applied in a TOD policy context. Third, the trade-off mechanism between wages and travel costs/housing costs have been examined, in order to identify the impact of travel time savings on the effective density premium, and its influence on train ridership. Methods used in this analysis comprised the use of probability density functions for curve-fitting of the trade-off relationship; the wage equations to model the relationship between wages and travel time; and the calculation of wage, land rent, and effective density premiums as a result of the trade-off.

3. 5 DETERMINATION OF MODEL PREDICTIONS

In terms of model prediction, (Handy, 1996) categorized three main groups of approaches (among others) to how the integration between urban form and travel behaviour may be explained: simulation studies, aggregate analyses, and disaggregate analyses.

Simulation studies, according to Handy, consist of techniques such as transport planning models and Lowry models. These models usually require an urban form component, such as the distribution of residential activities and employment, and a transport component, such as the structure of the transport network or travel patterns. One example application for simulation models was to predict and compare total travel between a typical suburban neighbourhood and the hypothetical "neo-traditional" neighbourhood.

Aggregate studies usually utilize aggregate measures for cities, neighbourhoods, or zones. These analyses provide a general understanding of the relationship between travel behaviour and urban form. Some techniques used in this type of study are, for example, multivariate regression, simple correlation, principal component analysis, and non-linear equations (Cervero, Murphy, Ferrel, Goguts, & Tsai, 2004). Applications of this type of study may be, for example, those which use aggregate measures to test the strength of statistical relationships between travel behaviour and urban form. Another example is to analyse the "compact city" policy, concerning the relationship between energy use and density (Holden, 2007).

Disaggregate studies involve data and analysis on the scale of individuals or households. They may use household/individual socio economic and and travel characteristics rather than city-level averages. For example, they may examine differences in travel choices of individuals between neighbourhoods and the relative importance of a variety of urban forms for those travel choices.

This thesis used the aggregate approach for train ridership prediction. Similar to the majority of aggregate studies (Gutiérrez et al., 2011; Holden, 2007; Taylor et al., 2009), multiple regression was used to predict outcomes based on multiple predictors. Regression analysis has been differentiated from correlation analysis, as regression is performed with the purpose of not only examining the existence of relationships (as in simple correlation), but also to examine the causality in relationships between two or more variables. Further, regression may also be used to estimate the value of one variable based on the other variable when these variables have been identified to have a high linear correlation coefficient (Evans et al., 2007, p. 33).

3. 6 DETERMINATION OF DEPENDENT VARIABLE

There are two alternative dependent variables, which have often been used in train ridership prediction:

Y = the absolute-number of daily train trip attractions or train trip productions

Y = the ratio or proportion of daily train trips out of total trip attraction (place of work model) or total employed residents (place of residence model)[1].

[1] The regression models had been conducted for all alternatives of dependent variables, including (1) Place of work dataset: (1.1) .the ratio of daily total train trip attraction per total trips attraction; (1.2). the ratio of daily total trip attraction per total jobs; (2). Place of residence dataset: (2.1). the ratio of daily total train trip production per total trips production; (2.2). The ratio of daily total trip production per total employed residents. Results showed the model (1.1) and model (2.2) gave the best results in terms of the possibility to compare between the LUTI and SETI-LUTI model.

The advantages and disadvantages of the first and second variables for the modelling are shown in table 3.2.

Table 3.2 Comparison of strengths and weaknesses for alternative dependent variables (after Taplin, 2016)

Alternative dependent variable	*Advantage*	*Disadvantage*
Absolute-number: the number of daily train trips	The interpretation of the model is straightforward.	Potential for ecological fallacy Potential for generating heteroscedasticity Dependent variable is not normally distributed.
Proportionate number: the ratio of the number of daily train trips to total workers or total jobs or total trips	Avoids ecological fallacy (eliminates bias from the degree of employment or total trip numbers) Variable is normally distributed.	The model interpretation is not straightforward.

This thesis has attempted to model with both alternatives for the dependent variable. Examples outcomes of the model are presented in the appendix of chapter 3. The decision was made to use the proportion of train ridership per total employed resident for residential data and the proportion of train ridership per total trip attraction for workplace data. This was found to be necessary to avoid large heteroscedasticity in the model and to obtain a full comparison between the LUTI and the SETI models, which required both land use density and its interaction term, or effective density and its interaction term, to be statistically significant and defined in the model.

Model variants that used the absolute-value of y resulted in much a higher variance (as indicated by the coefficient of determination) than variants with the proportional value of y. Furthermore, absolute-values appeared to be influenced by the variable of worker size to a large extent. For example, by removing the number of worker variable from the model with absolute y values, variance (as indicated by the coefficient of determination) was dropped by 30% (see appendix 2).

Furthermore, a log natural transformation of the dependent variable was performed to avoid the heteroscedasticity problem. Based on the requirement to fulfil this regression assumption (homogeneity), the results suggested that the best model was a log transformation of the proportional number of train ridership as its dependent variable. The natural log of proportion of train ridership was regarded to be free from the bias of worker or employment size and from the heteroscedasticity effect.

Using both the natural log of proportion of train trip production per total employed residents and the natural log of proportion of train trip attraction per total trip attraction have resulted in much better predictions and comparisons of the LUTI and SETI models than any other alternative *y* variables. The advantage of using two different ratios for the dependent variable was that each can omit the number of employed residents who work at home in the calculation of *y*. Failure to omit this factor may overestimate the size of train ridership in both models.

With log transformation *y*, the model adapted the LDV equation or the *limited dependent variable* equation (Taplin, 2016). What was important in this analysis was not the calculation themselves, but the alternative ways the problem could be approached and the interpretation of the results. Therefore, attention was paid to what the potential problems were and what interpretation might be assigned to the statistical results from the SPSS (Howell, 1997, p. 510).

Firstly, this thesis developed research hypotheses to determine the approach to the problem and the interpretation of the results. The use of "regression to predict *y* was performed on the assumption of simultaneous knowledge of all p predictors" (Howell, 1997, p. 509). For example, the research hypotheses require differentiating between models with and without agglomeration.

*Sec*ondly, by adapting the LDV equation of multiple linear regression, the emphasis of interpretation of model results was focussed on the multiplicative effects of the predictor variables rather than the additive effects (Taplin, 2016). Further explanation of this model is to be discussed in Chapter 6.

3. 7 DETERMINATION OF INDEPENDENT OR PREDICTOR VARIABLES

Within the framework of LUTI-based regression, this thesis categorised independent variables into three groups: people, place, and transport. For SETI-LUTI-based regression, a spatial-economic category was added, involving two measures: effective job density and effective employed resident density. Table 3.3 describes each variable and how each was used in previous research.

Table 3.3 Independent variables and theoretical justification

Component	Variable	References
Place component: Place was defined as the broad concept of land use, i.e. the generator of travel demand from various activities according to land use type.	Employed resident density	Density and the distance from the CBD correlate with the consistent increase in public transport share, independent of other factors (Rickwood & Glazebrook, 2009). Provided there were enough people to be carried on (mass) transit, the more effective and productive the transit would be. A more dense concentration of people translates to higher revenue and a more productive the transit system (Cushman, 1988).
	Job density	The areas with the highest employment densities such as the CBD usually have a concentration of transit hubs and a moderate level of population density due to high rents (Chen et al., 2008). The number of public transportation commuters were closely related to the number of CBD employees, rather than to overall metropolitan area size (Handy, 1996). Mixed land use and density were positively associated with journey to work times (Cervero & Duncan, 2006).
	Job housing balance	Job-housing balance refers to the distribution of employment relative to the distribution of workers within a given geographic area (Cervero et al., 1995). Plentiful jobs within four miles of home significantly reduced VMT and VHT for work trips. Housing-job proximity was the only land use variable negatively associated with commute time (Cervero & Duncan, 2006).
	Land value	Holding housing prices fixed, a real wage increase is required in the metropolitan market if the commuting time is increased. Elasticity of housing prices with respect to commuting time is larger than elasticity of wages (So, 2001).

People component: "People"was related to vertical segregation, consisting of the socio-demographic and socio-economic dimensions of travellers (Van Acker et al., 2007). .	Employed resident and Job structure by economic sector	Knowing the size and density of employment, along with the occupation structures and income level, and the characteristics of the transportation network has been important in predict trip generation (Hall, 1969). Elasticity of wages with respect to zonal commuting times represented the semi-elasticity of the hourly wage with respect to two additional minutes of commuting time, considering private and public sector workers (Timothy & Wheaton, 2001). Employment in the educational sector was positively influenced by monthly station boarding, higher than employment in the commercial sector. The use of employment types as a predictor, instead of employment density, attained higher explanatory power in ridership prediction (Gutiérrez et al., 2011).
	Employed resident and Job structure by occupation types	White collar workers (e.g. professionals and managers) have longer commuting times than blue collar workers, where white collar workers can afford higher travel costs while blue collar workers tend to cluster around their place of work (Wang & Chai, 2009).
	Vehicle ownership	Car ownership level is an important factor to influence the pattern of urban growth. For example, the growing of out of town shopping centres need to be complemented by higher car ownership levels (Hall, 1969) Regulation to restrict automobile ownership such as through tax on car acquisitions had a positive impact on transit patronage (de Grange et al., 2012). Blainey (2010) used a demographic variable consisting of car ownership and the number of jobs located within station catchments. Both variables were found to be statistically significant, with small improvements in model fit. Limited automobile ownership contributes to high rates of unemployment in the inner city as poor people living in inner city neighbourhoods are disadvantaged by the growing distance to suburban jobs. This was worse when the metropolitan area was deliberately designed for cars (Grengs, 2010).

	Income of residence based on occupation types Wage based on occupation types (at place of work) Wage of tertiary sector	Job decentralization created commuting time differentials that were capitalized into wage differences among competing employment sub-centres. Wages were found to vary up to 15 percent within metro areas and this variation is significantly related to the average travel times for workers in the various zones (Timothy & Wheaton, 2001). Population density, average income, bus service connectivity, distance to central stations, and service frequency were linked to trip production ridership (place of residence) (Chan & Miranda-Moreno, 2013). The variation in train ridership among urbanized areas can be explained by factors such as regional geography, metropolitan economy (income), population characteristics, and auto/highway characteristics (Taylor et al., 2009). The metropolitan wage level is closely related to the level of transportation. A wage premium is required to compensate for longer commuter time. Increase in commuting time lowers nonmetropolitan population (So, 2001). The higher the income, the less likely employees work and live at the same district and the more likely they spend longer times commuting (Wang & Chai, 2009). Income is negatively related to ridership production (Chan & Miranda-Moreno, 2013).
Transport component: The influence of public transport infrastructure combined with the attractiveness of public transport. Transportation factors appear to be more important than that of vehicle ownership and household income (J. Lin & Long, 2008).	Public transport supply such as train and bus network supplies	A household's decision to use a car (or not) was not only affected by internal factors (car type and the socio-economic/demographic profiles) but also by external factors in terms of public transport provision and the presence of parking restrictions (Nolan, 2010). Land use diversity, bus connectivity, and transfer stations were associated with station attraction ridership (place of work) (Chan & Miranda-Moreno, 2013). The policy of expanding metro or train networks stimulates the use of public transit (de Grange et al., 2012).
	Suburb's centrality by street network (travel distance)	Van de Coevering and Schwanen (2006) identified that the ratio of public transport to road and rail density and lower public parking in the CBD correlated positively with the total trip distance by public transport.

	Proximity or distance to train stations	Those who lived and worked in areas with good rail facilities used rail to get to work, combined with walking and cycling. This was true regardless of car ownership and work location (Commins & Nolan, 2011). The importance of railway stations, the proximity of living environments to station precincts, and the spatial accessibility of location increased the use of trains for commuters (Cervero, 1994; J. Lin & Long, 2008). Comparing Germany and the USA, it was found Germans living more than 1000 m from public transport have a higher share of trips by walking, biking and public transport (29% of all trips) compared to Americans living within 400 m of public transport (18% of all trips). People living next to a train station with a higher number of trains per hour were more likely to use public transport for their journey to work. The distance from the place of employment to the nearest train station was statistically related to travel behaviour (Ellis & Parolin, 2010).
	Network travel time	Whether the journey to work was faster by car or public transport was found to be significantly related to the modes used for working trips (Ellis & Parolin, 2010). Improvement in transportation that lowers commuting time increases non-metropolitan populations and will increase the number of nonmetropolitan commuters to metropolitan markets (So, 2001). Travel time saving is the most important attribute of public transport attractiveness. Route integration may require users to make transfers. Providing reliable and well connected routes was important to reduce total travel time (Chowdhury, Ceder, & Schwalger, 2015).
Spatial-economic dimension (SETI-LUTI model only): Agglomeration factor in terms of effective density	The concentration of activity near train stations, i.e clustering and distance decay patterns	Research Hypothesis 1. The dispersion of activity within a city region is a disadvantage for railway services. Rail service cannot accommodate a dispersion pattern of development as the strength of railway service relies on the ability to provide fast, reliable, and high capacity transport over a few main lines (Hall, 1969, p. 432).

	Effective job and employed resident density	Research Hypothesis 2. There is inconclusive evidence as to whether wage and commuting cost differences result from equilibrium agglomeration effects or from a disequilibrium distribution of employment (Timothy & Wheaton, 2001). If the distribution of transportation capacity is relatively uniform, then equal wages and commuting costs eventually imply some number of equal sized employment centre. Further, long run equilibrium dictates commuting costs and wages must equalize across different sized employment locations (Timothy & Wheaton, 2001).
	Station attribute: the interaction between distance of suburb to station (station access) and effective density determines the geographical extent of agglomeration	Research hypothesis 3. In general, stations farthest from downtown have higher trip production and lower trip attractions (Chan & Miranda-Moreno, 2013). Characteristics of stations were relevant for explaining ridership. The type of station, such as terminal, intermediate, interchange or intermodal was related to the number of riders (Gutiérrez et al., 2011). Station spacing influences the size of the catchment areas (the catchment radius) and additional spacing around a station would draw additional riders to that station (Kuby, et al., as cited in Gutiérrez et al., 2011, p. 1083).
	Job-housing balance (based on the *trade-off* perspective)	Research hypothesis 4 The joint decision of where to live and where to work was made by trading off wages, housing prices, and commuting costs (So, 2001). Wage variation across employment zones within a metro area is strongly correlated with the average commute time of the workers employed in that zone. Controlling for residential location, workers who commute farther should get paid more (Timothy & Wheaton, 2001).

3. 8 SUBURB AND FISHNET LEVEL OF DATASET

This thesis assumes that train ridership is not only a function of suburbs' socio-demographic/economic or land use characteristics, but also influenced by the stations' features. Therefore, the use of suburbs as study areas cannot be separated from the spatial distribution of stations in the Perth metropolitan region. This study proposed an additional way of modelling train ridership where one can accommodate the influence of station on train usage based on the catchment area of a station.

To accommodate the influence of station on a more detailed level, data in a suburb were divided into grid cells, i.e. a fishnet-based dataset, where the nearest station to each fishnet can be more precisely identified. Funderburg, Nixon, Boarnet, and Ferguson (2010), due to the need for finer geographic scale data, applied the grid or fishnet method. Their one-kilometre grid cells were developed as an aggregation of point-location microdata. In this thesis, however, the one-kilometre grid cells were developed as a disaggregation of aggregate suburb level data. A suburb level dataset was re-organised into smaller areas, which were in a form of one square kilometre grid cells using a geographic information system-mapping. A new value of each fishnet was assigned based on the area weighted calculation method (Dell, 2009).

The result of disaggregating suburb data into fishnet data resulted in an around 11,000 fishnet observations. An additional variable, i.e. the interaction term, was used to test any moderating effect of the suburb-station distance on the relationship between the agglomeration variable (effective density) and train ridership (hypothesis 3, chapter 7).

In addition, the fishnet-based dataset was used to examine the spatial attributes of distance decay of agglomeration (explored in Chapter 5). Fishnet-based data is used in some elements of analysis, such as in curve fitting of the probability density function for exploring the trade-off mechanism between distances to station, wages, land rents, and travel times. Fishnet-based data is also used in some wage equation models (hypothesis 4, chapter 8).

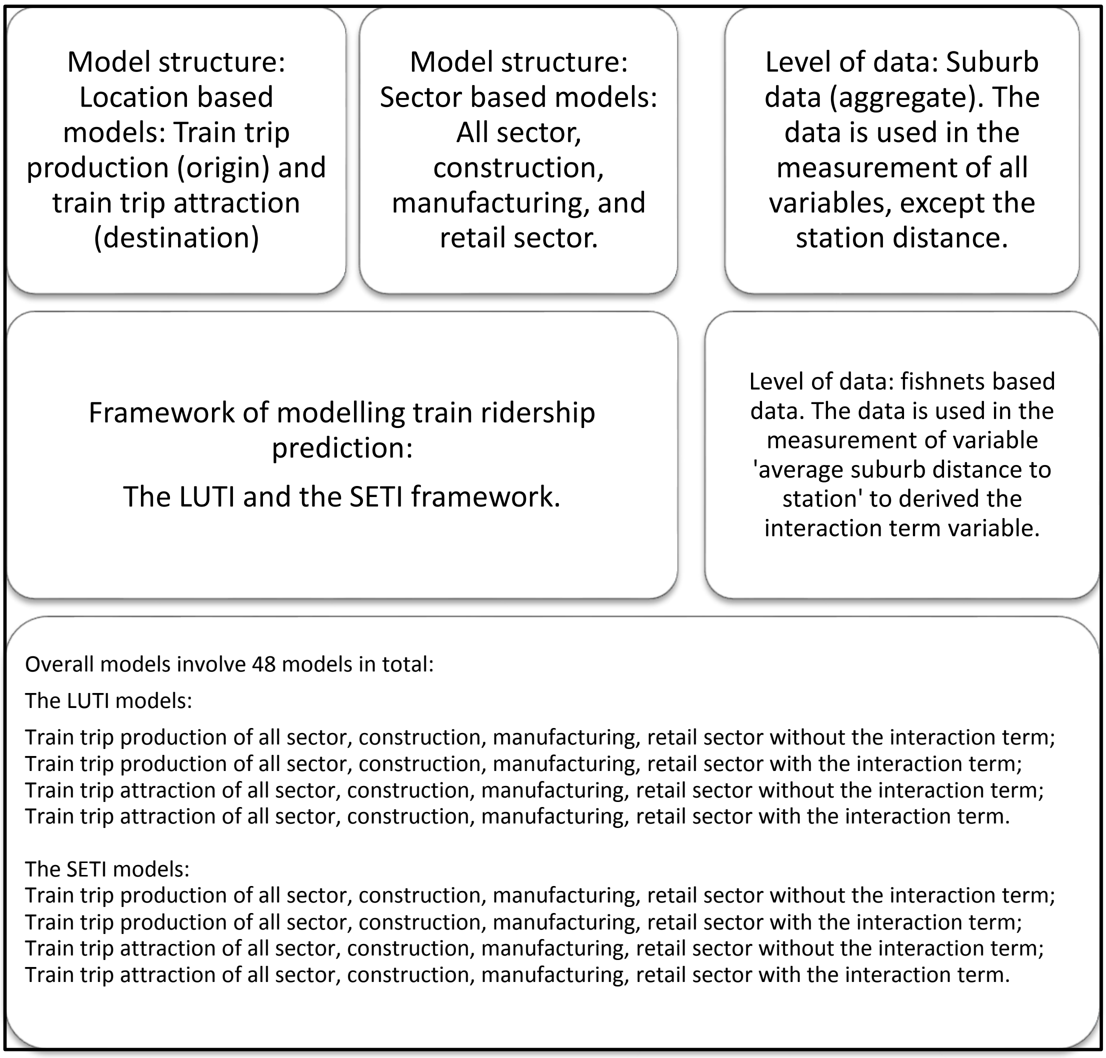


Figure 3.7 Model structure and data level

Furthermore, the model structure used in this thesis may be divided into location-based and sector-based analysis. Location-based analysis consists of a focus on origin or train trip production model and destination or train trip attraction model. The former model was build based on place of residence dataset, while the later model was based on place of work dataset. Sector-based analysis consists of the examination of the manufacturing, construction, and retail sectors, including a summary of all sectors. The comparison of the model with and without agglomeration (i.e. the LUTI and the SETI models) based on testing the hypothesis 2 (chapter 6) was conducted for all sector; while the testing of the hypothesis 3 was conducted by adding the interaction term based on sector construction, manufacturing, and retail sector (chapter 7). In the

LUTI model, the interaction between job density or employed resident density and the suburb average distance to train station was developed, while in the SETI model, the interaction term was developed between effective job or employed resident density and the suburb average distance to train station.

3. 9 STUDY AREA

The global trends in public transport development seem to be similarly reflected in the development of Perth. Before the 1960s, most employees and businesses considered multi-modal accessibility to be an important factor for location, so that neighbourhoods built before that time generally had good sidewalks, local services (shops, schools and parks) and frequent public transit services, particularly rail (such as the Fremantle station precinct).

Perth Metropolitan has committed to land use and transport integration policy since the formulation of its planning strategy in 1955, known as the 'compact city' scenario. The 1970 Corridor Plan and the 1990 Metroplan have also aimed to achieve land use transport integration (Curtis, 2005).

However, many neighbourhoods built between 1960 and 2000 were more automobile-dependent, coinciding with the construction of massive street networks and freeways in Perth from the 1960s to the late 1980s (Curtis, 2008). There were limited changes in land use around rail stations, due to the extension of the city boundary outward to a low density city, with the freedom of car use, a mechanism to cope with the developing urban structure (Curtis, 2005). The statistics reported by the ABS survey from 1976 to 2006 showed that Perth remains the city with the highest car use in Australia since the 1980s and until the early of 2000s, with the travel mode share for car drivers above 70%. Mode share for car passengers in Perth was also high, at above 11% in 1980s, although it continued to decrease to about 7% in 2000 and slightly increased in 2001. Concurrently, mode share for public transport (all types) in Perth was relatively low compared to other cities. It was about 12% in the 1980s, continued to decrease until 1996 to about 8%, then slightly increased above 10% from 2001 to 2006. Perth was the second lowest city in Australia in terms of public transport market share after

Canberra, especially during the period from the 1970s to 1980s (Mees, O'Connell, & Stone, 2008).

In recent years, there has been a growing demand for more transit-oriented development, both by individual consumers and by communities. This has been shown by the strategic policy changes in the Perth metropolitan region, which began in the 1980s to improve the LUTI performance by means of linking centres with a rapid transit system. A re-orientation toward public transport began with the revitalization of suburban rail lines in the 1980s in line with the TOD policy adhered to in state planning policy since 1988 (Curtis, 2008). The first revitalization began with the opening of the Fremantle line in 1983, followed by the electrification of suburban rail lines in 1987, and the construction of the northern suburban railway accompanied by the opening of a supporting feeder bus network in 1993 (Curtis, 2008). Recent investments made in Perth have been the construction of a 72 km railway line serving Perth's southern suburbs, which has become the most successful public transit system, in terms of its capability and comparative advantages compared to private transport. The average speed of 92 km/hour on the Perth-Mandurah rail line, compares favourably to that of the 60 km/hour speeds for the Perth-Clarkson rail line, and 47 km/hour for the Perth-Fremantle rail line. The system has also become the fastest rail in Australia, compared to those in other cities, such as Melbourne (46 km/hour), Sydney (60 km/hour), and Brisbane (68 km/hour) (Curtis, 2008). However, one may examine whether this extensive investment in public transport and the creation of new centres of activity by means of 68 (planned) TOD (D.C 1.6 1999 Development around Metropolitan Railway stations - Development Control: Policy Manual – Western Australian Planning Commission: 1998)2 has been the most effective solution for land use-transport integration in the Perth Metropolitan context.

[2] http://www.stirling.wa.gov.au/development/Projects/Glendalough%20Station%20Precinct%20Study/Glendalough%20Station%20Precinct%20Urban%20Design%20and%20Implementation%20Study.pdf

Table 3.4 Description of railway lines

	Rail line				
Station features and development	***Fremantle line***	***Joondalup line***	***Midland line***	***Armadale line***	***Mandurah line***
Development history (Curtis, 2008**)**	Opened since 1881 for Fremantle-Perth-Midland line Re-opened the line in 1983 after closed in 1979 and extended in 1993 northwards to Currumbine and in 2004 northwards to Clarkson.	Joondalup line opened in 1992.	Opened since 1881 for Fremantle-Perth-Midland line.	Opened since 1889 serving south-eastern suburbs Thornlie spur opened serving Thornlie station on Armadale line in 2005.	A 72 km railway line operational in December 2007 (Curtis, 2008) Serving South western suburbs: Perth, Esplanade, Canning Bridge, Bull Creek, Murdoch, Cockburn Central, Kwinana, Wellard, Rockingham, Warnbro, and Mandurah.
Recently experienced rail development/ expansion?	Yes In 2004 to Clarkson	No	No	Yes In 2005 for Thornlie	Yes In 2007 for all stations along Perth-Mandurah line
Activity centre	Shenton Park Subiaco	Glendalough Leederville Stirling Warwick	Bayswater East Perth Maylands	Burswood Cannington Thornlie	Bull Creek Cockburn Central Murdoch
Station transfer/ Interchange	Subiaco Claremont Fremantle	Glendalough Stirling Warwick Whitfords Joondalup Clarkson	Bayswater Bassendean Midland	Oats Street Cannington Thornlie Maddington Gosnells Kelmscott Armadale	Canning Bridge Bull Creek Murdoch Cockburn Central Kwinana Rockingham Warnbro Mandurah

Figure 3.8 presents the study area of the Perth metropolitan region. This thesis used the suburb or the state suburb (SSC) as an administrative boundary of the study area. The SSC was not part of the ABS structure. Instead, the areas were aligned closely with the Statistical Areas level 2 (SA2), where the SA2 are an aggregation of statistical

area level 1 (SA1). Both SA1 and SA2 are defined in the ABS structure (2011 Census Dictionary, p. 183 and p. 189). All suburbs were viewed as potentially both a place of residence and a place of work. For the place of residence dataset, the ABS census has provided almost all data especially in relation to the socio-demographic/economic data that are available at the suburb level. The place of residence dataset was derived from geographical areas based on place of usual residence. These areas were collected at the state suburb level (SSC) for Western Australia, and were chosen specifically for the Perth metropolitan region to include 339 suburbs.

The place of work dataset was derived from geographical areas based on Place of Work (POW) at the SA2 level or Statistical Area Level 2. Perth Metropolitan consisted of 159 SA2 areas. Some job data were also derived from Place of Work data, comprising 990 DZN or destination zones in the Perth metropolitan region. This was in order to derive a more detailed and precise job data set. The areas for SA2 and DZN were tractable using ArcGIS. Map layers provided by ABS were overlaid with the suburb boundaries in order to segregate the SA2 level and aggregate the DZN level into the suburb level. This included several stages of conversion procedures (further explained in Chapter 4). For all digital boundary maps files, the ABS provides all maps in a GIS file.[3]

[3] http://www.abs.gov.au/websitedbs/D3310114.nsf/home/ABS+Geography+Publications#DigitalGIS

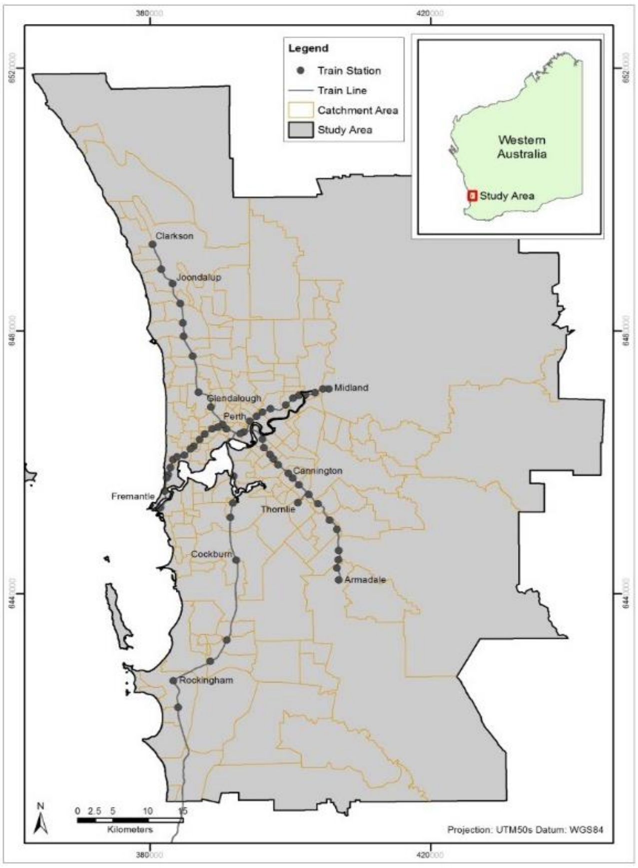


Figure 3.8 Map of study area

Amongst the 339 suburbs across the Perth Metropolitan Region, there were some suburbs that currently coincided with a national park or conversation areas that were unpopulated or had a very low population and number of employed residents. These areas were removed from the analysis. These suburbs were as follows:

1. SSC50026 Avon Valley National Park
2. SSC50126 Canning Mills
3. SSC50329 Herdsman
4. SSC50396 Keralup
5. SSC50441 Lesley
6. SSC50487 Melaleuca
7. SSC50753 Tamala Park
8. SSC50757 Tapping
9. SSC50762 The Lakes
10. SSC50763 The Spectacles
11. SSC50795 Walyunga National Park
12. SSC50838 Whiteman

Remote areas such as Garden Island and Rottnest Island were also excluded from the analysis. Other data limitations were related to the availability of property transaction data in 2011. Some suburbs with missing data for 2011 transactions were replaced by data from previous years or from after 2011. However, some suburbs that did not have any property transactions from between 2006 and 2013 may be representative of a less dynamic area of land use/residential development. Therefore, these suburbs were omitted as well. These suburbs were: Karrakup and Swan View. Three additional suburbs were removed since these suburbs have outlier data: Wooroloo, Mandogalup, and Keysbrook. In total, there were 316 suburbs adopted as the study area for train trip production and train trip attraction.

3. 10 THE SCOPE OF THE MODEL

The scope of the model has been limited to modelling home-to-work rail trips only, based on home-based trips (train trip production) and non-home end trips (train trip attraction) and modelling omitted non-work trips due to data limitation issues: the ABS Census only covered trips for working purposes and there were no non-work trips data available. However, the PARTS (Perth and Region Travel Survey) data based on 2006 surveys provided data on all trips, categorised by trip purpose. It found that the percentage of home-based trips were 37.12% out of all trips and that 38.5% of trips were working trips. In total, 11.18% of all surveyed trips were the trips where the main destinations were for the purpose of working. There were no data available for 2011 surveys.

In the context of agglomeration studies, such as in this thesis, the main indicator of agglomeration is the “effective density”, measured as job and employed resident effective density. These indicators were regarded as proxies for travel demand for working purposes. In summary, the home-to-work rail trips were regarded as the focus of this study, and not any non-work-related trips.

In addition, the focus of this thesis is on the utilization of travel time by public transport (park and ride) as the basis for the calculation of public transport induced agglomeration. The choice of public transport would be determined significantly by whether or not the journey to work was faster by public transport, as opposed to auto or car access (Ellis & Parolin, 2010). Furthermore, in the summary of the thesis findings (chapter 9 section 9.6 and appendix 60), this thesis re-ran the model based on whether the agglomeration was influenced by auto access (car travel time) and its impact on train ridership. The model showed that job agglomeration based on car travel time actually had an inverse relationship with the level of train ridership. Therefore, this thesis assumed that car travel time should be used in the prediction of the level of car use, and that the thesis would only focus on the influence of public transport park and ride travel time as the determinant of the level of public transport induced agglomeration.

3. 11 INDUSTRY SECTORS

Studies in agglomeration economies often assume that each industry type has a different effect on the economy. Specifically, the different impacts of transport investment on urban productivity partly depend on different economic sectors (Deng, 2013). In this way, the analysis of agglomeration usually allows for detailed sectoral coverage (Graham, 2007, p. 324). Therefore, agglomeration in terms of effective job density and employed resident density were both defined and classified based on industry types.

The measurement of agglomeration was subdivided further for employment based on industry sectors. This differentiation was necessary since different industry sectors are likely to have different elasticity with regard to their regression relation to train ridership. Therefore, the separate analyses were conducted for each sector.

ABS provided data on employment levels based on the 2011 Census – Employment, Income and Unpaid Work section. ABS defined a 1-digit category for industry of employment (INDP) using the Australian and New Zealand Standard Industrial Classification (ANZSIC). The categories were applicable to employed persons, and consisted of: (1) Agriculture, forestry and fishing, (2) Mining, (3) Manufacturing, (4) Electricity, Gas, Water and Waste services, (5) Construction, (6) Wholesale trade, (7) Retail trade, (8) Accommodation and food services, (9) Transport, Postal and Warehousing, (10) Information media and telecommunication, (11) Financial and Insurance services, (12) Rental, hiring and real estate services, (13) Professional, Scientific and technical services, (14) Administrative and support services, (15) Public administration and safety, (16) Education and training, (17) Health care and social assistance, (18) Arts and recreation services, (19) Other services. The geographical area based on the respondent's usual residence was chosen. This employment division was also used to define job division in this thesis, based on place of work geographical area.

Some sectors were orientated toward natural resources, encompassing primary industries: agriculture, forestry, fishing, and mining. These sectors were omitted from the analysis since primary industry contributes the least to urban working trips. Some industries in which their location might be influenced by the occurrence of natural resources, for example electricity, gas and water (Graham, 2007, p. 333) were also omitted from the analysis. Other sectors that consisted of manufacturing, construction and services industries were regarded as being orientated toward market or household consumption. Nevertheless, many agglomeration studies have found that the manufacturing and service sectors have different magnitudes and areas of influence or geographical spill-over effects (Dekle & Eaton, 1999; Graham, 2007). Therefore, this thesis chose industry sectors based on these considerations. The three sectors were chosen based on their potential influence on the magnitude of agglomeration, and their natural orientation towards market or natural resources. One sector from the general services sector was chosen, i.e. the retail sector, and two others from secondary industries, i.e. the construction and manufacturing sectors. The total 19 sectors (or the "all sectors" dataset) were used as the benchmark model.

Additional criteria for choosing sectors included the level of employment that may be sufficient to influence travel demand, and the contribution of the sector to urban productivity, as measured by employment income (Table 3.5).

Table 3.5 Description of sector economy based on size and productivity criteria

INDUSTRY NAME	*EMPLOYMENT DENSITY 2011*	*INDUSTRY NAME*	*WEEKLY EMPLOYMENT INCOME 2011*
Health Care and Social Assistance	22643.56	Professional, Scientific and Technical Services	107,257,291.00
Retail Trade	19883.93	Construction	106,319,916.00
Professional, Scientific and Technical Services	18825.04	Health Care and Social Assistance	90,555,760.00
Construction	18577.46	Manufacturing	81,090,266.00
Education and Training	17006.81	Mining	78,071,457.00
Manufacturing	15342.2	Education and Training	73,704,972.00
Public Administration and Safety	12592.43	Public Administration and Safety	69,610,568.00
Accommodation and Food Services	12165.25	Retail Trade	56,996,757.00
Mining	9425.81	Transport, Postal and Warehousing	42,422,604.00
Transport, Postal and Warehousing	8071.38	Wholesale Trade	36,784,407.00
Wholesale Trade	7407.61	Financial and Insurance Services	33,937,053.00
Other Services	7315.7	Other Services	29,102,664.00
Administrative and Support Services	6585.09	Accommodation and Food Services	27,482,399.00
Financial and Insurance Services	6224.44	Administrative and Support Services	25,642,500.00
Rental, Hiring and Real Estate Services	3619.65	Rental, Hiring and Real Estate Services	18,357,696.00
Arts and Recreation Services	3234.48	Electricity, Gas, Water and Waste Services	15,069,197.00
Information Media and Telecommunications	2620.01	Information Media and Telecommunications	11,523,021.00
Electricity, Gas, Water and Waste Services	2199.99	Arts and Recreation Services	10,267,490.00
Agriculture, Forestry and Fishing	894.72	Agriculture, Forestry and Fishing	4,267,601.00

The three sectors chosen, i.e. the retail, construction, and manufacturing sectors, were among the top 10 sectors in both criteria. These sectors were also chosen in order to have comparable results to other studies in agglomeration.

3. 12 SOFTWARE REQUIREMENTS

All data have been processed in Microsoft Excel to prepare all required calculations. All spatial data, specifically in a shapefile format, has been processed in the GIS (Geographical Information System) software developed by the Environmental

Systems Research Institute or ESRI. The processing of data included mapping all required information and variables into a visual format, including overlaying and buffering processes, for example to form the buffer zones around train stations. The conversion and aggregation of different geographical boundary levels were performed in GIS, specifically in ArcMap 10.2 (ESRI Inc., 2010). The analysis for the clustering of effective density was performed using the Getis-Ord G* tool in ArcGIS. The model for train ridership prediction was conducted using IBM SPSS Statistics 22 (IBM Inc., 2015), particularly for conducting multiple linear regression. All data preparation such as data cleaning and data transformation was performed either in Excel or SPSS.

CHAPTER 4. DATA PREPARATION

This section describes the process of data preparation prior to performing the analysis. The aim is to prepare the data by identifying and solving problems related to data format or content. These processes enable the extraction of more meaningful information by understanding the nature of the data (Blischke, Rezaul Karim, & Prabhakar Murthy, 2011). Descriptions of each of the dependent and independent variables are described below, followed by the stages involved in the data preparation.

4. 1 DESCRIPTION OF DEPENDENT AND INDEPENDENT VARIABLES

4.1.1 Dependent variable: method of travel to work

The data source for the number of working trips was the Australian Bureau of Statistics (ABS) "method of travel to work by persons" (MTWP) response from the 2011 Census. The dependent variable for the model developed in this study was the number of working trips by train. Chapter 3 described the initial analysis to test which measurement would give a better model fit between the number of trips to work by train, either as an absolute-number or as a proportion of the total workers, or total jobs, or total trips. The decision was made to use a proportion of the number of train trips but presented in the form of a limited dependent variable, thus requiring a log natural transformation of this variable to reduce the problem of homogeneity and outliers (Dhrymes, 1986; Taplin, 2016).

"Method of travel to work" data was provided by the ABS under a Curtin license by utilizing Table-Builder Pro database. This license allowed access to a database called '2011 Census – Counting Employed Persons, Place of Work'. The data included classifications for employment, income and unpaid work categories. The category of MTWP (Method of Travel to Work by Persons) consisted of 236 types of modes (both single and combinations of modes). The MTWP involving train travel consisted of the following modes: train; train and car as passenger; train and car as driver; train and taxi; train and bus; train and bicycle; train and ferry; combination of train and car as driver and car as passenger; train and bus and car as passenger; train and bus and car as driver; train and bus and bicycle; and train and others.

These data were collected in the format of a matrix of train trip production or home-based working trips by train (the place of residence model), train trip attraction data (the place of work model) and train trip distribution data of spatial flow between origin and destination (the gravity model). The first and the second data sets have been used as dependent variables in the train ridership modelling. The third data set has been used in the calculation of the time decay parameter of agglomeration, based on the Graham (2007)'s formula (see equation 3.1 and 5.1), where the level of spatial interaction was stated as a function of travel time between residential suburbs and workplace suburbs to derive or estimate the parameter of time decay (explained in detail in chapter 5).

4.1.2 Independent variable

4.1.2.1 Agglomeration measurement

This measurement was a function of the number of employed residents or job density and the network travel time by park and ride. Park and ride travel, based on home-work travel was around 21.85% of all travel involving trains (2011 ABS Census – MTWP data). The methodology has focussed on the use of travel time by park and ride in the agglomeration calculation due to data limitations. That is, while the percentage of bus and ride travel was around 25.2% and that of train only was 43.2% (2011 ABS Census – MTWP data), there were no travel time data available for all combination of mode types involving trains.

The STEM model output specified two types of public transport travel time matrices. Those included the travel time by walk and ride and also by park and ride. The travel time by park and ride was chosen to accommodate the component of auto accessibility (car travel time) in the agglomeration calculation, as this thesis assumed that both the auto accessibility to stations and train accessibility would influence the magnitude of agglomeration of an area.

The travel time was a direct way to assess any transport improvements and directly translates into improved accessibility or effective density (Graham & Melo, 2010). The difference in this study, as reflected in equation 3.1, is that this thesis used travel time instead of travel distance as an indicator of proximity between suburbs (further described in equation 5.1).

(i) Effective job density (*EJD*) is a function of the number of jobs in each suburb *i* and all other suburbs *j*, and the travel time by park and ride between suburb *i* and suburb *j*.
(ii) Effective employed resident density (*EDER*) is a function of the number of employed residents (resident employees) in each suburb *i* and all other suburbs *j*, and the travel time by park and ride between suburb *i* and suburb *j*.

It is important to note that this thesis mentions two types of decay variables: *first*, the time decay parameter, which is used in the calculation of effective density. This parameter is estimated by using the train trip distribution data in the gravity model (as discussed in section 5.1), where the resulting time decay parameter from the gravity model is applied as the time decay parameter in the calculation of effective density (equation 5.1). *Second*, the distance decay parameter is derived from the magnitude of effective density calculated for each suburb. The distribution of effective density across the study area is examined with respect to the distance from stations in aggregate. Curve fitting based on an exponential distance decay equation has been used to estimate the parameter for distance decay of the percentage of effective density as a function of distance from stations (discussed in chapter 5.3).

Thus, the time decay parameter measures the sensitivity of agglomeration magnitude to travel time between pairs of suburbs, where the distance decay parameter measures the sensitivity of agglomeration magnitude to the distance between train stations and the location in question (i.e. the fishnets centroids). The utilization of distance decay in the latter parameter is used as evidence to support the hypothesis of public transport induced agglomeration. As mentioned in chapter 3.3, the distance decay pattern of agglomeration is defined as the spatial attribute of agglomeration in this thesis. Meanwhile, chapter 3.2 has suggested the underlying assumption for the agglomeration concept, adapted from Venables (2004) and Graham (2007), is that a station is a focal point or point of reference, thus suggesting that the distance decay of agglomeration away from a station is an important attribute of agglomeration.

4.1.2.2 Travel time by park and ride

The travel time data used in equation 5.1 was based on park and ride services. Limitations on the data included the fact that travel time matrices were estimated from the STEM model (Department of Planning, 2011) which consisted of travel times between travel zones based on car journeys; walk and ride; and park and ride journeys. The travel zones were therefore converted into suburb matrices. The travel time by park and ride journey was choosed as the best alternative to represent the spatial interaction of the train trips.

4.1.2.3 Per hour per person wages and incomes level

Wage data were calculated by converting the place of work based income data for blue collar workers and for managers/professionals from the Statistical Area level 2 (the SA2) data into suburbs data. The calculation for wage level was similar to the calculation for other income levels. The ABS defined “total weekly personal income” (INCP) as the total income that the person in a place of residence usually receives every week. It was applicable to persons aged 15 years and over. Income per hour per person was calculated as the total weekly income summed up over all workers at a place of residence divided by the number of workers and by the total weekly hours (40 hours) to generate income per hour per person. Wage per hour per person was calculated as the total weekly income summed up over all workers at a place of work divided by the number of workers and by the total weekly hours (40 hours) to generate income per hour per person.

The ABS "INCP" was derived from the following categories (the median amounts appear in brackets):

1. Negative income
2. Nil Income
3. $1 - $199 ($80)
4. $200 - $299 ($263)
5. $300 - $399 ($349)
6. $400 - $599 ($487)
7. $600 - $799 ($698)
8. $800 - $999 ($896)
9. $1,000 - $1,249 ($1,107)
10. $1,250 - $1,499 ($1,363)
11. $1,500 - $1,999 ($1,695)
12. $2000 or more ($2,579)

Total weekly income of suburbs was defined as the amount of median income in each category times the number of workers of that income category in that suburb, then finally summed up over all the categories for each suburb. The rate of wage/income per person per hour was calculated from this total weekly income, divided by the total number of workers and the total hours per week (40 hours).

The income and wage was calculated for blue collar and managers/professionals occupation.

4.1.2.4 Land value of residential land uses and non-residential land uses

The data on land value were obtained from the mean of transaction price of properties calculated for all dwellings that were sold in each suburb across the Perth metropolitan region from 2006 to 2013. The land value was calculated by dividing the property price by the land area of the property. The calculation was then standardized based on housing characteristics of a 3 bedrooms-2 bathrooms property type by accommodate data of this property type and omit other data. This was to avoid bias based on housing characteristics. Property value data were obtained by permission of ©Western Australian Land Information Authority (Landgate, 2014). Data from all types of properties were comprised of 395.716 parcels from the 2006 to 2013 dataset in total.

All parcel data were aggregated into a suburb value; using the following calculation:

$$LV_{suburb_i} = \sum_1^N LV_{Parceli}/N \qquad \text{Equation 4.1}$$

Where:

LV_{suburb_i}= the average land value of all parcels for dwellings in suburb *i*, are calculated based on a 3 bedrooms-2 bathrooms property type.

N = the number of parcels in suburb *i*.

These data were further adjusted, based on a constant dollar value for 2011/2012, which is a CPI (Consumer Price Index) standard for housing that was issued by ABS (ABS, 2011)[4]. Any missing data for property transaction and/or suburbs in 2011 were replaced by transactions from other years between 2006 and 2013 and the prices were adjusted based on a constant dollar value as of 2011/2012. Land values for residential and non-residential land use were calculated as separate categories. The former category was used as a variable in the place of residence dataset for train trip production model, and the latter was used in the place of work dataset for train trip attraction model.

4.1.2.5 Public transport supply

The public transport supply index was a combination measurement of public transport service frequency (such as buses trips and train trips per week) and suburbs' access distance to bus stops and train stations, and was computed for each suburb using GIS software. The formula to calculate this index was specified as follows, after (Currie & Delbosc, 2010):

[4] A Guide to the Consumer Price Index: 16th Series (Document Number 6440.0), ABS 2011.

$$PTIsuburb_i = \sum_N \left(\frac{area_{Bn}}{area_{suburb\ i}} * SL_{Bn}\right) \qquad \text{Equation 4.2}$$

Where:

$PTIsuburb_i$ = the public transport supply based on buses and train network for suburb i,

N = the number of walk access buffers to buses stops and train stations in each suburb i,

Bn = buffer n for each stop for each suburb defined as a 400 meter radius from each stop, and an 800 meter radius from each station.

area suburb *i* = the the land area of the suburb in square kilometres,

SL = service level measure, i.e. the number of bus arrivals per week corresponding to each bus stop, and the train frequency per week for each station.

The procedure to conduct this calculation was as follows (Currie & Delbosc, 2010):

- Transperth of Western Australia provided maps of all bus stops and stations. Using GIS software, buffer areas were drawn from each stop out to a 400 m radius from the bus stop and 800 m from the station corresponding to each relevant suburb.
- The suburbs' access distance to each stop and station were calculated.
- The service frequency for each stop and station, measured as the total number of service arrivals in one week were calculated.
- A combined measure of service frequency and access distance was calculated based on the formula in equation 4.2 to obtain the index of public transport supply for each suburb.

4.1.2.6 Job housing balance index or job to worker ratio (JWR)

Based on the literature, job-housing balance refers to the distribution of job relative to the distribution of employed residents or workers within a geographic area in question (Giuliano, 1991). Some methods to calculate the job housing balance are as follows:

- Comparison of the number of resident workers with the number of jobs within a given area; i.e. E/P ratio or employment per population ratio (Giuliano, 1991). The index can be compared with the regional average of number of workers per occupied housing unit to identify job quality, for example, by categorising the area based on radial distance to downtown cores or central business district areas. The index can also be disaggregated by sectoral composition, such as manufacturing, service, retail trade, etc.

- Other comparisons such as jobs/households, jobs/total housing, jobs/labour force, jobs/population.
- Using a jobs/housing imbalance indicator, similar to the Gini index to calculate the equity of the jobs and housing availability (Bento, Cropper, Mobarak, & Vinha, 2003, 2005). This was simply conducted by displaying the cumulative proportion of employment and the cumulative proportion of population.
- The standard acceptable ratio used in local policy decision-making: is considered to be 1.3 to 1.7. Cervero (1996) suggested a value of 0.8 – 1.25 jobs to housing ratio, and a value of 0.75 to 1.5 is sufficient to represent balanced.

The operational definition of JWR used in this thesis was taken to be the ratio between the numbers of jobs to the number of employed residents. It was measured for each suburb across the Perth metropolitan region. The index was calculated to be a range of 0.75 – 1.5. The range below 0.75 represents employed resident-rich suburbs and the range above 1.5 represents job-rich suburbs.

4.1.2.7 The number of jobs

This thesis studied three industry sectors and a general index: (1) Construction, (2) Manufacturing, (3) Retail trade and (4) all 19 sectors as a benchmark. The number of jobs is the number of employments in place of work dataset.

The number of jobs was available from the ABS dataset at SA2 level and destination zones (DZ). This thesis used the Destination zones dataset that were the smallest areas in the hierarchical field for the Place of Work (ABS). A conversion from the destination zones to the suburb dataset was performed (see section 4.2.3).

4.1.2.8 Employed resident density

Employed resident density is the number of employed residents per area of a suburb. ABS recorded the number of employed residents based on a person's labour force status (LFSP) for the week prior to census night. "Employed resident" referred to persons aged 15 years and over who were either: (1) employed, and worked full-time, or were (2) employed and worked part-time, or were (3) employed, and were away from work[5]. Other categories of LFSP were not included in the analysis, such as the unemployed and those not in the labour force (Australian Bureau of Statistics (ABS), 2011).

4.1.2.9 Job density (Jobd)

Job density is the number of job per area of a suburb.

4.1.2.10 Blue collar and managers/professionals employed resident or job

The data for employed residents (employed persons at a place of residence) and jobs at a place of work were directly available from ABS, categorised by occupation. The variables chosen were (1). the proportion of blue collar and (2) the proportion of white collar (limited to professionals and managers) for each suburb.

The employment classification in ABS was based on occupation type and industry type. The occupation type divided employment into 8 categories: Professionals, Managers, Clerical and Administrative Workers, Technicians and Trades Workers, Community and Personal Service Workers, Sales Workers, Labourers, and Machinery Operators and Drivers. For this study, these 8 categories were re-classified into two categories of employment, i.e. blue collar and manager/professional employment for both place of residence and place of work data set.

The definition of blue collar and white collar workers in this thesis follows the ABS standard. ABS defines blue collar and white collar workers based on the type of occupation that were used for census data and other statistical data from 1996 to 2011

[5] includes persons who stated they worked but who did not state their number of hours worked or did not work any hours in the week prior to Census Night (ABS Census Dictionary 2901.0 2011 p.115).

for jobs classification. A blue-collar occupation refers to technicians, trade workers, machinery operators and drivers, and labours. White collar occupations are managers, professionals, community and personal service workers, clerical and administrative workers, and sales workers[6].

4.1.2.11 Vehicle ownership (average per household)

Vehicle ownership data was directly available from ABS. The data provided "the number of registered motor vehicles owned or used by household members, garaged, parked at or near private dwellings on census night" (ABS, 2010), in ranges. It included company-owned vehicles but excluded scooters and motorbikes. The published ranges for the number of motor vehicles were: (1). Zero motor vehicles, (2) one motor vehicle, (3) two motor vehicles, (4) three motor vehicles, (5). Four or more motor vehicles. All data applied to occupied private dwellings only. The number of occupied dwellings was calculated for this study to be the total occupied dwellings minus the reported number of "not stated" and "not applicable" responses. The average vehicle ownership per household in each suburb was calculated for each vehicle ownership category as this net number of occupied dwellings times the number of vehicles in that category divided by the number of households or occupied dwellings in that category.

4.1.2.12 The percentage of employed resident in each industry sector

This was simply the number of employed residents in each sector divided by the total number of employed residents in each suburb.

4.1.2.13 The percentage of job in each industry

This was calculated as the number of jobs in each sector divided by the total number of jobs in each suburb.

[6] (http://www.abs.gov.au/ AUSSTATS/ abs@.nsf/Previousproducts/ DB1BFFF6CB753F05CA257922000E2865?opendocument#Endnote12)

4.1.2.14 The average suburb distance to train station based on the fishnet dataset

For the purposes of determining the station catchment area and the average distance of a suburb to the nearest train station, a fishnet-based dataset (grid cells) was used. The average suburb distance to a train station was calculated over all *n* fishnets in suburb *i*, where the distance between each fishnet centroid to each station centroid was calculated using a Euclidian (straight-line) distance.

There are two types of distances between suburbs and stations that are used separately in train trip attraction model and train trip production model:

- The distance of the suburb from the boarding station, i.e. the distance from a residential suburb to its nearest station. This data was used in train trip production model.
- The distance of the suburb from the destination station, i.e. the distance from a workplace suburb to its nearest station. This data is used in the train trip attraction model.

4.1.2.15 Train station catchment radius

The effects of transportation improvements due to the Perth-Mandurah railway line extension have been ongoing since the commencement of development in 2004 and further following the opening of the railway in 2007. The dataset used in the model is that of 2011. It is assumed that the 7 years following the commencement of development and approximately 5 years following the opening of the railway translated into a substantial growth in residential and business development and shifts in land rent values, property values, and the number of jobs and population, especially in the areas closer to the railway line extension. Thus, comparing the levels of effective density, travel times, and train ridership between different railway lines and how they have changed following the extension may demonstrate the existence of public transport induced agglomeration. This thesis assumes that the current level of train ridership has been influenced by the agglomeration process. This implies that the determination of a catchment area would be dependent on the empirical data for train ridership.

The adopted catchment area definition, based on the observed train usage data, followed the assumption stated in Blainey (2010) and Cervero et al. (1995). The catchment station was delineated based on the assumption that people will always use the station closest to them. Although this assumption may be a simplification and people may choose the second or third closest station and so on, the limited policy tools available force an emphasis on the importance of bringing potential destinations and residences into greater proximity (Iacono, Krizek, & El-Geneidy, 2010). This assumption simplifies the modelling process (Cervero et al., 1995).

In order to determine the spatial boundary of a train station catchment area, this thesis used GIS software to allocate population and all the attributes of a suburb to their nearest station, and to delineate the catchment boundary of a station based on those identified suburbs. The allocation of population to their nearest station was applied using the distance via road network in GIS.

Further investigation into the delineation of catchment areas followed the method used in Cervero et al. (1995). "Catchment areas were defined as contiguous census tracks which encompass the origins of 90% of all access trips to (BART) stations (or destinations of 90% of all egress trips from station). The use of the 90 percent rank to demarcate catchment areas was chosen to represent the distance at which the vast majority of access trips are drawn. It was selected largely based on visual inspections of the cumulative distance of access trips to all (BART) stations. Beyond 90 percent, most access trips fall toward the extreme tail of the distance distribution. That is, some people make fairly long access trips of 20 or more miles from the exurban and rural fringes to reach suburban (BART) stations; however, these access trips represent statistical outliers" (Cervero et al., 1995, pp. 40-41).

This thesis differs in concept between the catchment definition offered in Cervero in the way that it measures the total catchment station averaged across all stations. Other adjustments to this catchment area definition for this thesis consist of: first, the mode of work travel taken into account was all modes involving trains. Therefore, only work travel involving train was used from the ABS dataset. Second, the train ridership distribution based on cumulative distance from a station and a distance cut off was selected based on no further statistically significant increase in train usage was observed in the data. In this way, a 95% trips to origin stations for home-based trips and 95% trips to destination stations was decided as cut off for catchment demarcation. It was worth noting that at 90% of trips (as adopted in Cervero's research), the cumulative graph still showed an increasing trend of ridership. Therefore, the 95% cut-off was deemed more relevant for the study area of this thesis. This 95% level of train ridership was the cut-off point for the maximum catchment radii summed over the whole metropolitan region. The cut-off point for the edge of the catchment coincides with a radius of approximately 16 km (or 10 miles) distance from a station.

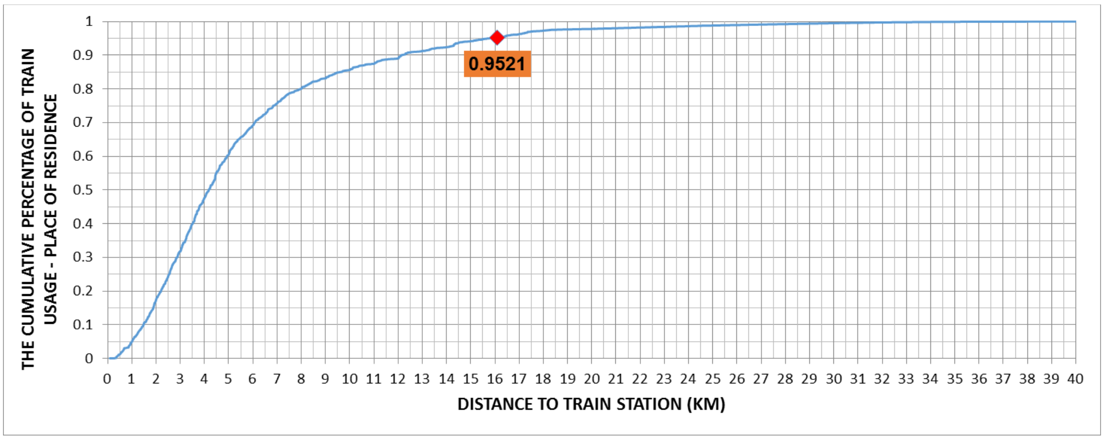


Figure 4.1 Place of Residence dataset: 95% trips at 16 kilometres or 10 miles.

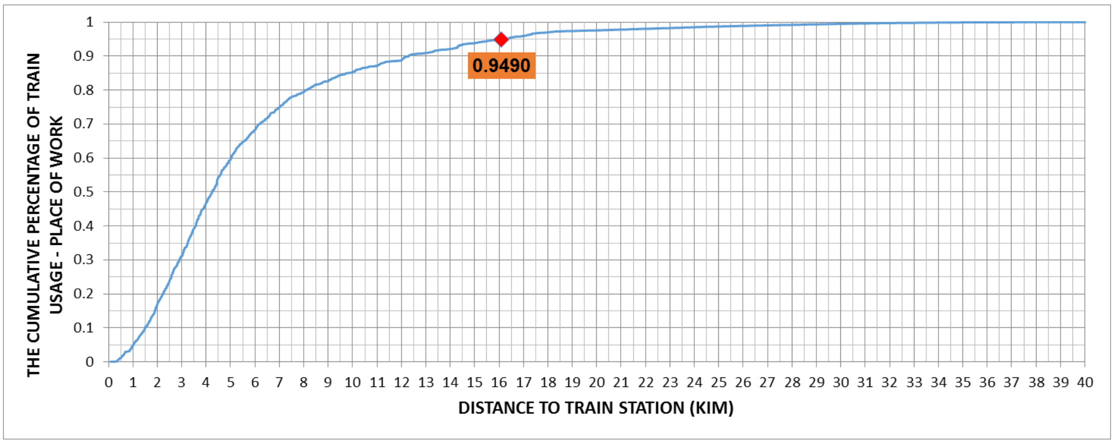


Figure 4.2 Place of Work dataset: 95% trips at 16 kilometres or 10 miles.

Based on the above-mentioned calculation, this thesis assumed the edge of the station catchment on average in the study area to be 16 km from train station. Thus, referring to figure 3.3 of section 3.2, this 16 km distance represents the maximum distance from train stations at which employed residents are still willing to commute by train (and by park and ride) and reach their job location. This means that 95% of commuters are willing to travel by train where the distance of the station from residential and workplace suburbs within 16 km. The bottom 5% of total ridership may be assigned as statistical outliers. Research in Olaru et al. (2014) found that the average distance of residential suburbs to their boarding stations for park and ride trips to work in the Perth metropolitan region is 6.07 km. The assumption of 16 km maximum distance to the boarding station in this thesis for 95% of train ridership is therefore a reasonable assumption.

4. 2 PREPARATION OF INDEPENDENT VARIABLES

The main objectives of this section are to describe the data preparation in relation to the independent variables that were input for various models. The process included: (1) data cleaning, which aims to remove any noise from the data such as any outliers or extreme values from the dataset (2) data transformation to prepare the data in order to remove these outliers. These steps included descriptive statistics and histogram displays for outlier detection and the transformation of continuous data into another continuous variable to reduce the impact of outliers. This is so that the statistical results can reflect all suburbs and avoid dependence on one or two suburbs; (3) Aggregation/segregation of unit analysis (observation) for consistency, due to that fact that data were obtained at different geographic levels or due to the presence of fractured data.

4.2.1 DATA CLEANING

Identification of any problems that might be embodied in the data was performed by checking the main data features such as centrality measures and dispersion measures of the data to reflect the normality distribution of data.

Central tendency of the data was examined using the sample mean and median. The comparison of mean and median indicates how symmetric the distribution of variables in the population is, relative to its centre. The difference in the two calculations may be used as an indication of skewness (Blischke et al., 2011). A positively skewed distribution refers to a distribution that is skewed to the right, indicating that most of the values were concentrated in the lower portion of the distribution. A negative skewed or the left skewed distribution means a greater number of higher values in the data set (Hensher, Rose, & Green, 2005). A non-normal distribution may exhibit data with potential outliers. Identifying outliers can be conducted by plotting a histogram of the data, then identifying whether a non-normal distribution is apparent from outlying or non-linear data.

Among the dataset, variables with an approximately normal distribution without any substantial skewness are displayed in table 4.1. For example, the proportion of jobs in managers/professionals occupations for a suburb in the Perth metropolitan region on average is 0.32 jobs per total (32%) and the median is 0.31. This variable had the most similarity between mean and median.

Table 4.1 The list of variables with a normal distribution

Variable	*Mean*	*Median*	*Std Dev*	*Skewness*
Proportion of train trip production (ptrain_wo)	0.0731 trips per total employed residents or 7.31%	0.0709	0.04	0.555
Proportion of job in managers/professionals occupation (p_manpr)	0.32 jobs per total jobs or 32%	0.31 jobs per total jobs	0.0867	0.856
Proportion of job in blue collar occupation (p_blu)	0.29 jobs per total jobs or 29%	0.265 jobs per total jobs	0.11	0.767
Hourly wages of employment of managers/ professionals job occupation (inc_mp)	$32.4 per person per hour	$32.1 per person per hour	$4.34	0.558
Hourly wages of employment of blue collar job occupation (inc_blu)	$21.27 per person per hour	$21 per person per hour	$3.8	0.946
Proportion of employed resident in managers/professionals occupation (pmanpr_r)	0.3396 employed resident per total or 33.96%	0.3211 employed resident per total	0.125	0.413
Proportion of employed resident in blue collar occupation (pblu_r)	0.3286 employed resident per total or 32.86%	0.33 employed resident per total	0.115	0.068
Proportion of employed resident in the construction sector (p_wocon)	0.1021 employed resident per total or 10.21%	0.1032 employed resident per total	0.035	0.020
The level of car ownership in a household (car_own)	1.88 cars per occupied dwellings	1.88 cars per occupied dwellings	0.414	-0.929

Employed resident density (wod)	609.6 employed resident per sq km of land area	634 employed resident per sq km of land area	462.3	0.497
Employed resident density in the construction sector (wodc)	58.18 employed resident in the construction sector per sq km of land area	57.16 employed resident in the construction sector per sq km of land area	44.9	0.477
Employed resident density in the manufacturing sector (wodm)	47.9 employed resident in the manufacturing sector per sq km of land area	48.2 employed resident in the manufacturing sector per sq km of land area	36	0.363
Employed resident density in the retail sector (wodr)	62.2 employed resident in the retail sector per sq km of land area	67.8 employed resident in the retail sector per sq km of land area	46.24	0.389
Effective employed resident density in 1000 units (edertt1000)	11.85 (in 1000 units) or 11,850 employed resident opportunities	12.217 (in 1000 units) or 12,217 employed resident opportunities	2.771 (in 1000 units)	-0.068
Effective employed resident density of the construction sector in 100 units (ederttc100)	6.02 (in 100 units) or 602 employed resident opportunities	6.05 (in 100 units) or 605 employed resident opportunities	1.94 (in 100 units)	0.180
Effective employed resident density of the manufacturing sector in 100 units (ederttm100)	9.32 (in 100 units) or 932 employed resident opportunities	9.40 (in 100 units) or 940 employed resident opportunities	1.98 (in 100 units)	-0.174
Effective employed resident density of the retail sector in 100 units (ederttr100)	5.04 (in 100 units) or 504 employed resident opportunities	5.21 (in 100 units) or 521 employed resident opportunities	1.87 (in 100 units)	0.141

On the other hand, some variables were distributed with a skew to the right or left, where the value of the median was either less than or greater than the value of the mean respectively. These variables are presented in table 4.2. A general rule was to consider any values that were not found in the interval between ± 2 standard deviations of the mean as potential outliers. This assumes that the non-normal distribution of data was largely influenced by outliers or extreme values in the data. A normal distribution has the property that the mean, mode, and median are coincident at the point of symmetry of the distribution. Approximately 95% of data lie between -2 and +2 standard deviations from the mean (Stopher, 2012, p. 46). This assumption simply states that 2.5% of data in the left-side tail of the distribution and 2.5% in the right-side tail may be regarded as "outliers".

Table 4.2 List of variable with skewed distribution and its related mean and median

Variable	*Mean*	*Median*	*Std Dev*	*Skewness*
The number of train trip attraction (train_w)	146.65 persons daily on average to a suburb	7.85	1476.66	16.936
Proportion of train trip attraction (ptrain_w)	0.0216 trips per total or 2.16%	0.0131	0.0348	5.329
The number of train trip production (train_r2)	171.7 persons daily on average from a suburb	108.5	186.78	2.252
All job number (from total 19 sectors) (all_job)	2010.7 jobs	722.5 jobs	5405.5	10.869
Proportion of job in the construction sector (p_jobcon)	0.0978 jobs per total or 9.78%	0.0828	0.006	2.307
Proportion of job in the manufacturing sector (p_jobman)	0.0665 jobs per total or 6.65%	0.035	0.0809	2.593
Proportion of job in the retail sector (p_jobret)	0.117 jobs per total or 11.7%	0.098	0.08	1.133
Public transport supply by buses network (ptiori)	893,395	411,960	1,353,034	4.841
Centrality suburb by road network travel distance (strio)	11,883 meters (11.883 km)	10,981 meters (10.981 km)	3,002 meters (3.002 km)	1.270
Average distance of a suburb from train stations (ave_dist)	5.87 km	4.26 km	4.9	2.390
Job density (jobd)	464 jobs per sq km of land area	191.5 jobs per sq km of land area	1042.09	6.029
Job density in the construction sector (jobdc)	31.05 jobs per sq km of land area	15.124	60.5	5.778
Job density in the manufacturing sector (jobdm)	28.8 jobs per sq km of land area	4.98	73.7	4.788
Job density in the retail sector (jobdr)	48.79 jobs per sq km of land area	19.3	82.18	3.317

Job to employed resident ratio (jwr)	19.84 jobs per one employed resident	0.38	171.2	11.049
Job to employed resident ratio in the construction sector (jwrc)	18.7 jobs per one employed resident	0.298	142.9	8.812
Job to employed resident ratio in the manufacturing sector (jwrm)	41.85 jobs per one employed resident	0.17	343.6	9.004
Job to employed resident ratio in the retail sector (jwrr)	10.3 jobs per one employed resident	0.39	90.87	11.581
Proportion of employed resident in the manufacturing sector (p_worker_man)	0.0889 employed resident in the manufacturing per total or 8.89%	0.0822	0.05	7.008
Proportion of employed resident in the retail sector (p_woret)	0.1025 employed resident in the retail per total or 10.25%	0.1031	0.027	1.661
Income of employed resident in the managers/professionals occupation (inc_pr_r)	$36.39 per person per hour	$37.5	$7.48	-2.873
Income of employed resident in the blue collar occupation (inc_blu_r)	$26.25 per person per hour	$26.5	$4	-2.811
Land value of residential land uses (lvr)	$1,301.7 per sq meters.	$1,057.4	$1092	1.699
Land value of non-residential land uses (lvnr)	$1,492.8 per sq meters.	$1,212.96	$1,195.9	1.266
Effective job density in 1000 units (ejdtt1000)	10.416 (in 1000 units) or 10,416 job opportunities	9.689 (in 1000 units) or 9,689 job opportunities	4.7 (in 1000 units)	5.267
Effective job density of the construction sector in 100 units (ejdttc100)	4.04 (in 100 units) or 404 job opportunities	3.47 (in 100 units) or 347 job opportunities	2.73 (in 100 units)	5.266
Effective job density of the manufacturing sector in 100 units (ejdttm100)	7.81 (in 100 units) or 781 job opportunities	7.04 (in 100 units) or 704 job opportunities	3.77 (in 100 units)	3.989
Effective job density of the retail sector in 100 units (ejdttr100)	4.72 (in 100 units) or 472 job opportunities	3.98 (in 100 units) or 398 job opportunities	3.32 (in 100 units)	3.095

A graphical method to perform this analysis was obtained using histogram analysis with normal distribution plots (appendix 6-8 of chapter 4). In order to solve the problem of non-normal distributions, a method to remedy the skewness of data for the above-mentioned variables was conducted by data transformation.

4.2.2 DATA TRANSFORMATION

Four transformation methods were used in this thesis, namely: (1). Natural logarithm transformation, (2). Square root transformation, (3). Square transformation, (4). Inverse transformation (Berg, Hoefsloot, Westerhuis, Smilde, & Werf, 2006). The predictor variables were examined graphically by histogram with normality plots before and after transformation to check if the skewness and non-normality of the distribution had been improved by the transformation method (appendix of chapter 4). Only two transformations were successful in reducing the effect of outlier data and improving the normality of the explanatory variables: the logarithm transformation and the square root transformation. The logarithm transformation and the square root transformation convert the original score data into the transformed data as per the following equations:

Natural logarithm transformation:

$$\tilde{x}_i = \ln(x_i) \qquad \text{Equation 4.3}$$

Square root transformation:

$$\tilde{x}_i = \sqrt[2]{x_i} \qquad \text{Equation 4.4}$$

Whereas:

$\tilde{x}_i$ = the value of variable *x* post-transformation for the observation i.

x_i = the original score of variable *x* for the observation *i*

Both transformations were found to be useful in making skewed distributions more symmetric, such that they reduced more of the higher values than the lower values, thus reducing the ranges and the standard deviation of a variable (Berg et al., 2006). For example, the original scores for the average distance of a suburb from train stations had a mean of 5.87 km, median of 4.26 km and a standard deviation of 4.9 km. Post-transformation, into log natural of average distance, the mean, median and standard deviation became 1.5, 1.45 and 0.687 in the log scale respectively. The resultant means, medians and standard deviations post-transformation for the skewed variables are presented in table 4.3. Compared to the original score, the normality distribution of these variables has been considerably improved.

The results of the histogram analysis are shown in the appendix 3 to 8 of chapter 4. In general, the natural logarithm transformatiom (ln) was able to correct the non-symmetric data into an approximately normal distribution. Only two variables, the *public transport supply* and *land value of non-residential land use* were transformed using the square-root transformation for further improved results. A few other variables remained in their original form due to (1) the original scores already having a normal distribution (listed in table 4.1), or (2) the failure of all methods in the transformation process (table 4.4). Table 4.3 presents the variables with successful transformations, resulting in approximately equal means and medians post-transformation.

Table 4.3 List of variable and its corresponding mean and median value post-transformation

Variable	***Mean***	***Median***	***Std Dev***	***Skewness***
Ln of the proportion of train trip attraction (ln_ptrainw22)	-4.565	-4.262	1.4	-0.490
Ln of the proportion of train trip production (ln_train_wo)	-2.78	-2.76	0.87	0.729
Square root of the public transport supply by buses network (sqrt_ptiori)	746.3	641.8	581	1.050
Ln of centrality suburb by road network travel distance (ln_strio)	9.35	9.30	0.23	0.795
Ln of average distance of a suburb from train stations (ln_avedist)	1.5	1.45	0.687	0.416
Ln of job density (ln_jobd)	4.96	5.25	1.75	-0.613
Ln of job density in the construction sector (ln_jobdc)	2.5	2.7	1.6	-0.743
Ln of job density in the manufacturing sector (ln_jobdm)	1.75	1.6	1.8	0.190
Ln of job density in the retail sector (ln_jobdr2)	2.85	3.0	1.58	-0.129

Ln of job to employed resident ratio (ln_jwr)	-0.675	-0.968	1.51	2.297
Ln of Job to employed resident ratio in the construction sector (ln_jwrc)	-0.87	-1.2	1.55	2.453
Ln of job to employed resident ratio in the manufacturing sector (ln_jwrm2)	-1.3	-1.7	1.92	1.980
Ln of job to employed resident ratio in the retail sector (ln_jwrr2)	-0.775	-0.9	1.56	1.365
Ln of land value of residential land uses (ln_lvr)	6.77	6.96	1.035	-1.021
Square root of land value of non-residential land uses (sr_lvnr)	35.4	34.8	15.4	0.267

Unsuccessful transformation mainly related to the data from the job database that were derived from the conversion of a destination zones classification into a suburbs classification. For these variables, their post-transformation distribution made the overall regression *model-fit* worse than without the transformation; such that the newly transformed *t-test* value for the corresponding variable was lower than that of the original variable. Therefore, in these cases, the original variables were included instead of the transformed variables.

Table 4.4 List of variable and its corresponding mean and median value in its original scores without the transformation

Variable	*Mean*	*Median*	*Std Dev*
All job number (from total 19 sectors) (all_job)	2,010.7 jobs	722.5 jobs	5405.5
Proportion of job in the construction sector (p_jobcon)	0.0978 jobs per total or 9.78%	0.0828	0.006
Proportion of job in the manufacturing sector (p_jobman)	0.0665 jobs per total or 6.65%	0.035	0.0809
Proportion of job in the retail sector (p_jobret)	0.117 jobs per total or 11.7%	0.098	0.08
Proportion of employed resident in the manufacturing sector (p_worker_man)	0.0889 employed resident in the manufacturing per total or 8.89%	0.0822	0.05
Proportion of employed resident in the retail sector (p_woret)	0.1025 employed resident in the retail per total or 10.25%	0.1031	0.027

Income of employed resident in the managers/professionals occupation (inc_pr_r)	$36.39 per person per hour	$37.5	$7.48
Income of employed resident in the blue collar occupation (inc_blu_r)	$26.25 per person per hour	$26.5	$4

The main reason that transformation could not improve some non-normal distributions was due to the presence of extreme outliers in those distributions. This required further data cleaning to remove outliers by: (1) removing the observations with extreme values, if this occurred in more than one variable for that observation; and (2) imputation methods, if the observation had an outlier value in only one variable. The outlier value was in this case replaced by the mean value or by the next lower or higher value in the data set.

There were some observations where the numbers of train ridership at place of residence and place of work were found to be zero, for example in suburbs of Bickley, Nowergup, Oldbury, Piesse Brook, O'Connor (WA), Henderson, Naval Base, Neerabup, Welshpool (WA), and Kwinana Town Centre. In these cases, the replacement method was used by adding the number of train ridership by one for all observations to eliminate the zero value while maintaining the level of variance in these data. The non-zero value was required since the dependent variable had to be transformed into log natural distribution.

The variable for public transport supply by bus network was affected by a high number of observations with zero values, involving approximately 20 observations and suburbs. Observations with zero values were either due to missing data or due to no public transport service available in the area. This thesis replaced the zero values with the next lowest value in the data set.

The two other variables with outlier values were the average distance of suburb to train station and the effective density. An extreme value for average distance to train station and other variables was recorded for the suburb Wooroloo, and suburbs Mandogalup and Keysbrook recorded extreme values for some agglomeration variables. These three suburbs were removed from the data set of both place of residence and place of work.

In conclusion, the overall model variables are listed in table 4.5.

Table 4.5 List of predictor variables used in the LUTI and the SETI-LUTI model

LUTI based Regression	*SETI-LUTI based Regression*
Proportion of employed resident in blue collar occupation (pblu_r)	Proportion of employed resident in blue collar occupation (pblu_r)
Proportion of job in blue collar occupation (p_blu)	Proportion of job in blue collar occupation (p_blu)
Proportion of employed resident in managers/professionals occupation (pmanpr_r)	Proportion of employed resident in managers/professionals occupation (pmanpr_r)
Proportion of job in managers/professionals occupation (p_manpr)	Proportion of job in managers/professionals occupation (p_manpr)
Proportion of employed resident in the construction sector (p_wocon)	Proportion of employed resident in the construction sector (p_wocon)
Proportion of employed resident in the manufacturing sector (p_worker_man)	Proportion of employed resident in the manufacturing sector (p_worker_man)
Proportion of employed resident in the retail sector (p_woret)	Proportion of employed resident in the retail sector (p_woret)
All job (all_job)	All job (all_job)
Proportion of job in the construction sector (p_jobcon)	Proportion of job in the construction sector (p_jobcon)
Proportion of job in the manufacturing sector (p_jobman)	Proportion of job in the manufacturing sector (p_jobman)
Proportion of job in the retail sector (p_jobret)	Proportion of job in the retail sector (p_jobret)
Income of employed resident with managers/professionals occupation (Incpr_r)	Income of employed resident with managers/professionals occupation (Incpr_r)
Income of employed resident with blue collar occupation (inc_blu_r)	Income of employed resident with blue collar occupation (inc_blu_r)
Income of employment with managers/professionals job occupation (Inc_mp)	Income of employment with managers/professionals job occupation (Inc_mp)
Income of employment with blue collar job occupation (inc_blu)	Income of employment with blue collar job occupation (inc_blu)
Average vehicle ownership (car_own)	Average vehicle ownership (car_own)
Square root of public transport (sqrt_ptiori)	Square root of public transport (sqrt_ptiori)
Log natural of road network travel distance (ln_strio)	Log natural of road network travel distance (ln_strio)

Log natural of average distance of a suburb to train stations **(ln_avedist)**	Log natural of average distance of a suburb to train stations (ln_avedist)
Ln of land value of residential land uses **(ln_lvr)**	Ln of land value of residential land uses (ln_lvr)
Square root of land value of non-residential land uses **(sr_lvnr)**	Square root of land value of non-residential land uses (sr_lvnr)
Employed resident density **(wod)** **The construction: (wodcon)** **The manufacturing: (wodman)** **The retail: (wodret)** **Log natural of job density** **(ln_Jobd)** **The construction: (ln_jobdcon)** **The manufacturing: (ln_jobdman)** **The retail: (ln_jobdret)**	Employed resident density (wod) The construction: (wodcon) The manufacturing: (wodman) The retail: (wodret) Log natural of job density (ln_Jobd) The construction: (ln_jobdcon) The manufacturing: (ln_jobdman) The retail: (ln_jobdret)
Log natural of Job to employed resident ratio **(ln_jwr)** **The construction: (ln_jwrc)** **The manufacturing: (ln_jwrm)** **The retail: (ln_jwrr)**	Log natural of Job to employed resident ratio (ln_jwr) The construction: (ln_jwrc) The manufacturing: (ln_jwrm) The retail: (ln_jwrr)
Interaction variables **Employed resident density and distance to train station** **(wod_ln)** **The construction: (wodc_ln)** **The manufacturing: (wodm_ln)** **The retail: (word_ln)** **Job density and distance to train station** **(lnjobd_ln)** **The construction: (lnjobdc_ln)** **The manufacturing: (lnjobdm_ln)** **The retail: (lnjobdm_ln)**	Interaction variables Employed resident density and distance to train station (wod_ln) The construction: (wodc_ln) The manufacturing: (wodm_ln) The retail: (word_ln) Job density and distance to train station (lnjobd_ln) The construction: (lnjobdc_ln) The manufacturing: (lnjobdm_ln) The retail: (lnjobdm_ln)
	Effective employed resident density (edertt1000) The construction: (ederttc100) The manufacturing: (ederttm100) The retail: (ederttr100) Effective job density (ejdtt1000) The construction: (ejdttc100) The manufacturing: (ejdttm100) The retail: (ejdttr100)
	Interaction variables Effective employed resident density and distance to train station (edertt_ln) The construction: (ederttc_ln) The manufacturing: (ederttm_ln) The retail: (ederttr_ln) Effective job density and distance to train station (ejdtt_ln) The construction: (ejdttc_ln)

	The manufacturing: (ejdttm_ln) The retail: (ejdttr_ln)

4.2.3 DATA AGGREGATION AND CONVERSION

The further problem of fractured data emerged, since the secondary data were collected from different sources and data were required to be classified with respect to administrative boundaries at the suburb level. This problem mainly occurred for data from the place of work dataset. "Fractured data" corresponds to the problems of data incompatibility or mismatch that arise from multiple sources or data obtained at different spatial levels (Blischke et al., 2011).

Most of place of work data were provided by the ABS for SA2 (Statistical Area Level 2) and destination zones (DZN) spatial categories. Conversion of spatial boundaries from the SA2 regions and DZN regions into suburb regions was performed using GIS software and simple arithmetic. In addition, travel time matrix data taken from the STEM model (Department of Planning of Western Australia) were subject to a conversion from traffic zones into suburb regions, using the same methods. Table 4.6 presents the spatially fractured datasets and the method required to convert the data.

The method of conversion and aggregation used was based on a weighting by land area (Dell, 2009). Generally, the maps to be overlaid are in the same projection systems where the unit of measure is in metres. Maps with different administrative boundaries are intersected using the Intersect tool. The centroids of suburbs in question are then matched to the nearest centroid of other administrative areas from the data sources. All zonal statistics data are exported to a .dbf file and also processed in Excel software for further calculation. A record of the conversion process for each source of data is shown in table 4.6 and the conversion process performed for variables is shown in equation 4.5.

Table 4.6 The fractured data and method of aggregation and conversion

Source of data	*GIS by AGGREGATION*
ABS Census data for Place of work	Match the SA2 and DZN boundary with the suburbs boundary in GIS mapping Result in159 SA2 to be converted into 339 matching suburb boundaries. Result in 990 DZN to be aggregated into 339 suburbs.

	Convert the value based on the calculated conversion ratio of land area between suburb and its paired SA2 and between DZN member of each suburbs and its corresponding suburb.
Travel time (matrices data) based on street network and railway line network data	Match each of the travel zone boundaries with the suburbs boundary in GIS mapping. Convert the value based on matching suburbs pairs and travel zones pairs. Several steps of GIS procedure were conducted in order to convert the spatial geography of traffic zones into suburbs. The conversion of travel time matrix from the 484 x 484 traffic zones matrix results in the 335x335suburbs[7] matrix of suburb geographical boundary (missing 4 suburbs in the dataset due to incomplete of paired traffic zones that corresponded to these four suburbs).

Variables that were subject to conversion consisted of all variables at place of work dataset, such as the number of jobs in all sectors and for each sector, job-housing balance, job density, wage level at place of work, the numbers of train trip attraction, and other socio-demographic variables such as occupation types. All these variables were converted from the destination zones dataset, except for that of wages, which were converted from the SA2 or Statistical area level 2. An example of a conversion (for the numbers of train trip attraction at suburb of place of work) sourced from destination zones data is as follows:

$$trainatt_s = \sum_{dz} landwt_{sdz} * trainatt_{dz} \quad \text{Equation 4.5}$$

Where:

$trainatt_s =$ The number of train trip attraction at suburb in question.

$landwt_{sdz}$ = The land area proportion of each suburb that pertains to the original administrative destination zone.

$trainatt_{dz} =$ The number of train trip attraction at destination zone.

The area weighted calculation expresses the percentage of the originally-classified area contained within the new area classification (ABS, 2014)[8]. Theis represents a ratio of the area of overlap calculated using a GIS system. For example, a 2011 suburb may be made up of 70% of one DZN plus 30% of another DZN.

[7] Only data from 316 suburbs are used
[8] Australian Statistical Geography Standard (ASGS) Correspondences (ABS, 2014).

It is essential that all datasets are in the same coordinate system before performing such calculations that involved multiple layers (Dell, 2009).

4. 3 CHAPTER SUMMARY

This chapter has described all essential aspects of model preparation. The multiple linear regression with *limited dependent variables* in a form of log natural transformation was used for the model predictions. The process of data cleaning and data transformation of dependent variables was mostly related to attempts to reduce the heteroscedasticity of the original datasets, and the data transformation of independent variables was mostly related to the problem of outliers. Thus, log natural transformation was used in many cases to correct this. The data aggregation to a suburb level for the place of work dataset was expected to create more manageable or interpretable results since the source of data was on a smaller scale (such as that for destination zones and traffic zones). Only one variable, i.e. wage data, was sourced from a higher level dataset; that of the Statistical area level 2. However, wage tends to be constant or similar between SA2 and suburb areas due to the fact that wage level is independent from the size of classification area.

The construction of a fishnet database was an attempt to build a more rigorous method for the calculation of distance between a suburb and train station. This variable is a proxy for train station accessibility, and thus essential in defining the geographical extent of agglomeration around a station, as the focal point of public transport induced agglomeration. In addition, the curve fitting of a distance decay equation is more precise by using the fishnet database, which provides more observations, with a greater sample size of more than 11,000 fishnets (compared to the 316 suburbs), to be fitted into the exponential distance decay equation.

CHAPTER 5. PUBLIC TRANSPORT INDUCED AGGLOMERATION

This chapter concerns the reasoning that links transportation system development to agglomeration and to travel behaviour, which should hold in principle if changes in these spatial interactions are due to public transport-induced agglomeration. Public transport-induced agglomeration may be defined as the concentration of economic activities and the clustering of offices, shops, entertainment centres, and other land-uses that emerge around public transportation stops (Weisbrod & Reno, 2009). This causal link is based on the following research assumptions:

(i) A new railway line extension would improve accessibility of a location; resulting in lower travel impedance between suburbs. It is expected to increase the willingness to travel and other spatial interactions. An increase in accessibility is associated with the increase in potential for transit demand (Melo et al., 2013).

(ii) A transport "shock" such as a railway line extension would change the accessibility of a location, and those changes imply that the station may be used as the focal point for spatial variations (Hensher et al., 2012). A location closer to the train station is expected to have a higher effective density. These spatial patterns form clusters of high effective density near the train station and imply a distance decay pattern away from the train station.

This thesis assumes that increasing train ridership before and after the extension and the emergence of the spatial decay patterns around stations are indicative of transport-induced agglomeration. Therefore, this thesis has quantified public transport-induced agglomeration using three measurements:

(i) A before-after railway line extension analysis. Changes in travel time, the effective density and train ridership before-and-after extension have been examined in areas around the Mandurah line, which was subject to the extension project. If the rate of change of agglomeration around the Mandurah line is much higher than that of the other lines, this may suggest that some of these changes may be due to the transport infrastructure development.

(ii) A spatial distribution of agglomeration in terms of clustering analysis.

(iii) A spatial distribution of agglomeration in terms of spatial decay phenomena. Clustering and spatial decay of high effective density away from station may indicate transport induced agglomeration.

These three measurements relate to first research hypothesis (section 3.2) which concerned the effect of transportation investment in terms of public transport-induced agglomeration and the spatial distribution of agglomeration. Preceding this investigation is the process of calculating effective density within different sectors of the economy and for all sectors.

5. 1 BEFORE-AFTER EXTENSION ANALYSIS

5.1.1 Method of calculating effective density

The magnitude of agglomeration in terms of effective density was measured for both effective job density and employed-resident effective density. Following the formula of Graham (2007, see equation 3.1), this thesis calculated the magnitude of agglomeration in each sector (rewritten in this section):

$$A_{im} = \frac{E_{im}}{\sqrt{(A_i/\pi)}} + \sum_{j}^{i \neq j} \frac{E_{jm}}{tt_{ij}^{\alpha}} \quad \text{Equation 5.1}$$

Where:

A_{im} = the effective density of job (employed resident) of suburb i in sector m. Where m= (all total sectors, construction, manufacturing, and retail).

E_{im} = the level of activity, i.e. the number of jobs (employed residents) of sector m available in suburb i.

E_{jm} = the level of activity, i.e the number of jobs (employed residents) of sector m available in suburb j where suburb $i \neq$ suburb j.

A_i = land area of suburb i.

tt_{ij} = travel time (by park and ride) between location i and j, as defined in section 4.1.2.2 of chapter 4. The travel time park and ride consisted of the time spent to access the origin station by cars, plus the time spend in-vehicle (in-vehicle time), and the egress time from the alighting station to the final destination. Thus, the calculation of agglomeration has considered the influence of both train and car accessibility.

α = the travel time decay parameter of agglomeration.

Whereas α was estimated separately based on the gravity model since the gravity model measures spatial interaction of i and j that are weighted by the cost of their interaction:

$$T_{ij} = v_i^{\mu} w_j^{\alpha} f_{ij} \qquad \text{Equation 5.2}$$

$$f_{ij} = \exp(-\beta C_{ij}) \qquad \text{Equation 5.3}$$

Where:

T_{ij} = the number of trips by train between suburb i and j.

v = the size of the employed resident of the residential suburb i. It is measured based on each sector of the economy in question.

w = the size of the jobs in workplace suburb j. It is measured based on each sector of the economy in question.

f_{ij}= the time decay function of the gravity model. It represents the impedance to travel or the deterrence in the interaction between locations.

C_{ij} = travel time by park-and-ride between suburb i and suburb j.

β = the level of deterrence or impedance to travel. It also represents the willingness of individuals to travel, given the distribution of activities and the conditions of the transportation network.

This thesis used a negative exponential form of gravity model as specified in equation 5.3.

The decay parameter stated in the travel impedance function $(-\beta)$ the gravity model is then used as the decay parameter α in equation 5.1 (Iacono et al., 2008). These β values are modelled separately based on the sector of economy (total of all sectors, manufacturing, construction and retail.

5.1.2 Result

The modelling resulting from equations 5.2 and 5.3 produced the alpha values for all sectors to be 1.1; for construction, 1.3; for manufacturing, 1.1; and for the retail sector, 1.4. In comparison, the α value proposed in Graham (2007) was 1.0 for the manufacturing sector, 1.8 for the consumer business and service sector, and 1.6 for the construction sector. Other research has attempted to apply different α values. For example, the measurement of employment effective density in assessing the wider economic impacts of transport infrastructure investment in the north-west rail link project in Sydney assumed an α value of 1.11 for all industry types (Hensher et al., 2012).

Based on these results of α estimations, the effective density based on travel times in 2006 (before extension) and in 2011 (after extension) were then calculated based on equation 5.1. This thesis assumes that any increase in the effective density post-extension was a result of the changes in travel time between suburbs, and thus assumes a constant number of jobs and employed residents in the equation. There was an issue of data availability for the job dataset for 2006, since the conversion of different geographical boundaries provided by the Austalian Bureau of Statistics (ABS) censuses for place of work data in 2006 for this thesis was not successful. However, transportation improvements should be able to be related to effective job density without necessarily following the changes in physical job numbers, provided that travel times are used to indicate the relative positions of these job numbers with respect to the reference point (Hensher et al., 2012). Thus, improving travel times between locations may be directly translated into accessibility improvements, and as such, the reduction of 20% travel times between locations in a particular direction will have an equivalent effect to moving the employment 20% closer in that direction (Graham & Melo, 2010). Therefore, by this assumption, the increases in the effective density from inferred travel time improvements post-extension are expected to have an equivalent effect to having higher job density, without literally moving jobs or employed residents in the study area.

The associations between the changes in travel time and effective density, due to the Perth-Mandurah railway line extension, to the changes in train ridership were examined by a correlation analysis, in order to test the possibility of public transport-induced agglomeration.

The Perth-Mandurah extension has improved travel time and accessibility between locations, especially those suburbs located along the railway line. Assessment of this involves the calculation of delta travel times, based on matrices of travel time data between travel zones, produced by Department of Planning (DoP) between 2006 and 2011. These matrices were subject to a conversion into suburb matrices for the purpose of this thesis. The results of these travel time conversions have been input into Graham's effective density formula for each of the 2006 and 2011 datasets. The change in effective job density and effective employed resident density between 2006 and 2011 was then obtained. Obviously, the correlation between travel time changes and effective density changes may be expected to be high as the effective density is a function of travel time. However, this thesis is further interested in analysing the correlation between travel time changes and train ridership for observed census data for 2006 and 2011, and between the effective density changes and train ridership within the same period.

On the overall figures, table 5.1 informed the average Park-and-Ride travel time for the Perth metropolitan region was 109.6 minutes in 2006 and decreased to 79.15 minutes in 2011. In detail, this refers to the time for one suburb in the Perth metropolitan region to reach or to be reached from any other suburb on average, based on the Park-and-Ride travel time network. Thus, by 2011, the overall travel time had improved by approximately 38.5% on average.

Based on the effective density figures in 2006 and 2011 in table 5.1, the average metropolitan region agglomeration of jobs increased by 2,226 units (+27%) for the total agglomeration of all jobs, 82 (25.6%) in construction, 161 (26%) in manufacturing, and 90 (23.5%) in retail. The changes in agglomeration of employed residents consisted of 2,980 (33.5%) for the total employed residents, 120 (24.8%) in construction, 214 (29.75%) in manufacturing, and 90 (21.7%) in the retail sector on average. The manufacturing sector gained the highest changes among the three sectors for the 2006-2011 periods, both in jobs and effective employed resident density.

The changes in travel time have been examined in more detail based on the suburbs most adjacent to railway stations. These were limited to an area within 3 km of the train station, except for Clarkson, which has a larger distance of approximately 5 km, and Neerabup (around 8 km distance), since those two areas have very large job catchment areas and are located at the end of Joondalup line. There are 12 suburbs directly adjacent to stations along the Mandurah railway line: these as Perth (WA), Bateman, Bertram, Bull Creek, Cockburn Central, Como (WA), Coolongup, Jandakot, Leeming, Murdoch, Permelia, and Wellard.Twelve suburbs are also directly adjacent to stations along the Fremantle line, 10 suburbs on the Midland line, 17 suburbs on the Joondalup line, and 16 suburbs on the Armadale/Thornlie line.

As reported in the table 5.1, the Perth-Mandurah railway line extension has reduced travel time to and from suburbs located adjacent to the line (stations), by an average of 42.84 minutes (or a 57.29% improvement). The improvement in travel time for the other railway lines ranged from 33 to 38%. Overall, the changes in travel time and effective job or employed resident density, before and after the railway line extension, were much higher for suburbs adjacent to stations along the Mandurah line than those suburbs adjacent to other lines or to the average for the Perth metropolitan area.

The two-sample unequal variance t-test was conducted to show these similarities/differences in the changing rates between railway lines (table 5.2). Table

5.2 shows that the t-test concluded that there is enough evidence to reject the null hypothesis, i.e. the changes in travel times for the Mandurah line are significantly different to the changes for other lines at the 99% level of confidence. Conversely, the changes of travel time between railway lines other than the Mandurah line were similar in magnitude (there was not enough evidence to reject the null hypothesis at 99% level of confidence), with a few exceptions, such as that for the Fremantle to Armadale line. In short, the travel time improvement for suburbs on the Mandurah line was approximately 1.33 times larger in absolute-terms compared to other suburbs and 1.5 times larger in percentage changes.

The increased effective densities of the suburbs adjacent to the Mandurah line are equivalent to moving jobs in construction 34.26% closer to these suburbs, and manufacturing and retail 49.5% and 34% closer, respectively. Similarly, the improvements in travel time increased the effective density in a way that was equivalent to moving employed residents in construction 48.45% closer to these suburbs, and 52.7% for manufacturing and 40.5% for retail. The reductions of travel times for other railway lines were in the range of between 25 to 28%. These travel time improvements translate into an increase in effective density in suburbs adjacent to stations of around 11-30%.

The increase in agglomeration of employed residents was higher than that for jobs. Armadale experienced the highest increase after Mandurah, while Joondalup experienced the lowest increase. In terms of agglomeration of jobs, Armadale also had the highest increase after the Mandurah line, and Fremantle experienced the lowest. In short, the increase of effective density of jobs in absolute-term for suburbs along the Mandurah line was 1.78 times larger than that of the average for the Perth metropolitan region, and 1.585 times larger for effective employed resident density.

Table 5.1 The percentage changes in travel time and agglomeration for the period 2006 - 2011

Railway line	*Ttime changes in minutes (%)*	*Agglomeration of Job changes (delta increase and %)*				*Agglomeration of Employed resident changes (delta increase and %)*			
	Average value	All	Con	Man	Ret	All	Con	Man	Ret
Mandurah	-42.84 (57.29%)	3,963 (36.7%)	146 (34.26%)	277 (49.5%)	152 (34%)	4,722 (57.25%)	203 (48.45%)	358 (52.7%)	159 (40.5%)
Fremantle	-25.44 (34.5%)	2,306 (13.16%)	79 (13.18%)	155 (19.5%)	96 (11.3%)	2,997 (28.38%)	117 (24.3%)	200 (27.5%)	91 (20.5%)
Midland	-26.72 (38.02%)	2,364 (14.3%)	91 (14.4%)	193 (20.6%)	99 (13.1%)	3,126 (28.45%)	122 (21.9%)	217 (26%)	93 (18.7%)
Joondalup	-25.9 (33.99%)	1,928 (14.6%)	76 (12.3%)	126 (16.8%)	90 (12.3%)	2,934 (26.2%)	122 (17%)	195 (23%)	84 (14.3%)
Armadale	-27.15 (37.61%)	2,389 (16.4%)	88 (13.8%)	158 (15.7%)	101 (13.06%)	3,166 (29.6%)	123 (21.5%)	222 (25%)	96 (19%)
Average Perth metropolitan region	-32.3 (38.5%)	2,226 (27.18%)	83 (25.6%)	162 (26%)	90 (23.5%)	2,980 (35.5%)	120 (24.8%)	214 (29.75%)	90 (21.74%)

Table 5.2 The two-sample unequal variance t-test between railway lines on the percentage changes in travel time and agglomeration for the period 2006 – 2011

Travel time changes	*Mandurah line to*	*Fremantle line to*	*Midland line to*	*Joondalup line to*
Fremantle	1.519E-06***			
Midland	3.611E-06***	4.394E-02**		
Joondalup	2.536E-06***	2.363E-01	9.562E-02	
Armadale	5.096E-06***	8.886E-03***	2.562E-01	1.514E-02**
Effective job density (all sector) changes	***Mandurah line to***	***Fremantle line to***	***Midland line to***	***Joondalup line to***
Fremantle	1.394E-05***			
Midland	2.794E-05***	4.050E-01		
Joondalup	8.967E-07***	4.555E-02**	3.406E-02**	
Armadale	2.769E-05***	3.663E-01	4.597E-01	2.774E-02**
Effective employed resident density (all sector) changes	***Mandurah line to***	***Fremantle line to***	***Midland line to***	***Joondalup line to***
Fremantle	2.365E-06***			
Midland	4.262E-05***	3.406E-01		
Joondalup	6.250E-07***	4.076E-01	2.649E-01	
Armadale	3.414E-06***	2.517E-01	4.452E-01	1.674E-01

****statistically significant at 99% level of confidence*

***statistically significant at 95% level of confidence*

All others are not statistically significant at 95% level of confidence (not enough evidences to reject the null hypothesis).

The next figure in table 5.3 illustrates the changes in train ridership in all areas and for each railway line. This investigation has been conducted only for train ridership production (place of residence data) due to limitations on place of work data (jobs data set) from the ABS Census for 2006. The proportion of train ridership on average for the Perth metropolitan region was 0.0418 per capita of employed residents in 2006, and became 0.0632 in 2011. There were 42 per 1000 employed residents using the train as their main journey to work in 2006, and this figure increased to 64 per 1000 in 2011. This means an increase of more than 50% or a multiplier of 1.5328. Breaking down these figures into separate railway line data indicated more extreme increases in the level of train ridership for suburbs along Mandurah line, which was an 11.9-fold increase between 2006 and 2011. Changes in other suburbs were more moderate, such as 1.26 for suburbs along the Fremantle line, 1.2 for the Midland line, 1.25 for the Joondalup line, and 1.153 for the Armadale line within the same period.

Table 5.3 The percentage of changes in train ridership for the period 2006 - 2011

Railway line	*Proportion of train ridership*	*Proportion of train ridership*	*Increases*	*% Increases*	*Multiplication changes*
	2006	2011			
Mandurah	0.0078	0.0927	0.0849	1093.041	11.930
Fremantle	0.0718	0.0904	0.0187	26.005	1.260
Midland	0.0804	0.0966	0.0162	20.1548	1.2015
Joondalup	0.0898	0.1125	0.0226	25.203	1.252
Armadale	0.0729	0.0840	0.0111	15.264	1.153
Average Perth metropolitan region	0.0418	0.0632	0.0214	51.18	1.512

The correlations between the percentage of changes in effective density and the percent changes in travel time and train ridership based on data observed for all suburbs in the study area have been calculated. The results are presented in the table 5.4.

Table 5.4 Correlation value between the percentage of changes in travel time and agglomeration; and between the percentage of changes in train ridership and agglomeration for the period 2006 - 2011.

Agglomeration variable	*Travel time*	*Proportion of train ridership*
Effective job density (all sector)	0.89	0.51
Effective job density construction	0.66	0.70
Effective job density manufacturing	0.50	0.68
Effective job density retail	0.55	0.58
Effective employed resident density (all sector)	0.66	0.63
Effective employed resident density construction	0.66	0.59
Effective employed resident density manufacturing	0.64	0.75
Effective employed resident density retail	0.47	0.58

The correlation between the changes in travel time and the changes in effective density ranges from 0.5 to 0.9, meaning that there is a medium to very strong correlation between an increase in accessibility due to public transport investment and the increase in effective density. This suggests that a causal relationship between public transport investment and effective density or agglomeration may exist. Correlation between the changes in effective density and the changes in train ridership production within the same period ranged from 0.5 to 0.8, also a medium to strong correlation. This suggests a causal relationship between agglomeration and train ridership due to public transport investment may exist.

5. 2 THE SPATIAL PATTERN OF AGGLOMERATION: HOT SPOTS ANALYSIS

5.2.1 Method

The objective of this section is to use spatial modelling techniques to assess the degree of agglomeration or concentration, by detecting clustering phenomena or the high concentration of certain variables distributed within the study space. It has been hypothesised that cluster events should be located around train stations; the hotspot analysis was conducted to investigate if such clustering exists. The local statistics Getis-Ord G* detects pockets of spatial association for the hot spots of high value agglomeration.

Getis-Ord Gi* statistics are used to identify clusters of suburbs with higher employment rates than one might expect to discover by random chance (defined as Complete Spatial Randomness or CSR). The Gi* statistic is actually equivalent to a z-score. A high z-score for a suburb indicates a spatial clustering of high values for employment rate. The higher the z-score, the more intense would be the clustering. Therefore, the magnitude of z-score may be used to indicate the level of spatial agglomeration for the employment rate. To reject the null hypothesis (CSR), a z-score greater than 1.96 (the 95% confidence level) is used.

The Gi*statistics assumes that spatial association is the summed product of the variable in locality *i* that interacts with variable in locality j_s. This structure of this modelling is referred to as spatial association. The modelling is based on "all pairs of values x_i and x_j such that *i* and *j* are within distance *d* of each other" will follow the statistic (Getis & Ord, 1992). The Gi* statistic can be defined as (Ord & Getis, 1995):

$$G_i^*(d) = \frac{\sum_j w_{ij}x_j - W_i^*\bar{X}}{s\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}}, all\ j. \qquad \text{Equation 5.4}$$

Where:

w_{ij} = a spatial weight between suburb *i* and *j*. x_j is the effective density for suburb *j*; *n* is total number of suburbs.

$$W_i^* = \sum_j w_{ij} \qquad \text{Equation 5.5}$$

$$\bar{X} = \frac{\sum_j x_j}{n} \qquad \text{Equation 5.6}$$

$$s = \sqrt{\frac{\sum_j x_j^2}{n} - (\bar{X})^2} \qquad \text{Equation 5.7}$$

$$S_{1i}^* = \sum_j w_{ij}^2 \qquad \text{Equation 5.8}$$

Where:

The z score is defined as $z_i = (x_i - \bar{x})/s$, x is the effective density; w_{ij} is a binary (one or zero) or non-binary spatial weight matric based on queen contiguity method; $\bar{x}$ and s is sample mean and variance of the employment rate; n is number of suburbs.

The Getis-Ord Gi* parameter involves local statistics, which exclude the suburbs outside of a certain distance band of area. It focuses on the spatial proximity, and therefore potentially the spatial dependence of the relationships, between the employment rate of individual suburbs and their nearby suburbs. The hot-spot clustering of high values may indicate externalities of agglomeration (positive production externalities).

The ArcGIS 9.3 desktop Help function provides some descriptions of conceptual spatial relationships in GIS. The polygon contiguity method used in this thesis, is a spatial relationship based on the adjacency of the spatial units to each other. Weights are given as binary code, where 1 is allocated to those spatial units in direct adjacency to the spatial unit in question, and 0 is allocated to all of others spatial units. This option in GIS is known as "Polygon Contiguity Edges Only". On the other hand, the inclusion of neighbour spatial units based on edges (a boundary) and corners (nodes) is named "Polygon Contiguity Edges and Corners", which is also used in this thesis. This method is best used when the modelling involves is dealing with continuous data that represented as polygons (ESRI Inc., 2010).

5.2.2 Result

5.2.2.1 Cluster location

This section discusses the results of cluster analysis on effective employed resident density (agglomeration) and effective job density (agglomeration), for all economic sectors. The magnitude of agglomeration for all sectors has been measured using a decay parameter alpha of 1.1. Map in figure 5.1 and 5.2 presents the spatial clustering of these data. Out of all suburbs in the study area (a total of 316 suburbs), 40 suburbs indicated clustering of effective job density (figure 5.1), 26 of which contained hotspots with 99% confidence levels, and 9 of which were statistically significant at the 95% level of confidence, with the remaining 5 suburbs statistically significant at the 90% level of confidence. Among these 40 suburbs, 29 suburbs (or 72.5%) are located directly adjacent to stations. The percentage of effective job density accounted for by these 29 suburbs is 77.8% of the agglomeration in the hot spot areas, and 16.6% of total effective job density in the Perth metropolitan region.

On the other hand, 91 suburbs exhibited statistically significant clustering of effective employed resident density (figure 5.2), with 47 suburbs with hotspots at the 99% level of confidence, 28 at the 95% of confidence, and the remaining 16 suburbs at the 90% level of confidence. Among these 91 suburbs, 52 suburbs (57%) are located in directly adjacent to stations. The percentage of effective employed resident density accounted for by these 52 suburbs is 58.32% of agglomeration in hotspots areas and 20.9% out of the total effective employed resident density in the Perth metropolitan region.

Furthermore, there are 25 suburbs that represent both of high value or hotspots in job and effective employed resident density. These suburbs are generally located not only near train stations, but also close to the Perth CBD. Two clusters are concentrated along the Joondalup and Armadale lines, with a few patches along the Midland and Fremantle lines.

Interestingly, the recently-built Mandurah line, has generated higher effective employed resident density but none of the adjacent suburbs of the Mandurah line have generated high values of effective job density. This may indicate that the high percentage of changes in effective job density among suburbs adjacent to stations on the Mandurah line only emerged in those suburbs that had not yet merged into clusters of neighbouring suburbs. This often means those suburbs with high rate of job growth are neighouring with those suburbs with a low rate of job growth, thus, no cluster of hot spots is detected.

Clustering of effective employed resident density potentially generated an amount of 15,080 total working person opportunities for every suburb on average, and clustering of effective job density potential brought 18,884 jobs opportunities for each suburb on average. The map in figure 5.1 and 5.2 shows the position of clusters with respect to distance from train station.

Table 5.5 The percentage of cluster of effective employed resident density and effective job density in the manufacturing sector

Sector Industry	***Description***	***Number of Suburban***	
		Effective employed resident density	*Effective job density*
All sector: alpha (decay parameter of agglomeration) is 1.1	Number of suburb with hotspots	91	40
	Number of suburbs with hotspots around station	52	29
	Cluster location	3 suburbs in the city area, 17 suburbs along Joondalup line, 6 in Midland, 11 in Mandurah, 10 in Armadale, 5 in Fremantle line.	3 Suburbs in the city area, 8 suburbs in Joondalup line, 3 in Midland line, 10 in Armadale line, 0 in Mandurah line, 5 in Fremantle line.
	Agglomeration magnitude	15,080 opportunities	18,884 opportunities

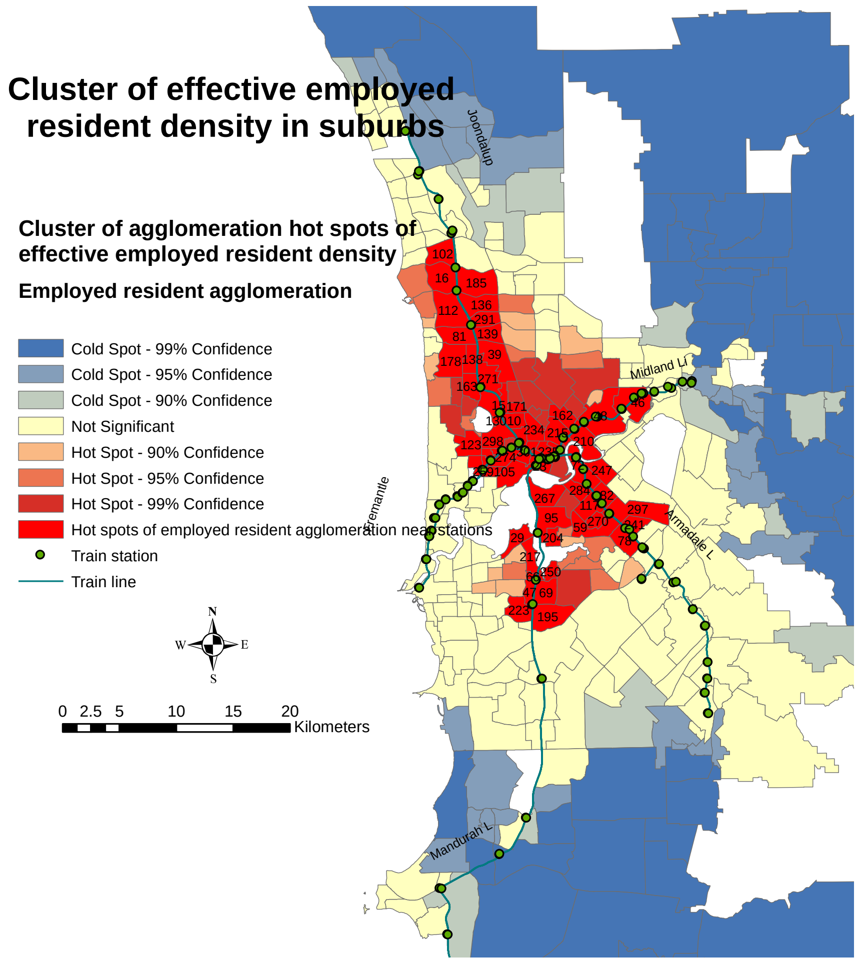


Figure 5.1 Cluster of agglomeration hot spots of effective job density in suburbs

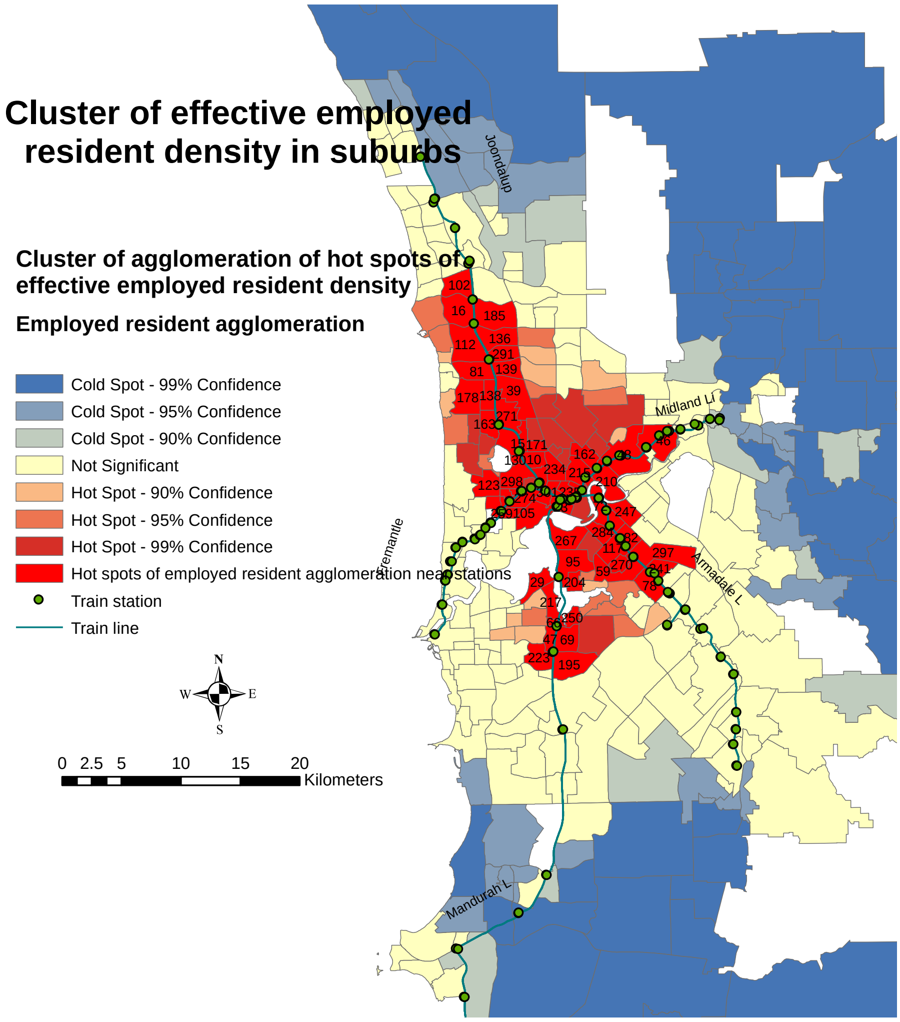


Figure 5.2 Cluster of agglomeration hot spots of effective employed resident density in suburbs

5.2.2.2 Effect on job-housing balance of location disparities in agglomeration clustering

The distributions of job and effective employed resident density hot spots have been further compared to the distribution of the main employment centres and main residential areas. Following the concept of job-housing balance (i.e. the ratio of jobs to employed residents), a 'main employment centre' is defined as the suburbs with job-housing balance above 1.5. Similarly, a'main residential area' is defined as the suburbs with job-housing balance below 0.75.

Map in figure 5.1 and 5.2 describes suburbs located around train station with high effective job density and high effective employed resident density. From 29 suburbs, there are 13 suburbs near train stations that contain effective job density hot spots that are also main residential areas, and 13 suburbs containing effective job density hot spots that are also main workplace areas. In addition, from 52 suburbs, 34 suburbs near train stations containing effective employed resident density hotspots are also main residential areas; and 14 suburbs near train stations containing effective employed resident density hotspots are also main workplace areas. These findings imply that the main residential areas near stations are likely to be high in both effective employed resident density and effective job density.

Similarly, main workplace areas near stations are likely to be high in both of agglomeration types. This also means that the determinant of train trip production in a residential suburb may consist of both the number of effective employed residents and the effective job density. Similarly, the determinant of train trip attraction in a workplace suburb may consist of both the effective job and effective employed resident density. These variables will be tested further for hypothesis 2 and 3 in chapter 6 and 7.

5. 3 THE SPATIAL PATTERN OF AGGLOMERATION: SPATIAL DECAY PHENOMENA

5.3.1 Method

Cluster analysis may be able to identify suburbs with high agglomeration values and show clustering with other surrounding suburbs. Geographical analysis of Getis-Ord Gi* presents the spatial distribution of hot spots to differentiate those that fall near stations from those that do not. However, cluster analysis cannot distinguish to what degree the spatial pattern of agglomeration values attenuates with distance from train stations.

As per the hypothesis for this thesis, the value of agglomeration of suburbs is expected to show distance decay away from stations. Since one suburb carries one value of agglomeration, it was difficult to assess this pattern due to lack of more high resolution on observations measured from one station catchment. Thus, the distributional pattern of agglomeration values of suburbs required a more detailed examination. Fishnet-based data have been used in order to allow for such an analysis.

Consider the illustration of fishnet-based data below, representing a buffer ring type dataset (Figure 5.3). The fishnet rings are constructed based on 1 km distance buffer rings from train stations along the catchment area (16 km) as determined in the section 4.1.2.15 in chapter 4. The number of fishnets (grid cells) and the agglomeration values in each fishnet are calculated for each ring. The cumulative distribution is calculated based on the probability that a random variable *X* located within a pre-specified distance *r* is observed to take on a value less than or equal to some known value for that distance ring. This thesis attempts to determine whether the distribution of random variable *X* (such as agglomeration) across *n* fishnets would fall within the known predetermined distance ring buffer as described in the following figure 5.3.

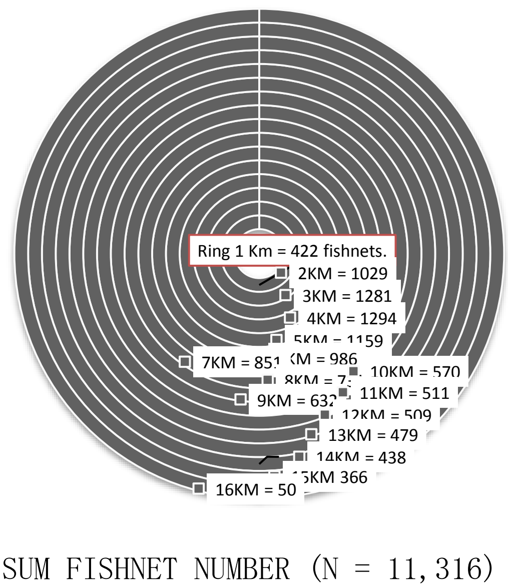


Figure 5.3 Illustration of fishnet were distributed across 1 km buffer rings from train stations (location zero) to the edge of the catchment (16 km).

Station proximity in this thesis is regarded as a continuous variable. The distance within each fishnet *i* was measured as the linear distance from the *i*th fishnet centroid to the nearest station centroid. This thesis assumes that railway station impacts occur up to a maximum distance of 16 km, on the basis of park-and-ride services data, as reported in chapter 3. The spatial decay model in this thesis is based on an exponential decay model, where the dependent variable y is the probability density function of the average agglomeration value of *n* number of fishnets located within a pre-specified distance ring *r*.

For each value of agglomeration in each fishnet, the probability density and cumulative probabililty density function are calculated. These are written in the following equations:

$$a = \sum_{i=1}^{R} a_i \quad \text{Equation 5.9}$$

$$b_i = {a_i}/{a} \quad \text{Equation 5.10}$$

$$f(r) = \sum_{i \leq r} b_i \quad \text{Equation 5.11}$$

$$F(r) = \sum_{i \leq r} a_i \quad \text{Equation 5.12}$$

Where:

a_i = the value of variable (effective density) of a fishnet located in ring *i*

b_i = the average value of fishnets located in ring *i*

a = the total value of effective density over all fishnet *i* over all ring *R*.

R = 16.

$f(r)$ = the probability density of fishnets in each ring *i*.

$F(r)$ = the cumulative of probability density over all ring *i* to *R*, where *R* = 16.

Equation 5.9 calculated the total value of agglomeration from each ring and summed over all 16 rings. Equation 5.10 calculated the average of agglomeration in each ring. Equation 5.11 calculated the probability density of agglomeration across each ring. Equation 5.12 calculated the cumulative density of the probability density function from all rings.

5.3.2 Result

An analysis was conducted for the all sector model, for the effective employed resident density (all sector) and the effective job density (all sector). The calculated results for the probability density function and each of its key variables are discussed below with reference to the following graphs.

Figures 5.4 and 5.5 showed the distribution and cumulative distribution of effective employed resident density. A location in ring 0–1 km from a train station contains 10.8% of the total agglomeration of employed residents. Likewise, a location in within 1-2 km from a train station covers 10.5% of agglomeration. Based on the probability density function, the average agglomeration of employed resident of the fishnets located within 0-1 km from a train station is 1.03 times higher than the next fring for 1-2 km distance. The cumulative density function also shows that fifty percent of the effective employed resident density is distributed around 0-5 km from train stations. The level of effective employed resident density gradually decreases as it moves further away from the train station. The density values sharply decline between 4 km and 9 km from the train station, and then more gradually decrease after 9km, and up to 15 km. Beyond 15 km, less than two% of effective employed resident density remains in the catchment area. The overall exponential distance decay is defined by a 9.5% decrease in agglomeration for every 1 km increase in distance from the nearest train station.

A similar pattern emerged for the distribution of effective job density. Figures 5.6 and 5.7 showed the distribution and cumulative distribution of effective job density. More than 13% of effective job density was found to be concentrated within the 0-1 km buffer. Likewise, the 1-2 km zone contained 12.5% of agglomeration. Based on the probability density function, the average agglomeration of job for fishnets *i* located within 0-1 km from a train station was 1.055 times higher than the next following ring of 1-2 km. Based on the cumulative distribution, 53% of effective job density is distributed around 0-5 km from a train station. The level of effective job density gradually decreases away from train station. The value sharply declines between 2 km and 9 km from the train station and is almost steady between 9 and 13 km. After 13 km, there is less than 2% of effective job density (1.8%) in the catchment area. The overall exponential distance decay is defined by a 11.2% decrease in effective job density for every 1 km increase in distance from the nearest train station.

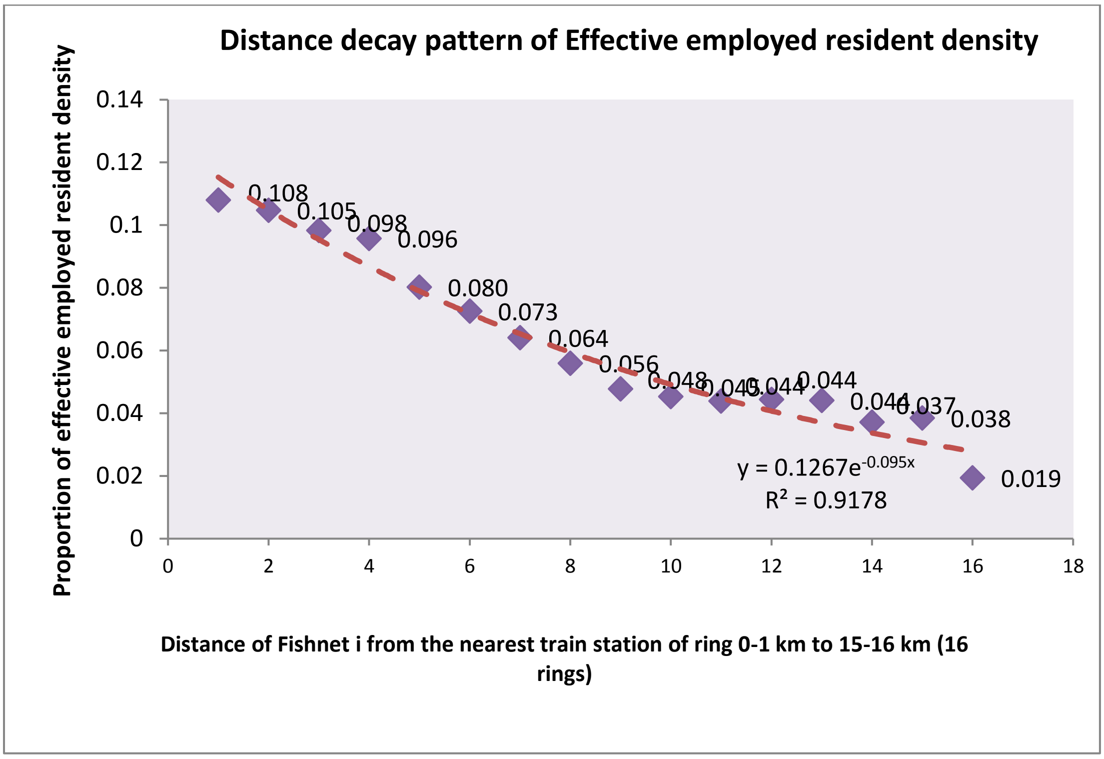


Figure 5.4 The spatial decay of effective employed resident density from train station

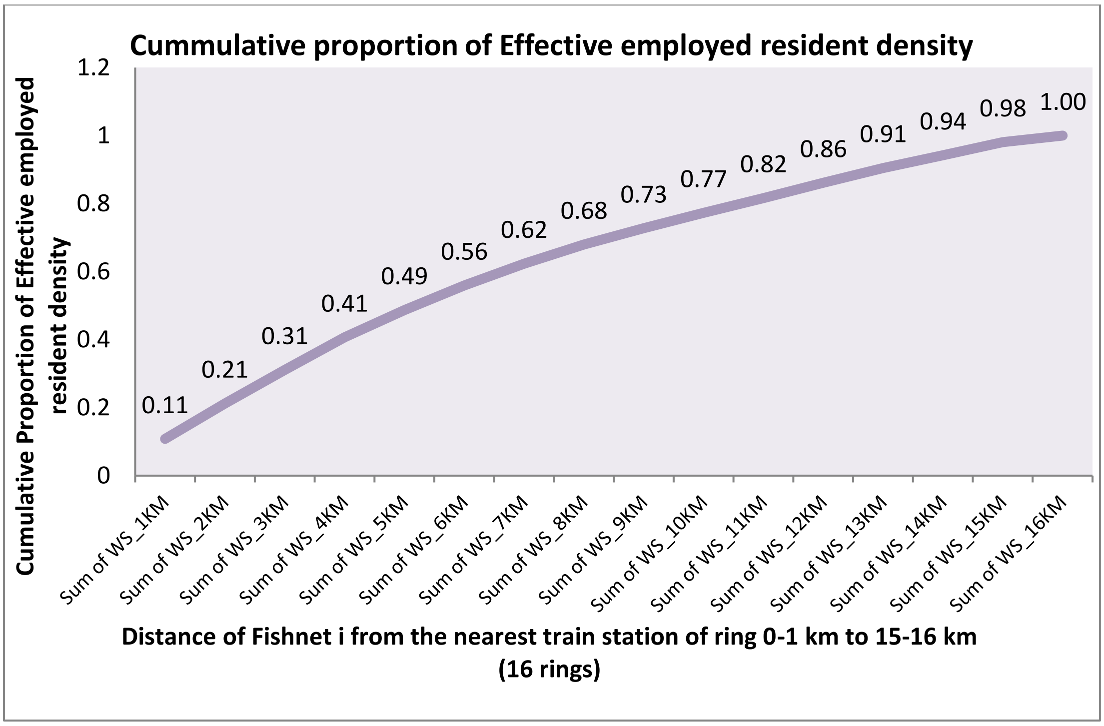


Figure 5.5 The Cumulative distribution of effective employed resident density

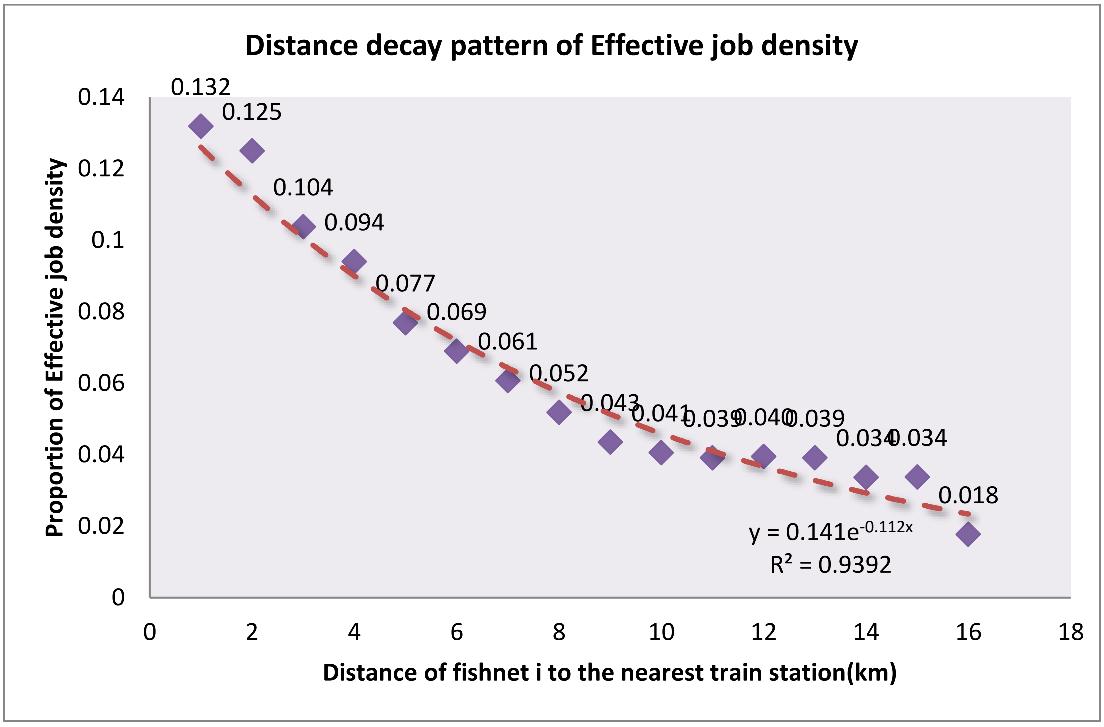


Figure 5.6 The spatial decay of effective job density from train station

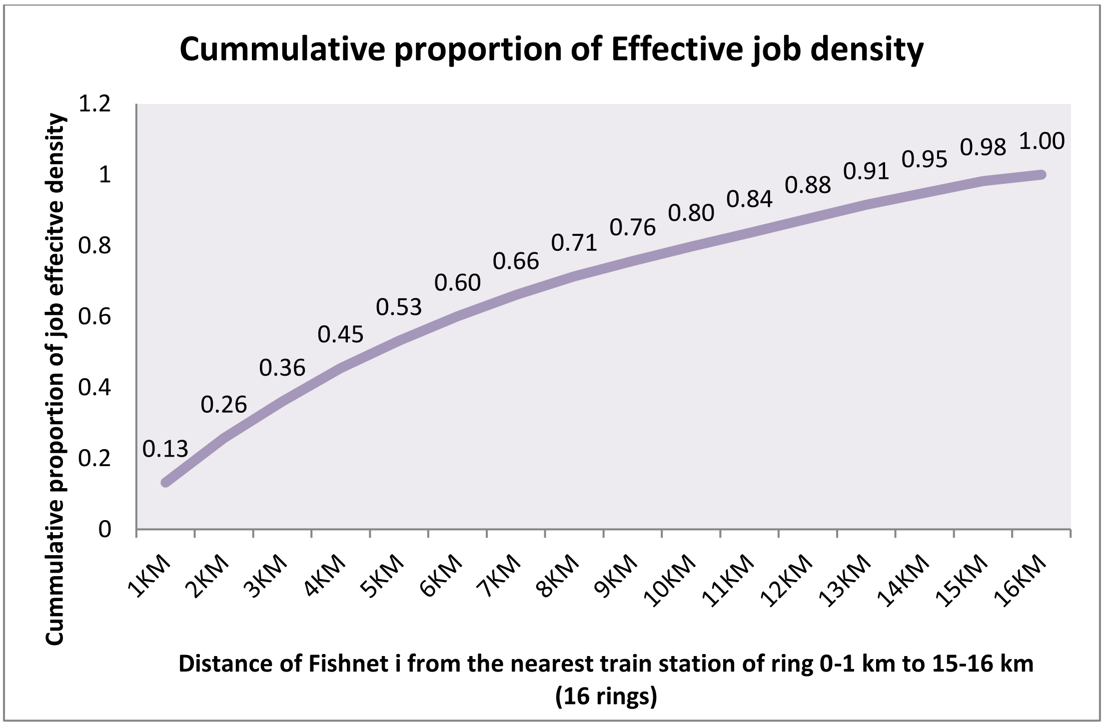


Figure 5.7 Cumulative distribution of effective job density

Overall, the spatial decay of effective job density was higher than that of effective employed resident density. Thus, based on the continuous distance measure, the impact of station proximity on the increase of effective density for every 1 km closer to the station was 11.2% for effective job density and 9.5% for effective employed resident density.

In relation to the impact of station proximity to the changes in effective density in within period 2006 – 2011, it was found that the railway line extension of Perth-Mandurah brought an increase in effective job density at the annual rate of 5.46% on average. The absolute-increase in effective job density from 2006 to 2011 was 2,226 on average per suburb. Suburbs near train stations experience higher annual rates of change. The exponential decay model, applied to these rates of change, indicated a 2.28% increase in the annual rate of change of effective job density for every 1 km closer to the train station, but an 11.9% increase in the absolute-changes in effective job density between 2006 and 2011 for every 1 km closer to the train station. In addition, the new extension produced an increase in effective employed resident density at the annual rate of 6.256% on average. The absolute increase in effective employed resident density from 2006 to 2011 was 2,980 on average per suburb. The exponential decay model, applied to these annual rates of changes, indicated a 2.21% increase in the annual rate of change of effective employed resident density for every 1 km closer to train station, but an 11.3% increase in the absolute-changes in effective employed resident density between 2006-2011 for every 1 km closer to train station. This indicates evidence in favour of the effect of transportation investment on public transport induced agglomeration, for both employed resident and effective job density. The railway line extension has brought a positive impact on faster growth of effective employed resident density, with slightly larger absolute-changes than that of effective job density during 2006 - 2011, and these impacts decay more rapidly with distance from the train stations for effective job density than that of employed residents.

This may suggest that the business sector was more responsive to increases in proximity to the train station than the residential sector. This may make sense, since locations close to stations offer higher accessibility and therefore become more attractive to business sector, although these locations also attract higher residential activities. For example, studies have shown that the impact of railway proximity on property values has been higher for commercial property than that of residential property. Measuring the impact of station proximity, a study on meta-analysis of railway station proximity effects on property value, based on 57 research observations, found that the impact of a station on property value or rent was 2.61% greater on average for every 250 m closer to the station, while the reduction of commercial property prices was 4.8% for every 250 m further from stations and that for single family residential property was 2.4% (Rietveld et al 2007).

5. 4 CHAPTER SUMMARY

This chapter has discussed the evidence for public transport induced agglomeration. In order to provide a base argument that public transport induced agglomeration has resulted from transport infrastructure developments such as railway line extension or station development, this chapter discussed three possible measures of public transport induced agglomeration. All three of the measurements suggest the influence of public transport induced agglomeration. This thesis found a strong correlation between the changes in travel time, the changes in effective density, and the changes in train ridership production over the period 2006-2011, and those changes have been substantially higher for suburbs located along the Perth-Mandurah railway line, where the extension has taken place. Clusterings and spatial distance decays for areas of high effective density were shown to exist near stations.

The findings regarding public transport induced agglomeration in this chapter, and their level of influence on train ridership were applied in the model of train ridership prediction and attraction (chapter 6 and 7). Further implications of model results are discussed in chapter 8.

CHAPTER 6. COMPARING THE LUTI AND SETI FRAMEWORKS FOR DERIVING TRAIN RIDERSHIP DETERMINANTS

6. 1 INTRODUCTION

This chapter discusses the second research hypothesis in relation to the model for train ridership prediction. The model is divided into two components: (1) the one-way train trip production (the working trips that are produced from residential suburbs that involve only train journeys or a combination of modes that involve train); and (2) the one-way train trip attraction (the working trips that are attracted to workplace suburbs, involving only train journeys or a combination of modes that involve train).

This section describes the pattern of train ridership in the study area and how this thesis contributes to the current discourse on the determinant factors of train ridership. The discussion links the research hypotheses (see Chapter 3) to the outputs of the models in understanding the key train ridership factors. This thesis may be differentiated from other research in train ridership modelling in the following ways:

- Train ridership modelling approaches in the literature have been divided into either the "system" approach using aggregate studies, such as at the suburb level (Guerra & Cervero, 2011), or the "station" or "corridor" approach using dis-aggregated modelling at the individual or personal choices level (Cervero, 2006; Cervero, Murakami, & Miller, 2009; Guerra & Cervero, 2011). The current research approach has been an aggregate or "system" approach that utilizes data on a suburban geographical unit of analysis. A system level analysis has the advantage of good availability of data across the system and over time, and allows for analysis with a wide range of parameters (Guerra & Cervero, 2011).

- Predictor variables consist of external factors for determining transit (such as population, land uses and the economy) instead of the internal factors of transit (such as train frequency services or the fare level). External factors have ben found to have greater effects on train ridership than the internal factors (Taylor et al., 2009).

The first part of this discussion involves factors that influence train ridership under the LUTI (without agglomeration) and the SETI model (with agglomeration) for both the train trip production and attraction model. This process has also been part of testing hypothesis 2. Factors are related to an individual/person (socio-demographic/economic) component, transportation (accessibility), place (land use) and agglomeration (effective density). Discussion in this section emphasizes the all-sector model, which covers total jobs and total employed residents from all 19 sectors.

The second part of the chapter summarises the determinant factors in train ridership prediction. This part of the chapter is preceded by a description of the current train ridership pattern, i.e. the proportion of employed residents who use train services in suburbs (train trip production); and the proportion of jobs in suburbs that are attracted working trips by train (train trip attraction).

The train trip production and train trip attraction models that are discussed in this chapter will be compared across different versions of the LUTI and SETI models, both are discussed based on the all-sector model (see chapter 4 for explanation of data and model structure):

(i) The LUTI model.

(ii) The LUTI model that incorporates job or employed resident density with the interaction term.

(iii) The SETI model h1, i.e. the model that incorporates effective employed resident density without the interaction term.

(iv) The SETI model h1b, i.e. the model that incorporates effective employed resident density with the interaction term (i.e. the interaction between effective employed resident density and average distance of suburb to train station).

(v) The SETI model h2, i.e. the model that incorporates effective job density without the interaction term.

(vi) The SETI model h2b, i.e., the model that incorporates effective job density with the interaction term

First, maps 6.1 and 6.2 present the pattern of train trip production (place of residence dataset) and train trip attraction (place of work dataset). The map of the proportion of employed residents travelling by trains leaving residential suburbs (map 6.1) shows that concentrations of train ridership have emerged not only within areas in adjacent to train stations, but also within areas that are 3-5 km from the station (such as in areas around Mandurah and Armadale line services) and a few areas beyond 5 km from the train station (such as at the northern area of Butler Station). Moreover, the level of train trip production out of the CBD area is generally lower than the average for the Perth metropolitan area. Murdoch Station, which has been well connected to bus services, has shown higher train trip production than the average, not only generated from the suburbs of Murdoch station, but also from its surrounding suburbs. A similar situation has been identified in areas around Wellard Station, Bull Creek Station, Cockburn Central Station, and Thornlie Station.

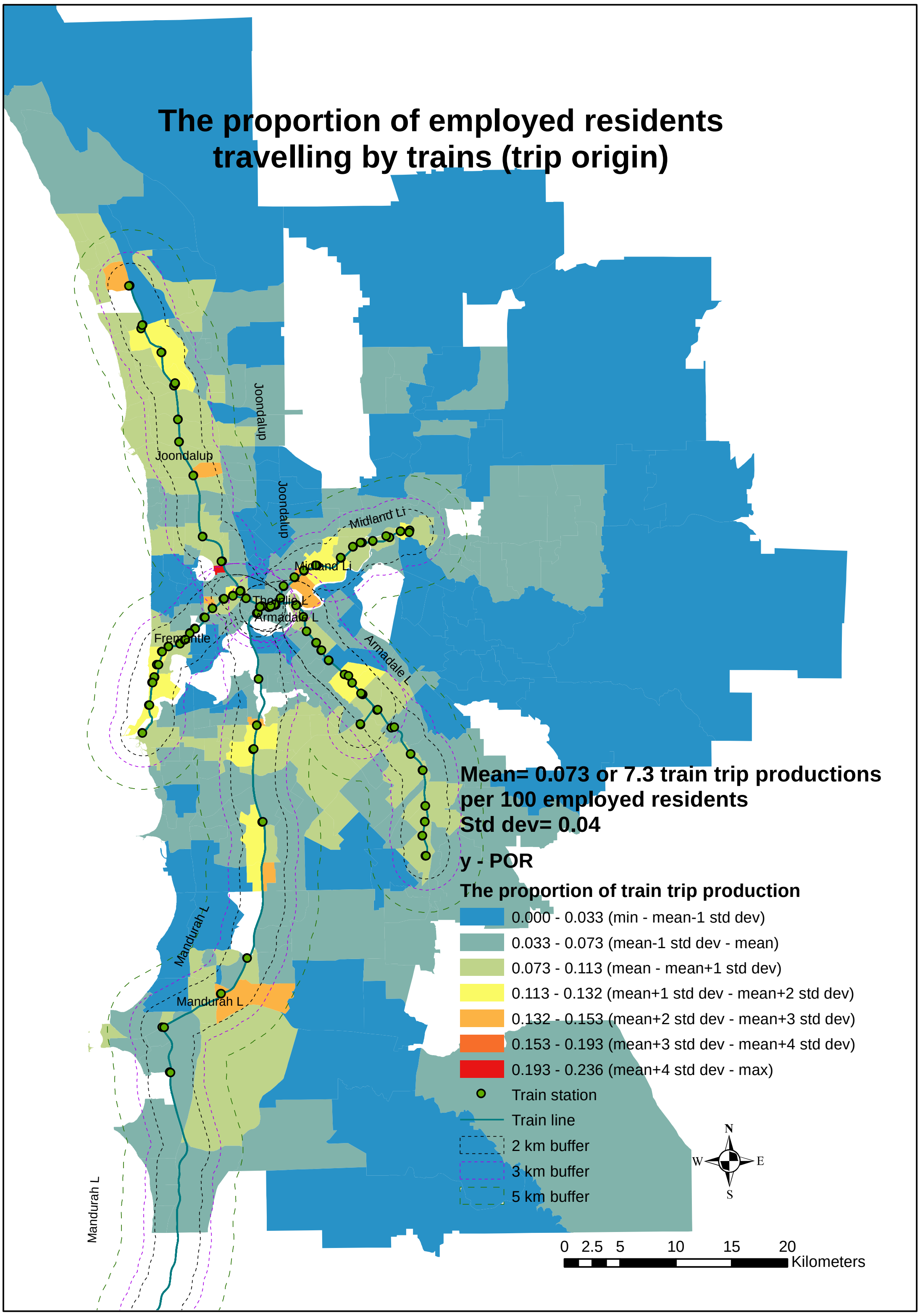


Figure 6.1 Map of the proportion of train trip production

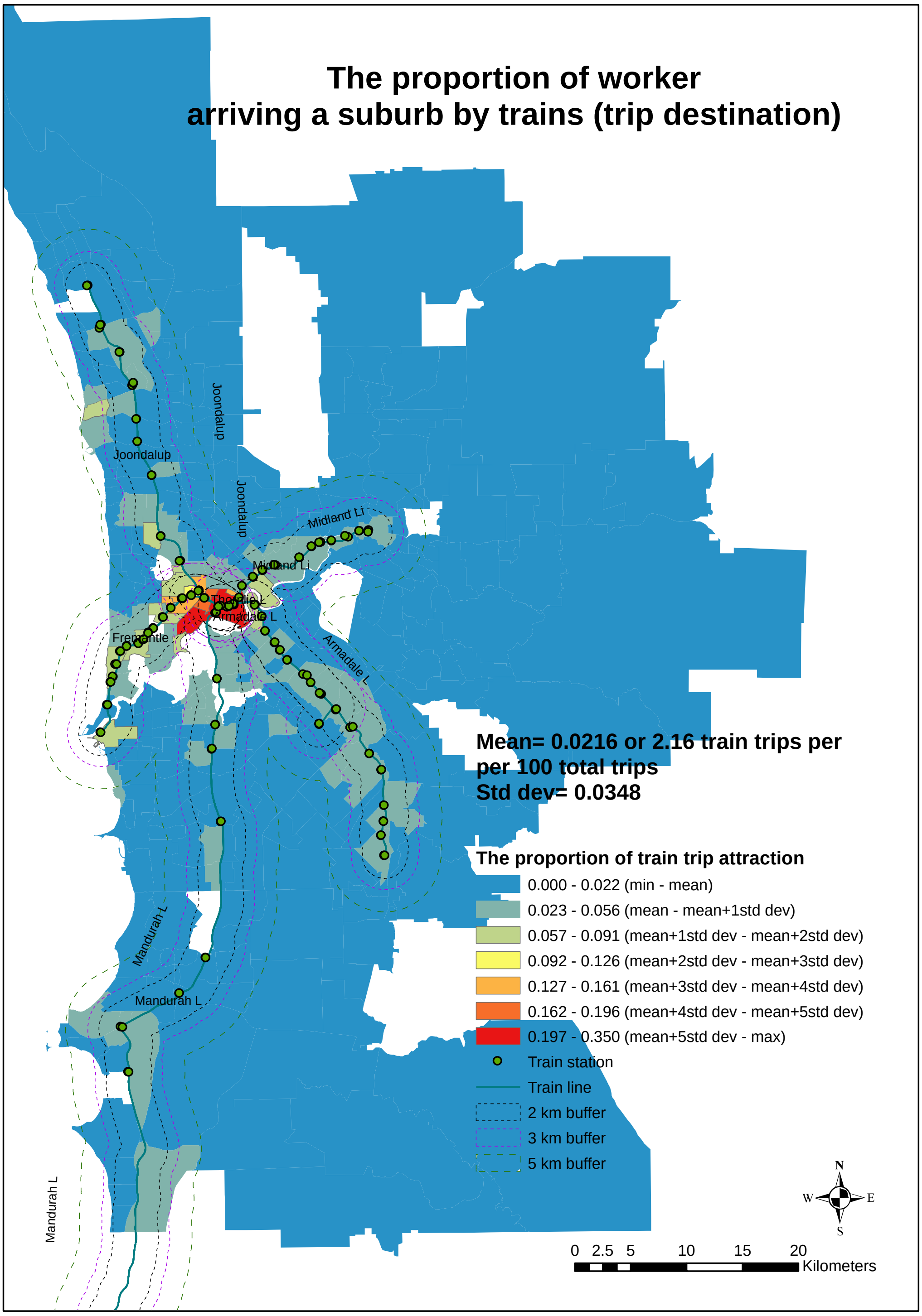


Figure 6.2 Map of the proportion of train trip attraction

Map 6.2 shows the proportion of train trips arriving at a destination suburb. The proportion of train ridership is generally higher to areas within the Perth CBD and along the Fremantle line, then gradually reduces in suburbs located adjacent to stations along the Midland and Armadale lines, around the inner suburbs of the Perth metropolitan area. The high attraction of Perth city together with West Perth, Northridge and East Perth is probably related to the current public transport planning policy of supporting increasing demands for access for area in the Perth CBD. The opening of the Perth-Mandurah railway line for the last 8 years, in accordance with the assignment of Rockhingham and Mandurah as strategic metropolitan centres and Murdoch and Jandakot stations as specialised centres, has contributed to a statistically significant proportion of train attraction above the overall mean for the metropolitan area of trips attracted to the Perth CBD.

6. 2 METHOD OF EXPLORATORY ANALYSIS: TRAIN RIDERSHIP FACTORS UNDER THE LUTI AND THE SETI MODEL (THE MULTIPLICATIVE EFFECT)

This section explores the relative strengths between predictor variables (X). It is necessary that the comparison of the strength of any relationship between each independent variable and dependent variable be held based on the value of the standardized beta coefficient. The standardized beta coefficient is a scale-free index and therefore can be compared across different variables. The coefficient can indicate the relative importance of the variables with which they are associated (Pedhazur, 1982). Standardized beta coefficient regression has been used to compare how important the role of certain variables compare to others based on their standardized value (z-score), in order to provide a basis of comparison between different types of independent variables.

The standardized value of y and x (or the z-score) is applied in the regression equation when the SPSS software calculates the standardized beta coefficient. The regression equation is written in a form of (Pedhazur, 1982):

$$z_y = \beta_1 z_{x1} + \beta_2 z_{x2} + \cdots + \beta_n z_{xn} \quad \text{Equation 6.1}$$

Whereas:

z_y = the standard score of dependent variable.

z_x= the standard scores of independent (predictor) variables.

n = the number of all predictor variables.

the standardized coefficient β = "the expected change in the dependent variable of y, expressed in standard scores, associated with a one standard deviation change in *x*, while holding the remaining variables constant" (Pedhazur, 1982).

Dependent variables have been specified based on the standard z-scores of the logarithm of proportion of train ridership (see chapter 4). The relationship of predictor variables on dependent variables has been discussed, based on the multiplicative effects of coefficients in the form of the logarithmic transformation in OLS regression (Taplin, 2016). Thus, the relative strengths between predictor variables may be compared based on the standardized beta coefficient in terms of the multiplicative effects of variables on train ridership instead of their marginal effects. The exponential of the standardized beta coefficient reflects the multiplicative effects on the proportion of train ridership (a dependent variable) of the associated predictor variables, holding all other variables constant. Therefore, equation 6.1 has been adjusted to accommodate the logarithm transformation of y and to allow for the multiplicative effect of *x* on *y*.

The OLS regression of the logarithm of y is given by:

$$\ln(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad \text{Equation 6.2}$$

Based on the equation 6.1, the equation 6.2 is adjusted into:

$$z_{\ln(y)} = \beta_1 z_{x_1} + \beta_2 z_{x_2} + \cdots + \beta_n z_{x_n} \quad \text{Equation 6.3}$$

The relative or the multiplicative effect on *y* from an increase in z_{x_1} to $z_{x_1 1}$+Δ_1 is given by:

$$\frac{z_y(z_{x_1}+\Delta_1, z_{x_2})}{z_y(z_{x_1}, z_{x_2})} = \exp {\beta_1}^{\Delta_1} \qquad \text{Equation 6.4}$$

Provided z_{x_1} increased by Δ_1, then z_y is multiplied by exp ${\beta_1}^{\Delta_1}$. The advantage of this concept is that the multiplication magnitude is not dependent on the magnitude of *x* or *y*. This condition has significantly simplified the interpretation of the results of the model. Taplin (2016) pointed out:"in many cases, interpreting coefficients in terms of multiplicative or relative effects is more appropriate and simpler" (Taplin, 2016, p.3) instead of the marginal effect in which "many of these (marginal effects) interpretations are incomplete or misleading" (Taplin, 2016, p. 6). More explanation about why the multiplicative effect is a superior interpretation and how to report model results and interpretation of regression coefficients for analysis of limited dependent variables can be found in Taplin (2016).

Overall comparison of the relative strengths between the socio-demographic, transportation, and land use variables were made based on the results of the regression model that produced the highest adjusted R-square. Specific comparison of the relative strengths between the traditional density variable and the effective density variable were also determined.

6. 3 RESULTS OF THE EXPLORATORY ANALYSIS: COMPARING THE LUTI AND THE SETI MODEL

This thesis proposes the method of interpreting beta coefficients in terms of the multiplicative or relative effect, instead of the marginal effect, as commonly found in other literatures. The meaning of the standardized beta coefficient then represents the degree of multiplicative effect on the proportion of train ridership (standard scores) for an increase of one standard deviation in the predictor variables. Therefore, all of model interpretations based on the standardized beta coefficient reported in this section are derived from equation 6.4.

The multiplicative effect on train ridership of increasing each predictor variable by one standard deviation is presented in tables 6.1 and 6.2. In the train trip production model, the interaction term in the SETI h1b model was not statistically significant, but was statistically significant in SETI h2b. Therefore, table 6.1 shows the SETI h1 instead of SETI h1b. Table 6.1 shows a comparison for train trip production under the LUTI, the SETI h1 and the SETI h2b model. For example, increasing one standard deviation of the proportion of employed resident in the manufacturing sector (st dev=0.05) has the effect of multiplying the proportion of train trip production by 1.143 (equals to an increase of 14.3%) in the LUTI model, and by 1.132 (equals to an increase of 13.2%) in SETI h1 and SETI h2b on its standard scores.

Table 6.1 The multiplicative effect of all statistically significant predictor variables on train ridership modelled in LUTI and SETI for place of residence (train ridership production). Detailed figures in appendix 11, 19, and 22.

Model train trip production: statistically significant variable	*Mean*	*Std. dev.*	*Standardized beta coefficient (Exponential of standardize beta coefficient)*		
			The LUTI model **Appendix 11 All sector LUTI**	***The SETI-with h1*** **Appendix 19 All sector SETI H1 effective employed resident density**	***The SETI-with h2b*** **Appendix 22 All sector SETI H2B effective job density with the interaction term**
Pblu_r	0.33	0.115	-0.257 (0.773)**	Not statistically significant	Not statistically significant
P_worker_man	0.0889	0.05	0.134 (1.143)**	0.124 (1.132)**	0.124 (1.132)**
Inc_pr_r	36.39	7.45	-0.42 (0.656)***	-0.407 (0.666)***	-0.369 (0.69)***
Car_own	1.88	0.4	-0.292 (0.747)***	-0.317 (0.728)***	-0.311 (0.7327)***
Sqrt_ptiori	746	581	Not statistically significant	-0.224 (0.799)***	Not statistically significant
Ln_strio	9.35	0.23	0.39 (1.477)***	0.507 (1.66)***	0.455 (1.576)***
Ln_avedist	1.515	0.687	-0.538 (0.583)***	-0.518 (0.596)***	-0.858 (0.424)***
Ln_lvr	6.77	1.035	-0.31 (0.7312)***	-0.308 (0.735)***	-0.34 (0.71)***
Ln_jwr	0.675	1.5	0.144 (1.15)**	0.145 (1.156)**	0.168 (1.184)**
Eder (in 1000 units)	11.856	2.77	n/a	0.361 (1.435)***	n/a
Ejd (in 1000 units)	10.416	4.7085	n/a	n/a	-0.24 (0.787) **
Interaction term of ejd (ejd_ln)	n/a	n/a	n/a	n/a	0.276 (1.318)**

Description:
**) all variables are statistically significant at p<0.05
***) all variables are statistically significant at P<0.01
Exponential value of the standardized beta coefficient is inside the bracket

In the SETI h1 model of train trip production, transportation variables such as road network travel distance and the distance of residential suburbs to train stations had a greater effect than the effective density variable and the land use variable. Only the proportion of employed residents in the manufacturing sector and the effective density produced a positive relationship with train ridership. Other variables, such as the proportion of employed residents in blue collar occupations, income, car ownership, public transport network supply coverage, road network travel distance, land rent, and job housing balance, produced negative relationships with train ridership.

In the train trip attraction model, the interaction term is statistically significant in both SETI h1b and SETI h2 b. Therefore, table 6.2 presents LUTI, SETI h1b and SETI h2b. For example, increasing one standard deviation of the number of all jobs across 19 sectors (std dev= 5,406 jobs) has the effect of multiplying the proportion of train trip attraction by 1.12 (equals to an increase of 12%) in the LUTI model and the SETI h1b model but has not statistically significant effect in the SETI h2b. This means that the standardized beta coefficient of the "all jobs" variable is significantly different to zero in LUTI and SETI h1b and that therefore the multiplicative effect on the proportion of train trip attraction is significantly different to 1. However, the standardized beta coefficient of the "all jobs" variable is not significantly different to zero in the SETI h2b model, thus the multiplicative effect is equal to 1.

The best model in terms of the highest adjusted R-square value of train ridership attraction has been shown by table 6.2 to be the SETI h1b model. Furthermore, table 6.2 indicates that the most important variable influencing ridership in SETI h1b is the effective density variable (shown as *Eder in the variables listing*). However, the real magnitude of influence of this variable is dependent on the distance between workplace suburbs and train stations. At maximum (assumed the minimum distance of destination suburb among the dataset to be 1.164 km for suburb Woodbridge (WA), the multiplicative effect of the effective employed resident density for this suburb is 1.4525 (or an increase of 45.25%), which is higher than the influence of job density and the transportation variables. Besides job density, other land use variables do not appear to be statistically significant. The influence of socio-demographic/economic variables is lower than the influence of the land use and transportation components.

Table 6.2 The multiplicative effect all statistically significant predictor variables on train ridership modelled in the LUTI and SETI for place of work (Train ridership attraction). Detailed figures in appendix 35, 44, and 46.

Model of train trip attraction:	*Mean*	*Std. Dev.*	*Standardized Beta coefficient (Exponential of standardize beta coefficient)*		
			The LUTI model ***Appendix 35 All sector LUTI POW model***	***The SETI-with h1b*** ***Appendix 44 All sector SETI H1B effective employed resident density with the interaction term – POW Model***	***The SETI-with h2b*** ***Appendix 46 All sector SETI H2B effective job density with the interaction term – POW Model***
All job	2010.74	5405.551	0.12 (1.1275)**	0.106 (1.112)**	Not statistically significant
Pjobret	.117	.082	Not statistically significant	0.139 (1.149)**	0.133 (1.142)**
ln_strio	9.355	.230	0.138 (1.148)**	0.151 (1.163)**	Not statistically significant
ln_avedist	1.515	.687	-0.341 (0.711)***	Not statistically significant	Not statistically significant
sr_lvnr			0.118 (1.1252)*	Not statistically significant	Not statistically significant
ln_jobd	4.959	1.745	0.292 (1.339)***	0.372 (1.451)***	0.326 (1.385)**
eder (in 1000 units)	11.857	2.771	n/a	0.446 (1.562) ***	n/a
Interaction terms with eder (eder_ln)	n/a	n/a	n/a	-0.479 (0.619)***	n/a
ejd (in 1000 units)	10.416	4.708	n/a	n/a	0.443 (1.557)***
Interaction terms with ejd (ejd_ln)	n/a	n/a	n/a	n/a	-0.272 (0.762)***

Description:
*) all variables are statistically significant at $p<0.10$
**) all variables are statistically significant at $p<0.05$
***) all variables are statistically significant at $P<0.01$
Exponential value of the standardized beta coefficient is inside the bracket

The process of comparing the LUTI model to the SETI model has demonstrated that the coefficients of certain land use and socio-demographic/economic variables may change in importance. For example, variable *pblu_r* (the proportion of blue collar employed residents) becomes statistically insignificant once the effective density variable has been added to construct the SETI model.

Each of the predictor variables, comprising the socio-demographic/economic, transportation, land use, and agglomeration variable types, are discussed in the separate sub-section below. The magnitude of the multiplicative effect is illustrated in plots of multiplicative effects. In these plots, the x axis represents increments of one standard deviation of the predictor variable (for example from *1 std dev* to *2 std dev*), and the y axis represents the magnitude of the multiplicative effect corresponding to that increase. Where $y = 1$ the coefficient of the predictor variable is not significantly different to zero (and therefore indicates no influence). When the predictor variable has a statistically significant positive influence on the proportion of train ridership, this influence is described by all plots higher than 1. On the other hand, all values plotting below 1 indicate a negative influence of the predictor variable on the proportion of train ridership. The variables associated with each of the LUTI and the SETI models are marked with different line styles, such as dotted lines to represent all of the LUTI variables and short lines to represent all of the SETI variables.

6.3.1 Socio-demographic/economic variables

Socio-demographic/economic variables are part of the vertical segregation that identifies persons with the same characteristics into the same category, for example, based on their income level or type of occupation. In the train trip production model (table 6.1), the proportion of employed residents in manufacturing (*p_worker_man)* had a statistically significant positive influence on train ridership in both the LUTI and SETI models. The remaining variables which had a negative contribution on train ridership were (in order of magnitude from the strongest one): the employed resident level of income for professionals/managers (*inc_pr_r*), car ownership level per dwelling (*car_own*) and the proportion of blue collar employed residents (*p_blu_r*).

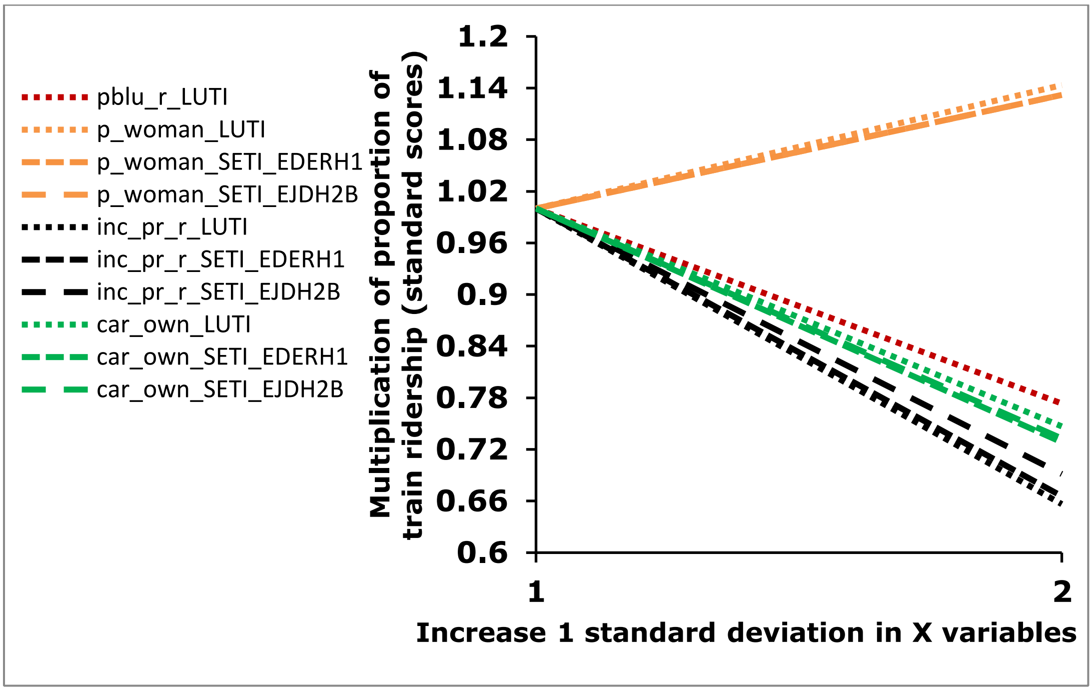


Figure 6.3 Plots of the multiplication effects based on the standardized value of predictor variables in socio-demographic/economic component on the proportion of train trip production (illustration).

Table 6.1 indicates that the multiplicative effects based on the SETI h1 model are the best model for explaining train trip production. Further, Figure 6.1 demonstrates that, based on the SETI h1 model, increasing one standard deviation of the proportion of employed residents in manufacturing (std dev=0.05) multiplies the proportion of train ridership production by 1.132. This equals to the increase of 13.2% in *y* variable. Increasing one standard deviation of income of employed residents with professionals/managers occupations (std dev= $7.45 per hour) and of car ownership (std dev=0.4) would reduce train ridership with multiplicative effects of 0.666 and 0.728 respectively, *cateris paribus*. These equal to the decrease of 33.4% and 27.2% repectively in the *y* variable.

Comparing between models, increasing one standard deviation of the proportion of employed residents in manufacturing would multiply the proportion of train ridership production by 1.1434 (an increase of 14.34%) for the LUTI, and 1.132 (an increase of 13.2%) for SETI h1 and SETI h2b, holding other variables constant. Increasing one standard deviation of income of employed residents with professionals/managers occupations would reduce the proportion of train ridership production, by factors of 0.793 for LUTI (a decrease of 20.7%), 0.6656 for SETI h1 (a decrease of 33%) and 0.69 for SETI h2b (a decrease of 31%), all else being equal. Similarly, increasing one standard deviation of car ownership would reduce train ridership by the factors of 0.747 for LUTI (a decrease of 25.3%), 0.728 for SETI h1 (a decrease of 27.2%) and 0.7327 for SETI h2b (a decrease of 26.7%), all else being equal.

While income, car ownership, and the proportion of manufacturing employed residents have positive or negative relationships as expected for their influence on train trip production, an unexpected observation is that of the negative influence of the proportion of blue collar employed residents. Blue collar occupations consist of machinery operators and drivers, technicians and trade workers, and labourers.

Map 6.4 shows that the distribution of jobs for blue collar occupations is mostly scattered throughout the study area and not concentrated in the Perth CBD area. Rather, the locations of jobs in this occupation type extend into the outer suburban areas where car ownership is often high (refer to map 6.8 of car ownership). This may suggest that a high number of blue collar employed residents live in suburbs underserved by railway services and would therefore be more likely to travel by cars. As a result, working trips by car would tend to be higher, resulting in the negative relationship between the proportion of blue collar jobs and the proportion of train trip production. By observing maps 6.4 and 6.5, the distribution of a high proportion of employed residents and jobs in blue collar occupations seems to correlate to each other (they have similar distribution patterns). This may suggest that blue collar workers tend to cluster around their place of work (Lowe, 1998 and Simmon and Koppelman, 2001 as cited in Wang & Chai, 2009). This may relate to the income level of blue collar workers, usually lower than white collar workers, thus encouraging them to avoid longer commuting times in order to avoid higher travel costs (Wang & Chai, 2009).

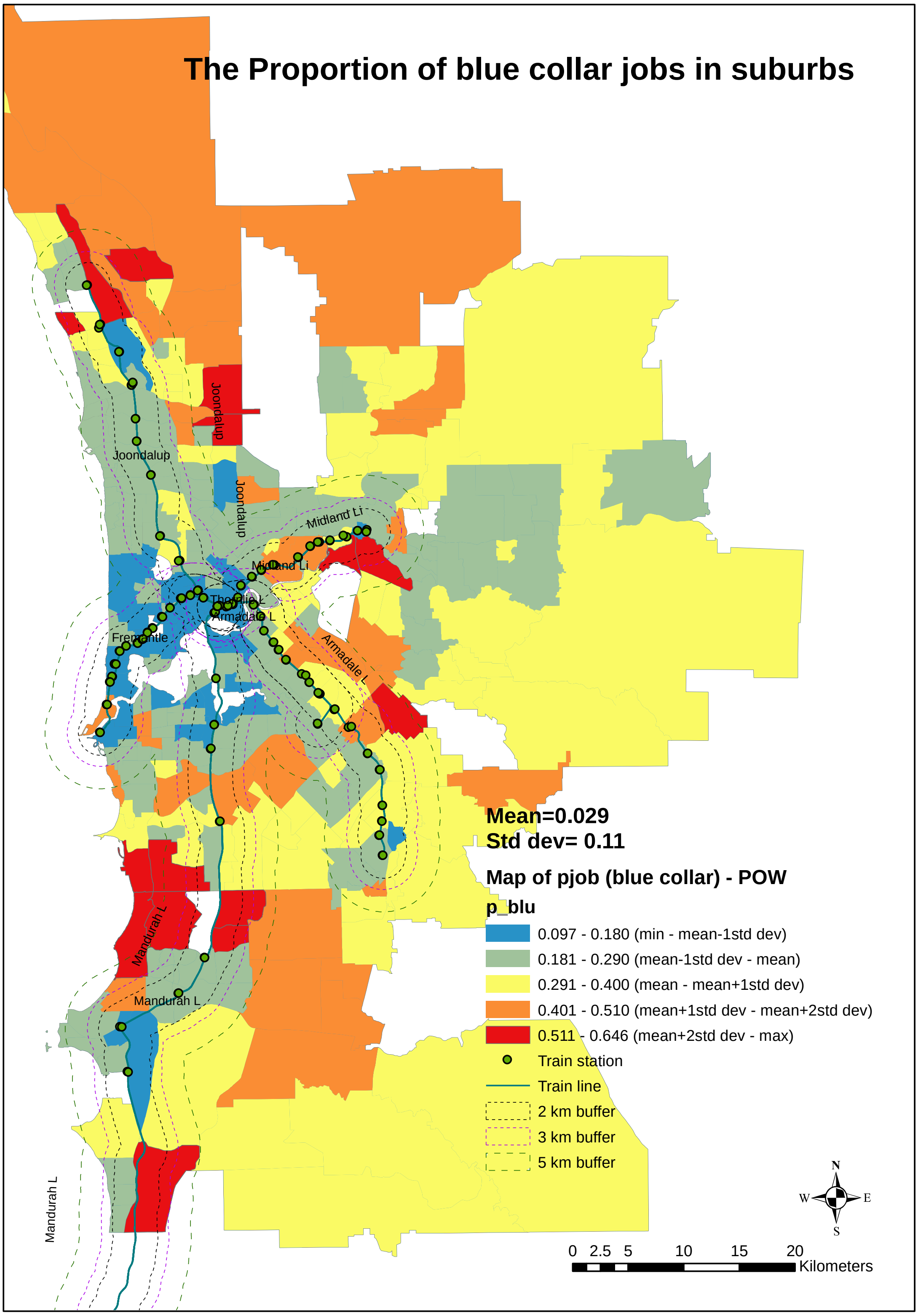


Figure 6.4 Map of the proportion of blue collar jobs in suburbs

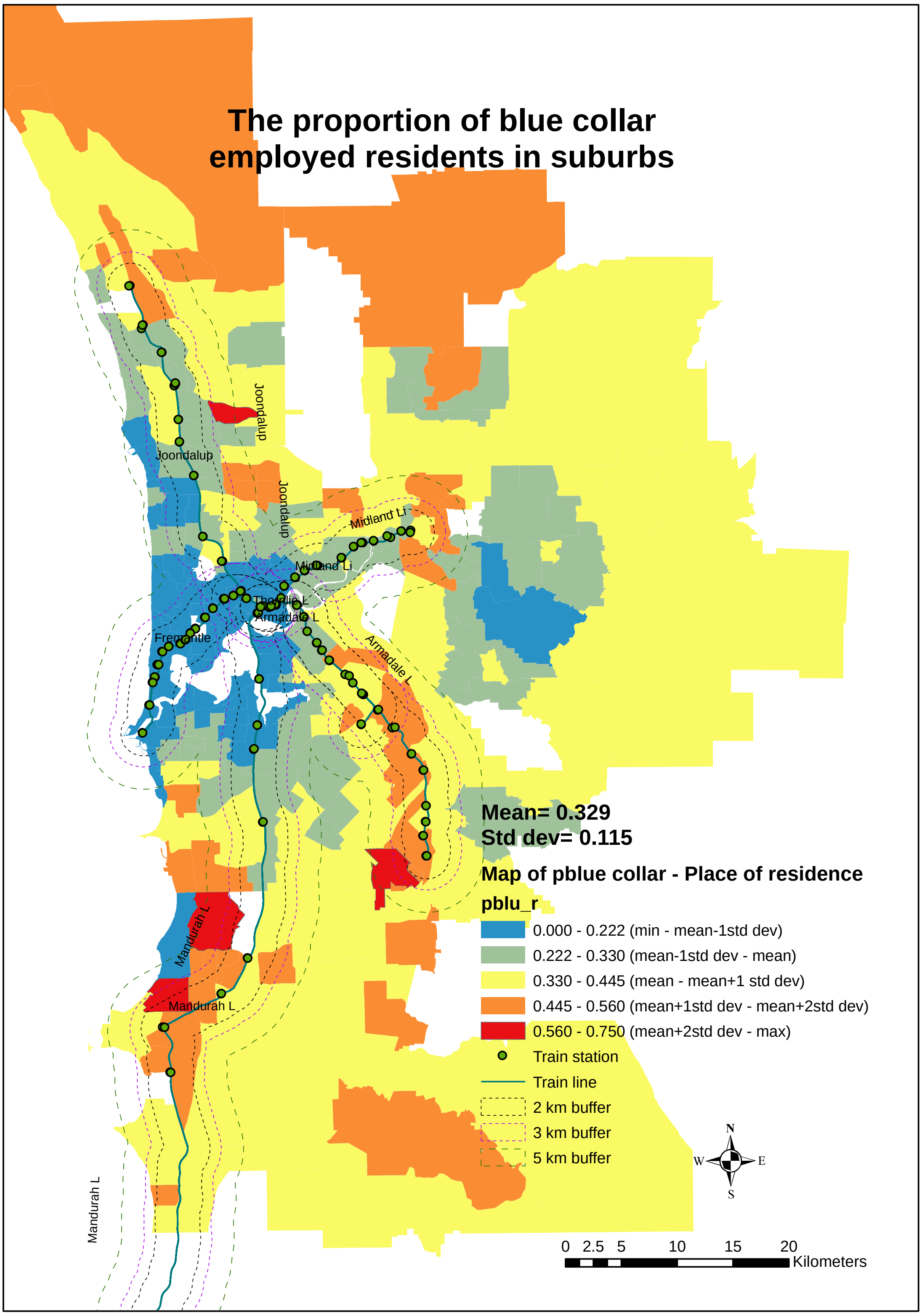


Figure 6.5 Map of the proportion of blue collar employed residents in suburbs

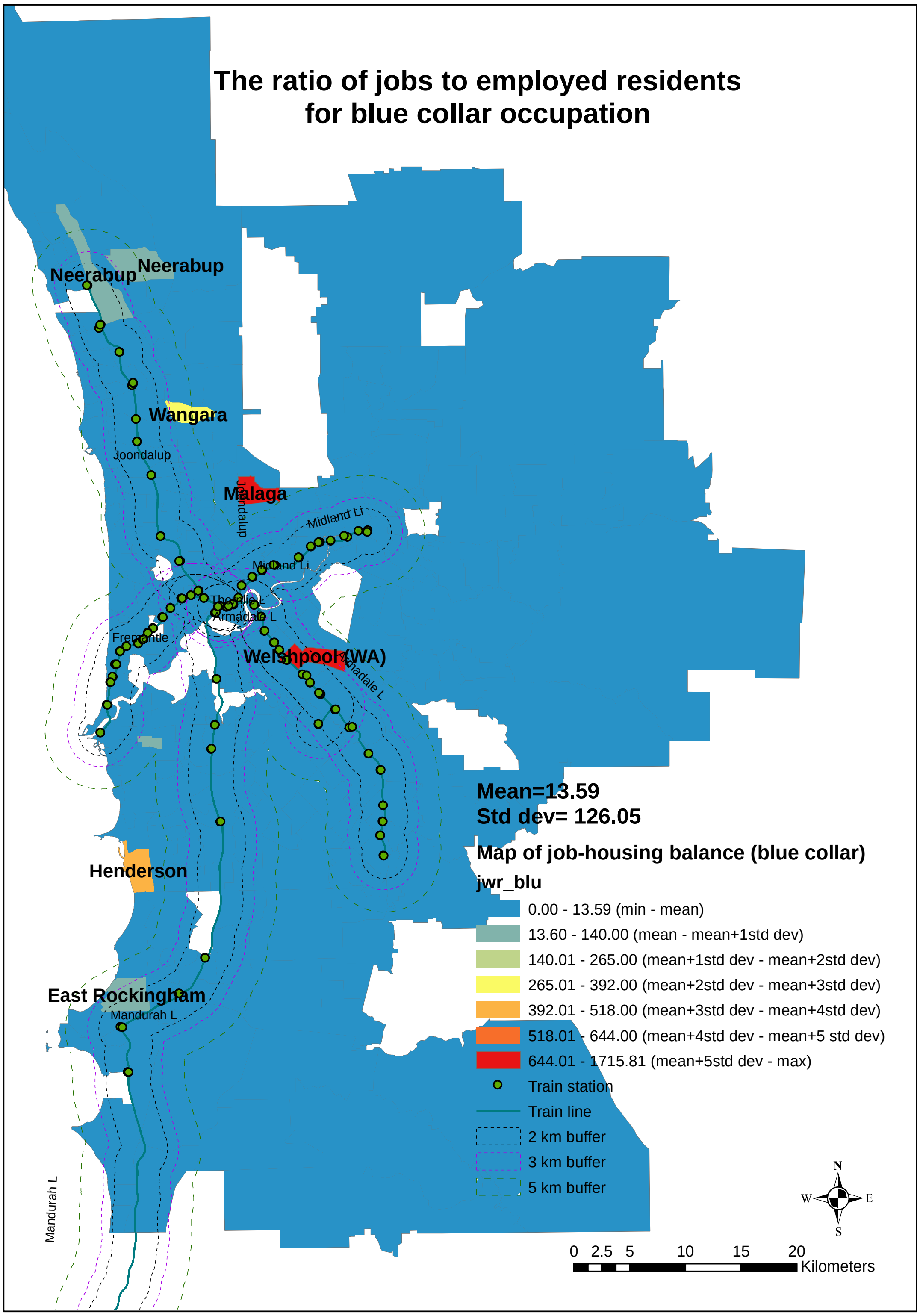


Figure 6.6 Map of the ratio of jobs to employed residents for blue collar occupation based on the standard deviation

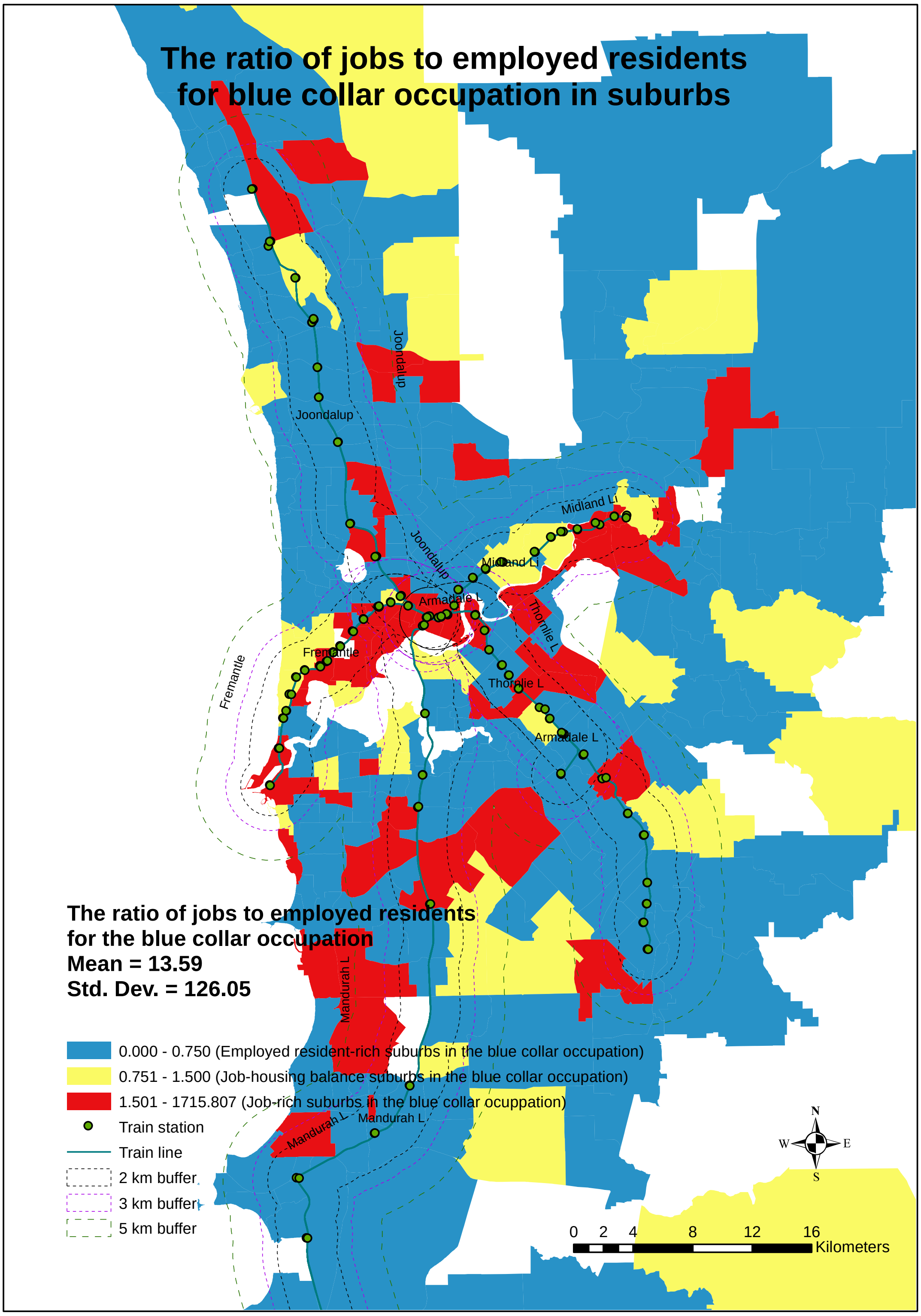


Figure 6.7 Map of the ratio of jobs to employed residents for blue collar occupation based on job-housing balance criteria

Map 6.6 and 6.7 illustrates the job-housing balance ratio of jobs to employed residents in blue collar occupations. Hypothesis 4 assumes that employed residents tend to work near their residential locations when this ratio is in balance. The yellow areas on map 6.7 refer to those areas with a high potential for self-containment, thereby lowering the overall number of trips and the number of trips produced by train. These may be observed to spread out along the Fremantle line, Midland line, and for some suburbs along the Joondalup and Armadale lines. Other yellow areas may be observed beyond the railway services.

The scattered spatial distribution of blue collar jobs (map 6.4) is likely to make railway trips difficult to those locations. Rail services cannot accommodate a dispersed pattern of development (Hall 1969). Some of these areas are apparently beyond the current railway line network services. There have been multiple plans for distributed light rail networks, but focus on the Perth city area. Thus, the negative influence of blue collar employed residents on train trip production may refer to two possibilities: (1). Employed residents work near their residences as suggested by Maps 6.4 and 6.5 have similar distribution patterns, and map 6.7 shows that areas of job-housing balance are distributed over the same areas where jobs are high (map 6.4) and employed residents are high (map 6.5). Employed residents clustered around their job locations tend to make fewer total trips or travel shorter distances; and short distance trips are less likely to be served by railways. (2). High proportions of blue collar employed residents are found in the outer suburbs that are underserved by railway services (map 6.5), and at the same time, these areas often have a high dependence on cars, as shown by the high levels of car ownership (map 6.8). The distribution of blue collar jobs spreads into some outer suburbs, and therefore the railway network has difficulty in serving and connecting these scattered job locations to their employed residents. As a result, an increase in the proportion of blue collar employed resident may lower the proportion of train trip production, cateris paribus.

In related to income and car ownership variable in the train trip production model, it has been widely known in the literatures that income level and car ownership are related with negative influence on train ridership. Table 6.1 showed based on the multiplicative effects of car ownership ranged between 0.7 to 0.8 in all models for every one unit increase in car ownership's standard score, this thesis found those who possess car with one standard deviation (equals to 1.5 cars per dwelling) use train for around 0.7 to 0.8 lower than those who not possess car. Put another way, those without a car make around 1.2 to 1.4[9] times more as many working trips by train to those with 1.5 cars per dwelling, independent of all other factors. The spatial distribution of household with high car ownership level are spread out on areas with low accessibility to or underserved by railway services (map 6.8). It seems households with car dominances are not concern to live in within the proximity to railway line services may implicitly reflect the self-selection effects (Bhat, Sen, & Eluru, 2009; Guo & Bhat, 2007; Pinjari et al., 2007).

The variable of income for managers and professionals contributed negatively to train ridership in the train ridership production model. Map 6.9 shows that high income earners in professional/manager occupations clustered around the northwest of the Perth metropolitan region, around the Perth CBD area, and around areas in the northeast of the Perth metropolitan region. This may be related to high wage jobs that are offered in the southern parts (around Henderson and its surrounding area) and northern parts of the metropolitan area (along the Fremantle coastal line and Neerabup) and around the CBD. This may suggest their mobility would be supported by intensive use of the Kwinana freeway running, and thus discourage the use of the train. It has been well known (Timothy & Wheaton, 2001) that high income employees will spend longer on commuting time and working in other areas , despite the fact the North-South freeway network may facilitate more trips for high income earners than the railway. On the other hand, high income earners residing around the Perth CBD area may not be users of trains as they may live and work in the city where there are free bus services within the CBD. The study suggested that those who are living in group households in the inner city area are likely to have jobs in the inner city as well, with

[9] This is calculated as: 1/0.8 = 1. 25 and 1/0.7 = 1.43

minimum travel requirements (AHURI, 2006). It is not likely train would be used as mode of travel to work.

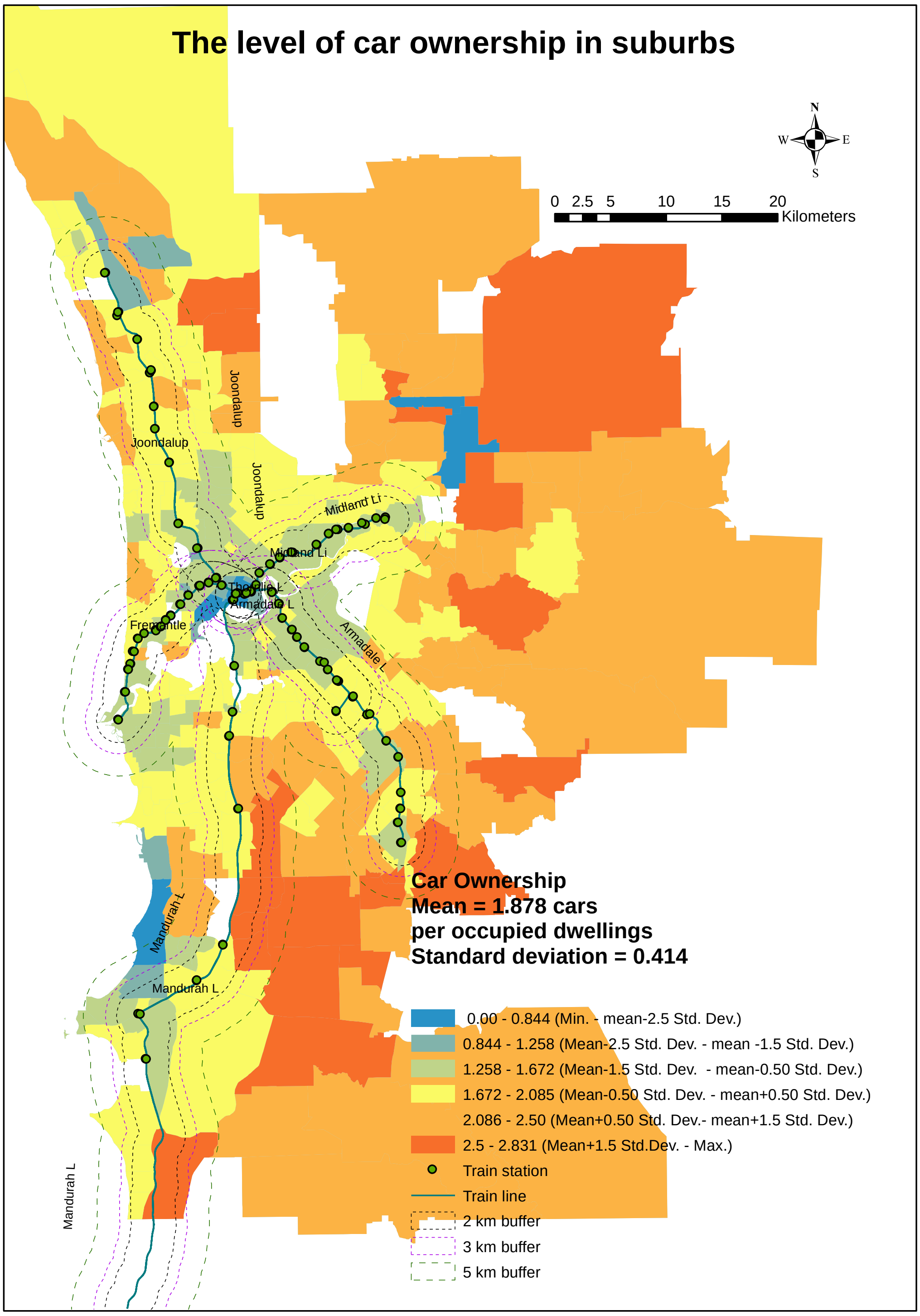


Figure 6.8 Map of the level of car ownership per occupied dwelling in suburbs

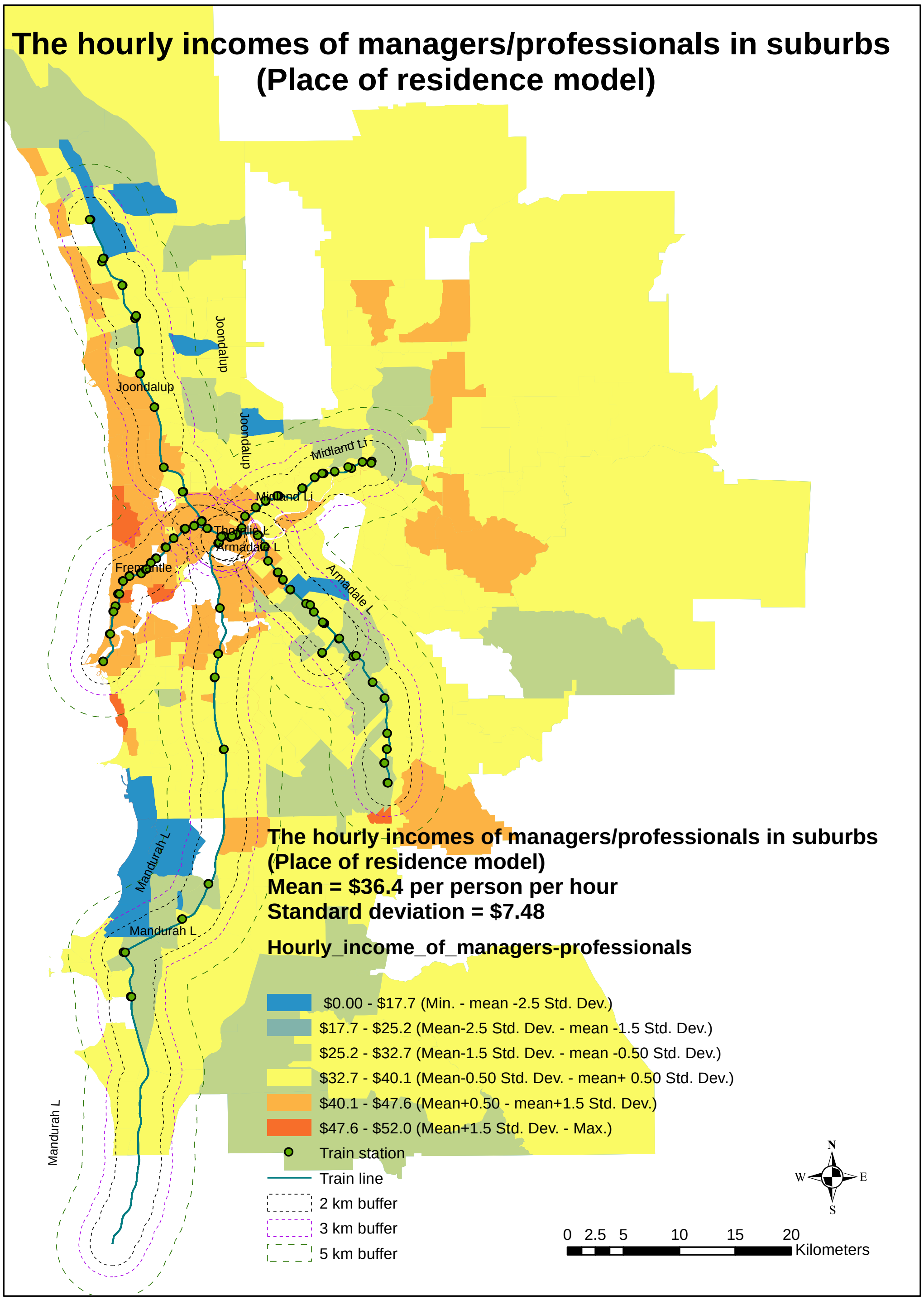


Figure 6.9 Map of income level of employed resident in managers/professionals occupation at residential suburbs

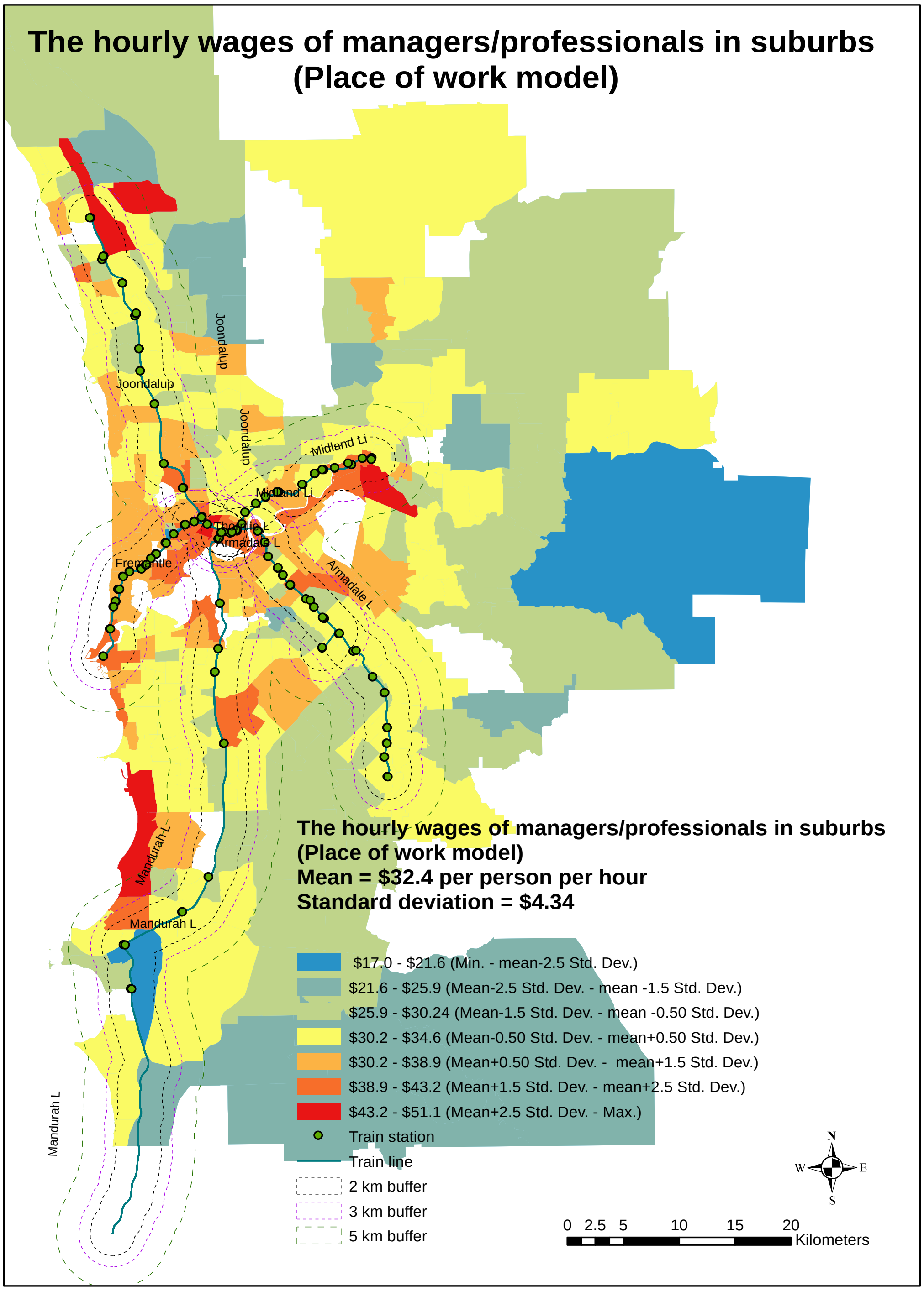


Figure 6.10 Map of the hourly wages of managers/professionals occupation in place of work suburbs

In the train trip attraction model, the results indicate only a few variables to be consistently statistically significant in all models, such as the proportion of jobs in the retail sector (*p_jobret*) and the number of all jobs (*all_job*). These variables have positive influences on train trip attraction.

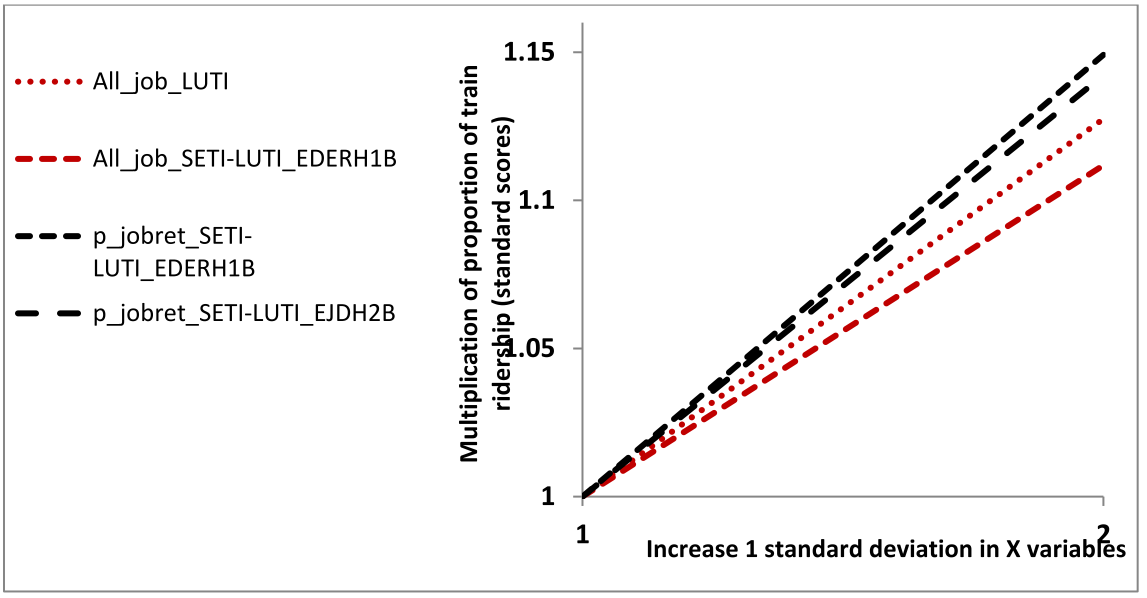


Figure 6.11 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in the socio-demographic/economic component on the proportion of train ridership (train trip attraction model)

Table 6.2 indicates that the multiplicative effects based on the SETI h1b model are the best model for explaining train trip attraction. In particular, table 6.2 shows that, based on the SETI h1b, increasing one standard deviation of the proportion of jobs in retail (std dev= 0.082) would multiply the expected value of the proportion of train trip attraction by 1.149 (an increase of 14.9%). Increasing one standard deviation of the number of all jobs by one standard deviation (std dev= 5,406) produces a lower multiplicative effect of 1.112 (an increase of 11.2%) than the effect from retail jobs.

Comparing between models, increasing one standard deviation of all jobs would multiply the proportion of train ridership by 1.127 for LUTI (an increase of 12.7%) and 1.112 for SETI h1b (an increase of 11.2%), holding other variables constant. The proportion of jobs in the retail sector has similar multiplicative effects in all models. Increasing one standard deviation of the proportion of jobs in the retail sector (st dev=0.08) would multiply the proportion of trip attraction by 1.149 for the SETI h1b model (an increase of 14.9%), and 1.142 for SETI h2b (an increase of 14.2%), all else being equal.

Map in figure 6.12 shows that the main workplace suburbs or job centres in the Perth metropolitan region are located in the Perth CBD, Malaga, Canning Vale, Welshpool, Fremantle, and Osborne Park. The Perth CBD alone accounts for 12.78% of the total jobs in the Perth metropolitan region, while the share of total jobs in other workplaces as mentioned above range from 2% to 3.5%.

This high concentration of jobs in the Perth CBD may contribute to the high train trip attraction to the Perth CBD. A study confirmed that the number of transit commutes was closely related to the number of jobs in the CBD rather than the overall metropolitan area size or total metropolitan workers (Hendrickson, 1986). Therefore, the CBD area often has a concentration of transit hubs to serve the highest job densities (Chen et al., 2008). Map 6.13 confirms the dominance of railway line service coverage in the Perth CBD areas. The radial form of five railway lines intersecting in the Perth CBD is an advantage for the high intensity of CBD jobs. That is, transit commuting from all directions is possible by train. Other research confirmed job numbers and density in the CBD areas were important for the consistent increase in transit share, independent of other factors (Rickwood & Glazebrook, 2009).

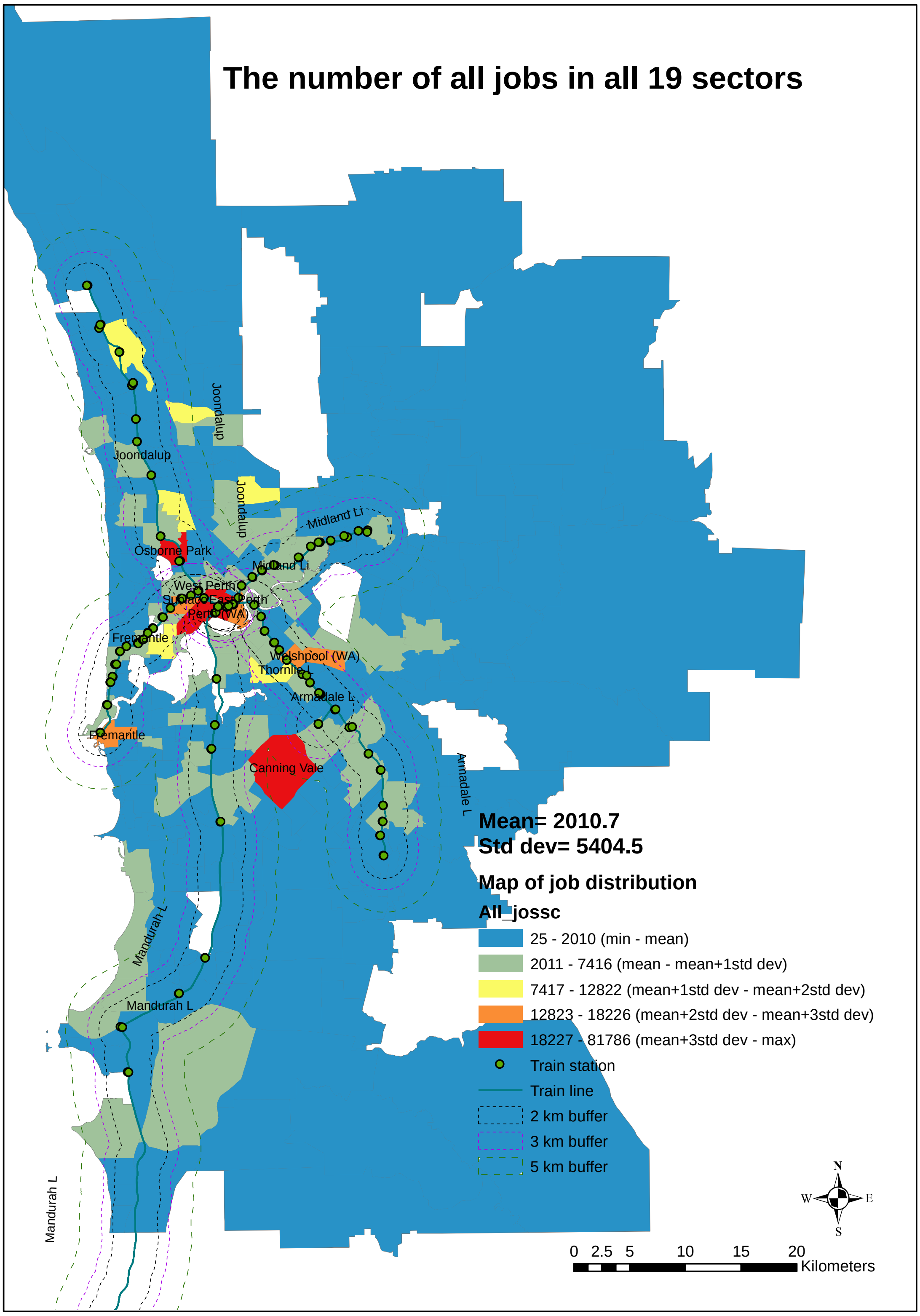


Figure 6.12 Map of the number of all jobs in all 19 sectors

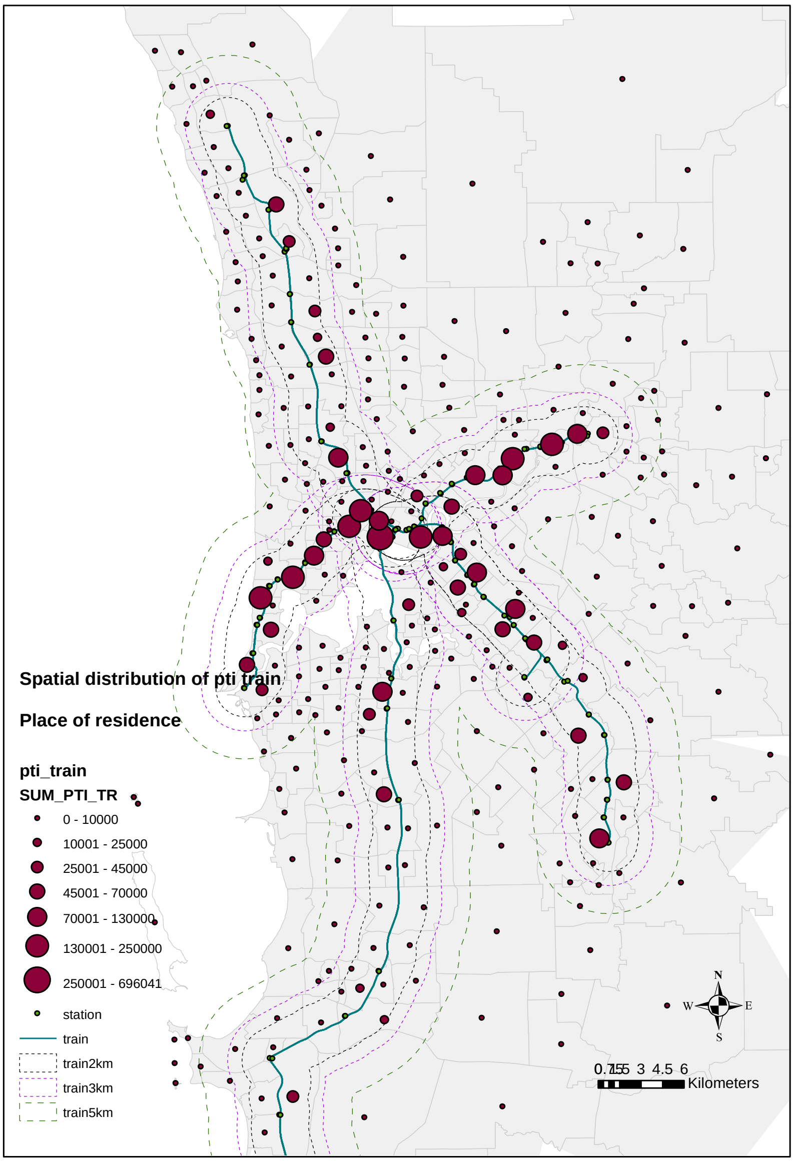


Figure 6.13 Map of the index of public transportation supply in suburbs

In the retail sector, both employed residents (at 90% level of confidence) and retail jobs (at 95% level of confidence) have been shown to contribute positively to higher train trip attraction. More retail jobs are concentrated near train stations (map 6.13), and the positive contribution of the proportion of retail jobs on train ridership may be facilitated by the distribution of these jobs around many train stations.

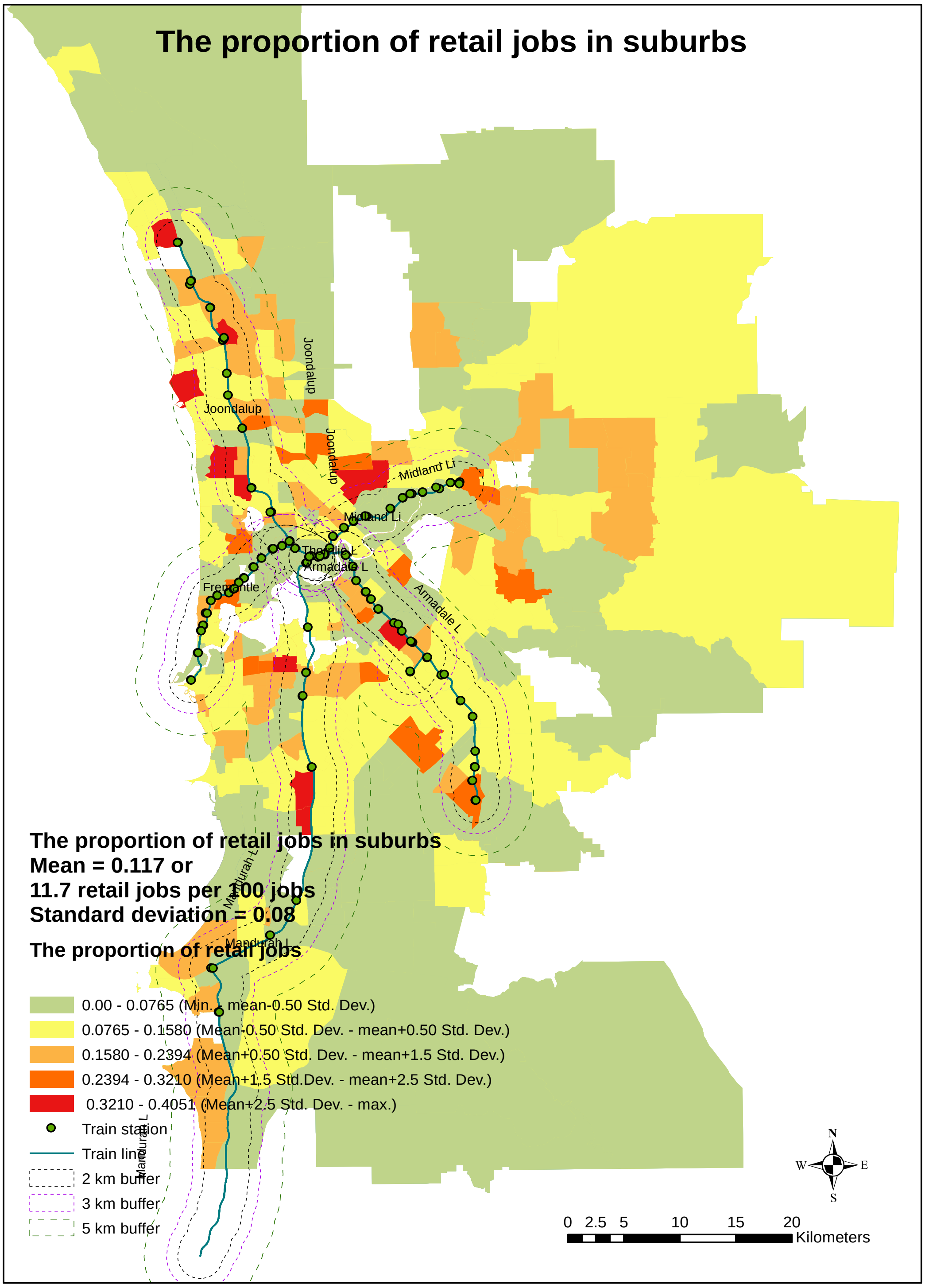


Figure 6.14 Map of jobs in the retail sector

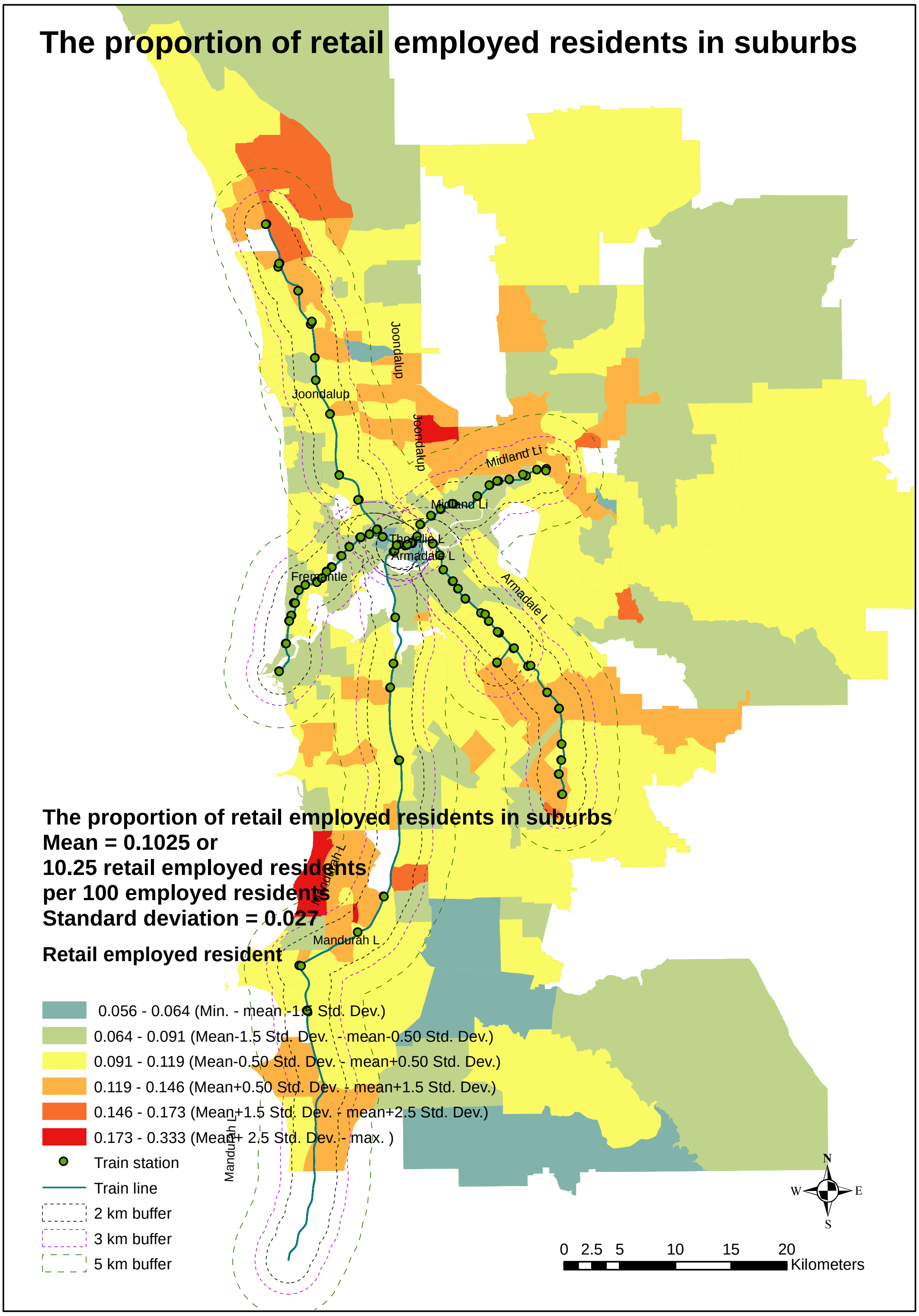


Figure 6.15 Map of spatial distribution of employed resident working in retail

The SETI h1b model (table 6.2) has shown that the influence of jobs in the retail sector on train trip attraction is higher than the influence of total jobs. In addition, as shown on map 6.14, the proportion of jobs in the retail sector in the Perth CBD area is below the mean of the Perth metropolitan area. A lower retail job proportion in the Perth CBD indicates that many of the working trips to reach retail jobs were not destined for the Perth CBD. Similarly, map 6.15 illustrates the distribution of employed residents in the retail sector, which is relatively low in the Perth CBD area. Interestingly, a high proportion of employed residents in retail live in areas in proximity to park and ride services (within 5 km of a train station) as shown by the orange and red areas inside the buffer lines of map 6.15. Map 6.14 confirms that the red and orange areas representing high proportions of retail jobs are located in relatively close proximity to train stations. A study found that abundant jobs within four miles of home reduce travel distance for work trips significantly (Cervero and Duncan, 2006). Thus, the housing-job proximity for the retail sector is facilitated by railway station proximity to both residential and retail workplace areas. It is possible that employed residents in the retail sector would cluster around retail job locations, allowing for a short distance travel by train using the same railway platform.

Other socio-demographic/economic variables, such as the proportion of manager and professional job occupations, the proportion of blue collar jobs, the proportion of jobs in the manufacturing and construction sectors, and wages per hour were not statistically significant in all models of train trip attraction (place of work model).

6.3.2 Transportation or accessibility variables

Transportation variables may be viewed as representing the level of accessibility of a suburb:

(i) Natural log of average distance between suburb and train station (*ln_avedist*) represents the accessibility of suburb to a train station. It may be seen for the accessibility of suburb by station access.

(ii) Natural log of road network travel distance (*ln_strio)* represents the accessibility or the centrality level of a suburb compared to all other suburbs, with respect to the road network. A higher travel distance by road network (a high value of *ln_strio*) represents lower level of accessibility, or low centrality. It may be seen as a proxy for accessibility by road access.

(iii)Square root of public transport supply (*sqrt_ptiori*) represents the level coverage of public transport in a suburb, in particular as measured by bus stops and train stations coverage and service frequencies. It may be seen as a proxy for accessibility of a suburb via public transport network.

In the train trip production model, increasing a suburb's centrality by road network, increasing the average distance of the suburb from a train station and increasing a suburb's public transport supply coverage, were all associated with a reduction in train trip production. This may indicate the possibility of substitution or complementary between train and bus, and between train and car. This would be mean subsitution given that increasing a suburb's accessibility based on these parameters would discourage the use of train, and complementary if vice versa.

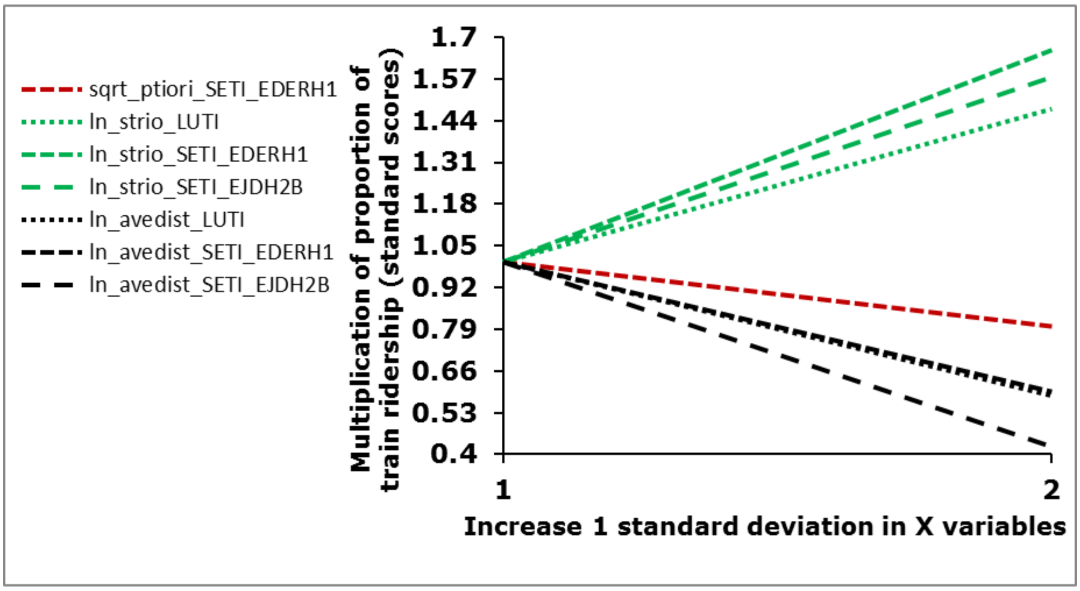


Figure 6.16 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in the transportation/accessibility component on the proportion of train ridership (train trip production model)

Table 6.1 shows the multiplicative effects on train trip production. In the SETI h1 model, increasing one standard deviation of *sqrt_ptiori* (std dev= 581 units) multiplies the proportion of train trip production by 0.7993 (a decrease of 19.01%). Increasing one standard deviation of *ln_strio* would increase train production, or multiply it by 1.66 (an increase of 66%).[10] Similarly, increasing one standard deviation of *ln_avedist*[11] would reduce train production or multiply it by 0.5957[12] (a decrease of 40%). The farther the distance of a residential suburb from a train station, the lower the train trip production by a factor of 0.5957.

Comparing between models, a relative increase in *strio* by 0.23% multiplies the proportion of train trip production by 1.477 in the LUTI model (an increase of 47.7%), by 1.66 in SETI h1 (an increase of 66%) and by 1.577 in SETI h2b (an increase of 58%). This means that when the location of a residential area makes it difficult to reach other suburbs by the road network, the centrality of that location is low. This variable may reveal the possibility of a modal switch between cars and trains when travel distance by car is increased (for example due to worsening of road congestion), and train ridership increases. Increasing *avedist* by 0.69% multiplies the proportion of train trip production by 0.583 in LUTI (a decrease of 41.7%), 0.5957 in SETI h1(a decrease of 40.43%), and 0.424 in SETI h2b (a decrease of 57.6%).

Note that the inverse relationship between train trip production and the level of public transport supply in a suburb (*sqrt_ptiori)* is not as expected. As the level of public transport supply was measured based on bus networks and railway networks, there may be bias in the dominance of the bus network over the railway network for this measurement. This can be observed in map 6.17. Separating the overall index into individual indices of public transport supply, the correlation between the overall index and the index of bus supply is 0.99; while the correlation between the overall index

[10] Note that increasing *ln_strio* by 1% of its standard deviation equals to 0.0023 (st dev = 0.23 units) is equivalent to multiplying *strio* by exp (0.0023) or1.002303. After rounding, a relative increase in *strio* of 0.23% (i.e. reducing the centrality of a suburb relative to all other suburbs by road network by 0.23% or increasing total km network distance by 0.23%) multiplies the proportion of train trip production by 1.66.

[11] Standard deviation of *ln_avedist* is 0.687, that is equals to average distance of 2 km

[12] Increasing *ln_avedist* by 0.00687 (st dev=0.687) is equivalent to multiplying *avedist* (distance of a suburb to the train station) by exp (0.00687) or 1.0069. After rounding, a relative increase in *avedist* by 0.69% multiplies the proportion of train trip production by 0.5957 (a decrease of 40%).

and the index of railway supply is 0.28. Thus, the measurement of bus supply is much more dominant than the measurement of the train supply and, as such, the overall index represents the network of buses supply (see map 6.17 (left) and similarities to map 6.19). Therefore, the inverse relationship between the public transport supply and train trip production may reflect the competition between bus services and railway services. Increasing the number of bus stops and the service frequency of buses would be expected to reduce the proportion of train trip production.

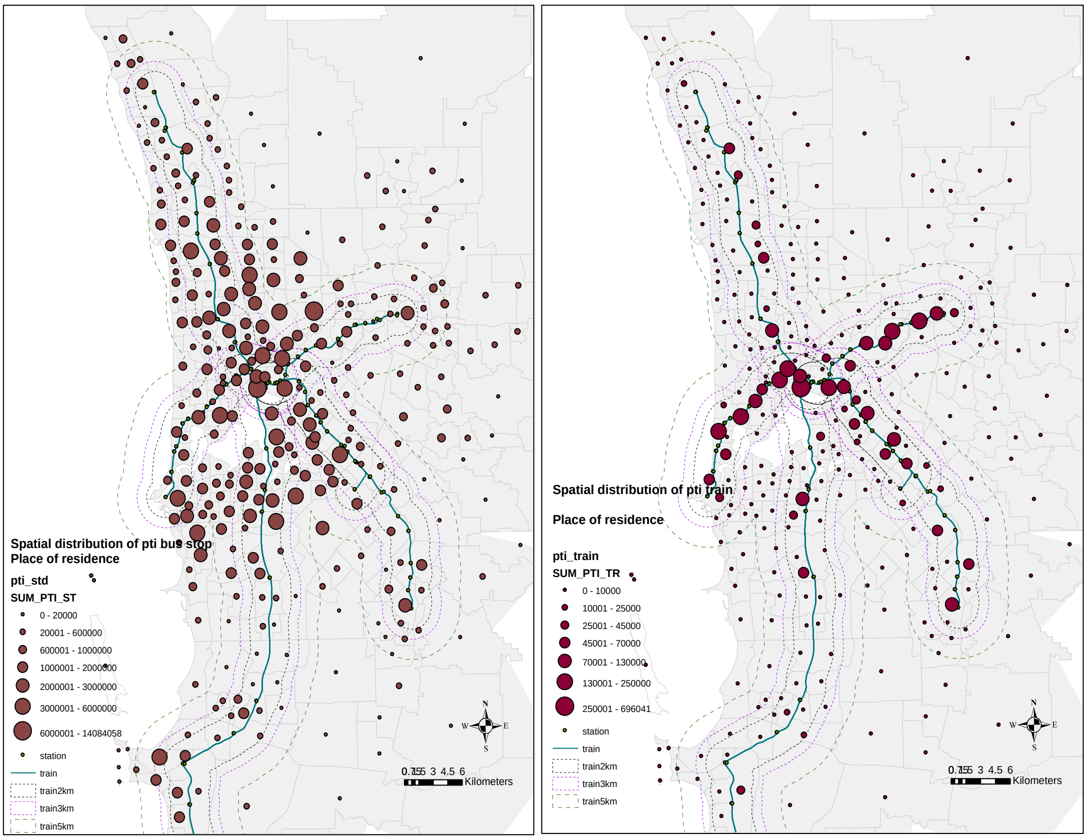


Figure 6.17 Map of the index of public transport supply measured by buses stops (left) and map of the index of public transport supply measured by train stations (right)

In the train trip attraction model, only two variables of transportation or accessibility were consistently statistically significant in the LUTI and the SETI model: the natural log of distance to the train station (*ln_avedist*) and natural log of the overall road network travel distance for the suburb (*ln_strio*). The magnitude of multiplicative effects from *ln_strio* is similar in all models. A relative increase in *strio* by 0.23% (i.e. reducing the centrality of the workplace suburb relative to all other suburbs by road network by 0.23%) multiplies the proportion of train trip attraction by 1.148 in the LUTI model (equals to an increase of 14.8%), and by 1.163 (equals to an increase of 16.3%) in SETI h1. Natural log of average distance to the train station was the strongest determinant of train ridership in an inverse sense, but was statistically significant only in the LUTI model. As distance of the workplace suburb from the train station increases, the proportion of ridership to reach the job location decreased. A relative increase in *avedist* of workplace suburbs by 0.69% multiplies the proportion of train trip attraction by 0.711 (equals to a reduction of 28.9%) in the LUTI model. Square root of public transport supply was not statistically significant in any model of train trip attraction.

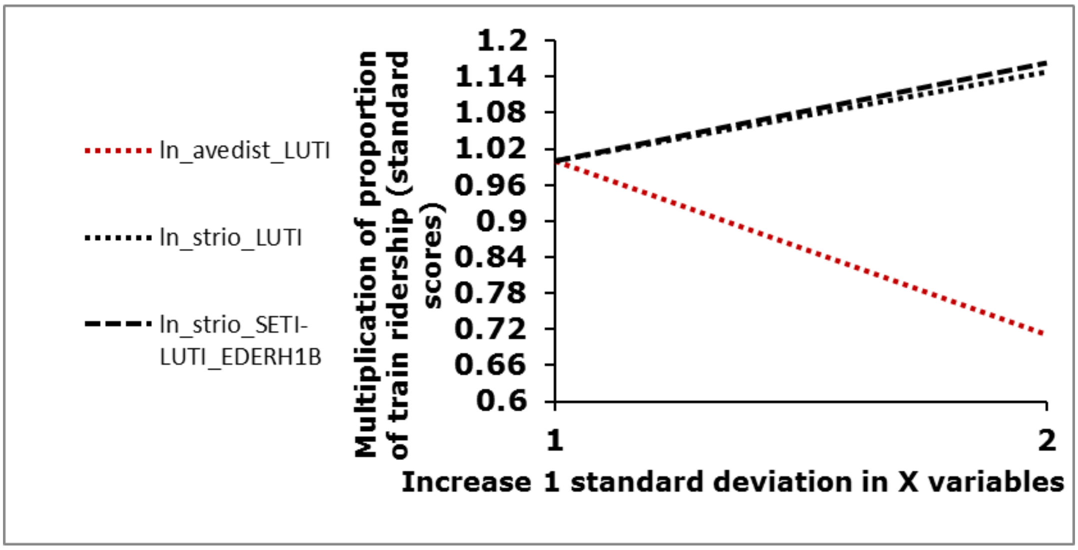


Figure 6.18 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in the transportation/accessibility component on the proportion of train ridership (train trip attraction model)

High values of road network travel distance relate to low suburb centrality and vice versa. Therefore, a positive coefficient indicates that increasing the road network travel distance (decreasing centrality by road network) is associated with an increase in train trip attraction. This has been clearly shown on map 6.20 which illustrates the level of road network travel distance. The Perth CBD and the inner suburbs of the Perth metropolitan area constitute the area with the highest centrality. Map 6.2 confirmed that the highest proportion of train trip attraction is toward the Perth CBD. However, increasing congestion in these areas would encourage more use of trains. The correlation is therefore a strong indication of competition between cars, or road-based modes, and trains. These effects are similar to the effects observe for train trip production. The decreasing of accessibility of a suburb by road access (such as via congestion) is likely to create higher levels of train ridership, both production and attraction.

Many studies have found that proximity to transport infrastructure will influence mode choices (Commin and Nollans, 2011). The outer suburb workplace areas, such as Malaga and Henderson (map 6.12), have been underserved by train stations and therefore may not attract train ridership. Likewise, when road network travel distance has been more influential positively (more travel distance means low centrality by cars) on train trip production than attraction, high density residential suburbs that are well-served by road network will rely more on cars. This also means that any changes to the system that reduces road network performance may influence the possibility of a mode switch from cars to train services.

On the other hand, train trip attraction is influenced by the competing transport infrastructure (or accessibility by road network) to a smaller extent than the influence from the accessibility of workplace suburbs to their destination train station. Thus, the accessibility of a workplace suburb to the train station has been more important than that of the residential suburb. This also suggests why the threshold distance from a train station for workplace suburbs was much smaller than that for residential suburbs (Chapter 7). Map 6.21 shows the level of accessibility of workplace suburbs to the destination station. Suburbs located along railway lines with shorter station spacing, such as along the Fremantle and Midland lines, have the highest accessibility level and may be expected to attract more ridership. Only a few stations along the Mandurah line are in proximity to employment centres, namely Kwinana, Cockburn Central, and Murdoch Stations, as well as a few places along the Joondalup line, such as those near the Neerabup employment centre, and along the Armadale line at Welshpool, Cannington and Armadale.

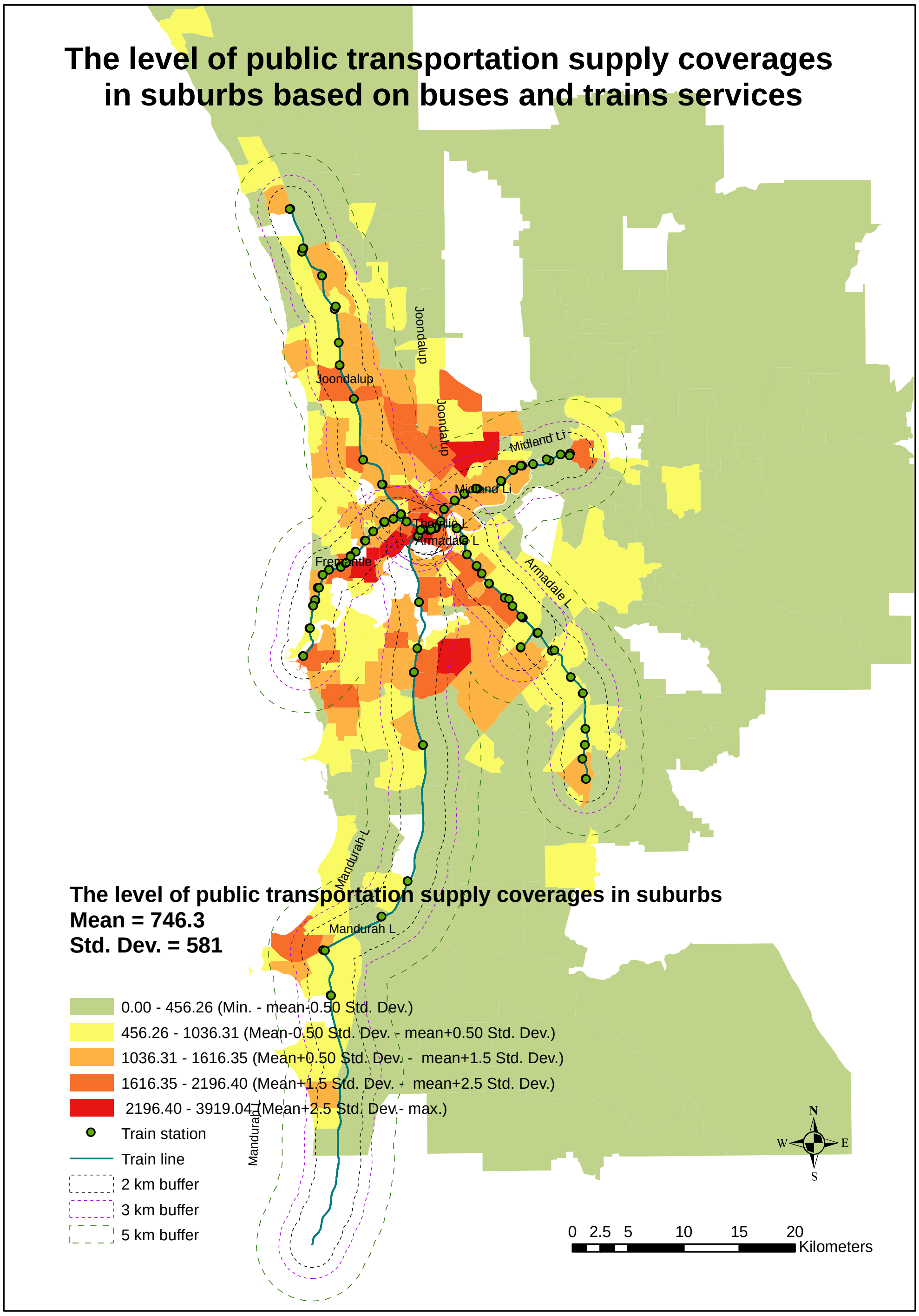


Figure 6.19 Map of public transport supply coverages index

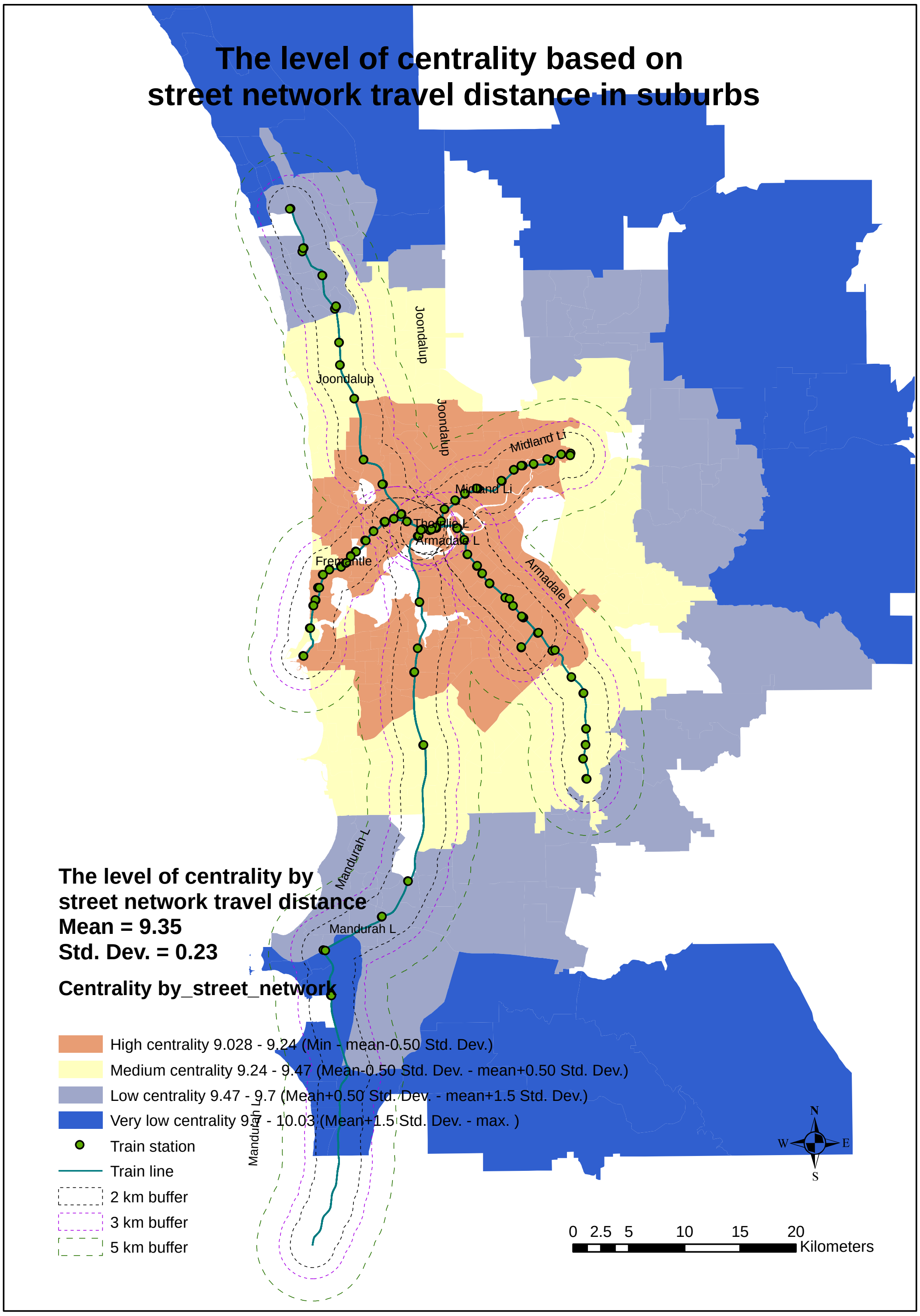


Figure 6.20 Map of road network travel distance

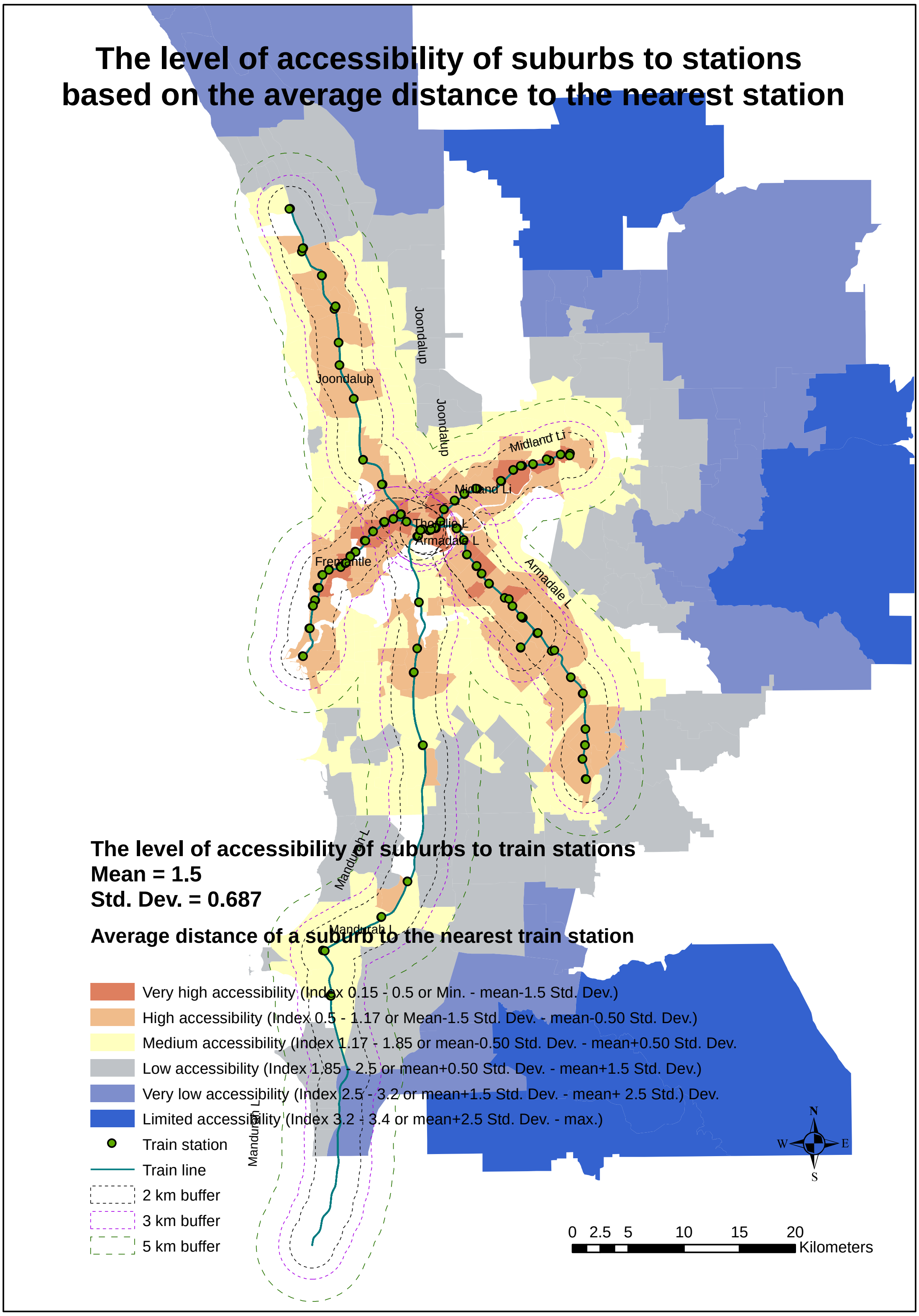


Figure 6.21 Map of the average suburb's distance to train station

6.3.3 Land use variables

In the train ridership production model, only two land use variables were found to be statistically significant: land value of residential land use (*ln_lvr*) and housing-job balance as defined by the natural log of job to employed resident ratio (*ln_jwr*). Employed resident density (*wod*) was only statistically significant at the 90% level of confidence, except in the SETI h2b model that is statistically significant at 95% level of confidence. While land value is negatively influence train trip production, job-housing balance has positive contribution to train ridership production, with land value having a substantially higher contribution than *ln_ jwr*.

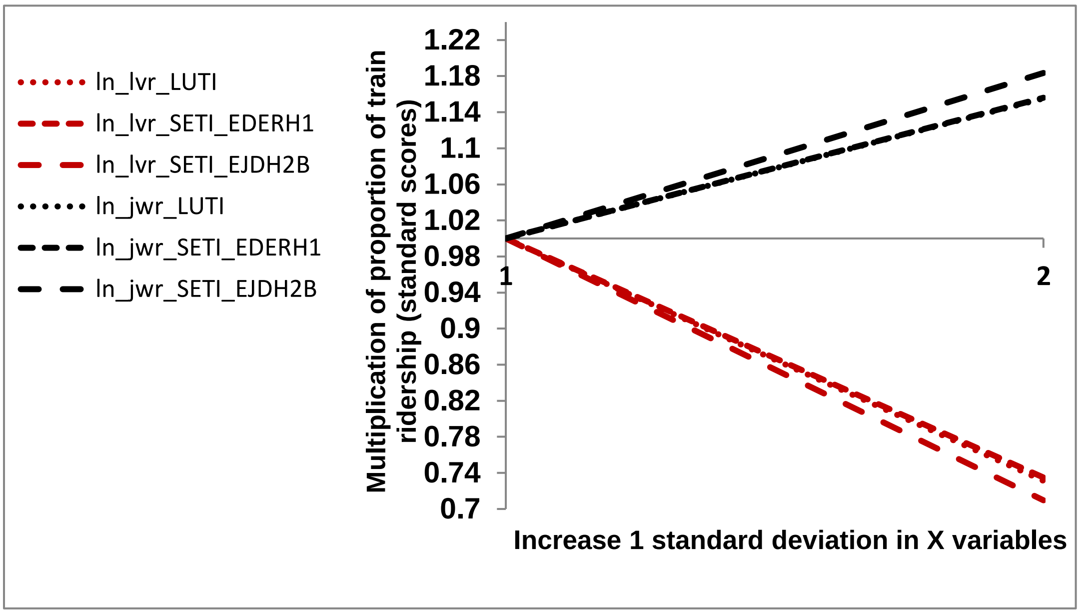


Figure 6.22 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in land use component on the proportion of train ridership (train trip production model)

Referring to table 6.1, it may be seen that increasing 1 percent of *ln_lvr* in its standard score (std dev= $1.037) by 0.01037 would multiply the proportion of train trip production by 0.73 (a decrease of 27%) in the LUTI model, by 0.735 (a decrease of 26.5%) in SETI h1 and by 0.71 (a decrease of 29%) in SETI h2b[13]. From

[13] Increasing 1 percent of *ln_lvr* in its standard score (std dev= $1.037) by 0.01037 is equivalent to multiplying *lvr* by exp (0.01037) by 1.010424. After rounding, a relative increase in the price of *lvr*

agglomeration theory, residential land rent is assumed to trade-off with transport costs. Theoretically, higher accessibility offers a land rent premium, which thus increases urban productivity due to higher agglomeration. If these assumptions are applied to the train trip production model, it may suggest that people would choose to live in an expensive area when they can avoid a high transport cost, such as by working in the same area as their home. In addition, the reduction in transport cost may also imply lower travel distances and thus discouraging the use of trains from residential suburbs. Instead, employed residents probably would use a more suitable mode for shorter travel distances such as by non-motorized transport or by bus. On the other hand, from the perspective of the LUTI framework, the negative influence of expensive land rents on train trip production is commonly viewed as a socio-economic impact, due to, for example, the common believe that rich people are not interested in using public transport.

The housing-job balance parameter, as measured by *ln_jwr* had a positive linear relationship with train trip production, when it might have been expected to give an inverse contribution. It may be suggested that job-rich suburbs still produce some degree of train trip production. From the SETI perspective, this means that suburbs with higher productivity are likely to attract more people and activity. Thus, although the proportion of jobs is higher than that of residences, increasing jobs in these suburbs may facilitate higher train ridership production as well.

While the negative influence of residential land values on train trip production has been as expected, the positive influence of the ratio of job to employed residents produced an opposite result to expectation. The negative influence of residential land value on train ridership production is plausible given higher land value have mostly emerged in the Perth CBD area and its inner suburbs in surrounding the CBD. High self-sufficient is likely to emerge in the CBD and inner suburbs areas, where employed residents would probably work in the same locations or the surrounding areas, as they

(land value of residential land use) of 1% multiplied the proportion of train trip production by 0.73 in the LUTI model, by 0.735 in SETI h1 and by 0.71 in SETI h2b trade-off paying for high land costs with lower transport costs. Thus, there is no incentive for increasing train trip production if the self-sufficiency of a suburb is high.

On the other hand, as seen in the place of work model, job density has been the only land use variable that contributes to train trip attraction for workplace suburbs. Map 6.24 shows that the higher values of the natural log of job density may be found to emerge around the railway line and closer to stations, especially those in Fremantle, Midland, and some parts of the Armadale and Joondalup lines.

The Transit Cooperative Research Program (TCRP) study found that an elasticity of commercial floor space on ridership was around 0.5, which is, a doubling of floor space (building activity) was associated with a 50% increase in the ridership of Metrorail (Cervero et al., 2004, p. 154). In terms of the multiplicative effects, this value was therefore equivalent to a factor of 1.5. In this research, the model found comparable results to that of the TCRP study, where the multiplicative effect of ln job density on train trip attraction was determined to be 1.339 in the LUTI model (an increase of 33.9%), 1.451 in the SETI h1b model (an increase of 45%) and 1.385 in the SETI h2b model (an increase of 38.5%).

In place of work model or train trip attraction (see table 6.2), the square root of land value for non-residential land use, or *sr_lvnr*, was found to be only statistically significant in the LUTI model at the 90% level of confidence, and was not statistically significant in both of the SETI models after adding the effective density variable. This may reflect the strong correlation between *sr_lvnr*, which is an indirect proxy or partial measurement of urban productivity, and the effective density. The correlation between *sr_lvnr* and effective job density was 0.55, while the correlation between *sr_lvnr* and effective employed resident density was 0.64. The addition of the agglomeration variable to the LUTI model has had the effect of rendering some of the land use variables statistically insignificant in the SETI model. Job-housing balance, i.e. the ratio of job to employed resident (*ln_jwr*), was statistically insignificant in all models of train trip attraction. However, natural log of job density was remained statistically significant and even stronger after adding effective density into the model. The influence of the job density variable outweighed the influence of transportation variables and all socio-economic/demographic. Furthermore, the influence of job density was stronger when modelled in SETI h1b and h2b than when modelled in LUTI.

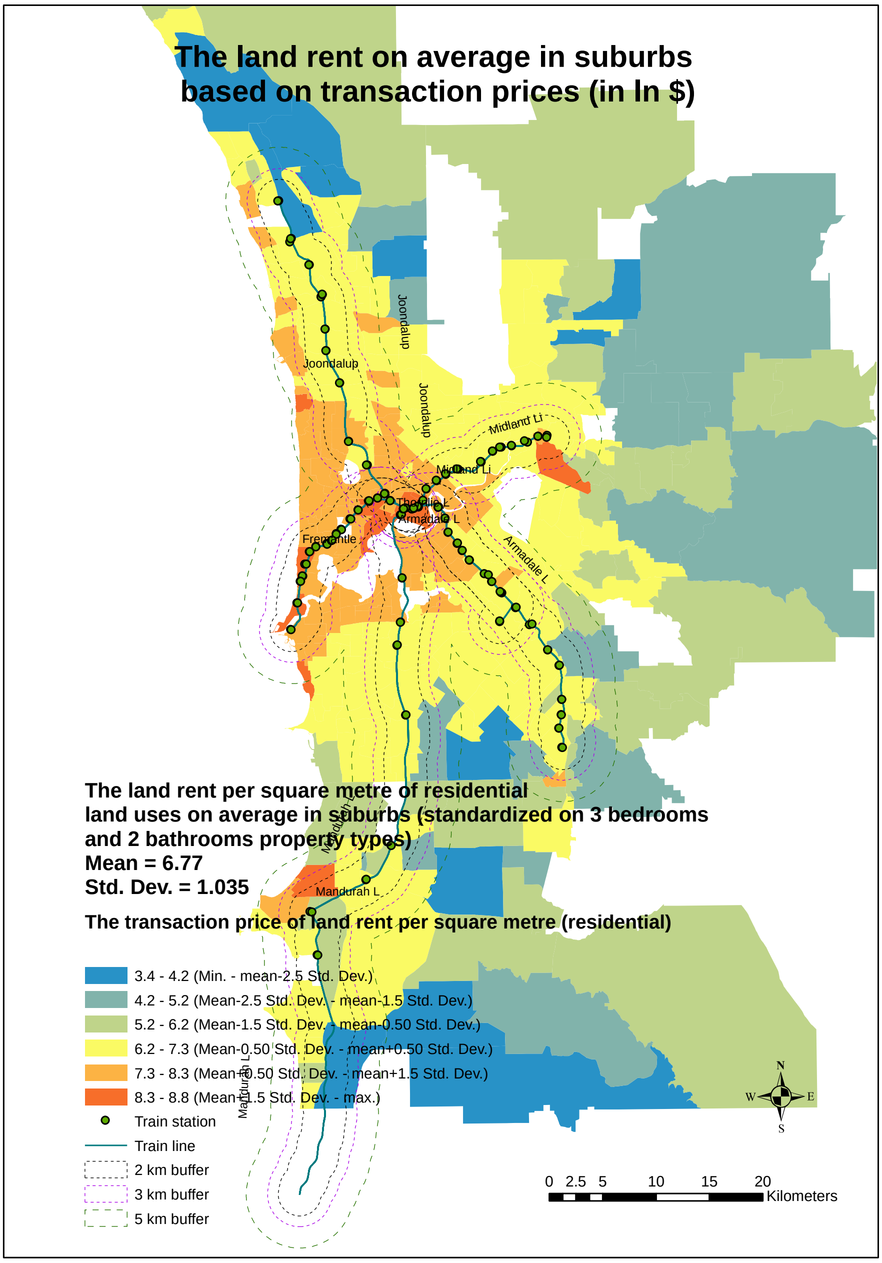


Figure 6.23 Map of the land rent for residential land use ($ per sq meter of land area)

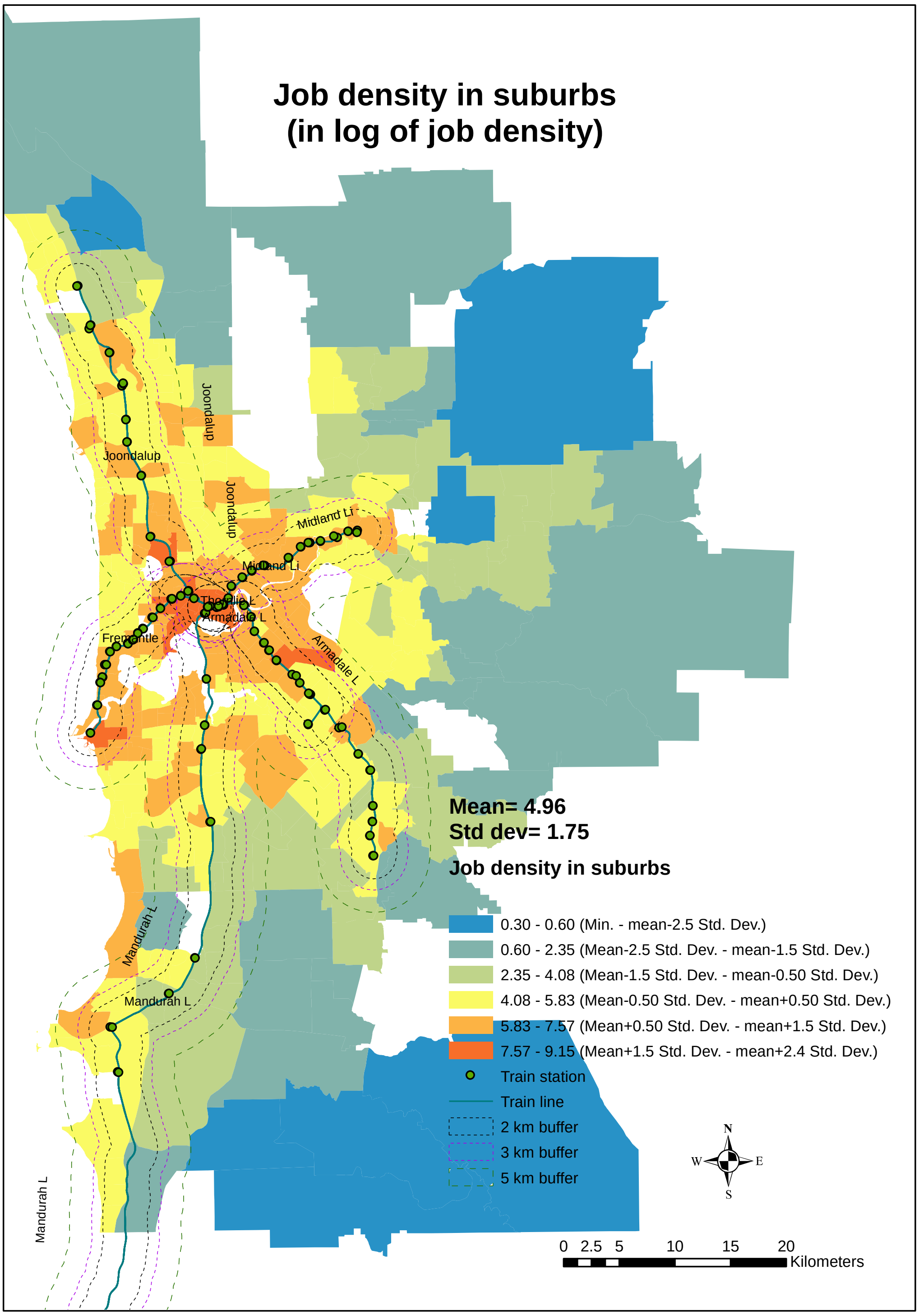


Figure 6.24 Map of land use of job density

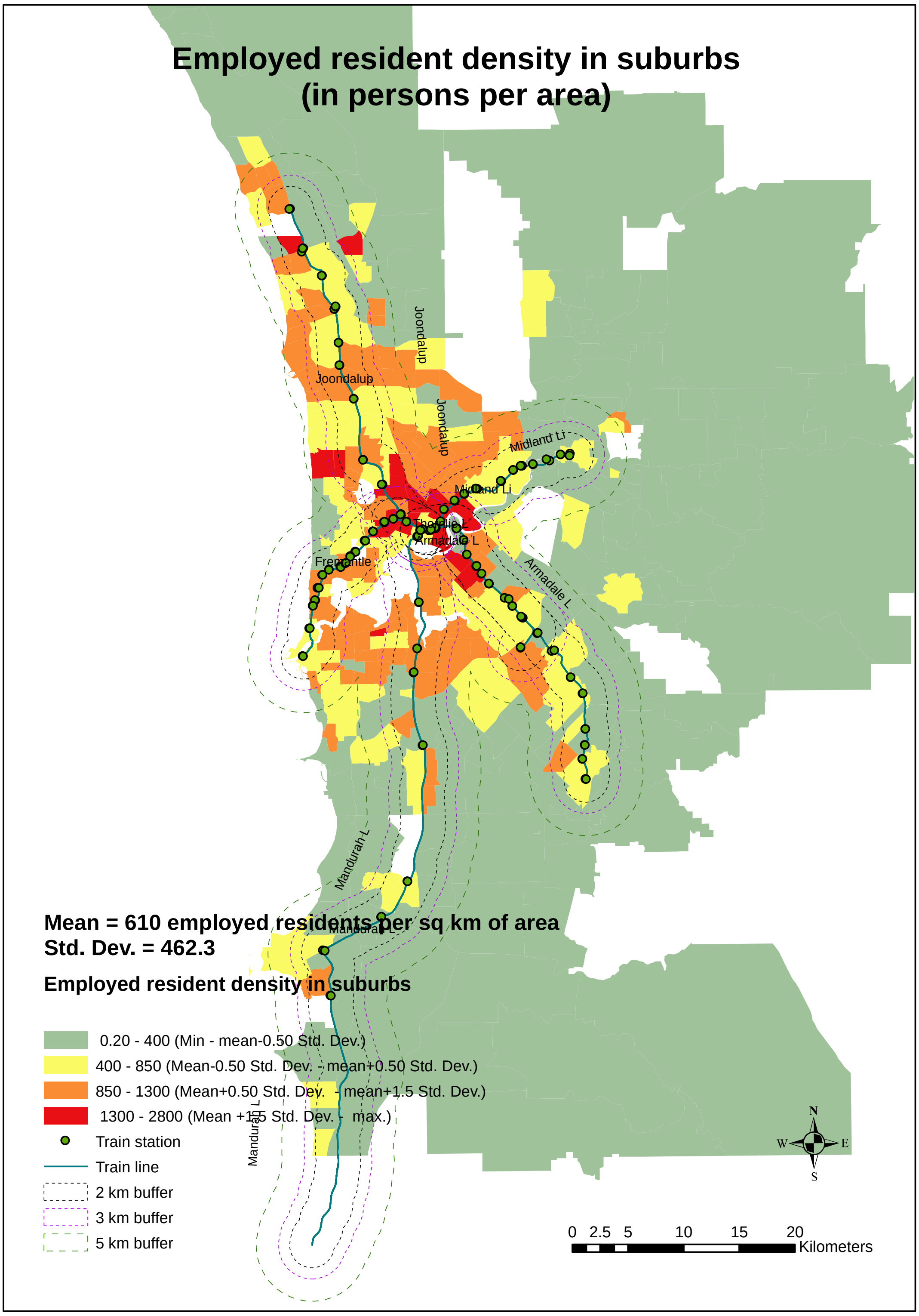


Figure 6.25 Map of land use of employed resident density

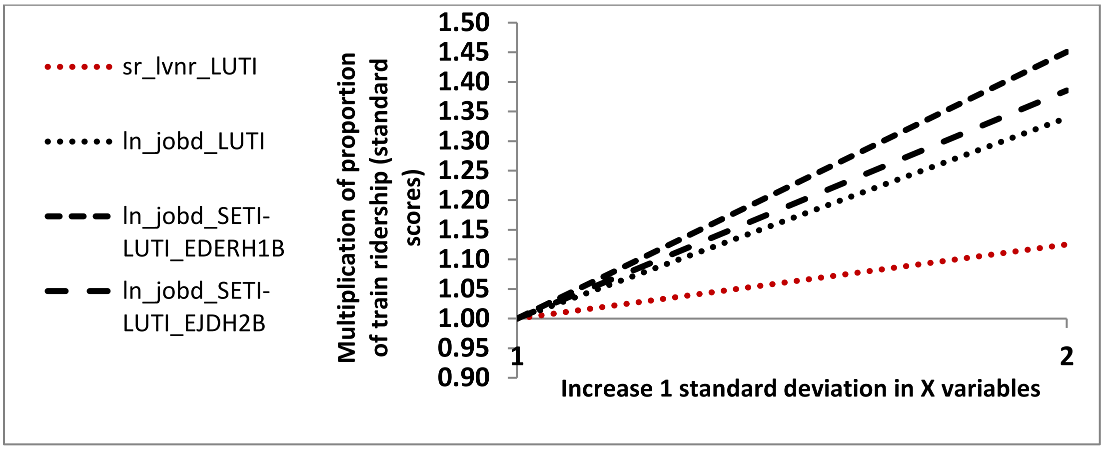


Figure 6.26 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in the land use component on the proportion of train ridership (train trip attraction model)

As shown in table 6.2, an increase of one standard deviation of *sr_lvnr* (std dev= $15.4 or equal to $237 per square meter land of non-residential land use) multiplied the proportion of train trip attraction by 1.125 (an increase of 12.5%) in the LUTI model. This result appears compatible with the agglomeration concept, as it applies to workplace suburbs. Land rent may be seen as a trade-off with transport cost. Higher accessibility implies higher land rent and this increases urban productivity due to higher agglomeration. Locations with higher agglomeration externalities were expected to attract more train ridership to workplace areas. Increasing *ln_jobd* (std dev= 1.745) by 1 percent of its standard score by (0.01745) multiplied the proportion of train trip attraction by 1.339 in the LUTI model (an increase of 33.9%), 1.451 in SETI h1b (an increase of 45.1%), and 1.385 in SETI h2b[14] (an increase of 38.5%).

The impact of the Perth-Mandurah railway line extension on land use development may exhibit risks of urban sprawl due to long urban lines, encouraging greater settlement at further distance from the central city. Ewing (1997) had defined the term “urban sprawl” based on four components: the expansion of leapfrog development,

[14] Increasing *ln_jobd* (std dev= 1.745) by 0.01745 is equivalent to multiplying job density by exp (0.01745) or by 1.0176. After rounding, a relative increase in job density of 1.76% multiplied the proportion of train trip attraction by 1.339 in the LUTI model, 1.451 in SETI h1b, and 1.385 in SETI h2b.

commercial strip development, low density development, and single-use development. Recently, Hamidi and Ewing (2014) pointed out the factor underlying those four components, i.e. poor accessibility. Poor accessibility may be observed in low density, single-use and leapfrog development where activities are segregated by large and many areas of vacant lands. Further negative implications of urban sprawl may emerge in terms of longer travel distances, which often are conducted by automobiles.

In this thesis, the tendency toward urban sprawl was measured by the percentage of land/property area sold following the rail extension of 2007-2013. The figures were compared between suburbs adjacent to the Mandurah railway line to the totals for the Perth metropolitan area. The land/property value transaction data included more than 300,000 transactions for the whole Perth metropolitan region during 2007 – 2013 (©Western Australian Land Information Authority - Landgate, 2014), accounting for a total 40,453.57 Ha of land area sold. In order to check for indications of urban sprawl, a three-kilometre buffer area was drawn along the Perth-Mandurah railway line, resulting in 67 suburbs within those buffer areas. There were 62,073 transactions from these areas which accounted for 4,882.768 ha land area sold. These figures represented 21% of the number of land transactions and 12% of the total land area sold in the whole Perth metropolitan region. In terms of total suburb land area, there were 446,398 ha of land area in the whole Perth metropolitan region, and the 3 km buffer area along the Perth-Mandurah railway line accounted for only 8.38%. The proportion of new land use development during 2007-2013 in the three-kilometre buffer areas to the total land area sold in Perth metropolitan was higher than the proportion of the buffer land area to the total Perth metropolitan land area. This may therefore be an indication of urban sprawl after the railway line extension along the Mandurah line.

An increase in unsustainable mobility patterns, such as travel distances and car use, most likely would be a poor outcome as a result of urban sprawl (García-Palomares, 2010). A comparison between the number of work trips (commuter) made by trains (all working trips involving trains) and cars (car drivers and car passengers) before-and-after the railway line extension was carried out. The number of modes were calculated for the whole Perth metropolitan region and for suburbs located adjacent to each of the railway lines. The results showed that the growth in the number of train uses was much higher than that of cars. The overall growth in the number of working trips made by train at the metropolitan level was 23.8% annually compared to that made by car, which was 3.02% annually over the period of 2006 – 2011. Deriving these figures at the suburb level and for each train line platform shows that the growth of train use for suburbs located in along the Mandurah line was 101.4% annually, compared to the growth figures for Fremantle, Midland, Joondalup, Armadale, and outside the platforms, which were 9.7%, 10.8%, 7.4%, 10.6% and 25.2% annually over the same period. Interestingly, the growth in car use was much lower than that of train use, and this growth is similar in the magnitude to that within platforms. For example, suburbs along the Mandurah line experienced growth in car use of 3% annually compared to Fremantle (1.9%), Midland (3.8%), Joondalup (1.5%), Armadale (4.7%) and suburbs located outside the platforms (3.7%). "If sprawl has any widely accepted outcome, it is automobile dependence and heavy automobile use" (Shima and Reid, 2014, p. 72). Referring to the travel data for the study area, there was an indication of lesser car dependence within suburbs adjacent to the Perth-Mandurah railway line, with the implication that the Perth-Mandurah railway line extension may be regarded to have had benefits of a more sustainable travel pattern.

6.3.4 Agglomeration or spatial economic variables

The interaction term between effective employed resident density and effective job density was added to the SETI models h1 and h2 to create the SETI h1b and h2b models. In the overall train trip production model, the contribution of the SETI h2b model (effective job density and its interaction term) to the total variance in the LUTI model was 2.1%. The interaction terms for effective employed resident density were not statistically significant, therefore model SETI h1b was not considered for this variable. The contribution of effective employed resident density without the interaction term in the SETI h1 was 3.05%. Therefore, the model of effective employed resident density without the interaction term (model SETI h1) was considered to be the best model for explaining train trip production. The relationship between effective employed resident density and train ridership was not influenced by the distance between residential suburbs and train stations.

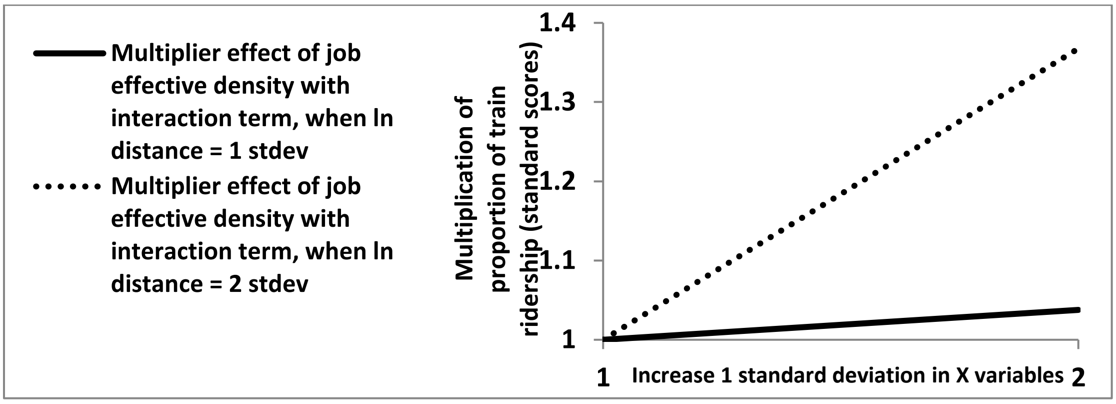


Figure 6.27 Illustration of plots of the multiplication effect based on the standardized value of predictor variables in the spatial-economic (effective density) component on the proportion of train ridership (Place of residence or train trip production model)

Including the interaction term, an increase of one standard deviation of effective job density (equal to 4,709 job opportunities) in residential suburbs when the distance between the suburb and train station is within one standard deviation of *ln_avedist* (or equal to 2 km) multiplies train ridership production by 1.0377 (an increase of 3.77%). However, when the distance of the suburb is within two standard deviations of *ln_avedist* (equal to 4 km), increasing one standard deviation of effective job density multiplied the proportion of train trip production by 1.3675 (an increase of 36.75%). Thus, the influence of effective job density in residential suburbs two standard deviations from the train station is 1.32 times greater than for the closer suburbs. In the model without the interaction term (SETI h1), increasing one standard deviation of effective employed resident density multiplied the proportion of train ridership by 1.4348 (an increase of 43.48%).

In the train trip attraction model, effective densities including the interaction term in the SETI model were found to be statistically significant, both in terms of effective employed resident density (h1b) and effective job density (h2b). The interaction effects produced a negative relationship with train ridership, while the original coefficient of effective density positively influenced ridership. The influence of effective density on ridership depends on the value of *ln_avedist*: This effect became less positive (or more negative) with an increase in the distance between workplace suburbs and train stations. The following graph (Figure 6.28) illustrates this. The influence of effective employed resident density for workplace suburbs placed one standard deviation closer to the train station was 1.6 times greater than the influence of effective employed resident density for the farther suburbs. Similarly, the influence of effective job density for workplace suburbs placed one standard deviation closer to the train station was 1.3 times greater than the influence of effective job density for the farther suburbs.

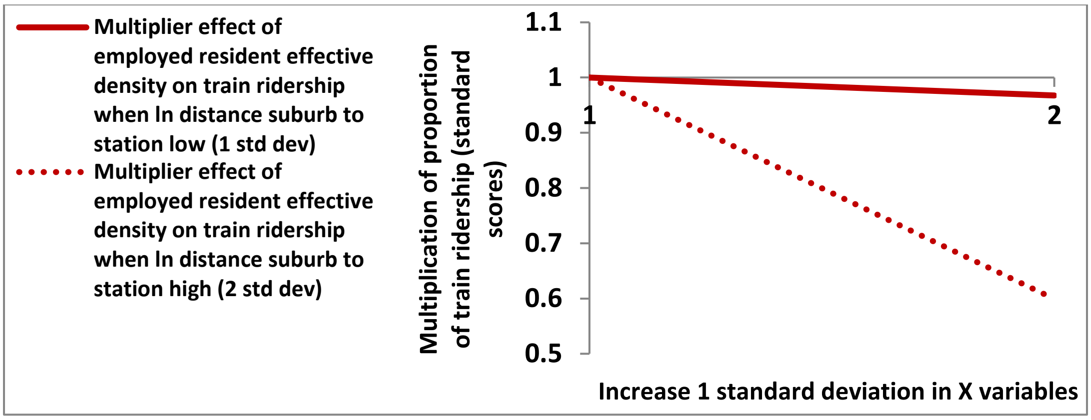


Figure 6.28 Illustration of plots of the multiplication effect based on the standardized value of predictor variables of the spatial-economic component (effective employed resident density with the interaction term) on the proportion of train trip attraction

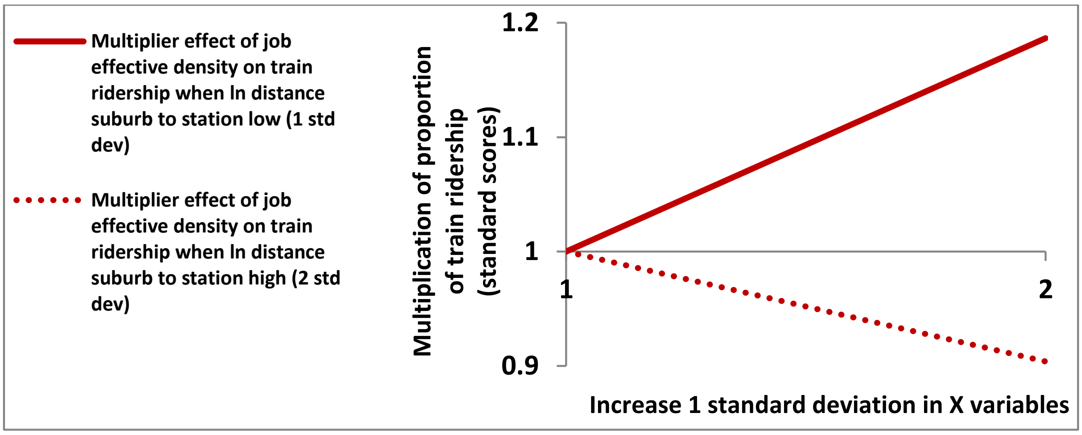


Figure 6.29 Illustration of plots of the multiplication effect based on the standardized value of predictor variables of the spatial-economic component (effective job density with the interaction term) on the proportion of train ridership attraction

6. 4 CHAPTER SUMMARY

This section discussed the comparisons between the LUTI and the SETI models for train trip production and attraction. The comparison aimed to determine the strength of each adjusted SETI model, with and without agglomeration, relative to the LUTI model (hypothesis 2).

This thesis has discussed variables for people, transport, land use components and effective density and their relationship to train ridership. The socio-demographic components being discussed were focused on job and employed resident structure, based on industry sector and type of occupation. Some socio-demographic variables that are part of the mainstream LUTI literature, such as income level and car ownership, were also used. Few studies have mentioned the importance of employment and occupation structures in train ridership (Hall 1969, Gutierrez 2011, Wang & Chai 2009).

Socio-economic variables remain an important determinant of train ridership in most train ridership modelling studies. This study found car ownership and income level negatively influence train trip production. Car ownership variables contain implications for the adjustment of behaviour to changing urban structure and the accessibility level of urban areas, as well as describing the socio-economic background of a household. Variables like income or wage level are often discussed as part of job location, residential location mobility, and trade-offs between wage and transport costs relationships, in addition to explaining the vertical segregation of households or individuals. Other study has found car ownership has a strong association with choices to use trains or cars for those who live in within station precincts (de Grange at al., 2012).

In terms of transportation or accessibility variables, this thesis found that train trip production and attraction have been more sensitive to the influence of road network travel distance, notably an indicator performance of competing mode, such as buses and cars. Note that the high dependence on car has been a trademark of Perth's city has made Perth's train ridership level depends much on the level of service of its competing modes. Low centrality by car corresponds with higher train ridership. This has been suggested in other study. The ratio of public transport to road supply has positive correlation with the total travel distance by public transport (Van de Coevering & Schwanen, 2006).

The level of influence of the transportation component is much stronger than the level of influence of socio-demographic/economic variables. Similarly, other study found that the important of accessibility parameter (by public transport) appeared to be more important than household income and vehicle ownership (J. Lin & Long, 2008).

The influence of land use variables were less than that of transportation variables but approximately similar to the influence of the socio-economic/demographic variables. Kain and Liu (1999) found that the accessibility factors is more important than land use factors such as land use mix and density in influence the travel behaviour. This has related with the direct effect of accessibility on ridership, such as the policy of expanding train networks will stimulate the use of public transport (de Grange et al., 2012).

The SETI h1 model was the preferred model for describing train trip production, and the SETI h1b model for the train trip attraction.

CHAPTER 7. THE GEOGRAPHICAL EXTENT OF AGGLOMERATION

7. 1 INTRODUCTION

This chapter discusses the third research hypothesis in relation to the model for train ridership prediction and the discussion on the overall model fits. The first part of discussion examines the geographical extent of the influence of agglomeration, in terms of effective density, on train trip production and attraction. This has been part of testing research hypothesis 3. The analyses have been conducted on the all-sector model and the sector-based model (the construction, manufacturing, and retail sectors) in both the LUTI model with the interaction term and the SETI model with the interaction term.

The second part of the discussion compares the LUTI model to the SETI model, specifically examining the extent to which effective density improves the prediction capability of the latter model. This relates to the testing of hypothesis 2. The section compares the model fits for the LUTI and the SETI models in terms of their prediction capability. *First,* the overall changes in model fit that are indicated by the adjusted R-square and the changes in R-square (*coefficient of determination)* between the LUTI to the SETI model are discussed. *Secondly,* the contributions of effective density variables to any improvements of the LUTI model are measured in terms of the *coefficient of partial determination* and their significances by the partial F-test statistic. *Thirdly,* both the LUTI and the SETI model are compared, in terms of their future prediction capability, based on external validation or back-testing using the 2006 data set. Discussion in this section also emphasizes the all-sector model.

7. 2 THE GEOGRAPHICAL EXTENT OF AGGLOMERATION

7.2.1 Method

The interaction effect was included in the model in order to examine the geographical extent of agglomeration, measured from train station as the focal point. The interaction term was constructed as a representation of the magnitude of agglomeration (effective density) and the average distance of a suburb from the nearest train station. The interaction terms incorporate the information about how much the influence of agglomeration on train ridership depends on the average distance of a suburb from the nearest train station, thus creating a measure of geographical extent of the influence of agglomeration. In this way, the interaction term may be able to show the threshold distance values beyond which agglomeration no longer will influence train ridership.

The examination of the geographical extent of agglomeration has based on on the value of the unstandardized beta coefficient. The model of train ridership prediction in the form of logarithmic-transformed variables of OLS regression incorporates the interaction term as followed: (Taplin, 2016)

$$\ln(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \beta_{n+1} x_m x_n + \varepsilon, where\ m \leq n \quad \text{Equation 7.1}$$

The number of predictor variable is represented by *n*. The *m* and *n* are different values from the set 1, 2, …*n*. The product of $x_m x_n$ denotes the interaction effect between the average distance of a suburb to the nearest train station (x_m) and the effective density (x_n). To simplify, the interaction terms may be written as a single predictor regression, hence the coefficient for x_n may account for the influence of x_m.

$$\ln(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_{n-1} x_{n-1} + (\beta_n + \beta_{n+1} x_m) x_n + \varepsilon \quad \text{Equation 7.2}$$

Therefore, the relationship between $\ln(y)$ and x_n now depends on the value x_m. The multiplicative effect from the interaction term is interpreted as: increasing x_n by Δ_n increases $ln(y)$ by $(\beta_n + \beta_{n+1}x_m)\Delta_n$ (Taplin, 2016).

The multiplicative effect of increasing x_n by Δ_n on the proportion of train ridership is calculated as follows (Taplin, 2016):

$$\frac{y(x_1,x_2,\dots,x_n+\Delta_n)}{y(x_1,x_2,\dots x_n)} = \exp(\beta_n + \beta_{n+1}x_m)^{\Delta_n} = (\exp(\beta_n)\,(\exp\beta_{n+1})^{x_m})^{\Delta_n} \quad \text{Equation 7.3}$$

The comparison of models in this section is performed for both the all-sector model and the sector-based models (for the construction, manufacturing and retail sector). In the sector-based models, some sector-specific variables in the LUTI model consisted of natural log of job density and natural log job to employed resident ratio (only for train trip attraction); employed resident density and natural log job to employed resident ratio (for train trip production). All other variables involving transportation (accessibility), socio-demographic/economic, and land value were used in any of the sector based models. In the SETI model, all of predictor variables in the LUTI model were included, with an addition of sector-specific variables for effective density, such as the sector based effective employed resident density and effective job density.

7.2.2 Results

7.2.2.1 Train trip production model

In the model of train trip production (residential suburbs), the agglomeration variable modelled in SETI with the interaction term was determined to be statistically significant for the all-sector model, and for the construction and manufacturing sectors of the SETI h2b (effective job density) model. Effective employed resident density was not statistically significant in any of the train trip production models.

There was a positive unstandardized coefficient of interaction in all of SETI, based on the effective employed resident density (model SETI h1), such that as distance between residential suburbs and train stations increased, the magnitude of influence of effective density on train ridership increased (table 7.1).

The model also shows the threshold value of distance beyond which effective job density will not influence train trip production. When a train station is located within 2.2 km from a residential suburb, then the variable of effective job density modelled in SETI h2b had no effect on train trip production. Specific sectors such as the construction and manufacturing sectors were found to have similar threshold lower bound distances, such as 2.7 km in the construction sector of the SETI h2b model and 1.73 km in the manufacturing sector of the SETI h2b model (table 7.1).

Table 7.1 Multiplicative effects of effective density with interaction term – train trip production (SETI h2b model)

Model	*Variable*	*Unst. beta coeff.*	*Stand. beta coef.*	*Sig*	*Exp (Unstd)*	*Exp (Std)*	*Multiplicative effect*	*Threshold lower bound distance*
The All sector model h1b (appendix 20 page 315)	Effective employed resident density	0.07	0.224	0.202	n/a		n/a	n/a
	Interaction effect	0.024	0.16	0.292	n/a			
The All sector model h2b (appendix 22 page 317	Effective job density	-0.044	-0.239	0.028	0.957	0.787	=[0.957*(1.045^ln_avedist)]^deltaejd	=0.044/0.044= 1 or = 2.7182 km
	Interaction effect	0.044	0.276	0.016	1.045	1.318		
The Construction model h1b (apeendix 24 page 318	Effective employed resident density	0.067	0.149	0.387	n/a		n/a	n/a
	Interaction effect	0.038	0.157	0.263	n/a			
The Construction model h2b (appendix 26 page 319)	Effective job density	-0.124	-0.389	0.002	0.883	0.678	=[0.883*(1.133^ln_avedist)]^deltaejdc	=0.124/0.125= 0.992 or = 2.7 km
	Interaction effect	0.125	0.495	0	1.133	1.640		
The Manufacturing model h1b (appendix 28 page 321)	Effective employed resident density	0.102	0.237	0.147	n/a		n/a	n/a
	Interaction effect	0.016	0.094	0.587	n/a			
The Manufacturing model h2b (appendix 30 page 322)	Effective job density	-0.04	-0.175	0.254	0.96	0.833	=[0.96*(1.076^ln_avedist)]^deltaejdm	=0.04/0.073= 0.548 or = 1.73 km
	Interaction effect	0.073	0.541	0.001	1.076	1.718		
The Retail model h1b (appendix 32 page 324)	Effective employed resident density	0.103	0.22	0.21	n/a		n/a	n/a

	Interaction effect	0.28	0.09	0.44	n/a		
The Retail model h2b (appendix 34 page 325)	Effective job density	-0.033	-0.128	0.274	n/a	n/a	n/a
	Interaction effect	0.044	0.187	0.09	n/a		

The increase in influence of effective density with distance from the train station might be explained in three ways:

First, the fact that effective density has been measured using the travel time for park-and-ride may influence the way in which effective job density influences train trip production. Park-and-ride serves residential suburbs when they are within a distance that cannot reasonably be served by other alternative modes of station access, such as by walking and bicycle. This implies that residential suburbs farther from the station may generate more train production provided the park-and-ride facilities are available. On the other hand, residential areas closer to the train station can easily access the station by walking or bicycle. Thus, the stronger influence of effective job density for more distant residential suburbs may refer to the possibility of this variable reflecting the dominance of park-and-ride travel behaviour.

Secondly, the spatial distribution of effective density has shown that the magnitude of effective density follows a distance-decay pattern. This means that the magnitude of effective density is lower in residential suburbs located further from the train station (see Chapter 5). Thus, increasing effective density in the areas where job opportunities were previously low may create a stronger multiplicative effect on train ridership than increasing effective density at the same level in other places. Increasing effective density implies also improvements in the accessibility of these areas, such that increasing accessibility may have caused greater impacts on the relative influence of effective job density on train ridership compared to other areas.

Thirdly, a lesser influence of effective job density on train trip production for residential suburb located near train stations might be due to the fact that effective job density had been very high for suburbs near stations hence the chance of attracting train trips was more likely than to produce trips. Employed residents living in this area may possibly have jobs within the same suburb, making use of the job intensity, thus fewer train trips would be produced or more short distance trips would be produced, resulting in lower train trip production.

Table 7.2 showed the multiplicative effects as they vary by distance, including those residential suburbs located within 100 meters, 1km, 10km and 20 km. The highlighted rows identify the cut-off distance at which effective density no longer influences train ridership (multiplicative factor = 1). The influences are compared to the influence of the land use density variable, i.e. employed resident density, within the same variable distances, on train trip production.

From the model results, increasing the same amount of effective job density in suburbs within 1km from a train station had negative (smaller) multiplicative effects on train ridership production than that of suburbs within farther distance rings. Since effective job density was counted in 1000 units, the regression model indicated that increasing 1 units of effective job density (1000) in the benchmark model (equalling 0.212 of its standard deviation), with other things being equal, multiplied train ridership by 0.957 for suburbs within 1 km distance (a decrease of 4.3%), but multiplied ridership by 1.004 (an increase of 0.4%) at a distance of 3km. An additional example is that increasing the effective job density by one standard deviation (equal to 4.717 job opportunities counted in 1000 units) resulted in the multiplicative effect on train ridership of 0.81276 (a decrease of 18.7%) when the suburb is located within 1 km from train station, while increasing effective job density by the same amount for residential suburbs within 3 km from stations multiplied ridership by 1.019 (an increase of 1.9%).

This pattern also applied to other sectors, where the construction sector was noted to generate the highest increases in train ridership production, relative to the manufacturing sector and the all-sector model (table 7.2).

Table 7.2 Multiplicative effects of effective density including the interaction term varied by distance – train trip production (SETI h2b model)

Variable (X)	*Distance of origin suburb to train station (km)*	*ln dist*	*Multiplicative effect of X on Y the proportion of train ridership*
Effective job density (the all sector model all sector in 1000 units)	0.1	-2.303	0.865
	1	0.000	0.957
	2.718282	1.000	1.000
	3	1.099	1.004
	5	1.609	1.027
	10	2.303	1.059
	20	2.996	1.092
Effective job density (the construction sector in 100 units)	0.1	-2.303	0.662
	1	0	0.88338
	2.696622	0.992	1
	3	1.098612	1.013416
	5	1.609438	1.080236
	10	2.302585	1.178006
	20	2.995732	1.284625
Effective job density (the manufacturing sector in 100 units)	0.1	-2.303	0.811
	1	0	0.96
	1.777739	0.575342	1
	3	1.098612	1.04
	5	1.609438	1.08
	10	2.302585	1.136
	20	2.995732	1.195

Figure 7.1 illustrates the relative strength of multiplicative effects between sectors on train trip production. The effective construction job density had the highest influence compared to the manufacturing sector and to all sectors. The pattern of lines indicates a stronger positive effect as residential suburbs are located further from the train station.

On the other hand, the LUTI model demonstrated that the interaction terms between employed resident density and distance to stations were statistically significant for the all-sector model, the construction, and the retail models. However, the interaction term relating the employed resident density to distance of residential suburb to the train station appears to have a limited contribution to train ridership. For example, increasing 1 employed resident per sq km (equal to 0.002159 of its standard deviation), resulted in a multiplicative effect of 1 in every location, i.e. there was no influence from employed resident density on train trip production everywhere (see the result in table 7.3 and 7.4).

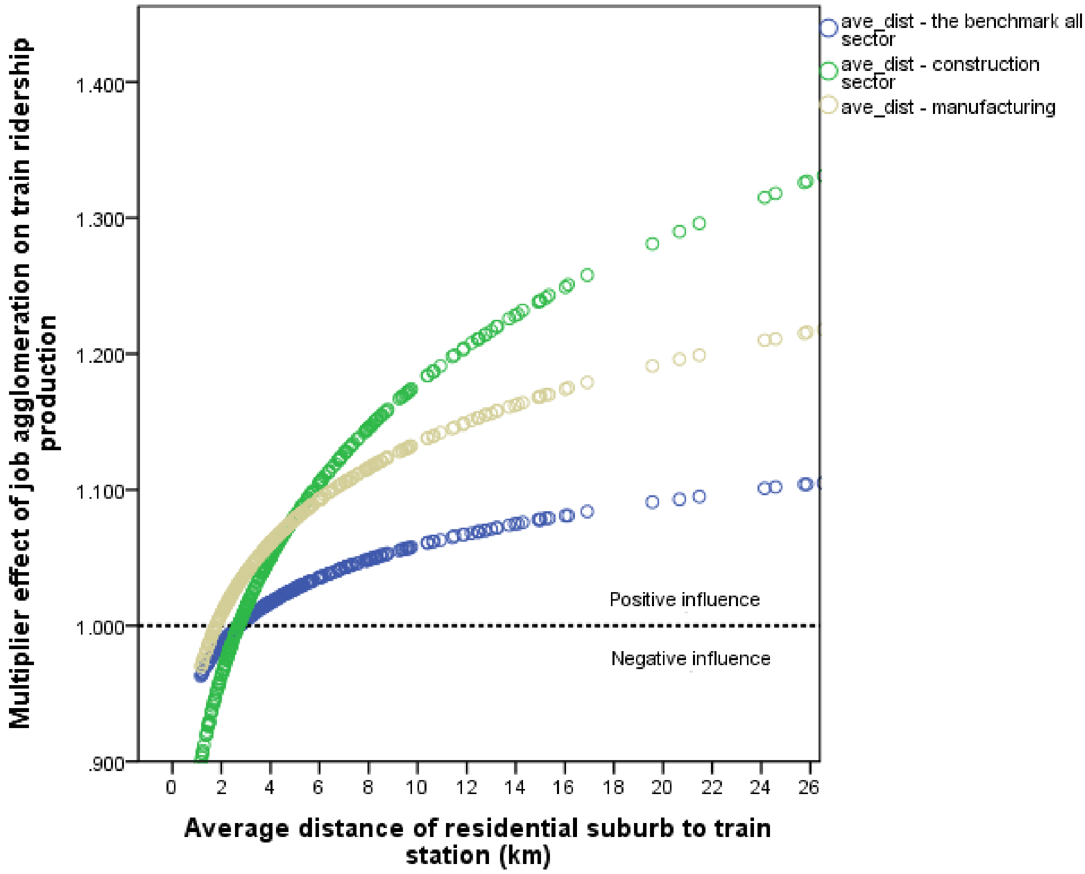


Figure 7.1 Multiplicative effect of effective job density on train trip production, compared between sectors (SETI h2b model)

Table 7.3 Multiplicative effect of employed resident density with interaction term varied by distance – Train trip production (LUTI model)

Model	*Variable*	*Unst beta*	*Stand beta*	*Sig*	*Exp (unst)*	*Exp (std)*	*Multiplicative effect*	*cut-off lower bound*
LUTI								
All sector (appendix 12 p. 310)	Employed resident density	-.00020	-.10434	.33728	1.000	0.901	=[1.000*(1.000 ^ln_avedist)]^ deltawod	=0.0002/ 0.00037 = 0.524 or 1.69 km
	Ln distance to train sta	-.85642	-.67709	.00000	0.425	0.508		
	Interaction effect	.00037	.25371	.00683	1.000	1.289		
Construction (appendix 14 p. 312)	Employed resident density	-.00083	-.04284	.74122	0.999	0.958	=[0.999* (1.003^ln_avedist)] ^deltawodc	=0.00083/ 0.00338= 0.245 or 1.277 km
	Ln distance to train sta	-.82434	-.65172	.00000	0.439	0.521		
	Interaction effect	.00338	.26634	.02545	1.003	1.305		
Retail (appendix 18 p. 314)	Employed resident density	-.00182	-.097	.398	0.998	0.907	=[0.998*(1.003 ^ln_avedist)]^ deltawodr	=0.00182/ 0.00289= 0.629 or 1.87 km
	Ln distance to train sta	-.839	-.663	.00000	0.432	0.515		
	Interaction effect	.00289	.21663	.036	1.003	1.242		

Table 7.4 Multiplicative effect of employed resident density with interaction term varied by distance – Train trip production (LUTI model)

Variable (X)	*Distance of origin suburb to train station (km)*	*ln dist*	*multiplicative effects of X on Y train ridership*
Employed resident density (the all sector)	1	0	0.999803898
	1.68903	0.524154	1
	10	2.302585	1.000665652
	20	2.995732	1.000925211
Employed resident density (construction)	1	0	0.999172745
	1.277343	0.244782	1
	10	2.302585	1.006981589

	20	2.995732	1.009344214
Employed resident density (retail)	1	0	0.998
	1.729274	0.547702	1
	10	2.302585	1.004
	20	2.995732	1.007

Thus, employed resident density with the interaction term (in LUTI model) was observed to have no influence on train trip production, while effective job density with the interaction term (in SETI h2b model) produced a statistically significant influence on train trip production. These influences level were variable with the distance of the residential suburb from the nearest train station. Therefore, policy that attempts to maximize the effect of effective job density on train trip production should consider the spatial distribution of stations relative to the distribution of residential suburbs in the system.

7.2.2.2 Train trip attraction model

In the train trip attraction model (for workplace suburbs), the effective density variable, modelled in SETI with the interaction terms h1b and h2b, was calculated to be statistically significant in all models. Thus, the comparison of the influence of the interaction terms for job density and effective density may be made for all models of SETI h1b and h2b.

Table 7.5 clearly shows negative unstandardized coefficients for interaction terms in all of the SETI models; also note that, as the distance of a workplace suburb from the train station increases, the magnitude of influence of effective density on train ridership was calculated to decrease. This means that the effective density and the proportion of train ridership were often higher when the workplace was closer to the train station. When the train station serving the workplace suburb is beyond 7 km distance, then the effective job and employed resident density were calculated to have no effect on train ridership attraction. Thus, locating jobs within the station precinct may be considered necessary in order to capture the maximum benefit of agglomeration in increasing train trip attraction.

Table 7.5 Multiplicative effect of effective density with interaction term – train trip attraction (SETI h1b and h2b model)

Model	Variable	Unst. beta coeff.	Stand. beta coef.	Sig	Exp (Unst)	Exp (Std)	Multiplicative effects	The threshold distance upper bound
The All sector h1b (appendix 44 page 331)	Effective employed resident density	0.226	0.446	0	1.254	1.562	=[1.254*(0.891^ln_ avedist)]^ deltaeder	=-0.226/-0.115 = 1.9652 =7.1365 km
	Interaction effect	-0.115	-0.479	0	0.891	0.619		
The All sector SETI h2b (appendix 46 p. 333)	Effective job density	0.132	0.443	0.002	1.141	1.557	=[1.141*(0.932^ln_ avedist)]^ deltaejd	=-0.132/-0.07 = 1.8857= 6.59 km
	Interaction effect	-0.07	-0.272	0.009	0.932	0.762		
The Construction model h1b (appendix 48 p. 334)	Effective employed resident density	0.229	0.316	0.009	1.257	1.372	=[1.257*(0.872^ln_ avedist)]^ deltaederc	=-0.229/-0.137 = 1.67 =5.32 km
	Interaction effect	-0.137	-0.352	0.006	0.872	0.703		
The Construction model h2b (appendix 50 p. 336)	Effective job density	0.161	0.314	0.013	1.175	1.369	=[1.175*(0.870^ln_ avedist)]^ deltaejdc	=-0.161/-0.139 = 1.158=3.184 km
	Interaction effect	-0.139	-0.341	0.004	0.870	0.711		
The Manufacturing model h1b (appendix 52 p. 337)	Effective employed resident density	0.279	0.395	0.003	1.322	1.484	=[1.322*(0.880^ln_ avedist)]^ deltaederm	=-0.279/-0.128= 2.1797 =8.84 km
	Interaction effect	-0.128	-0.456	0.003	0.880	0.634		
The Manufacturing model h2b (appendix 54 p. 339)	Effective job density	0.089	0.24	0.067	1.093	1.271	=[1.093*(0.939^ln_ avedist)]^ deltaejdm	=-0.089/-0.063 = 1.41=4.1 km
	Interaction effect	-0.063	-0.29	0.03	0.939	0.748		
The Retail model h1b (appendix 56 p. 340)	Effective employed resident density	0.203	0.272	0.035	1.225	1.313	=[1.225*(0.890^ln_ avedist)]^ deltaederr	=-0.203/-0.116= 1.75 =5.75 km
	Interaction effect	-0.116	-0.257	0.022	0.890	0.773		
The Retail model h2b (appendix 58 p. 342)	Effective job density	0.072	0.171	0.119	1.075	1.186	=[1.075*(0.921^ln_ avedist)]^ deltaejdr	=-0.072/-0.082 = 0.878=2.4 km
	Interaction effect	-0.082	-0.217	0.029	0.921	0.805		

Interestingly, specific sectors such as the construction, manufacturing, and retail sectors were found to have different threshold upper bound distances than those of the train production model. In order to capture the benefit of effective employed resident density on train ridership, the workplaces for construction should be located within 5 km of the station; and within 3 km to capture the benefit of effective job density. Similarly, for the manufacturing sector, train stations serving workplaces in manufacturing should be no more than 8 km and 4 km away, to capture the benefit of effective employed resident and job density respectively. The retail sector was the most sensitive sector in relation to the influence of train station distance to effective job density on train ridership. Workplaces for retail jobs should be no farther than 2 km from the train station and 5 km to capture the effective employed resident and job density respectively.

The explanation for why the effective job and employed resident density may have been stronger for workplace suburbs nearer stations may be explained in a way conversely to the explanation for the case of residential suburbs (section 7.2.2.1). *First*, the fact that effective density is measured using the travel time for park-and-ride, means that employed residents leave their car at the train station nearest their residential suburb. Once they arrive at their destination, they are likely to prefer to take the minimum egress time (i.e. the time it takes to walk from the destination station to their job locations), and therefore workplace locations closer to the train station become more attractive for train trips. This highlights the importance of minimizing the egress time for train users once they reach the destination stations.

Secondly, based on the distance decay pattern of effective density, the area within 0-2 km of the station are notably the densest areas in the metropolitan region. Increasing effective density in these highly developed areas of the metropolitan region create positive externalities on train ridership, probably due to two facts: the worse congestion in these areas and the limited car parking spaces deters people from using cars or other street based travel modes, including perhaps buses.

Thirdly, the train trip attraction model found that effective job density had a weaker relationship with train trip attraction than effective employed resident density, probably due to the fact that jobs in the nearby workplace suburbs will also attract ridership - there may be competition in attracting ridership between nearby workplace suburbs. On one hand, closer to the train station, this job competition in attracting trips may become greater, thus lowering the influence of effective job density for suburb in question. On the other hand, the strong influence of effective employed resident density also implies lower self-sufficiency among residential suburbs. Thus, people living in residential suburbs travel by train to their workplaces near stations more than to areas further from train stations.

Table 7.6 showing the multiplicative effect of the interaction terms, varying with distance for workplace suburbs located within 100 meters, 1 km, 10 km and 20 km of the station. The highlighted rows in table 7.6 indicate the threshold distance at which the effective density no longer influences train ridership (multiplicative factor = 1). These influences are compared to the influence of the land use density variable, i.e. job density, on train trip attraction.

Table 7.6 Multiplicative effect of effective density with interaction term varied by distance – train trip attraction (SETI h1b and h2b model)

Variable (X)	*distance of destination suburb to train station (km)*	*ln dist*	*multiplicative effect of X on Y train ridership*	*Variable (X)*	*distance of destination suburb to train station (km)*	*ln dist*	*multiplicative effect of X on Y train ridership*
Effective employed resident density (The all sector model in 1000 units)	0.1	-2.303	1.634	Effective job density (The benchmark sector in 1000 units)	0.1	-2.303	1.341
	1	0	1.253576		1	0	1.141108
	2	0.693147	1.15753		2	0.693147	1.087063
	3	1.098612	1.104795		3	1.098612	1.056643
	5	1.609438	1.041763		5	1.609438	1.019528
	7.136464	1.965217	1		6.591061	1.885714	1
	10	2.302585	0.961946		10	2.302585	0.971241
	20	2.995732	0.888244		20	2.995732	0.925241
Effective employed resident density (construction in 100 units)	0.1	-2.303	1.724	Effective job density (construction in 100 units)	0.1	-2.303	1.618
	1	0	1.257342		1	0	1.174685
	2	0.693147	1.143437		2	0.693147	1.066788
	3	1.098612	1.081653		3.18443	1.158273	1
	5.320317	1.671533	1		5	1.609438	0.939214
	10	2.302585	0.917178		10	2.302585	0.852946
	20	2.995732	0.834089		20	2.995732	0.774601
Effective employed resident density	0.1	-2.3026	1.7749	Effective job density (all	0.1	-2.3026	1.2637
	1	0	1.321807		1	0	1.093081
	2	0.693147	1.209585		2	0.693147	1.046375

(manufacturing in 100 units)	3	1.098612	1.148409	manufacturing in 100 units)	3	1.098612	1.019984
	5	1.609438	1.075722		4.107023	1.412698	1
	8.843542	2.179688	1		5	1.609438	0.987682
	10	2.302585	0.984392		10	2.302585	0.94548
	20	2.995732	0.900816		20	2.995732	0.905081
Effective employed resident density (retail in 100 units)	0.1	-2.303	1.600	Effective job density (retail in 100 units)	0.1	-2.303	1.298
	1	0	1.225072		1	0	1.074655
	2	0.693147	1.130426		2.4062	0.878049	1
	3	1.0986123	1.078489		3	1.0986123	0.982076
	5.754603	1.75	1		5	1.6094379	0.941789
	10	2.3025851	0.937911		10	2.3025851	0.889752
	20	2.995732	0.865451		20	2.9957323	0.840591

From the model results, increasing the effective employed resident density for workplace suburbs located within 1-2 km of the train station was found to have a larger multiplication effect on train ridership than increasing effective employed resident density by the same amount to location within farther distance rings. As effective employed resident density has been counted in 1000 units (0.361 standard deviation), the regression model showed that increasing 1 unit (1000) of total effective employed resident density for the benchmark sector in workplace suburbs located within 1-2 km, multiplied train ridership by 1.157 to 1.25 times (an increase of 15.7% – 25%), other things being equal, while increasing it by the same amount for locations within 3-5 km from station, multiplied ridership by 1.04 to 1.109 times (an increase of 4%-11%). Using one standard deviation increase in effective employed resident density (equal to 2,770 of employed resident opportunities) multiplied the proportion of train trip attraction by 1.497 to 1.855 (an increase of 49.7 – 85.5%) for suburbs within 1-2 km from train station; while increasing it by the same amount on suburbs located within 3-5 km from the station multiplied the proportion of train trip attraction by 1.1147 to 1.33 (an increase of 11.5% - 33%).

Moreover, referring to the TOD policy that utilises the concept of densification within 800 meters from train stations, increasing 1 unit (=1000) of effective employed resident density in the benchmark sector of the TOD precincts (800 meter radius around the train station), all other things being equal, would multiply ridership by 1.286 (an increase of 28.6%). Increasing this effective density by one standard deviation (equal to 2,770 of employed resident opportunities) within the TOD precinct would multiply the proportion of train trip attraction by 2.007 (an increase of 100.7%).

The same relationship applies to other sectors such as the construction, manufacturing, and retail sectors and to the other agglomeration type (effective job density).

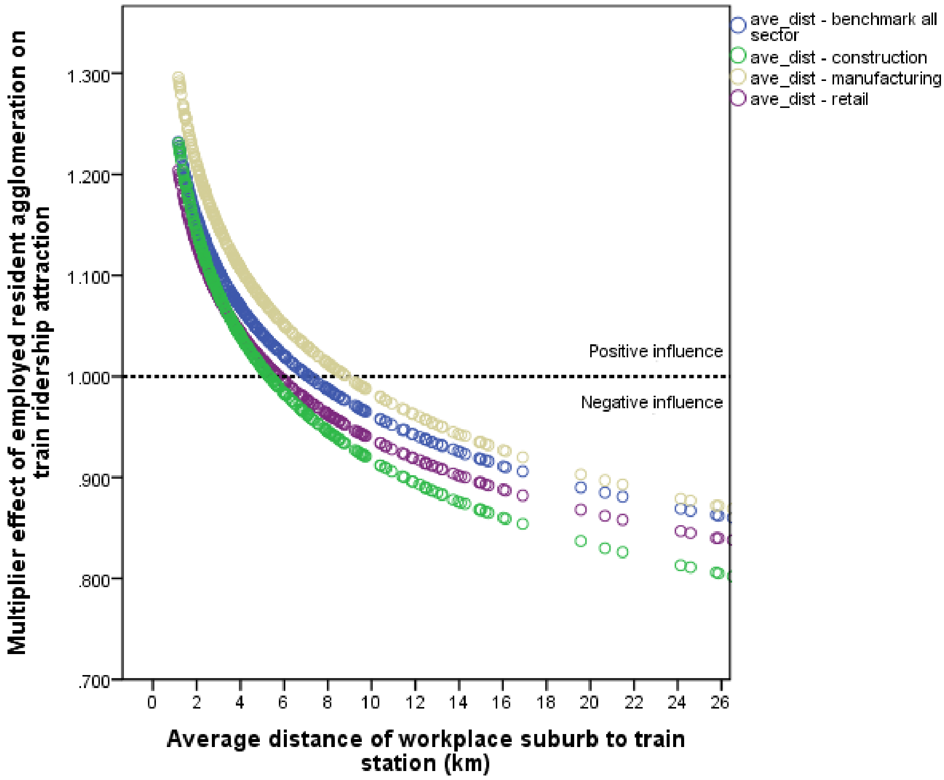


Figure 7.2 Multiplicative effect of effective employed resident density compared between sectors

Figure 7.2 illustrates the relative strength of multiplicative effects on train trip attraction between sectors. Effective employed resident density in the manufacturing sector had a higher influence than all sectors combined, the retail sector and the construction sector. The graph indicates a stronger negative effect as the distance from the train station increases.

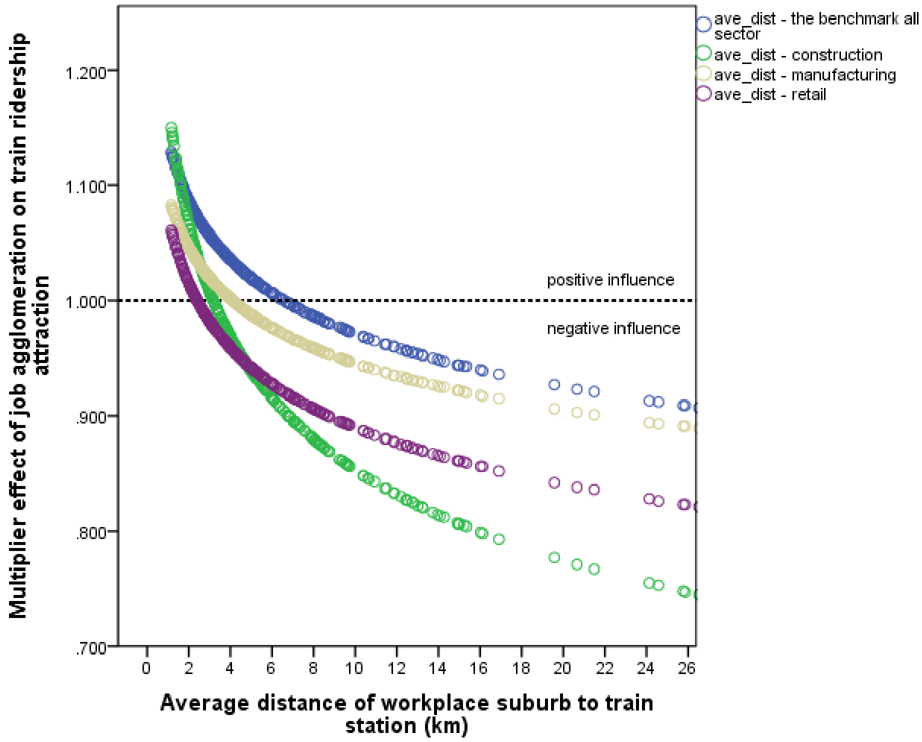


Figure 7.3 Multiplicative effect of effective job density compared between sectors

Similarly, figure 7.3 illustrates the relative strength of multiplicative effects between sectors for effective job density. Effective job density for the manufacturing and construction sector had a higher influence than the retail sector, whereas all sectors combined (the benchmark) were calculated to have the highest effect on train ridership after a certain distance. The graph indicates a stronger negative effect with increasing distance from the train station.

On the other hand, the multiplicative effect of job density with the interaction term on train ridership outweighs the influence of effective density with the interaction term. After including the interaction term, the contribution of natural log of density on ridership was always positive, and reduced with increasing distance of the workplace from the train station. Increasing 1 unit of natural log job density (0.573 standard deviations in the benchmark sector) for workplace suburbs within 1 km of the train station, the multiplicative effect of job density was higher than many other variables: 1.925 (an increase of 92.5%) for the construction sector, 1.5 (an increase of 50%) for manufacturing and 1.726 (an increase of 72.6%) for retail (see table 7.7). Increasing natural log job density by 1.745 (equal to one standard deviation) increased the proportion of train trip attraction by 3.187 (an increase of 218.7%) when the suburb is located within 1 km of the train station. In comparison, this multiplier is higher than the multiplier for increasing one standard deviation of effective employed resident density for the same distance range, which was 1.855 (an increase of 85.5%). However, these relatively high multipliers may be due to the high skewed of the number of job, where the standard deviation is high and 1 standard deviation of log job density refers to high number of job. Besides, since the influence on train ridership only becomes statistically insignificant if workplace suburbs are very far from the train station may represent an upwards bias. For example, the threshold distance of log job density was as at 30 km for construction jobs, 24 km for manufacturing jobs, and 142 km for retail jobs, yet the maximum distance of workplace suburbs in the study area is around 16 km (see table 7.7).

Table 7.7 Multiplicative effect of job density with interaction term– train trip attraction (LUTI model)

Model LUTI	*Variable*	*Unst beta coeff*	*Stand beta coeff*	*Sig*	*Exp (unst)*	*Exp (std)*	*Multiplicative effect*	*The threshold of upper bound distance*
All sector (appendix 36 p. 326)	Ln job density	.664	.827	.000	1.943	2.286	=[1.943*(0.825 ^ln_avedist)]^ deltalnjobd	=-0.664/ -0.193 = 3.4 or 31.3 km
	Interaction effect	-.193	-.373	.000	0.825	0.688		
Construction (appendix 38 p. 327)	Ln job density	.655	.741	.000	1.925	2.097	=[1.925*(0.823 ^ln_avedist)]^ deltalnjobdc	=-0.655/ -0.194 = 3.37 or 29.15 km
	Interaction effect	-.194	-.366	.000	0.823	0.694		
Manufacturing (appendix 40 p. 329)	Ln job density	.409	.524	.000	1.505	1.689	=[1.505*(0.879 ^ln_avedist)]^ deltalnjobdm	=-0.409/ -0.129 = 3.174 or 23.9 km
	Interaction effect	-.129	-.263	.010	0.879	0.769		
Retail (appendix 42 p. 330)	Ln job density	.546	.616	.000	1.726	1.852	=[1.726*(0.896 ^ln_avedist)]^ deltalnjobdr	=-0.546/ -0.110 = 4.95 or 141.5 km
	Interaction effect	-.110	-.176	.041	0.896	0.839		

Table 7.8 Multiplicative effect of job density with interaction term, varying by distance– train trip attraction (LUTI model)

Variable (X)	*Distance of destination suburb to train station (km)*	*ln dist*	*Multiplicative effect of X on Y train ridership*
Ln of job density (all sector)	1	0.000	1.943
	31.334	3.445	1.000
	50	3.912	0.914
Ln of job density (construction)	1	0.000	1.925
	29.150	3.372	1.000
	50	3.912	0.901
Ln of job density (manufacturing)	1	0.000	1.505
	23.908	3.174	1.000
	50	3.912	0.909
Ln of job density (retail)	1	0.000	1.726
	141.55	4.953	1.000
	150	5.011	0.994

Increasing density or effective density in an area located too far away from a train station may reduce train trip attraction due to the fact that job locations underserved by railway lines make car dependence higher, and therefore more jobs in these locations implies more travel demand that will be catered for by cars rather than by trains. Therefore, the policy of locating more jobs or more employed residents in outer suburbs far from the CBD should be accompanied by a policy to provide train stations serving these locations.

The multicolinnearity problem emerged as a result of the interaction between the variable distance and variable density or effective density. The indices of multicollinearity, in terms of the Variance Inflation Factor (or VIF indices) of over 10, indicate that these variables may exhibit a multicollinearity problem (Sung & Oh, 2011). Models including the interaction terms often had a VIF value above 10 (ranging from 10 to 25, as shown in the appendices that supported data in tables 7.5 and 7.7). However, this problem was unavoidable given that the interaction term was created based on the multiplication of the value of variable distance and the value of the variable effective density. By default, there would be multicollinearity issues either between the interaction variable with the variable distance; or between the interaction variable and the variable of effective density.

7. 3 OVERALL MODEL FIT AND THE CONTRIBUTION OF EFFECTIVE DENSITY VARIABLES IN THE MODEL PREDICTION

7.3.1 Method

This section involves the method for assessing model fit using linear regression with a logarithmic transformation. The coefficient of determination (R^2) measures the proportion of variation in the natural log of the proportion of train ridership that is explained by the set of predictor variables under the LUTI model and the SETI model. However, since the number of independent variables in the LUTI and the SETI models are different, an adjusted R^2 has been used to measure the prediction capability of both models, to account for the number of predictor variables and the sample size. The software SPSS that was used in modelling process calculates both the values of R^2 and the adjusted R^2. Following these analyses, the contribution of effective density variables to the LUTI model and the relative statistical fit of the SETI model were compared to the LUTI model. The contribution of each effective density variable in the SETI model was compared to the original LUTI model, by calculating the *coefficient of partial determination.* The formula is described as followed (Rose & Beck, 2007):

$$R_{agg}^2 = \frac{SSR\{(SETI\ model) - (LUTI\ model)\}}{SST\ (SETI\ model) - SSR\ (SETImodel) + SSR\ \{(SETI\ model) - (LUTI\ model)\}} \quad \text{Equation 7.4}$$

Whereas:

R_{agg}^2 = the coefficient of partial determination of the effective density variables that are added to the LUTI model to form the SETI model.

$SSR\{(SETI\ model) - LUTI\ model)\}$ = the sum of squares of the contribution of effective density variable in the SETI model given that the set of predictor variables in the LUTI model has been included in the SETI model.

$SST\ (SETI\ model)$ = the total sum of square for Y as in the SETI model.

$SSR\ (SETI\ model)$ = the regression sum of squares when all variables including land use variables have been included in the SETI model.

Testing portions of the model was conducted using the partial F-test at the 0.05 level of significance to determine whether the effective density variable (the SETI model) significantly improves the model given that all other variables (including the density variable) are included (the LUTI model). The testing hypothesis is as follows:

H0: effective density does not improve the model with the density variable included

H1: effective density does improve model.

The calculation of F-statistics is following the formula:

$$F_{STAT} = \frac{SSR\ SETI\ model - SSR\ LUTI\ model}{MSE\ SETI\ model} \qquad \text{Equation 7.5}$$

Whereas:

$SSR\ SETI\ model$ = the sum of squares of the SETI model.

$SSR\ LUTI\ model$ = the sum of squares of the LUTI model.

$MSE\ SETI\ model$ = the mean of squares of errors of the SETI model.

The results of F_{STAT} is compared to the results of F value at 5% level of significance (α= 0.05). If the value of $F_{STAT} > F_{0.05}$, there is enough evidence to reject the null hypothesis. This concludes that adding the effective density in the SETI model does improve the LUTI model.

Finally, the comparison between the LUTI and the SETI model was tested on another data set, i.e the 2006 data set. A simple residual analysis was performed to aid discussion of the models' relative performance.

7.3.2 Results

7.3.2.1 The overall model fit

Table 7.9 presents the overall model improvement achieved by changing from the LUTI to SETI models (h1 and h2b) for train trip production. There has been an improvement in terms of R-square increases, or of prediction capability, between the LUTI model and SETI h1 of approximately 4.3% relatively (from 39.1% to 40.8%). In addition, the SETI h2b is better correlated than the LUTI model by 2.87% relatively (from 39.1% to 40.0%). The model SETI h1 may therefore be preferred as the best model, followed by the SETI h2b, then finally the LUTI model. At maximum, 43.6% of the variation in the natural log proportion of train trip production may be explained by the SETI h1 model, when incorporating the adjustments for the number of independent variables and the sample sizes.

Table 7.9 The relative changes in the R-square among the LUTI and SETI models for train trip production (residential suburbs)

Model (Place of Residence)	*R^2*	*Adj R^2*	*AbsoluteR^2 changes*	*Relative R^2 changes*
Model	***%***	***%***	***SETI to LUTI model***	***SETI to LUTI model***
LUTI all sector	41.8	39.1	n/a	n/a
SETI all sector h1 (effective employed resident density without interaction effect)	43.6	40.8	1.8	4.3
SETI all sector h2b (effective job density with interaction effect)	43	40	1.2	2.87

The overall model improvement achieved by changing from the LUTI to the SETI (h1b and h2b) for train trip attraction (workplace suburbs) is presented in table 7.10. There has been an improvement in terms of the R-square increases between the LUTI model and the SETI h1b model of approximately 4.75% relatively (from 50.4% to 52.6%). In addition, model SETI h2b is better correlated than the LUTI model by 3.8% relatively (from 50.4% to 52.1%). The adjusted R-square value that reflects the number of independent variables also indicates that SETI is better correlated than the LUTI model. Model SETI h1b may therefore be preferred as the best model, followed by the SETI h2b, then the LUTI model. At maximum, 55.1% of the variation in the natural log proportion of train ridership attraction has been explained by the SETI h1b model, adjusted for the number of independent variables and the sample sizes.

Table 7.10 The relative changes in the R-square among the LUTI and SETI models in train trip attraction (workplace suburbs)

Model (Place of Work)	*R^2*	*Adj R^2*	*Absolute R^2changes*	*Relative R^2changes*
Model			***SETI to LUTI model***	***SETI to LUTI model***
LUTI all sector	52.6%	50.4%	n/a	n/a
SETI-LUTI all sector h1 b (eder with interaction effect)	55.1%	52.6%	2.5%	4.75%
SETI-LUTI all sector h2b (ejd with interaction effect)	54.6%	52.1%	2%	3.8%

7.3.2.2 The contribution of effective density variables

The contribution of the effective density and the interaction variable in the SETI model to the LUTI model has also been tested statistically. For the train trip production model, table 7.12 for the SETI h1 model shows a total sum of squares of 237.871, a sum of square of regression (SSR) for the SETI h1 of 103.696, and a regression sum of square for the LUTI model of 99.478 (table 7.11).

Table 7.11 ANOVA table of the LUTI model (Train trip production)

ANOVA[a]						
Model POR		Sum of Squares	df	Mean Square	F	Sig.
LUTI model All sector	SSR Regression	99.478	14	7.106	15.455	.000b
	SSE Residual	138.392	301	.460		
	SST Total	237.871	315			

a. Dependent Variable: ln_train_wo

b. Predictors: (Constant), ln_jwr, pblu_r, sqrt_ptiori, p_woret, ln_avedist, p_worker_man, inc_blu_r, car_own, p_wocon, ln_lvr, wod, ln_strio, inc_pr_r, pmanpr_r

Table 7.12 ANOVA table of the SETI h1 model (Train trip production)

ANOVA[a]						
Model POR		Sum of Squares	df	Mean Square	F	Sig.
Seti model All sector Hypothesis 1	SSR Regression	103.696	15	6.913	15.457	.000b
	SSE Residual	134.175	300	.447		
	SST Total	237.871	315			

a. Dependent Variable: ln_train_wo

b. Predictors: (Constant), edertt1000, p_woret, inc_blu_r, p_worker_man, car_own, ln_jwr, p_wocon, inc_pr_r, ln_avedist, ln_lvr, sqrt_ptiori, ln_strio, pmanpr_r, wod, pblu_r

The coefficient of partial determination of the effective employed resident density variable has been calculated as follows (equation 7.4):

$$R^2_{effetive\ density} = \frac{103.696 - 99.478}{237.871 - 99.478} = \frac{4.218}{138.393} = 0.03048\ or\ 3.05\%$$

The coefficient of partial determination of natural log of proportion of train ridership production with the additional effective employed resident density (while holding all other variables constant) is 3.05%.

Likewise, the contribution of effective job density and the interaction terms in the SETI h2b model has been calculated based on the values presented in table 7.13. The regression sum of square of model with the SETI h2b is 102.354.

$$R^2_{interaction} = \frac{102.354 - 99.478}{237.871 - 99.478} = \frac{2.876}{138.393} = 0.02078\ or\ 2.078\%$$

The coefficient of determination of variable effective job density and the interaction terms is 2.078%. Therefore, the coefficient of partial determination of effective employed resident density in the SETI h1 is higher than that of effective job density and its interaction term in the SETI h2b model. The additions to SETI h1 therefore improve correlation to a greater degree than the SETI h2b for train trip production. This thesis therefore concludes that the SETI h1 model is the best model for prediction of train trip production.

Table 7.13 ANOVA table of the SETI h2b model (Train trip production)

ANOVA[a]						
Model POR		Sum of Squares	Df	Mean Square	F	Sig.
Seti model All sector Hypothesis 2b	SSR Regression	102.354	16	6.397	14.114	.000b
	SSE Residual	135.517	299	.453		
	SST Total	237.871	315			

a. Dependent Variable: ln_train_wo

b. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, inc_pr_r, p_wocon, sqrt_ptiori, ln_jwr, ln_strio, inc_blu_r, ln_lvr, wod, ejdtt1000, pmanpr_r, pblu_r, ln_avedist

Based on the ANOVA table 7.11 to 7.13, the calculation of F_{STAT} is presented in table 7.14.

Table 7.14 The Partial F-Test of the SETI model (effective density variable) in the train trip production model

Model	*Degree of freedom (at α=0.05)*	$F_{0.05}$	F_{STAT}	*Conclusion*
SETI h1	15 and 300	1.66	(103.7 – 99.47) /0.447 = 9.46	Effective employed resident density has significantly improved the LUTI model
SETI h2b	16 and 299	1.66	(102.35 – 99.47) /0.453 = 6.4	Effective job density with the interaction term has significantly improved the LUTI model

In train trip attraction model, table 7.16 for the SETI h1b model shows a total sum of squares of 619.637, a sum of square of regression (SSR) of 341.136 and a regression sum of square of the LUTI model of 325.775 (table 7.15).

Table 7.15 ANOVA table of the LUTI model (Train trip attraction)

ANOVA[a]						
Model POW		Sum of Squares	df	Mean Square	F	Sig.
LUTI model All sector	SSR Regression	325.775	14	23.27	23.835	.000b
	SSE Residual	293.862	301	0.976		
	SST Total	619.637	315			

a. Dependent Variable: ln_ptrainw22

b. Predictors: (Constant), ln_jwr, p_manpr, sqrt_ptiori, inc_bl, ln_avedist, p_jobcon, All_job, p_jobret, sr_lvnr, p_jobman, inc_mp, ln_strio, ln_jobd, p_blu

Table 7.16 ANOVA table of the SETI h1b model (Train trip attraction)

ANOVA[a]						
Model POW		Sum of Squares	df	Mean Square	F	Sig.
Seti model All sector Hypothesis 1b	SSR Regression	341.136	16	21.321	22.89	.000b
	SSE Residual	278.501	299	0.931		
	SST Total	619.637	315			

a. Dependent Variable: ln_ptrainw22

b. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, inc_mp, p_manpr, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, Effective employed resident density of all sector (in 1000 units), ln_jobd, p_blu, ln_avedist

Therefore, the coefficient of partial determination for the effective employed resident density variable and the interaction terms is:

$$R^2_{interaction} = \frac{341.136 - 325.775}{619.637 - 325.775} = 0.05227 \; or \; 5.23\%$$

The coefficient of partial determination of natural log of proportion of train ridership with the additional effective employed resident density and the interaction terms (while holding all other variables constant), is 5.23%.

Likewise, the contribution of effective job density and the interaction terms in SETI h2b model has been calculated based on the value presented in the ANOVA table of 7.17. The regression sum of squares for the SETI h2b model is 338.16.

$$R^2_{interaction} = \frac{338.16 - 325.775}{619.637 - 325.775} = 0.042146 \; or \; 4.2\%$$

The coefficient determination of variable effective job density and the interaction terms is 4.2%. Therefore, the coefficient of partial determination for effective employed resident density and its interaction term is higher than that of effective job density and its interaction term. The SETI h1b model therefore improves correlation of the LUTI model to a greater degree than SETI h2a for train trip attraction. This thesis concludes that SETI h1b is the best model for predicting train trip attraction.

Table 7.17 ANOVA table of the SETI h2b model (Train trip attraction)

ANOVA[a]						
Model		Sum of Squares	Df	Mean Square	F	Sig.
Seti model All sector Hypothesis 2b	SSR Regression	338.16	16	21.135	22.451	.000b
	SSE Residual	281.477	299	0.941		
	SST Total	619.637	315			

a. Dependent Variable: ln_ptrainw22

b. Predictors: (Constant), ejdtt_ln, inc_bl, ln_jobd, p_manpr, p_jobcon, ln_jwr, p_jobret, ln_strio, All_job, p_jobman, inc_mp, sr_lvnr, sqrt_ptiori, p_blu, ln_avedist, Employment effective density of all sector (in 1000 units)

Based on the ANOVA table 7.15 to 7.17, the calculation of the F_{STAT} is performed in table 7.18.

Table 7.18 The Partial F-Test of the SETI model (effective density variable) in the train trip attraction model

Model	*Degree of freedom (at α=0.05)*	$F_{0.05}$	F_{STAT}	*Conclusion*
SETI h1b	16 and 299	1.66	(341.136–325.775) /0.931= 16.5	Effective employed resident density with the interaction term has significantly improved the LUTI model
SETI h2b	16 and 299	1.66	(338.16–325.775) /0.941= 13.16	Effective job density with the interaction term has significantly improved the LUTI model

7.3.2.3 Model validation (back-testing) based on the 2006 dataset

The level of performance of 2011 data in the SETI model may be compared to the LUTI model when both are validated using the observed 2006 data set for the study area. The model validation was calculated based on a simple procedure. First, by calculating the proportion of train trips from the model equation, the proportion of train trip production can be calculated for each observation (suburb) based on all predictor variables *X*s as of 2006. The result of the predicted production is then averaged over all observations and compared to the observed values. Secondly, a residual may be calculated, by taking the differences in the proportion of train ridership between the observed data and the predicted model.

This validation was conducted only for train trip production due to the constraints on data availability. These validation processes could not be conducted for train trip attraction. There was also anunavoidable degree of error associated with the conversion of geographical boundaries for place-of-work 2006 data in the ABS censuses to the 2011 suburb level data set.

The SETI h1 and h2b train trip production models were compared to the LUTI model in the validation process. The model equation that consisted of the unstandardized beta coefficients from all statistically significant predictor variables is shown in the following equations 7.6 – 7.8:

- The LUTI model of train trip production:

$$\widehat{\ln y} = -11.32 - 1.935 pblu_r + 2.284 pworker\_man - 0.049 incprr - 0.612 carown + 1.473 lnstrio - 0.682 lnavedist - 0.263 lnlvr + 0.083 lnjwr \quad \text{Equation 7.6}$$

- The SETI h1 model of train trip production:

$$\widehat{\ln y} = -16.566 + 2.112 pworker\_man - 0.047 incprr - 0.665 carown + 0.00034 sqrt_{ptiori} + 1.914 lnstrio - 0.655 lnavedist - 0.259 lnlvr + 0.083 lnjwr + 0.113 edertt1000 \quad \text{Equation 7.7}$$

- The SETI h2b model of train trip production:

$$\widehat{\ln y} = -13.374 + 1.474pmanprr + 2.116pworker\_man - 0.043incprr - 0.652carown + 1.719lnstrio - 1.086lnavedist - 0.288lnlvr + 0.097lnjwr - 0.044ejdtt1000 + 0.044ejdtt\_lndist \qquad \text{Equation 7.8}$$

All socio-demographic/economic 2006 data were retrieved from the 2006 ABS censuses, then each of the variable measurements were taken as the original database scores. These comprised the proportion of blue collar employed residents (*p_blu_r)*, the proportion of employed residents working in the manufacturing sector (*p_worker_man*), the income per hour of individual employed residents working as professionals/managers (*inc_pr_r*), and car ownership per dwelling (*car_own*).

Property data in terms of land value per square meter of residential land use (*ln_lvr*) was calculated from the 2006 Landgate database. The value of effective density *(ejdtt1000* and *edertt100*0) was recalculated based on 2006 travel time matrices (335x335 suburbs) of the STEM model. The road network travel distance (or centrality of suburbs based on road networks) was assumed to be the same as 2011 data. The public transport network supply (*sqrt_ptiori*) was assumed to be the same as the 2011 data. The average distance of residential suburbs to train stations as of 2006 was re-calculated using the assumption that the Perth-Mandurah railway line did not exist in the system. Therefore, all suburbs that had been associated with a station along the Mandurah line post-extension were associated with stations in Perth city or the Perth Underground station.

One limitation emerged in calculating housing-job balance (or the job to employed resident ratio) (*ln_jwr)* due to the fact that the number of jobs as of 2006 were not available. This thesis therefore referred to the BTRE (Bureau of Transportation and Regional Economics) report number 119 that stated: "*estimates from the land use and employment survey from 2002 to 2008, shown an average annual rate of jobs growth for the Outer sub-region at 7 per cent, with the Middle and Inner sub-regions at 3 and 2 per cent respectively" (Bureau of Transport and Regional Economics [BTRE], 2007, p. 232).* It was assumed that all suburbs within 0-5 km (or the "inner area") followed a 2 percent annual growth of jobs, all suburbs within 5-10 km (or the "middle area") followed a 3 percent growth, and the remaining suburbs followed a 7 per cent growth rate. The number of job as of 2006 was calculated based on these assumptions using the procedure mentioned in Xinhao and Rainervom, 2007 p. 38:

$$g_p = \left(\frac{Job_{t+n}}{Job_t}\right)^{\frac{1}{n}} - 1 \qquad \text{Equation 7.9}$$

Explained as:

g_p = represented the known average annual percent change of job in each area;

Job_t = represented the unknown number of job in 2006;

Job_{t+n} = represented the known number of job in 2011; and

n = the year differences

The model validation results are shown in table 7.19.

Table 7.19 The summary of model validation from the predicted model as of 2011 data on the observed data as of 2006

Observation	***Observed portion of train ridership***	***Predicted proportion of train ridership***			***Residuals (observed values minus predicted values)***		
		LUTI	**SETI h1**	**SETI h2b**	**LUTI**	**SETI h1**	**SETI h2b**
1	0.02410	0.02301	0.0402	0.034	0.00110	-0.0161	-0.0099
2	0.00572	0.01423	0.0129	0.013	-0.00851	-0.0071	-0.0077
3	0.00746	0.01352	0.0205	0.017	-0.00606	-0.0130	-0.0094
4	0.00357	0.01014	0.0090	0.010	-0.00656	-0.0054	-0.0064
5	0.00225	0.01408	0.0108	0.014	-0.01183	-0.0085	-0.0114
….							
297	0.01884	0.02549	0.0318	0.031	-0.00664	-0.0129	-0.0117
Average	0.04210	0.03369	0.0475	0.0477	0.00841	-0.0054	-0.0056

The closer the value of the predicted values to the observed values, the closer the average residual values will be to zero, and thus the better the model prediction. Table 7.19 shows that average predicted values from SETI h1 and h2b were closer to the average observed data than for LUTI, while the averages of the residual values from SETI h1 were also the smallest. The LUTI model prediction on average was not as accurate as other models. The positive values of the LUTI residuals indicate that LUTI tends to underestimate the observed values. Both of the SETI models produced negative average residual values indicating that the SETI models tend to overestimate the observed values, although to a lesser degree that the LUTI model. The SETI h1 model produced residuals equalling -0.0054 on average, while the LUTI produced residuals equalling 0.00841 on average. This indicates that the errors of SETI h1 represent 64.2% of the errors for the LUTI model on average.

7. 4 CHAPTER SUMMARY

The geographical extent of agglomeration has shown some interesting results such as:

(i) There was a geographical variability in the influence of density in the LUTI model, and of effective density in the SETI model, on the multiplicative effects on train ridership. This variability has implications for policy options, especially for policies aimed at capturing the maximum benefit of effective density on train ridership.

(ii) Effective job density clearly influenced train trip production, varying with distance from the train station. Increasing effective job density in residential suburbs near a station produced smaller multiplicative effects on train ridership than applying that same amount to residential suburbs at farther distances.

(iii) The LUTI model showed that the interaction term between employed resident density and distance of residential suburb from the train station does not have a statistically significant effect on train ridership. Employed resident density influences ridership at the same level (multiplicative =1) in all area (represents no influence at all), independent of the location of residential suburb from station.

(iv) The effect of effective job and employed resident density on train ridership was stronger for workplace suburbs near the station than for those farther away. Workplace suburbs would be better located at a shorter distance from the train station for the ability of suburbs to accumulate the benefits of effective job and employed resident density on train ridership. Jobs would be better added to residential suburbs once these suburbs were located near the train station. Main workplace areas that are underserved by stations may require sufficient alternatives for sustainable mobility, such as by adding new station. Park-and-ride facilities may be seen as necessary for providing service to residential suburbs located farther from train station.

(v) The effective density variable was more useful than job density for informing future policy decisions. The influence of effective density may be able to assess the extent to which maximum agglomeration benefits can be achieved or where no agglomeration benefit is apparent. Inclusion the interaction term to density variable does not provide this information as the results are not plausible empirically.

(vi) There appears to be a spatial competition between residential land use and non-residential land use for transit attraction. If an area directly adjacent to the station precinct is developed for jobs, it will attract train ridership, while residential densification in this area also implies a lower chance to develop jobs or non-residential land uses, and therefore the effect on transit use may be counter productive. This spatial competition is represented as the ratio of job to employed resident or job-housing balance. Discussion on the influence of effective density on train ridership from the job-housing balance perspective is discussed in chapter 8.

In terms of comparing the overall model fits between the model without agglomeration (effective density) or the LUTI model to the model with agglomeration or the SETI model, it was found that:

(vii) The SETI model outperformed other models in both train trip production and attraction model. The SETI model improved the prediction capability of the LUTI model by 2.87 and 4.5% for train trip production and by 3.8 and 4.75% for train trip attraction.

(viii) In the train trip production model, the employed effective density variable contributed to 3.05% of variances, while effective job density and its interaction terms contributed to 2.078%. This was significantly more than in the LUTI model. In train trip attraction model, the effective employed resident density and its interaction term contributed to 5.23% of variances, while effective job density and interaction term added 4.2%, which were also higher than in LUTI.

(ix) Model validation showed that the accuracy of the SETI model for train ridership prediction was greater than the LUTI model. The range of errors in the SETI model on average was lower than that of the LUTI model.

This thesis has also showed that effective density can be measured based on different sector of economy. The measurement of effective density based on travel time by park and ride for manufacturing sector has revealed that effective employed resident density contributes at significant degree in train ridership attraction; however, it emerged over larger geographical area than the other two sectors. The effective construction job density has contributed on almost similar level to manufacturing sector but over smaller geographical area than the other two sectors. The effective retail job density, on the other hand, contributes on much smaller multiplicative effect of train trip attraction over much smaller geographic area than the other two sectors.

In train ridership production model, the construction sector has the largest multiplicative effect of effective job density than the two other sectors. However, without considering the effect of distance of suburb from train station on the influence of effective density, the retail sector has the largest multiplicative effect of effective employed resident density than the other two sectors.

More details on model interpretations, policy implications, and the limitations of the model results will be discussed in chapter 8.

CHAPTER 8. POLICY IMPLICATIONS OF TRAIN RIDERSHIP PREDICTION

This chapter investigates the implications of the model for land use policies and for train ridership. The first section discusses the implications of the model findings for land use policy from the point of view of the importance of the geographical extent of agglomeration (hypothesis 3). The relationship between effective density and train ridership is evaluated based on the fundamental concept underlying the research hypothesis, namely the trade-off between wages/land rents and transport costs, as stated in Venables (2004) and Graham (2007). The second section of this chapter investigates the job-housing balance (hypothesis 4) from the perspective of these trade-offs and its implications for train ridership.

This thesis is part of the T-LU approach, as opposed to the LU-T approach, as mentioned in chapter 2 (section 2.6 literature gap). The thesis aim is to understand the impact of the transport system on land use (measured in terms of effective density) in the form of the job-based TOD scenario with train ridership as an outcome.

8. 1 IMPLICATIONS OF THE GEOGRAPHICAL EXTENT OF AGGLOMERATION ON TOD POLICY: EFFECTS OF ADDING MORE JOBS

Integrated land use-transport policy and planning is initiated from the position of deciding on desired future land use and then using transport as a part of the system to realize the future scenario. This thesis assumed that the future land use scenario involving the development of a job-intensive TOD scenario around the station precincts would be followed by an increase in rail use.

Specifically, this section discusses the application of model results in relation to the third research hypothesis. The third research hypothesis specifically mentions the geographical extent of agglomeration, i.e. how the distance of suburb from stations (in terms of place of residence or place of work) influences the strength of relationship of effective density on train ridership. This section aims to assess the application of this hypothesis to the effectiveness of TOD policies based on the scenario of developing a job-based TOD scenario.

TOD, according to Curtis (2008) is one approach within the LUTI mainstream for development, which is “based on the notion of precincts designed for high density residential development combined with high intensity commercial and retail development sitting within a high quality pedestrian environment” (Curtis, 2008, p. 287). Meanwhile, Botte and Olaru, 2011, define an operational definition of TOD as being “associated with moderate to high density development, located within an easy walk of approximately 800 meters of a major public transport stop (operating on highly synchronized and reliable timetables, at high frequency (5-15 min) and with extended operating hours), generally with a mix of residential, employment, and shopping opportunities, designed for pedestrians and cyclists, without excluding the automobile” (Botte & Olaru, 2011, p. 3).

There was a new focus on public transport that began with the revitalization of suburban rail lines in the 1980s in-line with TOD policy being adopted in Western Australian state planning from 1988 (Mees et al., 2008) with 68 station precincts in the Perth metropolitan are being assigned as TOD. However, the TOD concept has been subsequently challenged over the way its policies are translated into urban development implementation (Curtis & Mellor, 2011).

In the specific context of the Perth metropolitan region, while governments in most localities supported TOD development and incorporated this into land use and transport policies, there is still a huge gap between the stated policies and their implementation (Falconer et al., 2010; Curtis, 2012). Curtis (2012) stated that the gap between policies and practices occurred due to slow responses in land use changes and a slow response from business sectors toward these policies. Meanwhile, the efforts to conduct development by local government initiatives according to TOD principles were also constrained by the slow and inconsistent policy translation from the state to the local governments (Curtis and Mellor, 2011).

TOD precincts in Perth did not fulfil the density criteria outlined by state TOD policy in 2005 (i.e. 25 dwelling per ha or above) or even by some studies (i.e. 10,000 employees or residents in a TOD station precinct) (Newman and Kenworthy, 2006, in Curtis 2012). In practice, there is only one station precinct in the Perth metropolitan region with a density of 15 dwellings per ha, while 84% of the 68 TOD precincts have a density of less than 10 dwellings per ha, and only 8 out of 68 TOD precincts fall in line with the benchmark proposed by Newman and Kenworthy (Curtis, 2012). Similarly, Botte and Olaru, 2011, stated that in the Perth context, the significant challenge for TOD is due to the average population density being only 3.08 persons/ha, whereas the highest density TOD contains less than 40 dwelling/ha, less than a third of the recommended density to support the high capacity and rapid transit for TOD, based on the literature.

An insufficient population density would be regarded as a high risk from a business perspective and TOD development may be expected to gain low support from business sectors, making it more difficult to embrace its growth and to attract more mixed-use developments, employment, and residents (Curtis and Mellor, 2011). As a result, there is still relatively high car usage and a lower attraction for public transport use for residents living in TOD precincts (Falconer et al, 2010).

Understanding how to apply the results of this thesis, in relation to transit for TODs in the study area, may help to overcome these challenges. In particular, the Perth metropolitan region faces the problem of low density, which remains an obstacle to its success as a TOD to support higher train ridership levels.

In this section, the model results of the geographical extent of agglomeration (section 7.2) are investigated. From the model results, where regression involved the interaction terms, it has been shown that the level of train ridership is a function of both the changes in the effective density and the distance of suburbs from train stations. The implication of the densification/intensification of jobs within the scenario of job-based TOD, including its impacts on train use, would be assessed in this section. The implications for policy analysis of the modelled results have been investigated for each of the retail, manufacture and construction sectors, based on the train trip attraction results.

The *LDV* regression model in section 7.2 expressed the geographical extent of effective job density for each sector within which train ridership is able to capture the effects of agglomeration. In all sectors being studied, the geographical extent of effective job density has been consistently lower for the train trip attraction than train trip production, and the geographical extent of effective job density has been consistently smaller than the effective employed resident density in train trip attraction.

Center for Transit-Oriented Development has critized the emphasised on residential development of TOD as planning policies. “Connecting destinations to create ridership may seem like an obvious conclusion, but plans and policies have not reflected this approach. Most TOD policy have focused on residential development, rather than promoting agglomeration of jobs and commercial space in regional centers served by transit.” (The Centre for Transit-Oriented Development, 2009, p. 28). In addition, employment density is strongly associated with transit ridership more than residential density (Kolko, 2011) and that employment density in workplace location has been much more important than residential density in residential location (Arrington & Cervero, 2008). The TCRP’s study also found that the residential based TOD areas contributed to lower train ridership than the office and retail-based TOD areas (Evans et al., 2007). Similarly, rail commuting has been found to be stronger at stations that serve employment centres than those that do not (Blainey, 2010). Therefore, adding jobs to the station catchment or allocating stations to cater for employment centres is an example of an effective policy to increase train ridership.

From the model results in section 7.2, suburbs located within the threshold distance with low employment activities may develop higher levels of employment in order to achieve the formation of a job-based TOD. One option may be to increase effective job density by reaching the minimum ratio of 1.5 (refer to job-housing balance criteria, section 4.1.2.6 in chapter 4). Thus, one application of the train trip attraction model would be to add jobs in the suburbs which are currently low in employment activities but are located within the geographical extent of agglomeration. The expected outcomes following this policy would be to increase train trip attraction. This exercise required a revised calculation of the effective density variable, since the addition of new jobs to some suburbs has the effect of changing the effective density magnitude for all suburbs in the system.

In the construction sector, there are 90 suburbs that are located inside the threshold distance of 3.184 km of effective construction job density; in manufacturing, there are 128 suburbs within the threshold of 4.1 km; and for retail, there are 47 suburbs within 2.4062 km. Jobs were added to these suburbs until they reached the criteria for job-rich suburbs (at least 1.5 jobs is available for every 1 employed resident) to examine the effects on modelled ridership for these suburbs.

The effect on the modelling is as follows. First, the addition of new jobs has an impact on the value of effective density. Second, the increase or the delta of the effective job density from adding new jobs is calculated. The delta value is used to exponentiate the value of the multipliers (the coefficient of the model) to generate the new multiplicative effect. Referring to equation 7.3 and the model results from table 7.6 in chapter 7, the formula for calculating the multiplicative effects of effective job density on the proportion of train trip attraction upon adding new jobs is stated in equation 8.1 (an example for the retail sector only is carried out in equation 8.2). This equation has been modified to accommodate the impacts of additional effective job density accumulated due to the addition of new jobs, as follows (after Taplin, 2016):

$$\frac{Y(x_1+\Delta_1,x_2)}{Y(x_1,x_2)} = \exp(\beta_1 + \beta_3 x_2)^{\Delta_1} = (\exp(\beta_1)\exp(\beta_3)^{x_2})^{\Delta_1} \quad \text{Equation 8.1}$$

Where:

x_1= effective job density.

x_2= (natural log of) distance of suburbs from stations.

Δ_1= an increase in x_1.

$Y(x_1 + \Delta_1, x_2)$= the proportion of train trip attraction in suburbs after adding effective density as much as Δ_1, when suburbs are located in distance x_2 from stations.

$Y(x_1, x_2)$= the initial proportion of train trip attraction in suburbs.

$\exp(\beta_1 + \beta_3 x_2)^{\Delta_1}$ replaces the term $\exp(\beta_1)$= the coefficient regression of x_1, i.e. the effect of x_1 on Y when no interaction term exists.

$\exp(\beta_3)$= the coefficient of regression of the interaction terms, i.e. relating to the effect of x_1 on Y depends on x_2.

An example of equation 8.1 applied for the retail sector is as follows:

$$\frac{Y(x_1 + \Delta_1, x_2)}{Y(x_1, x_2)} = \exp(0.072 + -0.082x_2)^{\Delta_1} = (\exp(0.072)\exp(-0.082)^{x_2})^{\Delta_1}$$

$$=(1.075 * 0.921^{x_2})^{\Delta_1} \quad \text{Equation 8.2}$$

Where:

0.072 = the unstandardized beta coefficient of the effective job density variable in the retail sector (see table 7.5 in chapter 7).

-0.082 = the unstandardized beta coefficient of the interaction term between the effective job density in the retail sector and the distance to stations (see table 7.5 in chapter 7).

1.075 = the exponential of the unstandardized beta coefficient of the effective job density variable in the retail sector on the proportion of train trip attraction.

0.921 = the exponential of the unstandardized beta coefficient of the interaction terms

$x_2 = ln\_avedist$ = the natural log of the distance of workplace suburbs from the nearest train station.

$\Delta_1 = an\ increase\ in\ the\ effective\ job\ density$ = calculated as the new value of effective density after adding new jobs substracted by the old value of effective density as stated in Graham's formula in equation 5.1.

The component in the bracket, i.e. $(1.075 * 0.921^{x_2})$ in equation 8.2 refers to the value of the multiplicative effect for every one unit increase in effective job density in the retail sector on an increase in train ridership, all else being equal. In summary, this equation states that the multiplicative effect of adding effective density is equal to the value of multiplicative effects based on adding one unit of effective density exponentiated by the increase in effective job density. For example, the suburb of Applecross in table 8.3 is located within 2.22 km from the nearest train station ($x_2 = \ln(2.22)$). Current or initial retail job numbers equate to 232 jobs. In order to reach a job-housing balance ratio of 1.5 for Applecross, 181 new retail jobs are needed. These additional jobs would increase the effective job density where the increase in effective job density is 228.8 or 2.288 units (1 unit equates to an effective job density of 100). The multiplicative effect of adding 2.288 units of effective retail job density (due to adding 181 new retail jobs) is then calculated as follows:

$$(1.075 * 0.921^{x_2})^{\Delta_1} = (1.075 * 0.921^{(\ln 2.2)})^{2.288} = (1.075 * 0.9364)^{2.288}$$

$$= 1.01525 \qquad \text{Equation 8.3}$$

Therefore, the multiplicative effect of adding an effective job density of 228.8 to the retail sector (due to adding 181 new retail jobs) on the proportion of train trip attraction is a factor of 1.015, i.e. an increase of 1.5% (more complete results are displayed in table 8.3 for the retail sector).

The same method of calculation has been used for other sectors. For the construction sector, the results are shown in table 8.1, and for the manufacturing sector, table 8.2.

Table 8.1 The multiplicative effects of adding new construction jobs for the job-based TOD scenario

No.	Suburbs	Distance < 3.184 km	The number of construction Job	Add new Construction Jobs	New EJD Construction	Delta EJD Con	Delta EJD Con (in 100 units)	Multiplicative effects based on one unit increase in EJD	New multiplicative effects
1.	Applecross	2.2222	171.00	214.50	863.73	379.87	3.80	1.0513	1.2092
2.	Armadale (WA)	2.2723	124.51	682.49	955.36	639.25	6.39	1.0480	1.3497
3.	Ashfield (WA)	1.6536	9.96	71.04	652.97	319.02	3.19	1.0954	1.3372
4.	Atwell	2.2432	23.12	692.38	1116.23	802.70	8.03	1.0499	1.4784
5.	Bassendean (WA)	1.6188	248.08	377.42	919.32	428.52	4.29	1.0986	1.4963
6.	Bateman	2.6427	26.00	169.00	771.62	392.59	3.93	1.0263	1.1071
7.	Bayswater (WA)	2.1437	803.42	17.08	1037.99	209.66	2.10	1.0565	1.1222
8.	Beckenham	1.92	81.68	437.32	933.62	547.21	5.47	1.0729	1.4694
9.	Bedford	2.136	43.55	293.95	942.89	536.28	5.36	1.0571	1.3467
10.	Beldon	2.5931	52.00	467.00	1096.33	753.06	7.53	1.0290	1.2399
11.	Bentley (WA)	2.3465	197.31	80.19	827.08	291.23	2.91	1.0434	1.1316
12.	Brentwood (WA)	2.3787	19.50	82.50	757.51	349.05	3.49	1.0414	1.1520
13.	Bull Creek (WA)	1.7516	56.00	308.50	890.14	463.77	4.64	1.0866	1.4701
14.	Camillo	2.3959	16.50	289.50	770.98	509.07	5.09	1.0403	1.2230
15.	Cannington	1.7192	221.00	35.50	825.79	263.98	2.64	1.0895	1.2538
16.	Carine	2.9579	78.99	402.51	993.22	562.48	5.62	1.0103	1.0594
17.	Carlisle	1.2898	157.01	208.99	1005.67	456.23	4.56	1.1339	1.7738
18.	Claremont (WA)	1.4992	148.00	96.50	715.83	258.45	2.58	1.1104	1.3108
19.	Connolly	2.7775	52.00	312.50	895.10	542.49	5.42	1.0192	1.1086
20.	Coolbinia	2.803	14.08	68.42	841.99	378.98	3.79	1.0179	1.0695
21.	Cottesloe	1.7159	132.00	192.00	749.36	335.22	3.35	1.0898	1.3339
22.	Craigie (WA)	2.1538	58.00	611.00	1063.25	701.95	7.02	1.0559	1.4645
23.	Crawley	3.1051	27.52	5.48	494.56	160.11	1.60	1.0035	1.0056
24.	Daglish	1.2295	38.35	30.65	683.99	247.34	2.47	1.1414	1.3871
25.	Dalkeith	3.044	38.64	85.86	566.20	240.46	2.40	1.0063	1.0152
26.	Duncraig	3.0559	186.00	1044.00	1340.22	888.95	8.89	1.0057	1.0522
27.	East Cannington	1.9572	97.81	259.19	928.90	491.90	4.92	1.0700	1.3949
28.	East Fremantle	2.7215	132.00	238.50	770.73	377.40	3.77	1.0221	1.0859
29.	East Victoria Park	1.8402	190.87	386.63	1137.29	580.22	5.80	1.0792	1.5563
30.	Eden Hill	1.7988	31.83	229.17	825.94	454.36	4.54	1.0826	1.4344
31.	Edgewater	2.0613	19.58	397.42	903.54	605.63	6.06	1.0623	1.4422
32.	Embleton	2.4198	79.43	138.07	859.30	391.16	3.91	1.0389	1.1610
33.	Ferndale (WA)	2.712	33.40	224.60	763.57	425.49	4.25	1.0226	1.0996
34.	Floreat	3.0418	76.00	245.00	856.02	411.03	4.11	1.0064	1.0265
35.	Fremantle	3.1819	222.51	227.49	774.26	323.00	3.23	1.0001	1.0004
36.	Glendalough	1.9066	12.57	105.93	983.38	475.43	4.75	1.0739	1.4035
37.	Gosnells	3.067	127.00	1151.00	1028.48	699.10	6.99	1.0052	1.0372
38.	Greenwood (WA)	2.8697	235.00	683.00	1225.36	728.03	7.28	1.0146	1.1111
39.	Guildford (WA)	1.4108	60.31	34.19	537.66	190.11	1.90	1.1198	1.2400
40.	Heathridge	2.2095	89.00	736.00	1200.74	852.44	8.52	1.0521	1.5420
41.	Highgate (WA)	1.1953	10.72	83.78	887.60	456.91	4.57	1.1459	1.8632
42.	Inglewood (WA)	2.1942	54.70	260.30	882.71	466.92	4.67	1.0531	1.2735
43.	Joondalup	3.0572	480.97	363.53	1002.66	428.95	4.29	1.0057	1.0246

44.	Joondanna	2.6462	30.92	257.08	1037.63	593.03	5.93	1.0261	1.1649
45.	Karawara	2.3618	14.44	45.56	690.14	291.21	2.91	1.0424	1.1286
46.	Kelmscott	2.6526	220.00	485.00	741.99	396.36	3.96	1.0257	1.1059
47.	Kensington (WA)	2.3136	72.00	184.50	883.91	417.38	4.17	1.0454	1.2036
48.	Kenwick	2.6649	140.69	226.81	669.45	304.60	3.05	1.0251	1.0783
49.	Kiara	2.766	11.28	69.72	638.57	288.85	2.89	1.0198	1.0582
50.	Kinross	3.0197	82.00	651.50	1192.37	857.34	8.57	1.0074	1.0653
51.	Langford	2.4319	28.00	254.00	756.98	443.85	4.44	1.0382	1.1810
52.	Lathlain	1.3894	44.00	203.50	961.85	500.01	5.00	1.1222	1.7797
53.	Leederville	1.3737	98.00	104.50	946.64	382.09	3.82	1.1240	1.5629
54.	Leeming	2.9843	123.00	589.50	1028.14	587.47	5.87	1.0091	1.0544
55.	Lockridge	2.601	11.89	199.61	758.31	451.28	4.51	1.0285	1.1354
56.	Lynwood (WA)	2.838	18.60	168.90	743.15	407.13	4.07	1.0161	1.0673
57.	Manning	2.8852	28.48	202.52	886.58	470.86	4.71	1.0138	1.0667
58.	Maylands (WA)	2.0831	181.00	479.00	1104.17	586.66	5.87	1.0608	1.4135
59.	Menora	3.1337	20.92	67.58	778.81	337.48	3.37	1.0022	1.0076
60.	Midland	2.3136	211.10	67.90	666.37	209.32	2.09	1.0454	1.0974
61.	Mosman Park (WA)	1.9005	87.00	280.50	739.37	387.89	3.88	1.0744	1.3209
62.	Mount Claremont	2.3101	41.01	116.49	650.55	276.98	2.77	1.0456	1.1315
63.	Mount Hawthorn	1.8942	119.00	388.00	1276.79	683.42	6.83	1.0749	1.6380
64.	Mount Lawley	1.5865	104.27	441.73	1145.90	623.73	6.24	1.1017	1.8295
65.	Mount Nasura	2.0679	45.00	232.50	738.77	443.91	4.44	1.0618	1.3053
66.	Mount Pleasant (WA)	3.0139	114.39	292.11	977.25	496.63	4.97	1.0077	1.0387
67.	North Fremantle	2.3536	95.00	14.50	527.13	160.57	1.61	1.0429	1.0698
68.	North Lake	2.8949	14.36	95.14	631.96	309.68	3.10	1.0133	1.0419
69.	North Perth	2.2107	83.00	419.50	1190.03	666.31	6.66	1.0520	1.4022
70.	Padbury	2.8003	133.00	791.00	1197.62	785.82	7.86	1.0180	1.1507
71.	Parmelia	2.2581	34.64	386.86	760.90	467.41	4.67	1.0489	1.2502
72.	Queens Park (WA)	1.4585	96.19	203.81	915.51	449.96	4.50	1.1146	1.6297
73.	Rivervale	1.8713	276.75	266.25	1107.23	468.20	4.68	1.0767	1.4134
74.	Rossmoyne	2.7395	31.37	109.63	755.78	352.02	3.52	1.0211	1.0764
75.	Seville Grove	2.9966	66.00	687.00	1070.82	764.81	7.65	1.0085	1.0668
76.	Shenton Park	1.4905	97.07	19.93	633.59	195.81	1.96	1.1113	1.2295
77.	South Fremantle	2.3925	52.96	130.04	664.85	326.63	3.27	1.0405	1.1386
78.	South Guildford	2.5444	160.84	77.66	606.71	209.22	2.09	1.0317	1.0674
79.	St James (WA)	1.5844	37.73	190.27	882.59	466.22	4.66	1.1019	1.5720
80.	Stirling (WA)	2.5418	257.96	618.04	1395.86	758.49	7.58	1.0318	1.2683
81.	Swanbourne	2.3087	67.88	35.62	547.25	190.73	1.91	1.0457	1.0890
82.	Thornlie	2.9413	184.00	1460.00	1303.78	939.27	9.39	1.0111	1.1093
83.	Viveash	2.0809	11.07	27.93	473.78	186.35	1.86	1.0609	1.1165
84.	Warwick (WA)	2.5922	37.00	272.00	930.28	515.49	5.15	1.0290	1.1589
85.	Wembley	2.3809	145.67	395.83	1107.63	558.51	5.59	1.0412	1.2533
86.	West Leederville	1.22	58.75	166.25	990.60	484.71	4.85	1.1427	1.9087
87.	White Gum Valley	3.1365	44.36	152.14	721.45	388.39	3.88	1.0021	1.0082
88.	Winthrop	2.7806	32.00	274.00	813.30	438.13	4.38	1.0190	1.0861
89.	Woodbridge (WA)	1.1639	32.66	51.34	539.46	221.61	2.22	1.1502	1.3635
90.	Woodvale (WA)	2.8889	97.00	707.00	1023.25	666.55	6.67	1.0136	1.0944

Overall, the job-based TOD scenario provides further implications for increasing train use in the future. The model calculates new values for multiplicative effects from this analysis, with averages of 1.0617 (an increase of 6.2% in the proportion of train trip attraction) for adding jobs in the retail sector, 1.2516 (an increase of 25.16%) for the manufacturing sector, and 1.2518 (an increase of 25.18%) for the construction sector. Thus, by adding new jobs to acquire the criteria of a job-housing balance ratio of 1.5 in the job-based TOD scenario, the current proportion of train trip attraction increases on average by 6.172% for the effects of retail effective job density, 25.16% for the effects of manufacturing effective job density, and 25.18% for the effects of construction effective job density. Furthermore, the highest multiplicative effects from each of the three sectors emerged in the suburbs of Highgate (WA), with a 26.8% in the increase in the proportion of train trip attraction from the effects of retail effective job density, Queens Park (WA) with an 89% increase due to the effects of manufacturing effective job density, and West Leederville with a 90.9% increase from the effects of construction effective job density.

Table 8.2 The multiplicative effects of adding new manufacturing jobs for the job-based TOD scenario

	Suburbs	***Distance < 4.1 km***	***The number of manufacturing Job***	***Add new manufacturing jobs***	***New EJD manufacturing***	***Delta EJD man***	***Delta EJDman (in 100 units)***	***Multiplicative effects based on one unit increase in EJD***	***New multiplicative effects***
1.	Applecross	2.2222	69.00	165.50	1417.66	639.67	6.397	1.0395	1.2808
2.	Ardross	3.5474	41.00	124.50	1458.53	653.66	6.537	1.0093	1.0622
3.	Armadale (WA)	2.2723	114.16	780.34	1663.69	1003.66	10.037	1.0380	1.4539
4.	Ascot (WA)	3.1377	129.02	1.98	1402.79	499.16	4.992	1.0171	1.0883
5.	Ashby (WA)	3.9863	5.76	138.74	1151.83	623.06	6.231	1.0019	1.0118
6.	Ashfield (WA)	1.6536	6.26	69.24	1368.12	584.08	5.841	1.0590	1.3976
7.	Atwell	2.2432	7.27	659.23	1757.41	1068.22	10.682	1.0388	1.5023
8.	Bateman	2.6427	8.00	184.50	1530.73	745.75	7.458	1.0282	1.2302
9.	Beaconsfield (WA)	3.9837	101.00	171.00	1391.62	599.87	5.999	1.0019	1.0116
10.	Beckenham	1.92	88.40	455.10	1732.34	912.70	9.127	1.0491	1.5484
11.	Bedford	2.136	14.52	230.48	1603.80	758.04	7.580	1.0420	1.3664
12.	Beldon	2.5931	10.00	244.00	1374.29	766.99	7.670	1.0294	1.2488
13.	Brentwood (WA)	2.3787	7.39	87.61	1527.70	708.46	7.085	1.0350	1.2760
14.	Brookdale (WA)	3.3585	0.86	221.64	1418.17	844.04	8.440	1.0128	1.1129
15.	Bull Creek (WA)	1.7516	45.00	350.00	1716.17	855.39	8.554	1.0552	1.5829
16.	Burns Beach	3.9423	0.00	68.00	1038.10	496.03	4.960	1.0026	1.0129
17.	Camillo	2.3959	3.81	359.69	1493.46	887.04	8.870	1.0345	1.3514
18.	Cannington	1.7192	258.00	57.50	1714.34	659.75	6.597	1.0564	1.4361
19.	Carine	2.9579	7.99	306.01	1546.29	807.62	8.076	1.0209	1.1818

20.	Carlisle	1.2898	187.70	111.30	1734.85	698.14	6.981	1.0757	1.6643
21.	Caversham	3.457	131.99	89.01	1137.92	420.76	4.208	1.0109	1.0467
22.	Champion Lakes	3.3713	8.19	53.81	1243.27	553.63	5.536	1.0125	1.0713
23.	Churchlands	3.6697	10.63	82.87	1413.91	642.39	6.424	1.0071	1.0466
24.	Claremont (WA)	1.4992	83.00	76.50	1309.86	518.29	5.183	1.0655	1.3897
25.	Cloverdale	3.7197	33.21	480.29	1811.28	963.28	9.633	1.0063	1.0620
26.	Como (WA)	3.8018	33.00	551.00	1768.23	932.39	9.324	1.0049	1.0464
27.	Connolly	2.7775	10.00	209.50	1340.77	726.46	7.265	1.0249	1.1960
28.	Coolbinia	2.803	4.02	45.98	1564.46	680.03	6.800	1.0244	1.1778
29.	Cottesloe	1.7159	49.00	178.00	1306.73	583.82	5.838	1.0565	1.3785
30.	Craigie (WA)	2.1538	9.00	368.00	1467.79	810.67	8.107	1.0415	1.3905
31.	Crawley	3.1051	28.40	8.10	1113.04	420.16	4.202	1.0178	1.0768
32.	Currambine	3.2319	22.00	415.00	1450.82	866.77	8.668	1.0152	1.1398
33.	Daglish	1.2295	1.80	49.70	1326.98	577.72	5.777	1.0789	1.5511
34.	Dalkeith	3.044	18.48	82.52	1170.38	495.65	4.956	1.0190	1.0980
35.	Duncraig	3.0559	66.00	611.00	1643.25	921.99	9.220	1.0188	1.1873
36.	East Cannington	1.9572	13.61	355.89	1728.53	948.15	9.482	1.0478	1.5570
37.	East Fremantle	2.7215	31.00	236.50	1303.90	615.74	6.157	1.0263	1.1731
38.	East Victoria Park	1.8402	178.98	258.02	1830.84	815.87	8.159	1.0519	1.5108
39.	Eden Hill	1.7988	6.47	237.03	1566.18	741.81	7.418	1.0534	1.4708
40.	Edgewater	2.0613	2.61	245.39	1318.23	725.30	7.253	1.0444	1.3703
41.	Embleton	2.4198	70.55	100.95	1553.10	629.64	6.296	1.0339	1.2335
42.	Ferndale (WA)	2.712	5.14	280.36	1577.27	839.43	8.394	1.0265	1.2454
43.	Floreat	3.0418	26.00	141.00	1412.56	622.85	6.228	1.0191	1.1250
44.	Glendalough	1.9066	6.28	111.22	1737.38	831.09	8.311	1.0495	1.4945
45.	Gosnells	3.067	32.00	1537.50	1856.11	1182.82	11.828	1.0186	1.2431
46.	Greenwood (WA)	2.8697	63.00	515.00	1610.63	902.48	9.025	1.0228	1.2261
47.	Guildford (WA)	1.4108	45.23	25.77	1190.53	431.14	4.311	1.0696	1.3367
48.	Gwelup	3.7223	14.64	152.36	1447.35	691.51	6.915	1.0062	1.0438
49.	Hamersley	3.8619	9.00	237.50	1507.00	777.90	7.779	1.0039	1.0306
50.	Heathridge	2.2095	11.00	433.50	1432.49	848.02	8.480	1.0398	1.3926
51.	Highgate (WA)	1.1953	15.09	58.91	1621.83	719.50	7.195	1.0809	1.7498
52.	Hillman	3.7598	5.66	176.34	1215.79	636.50	6.365	1.0056	1.0361
53.	Hocking	3.9771	7.81	403.69	1607.51	971.05	9.710	1.0020	1.0199
54.	Huntingdale (WA)	3.8941	22.00	676.00	1697.95	1034.03	10.340	1.0034	1.0353
55.	Iluka (WA)	4.0504	10.00	239.50	1228.25	685.61	6.856	1.0009	1.0060
56.	Inglewood (WA)	2.1942	35.48	254.52	1615.32	769.36	7.694	1.0403	1.3551
57.	Innaloo	3.6884	29.00	316.50	1763.29	918.12	9.181	1.0068	1.0642
58.	Jolimont	1.4659	27.23	12.27	1298.15	496.73	4.967	1.0671	1.3804
59.	Joondalup	3.0572	253.53	285.47	1470.97	700.35	7.003	1.0188	1.1391
60.	Joondanna	2.6462	2.58	236.42	1667.09	877.33	8.773	1.0281	1.2750
61.	Karawara	2.3618	4.81	60.19	1529.84	686.49	6.865	1.0355	1.2703
62.	Kardinya	3.8771	90.88	478.12	1754.99	909.60	9.096	1.0036	1.0336
63.	Kelmscott	2.6526	416.00	286.50	1392.68	596.86	5.969	1.0279	1.1787
64.	Kensington (WA)	2.3136	12.00	167.00	1623.89	755.45	7.554	1.0368	1.3141
65.	Kenwick	2.6649	154.23	248.27	1402.03	626.07	6.261	1.0276	1.1860
66.	Kiara	2.766	2.35	146.65	1499.30	682.07	6.821	1.0252	1.1851
67.	Kingsley	3.2559	50.00	631.50	1631.71	931.54	9.315	1.0147	1.1460
68.	Kinross	3.0197	23.00	408.00	1423.37	858.91	8.589	1.0196	1.1811
69.	Langford	2.4319	16.00	445.00	1663.37	962.08	9.621	1.0336	1.3739
70.	Lathlain	1.3894	19.00	169.00	1646.97	783.55	7.835	1.0707	1.7075
71.	Leda	3.4007	6.00	306.50	1195.08	588.60	5.886	1.0120	1.0725
72.	Leederville	1.3737	45.00	59.00	1539.74	638.95	6.389	1.0714	1.5540
73.	Leeming	2.9843	49.00	643.00	1790.69	965.02	9.650	1.0203	1.2143
74.	Lockridge	2.601	2.48	314.52	1599.15	876.33	8.763	1.0292	1.2868

75.	Lynwood (WA)	2.838	2.86	239.14	1557.52	836.66	8.367	1.0236	1.2151
76.	Manning	2.8852	8.34	121.16	1556.52	721.44	7.214	1.0225	1.1741
77.	Marmion	3.9227	7.30	65.20	1267.65	605.25	6.053	1.0029	1.0177
78.	Maylands (WA)	2.0831	77.00	664.50	1944.46	1054.51	10.545	1.0437	1.5699
79.	Medina	3.5709	10.58	124.92	1116.09	537.75	5.377	1.0089	1.0485
80.	Menora	3.1337	5.98	60.52	1520.05	656.37	6.564	1.0172	1.1183
81.	Mosman Park (WA)	1.9005	47.00	267.00	1333.92	627.35	6.273	1.0497	1.3560
82.	Mount Claremont	2.3101	6.54	121.46	1307.01	567.13	5.671	1.0369	1.2282
83.	Mount Hawthorn	1.8942	100.00	191.50	1774.82	797.55	7.976	1.0500	1.4753
84.	Mount Lawley	1.5865	59.39	346.11	1817.22	885.36	8.854	1.0618	1.6998
85.	Mount Nasura	2.0679	3.00	201.50	1325.11	712.77	7.128	1.0442	1.3609
86.	Mount Pleasant (WA)	3.0139	30.39	265.61	1627.96	809.87	8.099	1.0197	1.1710
87.	Mount Richon	3.9691	1.45	156.55	1252.92	622.05	6.221	1.0022	1.0135
88.	Mullaloo	3.7287	27.00	258.50	1383.00	744.17	7.442	1.0061	1.0463
89.	Munster	3.904	167.60	183.90	1179.75	484.96	4.850	1.0032	1.0156
90.	Murdoch	2.435	120.45	10.55	1451.54	542.96	5.430	1.0335	1.1958
91.	Nedlands	2.3292	119.74	95.26	1256.80	493.34	4.933	1.0364	1.1928
92.	North Lake	2.8949	2.68	53.32	1242.76	543.49	5.435	1.0223	1.1272
93.	North Perth	2.2107	74.00	243.00	1793.68	832.86	8.329	1.0398	1.3840
94.	Northbridge (WA)	1.6833	47.00	3.00	1603.40	601.79	6.018	1.0578	1.4024
95.	Ocean Reef	4.024	29.00	438.00	1415.34	814.45	8.145	1.0013	1.0105
96.	Padbury	2.8003	24.00	509.00	1554.60	885.75	8.857	1.0244	1.2383
97.	Parmelia	2.2581	10.75	547.75	1546.06	895.76	8.958	1.0384	1.4015
98.	Pearsall	3.6336	4.88	201.12	1260.82	712.36	7.124	1.0077	1.0565
99.	Peppermint Grove	1.5098	14.00	27.00	1180.68	471.89	4.719	1.0651	1.3465
100.	Queens Park (WA)	1.4585	13.39	351.61	1812.84	983.75	9.838	1.0674	1.8996
101.	Rivervale	1.8713	250.73	213.27	1807.20	753.68	7.537	1.0508	1.4524
102.	Rossmoyne	2.7395	11.67	129.83	1546.41	739.17	7.392	1.0258	1.2075
103.	Safety Bay	4.0999	12.12	643.88	1489.52	907.25	9.072	1.0001	1.0010
104.	Seville Grove	2.9966	6.00	875.00	1849.05	1228.28	12.283	1.0201	1.2762
105.	Shenton Park	1.4905	111.54	17.96	1314.85	477.68	4.777	1.0659	1.3567
106.	Sorrento (WA)	3.9876	23.27	259.23	1549.01	800.01	8.000	1.0019	1.0150
107.	South Fremantle	2.3925	80.70	77.30	1262.25	493.21	4.932	1.0346	1.1828
108.	South Guildford	2.5444	155.77	24.73	1224.24	415.72	4.157	1.0306	1.1336
109.	South Perth	3.9637	61.00	398.50	1826.11	909.80	9.098	1.0022	1.0206
110.	St James (WA)	1.5844	32.90	209.10	1703.36	836.19	8.362	1.0618	1.6516
111.	Stirling (WA)	2.5418	116.08	388.42	1832.87	917.31	9.173	1.0307	1.3196
112.	Subiaco	1.4745	218.10	25.40	1394.55	470.04	4.700	1.0667	1.3544
113.	Swanbourne	2.3087	12.48	72.02	1184.61	486.76	4.868	1.0370	1.1932
114.	Thornlie	2.9413	41.00	1948.50	2183.06	1507.13	15.071	1.0213	1.3730
115.	Tuart Hill	3.5353	14.43	386.57	1750.43	987.24	9.872	1.0095	1.0977
116.	Victoria Park	2.0073	44.00	387.00	1840.12	979.60	9.796	1.0461	1.5555
117.	Warwick (WA)	2.5922	31.00	145.00	1497.13	718.36	7.184	1.0294	1.2315
118.	Waterford (WA)	3.7474	9.82	88.18	1562.10	703.87	7.039	1.0058	1.0415
119.	Wellard	3.7979	21.06	289.94	1088.25	505.22	5.052	1.0049	1.0252
120.	Wembley	2.3809	74.89	275.11	1674.21	780.52	7.805	1.0349	1.3075
121.	West Leederville	1.22	59.13	46.37	1533.07	628.58	6.286	1.0795	1.6172

122.	White Gum Valley	3.1365	58.71	88.79	1271.43	525.82	5.258	1.0171	1.0934
123.	Willetton	3.8614	282.00	789.50	1985.17	1044.89	10.449	1.0039	1.0414
124.	Wilson	3.2732	10.00	305.50	1687.80	870.05	8.701	1.0144	1.1325
125.	Winthrop	2.7806	19.00	271.00	1558.53	767.85	7.679	1.0249	1.2077
126.	Woodbridge (WA)	1.1639	25.45	24.55	1132.42	420.19	4.202	1.0827	1.3963
127.	Woodlands (WA)	3.66	21.88	103.12	1437.34	666.83	6.668	1.0073	1.0496
128.	Woodvale (WA)	2.8889	44.00	520.50	1475.83	830.86	8.309	1.0224	1.2022

Table 8.3 The multiplicative effects of adding new retail jobs for the job-based TOD scenario

	Suburbs	***Distance < 2.4062 km***	***The number of retail Job***	***Add new retail jobs***	***New EJD Retail***	***Delta EJD Retail***	***Delta EJDRet (in 100 units)***	***Multiplicative effects based on one unit increase in EJD***	***New multiplicative effects***
1.	Applecross	2.2222	232.00	181.00	773.28	228.80	2.288	1.0065	1.0150
2.	Armadale (WA)	2.2723	369.31	562.69	920.11	437.57	4.376	1.0047	1.0208
3.	Ashfield (WA)	1.6536	7.40	83.10	575.23	239.68	2.397	1.0312	1.0765
4.	Atwell	2.2432	14.06	668.94	976.98	672.89	6.729	1.0058	1.0395
5.	Bassendean (WA)	1.6188	235.33	402.67	828.58	337.46	3.375	1.0330	1.1159
6.	Bayswater (WA)	2.1437	327.16	583.84	973.65	414.13	4.141	1.0095	1.0400
7.	Beckenham	1.92	177.87	289.13	777.16	289.44	2.894	1.0187	1.0550
8.	Bedford	2.136	51.47	310.53	860.83	447.71	4.477	1.0098	1.0447
9.	Bentley (WA)	2.3465	469.47	83.03	920.84	143.31	1.433	1.0021	1.0030
10.	Brentwood (WA)	2.3787	33.98	103.02	692.54	250.79	2.508	1.0009	1.0024
11.	Bull Creek (WA)	1.7516	395.00	99.00	866.51	142.68	1.427	1.0264	1.0378
12.	Camillo	2.3959	8.88	371.12	731.25	470.10	4.701	1.0004	1.0017
13.	Carlisle	1.2898	171.29	205.71	899.24	311.61	3.116	1.0525	1.1727
14.	Cottesloe	1.7159	398.00	40.50	775.27	90.99	0.910	1.0281	1.0255
15.	Craigie (WA)	2.1538	89.00	414.00	788.49	378.86	3.789	1.0091	1.0350
16.	Daglish	1.2295	1.80	76.70	616.22	231.31	2.313	1.0566	1.1358
17.	East Cannington	1.9572	26.72	336.78	834.25	436.58	4.366	1.0171	1.0767
18.	East Perth	1.4896	231.67	170.83	857.84	245.88	2.459	1.0401	1.1015
19.	East Victoria Park	1.8402	399.59	238.41	1071.02	301.06	3.011	1.0222	1.0684
20.	Eden Hill	1.7988	22.66	292.84	790.55	420.38	4.204	1.0241	1.1055
21.	Edgewater	2.0613	137.41	259.09	759.48	313.82	3.138	1.0128	1.0406
22.	Glendalough	1.9066	49.56	121.94	907.86	326.26	3.263	1.0193	1.0642
23.	Guildford (WA)	1.4108	85.76	3.24	459.65	61.63	0.616	1.0448	1.0273
24.	Heathridge	2.2095	61.00	578.50	910.79	567.01	5.670	1.0070	1.0404
25.	Highgate (WA)	1.1953	46.08	114.92	953.17	413.86	4.139	1.0590	1.2680
26.	Inglewood (WA)	2.1942	254.30	136.20	846.51	223.11	2.231	1.0076	1.0170
27.	Jolimont	1.4659	54.83	8.67	577.73	84.70	0.847	1.0415	1.0350
28.	Karawara	2.3618	66.20	72.30	699.40	196.28	1.963	1.0015	1.0030
29.	Kensington (WA)	2.3136	34.00	185.50	707.89	278.48	2.785	1.0032	1.0090
30.	Lathlain	1.3894	19.00	179.50	758.06	332.01	3.320	1.0461	1.1613
31.	Maylands (WA)	2.0831	252.00	600.50	1139.14	560.29	5.603	1.0119	1.0685
32.	Mosman Park (WA)	1.9005	230.00	381.00	868.34	374.22	3.742	1.0195	1.0751

33.	Mount Claremont	2.3101	23.12	260.88	687.53	280.41	2.804	1.0033	1.0094
34.	Mount Hawthorn	1.8942	299.00	147.00	1029.32	235.33	2.353	1.0198	1.0473
35.	Mount Lawley	1.5865	491.66	386.34	1278.83	418.95	4.189	1.0347	1.1538
36.	Mount Nasura	2.0679	0.00	224.00	566.68	309.55	3.096	1.0125	1.0392
37.	Nedlands	2.3292	415.90	225.10	889.93	228.37	2.284	1.0027	1.0061
38.	North Fremantle	2.3536	113.00	22.50	476.49	69.96	0.700	1.0018	1.0013
39.	North Perth	2.2107	305.00	279.00	1110.42	361.06	3.611	1.0070	1.0254
40.	Parmelia	2.2581	21.50	454.50	707.17	428.32	4.283	1.0052	1.0226
41.	Queens Park (WA)	1.4585	26.28	355.22	915.36	468.13	4.681	1.0419	1.2119
42.	Rivervale	1.8713	133.25	396.75	972.19	459.72	4.597	1.0208	1.0994
43.	St James (WA)	1.5844	193.90	82.60	826.92	185.38	1.854	1.0349	1.0656
44.	Swanbourne	2.3087	81.14	124.86	556.64	153.46	1.535	1.0034	1.0052
45.	Viveash	2.0809	5.53	78.97	470.83	166.93	1.669	1.0120	1.0201
46.	Wembley	2.3809	218.91	552.59	1159.97	532.06	5.321	1.0009	1.0046
47.	West Leederville	1.22	70.43	177.57	883.17	336.16	3.362	1.0573	1.2059

Geographical extent of agglomeration refers to the radius or scope within which any benefit of agglomeration for train ridership may be realised by adding more jobs. This means that manufacturing jobs may be added at a further distance away from destination stations, and still contribute to the increase in train ridership, than the other two sectors.

The incorporation of distance variable to train station has influenced the relationship between effective density and train ridership. As shown in table 8.1 – 8.3, adding new jobs produces larger multiplicative effects. The magnitudes of these multiplicative effects are dependent on how many jobs are added and how far the suburb is from a station. Figures 8.1 – 8.3 may explain the extent to which the distance variable may influence the multiplicative effect of agglomeration on train ridership based on sectors.

Referring to table 8.1 and figure 8.1, suburbs 86 (West Leederville) and 89 (Woodbridge) are both located approximately 1 to 2 km from their nearest station, yet their multiplicative effects are very different. West Leederville has more construction job activities (59 jobs) than Woodbridge (33 jobs). As the number of employed residents is higher in West Leederville, more jobs would need to be added to acquire the ratio of 1.5 than in Woodbridge. Adding 166 new construction jobs in suburb 86 leads to a 1.9 multiplicative effect on train ridership (an increase of 90%), but adding 51 new construction jobs in suburb 89 leads to only a 1.36 multiplicative effect (an increase of 36%). Similarly, adding jobs to suburbs located at different distances from the nearest station, for example, suburbs 18 and 65 in figure 8.1, also results in different multiplicative effects. Similar conclusions may be applied to the other sectors (figures 8.2 and 8.3).

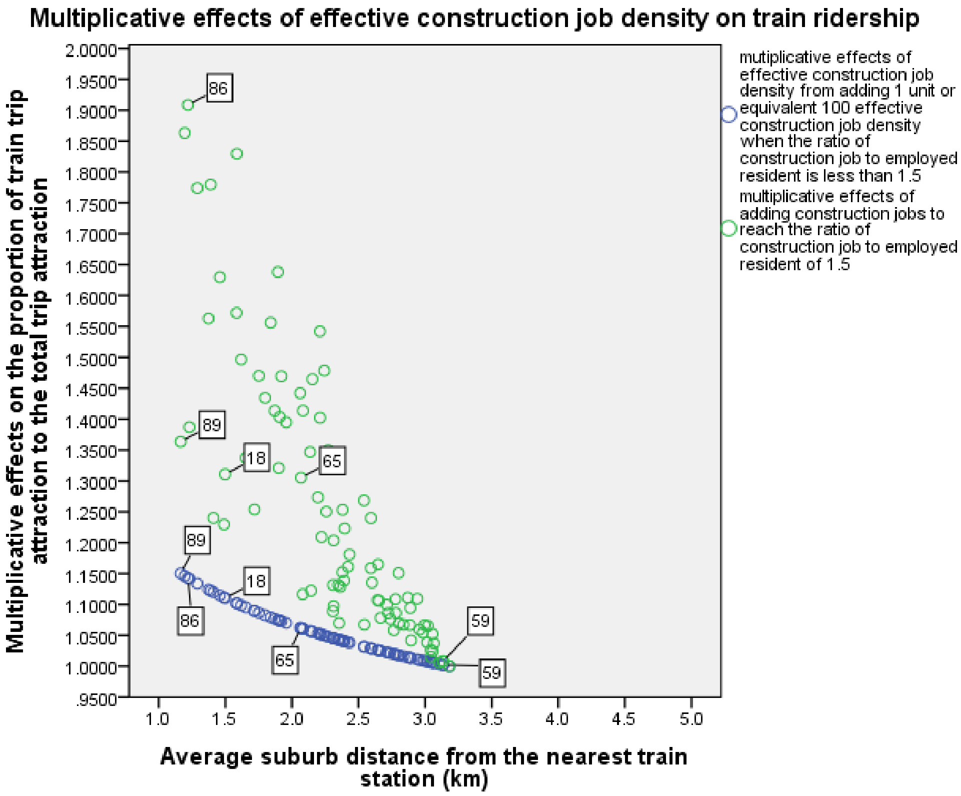


Figure 8.1 The multiplicative effects of adding new construction jobs on the proportion of train ridership attraction

In addition, while the magnitude of multiplicative effects is both a function of job number and distance, the results imply that different level of activities, as measured by current job numbers (West Leederville is much more developed than Woodbridge, for example) also contribute to different levels of multiplicative effects.

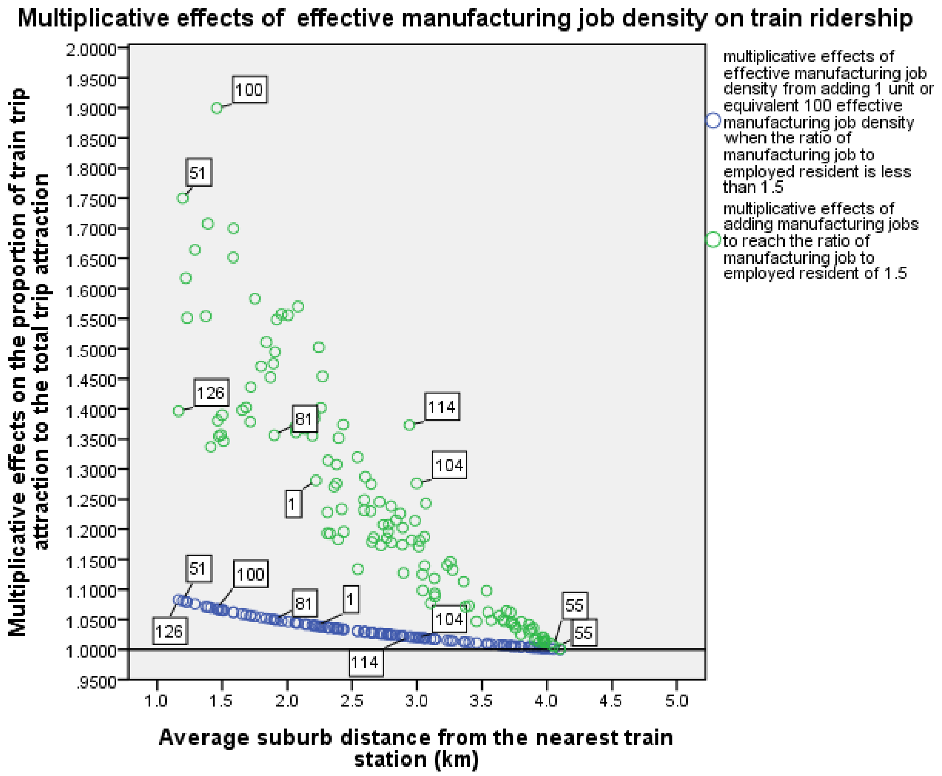


Figure 8.2 The multiplicative effects of adding new manufacturing jobs on the proportion of train ridership attraction

Figures 8.1 to 8.3 clearly demonstrate that the multiplicative effects on the proportion on train trip attraction are highest when the effects of effective job density are confined in a smaller geographic area. The construction sector (figure 8.1) and manufacturing sector (figure 8.2) showed almost similar multiplicative effects on train trip attraction. The multiplicative effect of effective density from the manufacturing sector however, it is confined in larger geographical extent (figure 8.2) than from the construction sector. On the other hand, the retail sector produced a smaller magnitude effect in much smaller geographical area (figure 8.3). As a steep curve over a short distance provides evidence of more concentration (Paez et al, 2011), this may suggest that the construction sector demonstrates the highest effects of effective job density on train ridership attraction.

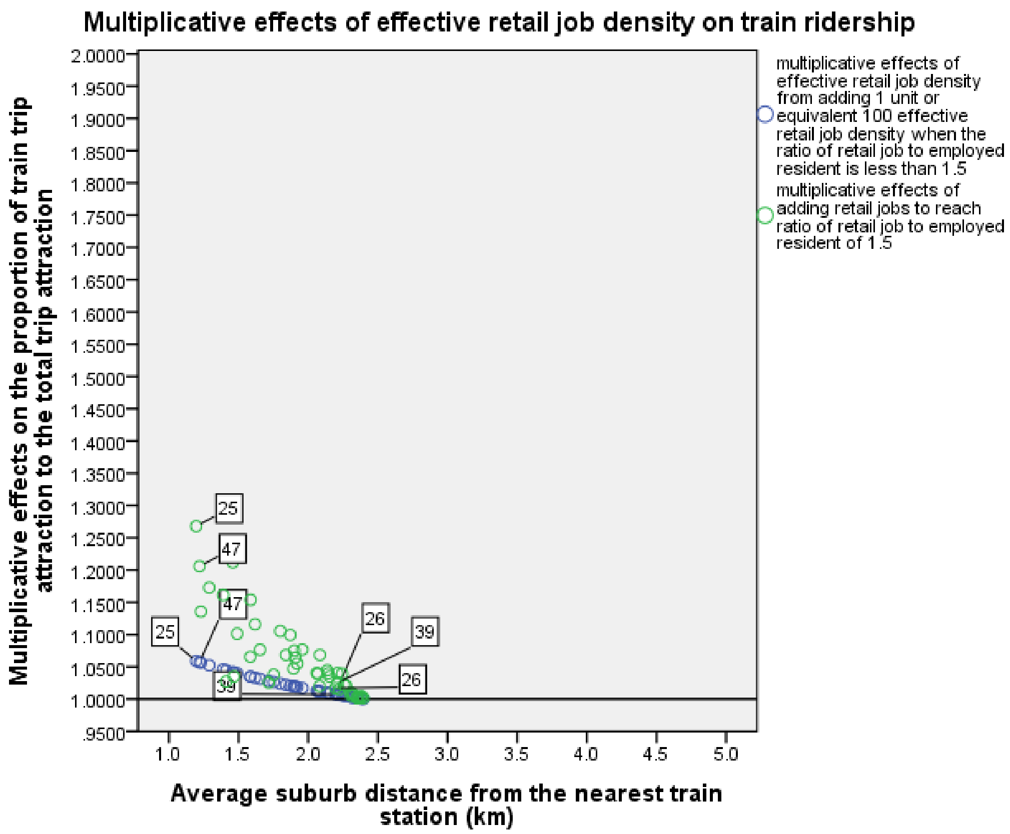


Figure 8.3 The multiplicative effects of adding new retail jobs on the proportion of train ridership attraction

8. 2 THE TRADE-OFF MECHANISM AND THE JOB-HOUSING BALANCE HYPOTHESIS

This section discusses the relationship between effective density and train ridership by considering the implications of the job-housing balance. This is discussed from the perspective of a wage/land rent and travel cost *trade-off*.

This thesis incorporates the concept of travel time saving, which has been used to explain the trade-off between wage, land rent or housing costs, and travel costs (Ryan, 2011, Cervero and Duncan, 2012, So, K 2011, Timothy and Wheaton, 2001). That is, lower paid workers in less-skilled jobs, such as in the retail sector or in blue collar occupations, tend to be more spread out. Therefore, they have more opportunities to live closer to work and spend less on transport costs. On the other hand, the higher paid jobs are often located in the city centre where houses are more expensive. There is a trade-off between paying more for a house to live closer to work and paying less for a house but travelling further. Wang and Chai (2009) suggested that wage differences between blue collar and white collar employed resident allowed for white collar workers with higher wages to live farther from employment centres and that blue collar worker with lower wages lived clustered near their workplaces.

The travel time saving concept suggests that once accessibility improved and travel time or travel costs reduced, these benefits are translated into higher real incomes (Venables, 2004). In addition, locations with improved accessibility following the railway line extension would also improve its centrality and the travel times to/from these locations, and travel time savings being capitalized into higher land rents (the land rent premium) (Ryan, 2011). Workers can choose to spend more on land rent; thus, more people can afford to live in a metropolitan area, on aggregate increasing the metropolitan size. The effective density scheme proposed in Venables (2004) and Graham (2007) applied the same principle of trade-off between wage, land rent or housing costs and travel costs, in order to determine the city size and, due to an increasing rate of return, to explain increases in urban productivity.

Thus, a strong correlation between travel time saving, wage levels, land rents, urban productivity and effective density, especially in relation to the impact of public transport investment, has been proposed. The decision on where to live and to work assumed to be resulting from this trade-off thereby determines the level of job-housing balance in the overall study area. Nevertheless, Timothy and Wheaton (2001) cautioned that there may an uncertainty as to whether these trade-offs were a result of the agglomeration equilibrium or the dis-equilibrium of employment distribution.

In this section, this study attempts to investigate whether the effect of effective density on train ridership can be explained by the concept of wages/travel costs trade-off. Where the job-housing balance has been a consequence of this trade-off, investigating the current job-housing balance characteristics may inform as to why effective density may be likely to influence train ridership.

The wage premium and land rent premium was calculated as a function of travel time saving following the research method proposed originally in Herbert and Steven and later in the research of Timothy and Wheaton, 2001 (Herbert and Steven, as cited in Timothy and Wheaton, 2001). In addition, the effective density premium can also be calculated as a function of travel time saving following the same procedure. Thus, the effective density premium can be compared to the wage and land rent premium to assess the characteristics of various suburbs.

Referring to the method applied in Timothy and Weathon, 2001, dummy variables were developed based on the job-housing balance category (instead of the employment centre dummy variable of Timothy and Weathon's study) and the level of job-housing balance in each suburb was determined. All suburbs were categorised into three groups based on the value of their job-housing balance, as explained in section 4.1.2.6 in chapter 4. These categories are dummy variables for the construction of log linear wage equation, and are mapped in figure 8.4:

1. Category 1: ratios of job-housing balance of between 0 and 0.75 represents suburbs where the number of jobs available are only limited compared to the number of employed residents. Suburbs in this category are named as "employed resident-rich" suburbs or suburb category 1.
2. Category 2: ratios of job housing balance of between 0.75 and 1.5 represent a balance between the number of jobs and the number of employed residents. Therefore, this dummy variable will be served as the reference category. Suburbs in this category are named as "job-housing balanced suburbs" or suburb category 2.
3. Category 3: ratios of job-housing balance above 1.5 represent suburbs where the number of jobs exceed the number of employed residents by at least 50%, where for every 1 employed resident there are 1.5 jobs available. Suburbs in this category are named as "job-rich suburbs" or suburb category 3.

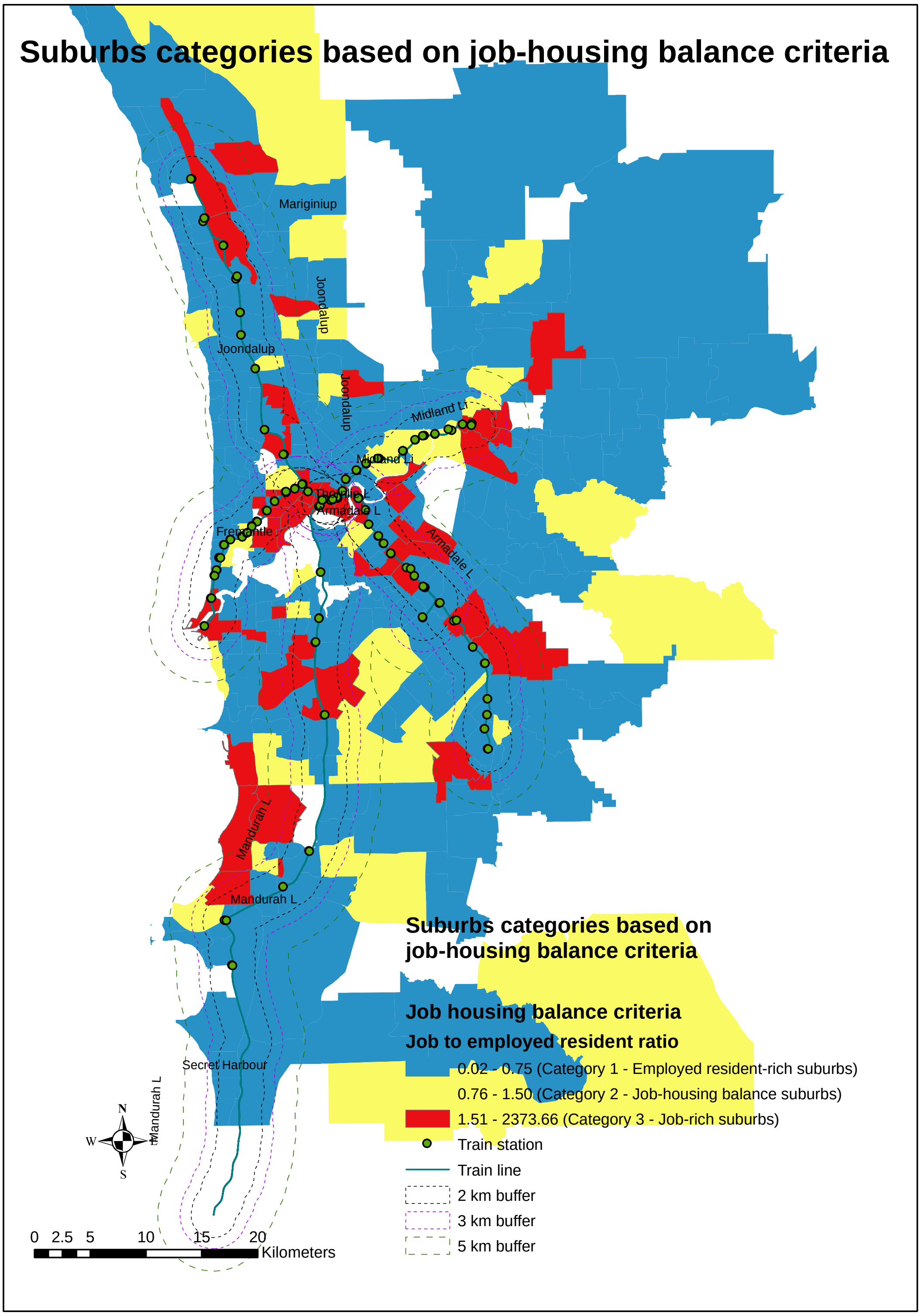


Figure 8.4 The category of suburbs based on job-housing balance criteria

Based on the study of Timothy and Wheaton 2001, wage differences between employment centres may be related to the average commuting costs for employment locations. Timothy and Wheaton, 2001, estimated the wage differences using the following log-linear wage equation:

$$\ln(W_j) = \alpha X_j + \beta Z_j, \quad j = 1, \ldots, J, \quad \text{Equation 8.4}$$

Where:

W_j = wage of individual worker *j*

X_j = a vector of work zone specific dummy variables

Z_j = a vector of individual worker characteristics.

Since Timothy and Wheaton, 2001, used a microdata sample dataset, the estimation of wages was able to be conducted at the individual worker *j* that worked at a work zone *j*. The *Z* parameter represented the control for individual characteristics in the model.

This thesis has used an aggregate approach based on the fishnet dataset, so that wages have been aggregated over suburb j from all n fishnets in suburb j, and wages have been specified for each occupation type, especially for managers/professionals and the blue-collar wages group. Therefore, the vector *Z* represents the control for various suburb characteristics, especially in terms of its job and occupation structures. Therefore, equation 8.4 may be modified as follows:

$$\ln(W_j) = \alpha X_j + \beta Z_j, \quad j = 1, \ldots, J, \quad \text{Equation 8.5}$$

Where:

W_j = the average hourly wages over all workers in *n* fishnets in suburb *j*.

X_j = a vector of work zone specific dummy variables based on the job-housing balance category: where the work zones are classified into category 1 (employed resident-rich suburbs, 2 (job-housing balance suburbs), and 3 (job-rich suburbs). This can take one of the 3 values 1, 2, 3, and category 2 is the reference category.

Z_j = a vector of suburb characteristics as control variables consisting of: the proportion of employed residents working as managers/professionals, the proportion of employed residents working in blue collar occupations, the proportion of employed residents working in the construction sector, the manufacturing sector and the retail sector, the proportion of manager/professional jobs, the proportion of blue collar jobs, the proportion of jobs in the construction, the manufacturing, and the retail sector.

j = a suburb over all *J* suburbs in the study area.

By the same procedure, the land rent differences and the effective density differences across suburb are estimated by the same model, as follows:

$$\ln(LV_j) = \alpha X_j + \beta Z_j, \;\; j = 1, \dots, J, \text{ Equation 8.6,}$$

And

$$\ln(ED_j) = \alpha X_j + \beta Z_j, \;\; j = 1, \dots, J, \text{ Equation 8.7}$$

Where

LV_j = the mean of land rents of suburb *j* over the *n* fishnets in suburb *j*.

ED_j = the mean of effective density (both in terms of employed resident and job) of suburb *j* over the *n* fishnets in suburb *j*.

X_j and Z_j are defined as mentioned above.

j = a suburb over all *J* suburbs in the study area.

All of log-linear equation results are presented in the appendix of 62-71 of chapter 8.

In place of work dataset or train trip attraction model, the addition of the agglomeration variable to the LUTI model has had the effect of rendering wage and land rent variables statistically insignificant in the SETI model. There may be multicollinearity issues between the indirect or partial urban productivity factors (wage, land rent and effective density) that result in a more significant effect of one factor and less significant effect of other factors on the dependent variable. For example, the correlation between the wages of managers/professionals and land rent is 0.5, the correlation between the wages of managers/professionals and effective job density is 0.4, and the correlation between land rent and effective employed resident density is 0.6.

In order to find the true influence of wage levels on the proportion of train trips, an additional model (without land use variables and effective density variables) has been developed (appendix 59 of chapter 8) to identify the direct effect of wages on train ridership. This model shows that the wages of managers/professionals has a statistically significant positive influence in attracting train trips, whereas the wages of blue collar workers remains insignificant. Therefore, the following nvestigation on wage premium equation only holds for the wages of manager/professional occupations.

The estimations from equations 8.5 to 8.7 were conducted allowing for diverse intercepts in each job-housing balance category through using dummy variables, with the reference category being the balanced category (category 2). The results have an adjusted R^2 range from around 0.30 to more than 0.60. Table 8.4 presents the associated wage premium, land rent premium, and effective density premium for each of the job-housing balance categories. The standardized beta coefficients represent the percentage difference in wages (or in land rent or effective density) between the job-rich suburbs (category 3) and the reference category (category 2); and between the employed resident-rich suburbs (category 1) and the reference category (category 2).

Table 8.4 The standardized beta coefficient of wage premium, land rent premium, and effective density premium for each job-housing balance category suburbs (Appendix 62-65)

Variable	***Category 1 job-housing ratio 0 – 0.75***	***Category 2 job-housing ratio 0.75-1.5 (balance)= reference category***	***Category 3 job housing ratio > 1.5***	***The adjusted R^2***
Wage premium (managers/professionals occupation)	-.051	0	.303	38.2%
Land rent premium	-.050	0	.079	30.8%
Effective employed resident density in all sector premium	.123	0	.104	49.1%
Effective job density in all sector premium	-.077	0	.297	57.1%

All variables in the table are statistically significant at the 99% level of confidence (p<0.001).

Based on table 8.4, wages for managers/professionals in the job-rich suburbs are 30.3% higher than that of the balanced job-housing suburbs. These results are in line with the results for the effective job density premium, where the job-rich suburbs are around 30% higher for effective job density compared to the balanced suburbs.

Similarly, the job-rich suburbs are 8% higher in land rent than balanced suburbs. For category 1, employed resident-rich suburbs have around 8% lower effective job density, 5% lower for land rent, and 5% to 7% lower wages compare to balanced suburbs. However, the effective employed resident density is 12% higher than that of the reference category. Note that effective job density and wages have a similar level of premium for category 3 suburbs, and that land rent premiums are lower.

The similarity of relative premiums for effective density and wages may suggest, based on Venables's convex curve for wages (Venables, 2004), that they both may exhibit similar patterns of distance decay relative to the distribution of stations. This explanation has been investigated based on the empirical data from 2011 in the study area.

To examine why suburbs in different job-housing balance categories have different characteristics for wages, land rents and effective density, a simple curve fitting model has been developed, based on the fishnet data set. The average value for wages, land rents and travel times in each buffer ring, as a function of the distance between suburbs and stations, have been constructed referred to equation 5.9 and 5.10 in chapter 5. Figure 8.5 -8.7 present the empirical data from 2011 based on the relationship for wages, land rents, and travel times respectively. The graphs are drawn based on the average values for ease of interpretation. The dependent variable is transformed into log natural, in order to perform the elasticity of wages on distance as a percentage.

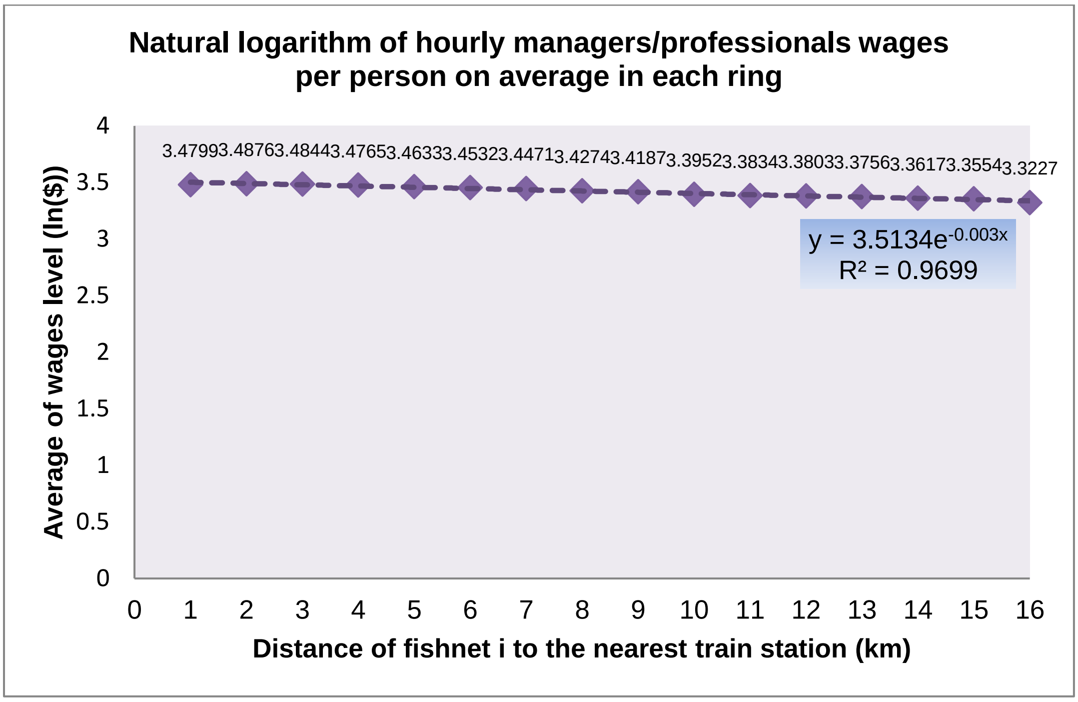


Figure 8.5 The association between wages and the suburb distance from train station. The *y* value in average ln of wages in $, referred to equation 5.10.

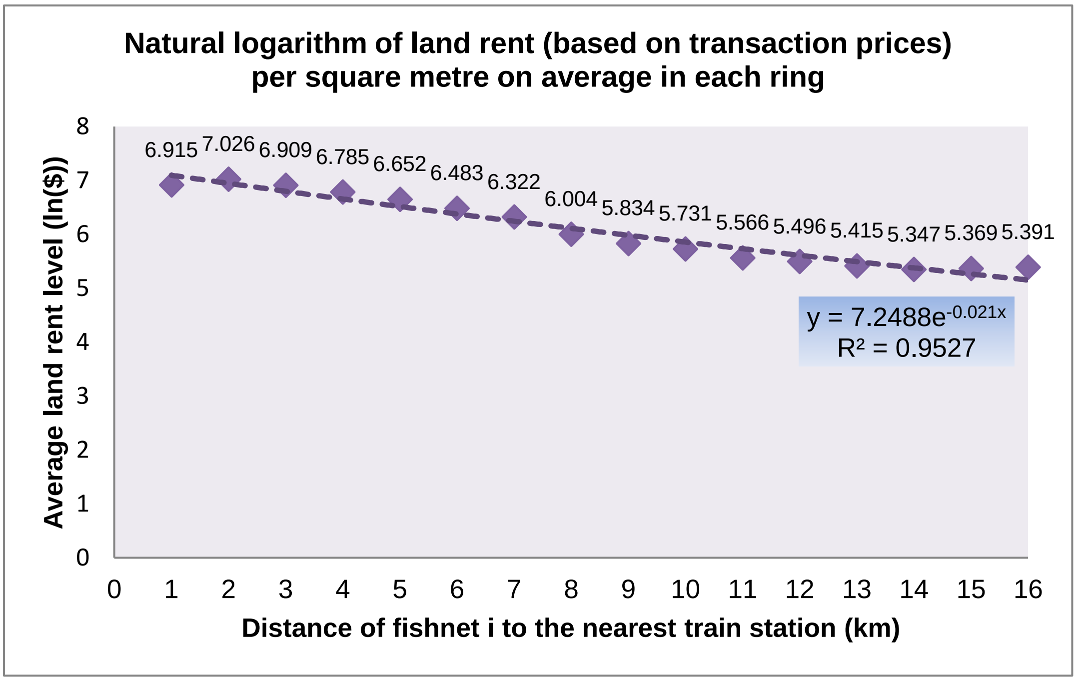


Figure 8.6 The association between land rents and the suburb distance from train station. The *y* value in average Ln of land rent in $.

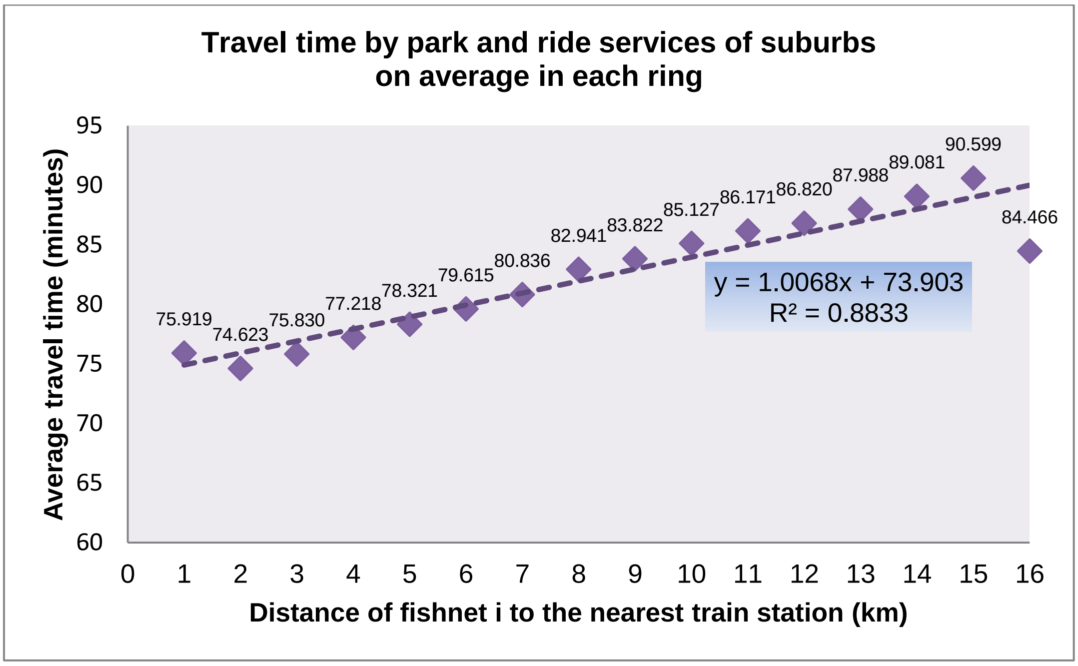


Figure 8.7 The association between travel costs (travel time) and the suburb distance from train station. The *y* value in average travel time in minutes.

A complementary graph showing the elasticity of wages and land rents on the changes in travel times is also shown in figure 8.8. The scatter graphs in figures 8.5-8.7 were developed based on equation 5.9 and 5.10. However, figure 8.8 was constructed from a variation of equation 5.9, but based on the same principle, where the fishnet data were divided into travel time buffer rings instead of distance buffer rings. Equations are rewritten as follows:

$$a = \sum_{i=1}^{R} a_i \qquad \text{Equation 8.8}$$

$$b_i = {}^{a_i}/_{a} \qquad \text{Equation 8.9}$$

Where:

a_i = the value of variable (effective density) of a fishnet located in ring *i*

b_i = the average value of fishnets located in ring *i*

a = the total value of effective density over all fishnet *i* over all ring *R*.

R = 16.

i = defined as travel time buffer rings from 1 to R=16 (instead of the station distance rings), where ring 1 represented travel time from 60 minutes to <=63 minutes, ring 2 represented travel time from 63 minutes to <=66 minutes, until it reachs ring 16 where the travel time is in between 105 and 108 minutes.

Based on Venables's wage curve model, this study has assumed the central location or the zero location to be stations instead of the CBD, and has calculated how the wages of managers/professionals, the land rents and the travel times are correlated with distance from the train station. Figure 8.5 and 8.6 show the results of the trade-off. In exponential form, the beta coefficient directly refers to the elasticity or the percentages changes in *y* as a function of *x*. However, in a linear model, such as that shown in figure 8.7, the coefficient beta value is the effect on *y* due to an increase of one unit in *x*.

These graphs may be interpreted both for the workplace suburbs and the residential suburbs. For example, applying them to workplace suburbs, a 2.1% increase in land rent density paid by firms, for every 1 km the workplace suburb is closer to the nearest station, is associated with a a reduction of 1.0068 minutes in the average travel time to suburbs located in the assigned ring buffer in question. Workplaces nearer to stations or located in more central locations will therefore be associated with higher wages, higher productivity, less travel costs, and a larger catchment size, and be expected to attract higher train ridership.

It is apparent from figure 8.6 that if firms choose to move to a location with an improvement of only 1 minute in accessibility, firms will then pay 2% more in land rent, and therefore would need to focus on activities with higher productivity to be able to relocate. If they choose to establish firms in a less accessible area, they will pay less for rent at a cost of their employees spending more on transport costs and potentially lower productivity.

Applying the results to the residential suburbs, for every 1 minute increase in travel time to reach workplace, employed residents need to receive 0.05% higher wages in order to compensate for the increased commute to work by train (figure 8.8); or, for every 1 km increase in distance of residential locations from stations, (figure 8.5), employed residents need to receive 0.3% higher wages, holding the house prices constant.

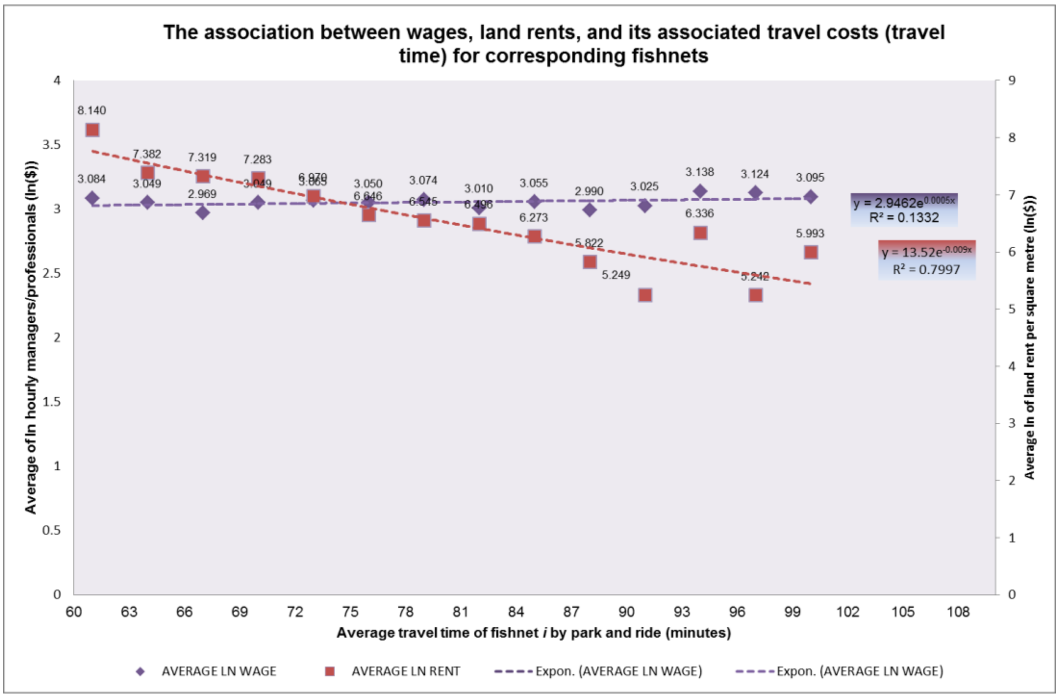


Figure 8.8 The association between travel times and the level of wages or land rents

It is apparent from figure 8.8 that, for every 1 minute saved in travel time, residents choosing to move to a location with an improvement of 1 minute in accessibility or centrality are willing to pay higher housing costs of 0.9% per square metre, holding the wages constant. Note that figure 8.8 shows a limited amount of variability in wages premium. The wages and land rent are much more sensitive on the increase in suburbs' distance to stations than on the increase in suburbs' average travel time.

The results from figures 8.5 and 8.6 may suggest why job-rich suburbs gain higher positive premium in wages, land rent and effective job density than the reference category (as shown in table 8.4). Job-rich suburbs, thus, are likely to be a typical of areas with higher accessibility that may accumulate higher travel time saving for workers and firms located there. Thus, locations near stations or locations with lower travel times may experience higher travel time saving and be capitalized into larger land rent premiums, and therefore higher effective job density compare to other areas (table 8.4).

The map of job-rich suburbs (category 3 on figure 8.4) shows that their distribution is essentially related to the areas around the Perth CBD (which have a high centrality), suburbs near train stations, inner suburbs and suburbs close to south-western coast line near the industrial centre of Henderson-Kwinana which is close to major highway/freeway routes.

The employed resident-rich suburbs (category 1) correspond to lower wage premiums, land rents, and effective job densities compared to the reference category, and therefore also in comparison to category 3. These areas are dominated mainly by residential activities, and eventually generate higher effective employed resident densities compared to the reference category. These areas may correspond to areas of the least degree of centrality, as suggested by the graphs of figures 8.8.

In order to further explore the implications of these job-housing balance characteristics, commuting patterns among these suburb categories have been examined.

Table 8.5 Differences in the numbers and proportions of train trips between suburb categories (appendix 66-69)

Variable	*Category 1 job-housing ratio 0 – 0.75*	*Category 2 job-housing ratio 0.75-1.5 (balance)= reference category*	*Category 3 job housing ratio > 1.5*	*The adjusted R^2*
The number of total trip attraction	-0.314	0	0.134	38.6%
The number of total trip production	0.077	0	-0.255	36.6%
The proportion of train trip attraction	.048	0	.178	19.5%
The proportion of train trip production	.198	0	.182	13.3%

The total number of trips and the proportion of trips by trains have been compared between each suburb category (table 8.5). Table 8.5 shows the employed resident-rich suburbs have higher total trips leaving those suburbs (production) than the reference category, and the job-rich suburbs have higher total trips arriving at those suburbs (attraction) than the reference category, which is not unexpected. However, in terms of the proportion of train trips, for both production and attraction, suburb categories 1 and 3 are both higher than the reference category. Job-rich suburbs have the highest proportion of train trip attraction, since these suburbs may be expected to attract considerably more train trips than the job-housing balanced suburbs. Similarly, employed resident suburbs have the highest proportion of train trips produced, since these suburbs are likely to produce considerably more train trips than the balanced suburbs.

However, to understand why the balanced suburbs produced and attracted the lowest levels of train ridership, a further log linear wage equation based on travel time saving was conducted, and a regression model with the natural log of hourly wages of managers/professionals as the dependent variable, and the travel time of park and ride, split by suburb category, as the independent variable was also performed. An interaction variable was constructed between the travel time and the category of suburbs to determine the elasiticity of wages on the changes in travel time for each suburb category, compared to the reference category. Results are as follows in table 8.6.

Table 8.6 The trade-off between hourly wages (managers/professionals occupation) as dependent variable and travel time park and ride for each suburb categories (appendix 70)

Independent variable	*Coefficient of travel time of ln wage of managers/ professional*	*Std errors*
Average travel time	-0.004	0.000
Employed resident-rich suburbs travel time (category 1)	0.002	0.000
Job rich suburbs travel time (category 3)	0.005	0.000
Adjusted R-square		30.2%

The value of travel time saving or *votts* is defined as "*the amount of money an individual is willing to outlay in order to save a unit of time spent traveling, cateris paribus*" (Hensher et al., 2005, p. 358). However, this concept is usually measured based on travel time spent on commuting. In this study, the travel time saving is measured from the centrality of workplace suburbs, based on the average travel time by park and ride to reach this workplace from all other (residential) suburbs on average. Thus, the regression coefficient represents the changes in hourly managers/professionals wage levels offered by each suburb category, relative to the reference category, for every 1 minute increase spent in travel to reach this workplace or for every 1 minute reduction in centrality of this suburb in question. By assuming 8 hours of working time, "a one minute of travel time represent a loss of 1/480[15] of the time spent working in an 8 hour day", or two minutes of two way journey travel time represent "a loss of 1/240 of the time spent working in a day[16]" (Timothy & Wheaton 2001, p. 353). As the data of trip journey represents one way journey, the coefficient in this equation would represent the impact of an increase in two minutes of travel time on the wage level.

Comparing the job-rich and employed resident-rich suburb categories to the reference category, a regression coefficient in table 8.6 represents the impact that increasing 1 minute of travel time would have on increasing or decreasing wages in job-rich and employed resident-rich suburb category, relative to the reference category. Dividing the estimated coefficient of each suburb category by 1/240, those who are working in employed rich suburbs (category 1) have a value of time of 0.48 of the wage rate for those who are working in the reference category[17]. Those who are working in the job-rich suburbs (category 3) have a value of time of 1.2[18] of the wage rate of those who are working in the reference category. Thus, the value of time of employed residents working in the job-rich suburbs in higher than the reference category and they get higher wages; and the value of time of employed resident working in the employed resident-rich suburbs is lower than the reference category and they get lower wages.

People who work in suburbs category 3 are working in more productive suburbs, and thus get paid more and have higher *votts* is consistent with the calculation refer to table 8.4. They may be willing to travel more than people working in job-housing balanced suburbs. It has been well known that worker received high wages tends to live at lower

[15] As 8 hours means 8 times 60 minutes or 480 minutes

[16] The trip pattern data from the ABS is based on the one way commuting journing.

[17] The coefficient of travel time of workers who work in employed rich suburbs category is 0.002. Thus 0.002 divided by 1/240 is equal to 0.48.

[18] The coefficient of travel time of workers who work in job-rich suburbs is 0.005. Thus 0.005 divided by 1/240 is equal to 1.2.

densities and commute farther (Timothy &Wheaton, 2001; Wang & Chai, 2009). This may suggest workplaces located near the city centre or stations of the job-rich suburbs have a larger catchment of train users due to workers being willing to undertake a longer commuting journey. Figure 8.9 shows that the centrality of suburbs, as measured by travel time by park and ride, is higher for suburbs located both near the Perth CBD and near to stations. This means, employment centre such as the Perth CBD, Fremantle, Armadale, Welshpool would have larger catchment by train than job-housing balance suburbs. Due to this larger catchment, these suburbs attract more train ridership or produce more train ridership than the reference category.

An explanation as to why job-housing balanced suburbs produce and attract a lower proportion of train ridership than other categories may be due to the spatial structure of the distribution of jobs and employed residents in the study area. Map in figure 8.4 showed that job-housing balanced suburbs are located in the areas in railway lines, or areas with limited railway services. These areas may also be associated with lower centrality and less development than other locations. This may suggest that limited land development, or low numbers of employed residents and jobs, where the ratio of jobs and housing remains in balance technically, may be related to lower travel demand for train ridership from/to these areas overall in comparison to other suburb categories.

When the accessibility to the railway services of suburbs is low, a policy of mixed land use (which provides an equal number of residential and employment land uses) may produce best results in reducing the total number of trips, but it will not increase train ridership. When the accessibility of suburbs is high, these areas are attractive for job development. Job-rich suburbs have the potential to increase both the train trip production and attraction, since these areas offer higher wages, working people in these suburbs have high *votts,* and thus the area has a larger catchment that can utilize both the effective employed resident and job densities to produce and attract train ridership. The high effective density premium generated from these areas may increase train ridership provided there is sufficient connectivity by railway services between residential suburbs and workplace suburbs. Employed resident-rich suburbs near train station may still attract train ridership if jobs are provided within the station precinct as working people work in these areas would have low *votts*. Employed resident-rich suburbs far from stations may still capture a higher effective employed resident density to generate train trip production (refer to the model results in section 7.2.2.1).

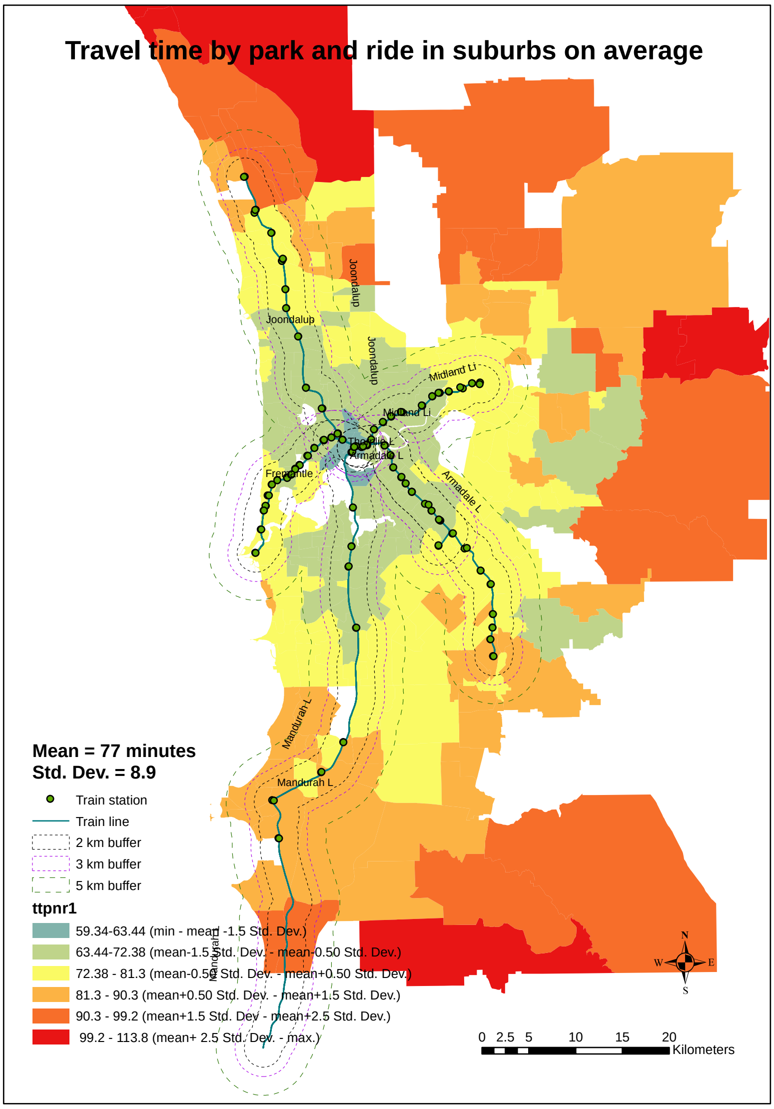


Figure 8.9 Map of average travel time by park and ride in suburbs.

8. 3 CHAPTER SUMMARY

Bartholomew (2007) had discussed a type of land use-transportation scenario planning based on labelling names and the variables employed. The labelling names based scenario attached either positive values or negative values associated with the names scenario, such as “Business as Usual” (adapted from Allen et al., as cited in Bartholomew, 2007, p. 402 – a negative scenario; Wise Growth Scenario – a positive value; Urban Sprawl Scenario – a negative value; Village/Town Centers Scenario – a less value labelled scenario). Based on the variables employed, the spatial pattern and urban from being the dominant motivational of formed scenario. This included location and density of growth, land use diversity, design, and including the job-housing balance. According to the categorization made in the meta-analysis from Bartholomew (2007), the policy scenario developed in this thesis applied the comparison between the ‘Business as Usual’ and the ‘Job-based TOD scenario’, which relied on the variable of job-housing balance. The results showed that the job-based TOD scenario had increased train ridership as much as 1.19 times higher (in the construction sector), and 1.215 and 1.04 times higher for the manufacturing and retail sector respectively than the ‘Business as Usual’ or no policy scenario.

This thesis has demonstrated that the construction sector possesses higher effective job density confined in smaller areas, and, thus produce larger agglomeration effects on train ridership attraction than the other two sectors. The manufacturing sector possesses a higher effective employed resident density but is confined across a larger area, and still produces larger agglomeration effects on train trip attraction than the other two sectors. The retail sector possesses the lowest effective job and employed resident density confined in the smallest area compare to the other two sectors, thus produces the lowest agglomeration effects on train trip attraction.

The geographical extent of agglomeration may be useful in assessing whether the policy of adding jobs will increase the effectiveness of TOD policy. Strategic decisions would involve choosing the suburbs that gave the largest multiplicative effects, which is a function of both the number of jobs added and the location of suburb relative to the distribution of stations and the current or initial level of activities (mixed land use and quantity).

The job-housing balance hypothesis has been investigated using the concept of a trade-off of wages and housing costs/travel costs. This thesis found that job-rich suburbs gave the highest premium in land rents, wages, and effective density than other suburb categories. The premiums for wages and effective density were similar and higher than the land rent premium. The levels of train trips produced and attracted to these suburbs are also higher than other suburbs. A travel time saving calculation has shown that those who work in job-rich suburbs have the highest value of travel time saving. They receive higher wages and are willing to pay more to save on travel time. The high accessibility of these suburbs, complemented by a high value of travel time saving, may suggest that workplaces situated in job-rich suburbs near stations should have larger catchment sizes. Thus, these workplaces would make most use of the multiplicative benefits of both effective employed resident density and job density in increasing train trip attraction.

The analysis of job-housing balance from the trade-off perspective and its implication on train ridership suggests that a policy of mixed land use, when applied to suburbs with low accessibility to the railway services, may produce best results in reducing the total number of trips, but it will not increase train ridership. When the accessibility of suburbs is high, these areas are attractive for job development. Job-rich suburbs have the potential to increase both the train trip production and attraction. Employed resident-rich suburbs near train station may still attract train ridership if jobs are provided within the station precinct as working people work in these areas would have low *votts*. Employed resident-rich suburbs far from stations may still capture higher effective employed resident density to generate train trip production.

The influence of effective density on train ridership may be more relevant in the train trip attraction model than that of train trip production model.

CHAPTER 9. CONCLUSIONS AND RECOMMENDATIONS

9. 1 INTRODUCTION

The previous chapter presented a policy analysis that incorporated the model interpretations and the evaluation of the various models based on hypothesis 3 and hypothesis 4. This chapter summarises the major findings of this thesis, discusses the limitations of the theories and discusses the methods and applications of modelling train ridership prediction. This chapter also addresses future directions that may arise from the outcomes of this thesis. The research hypotheses set out for this thesis are reiterated to show whether these have been supported or not.

9. 2 A MODIFICATION OR EXTENSION OF DENSITY INTO AN EFFECTIVE DENSITY CONCEPT FOR TRAIN RIDERSHIP PREDICTION

Background studies show that the massive transport investment allocated to the development of public transport infrastructure in large cities in the world has not yet reduced the level of car use to a very low level (Mohan, 2008). There are efforts in promoting public transit use, which have been undertaken with an understanding of a strong interaction between land use and transport, producing policy measures that prioritise public transport (Badoe & Miller, 2000). However, the current discourse on land use-transport studies from the T-LU side has lacked an understanding of the impact of public transport investment on the interlinked nature of spatial economic activity, land use, and travel behaviour. The wider economic benefits (WEBs) impacts assessment of public transport investment has emphasised the assessment of the agglomeration effect of public transport externalities on urban productivity in terms of public transport induced agglomeration (Graham, 2007). Limited studies on WEBs have attempted to examine the travel behavioural responses of public transport investment. The literature suggests that understanding land use-transport interaction involving the spatial economic-transportation system interaction would lead to a better understanding of travel behaviour effects (such as travel demand) on the transportation system (Russo & Musolino, 2012). .

This thesis has highlighted a modified or extended parameter based on density (based on the land use-transport interaction framework or the LUTI), namely the effective density (based on the spatial economic-transport interaction framework or the SETI), as one approach to understanding the impact of public transport induced agglomeration on train ridership. A simultaneous measurement of scale and proximity that has explicitly involved the accessibility improvement resulting from transport investment is one way in which the effective density parameter may better represent the market potential than the unmodified density parameter, since it may be able to measure the travel demand effects of the transportation system (such as railway line extension) in any model predictions.

In response to the problems and gaps stated in the literature mentioned above, this thesis aims to establish a model to understand the relationships between agglomeration and train ridership, by incorporating the influence of spatial economic-transportation interaction into ridership prediction. The effective density measurement, in terms of both employed residents and jobs (the SETI-based model), has been proposed to modify or extend the density measurement (the LUTI-based model). The train ridership prediction under the LUTI and the SETI frameworks have been compared.

The investigation was conducted for the Perth metropolitan region where the investment in the Perth-Mandurah railway line extension has taken place since 2004 and culminating in the opening in 2007. With eight years having past since railway opening, it is expected that land use development, densification and intensification of development have taken place subsequently in the study area. The impact of this transport project on the level of train ridership has been examined.

Four research objectives were derived. The first objective was to investigate the link between transportation investment and agglomeration by providing evidence that support the existence of public transport induced agglomeration. The second was to understand the influence of public transport induced agglomeration on train ridership in terms of both effective job density and effective employed resident density. The third was to investigate the geographical extent of agglomeration relative to the distribution of stations that were assumed to be the focal point orsource of agglomeration. The fourth was to investigate the implications of agglomeration for job-housing balance from the perspective of a trade-of between wages/land rents and travel costs and investigate the further implications for train ridership.

9. 3 SUMMARY OF RESEARCH FINDINGS

This section highlights the major findings of modelling train ridership prediction. These findings focus on restating the important role that effective density plays in train ridership prediction.

In summary, the main research findings from this thesis are as follows:

- This thesis has modelled train ridership based on models with and without agglomeration in order to understand the influence of effective density on train ridership. The model was conducted for both train trip production and train trip attraction. Comparing the sets of different models showed that the influence of effective density on train ridership is statistically significant. Thus, the model including agglomeration gave a more comprehensive prediction than the model without agglomeration. Both effective employed resident density and job density are statistically significant for the train trip attraction model, but only effective job density is statistically significant for train trip production.

- Comparison of independent variables determining train ridership indicated that the accessibility/transportation variable has a greater influence than land use and socio demographic/economic variables. Land use variables in turn have a greater influence than the socio demographic/economic variables. Job density was also observed to influence train ridership more than employed resident density.

- This thesis proposes a shorter threshold distance or the geographical extent of agglomeration in the train trip attraction model than in the train trip production model. This may suggest that station access is more important for train trip attraction than for train trip production. Thus, it is also proposed that, in order to enhance train ridership, jobs be added to suburbs located near the station precincts or within the TOD precincts, rather than in the more distant suburbs.

- The interaction of effective density and distance between suburbs and train stations has been shown to be statistically significant. This interaction term in the regression model also represents the distance decay of effective density around train station. Thus, increasing effective density without considering the distribution of residences or workplaces from train stations would not achieve optimum agglomeration benefits for train ridership. For example, increasing jobs in locations near the station would improve ridership attraction, but increasing jobs in locations further from the station would not be effective.

- In terms of business sector comparisons, the magnitude of effective density for the retail sector decays more rapidly with travel time between suburbs, than that for the construction and manufacturing sectors. This means that at increasingly greater travel time (longer commuting time), the spatial interaction between employed residents and their jobs retail sectors would tend to decrease more rapidly than that for workers in manufacturing and construction.

- Using the concept of trade-off between travel costs and housing costs as the underlying mechanism, this research has suggested that the accessibility improvement in terms of travel time savings is much greater for suburbs where jobs are concentrated (job-rich suburbs). This has been demonstrated by the higher premiums on land rents, wages, and effective density of job-rich suburbs compared to other suburbs. This greater travel time saving may explain why high premiums for effective job and employed resident densities in job rich suburbs area relate to higher train ridership.

- Wage equation modelling suggests that job-housing balanced suburbs have the lowest effect in both attracting and generating train ridership. The wage equation calculates the value of travel time saving. Due to higher values of travel time saving, those who work in job-rich suburbs are willing to pay more to save one minute of travel time. This also means that their ability to pay higher transport costs may encourage them to live farther from the centre than those who work in the reference category (job-housing balance). This also makes it possible to spend longer commuting by train. Conversely, those who work in employed resident-rich suburbs have the lowest values of travel time saving; hence they have less willingness to pay more on travel costs than those who work in the reference category. This may deter longer commuting by train.

9. 4 FACTORS DETERMINING TRAIN RIDERSHIP PREDICTION

This thesis has shown that train ridership is not only influenced by factors relating to land use, accessibility and socio-economic/demographic variables, as found in many previous studies, but also spatial economic factors, as measured by effective density, which have been omitted in recent modelling of train ridership prediction.

The effective density variable does not eliminate the influence of density, but has been shown to strengthen the influence of the density variable on train ridership. Model validation by back-testing has shown that predictions under the LUTI framework tend to underestimate the actual proportion of train ridership, while the SETI framework tends to overestimate the actual proportion of train ridership. However, the level of error of the model under the SETI framework is lower than that of the LUTI framework. Thus, the tendency of underestimating the magnitude of train ridership may be due to omitting the spatial economic factors of the effective density variable in modelling train ridership prediction.

In addition, job density is statistically significant in all models of train trip attraction. Comparing the relative effects of job density and effective employed resident density on train trip attraction, based on the LUTI model and the SETI h1b (effective employed resident density) model with the interaction term, job density was shown to have a greater influence than effective employed resident density. However, the influence of job density only dissipated beyond a considerable distance from stations (from 24 km to 142 km) which is beyond the constraints of the model, assumed to be 16 km maximum distance, and may not be an empirically real result. Thus, there is a limitation when train trip attraction is modelled under the LUTI framework to assess the influence of job density. The influence of job density will always appear to be strongly positive wherever the location of employment is relative to the stations, which is not evidenced in reality. The SETI model produced more realistic results by applying plausible threshold distances, where at this threshold distance the effective density will no longer influence the magnitude of train ridership attraction. Policy that relies on the SETI model will have spatial considerations for meeting certain constraints (such as distribution of residential areas and job centres relative to the location of stations) and may be able to optimize the benefits of policy scenarios based on the advantages of effective density on train ridership.

Other factors mentioned in the literature to be strong influences on train ridership have also been found to be important in this study. Transportation factors such as the street network travel distance variable and the average distance of suburbs from train stations have been very significant, where in some models, their influence outweighs other factors. Car ownership and income of managers/professional employed residents was found to negatively influence train trip production, which is also in line with the findings of other studies. However, the proportion of blue collar employed residents was also shown to inversely influence the proportion of train trip production. The analysis in chapter 6 indicated that the spatial distribution of blue collar employed residents in locations underserved by railway line services may be the cause of this negative relationship. Wages of managers/professionals positively influence train trip attraction, suggesting that places of work that offer higher wages tend to attract higher levels of train ridership. It appears that all the partial or indirect urban productivity variables (wage, commercial (non-residential) land rent, effective density) significantly influence train ridership, especially train trip attraction in the same direction of influence, where high productivity location would attract higher train ridership. This also has been shown by the consistency in the pattern of wage, land rent and effective density premiums, which tend to be much higher in place-of-work or job-rich suburbs than other suburb categories (employed resident-rich suburbs or job-housing balance suburbs).

9. 5 SUPPORT FOR RESEARCH HYPOTHESISES

Four research hypotheses have been developed to fulfil the aims of this thesis and address the problem statements. These hypotheses have been addressed in chapters 5, 6, and 7, and the policy implications have been discussed in chapter 8, from which the following conclusions are drawn:

In relation to the first hypothesis, there is evidence in support of the presence of public transport agglomeration, measured as the magnitude of effective density. The changes in effective density before-and-after the railway line extension in the study area are strongly correlated with the changes in travel times and the changes in the level of train ridership. These changes are also higher in suburbs located directly adjacent to the Perth-Mandurah line than those adjacent to other lines. Furthermore, clusters of high values of effective density, as shown by the Getis-Ord G_i* values, have emerged in suburbs near stations. On average, the percentage of effective density aggregated over suburbs from fishnet based data is higher in buffer rings near stations. These percentages reduce gradually with distance of suburbs from stations, thus showing a distance decay pattern of agglomeration (chapter 5).

In relation to the second research hypothesis, the effective density may be viewed as complementary to or an extension of the density parameter in measuring travel demand in terms of the opportunity of market potential. This has been tested and confirmed for train ridership prediction by comparing models with and without the agglomeration component. This thesis has demonstrated that both effective density and density parameters can be used to predict train ridership but the inclusion of an effective density variable increases the model accuracy (chapter 6).

In relation to the third research hypothesis, the distance of suburbs from stations has been shown to negatively influence the relationship between effective density and train ridership. The greater the distance, the less effective density tends to influence train ridership. This has been shown for the train ridership attraction model. However, in the train production model, this result was the opposite, suggesting a possible influence on travel times due to park and ride data in the model. The greater the distance, the more effective density tends to influence train production.

In order to determine whether or not there was a bias toward travel time by park and ride, an additional model was carried out to explore the influence of effective density when car accessibility only was used in the calculation of effective density, rather than the travel time by park and ride. The additional model had been shown in appendix 60 and an example was carried out for the model for all sectors SETI h1, SETI h2, and SETI h1B and h2B. The model with the interaction term was not significant (SETI h1B and SETI h2B). The model of effective employed resident density (SETI h1) was not significant. However, effective job density (calculated with car access or travel time by car) had a significant influence on train trip production. However, this variable has a negative or inverse relationship with train ridership with a standardized beta value of -0.17. Exponentiation of this number resulted in the value of 0.844. This means there was a reduction of 16% of the proportion of train ridership for every one-unit increase in effective job density based on car travel time. This additional model shows that there would be different conclusion when different travel times were used in the calculation of effective density. The results from the calculation of effective density by park and ride suggested that there is a positive influence of agglomeration on train use. On the other hand, the influence of agglomeration calculated based on travel time by car was negative on train ridership, thus, the use of car travel time based agglomeration would be best used in the prediction of car use as a dependent variable.

Focusing on the comparison of the influence of geographical extent of agglomeration between sectors, this thesis found that the manufacturing sector has a larger effective employed resident density, which is distributed over a larger geographical area (when measured from train stations), compared to the retail and the construction sectors. The retail sector, in contrast, contributes lower multiplicative effects on ridership, which are confined to the smallest geographical areas around stations. The comparison among the three sectors suggested that the construction sector has the greatest influence of effective job density on train trip attraction. Its multiplicative effects are similar to those of the manufacturing sector, but cover a much smaller geographical area than those the manufacturing sector (chapter 7).

In relation to the fourth research hypothesis, the level of train trip attraction is highest in job-rich suburbs, followed by employed resident-rich suburbs and the least in job-housing balanced suburbs. Thus, employed resident-rich suburbs may still generate train trip attraction as long as the location of these suburbs is very close to train stations, even though the number of jobs is low compared to the number of employed residents. However, the optimum benefits of agglomeration on train ridership may be achieved by retaining a policy of increasing the number of jobs in areas within station precincts. Thus, policy designed to convert residential activities or create more job activities in employed resident-rich suburbs located near stations may be effective for encouraging train usage.

On the other hand, train trip production is highest in the employed resident-rich suburbs, followed by job-rich suburbs and the lowest in job-housing balanced suburbs. Thus, job-rich suburbs may still generate a relatively higher number of train trips compared to job-housing balanced suburbs. Job-housing balanced suburbs are distributed widely over the study area, sometimes farther from stations and with less access to railway services. Suburbs with less development in both employed resident facilities and in jobs technically may also be among the job-housing balanced suburbs. Low access to stations in a geographical space and low levels of development for both jobs and residential activities may be reasons why these suburbs generate and attract the lowest train production and attraction compared to the other suburb categories.

9. 6 LIMITATIONS OF THE THESIS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Some caution in modelling train ridership under the influence of effective density should be noted, as arising from this study.

This thesis is limited in several aspects. For example, data availability for the place-of-work datasets were comprised of different geographical boundary units to the unit of analysis used in this research. This raised issues regarding conversion and aggregation problems that may have reduced the data accuracy. Careful consideration of the GIS methods and procedures were needed to ensure the accuracy of the conversion results.

The best available travel time dataset representing the dependent variable was limited to the measurement of travel time via park and ride services. The train trip production model seems to reflect the dominant influence of travel time park and ride instead of general train users' behaviour to some extent, where the influence of effective density on train ridership is larger with increased distance of the residential suburbs from stations. Future research may involve a detailed analysis of train ridership based on each particular mode of travel separately, such as train only, train plus car as drivers/passengers or park and ride, train plus buses or bus and ride, or travel time by car to predict the level of car use, and so on. Consideration of travel times that match the specific modes in question will increase the rigour of the model results.

The method of train ridership prediction which used limited dependent variables and log transformation of the dependent variable required non-zero values of the dependent variable. However, the dataset for the study area contains some suburbs with zero values. Deleting or eliminating these suburbs compromised the number of observations and not enough data as a result may incure a more serious problem. Problems of outliers also emerged in some variables such as for accessibility measurement and for the number of jobs, where the Perth CBD and nearby suburbs tend to have high values compared to other suburbs. Nevertheless, the utilization of limited dependent variables largely reduces the problem of outliers, and non-equal variances or homogeneity of data, if otherwise the dependent variable had not been transformed, although it does not completely eliminate this problem.

The fact that the study involves an aggregate approach may also represent a limitation of this thesis. The assumption is that there is a homogeneous distribution of variables within a suburb, for example, for distance to station and employment size. However, the accessibility to a station is not homogeneous within a suburb in reality, nor are the numbers of jobs. The approach of using fishnet-based dataset was designed reduce this problem. However, a more rigorous technique would be required to reveal any variability that represents a more accurate level of detail within a suburb area.

While the aggregate approach, by using administrative boundaries such as suburb boundaries, is common in train ridership prediction, the inclusion of an effective density variable may require that model prediction is better approached using a disaggregate method. A disaggregate approach would require a large amount of microdata for employment, travel and other variables accordingly, such as might be provided by a parcel-level property value dataset. Similarly, the identification of agglomeration based on many more divisions of employment may result in a more detailed analysis than the few sectors adopted in this thesis. In this regard, the emphasis on the design factor of land use components, such as stated in the TOD physical design requirements or guidelines (Renne, 2007), would be an important aspect to consider in addition to the density and mixed used (diversity) factors.

This thesis has assumed that the non-normal distribution of data was largely influence by outliers and that the decision to transform the non-normal distribution of data into normal distribution may introduce some weaknesses. A rule that ±2 standard deviations automatically be considered as outliers, as stated in Stopher (2012) may be viewed as being less conservative or less rigorous. Leys, Ley, Klein, Bernard, and Licata (2013) proposed the absolute deviation around the median as a better method to deal with the problem of outliers. Future research may consider this method for detecting outliers in a more rigorous way.

While normality in explanatory variables is an important property of data distribution, it is not the dominant consideration in determining variable functional form. Hypotheses concerning the relationship between the dependent and independent variable should also enter into the model specification. For example, it was difficult to make behavioural sense that the dependent variable should be regressed against a logarithmic or square-root transformation of an explanatory variable. The model has been re-tested by using the non-transformed data. The results, as shown in appendix 59, lead to a conclusion that the transformed data are preferable than the non-transformed data. The regression models with the non-transformed data for the train trip production model lead to high multicollinearity among the independent variables and a larger area of the geographical extent of agglomeration, which are non-intuitive with regards to TOD policy. The regression models with the non-transformed data for the train trip attraction model produced much lower adjusted R-square values, which indicated a lesser power of explanation in the model compared to the models with transformed data. Similarly, the train trip attraction model also produced larger areas for the geographical extent of agglomeration.

9. 7 CONCLUSION

The impact of transportation investment on train ridership, measured as agglomeration (effective density), suggests a strong link between the spatial economic system, transportation system, and land use. This thesis has demonstrated that the SETI framework can be used in a prediction of train ridership, accommodating this spatial economic dimension and thereby increasing model accuracy. The effects of agglomeration prevail predominantly in areas that gain an accessibility premium where jobs are concentrated and wages/land rent premiums emerge. Suburbs near stations gain more advantages and the effects of agglomeration externalities on increasing train ridership are likely to prevail when job development is prioritized over residential activities. This may also mean that the model of train trip prediction under the SETI framework can be of best use in predicting train trip attraction (to the place of work).

Effective density is a spatial economic extension of the density variable. This indicator of agglomeration cannot fully replace the density variable, and this study does not address alternative models in which effective density replaces density. The advantage of adding effective density is that it includes analysis of urban productivity and other spatial economic dimensions as determinants of train ridership, thus reducing the tendency of underestimate the impact of railway line extension on the increase in train ridership. This is similar to the use of WEBs as an alternative to the CBA in assessing the impact of transportation infrastructure investments, where the WEBs method includes more comprehensive monetary impacts, resulting in higher GDP effects from the transportation investments than the CBA method.

REFERENCES

AHURI. (2006). Are housing affordability problems creating labour shortages? *AHURI Research & Policy Bulletin,* (74). Retrieved from https://www.ahuri.edu.au/__data/assets/pdf_file/0024/2976/AHURI_RAP_Issue_74_Are_housing_affordability_problems_creating_labour_shortages.pdf

Alpkokin, P., Cheung, C., Black, J., & Hayashi, Y. (2008). Dynamics of clustered employment growth and its impacts on commuting patterns in rapidly developing cities. *Transportation Research Part A: Policy and Practice, 42*(3), 427-444. doi:DOI: 10.1016/j.tra.2007.11.002

Arrington, G. B., & Cervero, R. (2008). *TCRP Report 128: Effects of TOD on Housing, Parking, and Travel*. Retrieved from Washington, DC: http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_128.pdf

Australian Bureau of Statistics (ABS). (2010). *Survey of Motor Vehicle Use 12 Months Ended 31 October 2010*. (9208.0). Canberra, ACT.: ABS. Retrieved from www.abs.gov.au.

Australian Bureau of Statistics (ABS). (2011). ABS Census Dictionary 2901.0 2011. Retrieved from www.abs.gov.au.

Badoe, D. A., & Miller, E. J. (2000). Transportation–land-use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research Part D: Transport and Environment, 5*(4), 235-263. doi:http://dx.doi.org/10.1016/S1361-9209(99)00036-X

Bamford, G. (2009). Urban Form and Housing Density, Australian Cities and European Models: Copenhagen and Stockholm Reconsidered. *Urban Policy and Research, 27*(4), 337 - 356. Retrieved from http://www.informaworld.com/10.1080/08111140903108572

Bartholomew, K. (2007). Land use-transportation scenario planning: promise and reality. *Transportation, 34*(4), 397. doi:10.1007/s11116-006-9108-2

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2003). *The Impact of Urban Spatial Structure on Travel Demand in the United States*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=636369

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2005). The Effects of Urban Spatial Structure on Travel Demand in the United States. *Review of Economics and Statistics, 87*(3), 466-478. doi:10.1162/0034653054638292

Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics, 7*(1), 1.

Bertolini, L., Clercq, F. l., & Straatemeier, T. (2008). Urban transportation planning in transition. *Transport Policy, 15*(2), 69-72. doi:DOI: 10.1016/j.tranpol.2007.11.002

Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B: Methodological, 43*(7), 749-765. doi:DOI: 10.1016/j.trb.2009.02.001

Bhat, C. R., & Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels.

Transportation Research Part B: Methodological, 41(5), 506-526. doi:DOI: 10.1016/j.trb.2005.12.005

Bhat, C. R., Sen, S., & Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B: Methodological, 43*(1), 1-18. doi:DOI: 10.1016/j.trb.2008.06.009

Blainey, S. (2010). Trip end models of local rail demand in England and Wales. *Journal of Transport Geography, 18*(1), 153-165. doi:http://dx.doi.org/10.1016/j.jtrangeo.2008.11.002

Blischke, W., Rezaul Karim, M., & Prabhakar Murthy, D. N. (2011). Preliminary Data Analysis *Warranty Data Collection and Analysis* (pp. 159-189): Springer London.

Boarnet, M. G., & Crane, R. (2001). *Travel by Design*. New York: Oxford University Press.

Boustan, L. (2009). Job Decentralization and Residential Location. *Brookings-Wharton Papers on Urban Affairs, 2009*(1), 1-31.

Brinkman, J. (2011). *Essays on urban agglomeration.* (Doctor of Philosophy Economics and Public Policy Dissertation), Carnegie Mellon University, United States. (UMI Number 3519888)

Brons, M., Givoni, M., & Rietveld, P. (2009). Access to railway stations and its potential in increasing rail use. *Transportation Research Part A: Policy and Practice, 43*(2), 136-149. doi:DOI: 10.1016/j.tra.2008.08.002

Bureau of Transport and Regional Economics [BTRE]. (2007). *Estimating urban traffic and congestion cost trends for Australian cities*. Retrieved from Canberra ACT.: https://bitre.gov.au/publications/2007/files/wp_071.pdf

Buxton, M., & Taylor, E. (2010). Urban Land Supply, Governance and the Pricing of Land. *Urban Policy and Research, 29*(01), 5-22. Retrieved from http://dx.doi.org/10.1080/08111146.2011.537605

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009). The relationship between the built environment and nonwork travel: A case study of Northern California. *Transportation Research Part A: Policy and Practice, 43*(5), 548-559. doi:DOI: 10.1016/j.tra.2009.02.001

Catanese, A. J. (1988). Transit, Urban Life and Development in the Sunbelt. In W. Attoe (Ed.), *Transit, Land Use & Urban Form* (pp. 189-196). Austin, Texas: Centre for the Study of American Architecture.

Cervero, R. (1994). Transit-based housing in California: evidence on ridership impacts. *Transport Policy, 1*(3), 174-183. Retrieved from http://www.sciencedirect.com/science/article/B6VGG-4695XG2-4/2/a2b15b146ec21326721dcc836197f646

Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American Housing Survey. *Transportation Research Part A: Policy and Practice, 30*(5), 361-377. Retrieved from http://www.sciencedirect.com/science/article/B6VG7-3VV41D6-4/2/1691d493ffd69fbe7ecc5cc754c8d495

Cervero, R. (2002). Built environments and mode choice: toward a normative framework. *Transportation Research Part D: Transport and Environment, 7*(4), 265-284. Retrieved from http://www.sciencedirect.com/science/article/B6VH8-45CC0XM-2/2/c72e6843738db50b9211eae51815edc5

Cervero, R. (2006). Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth. *Journal of the American Planning Association, 72*(3), 285-295. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=21924125&site=ehost-live

Cervero, R., & Duncan, M. (2006). Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing. *Journal of American Planning Association, 72*(4), 475-490.

Cervero, R., & Kang, C. D. (2011). Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy, 18*(1), 102-116. doi:DOI: 10.1016/j.tranpol.2010.06.005

Cervero, R., & Landis, J. (1993). Assessing the impacts of urban rail transit on local real estate markets using quasi-experimental comparisons. *Transportation Research Part A: Policy and Practice, 27*(1), 13-22. Retrieved from http://www.sciencedirect.com/science/article/B6VG7-469GHYJ-1W/2/7c60c2ea15139c7f00f1bdd41fdf6ae0

Cervero, R., & Landis, J. (1997). Twenty years of the Bay Area Rapid Transit system: Land use and development impacts. *Transportation Research Part A: Policy and Practice, 31*(4), 309-333. Retrieved from http://www.sciencedirect.com/science/article/B6VG7-3SWXY3H-4/2/4130782ace6540bfa809718a681f4504

Cervero, R., Murakami, J., & Miller, M. A. (2009). Direct Ridership Model of Bus Rapid Transit in Los Angeles County. Retrieved from http://www.its.berkeley.edu/sites/default/files/publications/UCB/2009/VWP/UCB-ITS-VWP-2009-3.pdf

Cervero, R., Murphy, S., Ferrel, C., Goguts, N., & Tsai, Y.-H. (2004). *TCRP Report 102: Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects* (102). Retrieved from Washington, DC: http://www.trb.org/Publications/Blurbs/154989.aspx

Cervero, R., Round, A., Goldman, T., & Wu, K.-L. (1995). Rail Access Modes and Catchment Areas for the BART System. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.1103&rep=rep1&type=pdf

Chan, S., & Miranda-Moreno, L. (2013). A station-level ridership model for the metro network in Montreal, Quebec. *Canadian Journal of Civil Engineering, 40*(3), 254-262. doi:10.1139/cjce-2011-0432

Chang, J. S., & Mackett, R. L. (2006). A bi-level model of the relationship between transport and residential location. *Transportation Research Part B: Methodological, 40*(2), 123-146. doi:DOI: 10.1016/j.trb.2005.02.002

Chatman, D. G. (2008). Deconstructing development density: Quality, quantity and price effects on household non-work travel. *Transportation Research Part A: Policy and Practice, 42*(7), 1008-1030. doi:DOI: 10.1016/j.tra.2008.02.003

Chen, C., Gong, H., & Paaswell, R. (2008). Role of the built environment on mode choice decisions: additional evidence on the impact of density. *Transportation, 35*(3), 285-299. doi:10.1007/s11116-007-9153-5

Chen, C., Varley, D., & Chen, J. (2011). What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour. *Urban Studies, 48*(9), 1893-1908. doi:10.1177/0042098010379280

Chiang, W.-C., Russell, R. A., & Urban, T. L. (2011). Forecasting ridership for a metropolitan transit authority. *Transportation Research Part A: Policy and Practice, 45*(7), 696-705. doi:http://dx.doi.org/10.1016/j.tra.2011.04.016

Cho, S.-H., Poudyal, N., & Lambert, D. M. (2008). Estimating spatially varying effects of urban growth boundaries on land development and land value. *Land Use Policy, 25*(3), 320-329. doi:DOI: 10.1016/j.landusepol.2007.08.004

Choocharukul, K., Van, H. T., & Fujii, S. (2008). Psychological effects of travel behavior on preference of residential location choice. *Transportation Research Part A: Policy and Practice, 42*(1), 116-124. doi:DOI: 10.1016/j.tra.2007.06.008

Chowdhury, S., Ceder, A., & Schwalger, B. (2015). The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. *Journal of Transport Geography, 43*(0), 151-159. doi:http://dx.doi.org/10.1016/j.jtrangeo.2015.01.009

Commins, N., & Nolan, A. (2011). The determinants of mode of transport to work in the Greater Dublin Area. *Transport Policy, 18*(1), 259-268. doi:DOI: 10.1016/j.tranpol.2010.08.009

Currie, G., & Delbosc, A. (2010). Modelling the social and psychological impacts of transport disadvantage. *Transportation, 37*(6), 953. Retrieved from http://proquest.umi.com.dbgw.lis.curtin.edu.au/pqdweb?did=2191412271&Fmt=7&clientId=22212&RQT=309&VName=PQD

Curtis, C. (2005). The windscreen world of land use transport integration: experiences from Perth, a dispersed city. *Town Planning Review, 76*(4), 423-453.

Curtis, C. (2008). Evolution of the Transit-oriented Development Model for Low-density Cities: A Case Study of Perth's New Railway Corridor. *Planning Practice and Research, 23*(3), 285 - 302. Retrieved from http://www.informaworld.com/10.1080/02697450802423559

Curtis, C. (2012). Transitioning to Transit-Oriented Development: The Case of Perth, Western Australia. *Urban Policy and Research, 30*(3), 275-292. doi:10.1080/08111146.2012.665364

Curtis, C., & Mellor, R. (2011). Anticipating a New Railway: The Relevance of Transit Oriented Development to Bussiness in Car Oriented City. *Urban Policy and Research, 29*(1), 1-26.

Curtis, C., & Olaru, D. (2011). *The relationship between accessibility and travel patterns of residents in three new station precincts*. Paper presented at the The 3rd World Planning Schools Congress, Perth (Western Australia), 4-8 July 2011.

Cushman, K. (1988). Exploring the Land Development and Transit Connection. In W. Attoe (Ed.), *Transit, Land Use & Urban Form* (pp. 9-36). Austin, Texas: Centre for the Study of American Architecture.

Dargay, J. (2007). The effect of prices and income on car travel in the UK. *Transportation Research Part A: Policy and Practice, 41*(10), 949-960. doi:DOI: 10.1016/j.tra.2007.05.005

de Grange, L., Troncoso, R., & González, F. (2012). An empirical evaluation of the impact of three urban transportation policies on transit use. *Transport Policy, 22*(0), 11-19. doi:http://dx.doi.org/10.1016/j.tranpol.2012.04.003

Dekle, R., & Eaton, J. (1999). Agglomeration and Land Rents: Evidence from the Prefectures. *JOURNAL OF URBAN ECONOMICS, 46*(2), 200-214. doi:10.1006/juec.1998.2118

Dell, M. (2009). GIS Analysis for Applied Economists: Manuscript. Retrieved from http://scholar.harvard.edu/files/dell/files/090110combined_gis_notes.pdf

Deng, T. (2013). Impacts of Transport Infrastructure on Productivity and Economic Growth: Recent Advances and Research Challenges. *Transport Reviews, 33*(6), 686-699. doi:10.1080/01441647.2013.851745

Department of Transport (DOT) of Western Australia. (2011). *Public Transport for Perth in 2031: Mapping out the future for Perth's public transport network*. Perth, WA. Retrieved from www.transport.wa.gov.au.

Department of Transport, S. G. V. (2012). *Job density, productivity and the role of transport: an overview of agglomeration benefits from investments and implications for the transport portfolio.* . Retrieved from http://www.transport.vic.gov.au/__data/assets/pdf_file/0005/74228/Job-density-productivity-and-the-role-of-transport.pdf

Dhrymes, P. J. (1986). Limited Dependent Variables. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of Econometrics* (Vol. III): Elsevier Science Publishers BV.

Di Addario, S., & Patacchini, E. (2008). Wages and the City. Evidence from Italy. *Labour Economics, 15*(5), 1040-1061. doi:http://dx.doi.org/10.1016/j.labeco.2007.09.003

Diana, M., & Mokhtarian, P. L. (2009). Grouping travelers on the basis of their different car and transit levels of use. *Transportation, 36*(4), 455-467. doi:10.1007/s11116-009-9207-y

Domencich, T. A., & McFadden, D. (1975). *Urban Travel Demand: A behavioural analysis*. Amsterdam . Oxford . New York: North-Holland Publishing Company, American Elsevier Publishing Company.

Du, H., & Mulley, C. (2007). The short-term land value impacts of urban rail transit: Quantitative evidence from Sunderland, UK. *Land Use Policy, 24*(1), 223-233. doi:DOI: 10.1016/j.landusepol.2005.12.003

Ellis, C., & Parolin, B. (2010, 29 September-1 October 2010). *Does increased residential density around train stations encourage more environmentally sustainable travel behaviour?* Paper presented at the Australasian Transport Research Forum, Canberra, Australia.

Eluru, N., Bhat, C., Pendyala, R., & Konduri, K. (2010). A joint flexible econometric model system of household residential location and vehicle fleet composition/usage choices. *Transportation, 37*(4), 603. Retrieved from http://proquest.umi.com.dbgw.lis.curtin.edu.au/pqdweb?did=2045910571&Fmt=7&clientId=22212&RQT=309&VName=PQD

Evans, J. E. J., Pratt, R. H., Stryker, A., & Kuzmyak, J. R. (2007). *TCRP Report 95, Chapter 17: Transit Oriented Development*. Retrieved from Washington, DC: http://www.trb.org/Publications/Blurbs/159049.aspx

Ewing, R. (1997). Is Los Angeles-Style Sprawl Desirable? *Journal of American Planning Association, 63*(1), 107 - 126.

Ewing, R., & Cervero, R. (2001). Travel and the Built Environment: A Synthesis. *Transportation Research Record: Journal of the Transportation Research Board, 1780*, 87-114. doi:10.3141/1780-10

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association, 76*(3), 265-294. doi:10.1080/01944361003766766

Falconer, R., Newman, P., & Giles-Corti, B. (2010). Is practice aligned with the principles? Implementing New Urbanism in Perth, Western Australia. *Transport Policy, 17*(5), 287-294. doi:DOI: 10.1016/j.tranpol.2010.01.014

Ferguson, E. (2000). *Travel Demand Management and Public Policy*. England

USA: Ashgate Publishing Ltd, Ashgate Publishing Company.

Fleishman, D., Wensley, J., Halvorsen, R., Englisher, L., Falbel, S., Schweiger, C., . . . Crikelair, T. (2007). *TCRP Report 111: Elements Needed to Create High Ridership Transit Systems*. Retrieved from Washington, DC: http://www.tcrponline.org/PDFDocuments/TCRP_RPT_111.pdf

Fujita, M., & Thisse, J.-F. (2002). *Economics of agglomeration*. The United KIngdom: The University Press, Cambridge.

Funderburg, R. G., Nixon, H., Boarnet, M. G., & Ferguson, G. (2010). New highways and land use change: Results from a quasi-experimental research design. *Transportation Research Part A: Policy and Practice, 44*(2), 76-98. doi:http://dx.doi.org/10.1016/j.tra.2009.11.003

García-Palomares, J. C. (2010). Urban sprawl and travel to work: the case of the metropolitan area of Madrid. *Journal of Transport Geography, 18*(2), 197-213. doi:DOI: 10.1016/j.jtrangeo.2009.05.012

Getis, A., & Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis, 24*(3), 189-206. doi:10.1111/j.1538-4632.1992.tb00261.x

Geurs, K., Zondag, B., de Jong, G., & de Bok, M. (2010). Accessibility appraisal of land-use/transport policy strategies: More than just adding up travel-time savings. *Transportation Research Part D: Transport and Environment, 15*(7), 382-393. doi:DOI: 10.1016/j.trd.2010.04.006

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography, 12*(2), 127-140. doi:http://dx.doi.org/10.1016/j.jtrangeo.2003.10.005

Giuliano, G. (1991). Is Jobs-Housing Balance a Transportation Issue? *Transportation Research Record, 1305*, 305-312.

Gonçalves, J. A. M., Portugal, L. d. S., & Nassi, C. D. (2009). Centrality indicators as an instrument to evaluate the integration of urban equipment in the area of influence of a rail corridor. *Transportation Research Part A: Policy and Practice, 43*(1), 13-25. doi:DOI: 10.1016/j.tra.2008.06.010

Graham, D. J. (2007). Agglomeration, Productivity and Transport Investment. *Journal of Transport Economics and Policy, 41*(3), 317-343. doi:10.2307/20054024

Graham, D. J., Gibbons, S., & Martin, R. (2009). Transport investment and the distance decay of agglomeration benefits. Retrieved from http://personal.lse.ac.uk/gibbons/Papers/Agglomeration%20and%20Distance%20Decay%20Jan%202009.pdf

Graham, D. J., & Melo, P. C. (2010). *Advice on the Assessment of Wider Economic Impacts: a report for HS2*. Retrieved from http://webarchive.nationalarchives.gov.uk/+/http:/www.dft.gov.uk/pgr/rail/pi/highspeedrail/hs2ltd/appraisalmaterial/pdf/wídereconomicreport.pdf

Grant, J. L. (2009). Theory and Practice in Planning the Suburbs: Challenges to Implementing New Urbanism, Smart Growth, and Sustainability Principles. *Planning Theory & Practice, 10*(1), 11 - 33. Retrieved from http://www.informaworld.com/10.1080/14649350802661683

Greenberg, K. (1988). Transit and Changing Urban Character. In W. Attoe (Ed.), *Transit, Land Use & Urban Form* (pp. 47-54). Austin, Texas: Center for the Study of American Architecture.

Grengs, J. (2010). Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography, 18*(1), 42-54. doi:http://dx.doi.org/10.1016/j.jtrangeo.2009.01.012

Guerra, E., & Cervero, R. (2011). Cost of a Ride. *Journal of the American Planning Association, 77*(3), 267-290. doi:10.1080/01944363.2011.589767

Guillain, R., & Le Gallo, J. (2007). Agglomeration and dispersion of economic activities in Paris and its surroundings : an Exploratory spatial data analysis. Retrieved from http://leg2.u-bourgogne.fr/documents-de-travail/e2007-01.pdf

Guo, J. Y., & Bhat, C. R. (2007). Operationalizing the concept of neighborhood: Application to residential location choice analysis. *Journal of Transport Geography, 15*(1), 31-45. doi:DOI: 10.1016/j.jtrangeo.2005.11.001

Gutiérrez, J., Cardozo, O. D., & García-Palomares, J. C. (2011). Transit ridership forecasting at station level: an approach based on distance-decay weighted regression. *Journal of Transport Geography, 19*(6), 1081-1092. doi:http://dx.doi.org/10.1016/j.jtrangeo.2011.05.004

Hall, P. (1969). Transportation. *Urban Studies, 6*(3), 408-435. doi:10.1080/00420986920080511

Hamidi, S., & Ewing, R. (2014). A Longitudinal Study of Changes in Urban Sprawl between 2000 and 2010 in the United States. *Landscape and Urban Planning*. doi:10.1016/j.landurbplan.2014.04.021

Handy, S. (1996). Methodologies for exploring the link between urban form and travel behavior. *Transportation Research Part D: Transport and Environment, 1*(2), 151-165. doi:http://dx.doi.org/10.1016/S1361-9209(96)00010-7

Handy, S. (2008). Regional transportation planning in the US: An examination of changes in technical aspects of the planning process in response to changing goals. *Transport Policy, 15*(2), 113-126. doi:DOI: 10.1016/j.tranpol.2007.10.006

Hensher, D. A., & Puckett, S. M. (2007). Congestion and variable user charging as an effective travel demand management instrument. *Transportation Research Part A: Policy and Practice, 41*(7), 615-626. doi:DOI: 10.1016/j.tra.2006.07.002

Hensher, D. A., Rose, J. M., & Green, W. H. (2005). *Applied Choice Analysis: A Primer*. Cambridge, UK: Cambridge University Press.

Hensher, D. A., & Ton, T. (2002). TRESIS: A transportation, land use and environmental strategy impact simulator for urban areas. *Transportation, 29*(4), 439-457. doi:10.1023/A:1016335814417

Hensher, D. A., Truong, T. P., Mulley, C., & Ellison, R. (2012). *Assessing the wider economy impacts of transport infrastructure investment with an illustrative application to the nort-west rail link project in Sydney, Australia.* (ISSN 1832-570X). Retrieved from http://sydney.edu.au/business/__data/assets/pdf_file/0004/127786/ITLS-WP-12-05.pdf

Hess, S., Rose, J. M., & Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics and Transportation Review, 44*(5), 847-863. doi:DOI: 10.1016/j.tre.2007.06.002

Hickman, R., & Banister, D. (2005). Reducing Travel by Design: What About Change Over Time? In K. Williams (Ed.), *Sptaial planning, urban form and sustainable transport*. England: Ashgate Publishing Limited.

Holden, E. (2007). *Achieving Sustainable Mobility: Everyday and Leisure-time travel in the EU*. Aldershot: Ashgate.

Howell, D. C. (1997). *Statistical Methods for Psychology 4th ed.* the United States of America: Wadsworth Publishing Company, a Division of International Thomson Publishing Inc.

Iacono, M., Krizek, K. J., & El-Geneidy, A. (2010). Measuring non-motorized accessibility: issues, alternatives, and execution. *Journal of Transport Geography, 18*(1), 133-140. doi:DOI: 10.1016/j.jtrangeo.2009.02.002

Jonsson, R. D. (2008). Analysing sustainability in a land-use and transport system. *Journal of Transport Geography, 16*(1), 28-41. doi:DOI: 10.1016/j.jtrangeo.2007.02.006

Kain, J. F., & Liu, Z. (1999). Secrets of success: assessing the large increases in transit ridership achieved by Houston and San Diego transit providers. *Transportation Research Part A: Policy and Practice, 33*(7–8), 601-624. doi:http://dx.doi.org/10.1016/S0965-8564(99)00009-9

Kim, C. (2008). Commuting time stability: A test of a co-location hypothesis. *Transportation Research Part A: Policy and Practice, 42*(3), 524-544. doi:DOI: 10.1016/j.tra.2008.01.001

Kim, S., & Ulfarsson, G. (2008). Curbing automobile use for sustainable transportation: analysis of mode choice on short home-based trips. *Transportation, 35*(6), 723-737. doi:10.1007/s11116-008-9177-5

Kolko, J. (2011). *Making the Most of Transit: Density, Employment Growth, and Ridership around New Stations*. Retrieved from http://www.ppic.org/content/pubs/report/R_211JKR.pdf

Kryvobokov, M. (2007). Analysing location attributes with a hedonic model for apartment prices in Donetsk, Ukraine.(Report). *International journal of strategic property management, 11*(3), 157.

Lakshmanan, T. R. (2011). The broader economic consequences of transport infrastructure investments. *Journal of Transport Geography, 19*(1), 1-12. doi:http://dx.doi.org/10.1016/j.jtrangeo.2010.01.001

Lee, B., & Waddell, P. (2010). Residential mobility and location choice: a nested logit model with sampling of alternatives. *Transportation, 37*(4), 587. Retrieved from http://proquest.umi.com.dbgw.lis.curtin.edu.au/pqdweb?did=2045910561&Fmt=7&clientId=22212&RQT=309&VName=PQD

Levine, J., & Frank, L. (2007). Transportation and land-use preferences and residents' neighborhood choices: the sufficiency of compact development in the Atlanta region. *Transportation, 34*(2), 255-274. doi:10.1007/s11116-006-9104-6

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764-766. doi:http://dx.doi.org/10.1016/j.jesp.2013.03.013

Limtanakool, N., Dijst, M., & Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography, 14*(5), 327-341. doi:DOI: 10.1016/j.jtrangeo.2005.06.004

Lin, J., & Long, L. (2008). What neighborhood are you in? Empirical findings of relationships between household travel and neighborhood characteristics. *Transportation, 35*(6), 739-758. doi:10.1007/s11116-008-9167-7

Lin, T., & Truong, T. P. (2012). Transport improvement, agglomeration effect and urban productivity: The case of Chinese cities. Retrieved from http://sydney.edu.au/business/__data/assets/pdf_file/0009/136188/ITLS-WP-12-12.docx.pdf

Lindsey, M., Schofer, J. L., Durango-Cohen, P., & Gray, K. A. (2011). The effect of residential location on vehicle miles of travel, energy consumption and greenhouse gas emissions: Chicago case study. *Transportation Research Part D: Transport and Environment, 16*(1), 1-9. doi:DOI: 10.1016/j.trd.2010.08.004

Liu, C. S., Qing. (2011). An empirical analysis of the influence of urban form on household travel and energy consumption. *Computers, Environment and Urban Systems, 35*, 347-357. doi:http://dx.doi.org/10.1016/j.compenvurbsys.2011.05.006

Lois, D., & López-Sáez, M. (2009). The relationship between instrumental, symbolic and affective factors as predictors of car use: A structural equation modeling approach. *Transportation Research Part A: Policy and Practice, 43*(9-10), 790-799. doi:DOI: 10.1016/j.tra.2009.07.008

Manaugh, K., Miranda-moreno, L., & El-geneidy, A. (2010). The effect of neighbourhood characteristics, accessibility, home-work location, and demographics on commuting distances. *Transportation, 37*(4), 627. Retrieved from http://proquest.umi.com.dbgw.lis.curtin.edu.au/pqdweb?did=2045910611&Fmt=7&clientId=22212&RQT=309&VName=PQD

Mare, D. C., & Graham, D. J. (2009). *Agglomeration elasticities in New Zealand*. Retrieved from New Zealand: http://www.nzta.govt.nz/resources/research/reports/376/docs/376.pdf

Marshall, A. (1920). *Principles of Economics: An Introductory Volume* (8th ed.). London: Macmillan.

McFadden, D. (2007). The behavioral science of transportation. *Transport Policy, 14*(4), 269-274. doi:DOI: 10.1016/j.tranpol.2007.07.001

McIntosh, J., Newman, P., Crane, T., & Mourtiz, M. (2011). *Discussion paper - Alternative Funding Mechanisms for Public Transport in Perth: The Potential Role of Value Capture*. Perth: Curtin University Sustainability Policy Institute, Corview Group and Committee for Perth. Retrieved from http://www.committeeforperth.com.au/pdf/Advocacy/Report%20-%20AlternativeFundingforPublicTransportinPerthDecember2011.pdf.

Mees, P., O'Connell, G., & Stone, J. (2008). Travel to Work in Australian Capital Cities, 1976–2006. *Urban Policy and Research, 26*(3), 363 - 378. Retrieved from http://www.informaworld.com/10.1080/08111140802311236

Melo, P. C., Graham, D. J., Levinson, D., & Aarabi, S. (2013). *Agglomeration, accessibility, and productivity: evidence for urbanized areas in the US*. Paper presented at the The Transportation Resarch Board 92 Annual Meeting. http://nexus.umn.edu/Papers/Agglomeration.pdf

Mohan, D. (2008). Mythologies, Metro Rail Systems and Future Urban Transport. *Economic and Political Weekly, 43*(4), 41-53

Mokhtarian, P. L., & Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B: Methodological, 42*(3), 204-228. doi:DOI: 10.1016/j.trb.2007.07.006

Næss, P. (2009). Relationships between residential location, transportation rationales and environmental consequences. *IOP conference series. Earth and environmental science, 6*(26), 262010.

Naess, P. (2006). *Urban structure matters: residential location, car dependence and travel behaviour*. New York/London: Routledge.

Nolan, A. (2010). A dynamic analysis of household car ownership. *Transportation Research Part A: Policy and Practice, 44*(6), 446-455. doi:DOI: 10.1016/j.tra.2010.03.018

Nurlaela, S., & Curtis, C. (2012). Modeling household residential location choice and travel behavior and its relationship with public transport accessibility. *Procedia Social and Behavioral Sciences, 54*, 56-64. doi:10.1016/j.sbspro.2012.09.725.

Ohnmacht, T., Götz, K., & Schad, H. (2009). Leisure mobility styles in Swiss conurbations: construction and empirical analysis. *Transportation, 36*(2), 243-265. doi:10.1007/s11116-009-9198-8

Olaru, D., Smith, B., Xia, J. C., & Lin, T. G. (2014). Travellers' Attitudes Towards Park-and-Ride (PnR) and Choice of PnR Station: Evidence from Perth, Western Australia. *Procedia-Social and Behavioral Sciences, 162*, 101-110.

Ord, J. K., & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis, 27*(4), 286-306. doi:10.1111/j.1538-4632.1995.tb00912.x

Ory, D. T., & Mokhtarian, P. L. (2009). Modeling the structural relationships among short-distance travel amounts, perceptions, affections, and desires. *Transportation Research Part A: Policy and Practice, 43*(1), 26-43. doi:DOI: 10.1016/j.tra.2008.06.004

Páez, A. (2006). Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. *Journal of Transport Geography, 14*(3), 167-176. doi:DOI: 10.1016/j.jtrangeo.2005.11.002

Pedhazur, E. J. (1982). *Multiple regression in behavioral research : explanation and prediction / Elazar J. Pedhazur* (2nd ed.. ed.). New York: New York : Holt, Rinehart, and Winston.

Pinjari, A. R., Bhat, C. R., & Hensher, D. A. (2009). Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B: Methodological, 43*(7), 729-748. doi:DOI: 10.1016/j.trb.2009.02.002

Pinjari, A. R., Pendyala, R., Bhat, C., & Waddell, P. (2007). Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation, 34*(5), 557-573. doi:10.1007/s11116-007-9127-7

Rawnsley, T., Davies, B., Szafraniec, J., & Ratnam, K. (2014). *Economic and land use impacts of transport projects*. Paper presented at the AITPM 2014 National Conference.

Renne, J. L. (2007). *Measuring the Performance of Transit-Oriented Developments in Western Australia*. Retrieved from Perth, WA: http://www.vtpi.org/renne_tod_performance.pdf

Rickwood, P., & Glazebrook, G. (2009). Urban Structure and Commuting in Australian Cities. *Urban Policy and Research, 27*(2), 171 - 188. Retrieved from http://www.informaworld.com/10.1080/08111140802433378

Rodríguez, D. A., & Mojica, C. H. (2009). Capitalization of BRT network expansions effects into prices of non-expansion areas. *Transportation Research Part A: Policy and Practice, 43*(5), 560-571. doi:DOI: 10.1016/j.tra.2009.02.003

Rose, J., & Beck, M. (2007). *Basic Quantitative Analysis for Management*. University of Sydney: Pearson Custom Publishing.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on The Nature and Sources of Agglomeration Economies. In J. V. H. a. J. F. Thisse (Ed.), Handbook of Regional and Urban Economics (Vol. 4, pp. 2120-2167): Elsevier B.V. doi:10.1016/S0169-7218(04)07049-2

Russo, F., & Musolino, G. (2012). A unifying modelling framework to simulate the Spatial Economic Transport Interaction process at urban and national scales. *Journal of Transport Geography, 24*(0), 189-197. doi:http://dx.doi.org/10.1016/j.jtrangeo.2012.02.003

Senbil, M., Kitamura, R., & Mohamad, J. (2009). Residential location, vehicle ownership and travel in Asia: a comparative analysis of Kei-Han-Shin and Kuala Lumpur metropolitan areas. *Transportation, 36*(3), 325-350. doi:10.1007/s11116-009-9195-y

Shiftan, Y., Outwater, M. L., & Zhou, Y. (2008). Transit market research using structural equation modeling and attitudinal market segmentation. *Transport Policy, 15*(3), 186-195. doi:DOI: 10.1016/j.tranpol.2008.03.002

So, K. (2001). The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices. *American journal of agricultural economics, 83*(4), 1036-1048.

Southworth, F. (2001). On the potential impacts of land use change policies on automobile vehicle miles of travel. *Energy Policy, 29*, 1271-1283.

Stopher, P. R. (2012). *Collecting, Managing, and Assessing Data Using Sample Surveys*. Cambridge, UK: Cambridge University Press.

Sung, H., & Oh, J.-T. (2011). Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities, 28*(1), 70-82. doi:http://dx.doi.org/10.1016/j.cities.2010.09.004

Taplin, R. H. (2016). Research involving limited dependent variables: issues in the literature and recommendations for improvement. *Quality and Quantity, 50*(5), 2121-2140. doi:10.1007/s11135-015-0254-5

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. *Transportation Research Part A: Policy and Practice, 43*(1), 60-77. doi:http://dx.doi.org/10.1016/j.tra.2008.06.007

Tennøy, A. (2010). Why we fail to reduce urban road traffic volumes: Does it matter how planners frame the problem? *Transport Policy, 17*(4), 216-223. doi:DOI: 10.1016/j.tranpol.2010.01.011

The Centre for Transit-Oriented Development. (2009). *Destinations Matter: Building Transit Success* Retrieved from Oakland, CA: http://www.reconnectingamerica.org/assets/Uploads/ctodwp_destinations_matter.pdf

Timothy, D., & Wheaton, W. C. (2001). Intra-Urban Wage Variation, Employment Location, and Commuting Times. *JOURNAL OF URBAN ECONOMICS, 50*(2), 338-366. doi:http://dx.doi.org/10.1006/juec.2001.2220

Van Acker, V., & Witlox, F. (2009). Why land use patterns affect travel behaviour (or not). *I*. Retrieved from http://belgeo.revues.org/8777 website:

Van Acker, V., Witlox, F., & Van Wee, B. (2007). The Effects of the Land Use System on Travel Behavior: A Structural Equation Modeling Approach. *Transportation Planning & Technology, 30*(4), 331-353. doi:10.1080/03081060701461675

Van de Coevering, P., & Schwanen, T. (2006). Re-evaluating the impact of urban form on travel patternsin Europe and North-America. *Transport Policy, 13*(3), 229-239. doi:DOI: 10.1016/j.tranpol.2005.10.001

Vance, C., & Hedel, R. (2007). The impact of urban form on automobile travel: disentangling causation from correlation. *Transportation, 34*(5), 575-588. doi:10.1007/s11116-007-9128-6

Vega, A., & Reynolds-Feighan, A. (2009). A methodological framework for the study of residential location and travel-to-work mode choice under central and suburban employment destination patterns. *Transportation Research Part A: Policy and Practice, 43*(4), 401-419. doi:DOI: 10.1016/j.tra.2008.11.011

Venables, A. J. (2004). Evaluating Urban Transport Improvements: Cost-Benefit Analysis in the Presence of Agglomeration and Income Taxation. *Centre for Economic Performance (CEP) Discussion Paper No 651*.

Waddell, P., Ulfarsson, G. F., Franklin, J. P., & Lobb, J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice, 41*(5), 382-410. doi:DOI: 10.1016/j.tra.2006.09.008

Wagener, M. (2004). Overview of Land-Use Transport Models. In D. A. Hensher & K. J. Button (Eds.), *Transport Gegoraphy and Spatial Systems*. (pp. 127-146). Kidlington, UK: Pergamon/Elsevier Science.

Wang, D., & Chai, Y. (2009). The jobs–housing relationship and commuting in Beijing, China: the legacy of Danwei. *Journal of Transport Geography, 17*(1), 30-38. doi:http://dx.doi.org/10.1016/j.jtrangeo.2008.04.005

Wardman, M. (2006). Demand for rail travel and the effects of external factors. *Transportation Research Part E: Logistics and Transportation Review, 42*(3), 129-148. doi:http://dx.doi.org/10.1016/j.tre.2004.07.003

Wardman, M., Lythgoe, W., & Whelan, G. (2007). Rail Passenger Demand Forecasting: Cross-Sectional Models Revisited. *Research in Transportation Economics, 20*(0), 119-152. doi:http://dx.doi.org/10.1016/S0739-8859(07)20005-8

Watkins, C. (1999). Property valuation and the structure of urban housing markets. *Journal of property investment and finance, 17*(2), 157-175.

Weisbrod, G., & Reno, A. (2009). *Economic impact of public transportation investment.* Retrieved from http://www.apta.com/resources/reportsandpublications/Documents/Economic-Impact-Public-Transportation-Investment-APTA.pdf.

Williams, K. (2005). Spatial Planning, Urban Form and Sustainable Transport: An Introduction. In K. Williams (Ed.), *Spatial Planning, Urban Form and Sustainable Transport* (pp. 1-16). England: Ashgate Publishing Limited.

Zemp, S., Stauffacher, M., Jang, D. J., & Scholz, R. W. (2011). Classifying railway stations for strategic transport and land use planning: Context matters! *Journal of Transport Geography, 19*, 670-679.

APPENDIX CHAPTER 3

APPENDIX 1 THE OPERATIONAL DEFINITION OF VARIABLES

Factor - Variables	*Operational definition*	*Sources*	*Further References*
Self-sufficiency or self-containment **Job housing balance index**	The ratio of the number of jobs to the number of workers. The balance index is 0.75 – 1.5; job-rich suburbs will have ratio >1.5 and residential-rich suburbs will have ratio <0.75.	Calculated from the ABS Census 2011	(Giuliano, 1991) (Cervero & Duncan, 2006) (Chan and Moreno 2013)
Density **Employed resident density and job density**	Employed resident density was the number of resident employees per land area of suburb. Job density was the number of jobs per land area of suburb.	Calculated from ABS Census 2011	Scheiner, J, 2010; Eluru, N ,C. Bhat et al 2010, Cervero, R. and J. Landis, 1997). (Rickwood & Glazebrook, 2009 (Chen et al., 2008)
Public transport supply index by bus network	Public transport index was a measurement of public transport for each suburb. The index was a combined measure of service frequency (such as bus trips per week) and access distance was computed for each suburb using GIS software. The higher the index, the better the level of public transport in that suburb.	Calculation based on Currie and Delbosc's formula	(Currie & Delbosc, 2010) (Van de Coevering & Schwanen, 2006)
Suburb's centrality by road network travel distance	Suburb centrality index was a measurement of private transport supply for a suburb. It referred to the total road network travel distance from a suburb to all other suburbs based on the shortest routes. The data was computed for each suburb using GIS software. The higher the index, the lower the level of accessibility or centrality by private car in that suburb.	GIS calculation	(Van de Coevering & Schwanen, 2006)
Proximity or Distance of suburb to train station	The distance from the centroid of the suburb to the nearest train station as measured by: The distance from the fishnet centroid to the nearest train station, calculated as a mean value across all fishnets in each suburb. Calculation was obtained by GIS.	GIS establishment on the fishnet dataset GIS calculation of Euclidian distance	Cervero and Duncan, 2010; Commins, N and A. Nolan, 2011 (Ellis & Parolin, 2010)

Travel time by park and ride services	The average daily travel in minutes measured between each suburb across Perth metropolitan region (matrices data set). The amount of time between suburbs was calculated based on park and ride services. The data has been estimated in the STEM model produced by Department of Planning.	STEM model (Dept of Planning of Western Australia) And GIS calculation	Guiterrez (2011) Graham and Melo (2010)
The number of working trips by train	The number of daily trips for journeys to work involving all train combinations as the mode of travel, for each suburb.	ABS Census 2011 ‘	Taylor et al 2009
The ratio of the number of train trips in residential suburbs to the number of employed residents	The number of daily trips for journeys to work involving all train combinations as the mode of travel for each suburb, divided by total employed residents.	ABS Census 2011 ‘	Taylor et al 2009
The ratio of the number of train trips to workplace suburbs to the total trips	The number of daily trips for journeys to work involving all train combinations as the mode of travel, for each suburb, divided by total trips	ABS Census 2011 ‘	Taylor et al 2009
Agglomeration factor	The effective job density as a function of: Total jobs in each suburb i and all other suburbs j, The travel time by park and ride services between suburb i and suburbs j.	Calculation was made based on Graham's formula. Travel time matrix between suburb from STEM model (DoP) A GIS calculation was made for the conversion to travel times. The number of employed residents and jobs were obtained from the ABS Census Area of suburb (ABS Census)	(Graham, 2007) (Graham, Gibbons, & Martin, 2009) (Graham & Melo, 2010).

	The effective employed resident density as a function of: Total resident employees in each suburb i and all other suburbs j, The travel time by park and ride services between suburb i and suburbs j.	Calculation was made based on Graham's formula. Travel time matrix between suburb from STEM model (DoP) A GIS calculation was made for the conversion to travel times. The number of workers and jobs were obtained from ABS Census Area of suburb (ABS Census)	(Graham, 2007) (Graham et al., 2009) (Graham & Melo, 2010).
Land value for residential land uses and non-residential land uses	The mean price calculated across all dwellings that were sold in each suburb across Perth metropolitan region in 2011 transactions. The land value was calculated based on the property price divided by the area of land of the property. The calculation was standardized based on housing characteristics of a 3 bedrooms-2 bathrooms property type. This was to avoid bias from housing characteristics.	©Western Australian Land Information Authority (Landgate, 2014).	Tillema T et al, 2010, Ryan 1999 Rodriguea, D. A and C.H. Mojica, 2009; Du H and C. Mulley, 2007, Munoz-Raskin, R. 201 (So et al 2009).
Personal wage of non-resident employees for blue collar and managers/professional employment	Total weekly income based on weekly personal income in place of work (SA2 level). The total income was summed over all non-resident employees at the place of work divided by the total non-resident employees and by total weekly hours (40 hours). A conversion of data for income from SA2 level into suburbs was conducted (Chapter 4).	Calculated from ABS Census	(So, 2001)
Personal income of resident employees based on blue collar and managers/professionals occupations	Total weekly income summed over all employed residents (person) at the place of residence divided by the total employed residents and by total weekly hours (40 hours).	Calculated from ABS Census	(Wardman, 2006)
The level of car ownership per household	The average number of cars owned in each household, i.e. total cars divided by total dwellings.	ABS Census	Dargay, J., 2007; Cervero and Duncan, 2010 (Wardman, 2006) Taylor 2009

The proportion of manager/professional employed residents	The proportion of employees that were working as managers or professional occupation in the place of residence	Calculated from ABS Census	(Van de Coevering & Schwanen, 2006) Hall (p. 424)
The proportion of blue collar employed residents	The proportion of employees that were working in blue collar occupation types in the place of residence Blue collar occupations refer to technicians, trade workers, machinery operators and drivers, and labours.	Calculated from ABS Census	(Van de Coevering & Schwanen, 2006) Hall (p. 424)
The proportion of manager/professional jobs	The proportion of employment in managerial or professional occupations at the place of work	Calculated from ABS Census	(Van de Coevering & Schwanen, 2006) Hall (1969)
The proportion of blue collar jobs	The proportion of employment in blue collar occupations in the place of work	Calculated from ABS Census	(Van de Coevering & Schwanen, 2006) Hall (1969)
The proportion of employed residents in the construction, manufacturing, retail and total 19 sectors.	Employed residents by industry sector was calculated based on the number of employed residents in that sector divided by total employed residents	Calculated from ABS Census	Hall (1969) (Timothy and Wheaton 2001). (J Gutierrez et al 2011).
The proportion of jobs in the construction, manufacturing, retail and total 19 sectors.	Jobs by industry sector was calculated based on the number of jobs in that sector divided by the total jobs.	Calculated from ABS Census	Hall (1969) (Timothy and Wheaton 2001). (J Gutierrez et al 2011).
Catchment radius of stations (overall)	The maximum distance of the farthest suburbs that captured 95% of total train ridership.	GIS (centroid to centroid distance)	(Zemp, Stauffacher, Jang, & Scholz, 2011) Cervero (1992)

APPENDIX 2 DETERMINATION OF THE DEPENDENT VARIABLE (MODEL SPECIFICATION)

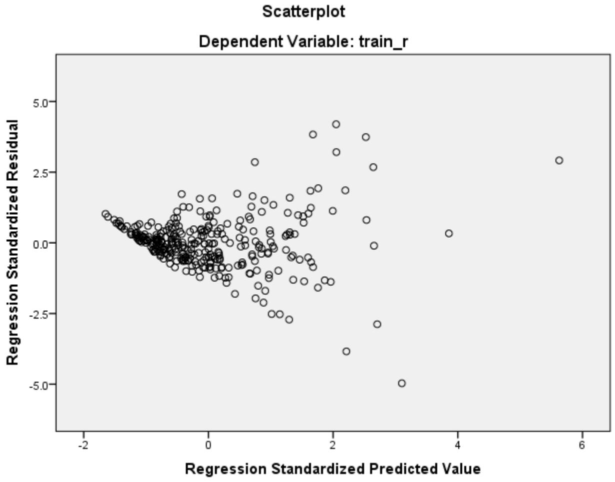
1. Absolutenumber train trip production

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.863[a]	.745	.730	97.04722

a. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, worker, inc_pr_r, p_wocon, wod_ln, ln_strio, inc_blu_r, ln_lvr, sqrt_ptiori, ln_jwr, ejdtt1000, pmanpr_r, pblu_r, wod, ln_avedist

b. Dependent Variable: train_r



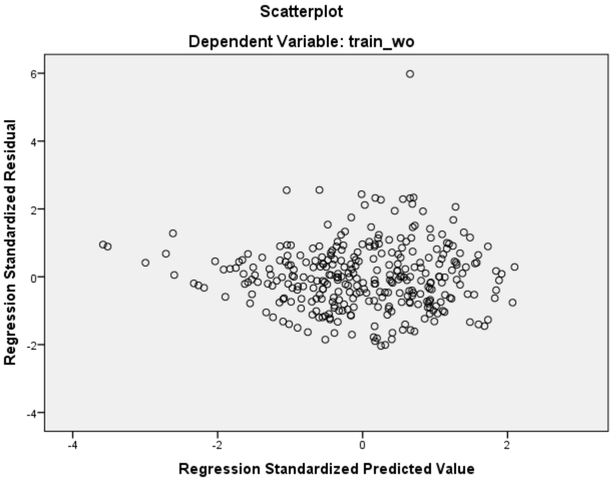
2. Proportion of train trip production

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.685[a]	.470	.438	.02646

a. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, worker, inc_pr_r, p_wocon, wod_ln, ln_strio, inc_blu_r, ln_lvr, sqrt_ptiori, ln_jwr, ejdtt1000, pmanpr_r, pblu_r, wod, ln_avedist

b. Dependent Variable: train_wo



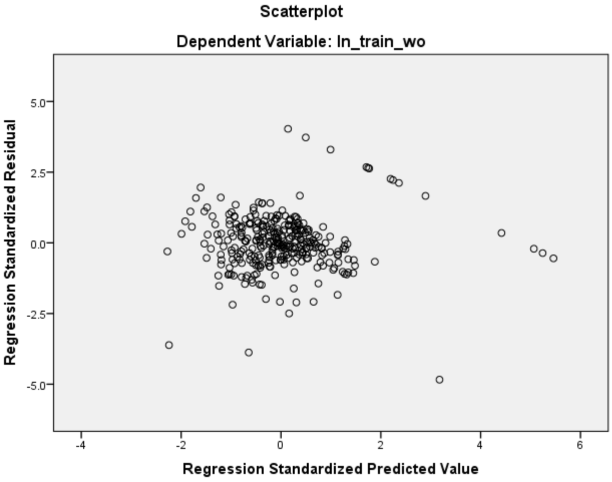
3. Ln of proportion of train trip production

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.663[a]	.440	.406	.66960

a. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, worker, inc_pr_r, p_wocon, wod_ln, ln_strio, inc_blu_r, ln_lvr, sqrt_ptiori, ln_jwr, ejdtt1000, pmanpr_r, pblu_r, wod, ln_avedist

b. Dependent Variable: ln_train_wo



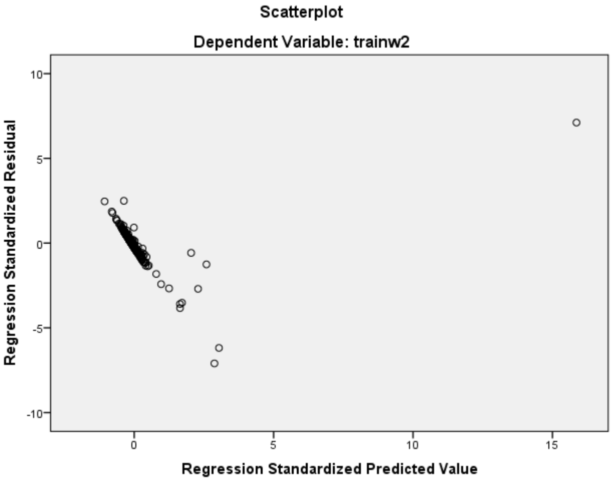
4. Absolutenumber of train trip attraction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.935[a]	.875	.868	536.59266

a. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, inc_mp, p_manpr, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, lnjobd_ln, edertt1000, p_blu, ln_avedist, ln_jobd

b. Dependent Variable: trainw2



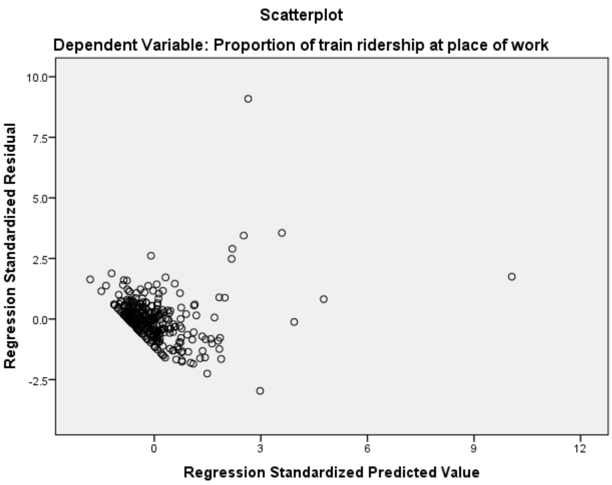
5. Proportion of train trip attraction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.844[a]	.712	.695	.01923

a. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, inc_mp, p_manpr, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, lnjobd_ln, edertt1000, p_blu, ln_avedist, ln_jobd

b. Dependent Variable: Proportion of train ridership at place of work



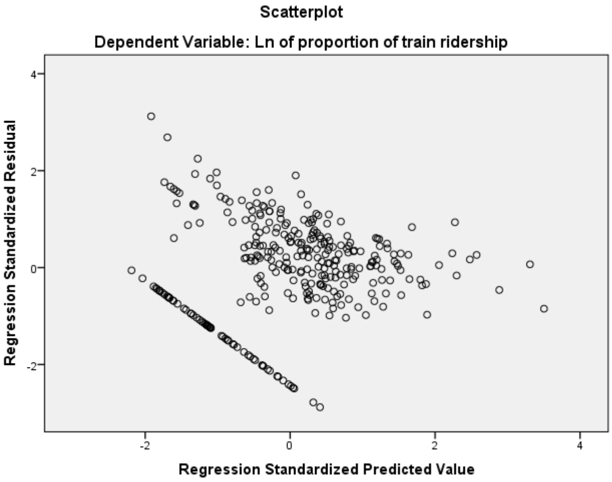
6. Ln of proportion of train trip attraction

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.745[a]	.555	.529	.96218

a. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, inc_mp, p_manpr, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, lnjobd_ln, edertt1000, p_blu, ln_avedist, ln_jobd

b. Dependent Variable: Ln of proportion of train ridership



APPENDIX CHAPTER 4

APPENDIX 3 CENTRAL TENDENCY VALUES WITH NORMAL DISTRIBUTION

Statistics

		train_wo	p_manpr	p_blu	inc_mp	inc_bl	pmanpr_r	pblu_r	p_wocon	car_own	wod	wodc	wodm	wodr	edertt1000	ederttc100	edertm100	ederttr100
N	Valid	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316
	Missing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean		0.062	0.325	0.290	32.407	21.275	0.340	0.329	0.102	1.879	609.612	58.183	47.895	62.224	11.857	6.023	9.319	5.037
Std. Error of Mean		0.002	0.005	0.006	0.244	0.214	0.007	0.006	0.002	0.023	26.008	2.531	2.030	2.601	0.156	0.109	0.111	0.105
Median		0.060	0.311	0.265	32.116	21.007	0.321	0.333	0.103	1.884	635.981	57.159	48.193	67.808	12.217	6.056	9.403	5.209
Mode		0.000	0.269	.322[a]	29.697	22.504	0.250	0.500	0.000	0.000	.179[a]	0.000	0.000	0.000	5.287[a]	1.987[a]	4.359[a]	1.349[a]
Std. Deviation		0.035	0.087	0.111	4.342	3.812	0.125	0.115	0.035	0.415	462.327	44.983	36.079	46.239	2.771	1.936	1.981	1.873
Variance		0.001	0.008	0.012	18.851	14.529	0.016	0.013	0.001	0.172	213746.5	2023.5	1301.7	2138.1	7.680	3.750	3.924	3.510
Skewness		0.555	0.856	0.767	0.558	0.946	0.413	0.068	0.020	-0.929	0.497	0.477	0.363	0.389	-0.068	0.180	-0.174	0.141
Std. Error of Skewness		0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137
Kurtosis		0.972	0.430	0.111	2.315	3.664	-0.263	0.034	0.919	4.174	0.435	-0.460	-0.768	0.185	-0.724	-0.636	-0.486	-0.619
Std. Error of Kurtosis		0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273
Maximum		0.240	0.615	0.646	51.129	41.385	0.670	0.750	0.210	2.830	2794.8	194.3	144.3	275.9	17.9	12.0	14.4	10.1

a. Multiple modes exist. The smallest value is shown

APPENDIX 4 CENTRAL TENDENCY VALUES WITH NON-NORMAL DISTRIBUTION

Statistics

		train_w	ptrainw2	train_r2	All_job	p_jobcon	p_jobman	p_jobret	ptiori	strio	ave_dist	jobd	jobdc	jobdm	jobdr	jwr	jwrc	jwrm	jwrr
N	Valid	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316	316
	Missing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Mean	146.65	0.0226	171.70	2010.74	0.098	0.066	0.117	893395.08	11883.29	5.87	464.00	31.05	28.82	48.79	19.84	18.69	41.85	10.32
Std. Error of Mean	83.07	0.0020	10.51	304.09	0.003	0.005	0.005	76114.16	168.89	0.28	58.62	3.40	4.15	4.62	9.63	8.04	19.33	5.11
Median	7.86	0.0141	108.50	722.48	0.083	0.035	0.099	411960.46	10981.46	4.27	191.47	15.12	4.98	19.32	0.38	0.30	0.17	0.39
Mode	0.00	0.0010	1.00	25.07[a]	0.078	0.158	.00[a]	0.00	8337.3[a]	2.31	1.43	.225[a]	.6316[a]	.00[a]	.016[a]	.0128[a]	.00[a]	0.00
Std. Deviation	1476.66	0.0348	186.78	5405.55	0.061	0.081	0.082	1353034.91	3002.19	4.96	1042.09	60.48	73.72	82.19	171.21	142.91	343.63	90.86
Variance	2180517.5	0.0012	34888.2	29219983.3	0.004	0.007	0.007	1830703465653.6	9013120.5	24.6	1085960.2	3657.8	5434.5	6754.8	29311.5	20423.6	118083.1	8256.3
Skewness	16.94	5.24	2.25	10.87	2.307	2.593	1.133	4.84	1.27	2.39	6.03	5.78	4.79	3.32	11.05	8.81	9.00	11.58
Std. Error of Skewness	0.14	0.14	0.14	0.14	0.137	0.137	0.137	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Kurtosis	295.06	37.36	7.45	152.20	12.659	7.946	1.188	42.10	1.24	6.76	43.25	45.60	26.87	12.70	133.28	80.55	81.96	139.64
Std. Error of Kurtosis	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27
Maximum	25871.07	0.35	1362.00	81785.94	0.58	0.52	0.41	15358826.49	22666.50	29.54	9342.28	664.47	571.19	531.26	2373.66	1564.62	3545.99	1198.00

a. Multiple modes exist. The smallest value is shown

Statistics

		p_worker_man	p_woret	inc_pr_r	inc_blu_r	lvr	lvnr	ejdtt1000	ejdttc100	ejdttm100	ejdttr100
N	Valid	316	316	316	316	316	316	316	316	316	316
	Missing	0	0	0	0	0	0	0	0	0	0
Mean		.0889	.1025	36.3895	26.2531	1301.778	1492.806	10.416	4.041	7.815	4.721
Std. Error of Mean		.00287	.00153	.42115	.22509	61.434	67.277	0.265	0.154	0.212	0.187
Median		.0822	.1031	37.4937	26.5352	1057.440	1212.963	9.689	3.472	7.039	3.978
Mode		0.00	0.00	0.00	0.00	107.169	107.169	4.426[a]	1.322[a]	3.647[a]	1.224[a]

Std. Deviation	.05107	.02711	7.48653	4.00136	1092.083	1195.943	4.708	2.731	3.777	3.326
Variance	.003	.001	56.048	16.011	1192644.6	1430280.2	22.170	7.459	14.264	11.062
Skewness	7.008	1.661	-2.873	-2.811	1.699	1.266	5.267	5.266	3.989	3.095
Std. Error of Skewness	.137	.137	.137	.137	0.137	0.137	0.137	0.137	0.137	0.137
Kurtosis	88.650	20.784	12.183	16.484	3.408	1.215	48.432	40.455	21.603	13.030
Std. Error of Kurtosis	.273	.273	.273	.273	0.273	0.273	0.273	0.273	0.273	0.273
Maximum	.75	.33	52.01	35.69	6448.999	5331.517	62.002	31.193	37.546	25.487

a. Multiple modes exist. The smallest value is shown

APPENDIX 5 CENTRAL TENDENCY POST-TRANSFORMATION

Statistics

		Ln of proportion of train ridership	ln_train_wo	sqrt_ptiori	ln_strio	ln_avedist	ln_jobd	ln_jobdc	ln_jobdm	ln_jobdr2
N	Valid	316	316	316	316	316	316	316	315	316
	Missing	0	0	0	0	0	0	0	1	0
Mean		-4.5653	-2.7806	746.286	9.355	1.515	4.959	2.456	1.750	2.854
Std. Error of Mean		.07890	.04888	32.682	0.013	0.039	0.098	0.089	0.101	0.089
Median		-4.2621	-2.7648	641.836	9.304	1.451	5.255	2.716	1.610	3.011
Mode		-6.91	0.00	0.000	9.028[a]	0.839	0.355	-1.489[a]	-2.451	0.193
Std. Deviation		1.40253	.86899	580.965	0.230	0.687	1.745	1.586	1.794	1.583
Variance		1.967	.755	337520.3	0.053	0.472	3.046	2.515	3.219	2.505
Skewness		-.490	.729	1.050	0.795	0.416	-0.613	-0.743	0.190	-0.129
Std. Error of Skewness		.137	.137	0.137	0.137	0.137	0.137	0.137	0.137	0.137
Kurtosis		-.556	3.979	2.173	-0.106	-0.139	0.299	0.840	-0.017	-0.810

Std. Error of Kurtosis	.273	.273	0.273	0.273	0.273	0.273	0.273	0.274	0.273
Maximum	-1.04	0.00	3919.034	10.029	3.386	9.142	6.499	6.348	6.277

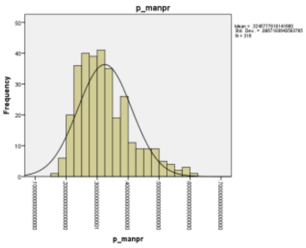
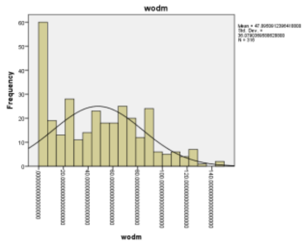
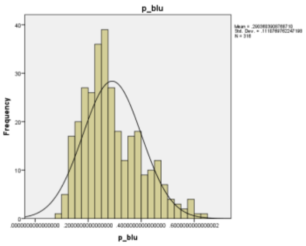
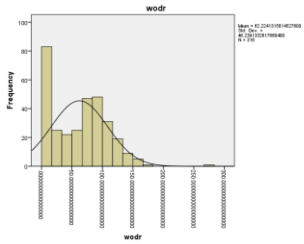
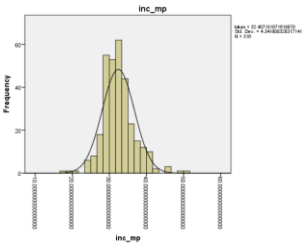
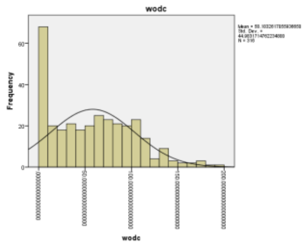
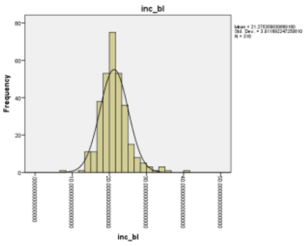
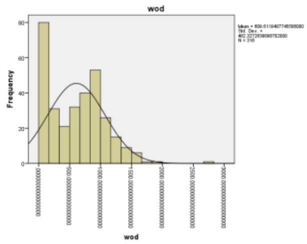
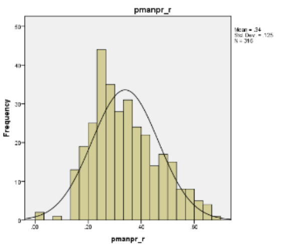
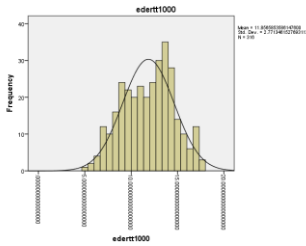
a. Multiple modes exist. The smallest value is shown

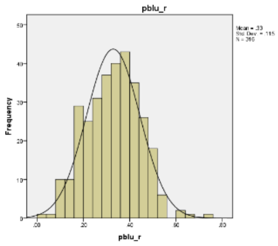
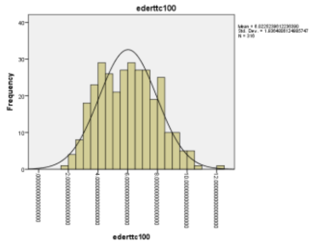
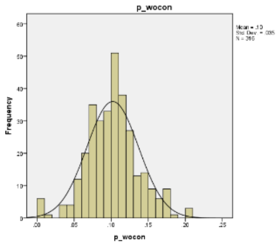
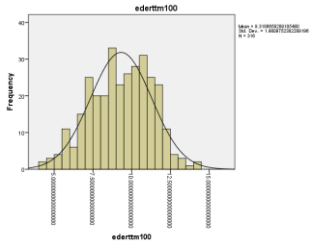
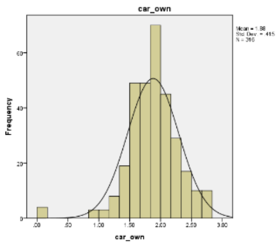
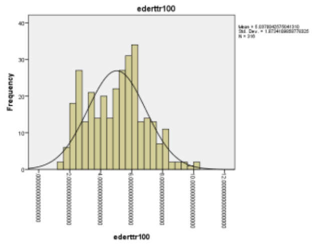
Statistics

		ln_jwr	ln_jwrc	ln_jwrm2	ln_jwrr2	ln_lvr	sr_lvnr
N	Valid	316	316	316	316	316	316
	Missing	0	0	0	0	0	0
Mean		-.675	-.867	-1.305	-.775	6.772	35.443
Std. Error of Mean		.085	.087	.108	.088	.058	.867
Median		-.968	-1.212	-1.724	-.911	6.964	34.827
Mode		-4.128[a]	-4.357[a]	-4.605[a]	-4.605	4.674	10.352
Std. Deviation		1.510	1.548	1.920	1.566	1.035	15.407
Variance		2.281	2.395	3.687	2.451	1.071	237.371
Skewness		2.297	2.453	1.980	1.365	-1.021	.267
Std. Error of Skewness		.137	.137	.137	.137	.137	.137
Kurtosis		9.311	9.649	6.665	5.184	1.140	-.201
Std. Error of Kurtosis		.273	.273	.273	.273	.273	.273
Maximum		7.772	7.355	8.174	7.088	8.772	73.017

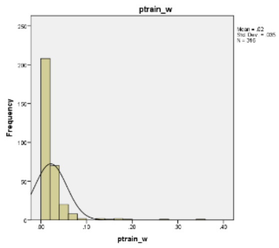
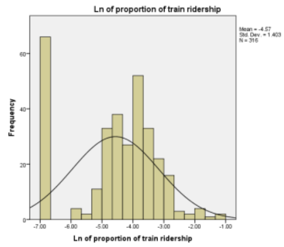
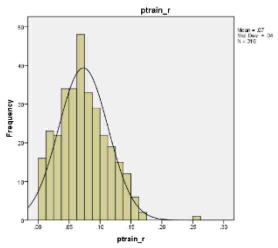
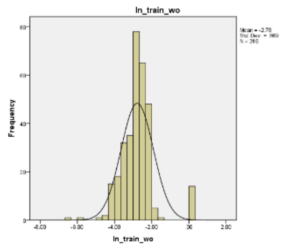
a. Multiple modes exist. The smallest value is shown

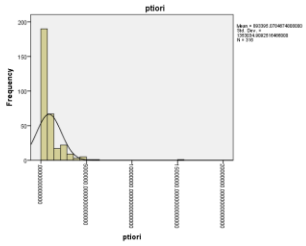
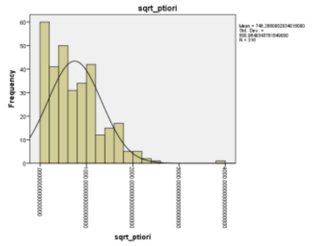
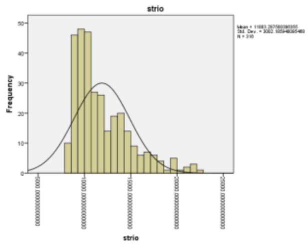
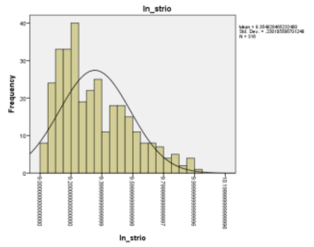
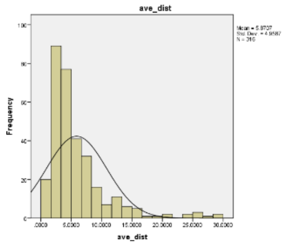
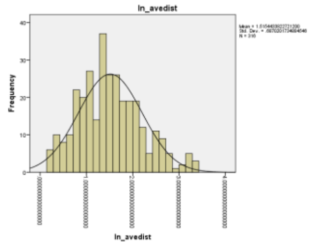
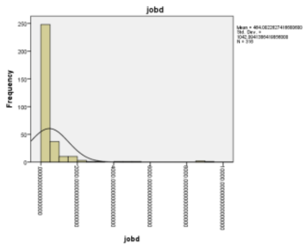
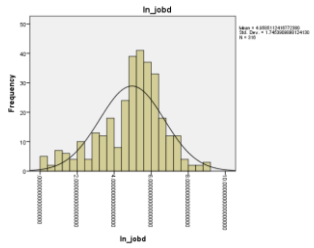
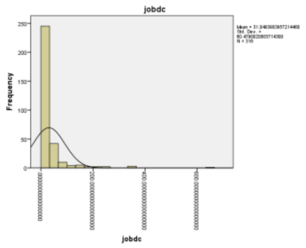
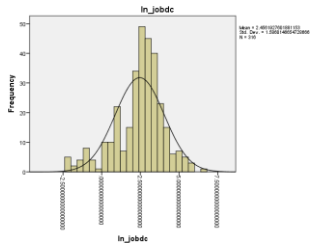
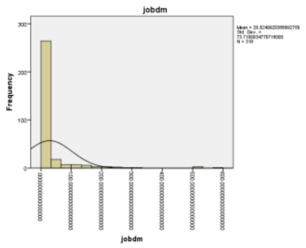
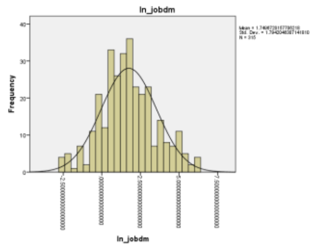
APPENDIX 6 HISTOGRAM OF VARIABLES WITHOUT TRANSFORMATION

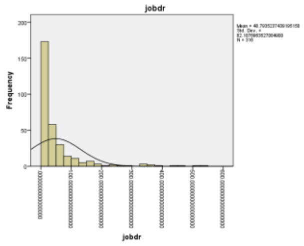
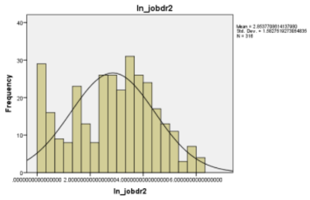
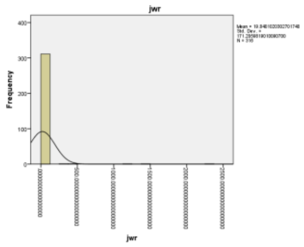
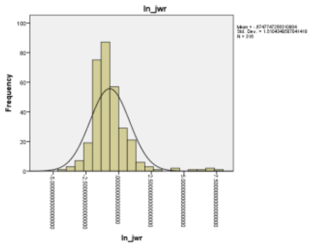
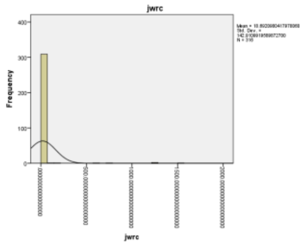
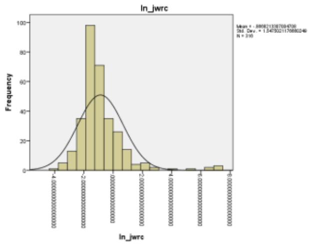
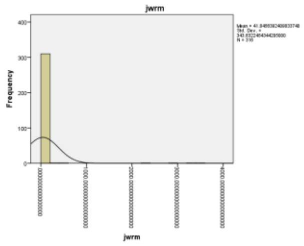
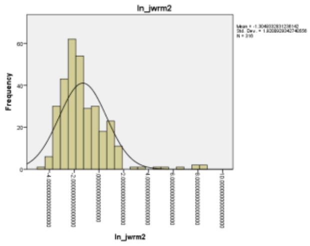
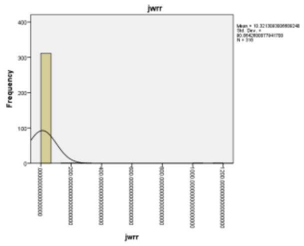
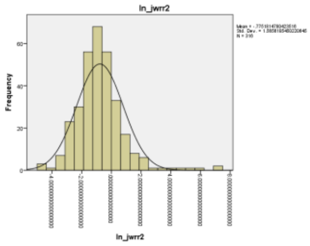
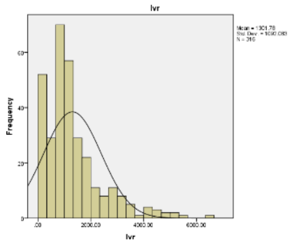
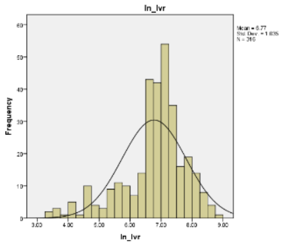
Variable	*Histogram*	*Variable*	*Histogram*
Proportion of job in managers/professionals occupation **(p_manpr)**		Employed resident density in the manufacturing sector (wodm)	
Proportion of job in blue collar occupation **(p_blu)**		Employed resident density in the retail sector (wodr)	
Income of employment of managers/ professionals job (inc_mp)		Employed resident density in the construction sector (wodc)	
Income of employment of blue collar job **(inc_blu)**		Employed resident density (wod)	
Proportion of employed resident in managers/professionals occupation **(pmanpr_r)**		Effective employed resident density in 1000 units (edertt1000)	

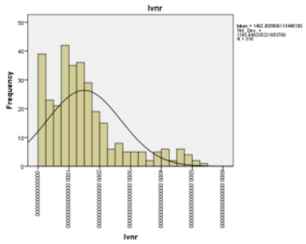
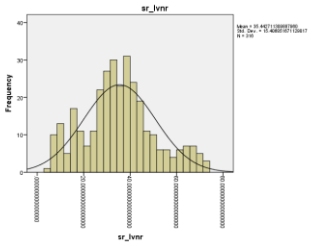
Proportion of employed resident in blue collar occupation **(pblu_r)**		Effective employed resident density of the construction sector in 100 units (ederttc100)	
Proportion of employed resident in the construction sector (p_wocon)		Effective employed resident density of the manufacturing sector in 100 units (ederttm100)	
The level of car ownership in a household **(car_own)**		Effective employed resident density of the retail sector in 100 units (ederttr100)	

APPENDIX 7 HISTOGRAM OF VARIABLE WITH SUCCESSFUL TRANSFORMATIONS

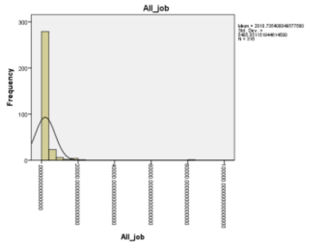
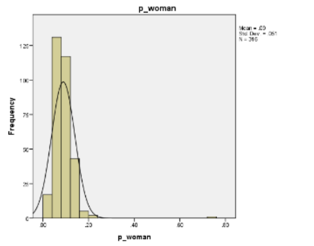
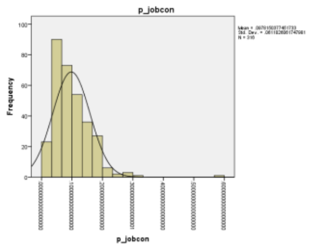
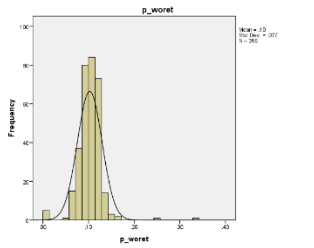
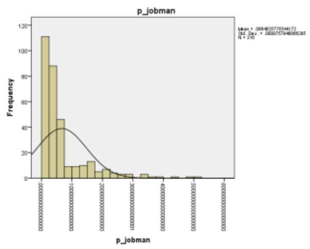
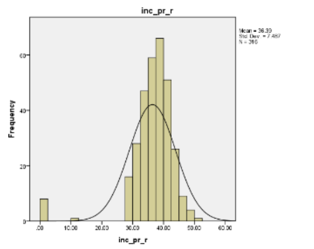
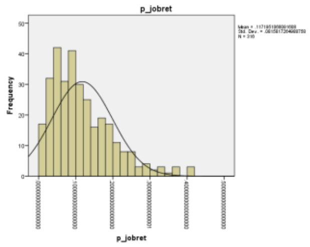
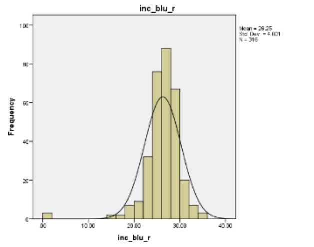
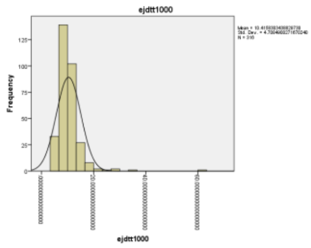
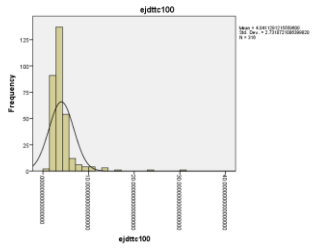
Variable (Before transformation)	*Histogram (Before transformation)*	*Variable (After transformation)*	*Histogram (After transformation)*
Proportion of train trip attraction **(ptrain_w)**		Ln of the proportion of train trip attraction (ln_ptrainw22)	
Proportion of train trip production **(ptrain_r2)**		Ln of the proportion of train trip production (ln_train_wo)	

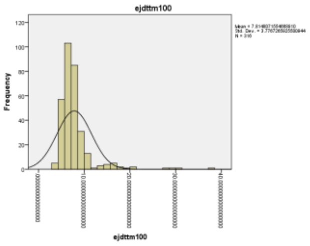
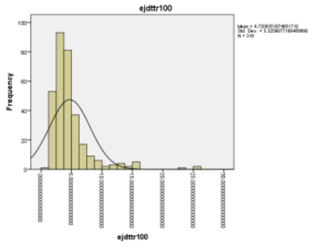
Public transport supply by buses network **(ptiori)**		Square root of the public transport supply by buses network (sqrt_ptiori)	
Centrality suburb by road network travel distance **(strio)**		Ln of centrality suburb by road network travel distance (ln_strio)	
Average distance of a suburb from train stations **(ave_dist)**		Ln of average distance of a suburb from train stations (ln_avedist)	
Job density **(jobd)**		Ln of job density (ln_jobd)	
Job density in the construction sector **(jobdc)**		Ln of job density in the construction sector (ln_jobdc)	
Job density in the manufacturing sector **(jobdm)**		Ln of job density in the manufacturing sector (ln_jobdm)	

Job density in the retail sector **(jobdr)**		Ln of job density in the retail sector (ln_jobdr2)	
Job to employed resident ratio **(jwr)**		Ln of job to employed resident ratio (ln_jwr)	
Job to employed resident ratio in the construction sector **(jwrc)**		Ln of Job to employed resident ratio in the construction sector (ln_jwrc)	
Job to employed resident ratio in the manufacturing sector **(jwrm)**		Ln of job to employed resident ratio in the manufacturing sector (ln_jwrm2)	
Job to employed resident ratio in the retail sector **(jwrr)**		Ln of job to employed resident ratio in the retail sector (ln_jwrr2)	
Land value of residential land uses **(lvr)**		Ln of land value of residential land uses (ln_lvr)	

Land value of non-residential land uses **(lvnr)**		Square root of land value of non-residential land uses (sr_lvnr)	

APPENDIX 8 HISTOGRAM OF VARIABLES THAT ARE REMAINED IN THEIR ORIGINAL SCORES

Variable	*Histogram*	*Variable*	*Histogram*
All job **(all_job)**		Proportion of employed resident in the manufacturing sector (p_worker_man)	
Proportion of job in the construction sector **(p_jobcon)**		Proportion of employed resident in the retail sector (p_woret)	
Proportion of job in the manufacturing sector **(p_jobman)**		Income of managers/ professionals employed resident (inc_pr_r)	
Proportion of job in the retail sector **(p_jobret)**		Income of blue collar employed resident	
Effective job density in 1000 units (ejdtt1000)		Effective job density of the construction sector in 100 units (ejdttc100)	

Effective job density of the manufacturing sector in 100 units (ejdttm100)		Effective job density of the retail sector in 100 units (ejdttr100)	

APPENDIX CHAPTER 5

APPENDIX 9 GETIS-ORD* STATISTICS OF VARIABLE WITH *EDER* ALL SECTOR

sscname	*ave_dist*	*SOURCE_ID*	*edertt1000*	*GiZScore*	*GiPValue*	*Gi_Bin*
Mount Hawthorn	1.8942	10	17.084	3.893	0.000	3
Osborne Park	2.7184	15	14.319	4.040	0.000	3
Padbury	2.8003	16	14.164	2.943	0.003	3
Perth (WA)	2.4974	23	17.213	4.169	0.000	3
Balcatta	3.8854	39	14.437	3.594	0.000	3
Bateman	2.6427	47	13.762	2.283	0.022	2
Bayswater (WA)	2.1437	48	15.606	2.488	0.013	2
Bentley (WA)	2.3465	59	14.963	3.099	0.002	3
Brentwood (WA)	2.3787	66	13.737	2.106	0.035	2
Bull Creek (WA)	1.7516	69	15.359	2.718	0.007	3
Burswood	2.0104	72	13.183	2.284	0.022	2
Cannington	1.7192	78	13.957	1.687	0.092	1
Carine	2.9579	81	14.857	2.481	0.013	2
Carlisle	1.2898	82	15.073	2.651	0.008	3
Como (WA)	3.8018	95	17.017	3.145	0.002	3
Craigie (WA)	2.1538	102	13.277	2.015	0.044	2
Duncraig	3.0559	112	16.366	3.102	0.002	3
East Victoria Park	1.8402	117	16.428	3.035	0.002	3
Floreat	3.0418	123	14.345	1.784	0.074	1
Glendalough	1.9066	130	15.386	2.988	0.003	3
Greenwood (WA)	2.8697	136	14.875	2.904	0.004	3
Gwelup	3.7223	138	13.477	2.885	0.004	3
Hamersley	3.8619	139	14.255	2.106	0.035	2
Highgate (WA)	1.1953	151	15.284	3.099	0.002	3
Inglewood (WA)	2.1942	162	14.421	3.826	0.000	3
Innaloo	3.6884	163	16.922	3.152	0.002	3
Jolimont	1.4659	169	12.138	1.936	0.053	1
Joondanna	2.6462	171	15.230	3.745	0.000	3

Karrinyup	4.6538	178	14.085	3.063	0.002	3
Kingsley	3.2559	185	15.597	2.368	0.018	2
Lathlain	1.3894	192	14.143	3.029	0.002	3
Leederville	1.3737	194	14.788	4.080	0.000	3
Leeming	2.9843	195	15.961	2.342	0.019	2
Manning	2.8852	204	14.408	1.990	0.047	2
Maylands (WA)	2.0831	210	17.153	3.226	0.001	3
Mount Lawley	1.5865	215	17.877	4.459	0.000	3
Mount Pleasant (WA)	3.0139	217	15.246	1.684	0.092	1
Murdoch	2.435	223	12.990	2.130	0.033	2
North Perth	2.2107	234	17.324	4.586	0.000	3
Northbridge (WA)	1.6833	235	14.431	2.679	0.007	3
Rivervale	1.8713	247	15.656	2.078	0.038	2
Rossmoyne	2.7395	250	13.786	2.484	0.013	2
Shenton Park	1.4905	259	12.933	2.048	0.041	2
South Perth	3.9637	267	17.877	3.470	0.001	3
St James (WA)	1.5844	270	14.438	1.826	0.068	1
Stirling (WA)	2.5418	271	16.760	2.705	0.007	3
Subiaco	1.4745	274	14.899	3.160	0.002	3
Victoria Park	2.0073	284	17.022	3.496	0.000	3
Warwick (WA)	2.5922	291	14.153	2.893	0.004	3
Welshpool (WA)	2.7892	297	11.651	1.975	0.048	2
Wembley	2.3809	298	17.094	3.472	0.001	3
West Perth	1.2481	301	16.724	3.854	0.000	3

APPENDIX 10 GETIS-ORD* STATISTICS OF VARIABLE WITH E*JD* ALL SECTOR

sscname	*ave_dist*	*SOURCE_ID*	*ejdttt1000*	*GiZScore*	*GiPValue*	*Gi_Bin*
Mount Hawthorn	1.8942	10	14.498	3.631	0.000	3
Osborne Park	2.7184	15	29.129	3.161	0.002	3
Perth (WA)	2.4974	23	62.002	9.982	0.000	3
Bentley (WA)	2.3465	59	18.180	2.677	0.007	3
Burswood	2.0104	72	18.284	1.929	0.054	1
Cannington	1.7192	78	16.947	1.864	0.062	1
Carlisle	1.2898	82	12.948	2.458	0.014	2
Daglish	1.2295	105	12.215	2.951	0.003	3
East Perth	1.4896	115	26.428	8.966	0.000	3
East Victoria Park	1.8402	117	13.878	2.888	0.004	3
Glendalough	1.9066	130	13.400	3.394	0.001	3
Highgate (WA)	1.1953	151	14.494	7.496	0.000	3
Innaloo	3.6884	163	14.607	2.589	0.010	3

Jolimont	1.4659	169	14.754	3.142	0.002	3
Joondanna	2.6462	171	11.596	2.670	0.008	3
Lathlain	1.3894	192	12.432	2.105	0.035	2
Leederville	1.3737	194	18.051	4.521	0.000	3
Mount Lawley	1.5865	215	15.591	6.657	0.000	3
Northbridge (WA)	1.6833	235	22.893	10.929	0.000	3
Queens Park (WA)	1.4585	241	11.495	2.580	0.010	3
Shenton Park	1.4905	259	15.144	6.932	0.000	3
St James (WA)	1.5844	270	12.235	2.606	0.009	3
Stirling (WA)	2.5418	271	14.653	3.302	0.001	3
Subiaco	1.4745	274	27.167	8.745	0.000	3
Victoria Park	2.0073	284	15.836	2.403	0.016	2
Welshpool (WA)	2.7892	297	21.750	2.279	0.023	2
Wembley	2.3809	298	16.397	3.631	0.000	3
West Leederville	1.22	300	15.563	5.760	0.000	3
West Perth	1.2481	301	35.062	9.936	0.000	3

APPENDIX CHAPTER 6 AND CHAPTER 7
PLACE OF RESIDENCE (POR) MODEL

APPENDIX 11 ALL SECTOR LUTI

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.647[a]	.418	.391	.67807

a. Predictors: (Constant), ln_jwr, pblu_r, sqrt_ptiori, p_woret, ln_avedist, p_worker_man, inc_blu_r, car_own, p_wocon, ln_lvr, wod, ln_strio, inc_pr_r, pmanpr_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	99.478	14	7.106	15.455	.000[b]
	Residual	138.392	301	.460		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ln_jwr, pblu_r, sqrt_ptiori, p_woret, ln_avedist, p_worker_man, inc_blu_r, car_own, p_wocon, ln_lvr, wod, ln_strio, inc_pr_r, pmanpr_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-11.31994	2.82593		-4.00573	.00008		
	pmanpr_r	1.10166	.75182	.15859	1.46532	.14388	.16500	6.06045
	pblu_r	-1.93486	.85456	-.25684	-2.26417	.02427	.15021	6.65749
	p_worker_man	2.28412	.98184	.13424	2.32637	.02066	.58047	1.72275
	p_wocon	3.24573	1.88023	.13072	1.72624	.08533	.33705	2.96695
	p_woret	2.87018	1.75952	.08954	1.63122	.10389	.64151	1.55881
	inc_pr_r	-.04885	.01067	-.42083	-4.57670	.00001	.22861	4.37422
	inc_blu_r	-.00089	.01445	-.00408	-.06128	.95118	.43650	2.29093
	car_own	-.61236	.14095	-.29214	-4.34443	.00002	.42746	2.33942
	sqrt_ptiori	-.00015	.00009	-.10355	-1.68668	.09270	.51286	1.94986
	ln_strio	1.47332	.29218	.39013	5.04253	.00000	.32291	3.09684
	ln_avedist	-.68150	.09041	-.53879	-7.53820	.00000	.37835	2.64304
	ln_lvr	-.26263	.06120	-.31272	-4.29145	.00002	.36399	2.74732
	wod	.00020	.00014	.10824	1.41936	.15683	.33234	3.00895
	ln_jwr	.08307	.04118	.14439	2.01712	.04457	.37721	2.65102

a. Dependent Variable: ln_train_wo

APPENDIX 12 ALL SECTOR – LUTI WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.657[a]	.432	.404	.67095

a. Predictors: (Constant), wod_ln, p_wocon, p_woret, p_worker_man, inc_pr_r, ln_avedist, car_own, sqrt_ptiori, inc_blu_r, ln_jwr, ln_lvr, ln_strio, pmanpr_r, wod, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	102.818	15	6.855	15.226	.000[b]
	Residual	135.053	300	.450		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), wod_ln, p_wocon, p_woret, p_worker_man, inc_pr_r, ln_avedist, car_own, sqrt_ptiori, inc_blu_r, ln_jwr, ln_lvr, ln_strio, pmanpr_r, wod, pblu_r

Coefficients[a]

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics

		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-11.75072	2.80075		-4.19557	.00004		
	pmanpr_r	1.37847	.75084	.19844	1.83590	.06736	.16198	6.17355
	pblu_r	-1.54589	.85756	-.20521	-1.80266	.07245	.14604	6.84738
	p_worker_man	1.87599	.98302	.11026	1.90840	.05729	.56698	1.76373
	p_wocon	2.75317	1.86927	.11089	1.47286	.14184	.33389	2.99498
	p_woret	2.26181	1.75533	.07056	1.28854	.19855	.63113	1.58447
	inc_pr_r	-.04401	.01071	-.37916	-4.10958	.00005	.22233	4.49789
	inc_blu_r	.00043	.01431	.00199	.03014	.97598	.43601	2.29355
	car_own	-.62187	.13952	-.29667	-4.45726	.00001	.42719	2.34089
	sqrt_ptiori	-.00020	.00009	-.13347	-2.16208	.03140	.49661	2.01364
	ln_strio	1.55663	.29073	.41219	5.35427	.00000	.31934	3.13150
	ln_avedist	-.85642	.11012	-.67709	-7.77694	.00000	.24967	4.00524
	ln_lvr	-.31115	.06312	-.37050	-4.92935	.00000	.33500	2.98508
	wod	-.00020	.00020	-.10434	-.96110	.33728	.16057	6.22788
	ln_jwr	.11124	.04204	.19335	2.64584	.00858	.35439	2.82174
	wod_ln	.00037	.00014	.25371	2.72369	.00683	.21811	4.58474

a. Dependent Variable: ln_train_wo

APPENDIX 13 THE CONSTRUCTION SECTOR - LUTI

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.665[a]	.443	.417	.66361

a. Predictors: (Constant), ln_jwrc, ln_lvr, p_worker_man, p_woret, p_wocon, sqrt_ptiori, inc_pr_r, ln_avedist, car_own, inc_blu_r, wodc, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	105.315	14	7.522	17.082	.000[b]
	Residual	132.556	301	.440		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ln_jwrc, ln_lvr, p_worker_man, p_woret, p_wocon, sqrt_ptiori, inc_pr_r, ln_avedist, car_own, inc_blu_r, wodc, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-11.82154	2.75696		-4.28789	.00002		
	pmanpr_r	1.26900	.73427	.18268	1.72823	.08497	.16569	6.03537
	pblu_r	-1.32954	.84845	-.17649	-1.56703	.11816	.14595	6.85158
	p_worker_man	2.12805	.95765	.12507	2.22216	.02702	.58443	1.71108
	p_wocon	1.69196	1.92027	.06815	.88111	.37896	.30951	3.23091
	p_woret	3.00738	1.71753	.09382	1.75100	.08096	.64488	1.55068
	inc_pr_r	-.04020	.01034	-.34635	-3.88624	.00013	.23309	4.29011
	inc_blu_r	.00112	.01414	.00515	.07902	.93707	.43653	2.29079
	car_own	-.57727	.13593	-.27540	-4.24694	.00003	.44027	2.27133
	sqrt_ptiori	-.00015	.00009	-.09901	-1.67898	.09419	.53234	1.87850
	ln_strio	1.48275	.28523	.39263	5.19843	.00000	.32455	3.08121
	ln_avedist	-.66038	.08798	-.52210	-7.50567	.00000	.38262	2.61354
	ln_lvr	-.29231	.06012	-.34806	-4.86245	.00000	.36131	2.76769
	wodc	.00407	.00124	.21079	3.29290	.00111	.45180	2.21336
	ln_jwrc	.14083	.03661	.25079	3.84681	.00015	.43559	2.29575

a. Dependent Variable: ln_train_wo

APPENDIX 14 THE CONSTRUCTION SECTOR – LUTI WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.672[a]	.452	.425	.65920

a. Predictors: (Constant), wodc_ln, ln_strio, p_worker_man, p_woret, inc_pr_r, car_own, ln_jwrc, sqrt_ptiori, inc_blu_r, p_wocon, ln_avedist, ln_lvr, pmanpr_r, pblu_r, wodc
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	107.507	15	7.167	16.493	.000[b]
	Residual	130.364	300	.435		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), wodc_ln, ln_strio, p_worker_man, p_woret, inc_pr_r, car_own, ln_jwrc, sqrt_ptiori, inc_blu_r, p_wocon, ln_avedist, ln_lvr, pmanpr_r, pblu_r, wodc

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-11.30964	2.74809		-4.11545	.00005		
	pmanpr_r	1.41895	.73244	.20427	1.93730	.05365	.16431	6.08594
	pblu_r	-1.19699	.84487	-.15889	-1.41677	.15759	.14524	6.88519
	p_worker_man	1.81482	.96145	.10666	1.88759	.06005	.57213	1.74786
	p_wocon	1.65571	1.90757	.06669	.86797	.38610	.30949	3.23114
	p_woret	2.38961	1.72814	.07455	1.38277	.16776	.62854	1.59099
	inc_pr_r	-.03906	.01029	-.33648	-3.79616	.00018	.23252	4.30067
	inc_blu_r	.00212	.01406	.00977	.15090	.88016	.43609	2.29310
	car_own	-.59888	.13536	-.28570	-4.42415	.00001	.43805	2.28286
	sqrt_ptiori	-.00016	.00009	-.10831	-1.84423	.06614	.52970	1.88787
	ln_strio	1.48578	.28334	.39343	5.24389	.00000	.32454	3.08128
	ln_avedist	-.82434	.11388	-.65172	-7.23869	.00000	.22537	4.43716
	ln_lvr	-.33296	.06240	-.39646	-5.33597	.00000	.33091	3.02193
	wodc	-.00083	.00250	-.04284	-.33054	.74122	.10875	9.19518
	ln_jwrc	.14348	.03638	.25551	3.94332	.00010	.43513	2.29816
	wodc_ln	.00338	.00151	.26634	2.24579	.02545	.12988	7.69916

a. Dependent Variable: ln_train_wo

APPENDIX 15 MANUFACTURING SECTOR - LUTI

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.651[a]	.424	.397	.67457

a. Predictors: (Constant), ln_jwrm2, ln_avedist, p_worker_man, p_woret, sqrt_ptiori, inc_blu_r, car_own, p_wocon, inc_pr_r, wodm, ln_lvr, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	100.902	14	7.207	15.839	.000[b]
	Residual	136.969	301	.455		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ln_jwrm2, ln_avedist, p_worker_man, p_woret, sqrt_ptiori, inc_blu_r, car_own, p_wocon, inc_pr_r, wodm, ln_lvr, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics

		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-12.74020	2.90380		-4.38742	.00002		
	pmanpr_r	1.45843	.75949	.20995	1.92028	.05577	.16003	6.24893
	pblu_r	-2.05785	.83863	-.27317	-2.45383	.01470	.15436	6.47822
	p_worker_man	2.39923	.96816	.14101	2.47814	.01376	.59084	1.69250
	p_wocon	3.45926	1.86022	.13933	1.85960	.06392	.34080	2.93430
	p_woret	3.11332	1.75707	.09712	1.77188	.07743	.63669	1.57062
	inc_pr_r	-.04825	.01036	-.41565	-4.65634	.00000	.24008	4.16525
	inc_blu_r	-.00363	.01433	-.01673	-.25361	.79997	.43937	2.27600
	car_own	-.62960	.13715	-.30036	-4.59064	.00001	.44686	2.23784
	sqrt_ptiori	-.00016	.00009	-.10770	-1.76887	.07793	.51604	1.93782
	ln_strio	1.62695	.29760	.43081	5.46697	.00000	.30806	3.24611
	ln_avedist	-.71037	.08829	-.56162	-8.04627	.00000	.39266	2.54671
	ln_lvr	-.26549	.06047	-.31612	-4.39006	.00002	.36892	2.71058
	wodm	.00361	.00171	.14997	2.11450	.03529	.38028	2.62962
	ln_jwrm2	.08370	.03180	.18494	2.63196	.00893	.38744	2.58105

a. Dependent Variable: ln_train_wo

APPENDIX 16 MANUFACTURING SECTOR – LUTI WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656[a]	.430	.402	.67226

a. Predictors: (Constant), wodm_ln, ln_avedist, inc_pr_r, p_woret, p_worker_man, p_wocon, sqrt_ptiori, car_own, inc_blu_r, ln_jwrm2, ln_lvr, ln_strio, pmanpr_r, pblu_r, wodm
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	102.292	15	6.819	15.090	.000[b]
	Residual	135.579	300	.452		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), wodm_ln, ln_avedist, inc_pr_r, p_woret, p_worker_man, p_wocon, sqrt_ptiori, car_own, inc_blu_r, ln_jwrm2, ln_lvr, ln_strio, pmanpr_r, pblu_r, wodm

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-12.57016	2.89547		-4.34132	.00002		
	pmanpr_r	1.52881	.75795	.22009	2.01704	.04458	.15958	6.26650
	pblu_r	-1.90667	.84018	-.25310	-2.26935	.02396	.15274	6.54714
	p_worker_man	2.10105	.97971	.12348	2.14458	.03279	.57305	1.74506
	p_wocon	3.29883	1.85609	.13286	1.77730	.07653	.33997	2.94144
	p_woret	2.62716	1.77285	.08196	1.48188	.13942	.62112	1.60999
	inc_pr_r	-.04633	.01038	-.39915	-4.46203	.00001	.23743	4.21184
	inc_blu_r	-.00343	.01428	-.01579	-.24013	.81040	.43934	2.27616
	car_own	-.64607	.13700	-.30822	-4.71582	.00000	.44476	2.24840
	sqrt_ptiori	-.00018	.00009	-.12033	-1.96937	.04983	.50887	1.96515
	ln_strio	1.65554	.29702	.43838	5.57378	.00000	.30713	3.25592
	ln_avedist	-.84925	.11837	-.67142	-7.17431	.00000	.21692	4.60994
	ln_lvr	-.29654	.06281	-.35310	-4.72087	.00000	.33961	2.94455
	wodm	-.00127	.00326	-.05283	-.38977	.69699	.10343	9.66808
	ln_jwrm2	.08767	.03177	.19372	2.75932	.00615	.38547	2.59424
	wodm_ln	.00340	.00194	.21883	1.75371	.08050	.12202	8.19539

a. Dependent Variable: ln_train_wo

APPENDIX 17 RETAIL SECTOR - LUTI

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.654[a]	.427	.401	.67268

a. Predictors: (Constant), ln_jwrr2, p_worker_man, ln_avedist, p_woret, inc_blu_r, sqrt_ptiori, p_wocon, car_own, inc_pr_r, wodr, ln_lvr, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	101.669	14	7.262	16.049	.000[b]
	Residual	136.202	301	.452		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ln_jwrr2, p_worker_man, ln_avedist, p_woret, inc_blu_r, sqrt_ptiori, p_wocon, car_own, inc_pr_r, wodr, ln_lvr, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-10.65856	2.76690		-3.85216	.00014		
	pmanpr_r	1.29980	.74697	.18712	1.74009	.08287	.16451	6.07868
	pblu_r	-1.89865	.83714	-.25204	-2.26801	.02404	.15404	6.49168
	p_worker_man	2.24547	.97114	.13197	2.31219	.02144	.58393	1.71253
	p_wocon	3.50667	1.86371	.14123	1.88155	.06086	.33762	2.96192
	p_woret	2.66258	1.74696	.08306	1.52412	.12853	.64047	1.56134
	inc_pr_r	-.05068	.00975	-.43662	-5.20026	.00000	.26984	3.70585
	inc_blu_r	.00093	.01436	.00426	.06444	.94867	.43485	2.29964
	car_own	-.61199	.13738	-.29196	-4.45465	.00001	.44284	2.25815
	sqrt_ptiori	-.00020	.00009	-.13376	-2.14680	.03261	.49003	2.04070
	ln_strio	1.41158	.28904	.37378	4.88368	.00000	.32474	3.07938
	ln_avedist	-.69496	.08875	-.54943	-7.83062	.00000	.38641	2.58794
	ln_lvr	-.27287	.06100	-.32491	-4.47340	.00001	.36059	2.77323
	wodr	.00195	.00121	.10372	1.61182	.10805	.45941	2.17670
	ln_jwrr2	.09836	.03301	.17721	2.97975	.00312	.53784	1.85929

a. Dependent Variable: ln_train_wo

APPENDIX 18 RETAIL SECTOR – LUTI WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.660[a]	.436	.408	.66887

a. Predictors: (Constant), wodr_ln, p_worker_man, p_wocon, p_woret, inc_pr_r, ln_avedist, ln_jwrr2, car_own, inc_blu_r, sqrt_ptiori, ln_lvr, ln_strio, pmanpr_r, pblu_r, wodr
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	103.655	15	6.910	15.446	.000[b]
	Residual	134.215	300	.447		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), wodr_ln, p_worker_man, p_wocon, p_woret, inc_pr_r, ln_avedist, ln_jwrr2, car_own, inc_blu_r, sqrt_ptiori, ln_lvr, ln_strio, pmanpr_r, pblu_r, wodr

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-10.67578	2.75123		-3.88036	.00013		
	pmanpr_r	1.49072	.74824	.21460	1.99229	.04725	.16210	6.16912

	pblu_r	-1.74249	.83569	-.23131	-2.08509	.03791	.15283	6.54313
	p_worker_man	1.94283	.97626	.11419	1.99008	.04749	.57130	1.75040
	p_wocon	3.24347	1.85735	.13063	1.74629	.08178	.33609	2.97537
	p_woret	2.16186	1.75324	.06744	1.23307	.21852	.62871	1.59056
	inc_pr_r	-.04917	.00972	-.42360	-5.06004	.00000	.26837	3.72616
	inc_blu_r	.00042	.01428	.00192	.02924	.97669	.43473	2.30030
	car_own	-.62893	.13684	-.30004	-4.59607	.00001	.44131	2.26596
	sqrt_ptiori	-.00023	.00009	-.15200	-2.42992	.01569	.48064	2.08056
	ln_strio	1.46213	.28840	.38717	5.06977	.00000	.32249	3.10083
	ln_avedist	-.83903	.11163	-.66333	-7.51616	.00000	.24147	4.14123
	ln_lvr	-.30606	.06266	-.36443	-4.88413	.00000	.33781	2.96022
	wodr	-.00182	.00216	-.09705	-.84564	.39843	.14281	7.00243
	ln_jwrr2	.10231	.03288	.18433	3.11202	.00204	.53609	1.86535
	wodr_ln	.00289	.00137	.21663	2.10735	.03592	.17798	5.61854

a. Dependent Variable: ln_train_wo

APPENDIX 19 ALL SECTOR SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.660[a]	.436	.408	.66877

a. Predictors: (Constant), edertt1000, p_woret, inc_blu_r, p_worker_man, car_own, ln_jwr, p_wocon, inc_pr_r, ln_avedist, ln_lvr, sqrt_ptiori, ln_strio, pmanpr_r, wod, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	103.696	15	6.913	15.457	.000[b]
	Residual	134.175	300	.447		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), edertt1000, p_woret, inc_blu_r, p_worker_man, car_own, ln_jwr, p_wocon, inc_pr_r, ln_avedist, ln_lvr, sqrt_ptiori, ln_strio, pmanpr_r, wod, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-16.566	3.269		-5.068	.000		
	pmanpr_r	1.290	.744	.186	1.734	.084	.164	6.102
	pblu_r	-1.542	.853	-.205	-1.808	.072	.147	6.811
	p_worker_man	2.112	.970	.124	2.178	.030	.579	1.729
	p_wocon	2.639	1.865	.106	1.415	.158	.333	3.001
	p_woret	2.992	1.736	.093	1.723	.086	.641	1.560
	inc_pr_r	-.047	.011	-.407	-4.488	.000	.228	4.384
	inc_blu_r	-.002	.014	-.011	-.164	.870	.436	2.293
	car_own	-.665	.140	-.317	-4.750	.000	.421	2.376
	sqrt_ptiori	.000	.000	-.224	-3.106	.002	.361	2.770
	ln_strio	1.914	.322	.507	5.945	.000	.259	3.863
	ln_avedist	-.655	.090	-.518	-7.312	.000	.375	2.668
	ln_lvr	-.259	.060	-.308	-4.282	.000	.364	2.749
	wod	.000	.000	-.061	-.657	.512	.216	4.629
	ln_jwr	.083	.041	.145	2.055	.041	.377	2.651
	edertt1000	.113	.037	.361	3.071	.002	.136	7.350

a. Dependent Variable: ln_train_wo

APPENDIX 20 ALL SECTOR SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.662[a]	.438	.408	.66864

a. Predictors: (Constant), edertt_ln, inc_pr_r, p_woret, sqrt_ptiori, p_worker_man, p_wocon, wod, car_own, inc_blu_r, ln_jwr, ln_strio, ln_lvr, pmanpr_r, pblu_r, edertt1000, ln_avedist
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.195	16	6.512	14.566	.000[b]
	Residual	133.676	299	.447		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), edertt_ln, inc_pr_r, p_woret, sqrt_ptiori, p_worker_man, p_wocon, wod, car_own, inc_blu_r, ln_jwr, ln_strio, ln_lvr, pmanpr_r, pblu_r, edertt1000, ln_avedist

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-16.38648	3.27288		-5.00675	.00000		
	pmanpr_r	1.37004	.74773	.19723	1.83227	.06791	.16221	6.16496
	pblu_r	-1.42843	.85905	-.18962	-1.66280	.09740	.14453	6.91883
	p_worker_man	1.96653	.97958	.11558	2.00752	.04559	.56704	1.76354
	p_wocon	2.49726	1.86940	.10058	1.33587	.18261	.33155	3.01615
	p_woret	2.65809	1.76401	.08292	1.50684	.13291	.62063	1.61127
	inc_pr_r	-.04522	.01072	-.38961	-4.21923	.00003	.22042	4.53678
	inc_blu_r	-.00212	.01426	-.00976	-.14862	.88195	.43593	2.29394
	car_own	-.66659	.14007	-.31801	-4.75911	.00000	.42093	2.37568
	sqrt_ptiori	-.00032	.00011	-.21424	-2.94485	.00348	.35510	2.81611
	ln_strio	1.94747	.32340	.51568	6.02194	.00000	.25630	3.90170
	ln_avedist	-.91316	.26028	-.72194	-3.50838	.00052	.04439	22.52917
	ln_lvr	-.27514	.06238	-.32762	-4.41059	.00001	.34064	2.93566
	wod	-.00006	.00018	-.03361	-.34694	.72888	.20026	4.99348
	ln_jwr	.09240	.04148	.16060	2.22755	.02665	.36157	2.76574
	edertt1000	.07017	.05493	.22378	1.27741	.20245	.06124	16.32805
	edertt_ln	.02385	.02258	.16029	1.05623	.29172	.08161	12.25330

a. Dependent Variable: ln_train_wo

APPENDIX 21 ALL SECTOR SETI H2 EFFECTIVE JOB DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.647[a]	.419	.390	.67866

a. Predictors: (Constant), ejdtt1000, p_worker_man, p_woret, inc_pr_r, p_wocon, wod, car_own, inc_blu_r, ln_avedist, sqrt_ptiori, ln_lvr, ln_jwr, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	99.696	15	6.646	14.430	.000[b]
	Residual	138.175	300	.461		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdtt1000, p_worker_man, p_woret, inc_pr_r, p_wocon, wod, car_own, inc_blu_r, ln_avedist, sqrt_ptiori, ln_lvr, ln_jwr, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-10.77700	2.93666		-3.66982	.00029		

	pmanpr_r	1.18700	.76266	.17088	1.55640	.12067	.16063	6.22545
	pblu_r	-1.86851	.86074	-.24803	-2.17083	.03073	.14832	6.74231
	p_worker_man	2.31197	.98353	.13588	2.35067	.01939	.57948	1.72568
	p_wocon	3.38176	1.89226	.13620	1.78715	.07492	.33336	2.99977
	p_woret	2.87416	1.76108	.08966	1.63205	.10372	.64151	1.55883
	inc_pr_r	-.04769	.01081	-.41088	-4.41041	.00001	.22310	4.48238
	inc_blu_r	.00058	.01462	.00265	.03935	.96864	.42728	2.34041
	car_own	-.63316	.14429	-.30206	-4.38823	.00002	.40865	2.44709
	sqrt_ptiori	-.00012	.00010	-.08339	-1.22471	.22164	.41768	2.39420
	ln_strio	1.41685	.30376	.37518	4.66438	.00000	.29928	3.34131
	ln_avedist	-.68587	.09071	-.54225	-7.56124	.00000	.37649	2.65608
	ln_lvr	-.26710	.06160	-.31805	-4.33629	.00002	.35993	2.77829
	wod	.00021	.00014	.10918	1.43018	.15371	.33224	3.00991
	ln_jwr	.09675	.04577	.16817	2.11370	.03537	.30589	3.26914
	ejdtt1000	-.00969	.01410	-.05251	-.68730	.49243	.33176	3.01427

a. Dependent Variable: ln_train_wo

APPENDIX 22 ALL SECTOR SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656[a]	.430	.400	.67323

a. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, inc_pr_r, p_wocon, sqrt_ptiori, ln_jwr, ln_strio, inc_blu_r, ln_lvr, wod, ejdtt1000, pmanpr_r, pblu_r, ln_avedist
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	102.354	16	6.397	14.114	.000[b]
	Residual	135.517	299	.453		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdtt_ln, p_woret, p_worker_man, car_own, inc_pr_r, p_wocon, sqrt_ptiori, ln_jwr, ln_strio, inc_blu_r, ln_lvr, wod, ejdtt1000, pmanpr_r, pblu_r, ln_avedist

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-13.37449	3.10435		-4.30830	.00002		
	pmanpr_r	1.47420	.76579	.21223	1.92508	.05517	.15678	6.37846
	pblu_r	-1.41603	.87405	-.18797	-1.62008	.10627	.14154	7.06522
	p_worker_man	2.11554	.97902	.12434	2.16087	.03150	.57550	1.73761
	p_wocon	2.93461	1.88617	.11819	1.55586	.12080	.33016	3.02880
	p_woret	2.21598	1.76799	.06913	1.25339	.21104	.62635	1.59656
	inc_pr_r	-.04283	.01091	-.36902	-3.92501	.00011	.21555	4.63921
	inc_blu_r	.00070	.01450	.00321	.04808	.96169	.42727	2.34044
	car_own	-.65183	.14334	-.31097	-4.54747	.00001	.40747	2.45419
	sqrt_ptiori	-.00016	.00010	-.10842	-1.58676	.11362	.40812	2.45029
	ln_strio	1.71882	.32611	.45514	5.27066	.00000	.25552	3.91359
	ln_avedist	-1.08579	.18807	-.85842	-5.77318	.00000	.08618	11.60345
	ln_lvr	-.28846	.06174	-.34348	-4.67246	.00000	.35258	2.83620
	wod	.00024	.00014	.12903	1.69398	.09131	.32839	3.04519
	ln_jwr	.09693	.04541	.16849	2.13478	.03359	.30589	3.26915
	ejdtt1000	-.04413	.01995	-.23909	-2.21226	.02770	.16312	6.13037
	ejdtt_ln	.04432	.01830	.27641	2.42151	.01605	.14623	6.83836

a. Dependent Variable: ln_train_wo

APPENDIX 23 CONSTRUCTION SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.674[a]	.455	.427	.65763

a. Predictors: (Constant), ederttc100, p_wocon, p_woret, p_worker_man, inc_pr_r, car_own, ln_jwrc, ln_avedist, inc_blu_r, sqrt_ptiori, ln_lvr, ln_strio, pmanpr_r, wodc, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	108.126	15	7.208	16.667	.000[b]
	Residual	129.745	300	.432		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttc100, p_wocon, p_woret, p_worker_man, inc_pr_r, car_own, ln_jwrc, ln_avedist, inc_blu_r, sqrt_ptiori, ln_lvr, ln_strio, pmanpr_r, wodc, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-13.63463	2.82315		-4.82958	.00000		
	pmanpr_r	1.55451	.73622	.22379	2.11147	.03556	.16186	6.17832
	pblu_r	-1.07774	.84658	-.14306	-1.27305	.20399	.14397	6.94613
	p_worker_man	2.06002	.94939	.12107	2.16982	.03080	.58396	1.71243
	p_wocon	1.18725	1.91323	.04782	.62054	.53537	.30620	3.26588
	p_woret	2.97299	1.70210	.09275	1.74666	.08172	.64484	1.55078
	inc_pr_r	-.04085	.01025	-.35191	-3.98341	.00009	.23295	4.29273
	inc_blu_r	.00082	.01402	.00377	.05848	.95341	.43650	2.29095
	car_own	-.64963	.13766	-.30992	-4.71913	.00000	.42156	2.37217
	sqrt_ptiori	-.00029	.00010	-.19576	-2.80936	.00529	.37447	2.67048
	ln_strio	1.63230	.28868	.43223	5.65432	.00000	.31115	3.21392
	ln_avedist	-.65655	.08720	-.51907	-7.52890	.00000	.38251	2.61432
	ln_lvr	-.29145	.05957	-.34704	-4.89215	.00000	.36130	2.76778
	wodc	-.00009	.00204	-.00468	-.04430	.96470	.16281	6.14211
	ln_jwrc	.12780	.03664	.22759	3.48833	.00056	.42712	2.34128
	ederttc100	.13117	.05145	.29230	2.54946	.01129	.13832	7.22982

a. Dependent Variable: ln_train_wo

APPENDIX 24 CONSTRUCTION SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.676[a]	.457	.428	.65735

a. Predictors: (Constant), ederttc_ln, inc_pr_r, sqrt_ptiori, p_woret, p_worker_man, wodc, p_wocon, ln_jwrc, car_own, inc_blu_r, ln_strio, ln_lvr, pmanpr_r, pblu_r, ln_avedist, ederttc100
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	108.670	16	6.792	15.718	.000[b]
	Residual	129.200	299	.432		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttc_ln, inc_pr_r, sqrt_ptiori, p_woret, p_worker_man, wodc, p_wocon, ln_jwrc, car_own, inc_blu_r, ln_strio, ln_lvr, pmanpr_r, pblu_r, ln_avedist, ederttc100

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-13.39004	2.83034		-4.73090	.00000		

	pmanpr_r	1.60346	.73720	.23083	2.17508	.03041	.16129	6.20003
	pblu_r	-1.02533	.84750	-.13611	-1.20982	.22730	.14353	6.96728
	p_worker_man	1.90862	.95852	.11217	1.99121	.04737	.57240	1.74704
	p_wocon	1.19248	1.91241	.04803	.62355	.53340	.30619	3.26589
	p_woret	2.61930	1.73031	.08171	1.51377	.13114	.62345	1.60399
	inc_pr_r	-.03976	.01030	-.34253	-3.86158	.00014	.23088	4.33122
	inc_blu_r	.00106	.01401	.00486	.07530	.94003	.43640	2.29146
	car_own	-.65367	.13765	-.31185	-4.74890	.00000	.42127	2.37379
	sqrt_ptiori	-.00028	.00011	-.18447	-2.62128	.00921	.36681	2.72618
	ln_strio	1.65640	.28935	.43861	5.72445	.00000	.30943	3.23172
	ln_avedist	-.86194	.20271	-.68145	-4.25212	.00003	.07073	14.13854
	ln_lvr	-.31177	.06224	-.37124	-5.00908	.00000	.33073	3.02365
	wodc	.00014	.00205	.00721	.06787	.94593	.16119	6.20386
	ln_jwrc	.13119	.03675	.23363	3.57023	.00042	.42423	2.35719
	ederttc100	.06674	.07707	.14872	.86590	.38724	.06158	16.23929
	ederttc_ln	.03785	.03372	.15732	1.12229	.26264	.09245	10.81666

a. Dependent Variable: ln_train_wo

APPENDIX 25 CONSTRUCTION SETI H2 EFFECTIVE JOB DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.665[a]	.443	.415	.66463

a. Predictors: (Constant), ejdttc100, inc_pr_r, p_woret, p_worker_man, wodc, p_wocon, car_own, ln_avedist, sqrt_ptiori, inc_blu_r, ln_lvr, ln_jwrc, ln_strio, pmanpr_r, pblu_r

b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	105.351	15	7.023	15.900	.000[b]
	Residual	132.520	300	.442		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo

b. Predictors: (Constant), ejdttc100, inc_pr_r, p_woret, p_worker_man, wodc, p_wocon, car_own, ln_avedist, sqrt_ptiori, inc_blu_r, ln_lvr, ln_jwrc, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-11.99785	2.83016		-4.23928	.00003		
	pmanpr_r	1.26959	.73540	.18277	1.72639	.08531	.16569	6.03541
	pblu_r	-1.34755	.85211	-.17888	-1.58143	.11483	.14514	6.88975
	p_worker_man	2.13670	.95960	.12558	2.22666	.02671	.58384	1.71281
	p_wocon	1.63109	1.93512	.06569	.84289	.39996	.30571	3.27106
	p_woret	2.99157	1.72106	.09333	1.73822	.08320	.64420	1.55231
	inc_pr_r	-.04048	.01041	-.34876	-3.88974	.00012	.23100	4.32905
	inc_blu_r	.00060	.01428	.00277	.04205	.96649	.42943	2.32865
	car_own	-.57508	.13635	-.27435	-4.21753	.00003	.43886	2.27865
	sqrt_ptiori	-.00016	.00009	-.10499	-1.67440	.09509	.47231	2.11724
	ln_strio	1.50243	.29396	.39784	5.11095	.00000	.30648	3.26280
	ln_avedist	-.66064	.08812	-.52230	-7.49669	.00000	.38258	2.61381
	ln_lvr	-.29171	.06024	-.34735	-4.84216	.00000	.36087	2.77107
	wodc	.00402	.00125	.20799	3.20627	.00149	.44132	2.26593
	ln_jwrc	.13322	.04543	.23723	2.93260	.00362	.28378	3.52385
	ejdttc100	.00597	.02103	.01877	.28390	.77668	.42495	2.35322

a. Dependent Variable: ln_train_wo

APPENDIX 26 CONSTRUCTION SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.684[a]	.467	.439	.65093

a. Predictors: (Constant), ejdttc_ln, ln_strio, p_woret, p_worker_man, wodc, inc_blu_r, car_own, sqrt_ptiori, inc_pr_r, p_wocon, ln_lvr, ln_jwrc, ln_avedist, pmanpr_r, pblu_r, ejdttc100
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	111.180	16	6.949	16.400	.000[b]
	Residual	126.690	299	.424		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttc_ln, ln_strio, p_woret, p_worker_man, wodc, inc_blu_r, car_own, sqrt_ptiori, inc_pr_r, p_wocon, ln_lvr, ln_jwrc, ln_avedist, pmanpr_r, pblu_r, ejdttc100

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-14.41108	2.84716		-5.06156	.00000		
	pmanpr_r	2.05237	.75052	.29546	2.73458	.00662	.15259	6.55356
	pblu_r	-.59023	.85916	-.07835	-.68699	.49262	.13695	7.30212
	p_worker_man	2.45944	.94384	.14455	2.60578	.00963	.57888	1.72749
	p_wocon	1.31817	1.89712	.05309	.69483	.48770	.30511	3.27755
	p_woret	2.06582	1.70396	.06445	1.21236	.22633	.63038	1.58634
	inc_pr_r	-.03123	.01049	-.26909	-2.97659	.00315	.21796	4.58804
	inc_blu_r	.00179	.01399	.00825	.12802	.89822	.42921	2.32988
	car_own	-.69083	.13714	-.32957	-5.03733	.00000	.41613	2.40307
	sqrt_ptiori	-.00014	.00009	-.09244	-1.50300	.13389	.47088	2.12367
	ln_strio	1.79192	.29830	.47449	6.00719	.00000	.28551	3.50256
	ln_avedist	-1.10293	.14720	-.87197	-7.49284	.00000	.13153	7.60292
	ln_lvr	-.34378	.06065	-.40935	-5.66829	.00000	.34154	2.92790
	wodc	.00332	.00124	.17186	2.67392	.00791	.43119	2.31917
	ln_jwrc	.10106	.04533	.17997	2.22961	.02652	.27340	3.65765
	ejdttc100	-.12393	.04063	-.38949	-3.05018	.00249	.10924	9.15407
	ejdttc_ln	.12522	.03376	.49522	3.70927	.00025	.09993	10.00676

a. Dependent Variable: ln_train_wo

APPENDIX 27 MANUFACTURING SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.663[a]	.440	.412	.66647

a. Predictors: (Constant), ederttm100, p_worker_man, p_woret, inc_blu_r, car_own, ln_jwrm2, pmanpr_r, ln_avedist, p_wocon, ln_lvr, sqrt_ptiori, inc_pr_r, ln_strio, wodm, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.618	15	6.975	15.702	.000[b]
	Residual	133.253	300	.444		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttm100, p_worker_man, p_woret, inc_blu_r, car_own, ln_jwrm2, pmanpr_r, ln_avedist, p_wocon, ln_lvr, sqrt_ptiori, inc_pr_r, ln_strio, wodm, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-17.62780	3.32961		-5.29425	.00000		

	pmanpr_r	1.80234	.75973	.25946	2.37236	.01831	.15611	6.40584
	pblu_r	-1.85631	.83148	-.24641	-2.23255	.02632	.15328	6.52404
	p_worker_man	2.13979	.96072	.12576	2.22726	.02667	.58569	1.70738
	p_wocon	3.53355	1.83805	.14232	1.92245	.05550	.34073	2.93487
	p_woret	3.41149	1.73902	.10643	1.96173	.05072	.63445	1.57616
	inc_pr_r	-.04739	.01024	-.40824	-4.62708	.00001	.23988	4.16876
	inc_blu_r	-.00848	.01426	-.03905	-.59491	.55235	.43330	2.30790
	car_own	-.68862	.13703	-.32852	-5.02536	.00000	.43695	2.28859
	sqrt_ptiori	-.00032	.00010	-.21069	-3.01399	.00280	.38214	2.61682
	ln_strio	2.03858	.32665	.53981	6.24079	.00000	.24958	4.00668
	ln_avedist	-.68531	.08765	-.54180	-7.81824	.00000	.38882	2.57185
	ln_lvr	-.25722	.05982	-.30629	-4.30026	.00002	.36808	2.71678
	wodm	.00007	.00209	.00279	.03225	.97429	.24898	4.01635
	ln_jwrm2	.08735	.03144	.19300	2.77776	.00582	.38682	2.58521
	ederttm100	.13319	.04605	.30362	2.89230	.00410	.16945	5.90162

a. Dependent Variable: ln_train_wo

APPENDIX 28 MANUFACTURING SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.664[a]	.440	.410	.66724

a. Predictors: (Constant), ederttm_ln, ederttm100, p_woret, p_worker_man, inc_pr_r, p_wocon, car_own, ln_jwrm2, inc_blu_r, ln_lvr, sqrt_ptiori, ln_strio, wodm, pmanpr_r, pblu_r, ln_avedist
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.751	16	6.547	14.705	.000[b]
	Residual	133.119	299	.445		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttm_ln, ederttm100, p_woret, p_worker_man, inc_pr_r, p_wocon, car_own, ln_jwrm2, inc_blu_r, ln_lvr, sqrt_ptiori, ln_strio, wodm, pmanpr_r, pblu_r, ln_avedist

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-17.43920	3.35123		-5.20382	.00000		
	pmanpr_r	1.81580	.76101	.26140	2.38605	.01765	.15594	6.41253
	pblu_r	-1.80916	.83688	-.24016	-2.16179	.03143	.15166	6.59373
	p_worker_man	2.06253	.97213	.12122	2.12167	.03469	.57337	1.74407
	p_wocon	3.48039	1.84275	.14018	1.88870	.05990	.33979	2.94303
	p_woret	3.24396	1.76769	.10120	1.83515	.06748	.61548	1.62476
	inc_pr_r	-.04654	.01037	-.40095	-4.48856	.00001	.23456	4.26329
	inc_blu_r	-.00852	.01427	-.03924	-.59697	.55098	.43328	2.30795
	car_own	-.69240	.13736	-.33032	-5.04071	.00000	.43584	2.29439
	sqrt_ptiori	-.00031	.00011	-.20544	-2.90843	.00390	.37512	2.66579
	ln_strio	2.05079	.32779	.54304	6.25633	.00000	.24843	4.02527
	ln_avedist	-.82876	.27610	-.65521	-3.00165	.00291	.03928	25.45731
	ln_lvr	-.26551	.06176	-.31615	-4.29872	.00002	.34603	2.88989
	wodm	.00019	.00210	.00780	.08942	.92881	.24625	4.06089
	ln_jwrm2	.08888	.03161	.19639	2.81221	.00525	.38377	2.60575
	ederttm100	.10404	.07040	.23716	1.47775	.14053	.07267	13.76122
	ederttm_ln	.01646	.03004	.09393	.54797	.58412	.06371	15.69718

a. Dependent Variable: ln_train_wo

APPENDIX 29 MANUFACTURING SETI H2 EFFECTIVE JOB DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.665[a]	.443	.415	.66476

a. Predictors: (Constant), ejdttm100, pmanpr_r, wodm, inc_blu_r, p_woret, car_own, p_worker_man, ln_avedist, sqrt_ptiori, p_wocon, ln_lvr, ln_strio, inc_pr_r, ln_jwrm2, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	105.301	15	7.020	15.886	.000[b]
	Residual	132.570	300	.442		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttm100, pmanpr_r, wodm, inc_blu_r, p_woret, car_own, p_worker_man, ln_avedist, sqrt_ptiori, p_wocon, ln_lvr, ln_strio, inc_pr_r, ln_jwrm2, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-15.61103	3.00274		-5.19893	.00000		
	pmanpr_r	1.94558	.76420	.28008	2.54592	.01140	.15349	6.51489
	pblu_r	-1.34031	.85715	-.17792	-1.56369	.11894	.14350	6.96882
	p_worker_man	1.12957	1.03547	.06639	1.09088	.27620	.50160	1.99361
	p_wocon	3.50095	1.83320	.14100	1.90975	.05712	.34078	2.93445
	p_woret	3.18458	1.73165	.09935	1.83904	.06690	.63658	1.57089
	inc_pr_r	-.04387	.01030	-.37795	-4.25751	.00003	.23573	4.24208
	inc_blu_r	-.00706	.01416	-.03251	-.49844	.61854	.43678	2.28946
	car_own	-.66753	.13569	-.31846	-4.91967	.00000	.44335	2.25555
	sqrt_ptiori	-.00021	.00009	-.13913	-2.28753	.02286	.50220	1.99124
	ln_strio	1.85874	.30233	.49219	6.14810	.00000	.28987	3.44984
	ln_avedist	-.69740	.08710	-.55136	-8.00695	.00000	.39179	2.55240
	ln_lvr	-.28086	.05979	-.33443	-4.69713	.00000	.36648	2.72870
	wodm	.00272	.00171	.11312	1.59625	.11149	.36995	2.70308
	ln_jwrm2	.00241	.04057	.00532	.05937	.95270	.23117	4.32574
	ejdttm100	.05891	.01867	.25604	3.15503	.00177	.28208	3.54510

a. Dependent Variable: ln_train_wo

APPENDIX 30 MANUFACTURING SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.680[a]	.462	.433	.65413

a. Predictors: (Constant), ejdttm_ln, p_woret, p_wocon, p_worker_man, wodm, inc_pr_r, car_own, sqrt_ptiori, inc_blu_r, ln_lvr, ln_strio, ln_jwrm2, ln_avedist, pmanpr_r, pblu_r, ejdttm100
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	109.931	16	6.871	16.057	.000[b]
	Residual	127.940	299	.428		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttm_ln, p_woret, p_wocon, p_worker_man, wodm, inc_pr_r, car_own, sqrt_ptiori, inc_blu_r, ln_lvr, ln_strio, ln_jwrm2, ln_avedist, pmanpr_r, pblu_r, ejdttm100

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-17.12076	2.99019		-5.72564	.00000		

	pmanpr_r	2.68384	.78476	.38636	3.41994	.00071	.14094	7.09512
	pblu_r	-.60964	.87221	-.08093	-.69897	.48512	.13419	7.45209
	p_worker_man	2.04131	1.05595	.11997	1.93315	.05416	.46705	2.14112
	p_wocon	3.61288	1.80423	.14551	2.00245	.04614	.34066	2.93549
	p_woret	2.57520	1.71402	.08034	1.50243	.13404	.62915	1.58945
	inc_pr_r	-.03478	.01051	-.29968	-3.31002	.00105	.21945	4.55681
	inc_blu_r	-.00676	.01394	-.03113	-.48506	.62799	.43677	2.28955
	car_own	-.73684	.13517	-.35152	-5.45118	.00000	.43258	2.31171
	sqrt_ptiori	-.00017	.00009	-.11350	-1.88048	.06101	.49383	2.02500
	ln_strio	2.04692	.30295	.54202	6.75668	.00000	.27953	3.57738
	ln_avedist	-1.23291	.18397	-.97474	-6.70159	.00000	.08503	11.76047
	ln_lvr	-.31403	.05970	-.37393	-5.26054	.00000	.35602	2.80886
	wodm	.00250	.00168	.10369	1.48569	.13841	.36932	2.70764
	ln_jwrm2	.00136	.03992	.00300	.03405	.97286	.23116	4.32601
	ejdttm100	-.04035	.03533	-.17535	-1.14207	.25434	.07631	13.10505
	ejdttm_ln	.07349	.02234	.54109	3.28963	.00112	.06649	15.04026

a. Dependent Variable: ln_train_wo

APPENDIX 31 RETAIL SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.666[a]	.443	.415	.66453

a. Predictors: (Constant), ederttr100, inc_blu_r, pmanpr_r, ln_jwrr2, p_woret, p_worker_man, car_own, ln_avedist, p_wocon, ln_lvr, sqrt_ptiori, ln_strio, inc_pr_r, wodr, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	105.390	15	7.026	15.910	.000[b]
	Residual	132.480	300	.442		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttr100, inc_blu_r, pmanpr_r, ln_jwrr2, p_woret, p_worker_man, car_own, ln_avedist, p_wocon, ln_lvr, sqrt_ptiori, ln_strio, inc_pr_r, wodr, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-12.57330	2.81183		-4.47157	.00001		
	pmanpr_r	1.71112	.75140	.24633	2.27724	.02348	.15866	6.30279
	pblu_r	-1.61374	.83281	-.21422	-1.93772	.05360	.15190	6.58310
	p_worker_man	2.14351	.96002	.12598	2.23278	.02630	.58315	1.71482
	p_wocon	3.23806	1.84346	.13042	1.75652	.08002	.33677	2.96940
	p_woret	2.75315	1.72608	.08589	1.59503	.11176	.64026	1.56185
	inc_pr_r	-.05121	.00963	-.44121	-5.31833	.00000	.26975	3.70719
	inc_blu_r	-.00047	.01420	-.00217	-.03327	.97348	.43435	2.30229
	car_own	-.70458	.13942	-.33613	-5.05380	.00000	.41967	2.38284
	sqrt_ptiori	-.00039	.00011	-.26232	-3.45948	.00062	.32288	3.09714
	ln_strio	1.57998	.29137	.41837	5.42257	.00000	.31187	3.20646
	ln_avedist	-.69301	.08768	-.54789	-7.90426	.00000	.38638	2.58810
	ln_lvr	-.27591	.06027	-.32854	-4.57812	.00001	.36048	2.77407
	wodr	-.00196	.00180	-.10420	-1.08814	.27741	.20244	4.93981
	ln_jwrr2	.09489	.03263	.17096	2.90787	.00391	.53712	1.86179
	ederttr100	.15186	.05231	.32739	2.90306	.00397	.14598	6.85040

a. Dependent Variable: ln_train_wo

APPENDIX 32 RETAIL SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.666[a]	.444	.414	.66499

a. Predictors: (Constant), ederttr_ln, ln_lvr, inc_blu_r, p_woret, p_worker_man, ln_jwrr2, p_wocon, car_own, sqrt_ptiori, inc_pr_r, wodr, ln_strio, pmanpr_r, ln_avedist, pblu_r, ederttr100
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	105.650	16	6.603	14.932	.000[b]
	Residual	132.221	299	.442		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttr_ln, ln_lvr, inc_blu_r, p_woret, p_worker_man, ln_jwrr2, p_wocon, car_own, sqrt_ptiori, inc_pr_r, wodr, ln_strio, pmanpr_r, ln_avedist, pblu_r, ederttr100

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-12.51484	2.81481		-4.44607	.00001		
	pmanpr_r	1.74574	.75328	.25132	2.31753	.02115	.15809	6.32557
	pblu_r	-1.57956	.83457	-.20968	-1.89265	.05937	.15147	6.60199
	p_worker_man	2.02995	.97206	.11931	2.08829	.03762	.56958	1.75569
	p_wocon	3.17796	1.84640	.12800	1.72117	.08625	.33616	2.97477
	p_woret	2.50722	1.75688	.07822	1.42709	.15460	.61887	1.61585
	inc_pr_r	-.05045	.00969	-.43466	-5.20823	.00000	.26691	3.74653
	inc_blu_r	-.00071	.01421	-.00325	-.04968	.96041	.43415	2.30335
	car_own	-.70383	.13952	-.33578	-5.04485	.00000	.41965	2.38295
	sqrt_ptiori	-.00038	.00011	-.25319	-3.29619	.00110	.31509	3.17371
	ln_strio	1.60596	.29354	.42525	5.47102	.00000	.30770	3.24990
	ln_avedist	-.81700	.18416	-.64592	-4.43639	.00001	.08770	11.40258
	ln_lvr	-.28896	.06267	-.34407	-4.61086	.00001	.33385	2.99535
	wodr	-.00175	.00182	-.09286	-.95760	.33904	.19771	5.05792
	ln_jwrr2	.09616	.03270	.17325	2.94104	.00353	.53573	1.86661
	ederttr100	.10310	.08243	.22226	1.25065	.21204	.05886	16.98870
	ederttr_ln	.02782	.03633	.09967	.76574	.44443	.10974	9.11250

a. Dependent Variable: ln_train_wo

APPENDIX 33 RETAIL SETI H2 EFFECTIVE JOB DENSITY

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.654[a]	.428	.399	.67365

a. Predictors: (Constant), ejdttr100, p_woret, p_worker_man, inc_pr_r, p_wocon, wodr, car_own, ln_avedist, inc_blu_r, ln_jwrr2, ln_lvr, sqrt_ptiori, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	101.729	15	6.782	14.944	.000[b]
	Residual	136.142	300	.454		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttr100, p_woret, p_worker_man, inc_pr_r, p_wocon, wodr, car_own, ln_avedist, inc_blu_r, ln_jwrr2, ln_lvr, sqrt_ptiori, ln_strio, pmanpr_r, pblu_r

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-10.80947	2.80183		-3.85801	.00014		
	pmanpr_r	1.27988	.75005	.18425	1.70638	.08897	.16363	6.11131
	pblu_r	-1.92231	.84088	-.25518	-2.28608	.02295	.15312	6.53080
	p_worker_man	2.24215	.97259	.13178	2.30534	.02183	.58388	1.71268
	p_wocon	3.42262	1.88068	.13785	1.81989	.06977	.33252	3.00738
	p_woret	2.60691	1.75618	.08133	1.48442	.13875	.63560	1.57331
	inc_pr_r	-.05105	.00981	-.43982	-5.20238	.00000	.26692	3.74649
	inc_blu_r	.00051	.01443	.00234	.03516	.97197	.43209	2.31435
	car_own	-.61136	.13759	-.29166	-4.44325	.00001	.44277	2.25851
	sqrt_ptiori	-.00022	.00011	-.14599	-2.05930	.04033	.37959	2.63444
	ln_strio	1.42921	.29349	.37845	4.86967	.00000	.31588	3.16580
	ln_avedist	-.69183	.08929	-.54696	-7.74789	.00000	.38282	2.61221
	ln_lvr	-.27162	.06118	-.32343	-4.43945	.00001	.35945	2.78202
	wodr	.00195	.00121	.10386	1.61160	.10810	.45939	2.17678
	ln_jwrr2	.09061	.03933	.16325	2.30359	.02193	.37987	2.63248
	ejdttr100	.00693	.01906	.02652	.36351	.71648	.35856	2.78893

a. Dependent Variable: ln_train_wo

APPENDIX 34 RETAIL SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.658[a]	.433	.403	.67162

a. Predictors: (Constant), ejdttr_ln, wodr, inc_blu_r, pmanpr_r, car_own, p_woret, p_worker_man, ln_strio, ln_jwrr2, sqrt_ptiori, p_wocon, ln_lvr, inc_pr_r, ln_avedist, pblu_r, ejdttr100
b. Dependent Variable: ln_train_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	103.000	16	6.438	14.272	.000[b]
	Residual	134.870	299	.451		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttr_ln, wodr, inc_blu_r, pmanpr_r, car_own, p_woret, p_worker_man, ln_strio, ln_jwrr2, sqrt_ptiori, p_wocon, ln_lvr, inc_pr_r, ln_avedist, pblu_r, ejdttr100

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-11.91105	2.86938		-4.15109	.00004		
	pmanpr_r	1.42031	.75245	.20447	1.88757	.06005	.16161	6.18776
	pblu_r	-1.79339	.84185	-.23806	-2.13031	.03396	.15185	6.58558
	p_worker_man	2.15705	.97098	.12678	2.22152	.02706	.58229	1.71736
	p_wocon	3.12845	1.88317	.12600	1.66127	.09771	.32964	3.03364
	p_woret	2.07229	1.77959	.06465	1.16448	.24516	.61526	1.62534
	inc_pr_r	-.04894	.00986	-.42160	-4.96097	.00000	.26256	3.80861
	inc_blu_r	-.00138	.01443	-.00637	-.09590	.92366	.42945	2.32853
	car_own	-.64241	.13842	-.30647	-4.64107	.00001	.43487	2.29956
	sqrt_ptiori	-.00023	.00011	-.15246	-2.15384	.03205	.37846	2.64226
	ln_strio	1.58270	.30655	.41909	5.16289	.00000	.28779	3.47479
	ln_avedist	-.86665	.13699	-.68517	-6.32647	.00000	.16167	6.18538
	ln_lvr	-.28688	.06167	-.34160	-4.65177	.00000	.35164	2.84383
	wodr	.00166	.00122	.08858	1.36502	.17327	.45036	2.22045
	ln_jwrr2	.07570	.04021	.13638	1.88263	.06072	.36134	2.76751
	ejdttr100	-.03361	.03072	-.12863	-1.09391	.27487	.13715	7.29102
	ejdttr_ln	.04399	.02620	.18686	1.67906	.09418	.15312	6.53086

a. Dependent Variable: ln_train_wo

APPENDIX CHAPTER 6 AND CHAPTER 7
PLACE OF WORK (POW) MODEL

LUTI MODEL – POW MODEL

APPENDIX 35 ALL SECTOR LUTI POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.725[a]	.526	.504	.98807

a. Predictors: (Constant), ln_jwr, p_manpr, sqrt_ptiori, wage_bl, ln_avedist, p_jobcon, All_job, p_jobret, sr_lvnr, p_jobman, wage_mp, ln_strio, ln_jobd, p_blu

b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	325.775	14	23.270	23.835	.000[b]
	Residual	293.862	301	.976		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22

b. Predictors: (Constant), ln_jwr, p_manpr, sqrt_ptiori, wage_bl, ln_avedist, p_jobcon, All_job, p_jobret, sr_lvnr, p_jobman, wage_mp, ln_strio, ln_jobd, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.663	3.832		-3.565	.000		
	All_job	3.125E-5	.000	.120	2.272	.024	.561	1.784
	p_manpr	1.262	1.228	.078	1.028	.305	.274	3.656
	p_blu	-1.178	1.353	-.093	-.871	.385	.137	7.287
	p_jobman	1.260	1.260	.073	1.000	.318	.298	3.357
	p_jobcon	-.278	1.282	-.012	-.217	.828	.505	1.978
	p_jobret	1.905	1.025	.111	1.859	.064	.443	2.255
	wage_mp	.020	.021	.061	.937	.350	.373	2.682
	wage_bl	-.013	.022	-.036	-.591	.555	.425	2.353
	sqrt_ptiori	-8.563E-5	.000	-.035	-.544	.587	.371	2.698
	ln_strio	.843	.388	.138	2.171	.031	.388	2.576
	ln_avedist	-.696	.144	-.341	-4.844	.000	.318	3.140
	sr_lvnr	.011	.006	.118	1.818	.070	.373	2.681
	ln_jobd	.235	.082	.292	2.875	.004	.153	6.544
	ln_jwr	-.028	.058	-.030	-.483	.629	.405	2.466

a. Dependent Variable: ln_ptrainw22

APPENDIX 36 ALL SECTOR LUTI POW MODEL WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.742[a]	.550	.528	.96387

a. Predictors: (Constant), lnjobd_ln, p_blu, All_job, inc_mp, p_jobret, ln_strio, p_jobcon, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, p_jobman, ln_avedist, p_manpr, ln_jobd

b. Dependent Variable: Ln of proportion of train ridership

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	340.924	15	22.728	24.464	.000[b]
	Residual	278.713	300	.929		
	Total	619.637	315			

a. Dependent Variable: Ln of proportion of train ridership

b. Predictors: (Constant), lnjobd_ln, p_blu, All_job, inc_mp, p_jobret, ln_strio, p_jobcon, ln_jwr, sr_lvnr, inc_bl, sqrt_ptiori, p_jobman, ln_avedist, p_manpr, ln_jobd

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
1.00000	(Constant)	-14.57543	3.74531		-3.89165	.00012		
	All_job	.00002	.00001	.08992	1.72045	.08638	.54886	1.82197
	p_manpr	.76907	1.20368	.04755	.63893	.52336	.27069	3.69427
	p_blu	-1.40021	1.32100	-.11089	-1.05996	.29001	.13698	7.30011
	p_jobman	1.84629	1.23744	.10660	1.49203	.13674	.29374	3.40434
	p_jobcon	.82710	1.27976	.03603	.64630	.51858	.48234	2.07324
	p_jobret	2.30964	1.00475	.13435	2.29871	.02221	.43896	2.27813
	inc_mp	.02961	.02063	.09165	1.43496	.15234	.36753	2.72085
	inc_bl	-.01632	.02187	-.04436	-.74647	.45597	.42449	2.35575
	sqrt_ptiori	-.00014	.00015	-.05730	-.89760	.37012	.36793	2.71793
	ln_strio	.69767	.38050	.11446	1.83359	.06771	.38474	2.59914
	ln_avedist	.23746	.27023	.11632	.87872	.38026	.08557	11.68665
	sr_lvnr	.00451	.00597	.04954	.75480	.45096	.34804	2.87323
	ln_jobd	.66435	.13289	.82675	4.99909	.00000	.05482	18.24199
	ln_jwr	-.11270	.06024	-.12137	-1.87093	.06233	.35627	2.80683
	lnjobd_ln	-.19286	.04776	-.37333	-4.03799	.00007	.17540	5.70121

a. Dependent Variable: Ln of proportion of train ridership

APPENDIX 37 CONSTRUCTION LUTI POW

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.726[a]	.527	.505	.98666

a. Predictors: (Constant), ln_jwrc, sqrt_ptiori, p_manpr, p_jobcon, ln_avedist, wage_bl, All_job, p_jobret, sr_lvnr, p_jobman, wage_mp, ln_strio, ln_jobdc, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	326.617	14	23.330	23.965	.000[b]
	Residual	293.020	301	.973		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ln_jwrc, sqrt_ptiori, p_manpr, p_jobcon, ln_avedist, wage_bl, All_job, p_jobret, sr_lvnr, p_jobman, wage_mp, ln_strio, ln_jobdc, p_blu

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	VIF
1	(Constant)	-13.249	3.744		-3.539	.000		
	All_job	3.031E-5	.000	.117	2.215	.027	.565	1.770
	p_manpr	1.404	1.235	.087	1.136	.257	.269	3.713
	p_blu	-1.179	1.347	-.093	-.875	.382	.138	7.246
	p_jobman	1.111	1.261	.064	.881	.379	.297	3.372
	p_jobcon	-2.061	1.442	-.090	-1.430	.154	.398	2.511
	p_jobret	1.880	1.024	.109	1.837	.067	.443	2.257
	wage_mp	.019	.021	.059	.915	.361	.377	2.653
	wage_bl	-.015	.022	-.041	-.681	.496	.428	2.334
	sqrt_ptiori	-6.794E-5	.000	-.028	-.435	.664	.376	2.661
	ln_strio	.884	.387	.145	2.285	.023	.390	2.562
	ln_avedist	-.713	.140	-.349	-5.104	.000	.336	2.979
	sr_lvnr	.011	.006	.123	1.928	.055	.385	2.599
	ln_jobdc	.235	.078	.266	3.011	.003	.202	4.962
	ln_jwrc	-.014	.055	-.015	-.245	.807	.423	2.364

a. Dependent Variable: ln_ptrainw22

APPENDIX 38 CONSTRUCTION LUTI POW WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.741[a]	.549	.526	.96563

a. Predictors: (Constant), lnjobdc_ln, p_blu, All_job, p_jobret, ln_avedist, inc_mp, ln_jwrc, p_jobcon, sr_lvnr, inc_bl, ln_strio, sqrt_ptiori, p_jobman, p_manpr, ln_jobdc
b. Dependent Variable: Ln of proportion of train ridership

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	339.907	15	22.660	24.302	.000[b]
	Residual	279.730	300	.932		
	Total	619.637	315			

a. Dependent Variable: Ln of proportion of train ridership
b. Predictors: (Constant), lnjobdc_ln, p_blu, All_job, p_jobret, ln_avedist, inc_mp, ln_jwrc, p_jobcon, sr_lvnr, inc_bl, ln_strio, sqrt_ptiori, p_jobman, p_manpr, ln_jobdc

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-13.38082	3.66463		-3.65134	.00031		
	All_job	.00002	.00001	.08482	1.62158	.10594	.55006	1.81799
	p_manpr	1.12592	1.21117	.06962	.92961	.35332	.26833	3.72682
	p_blu	-1.43595	1.32025	-.11372	-1.08764	.27763	.13764	7.26524
	p_jobman	1.56381	1.23960	.09029	1.26154	.20809	.29379	3.40381
	p_jobcon	-1.55831	1.41737	-.06789	-1.09944	.27246	.39466	2.53382
	p_jobret	2.22350	1.00598	.12934	2.21029	.02784	.43949	2.27538
	inc_mp	.03132	.02067	.09697	1.51576	.13063	.36768	2.71972
	inc_bl	-.01874	.02183	-.05092	-.85841	.39135	.42768	2.33818
	sqrt_ptiori	-.00011	.00015	-.04429	-.69827	.48555	.37409	2.67314
	ln_strio	.76690	.37975	.12582	2.01948	.04433	.38766	2.57955
	ln_avedist	-.27389	.17945	-.13417	-1.52632	.12798	.19476	5.13463
	sr_lvnr	.00607	.00585	.06669	1.03711	.30052	.36393	2.74777
	ln_jobdc	.65490	.13492	.74058	4.85384	.00000	.06464	15.46992
	ln_jwrc	-.08274	.05708	-.09130	-1.44961	.14821	.37938	2.63591
	lnjobdc_ln	-.19419	.05144	-.36562	-3.77527	.00019	.16044	6.23289

a. Dependent Variable: Ln of proportion of train ridership

APPENDIX 39 MANUFACTURING LUTI POW

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.724[a]	.524	.502	.98698

a. Predictors: (Constant), ln_jwrm2, ln_avedist, p_jobret, p_jobcon, All_job, wage_mp, p_manpr, sr_lvnr, wage_bl, sqrt_ptiori, ln_strio, p_jobman, ln_jobdm, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	321.895	14	22.993	23.603	.000[b]
	Residual	292.237	300	.974		
	Total	614.133	314			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ln_jwrm2, ln_avedist, p_jobret, p_jobcon, All_job, wage_mp, p_manpr, sr_lvnr, wage_bl, sqrt_ptiori, ln_strio, p_jobman, ln_jobdm, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-13.226	3.792		-3.488	.001		
	All_job	3.048E-5	.000	.118	2.239	.026	.571	1.750
	p_manpr	1.140	1.232	.071	.926	.355	.273	3.669

	p_blu	-1.791	1.327	-.141	-1.350	.178	.145	6.912
	p_jobman	-.621	1.563	-.036	-.397	.692	.193	5.172
	p_jobcon	1.020	1.418	.040	.720	.472	.516	1.939
	p_jobret	1.948	1.024	.114	1.901	.058	.444	2.251
	wage_mp	.025	.020	.077	1.222	.222	.397	2.517
	wage_bl	-.022	.022	-.059	-.974	.331	.432	2.313
	sqrt_ptiori	-5.600E-5	.000	-.023	-.355	.723	.369	2.708
	ln_strio	.917	.390	.151	2.350	.019	.385	2.595
	ln_avedist	-.782	.138	-.385	-5.666	.000	.344	2.904
	sr_lvnr	.013	.006	.145	2.285	.023	.394	2.540
	ln_jobdm	.184	.077	.236	2.402	.017	.165	6.076
	ln_jwrm2	.012	.059	.017	.212	.832	.247	4.047

a. Dependent Variable: ln_ptrainw22

APPENDIX 40 MANUFACTURING LUTI POW WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.731[a]	.535	.511	.97757

a. Predictors: (Constant), lnjobdm_ln, p_jobcon, All_job, p_jobret, ln_avedist, inc_bl, p_manpr, ln_jwrm2, sr_lvnr, sqrt_ptiori, inc_mp, ln_strio, p_jobman, p_blu, ln_jobdm
b. Dependent Variable: Ln of proportion of train ridership

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	328.396	15	21.893	22.909	.000[b]
	Residual	285.737	299	.956		
	Total	614.133	314			

a. Dependent Variable: Ln of proportion of train ridership
b. Predictors: (Constant), lnjobdm_ln, p_jobcon, All_job, p_jobret, ln_avedist, inc_bl, p_manpr, ln_jwrm2, sr_lvnr, sqrt_ptiori, inc_mp, ln_strio, p_jobman, p_blu, ln_jobdm

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1.00000	(Constant)	-12.68810	3.76155		-3.37311	.00084		
	All_job	.00003	.00001	.10949	2.09379	.03712	.56906	1.75730
	p_manpr	.90919	1.22333	.05630	.74320	.45794	.27114	3.68818
	p_blu	-2.21012	1.32392	-.17442	-1.66938	.09609	.14254	7.01572
	p_jobman	.38252	1.59559	.02216	.23973	.81070	.18212	5.49084
	p_jobcon	1.68580	1.42737	.06594	1.18105	.23852	.49921	2.00316
	p_jobret	2.21586	1.01977	.12927	2.17289	.03057	.43967	2.27443
	inc_mp	.03599	.02061	.11168	1.74601	.08184	.38035	2.62915
	inc_bl	-.02397	.02205	-.06528	-1.08729	.27779	.43163	2.31681
	sqrt_ptiori	-.00008	.00016	-.03366	-.51761	.60512	.36795	2.71774
	ln_strio	.78989	.38957	.12988	2.02760	.04349	.37926	2.63670
	ln_avedist	-.61345	.15110	-.30182	-4.05999	.00006	.28158	3.55144
	sr_lvnr	.01090	.00576	.12028	1.89186	.05948	.38499	2.59745
	ln_jobdm	.40859	.11477	.52420	3.56017	.00043	.07178	13.93218
	ln_jwrm2	-.05063	.06282	-.06930	-.80605	.42086	.21051	4.75033
	lnjobdm_ln	-.12872	.04935	-.26298	-2.60812	.00956	.15306	6.53354

a. Dependent Variable: Ln of proportion of train ridership

APPENDIX 41 RETAIL LUTI POW

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.731[a]	.534	.512	.97935

a. Predictors: (Constant), ln_jwrr2, p_blu, wage_bl, ln_strio, All_job, p_jobret, p_jobcon, sr_lvnr, wage_mp, ln_avedist, sqrt_ptiori, p_jobman, p_manpr, ln_jobdr2

b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	330.937	14	23.638	24.646	.000[b]
	Residual	288.700	301	.959		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ln_jwrr2, p_blu, wage_bl, ln_strio, All_job, p_jobret, p_jobcon, sr_lvnr, wage_mp, ln_avedist, sqrt_ptiori, p_jobman, p_manpr, ln_jobdr2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.271	3.751		-3.538	.000		
	All_job	2.593E-5	.000	.100	1.898	.059	.559	1.790
	p_manpr	1.256	1.208	.078	1.040	.299	.277	3.605
	p_blu	-1.442	1.306	-.114	-1.104	.270	.145	6.907
	p_jobman	1.170	1.246	.068	.938	.349	.299	3.345
	p_jobcon	-.028	1.247	-.001	-.022	.982	.524	1.907
	p_jobret	-.317	1.226	-.018	-.259	.796	.304	3.286
	wage_mp	.019	.020	.058	.920	.359	.392	2.551
	wage_bl	-.003	.023	-.008	-.136	.892	.410	2.440
	sqrt_ptiori	.000	.000	-.060	-.910	.364	.353	2.829
	ln_strio	.843	.384	.138	2.196	.029	.390	2.562
	ln_avedist	-.700	.136	-.343	-5.133	.000	.347	2.880
	sr_lvnr	.008	.006	.093	1.430	.154	.365	2.740
	ln_jobdr2	.365	.104	.412	3.505	.001	.112	8.919
	ln_jwrr2	-.038	.061	-.042	-.618	.537	.337	2.965

a. Dependent Variable: ln_ptrainw22

APPENDIX 42 RETAIL LUTI POW WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.735[a]	.541	.518	.97419

a. Predictors: (Constant), lnjobdr2_ln, p_jobman, ln_strio, All_job, p_jobcon, inc_mp, p_manpr, ln_jwrr2, sr_lvnr, inc_bl, sqrt_ptiori, p_jobret, ln_avedist, p_blu, ln_jobdr2
b. Dependent Variable: Ln of proportion of train ridership

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	334.925	15	22.328	23.527	.000[b]
	Residual	284.712	300	.949		
	Total	619.637	315			

a. Dependent Variable: Ln of proportion of train ridership
b. Predictors: (Constant), lnjobdr2_ln, p_jobman, ln_strio, All_job, p_jobcon, inc_mp, p_manpr, ln_jwrr2, sr_lvnr, inc_bl, sqrt_ptiori, p_jobret, ln_avedist, p_blu, ln_jobdr2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-13.13570	3.73191		-3.51983	.00050		
	All_job	.00002	.00001	.09254	1.76293	.07893	.55587	1.79897
	p_manpr	1.04204	1.20623	.06443	.86388	.38834	.27535	3.63178
	p_blu	-1.74136	1.30691	-.13791	-1.33243	.18373	.14297	6.99462
	p_jobman	1.53894	1.25279	.08885	1.22841	.22026	.29276	3.41579
	p_jobcon	.43598	1.26088	.01899	.34577	.72976	.50758	1.97013
	p_jobret	-.03648	1.22733	-.00212	-.02972	.97631	.30052	3.32761
	inc_mp	.02256	.02028	.06983	1.11220	.26694	.38859	2.57343

	inc_bl	-.00471	.02251	-.01281	-.20949	.83421	.40935	2.44288
	sqrt_ptiori	-.00016	.00016	-.06542	-.99317	.32143	.35296	2.83319
	ln_strio	.77119	.38339	.12652	2.01150	.04517	.38712	2.58321
	ln_avedist	-.43816	.18617	-.21463	-2.35348	.01924	.18416	5.43004
	sr_lvnr	.00788	.00590	.08656	1.33451	.18305	.36405	2.74687
	ln_jobdr2	.54588	.13607	.61602	4.01177	.00008	.06496	15.39460
	ln_jwrr2	-.06786	.06216	-.07575	-1.09177	.27581	.31813	3.14339
	lnjobdr2_ln	-.11022	.05377	-.17600	-2.04976	.04126	.20775	4.81341

a. Dependent Variable: Ln of proportion of train ridership

ALL SECTOR - SETI POW MODEL

APPENDIX 43 ALL SECTOR SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.727[a]	.529	.506	.98618

a. Predictors: (Constant), Effective employed resident density of all sector (in 1000 units), ln_jwr, p_jobret, p_jobcon, All_job, wage_bl, p_manpr, sr_lvnr, p_jobman, ln_avedist, wage_mp, ln_strio, sqrt_ptiori, ln_jobd, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	327.874	15	21.858	22.475	.000[b]
	Residual	291.763	300	.973		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Effective employed resident density of all sector (in 1000 units), ln_jwr, p_jobret, p_jobcon, All_job, wage_bl, p_manpr, sr_lvnr, p_jobman, ln_avedist, wage_mp, ln_strio, sqrt_ptiori, ln_jobd, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-16.814	4.385		-3.834	.000		
	All_job	3.152E-5	.000	.121	2.296	.022	.560	1.784
	p_manpr	1.318	1.226	.082	1.075	.283	.273	3.660
	p_blu	-1.237	1.351	-.098	-.916	.361	.137	7.294
	p_jobman	1.346	1.259	.078	1.069	.286	.297	3.365
	p_jobcon	-.524	1.290	-.023	-.406	.685	.497	2.012
	p_jobret	1.922	1.023	.112	1.878	.061	.443	2.256
	wage_mp	.019	.021	.059	.913	.362	.373	2.683
	wage_bl	-.010	.023	-.026	-.423	.673	.420	2.383
	sqrt_ptiori	.000	.000	-.085	-1.160	.247	.292	3.424
	ln_strio	1.116	.430	.183	2.597	.010	.316	3.169
	ln_avedist	-.674	.144	-.330	-4.682	.000	.315	3.172
	sr_lvnr	.010	.006	.110	1.687	.093	.370	2.701
	ln_jobd	.209	.083	.260	2.504	.013	.146	6.852
	ln_jwr	6.134E-5	.061	.000	.001	.999	.366	2.735
	eder1000	.066	.045	.130	1.469	.143	.201	4.981

a. Dependent Variable: ln_ptrainw22

APPENDIX 44 ALL SECTOR SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.742[a]	.551	.526	.96511

a. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, wage_mp, p_manpr, ln_jwr, sr_lvnr, wage_bl, sqrt_ptiori, Effective employed resident density of all sector (in 1000 units), ln_jobd, p_blu, ln_avedist
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	341.136	16	21.321	22.890	.000[b]
	Residual	278.501	299	.931		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), edertt_ln, p_jobman, All_job, p_jobret, p_jobcon, ln_strio, wage_mp, p_manpr, ln_jwr, sr_lvnr, wage_bl, sqrt_ptiori, Effective employed resident density of all sector (in 1000 units), ln_jobd, p_blu, ln_avedist

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-17.592	4.296		-4.094	.000		
	All_job	2.740E-5	.000	.106	2.032	.043	.557	1.796
	p_manpr	1.365	1.200	.084	1.138	.256	.273	3.660
	p_blu	-.954	1.324	-.076	-.721	.472	.137	7.317
	p_jobman	1.485	1.232	.086	1.205	.229	.297	3.368
	p_jobcon	-.224	1.265	-.010	-.177	.859	.495	2.020
	p_jobret	2.395	1.009	.139	2.374	.018	.436	2.291
	wage_mp	.024	.021	.073	1.146	.253	.372	2.692
	wage_bl	-.009	.022	-.024	-.400	.690	.420	2.383
	sqrt_ptiori	.000	.000	-.108	-1.500	.135	.290	3.449
	ln_strio	.919	.424	.151	2.169	.031	.311	3.218
	ln_avedist	.623	.372	.305	1.677	.095	.045	22.055
	sr_lvnr	.007	.006	.077	1.197	.232	.363	2.752
	ln_jobd	.299	.085	.372	3.518	.001	.134	7.442
	ln_jwr	-.057	.061	-.061	-.926	.355	.344	2.911
	eder1000	.226	.061	.446	3.702	.000	.104	9.643
	edertt_ln	-.115	.030	-.479	-3.773	.000	.093	10.726

a. Dependent Variable: ln_ptrainw22

APPENDIX 45 ALL SECTOR SETI H2 EFFECTIVE JOB DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.732[a]	.535	.512	.97974

a. Predictors: (Constant), Employment effective density of all sector (in 1000 units), p_jobret, p_jobman, p_jobcon, ln_avedist, wage_mp, ln_jwr, p_manpr, wage_bl, sr_lvnr, sqrt_ptiori, ln_strio, ln_jobd, p_blu, All_job
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	331.669	15	22.111	23.035	.000[b]
	Residual	287.968	300	.960		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Employment effective density of all sector (in 1000 units), p_jobret, p_jobman, p_jobcon, ln_avedist, wage_mp, ln_jwr, p_manpr, wage_bl, sr_lvnr, sqrt_ptiori, ln_strio, ln_jobd, p_blu, All_job

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-19.653	4.504		-4.364	.000		

	All_job	-3.118E-5	.000	-.120	-1.088	.277	.127	7.870
	p_manpr	1.000	1.222	.062	.819	.414	.271	3.684
	p_blu	-1.090	1.342	-.086	-.812	.417	.137	7.293
	p_jobman	1.718	1.263	.099	1.361	.175	.291	3.431
	p_jobcon	-.398	1.272	-.017	-.313	.755	.505	1.981
	p_jobret	2.041	1.018	.119	2.006	.046	.442	2.262
	wage_mp	.021	.021	.066	1.026	.306	.372	2.685
	wage_bl	-.013	.022	-.035	-.577	.564	.425	2.353
	sqrt_ptiori	.000	.000	-.046	-.714	.476	.369	2.711
	ln_strio	1.386	.443	.227	3.129	.002	.293	3.410
	ln_avedist	-.651	.144	-.319	-4.532	.000	.313	3.192
	sr_lvnr	.009	.006	.096	1.483	.139	.366	2.730
	ln_jobd	.218	.081	.271	2.687	.008	.152	6.588
	ln_jwr	-.060	.059	-.065	-1.020	.308	.386	2.592
	ejd1000	.103	.042	.347	2.478	.014	.079	12.679

a. Dependent Variable: ln_ptrainw22

APPENDIX 46 ALL SECTOR SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.739[a]	.546	.521	.97025

a. Predictors: (Constant), ejdtt_ln, wage_bl, ln_jobd, p_manpr, p_jobcon, ln_jwr, p_jobret, ln_strio, All_job, p_jobman, wage_mp, sr_lvnr, sqrt_ptiori, p_blu, ln_avedist, Employment effective density of all sector (in 1000 units)
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	338.160	16	21.135	22.451	.000[b]
	Residual	281.477	299	.941		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ejdtt_ln, wage_bl, ln_jobd, p_manpr, p_jobcon, ln_jwr, p_jobret, ln_strio, All_job, p_jobman, wage_mp, sr_lvnr, sqrt_ptiori, p_blu, ln_avedist, Employment effective density of all sector (in 1000 units)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-15.071	4.789		-3.147	.002		
	All_job	-1.244E-5	.000	-.048	-.425	.671	.120	8.368
	p_manpr	1.015	1.210	.063	.839	.402	.271	3.684
	p_blu	-1.012	1.329	-.080	-.761	.447	.137	7.296
	p_jobman	1.956	1.254	.113	1.560	.120	.290	3.449
	p_jobcon	.140	1.276	.006	.110	.913	.492	2.034
	p_jobret	2.289	1.012	.133	2.261	.024	.438	2.282
	wage_mp	.023	.021	.072	1.126	.261	.372	2.688
	wage_bl	-.022	.022	-.059	-.982	.327	.415	2.411
	sqrt_ptiori	-9.664E-5	.000	-.040	-.623	.534	.368	2.714
	ln_strio	.844	.485	.139	1.742	.083	.240	4.164
	ln_avedist	-.016	.280	-.008	-.057	.954	.081	12.413
	sr_lvnr	.008	.006	.084	1.296	.196	.364	2.746
	ln_jobd	.262	.082	.326	3.190	.002	.146	6.871
	ln_jwr	-.062	.058	-.067	-1.062	.289	.386	2.592
	ejd1000	.132	.043	.443	3.087	.002	.074	13.552
	ejdtt_ln	-.070	.027	-.272	-2.626	.009	.142	7.041

a. Dependent Variable: ln_ptrainw22

CONSTRUCTION SECTOR - SETI POW MODEL

APPENDIX 47 CONSTRUCTION SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.726[a]	.527	.504	.98803

a. Predictors: (Constant), Effective employed resident density of construction sector (in 100 units), p_manpr, wage_bl, All_job, p_jobcon, ln_strio, p_jobman, p_jobret, wage_mp, sr_lvnr, ln_jwrc, ln_avedist, sqrt_ptiori, ln_jobdc, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	326.777	15	21.785	22.316	.000[b]
	Residual	292.860	300	.976		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Effective employed resident density of construction sector (in 100 units), p_manpr, wage_bl, All_job, p_jobcon, ln_strio, p_jobman, p_jobret, wage_mp, sr_lvnr, ln_jwrc, ln_avedist, sqrt_ptiori, ln_jobdc, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.318	3.753		-3.548	.000		
	All_job	3.051E-5	.000	.118	2.225	.027	.564	1.773
	p_manpr	1.424	1.238	.088	1.150	.251	.269	3.719
	p_blu	-1.220	1.353	-.097	-.902	.368	.137	7.287
	p_jobman	1.134	1.264	.065	.897	.370	.296	3.379
	p_jobcon	-2.109	1.449	-.092	-1.456	.147	.396	2.528
	p_jobret	1.877	1.025	.109	1.831	.068	.443	2.257
	wage_mp	.019	.021	.059	.915	.361	.377	2.653
	wage_bl	-.015	.022	-.040	-.656	.512	.427	2.341
	sqrt_ptiori	-9.401E-5	.000	-.039	-.556	.579	.321	3.114
	ln_strio	.884	.387	.145	2.283	.023	.390	2.563
	ln_avedist	-.712	.140	-.349	-5.094	.000	.336	2.979
	sr_lvnr	.011	.006	.124	1.932	.054	.385	2.600
	ln_jobdc	.219	.088	.247	2.489	.013	.159	6.274
	ln_jwrc	.000	.064	.000	-.006	.995	.315	3.175
	Effective employed resident density of construction sector (in 100 units)	.021	.053	.029	.404	.686	.299	3.350

a. Dependent Variable: ln_ptrainw22

APPENDIX 48 CONSTRUCTION SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.733[a]	.538	.513	.97848

a. Predictors: (Constant), ederttc_ln, sqrt_ptiori, p_jobman, p_jobret, p_jobcon, wage_mp, All_job, ln_strio, p_manpr, sr_lvnr, wage_bl, ln_jwr, Effective employed resident density of construction sector (in 100 units), ln_jobd, p_blu, ln_avedist
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	333.369	16	20.836	21.762	.000[b]
	Residual	286.268	299	.957		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ederttc_ln, sqrt_ptiori, p_jobman, p_jobret, p_jobcon, wage_mp, All_job, ln_strio, p_manpr, sr_lvnr, wage_bl, ln_jwr, Effective employed resident density of construction sector (in 100 units), ln_jobd, p_blu, ln_avedist

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-15.111	3.833		-3.942	.000		
	All_job	2.764E-5	.000	.107	2.019	.044	.555	1.803
	p_manpr	1.406	1.220	.087	1.153	.250	.272	3.682
	p_blu	-1.019	1.344	-.081	-.759	.449	.136	7.329
	p_jobman	1.304	1.248	.075	1.045	.297	.298	3.360
	p_jobcon	-.268	1.291	-.012	-.208	.836	.488	2.049
	p_jobret	2.162	1.020	.126	2.121	.035	.439	2.277
	wage_mp	.021	.021	.065	1.005	.316	.372	2.685
	wage_bl	-.008	.022	-.023	-.377	.706	.422	2.370
	sqrt_ptiori	.000	.000	-.070	-1.001	.318	.315	3.170
	ln_strio	.793	.385	.130	2.060	.040	.387	2.582
	ln_avedist	.094	.320	.046	.292	.770	.063	15.917
	sr_lvnr	.010	.006	.107	1.650	.100	.371	2.695
	ln_jobd	.299	.093	.373	3.219	.001	.115	8.675
	ln_jwr	-.053	.067	-.057	-.786	.433	.295	3.388
	Effective employed resident density of construction sector (in 100 units)	.229	.087	.316	2.617	.009	.106	9.425
	ederttc_ln	-.137	.050	-.352	-2.742	.006	.094	10.678

a. Dependent Variable: ln_ptrainw22

APPENDIX 49 CONSTRUCTION SETI H2 EFFECTIVE JOB DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.726[a]	.527	.504	.98817

a. Predictors: (Constant), Employment effective density of construction sector (in 100 units), p_jobcon, p_jobret, p_jobman, ln_avedist, wage_mp, sqrt_ptiori, p_manpr, wage_bl, sr_lvnr, ln_jwrc, ln_strio, All_job, ln_jobdc, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	326.692	15	21.779	22.304	.000[b]
	Residual	292.945	300	.976		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Employment effective density of construction sector (in 100 units), p_jobcon, p_jobret, p_jobman, ln_avedist, wage_mp, sqrt_ptiori, p_manpr, wage_bl, sr_lvnr, ln_jwrc, ln_strio, All_job, ln_jobdc, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.504	3.861		-3.497	.001		
	All_job	2.727E-5	.000	.105	1.552	.122	.344	2.909
	p_manpr	1.403	1.237	.087	1.134	.258	.269	3.713
	p_blu	-1.172	1.349	-.093	-.868	.386	.138	7.248
	p_jobman	1.085	1.266	.063	.857	.392	.295	3.390
	p_jobcon	-2.113	1.456	-.092	-1.451	.148	.392	2.553
	p_jobret	1.891	1.026	.110	1.843	.066	.443	2.260
	wage_mp	.019	.021	.059	.916	.361	.377	2.653

	wage_bl	-.014	.023	-.039	-.638	.524	.421	2.375
	sqrt_ptiori	-6.707E-5	.000	-.028	-.429	.668	.376	2.662
	ln_strio	.906	.396	.149	2.289	.023	.374	2.677
	ln_avedist	-.715	.140	-.350	-5.104	.000	.335	2.985
	sr_lvnr	.011	.006	.124	1.937	.054	.384	2.605
	ln_jobdc	.229	.082	.259	2.805	.005	.185	5.394
	ln_jwrc	-.017	.057	-.019	-.303	.762	.400	2.502
	Employment effective density of construction sector (in 100 units)	.011	.040	.022	.277	.782	.259	3.858

a. Dependent Variable: ln_ptrainw22

APPENDIX 50 CONSTRUCTION SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.735[a]	.540	.516	.97587

a. Predictors: (Constant), ejdttc_ln, ln_strio, p_jobret, p_jobcon, wage_mp, p_jobman, All_job, p_manpr, sr_lvnr, ln_jwrc, wage_bl, sqrt_ptiori, ln_avedist, ln_jobdc, p_blu, Employment effective density of construction sector (in 100 units)
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	334.894	16	20.931	21.979	.000[b]
	Residual	284.743	299	.952		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ejdttc_ln, ln_strio, p_jobret, p_jobcon, wage_mp, p_jobman, All_job, p_manpr, sr_lvnr, ln_jwrc, wage_bl, sqrt_ptiori, ln_avedist, ln_jobdc, p_blu, Employment effective density of construction sector (in 100 units)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-12.392	3.832		-3.234	.001		
	All_job	1.916E-5	.000	.074	1.091	.276	.335	2.982
	p_manpr	1.286	1.222	.079	1.052	.294	.269	3.717
	p_blu	-1.266	1.333	-.100	-.949	.343	.138	7.253
	p_jobman	1.940	1.284	.112	1.511	.132	.280	3.574
	p_jobcon	-1.769	1.443	-.077	-1.226	.221	.389	2.570
	p_jobret	2.048	1.015	.119	2.019	.044	.441	2.266
	wage_mp	.020	.021	.062	.966	.335	.377	2.653
	wage_bl	-.023	.022	-.064	-1.045	.297	.413	2.421
	sqrt_ptiori	-7.664E-5	.000	-.032	-.496	.620	.375	2.663
	ln_strio	.731	.396	.120	1.847	.066	.365	2.740
	ln_avedist	-.228	.216	-.112	-1.055	.292	.137	7.278
	sr_lvnr	.010	.006	.110	1.736	.084	.382	2.620
	ln_jobdc	.286	.083	.324	3.455	.001	.175	5.715
	ln_jwrc	-.014	.056	-.016	-.257	.797	.400	2.502
	Employment effective density of construction sector (in 100 units)	.161	.065	.314	2.494	.013	.097	10.317
	ejdttc_ln	-.139	.047	-.341	-2.935	.004	.114	8.769

a. Dependent Variable: ln_ptrainw22

MANUFACTURING SECTOR - SETI POW MODEL

APPENDIX 51 MANUFACTURING SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.725[a]	.526	.503	.98641

a. Predictors: (Constant), Effective employed resident density of manufacturing sector (in 100 units), p_manpr, wage_mp, All_job, p_jobcon, p_jobman, p_jobret, ln_avedist, sr_lvnr, wage_bl, ln_strio, sqrt_ptiori, ln_jwrm2, ln_jobdm, p_blu
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	323.206	15	21.547	22.145	.000[b]
	Residual	290.927	299	.973		
	Total	614.133	314			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Effective employed resident density of manufacturing sector (in 100 units), p_manpr, wage_mp, All_job, p_jobcon, p_jobman, p_jobret, ln_avedist, sr_lvnr, wage_bl, ln_strio, sqrt_ptiori, ln_jwrm2, ln_jobdm, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-15.637	4.322		-3.618	.000		
	All_job	2.949E-5	.000	.114	2.163	.031	.569	1.757
	p_manpr	1.274	1.237	.079	1.031	.304	.270	3.701
	p_blu	-1.851	1.327	-.146	-1.395	.164	.144	6.923
	p_jobman	-.912	1.583	-.053	-.576	.565	.188	5.305
	p_jobcon	.878	1.422	.034	.617	.537	.512	1.954
	p_jobret	1.930	1.024	.113	1.885	.060	.444	2.252
	wage_mp	.025	.020	.078	1.232	.219	.397	2.517
	wage_bl	-.019	.022	-.051	-.842	.401	.427	2.341
	sqrt_ptiori	.000	.000	-.061	-.835	.404	.296	3.378
	ln_strio	1.114	.425	.183	2.619	.009	.324	3.089
	ln_avedist	-.761	.139	-.374	-5.469	.000	.338	2.955
	sr_lvnr	.014	.006	.149	2.345	.020	.393	2.547
	ln_jobdm	.162	.079	.208	2.065	.040	.156	6.428
	ln_jwrm2	.054	.069	.074	.790	.430	.179	5.591
	Effective employed resident density of manufacturing sector (in 100 units)	.067	.058	.095	1.160	.247	.237	4.216

a. Dependent Variable: ln_ptrainw22

APPENDIX 52 MANUFACTURING SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.735[a]	.540	.515	.97394

a. Predictors: (Constant), ederttm_ln, Effective employed resident density of manufacturing sector (in 100 units), All_job, p_jobman, p_jobret, p_jobcon, wage_mp, p_manpr, sr_lvnr, wage_bl, ln_strio, sqrt_ptiori, ln_jwrm2, ln_jobdm, p_blu, ln_avedist
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	331.461	16	20.716	21.840	.000[b]
	Residual	282.672	298	.949		
	Total	614.133	314			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ederttm_ln, Effective employed resident density of manufacturing sector (in 100 units), All_job, p_jobman, p_jobret, p_jobcon, wage_mp, p_manpr, sr_lvnr, wage_bl, ln_strio, sqrt_ptiori, ln_jwrm2, ln_jobdm, p_blu, ln_avedist

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF

1	(Constant)	-17.728	4.326		-4.098	.000		
	All_job	2.759E-5	.000	.107	2.047	.041	.568	1.761
	p_manpr	1.547	1.224	.096	1.264	.207	.269	3.722
	p_blu	-1.688	1.311	-.133	-1.287	.199	.144	6.935
	p_jobman	-.840	1.563	-.049	-.538	.591	.188	5.307
	p_jobcon	.789	1.405	.031	.562	.575	.512	1.955
	p_jobret	2.338	1.020	.136	2.291	.023	.436	2.294
	wage_mp	.031	.020	.097	1.547	.123	.393	2.545
	wage_bl	-.015	.022	-.041	-.687	.492	.426	2.348
	sqrt_ptiori	.000	.000	-.095	-1.296	.196	.289	3.463
	ln_strio	1.068	.420	.176	2.541	.012	.323	3.093
	ln_avedist	.385	.412	.189	.935	.351	.038	26.588
	sr_lvnr	.013	.006	.144	2.289	.023	.392	2.549
	ln_jobdm	.215	.080	.276	2.699	.007	.148	6.767
	ln_jwrm2	.010	.070	.014	.145	.884	.171	5.864
	Effective employed resident density of manufacturing sector (in 100 units)	.279	.092	.395	3.042	.003	.092	10.914
	ederttm_ln	-.128	.044	-.456	-2.950	.003	.065	15.453

a. Dependent Variable: ln_ptrainw22

APPENDIX 53 MANUFACTURING SETI H2 EFFECTIVE JOB DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.877[a]	.769	.757	.90349

a. Predictors: (Constant), Employment effective density of manufacturing sector (in 100 units), p_jobcon, p_jobret, sr_lvnr, All_job, wage_bl, ln_avedist, p_manpr, wage_mp, sqrt_ptiori, p_jobman, ln_strio, ln_jwrm2, ln_jobdm, p_blu
b. Dependent Variable: ln_ptrainw2

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	812.379	15	54.159	66.346	.000[b]
	Residual	244.074	299	.816		
	Total	1056.453	314			

a. Dependent Variable: ln_ptrainw2
b. Predictors: (Constant), Employment effective density of manufacturing sector (in 100 units), p_jobcon, p_jobret, sr_lvnr, All_job, wage_bl, ln_avedist, p_manpr, wage_mp, sqrt_ptiori, p_jobman, ln_strio, ln_jwrm2, ln_jobdm, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-11.411	3.697		-3.087	.002		
	All_job	6.847E-5	.000	.202	5.184	.000	.509	1.967
	p_manpr	1.007	1.131	.048	.891	.374	.271	3.688
	p_blu	.001	1.218	.000	.001	1.000	.144	6.951
	p_jobman	-5.515	1.447	-.244	-3.812	.000	.189	5.283
	p_jobcon	-1.136	1.317	-.034	-.863	.389	.501	1.995
	p_jobret	1.565	.940	.070	1.665	.097	.442	2.261
	wage_mp	.024	.019	.056	1.261	.208	.397	2.517
	wage_bl	-.017	.021	-.036	-.838	.403	.422	2.370
	sqrt_ptiori	.000	.000	.154	3.374	.001	.369	2.708
	ln_strio	1.317	.372	.165	3.540	.000	.355	2.814
	ln_avedist	-.586	.126	-.220	-4.636	.000	.343	2.913
	sr_lvnr	.020	.005	.168	3.764	.000	.387	2.585
	ln_jobdm	.255	.071	.250	3.577	.000	.158	6.318
	ln_jwrm2	.255	.056	.266	4.555	.000	.226	4.420

	Employment effective density of manufacturing sector (in 100 units)	.063	.024	.131	2.596	.010	.305	3.283

a. Dependent Variable: ln_ptrainw2

APPENDIX 54 MANUFACTURING SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION EFFECT – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.729[a]	.532	.507	.98244

a. Predictors: (Constant), ejdttm_ln, p_jobret, ln_strio, All_job, p_jobcon, wage_mp, p_manpr, wage_bl, sr_lvnr, ln_jwrm2, sqrt_ptiori, p_jobman, ln_avedist, ln_jobdm, p_blu, Employment effective density of manufacturing sector (in 100 units)
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	326.507	16	20.407	21.143	.000[b]
	Residual	287.626	298	.965		
	Total	614.133	314			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ejdttm_ln, p_jobret, ln_strio, All_job, p_jobcon, wage_mp, p_manpr, wage_bl, sr_lvnr, ln_jwrm2, sqrt_ptiori, p_jobman, ln_avedist, ln_jobdm, p_blu, Employment effective density of manufacturing sector (in 100 units)

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	-13.006	4.022		-3.234	.001		
	All_job	2.580E-5	.000	.100	1.778	.076	.498	2.007
	p_manpr	1.119	1.229	.069	.910	.364	.271	3.688
	p_blu	-2.022	1.329	-.160	-1.522	.129	.143	6.996
	p_jobman	.108	1.609	.006	.067	.947	.181	5.525
	p_jobcon	1.514	1.450	.059	1.044	.297	.489	2.046
	p_jobret	2.188	1.028	.128	2.130	.034	.437	2.286
	wage_mp	.029	.020	.090	1.431	.153	.394	2.541
	wage_bl	-.025	.022	-.067	-1.101	.272	.420	2.380
	sqrt_ptiori	-5.419E-5	.000	-.023	-.345	.730	.369	2.708
	ln_strio	.812	.407	.134	1.993	.047	.350	2.856
	ln_avedist	-.345	.242	-.170	-1.423	.156	.110	9.052
	sr_lvnr	.013	.006	.145	2.272	.024	.387	2.585
	ln_jobdm	.175	.078	.224	2.248	.025	.158	6.334
	ln_jwrm2	.007	.061	.009	.111	.912	.226	4.426
	Employment effective density of manufacturing sector (in 100 units)	.089	.048	.240	1.841	.067	.093	10.772
	ejdttm_ln	-.063	.029	-.290	-2.186	.030	.089	11.174

a. Dependent Variable: ln_ptrainw22

RETAIL SECTOR - SETI POW

APPENDIX 55 RETAIL SETI H1 EFFECTIVE EMPLOYED RESIDENT DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.731[a]	.534	.511	.98061

a. Predictors: (Constant), Effective employed resident density of retail sector (in 100 units), ln_jwrr2, p_manpr, wage_bl, All_job, p_jobcon, ln_avedist, p_jobman, sr_lvnr, wage_mp, p_jobret, ln_strio, sqrt_ptiori, p_blu, ln_jobdr2
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	331.156	15	22.077	22.959	.000[b]
	Residual	288.481	300	.962		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Effective employed resident density of retail sector (in 100 units), ln_jwrr2, p_manpr, wage_bl, All_job, p_jobcon, ln_avedist, p_jobman, sr_lvnr, wage_mp, p_jobret, ln_strio, sqrt_ptiori, p_blu, ln_jobdr2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.553	3.802		-3.565	.000		
	All_job	2.650E-5	.000	.102	1.930	.054	.554	1.804
	p_manpr	1.300	1.213	.080	1.072	.285	.276	3.626
	p_blu	-1.455	1.308	-.115	-1.113	.267	.145	6.910
	p_jobman	1.162	1.248	.067	.931	.352	.299	3.346
	p_jobcon	-.118	1.263	-.005	-.094	.925	.513	1.951
	p_jobret	-.340	1.229	-.020	-.276	.782	.304	3.291
	wage_mp	.018	.020	.057	.903	.367	.392	2.553
	wage_bl	-.002	.023	-.007	-.107	.915	.408	2.449
	sqrt_ptiori	.000	.000	-.077	-1.026	.306	.275	3.641
	ln_strio	.868	.388	.142	2.237	.026	.383	2.608
	ln_avedist	-.697	.137	-.342	-5.106	.000	.347	2.884
	sr_lvnr	.008	.006	.093	1.430	.154	.365	2.740
	ln_jobdr2	.346	.112	.390	3.093	.002	.097	10.257
	ln_jwrr2	-.020	.071	-.023	-.290	.772	.250	3.992
	Effective employed resident density of retail sector (in 100 units)	.028	.059	.038	.476	.634	.248	4.026

a. Dependent Variable: ln_ptrainw22

APPENDIX 56 RETAIL SETI H1B EFFECTIVE EMPLOYED RESIDENT DENSITY WITH THE INTERACTION TERM

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.737[a]	.543	.518	.97362

a. Predictors: (Constant), ederttr_ln, p_blu, All_job, wage_mp, p_jobret, ln_strio, p_jobcon, ln_jwrr2, sr_lvnr, wage_bl, sqrt_ptiori, p_jobman, p_manpr, ln_avedist, ln_jobdr2, Effective employed resident density of retail sector (in 100 units)
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	336.202	16	21.013	22.167	.000[b]
	Residual	283.435	299	.948		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ederttr_ln, p_blu, All_job, wage_mp, p_jobret, ln_strio, p_jobcon, ln_jwrr2, sr_lvnr, wage_bl, sqrt_ptiori, p_jobman, p_manpr, ln_avedist, ln_jobdr2, Effective employed resident density of retail sector (in 100 units)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-13.967	3.779		-3.696	.000		
	All_job	2.575E-5	.000	.099	1.888	.060	.554	1.805
	p_manpr	1.187	1.206	.073	.985	.325	.275	3.632

	p_blu	-1.585	1.299	-.126	-1.220	.223	.144	6.923
	p_jobman	1.359	1.242	.078	1.094	.275	.297	3.361
	p_jobcon	.082	1.257	.004	.065	.948	.510	1.960
	p_jobret	-.101	1.224	-.006	-.083	.934	.302	3.315
	wage_mp	.023	.020	.071	1.124	.262	.388	2.576
	wage_bl	.001	.023	.002	.039	.969	.407	2.459
	sqrt_ptiori	.000	.000	-.096	-1.284	.200	.271	3.686
	ln_strio	.787	.387	.129	2.036	.043	.380	2.630
	ln_avedist	-.173	.265	-.085	-.653	.514	.091	10.991
	sr_lvnr	.008	.006	.088	1.363	.174	.365	2.743
	ln_jobdr2	.388	.113	.438	3.452	.001	.095	10.542
	ln_jwrr2	-.055	.072	-.062	-.773	.440	.239	4.178
	Effective employed resident density of retail sector (in 100 units)	.203	.096	.272	2.118	.035	.093	10.748
	ederttr_ln	-.116	.050	-.257	-2.307	.022	.123	8.125

a. Dependent Variable: ln_ptrainw22

APPENDIX 57 RETAIL SETI H2 EFFECTIVE JOB DENSITY – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.731[a]	.534	.511	.98097

a. Predictors: (Constant), Employment effective density of retail sector (in 100 units), wage_bl, p_manpr, p_jobcon, ln_avedist, p_jobman, sr_lvnr, ln_jwrr2, All_job, wage_mp, sqrt_ptiori, ln_strio, p_jobret, p_blu, ln_jobdr2
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	330.945	15	22.063	22.927	.000[b]
	Residual	288.692	300	.962		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), Employment effective density of retail sector (in 100 units), wage_bl, p_manpr, p_jobcon, ln_avedist, p_jobman, sr_lvnr, ln_jwrr2, All_job, wage_mp, sqrt_ptiori, ln_strio, p_jobret, p_blu, ln_jobdr2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-13.326	3.808		-3.499	.001		
	All_job	2.514E-5	.000	.097	1.535	.126	.390	2.565
	p_manpr	1.253	1.211	.077	1.035	.302	.277	3.609
	p_blu	-1.446	1.309	-.115	-1.105	.270	.145	6.917
	p_jobman	1.177	1.251	.068	.941	.348	.298	3.359
	p_jobcon	-.035	1.252	-.002	-.028	.978	.522	1.914
	p_jobret	-.338	1.252	-.020	-.270	.787	.293	3.413
	wage_mp	.019	.020	.058	.922	.358	.391	2.559
	wage_bl	-.003	.023	-.009	-.140	.889	.409	2.445
	sqrt_ptiori	.000	.000	-.061	-.912	.363	.347	2.880
	ln_strio	.848	.389	.139	2.178	.030	.380	2.628
	ln_avedist	-.699	.137	-.343	-5.122	.000	.347	2.881
	sr_lvnr	.009	.006	.094	1.431	.154	.363	2.754
	ln_jobdr2	.363	.107	.409	3.386	.001	.106	9.414
	ln_jwrr2	-.038	.062	-.043	-.623	.534	.329	3.036
	Employment effective density of retail sector (in 100 units)	.003	.034	.007	.087	.930	.241	4.151

a. Dependent Variable: ln_ptrainw22

APPENDIX 58 RETAIL SETI H2B EFFECTIVE JOB DENSITY WITH THE INTERACTION TERM – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.736[a]	.541	.517	.97483

a. Predictors: (Constant), ejdttr_ln, p_blu, wage_mp, ln_strio, p_jobret, p_jobcon, All_job, ln_jwrr2, sr_lvnr, wage_bl, sqrt_ptiori, p_jobman, p_manpr, ln_avedist, Employment effective density of retail sector (in 100 units), ln_jobdr2
b. Dependent Variable: ln_ptrainw22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	335.498	16	20.969	22.065	.000[b]
	Residual	284.139	299	.950		
	Total	619.637	315			

a. Dependent Variable: ln_ptrainw22
b. Predictors: (Constant), ejdttr_ln, p_blu, wage_mp, ln_strio, p_jobret, p_jobcon, All_job, ln_jwrr2, sr_lvnr, wage_bl, sqrt_ptiori, p_jobman, p_manpr, ln_avedist, Employment effective density of retail sector (in 100 units), ln_jobdr2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-11.655	3.860		-3.019	.003		
	All_job	2.676E-5	.000	.103	1.643	.101	.389	2.570
	p_manpr	1.142	1.204	.071	.949	.344	.277	3.615
	p_blu	-1.437	1.300	-.114	-1.105	.270	.145	6.917
	p_jobman	1.465	1.250	.085	1.172	.242	.294	3.397
	p_jobcon	.290	1.253	.013	.231	.817	.515	1.942
	p_jobret	.051	1.256	.003	.041	.968	.287	3.482
	wage_mp	.019	.020	.060	.954	.341	.391	2.560
	wage_bl	-.008	.023	-.021	-.349	.727	.405	2.468
	sqrt_ptiori	.000	.000	-.059	-.895	.371	.347	2.881
	ln_strio	.627	.400	.103	1.568	.118	.356	2.808
	ln_avedist	-.380	.199	-.186	-1.910	.057	.161	6.205
	sr_lvnr	.008	.006	.089	1.364	.174	.363	2.757
	ln_jobdr2	.405	.108	.458	3.746	.000	.103	9.729
	ln_jwrr2	-.049	.061	-.054	-.793	.429	.327	3.054
	Employment effective density of retail sector (in 100 units)	.072	.046	.171	1.562	.119	.128	7.807
	ejdttr_ln	-.082	.038	-.217	-2.189	.029	.157	6.384

a. Dependent Variable: ln_ptrainw22

APPENDIX 59 EXAMPLE OF MODELS WITH NON-TRANSFORMED VARIABLES PLACE OF RESIDENCE (POR MODEL) FOR THE ALL SECTOR SETI H1

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.662a	.438	.410	.66738

a. Predictors: (Constant), ejdtt1000, jwr, p_woret, p_worker_man, inc_blu_r, wod, ave_dist, p_wocon, car_own, lvr, ptiori, inc_pr_r, strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.252	15	6.950	15.605	.000b
	Residual	133.618	300	.445		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo

b. Predictors: (Constant), ejdtt1000, jwr, p_woret, p_worker_man, inc_blu_r, wod, ave_dist, p_wocon, car_own, lvr, ptiori, inc_pr_r, strio, pmanpr_r, pblu_r

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-5.33350	1.23756		-4.30967	.00002		
	pmanpr_r	5.76410	1.13377	.82979	5.08399	.00000	.07029	14.22749
	pblu_r	4.07009	1.55843	.54028	2.61165	.00946	.04375	22.85633
	p_worker_man	-1.19382	1.17951	-.07016	-1.01213	.31229	.38963	2.56655
	p_wocon	1.98698	1.85300	.08003	1.07231	.28444	.33617	2.97468
	p_woret	4.62407	1.87717	.14425	2.46332	.01433	.54600	1.83151
	inc_pr_r	-.01704	.01334	-.14679	-1.27773	.20233	.14187	7.04894
	inc_blu_r	-.00492	.01411	-.02265	-.34855	.72767	.44349	2.25486
	car_own	-.67484	.13835	-.32194	-4.87784	.00000	.42983	2.32649
	ptiori	.00000	.00000	-.11353	-1.71282	.08778	.42622	2.34621
	strio	.00012	.00002	.42471	5.24368	.00000	.28542	3.50357
	ave_dist	-.09078	.01220	-.51799	-7.44345	.00000	.38664	2.58638
	lvr	-.00029	.00005	-.36284	-5.36413	.00000	.40924	2.44355
	wod	.00007	.00011	.03482	.57458	.56600	.50973	1.96183
	jwr	.00217	.00054	.42821	4.01198	.00008	.16436	6.08407
	ejdtt1000	.01289	.01471	.06985	.87672	.38134	.29494	3.39047

a. Dependent Variable: ln_train_wo

REMOVING THE MULTICOLLINEARITY ISSUES FOR MODEL ALL SECTOR SETI H1

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.652a	.425	.398	.67433

a. Predictors: (Constant), edertt1000, p_woret, jwr, p_worker_man, inc_blu_r, car_own, ptiori, lvr, p_wocon, ave_dist, inc_pr_r, strio, wod, pmanpr_r
b. Dependent Variable: ln_train_wo

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	101.000	14	7.214	15.865	.000b
	Residual	136.871	301	.455		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), edertt1000, p_woret, jwr, p_worker_man, inc_blu_r, car_own, ptiori, lvr, p_wocon, ave_dist, inc_pr_r, strio, wod, pmanpr_r

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-2.60620	.62778		-4.15148	.00004		
	pmanpr_r	3.49789	.67518	.50355	5.18070	.00000	.20234	4.94208
	p_worker_man	.82790	.92585	.04866	.89420	.37193	.64562	1.54889
	p_wocon	2.87426	1.84557	.11576	1.55738	.12043	.34598	2.89034
	p_woret	2.84208	1.77864	.08866	1.59789	.11112	.62090	1.61057
	inc_pr_r	-.04243	.00949	-.36558	-4.47204	.00001	.28606	3.49578
	inc_blu_r	-.01306	.01396	-.06012	-.93506	.35051	.46236	2.16281
	car_own	-.71210	.13648	-.33972	-5.21757	.00000	.45092	2.21769
	ptiori	.00000	.00000	-.10751	-1.80438	.07217	.53843	1.85726
	strio	.00014	.00002	.47968	5.67362	.00000	.26744	3.73917

	ave_dist	-.08658	.01251	-.49407	-6.92040	.00000	.37505	2.66629
	lvr	-.00025	.00005	-.31303	-4.63382	.00001	.41889	2.38726
	wod	-.00013	.00016	-.06766	-.77494	.43898	.25078	3.98756
	jwr	.00100	.00028	.19751	3.51880	.00050	.60674	1.64816
	edertt1000	.04270	.03563	.13619	1.19868	.23160	.14809	6.75260

a. Dependent Variable: ln_train_wo

PLACE OF WORK (POW MODEL) FOR THE ALL SECTOR SETI H1

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.710a	.505	.480	1.01142

a. Predictors: (Constant), edertt1000, jwr, p_jobret, All_job, p_jobcon, wage_mp, p_jobman, lvnr, ave_dist, p_manpr, wage_bl, jobd, strio, ptiori, p_blu
b. Dependent Variable: Ln of proportion of train ridership

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	312.747	15	20.850	20.382	.000b
	Residual	306.890	300	1.023		
	Total	619.637	315			

a. Dependent Variable: Ln of proportion of train ridership
b. Predictors: (Constant), edertt1000, jwr, p_jobret, All_job, p_jobcon, wage_mp, p_jobman, lvnr, ave_dist, p_manpr, wage_bl, jobd, strio, ptiori, p_blu

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-7.62535	1.25269		-6.08717	.00000		
	All_job	.00001	.00002	.04554	.55917	.57647	.24889	4.01787
	p_manpr	.75247	1.24353	.04653	.60511	.54556	.27926	3.58090
	p_blu	-2.28026	1.32302	-.18059	-1.72353	.08582	.15037	6.65015
	p_jobman	2.50199	1.24868	.14445	2.00370	.04600	.31764	3.14822
	p_jobcon	-.57287	1.29897	-.02496	-.44102	.65952	.51551	1.93984
	p_jobret	2.24413	1.04432	.13054	2.14889	.03244	.44740	2.23514
	wage_mp	.03025	.02029	.09365	1.49076	.13707	.41831	2.39059
	wage_bl	-.01031	.02301	-.02802	-.44815	.65437	.42230	2.36798
	ptiori	.00000	.00000	-.08577	-1.12908	.25976	.28611	3.49514
	strio	.00009	.00003	.19708	2.75533	.00622	.32268	3.09909
	ave_dist	-.08324	.01821	-.29430	-4.57197	.00001	.39843	2.50985
	lvnr	.00020	.00007	.17445	2.92490	.00371	.46407	2.15483
	jobd	.00028	.00009	.20840	3.05537	.00245	.35486	2.81801
	jwr	-.00018	.00039	-.02148	-.45067	.65255	.72696	1.37559
	edertt1000	.11485	.03938	.22694	2.91647	.00381	.27266	3.66762

a. Dependent Variable: Ln of proportion of train ridership

APPENDIX 60 EXAMPLE OF MODELS WHEN THE EFFECTIVE DENSITY IS CALCULATED BASED ON THE CAR ACCESSIBILITY OR CAR TRAVEL TIME – PLACE OF RESIDENCE (POR) MODEL
ALL SECTOR SETI H1

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.647a	.418	.389	.67906

a. Predictors: (Constant), eder_ttcar, ln_jwr, p_worker_man, p_woret, car_own, inc_blu_r, p_wocon, ln_avedist, sqrt_ptiori, inc_pr_r, ln_lvr, wod, pmanpr_r, ln_strio, pblu_r
b. Dependent Variable: ln_train_wo

ANOVAa

Model	Sum of Squares	df	Mean Square	F	Sig.

1	Regression	99.533	15	6.636	14.390	.000b
	Residual	138.337	300	.461		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), eder_ttcar, ln_jwr, p_worker_man, p_woret, car_own, inc_blu_r, p_wocon, ln_avedist, sqrt_ptiori, inc_pr_r, ln_lvr, wod, pmanpr_r, ln_strio, pblu_r

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-10.438	3.814		-2.737	.007		
	pmanpr_r	1.095	.753	.158	1.454	.147	.165	6.064
	pblu_r	-1.948	.857	-.259	-2.274	.024	.150	6.670
	p_worker_man	2.268	.984	.133	2.304	.022	.579	1.727
	p_wocon	3.314	1.893	.133	1.750	.081	.333	2.999
	p_woret	2.881	1.762	.090	1.635	.103	.641	1.559
	inc_pr_r	-.048	.011	-.417	-4.490	.000	.225	4.445
	inc_blu_r	-.001	.015	-.007	-.098	.922	.432	2.317
	car_own	-.613	.141	-.292	-4.341	.000	.427	2.340
	sqrt_ptiori	.000	.000	-.096	-1.456	.147	.450	2.225
	ln_strio	1.388	.383	.368	3.624	.000	.188	5.306
	ln_avedist	-.680	.091	-.538	-7.503	.000	.378	2.649
	ln_lvr	-.264	.061	-.314	-4.298	.000	.363	2.756
	wod	.000	.000	.119	1.445	.150	.287	3.479
	ln_jwr	.085	.041	.147	2.040	.042	.373	2.683
	eder_ttcar	-.004	.011	-.037	-.345	.730	.165	6.052

a. Dependent Variable: ln_train_wo

ALL SECTOR SETI H1B

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650a	.423	.392	.67752

a. Predictors: (Constant), ederttcar_ln, ln_strio, p_woret, p_worker_man, ln_jwr, car_own, inc_blu_r, sqrt_ptiori, p_wocon, inc_pr_r, ln_lvr, wod, pmanpr_r, pblu_r, ln_avedist, eder_ttcar
b. Dependent Variable: ln_train_wo

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	100.620	16	6.289	13.700	.000b
	Residual	137.251	299	.459		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ederttcar_ln, ln_strio, p_woret, p_worker_man, ln_jwr, car_own, inc_blu_r, sqrt_ptiori, p_wocon, inc_pr_r, ln_lvr, wod, pmanpr_r, pblu_r, ln_avedist, eder_ttcar

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-10.945	3.820		-2.865	.004		
	pmanpr_r	1.265	.759	.182	1.666	.097	.161	6.194
	pblu_r	-1.694	.870	-.225	-1.946	.053	.145	6.918
	p_worker_man	2.065	.991	.121	2.084	.038	.569	1.758
	p_wocon	3.085	1.895	.124	1.628	.105	.331	3.018
	p_woret	2.357	1.791	.074	1.316	.189	.618	1.618
	inc_pr_r	-.045	.011	-.388	-4.109	.000	.216	4.624

	inc_blu_r	.000	.015	-.002	-.031	.975	.431	2.322
	car_own	-.630	.141	-.300	-4.457	.000	.425	2.354
	sqrt_ptiori	.000	.000	-.093	-1.426	.155	.449	2.226
	ln_strio	1.483	.387	.393	3.831	.000	.184	5.445
	ln_avedist	-.938	.191	-.742	-4.920	.000	.085	11.782
	ln_lvr	-.282	.062	-.336	-4.524	.000	.349	2.863
	wod	.000	.000	.148	1.756	.080	.273	3.663
	ln_jwr	.101	.043	.175	2.357	.019	.351	2.851
	eder_ttcar	-.020	.015	-.202	-1.329	.185	.083	12.009
	ederttcar_ln	.011	.007	.172	1.539	.125	.154	6.500

a. Dependent Variable: ln_train_wo

ALL SECTOR SETI H2

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656a	.430	.401	.67241

a. Predictors: (Constant), ejd_ttcar, inc_blu_r, p_woret, p_worker_man, wod, car_own, p_wocon, inc_pr_r, sqrt_ptiori, ln_avedist, ln_lvr, ln_jwr, ln_strio, pmanpr_r, pblu_r
b. Dependent Variable: ln_train_wo

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	102.231	15	6.815	15.074	.000b
	Residual	135.640	300	.452		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejd_ttcar, inc_blu_r, p_woret, p_worker_man, wod, car_own, p_wocon, inc_pr_r, sqrt_ptiori, ln_avedist, ln_lvr, ln_jwr, ln_strio, pmanpr_r, pblu_r

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.00000	(Constant)	-9.21369	2.92949		-3.14515	.00183		
	pmanpr_r	1.40504	.75562	.20227	1.85946	.06394	.16063	6.22530
	pblu_r	-1.60314	.85802	-.21281	-1.86840	.06268	.14652	6.82507
	p_worker_man	2.19687	.97429	.12912	2.25485	.02486	.57970	1.72503
	p_wocon	3.42029	1.86588	.13776	1.83307	.06778	.33656	2.97122
	p_woret	2.78121	1.74521	.08676	1.59362	.11207	.64124	1.55948
	inc_pr_r	-.04336	.01081	-.37358	-4.00954	.00008	.21895	4.56717
	inc_blu_r	.00243	.01439	.01119	.16878	.86608	.43270	2.31107
	car_own	-.68626	.14295	-.32739	-4.80069	.00000	.40869	2.44683
	sqrt_ptiori	-.00010	.00009	-.06750	-1.07814	.28184	.48493	2.06216
	ln_strio	1.24328	.30437	.32922	4.08474	.00006	.29261	3.41751
	ln_avedist	-.68747	.08968	-.54351	-7.66541	.00000	.37808	2.64496
	ln_lvr	-.28159	.06117	-.33530	-4.60318	.00001	.35825	2.79135
	wod	.00024	.00014	.12910	1.69646	.09084	.32824	3.04652
	ln_jwr	.11901	.04336	.20686	2.74473	.00642	.33464	2.98825
	ejd_ttcar	-.00893	.00362	-.17012	-2.46728	.01417	.39980	2.50128

a. Dependent Variable: ln_train_wo

ALL SECTOR SETI H2B

Model Summaryb

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.656a	.430	.399	.67348

a. Predictors: (Constant), ejdttcar_ln, p_woret, p_worker_man, car_own, inc_pr_r, p_wocon, wod, inc_blu_r, sqrt_ptiori, ln_jwr, ln_lvr, ln_strio, ejd_ttcar, pmanpr_r, ln_avedist, pblu_r
b. Dependent Variable: ln_train_wo

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	102.252	16	6.391	14.090	.000b
	Residual	135.619	299	.454		
	Total	237.871	315			

a. Dependent Variable: ln_train_wo
b. Predictors: (Constant), ejdttcar_ln, p_woret, p_worker_man, car_own, inc_pr_r, p_wocon, wod, inc_blu_r, sqrt_ptiori, ln_jwr, ln_lvr, ln_strio, ejd_ttcar, pmanpr_r, ln_avedist, pblu_r

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-9.53671	3.29091		-2.89790	.00403		
	pmanpr_r	1.42397	.76184	.20499	1.86911	.06258	.15852	6.30819
	pblu_r	-1.56966	.87315	-.20836	-1.79770	.07323	.14194	7.04541
	p_worker_man	2.18297	.97794	.12830	2.23221	.02634	.57721	1.73247
	p_wocon	3.39302	1.87308	.13666	1.81146	.07107	.33504	2.98468
	p_woret	2.71018	1.77844	.08455	1.52391	.12859	.61947	1.61428
	inc_pr_r	-.04320	.01086	-.37219	-3.97902	.00009	.21793	4.58857
	inc_blu_r	.00268	.01446	.01233	.18520	.85320	.42997	2.32575
	car_own	-.68600	.14318	-.32727	-4.79105	.00000	.40866	2.44701
	sqrt_ptiori	-.00010	.00010	-.07001	-1.09789	.27314	.46897	2.13235
	ln_strio	1.27839	.34521	.33851	3.70321	.00025	.22820	4.38214
	ln_avedist	-.71082	.14027	-.56197	-5.06750	.00000	.15505	6.44963
	ln_lvr	-.28263	.06146	-.33654	-4.59867	.00001	.35603	2.80875
	wod	.00025	.00015	.13176	1.70666	.08892	.31990	3.12599
	ln_jwr	.11958	.04351	.20784	2.74844	.00635	.33343	2.99912
	ejd_ttcar	-.00959	.00472	-.18258	-2.03235	.04300	.23627	4.23237
	ejdttcar_ln	.00110	.00508	.01847	.21676	.82854	.26268	3.80696

a. Dependent Variable: ln_train_wo

APPENDIX CHAPTER 8 WAGE EQUATIONS (THE TRADE-OFF MODEL)

APPENDIX 61 THE ALL SECTOR MODEL WITHOUT LAND USE AND EFFECTIVE DENSITY VARIABLE – POW MODEL

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.705[a]	.498	.479	1.01191

a. Predictors: (Constant), ln_avedist, inc_bl, All_job, p_manpr, p_jobcon, p_jobman, p_jobret, inc_mp, sqrt_ptiori, ln_strio, p_blu
b. Dependent Variable: Ln of proportion of train ridership

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	308.353	11	28.032	27.376	.000[b]
	Residual	311.284	304	1.024		
	Total	619.637	315			

a. Dependent Variable: Ln of proportion of train ridership
b. Predictors: (Constant), ln_avedist, inc_bl, All_job, p_manpr, p_jobcon, p_jobman, p_jobret, inc_mp, sqrt_ptiori, ln_strio, p_blu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1.00000	(Constant)	-9.15989	3.70313		-2.47355	.01392		
	All_job	.00004	.00001	.14602	2.89032	.00413	.64747	1.54448
	p_manpr	.84114	1.23209	.05201	.68269	.49532	.28474	3.51194

	p_blu	-2.58329	1.32524	-.20459	-1.94930	.05218	.15002	6.66598
	p_jobman	1.61602	1.19052	.09330	1.35740	.17566	.34977	2.85900
	p_jobcon	.66551	1.26424	.02899	.52642	.59898	.54475	1.83570
	p_jobret	2.50804	1.03927	.14589	2.41326	.01640	.45220	2.21143
	inc_mp	.05577	.01926	.17265	2.89582	.00406	.46492	2.15091
	inc_bl	-.01768	.02264	-.04805	-.78082	.43551	.43632	2.29192
	sqrt_ptiori	.00010	.00015	.04209	.67053	.50303	.41934	2.38472
	ln_strio	.48035	.38454	.07881	1.24915	.21257	.41517	2.40865
	ln_avedist	-.97110	.12702	-.47569	-7.64532	.00000	.42687	2.34264

a. Dependent Variable: Ln of proportion of train ridership

APPENDIX 62 WAGE EQUATION: LOG LINEAR REGRESSION ON WAGES PREMIUM

Syntax:

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_wagemanpr
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_wagemanpr	3.444065764402756	.152428252984216	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.619[a]	.383	.382	.119783198475039

a. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr
b. Dependent Variable: ln_wagemanpr

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	104.693	7	14.956	1042.383	.000[b]
	Residual	168.776	11763	.014		
	Total	273.469	11770			

a. Dependent Variable: ln_wagemanpr
b. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	3.707	.016		237.896	.000		
	jwr_0_0.75	-.016	.003	-.051	-5.240	.000	.554	1.804
	jwr_1.5	.122	.004	.303	30.418	.000	.530	1.887
	p_jobman	.947	.019	.549	50.851	.000	.450	2.224
	p_jobcon	.016	.023	.006	.674	.500	.681	1.469
	p_jobret	-.072	.022	-.036	-3.304	.001	.448	2.232

	p_jobmanpr	-.373	.025	-.204	-14.660	.000	.270	3.703
	p_jobblu	-.691	.023	-.532	-30.551	.000	.173	5.770

a. Dependent Variable: ln_wagemanpr

APPENDIX 63 WAGE EQUATION: LOG LINEAR REGRESSION ON LAND RENTS PREMIUM

Syntax:

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_lvr
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 p_worker_man p_wocon p_woret p_worker_manpr p_worker_manblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_lvr	6.3639	1.14854	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771
p_worker_man	.0945	.05223	11771
p_wocon	.1095	.03845	11771
p_woret	.1009	.02633	11771
p_worker_manpr	.3172	.11622	11771
p_worker_manblu	.3585	.11452	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.556[a]	.309	.308	.95539

a. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu

b. Dependent Variable: ln_lvr

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4793.816	12	399.485	437.658	.000[b]
	Residual	10732.441	11758	.913		
	Total	15526.257	11770			

a. Dependent Variable: ln_lvr

b. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	10.242	.169		60.769	.000		
	jwr_0_0.75	-.121	.026	-.050	-4.697	.000	.525	1.906
	jwr_1.5	.241	.034	.079	7.130	.000	.480	2.085
	p_jobman	1.571	.178	.121	8.804	.000	.312	3.206
	p_jobcon	-.096	.206	-.005	-.469	.639	.552	1.810
	p_jobret	1.318	.178	.087	7.404	.000	.423	2.364

	p_jobmanpr	-2.564	.218	-.186	-11.766	.000	.234	4.269
	p_jobblu	-4.870	.185	-.497	-26.301	.000	.164	6.079
	p_worker_man	1.043	.204	.047	5.114	.000	.684	1.463
	p_wocon	-.508	.356	-.017	-1.428	.153	.414	2.414
	p_woret	3.587	.421	.082	8.530	.000	.632	1.581
	p_worker_manpr	-1.815	.140	-.184	-12.987	.000	.294	3.400
	p_worker_manblu	-4.295	.139	-.428	-30.877	.000	.306	3.273

a. Dependent Variable: ln_lvr

APPENDIX 64 WAGE EQUATION: LOG LINEAR REGRESSION ON EFFECTIVE EMPLOYED RESIDENT DENSITY PREMIUM

Syntax:

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_edertt1000
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 p_worker_man p_wocon p_woret p_worker_manpr p_worker_manblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_edertt1000	2.3548	.26732	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771
p_worker_man	.0945	.05223	11771
p_wocon	.1095	.03845	11771
p_woret	.1009	.02633	11771
p_worker_manpr	.3172	.11622	11771
p_worker_manblu	.3585	.11452	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.701[a]	.491	.491	.19077

a. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu

b. Dependent Variable: ln_edertt1000

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	413.180	12	34.432	946.124	.000[b]
	Residual	427.901	11758	.036		
	Total	841.081	11770			

a. Dependent Variable: ln_edertt1000

b. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.772	.034		82.360	.000		
	jwr_0_0.75	.069	.005	.123	13.498	.000	.525	1.906
	jwr_1.5	.073	.007	.104	10.907	.000	.480	2.085
	p_jobman	.788	.036	.261	22.132	.000	.312	3.206

	p_jobcon	.167	.041	.036	4.076	.000	.552	1.810
	p_jobret	.517	.036	.147	14.555	.000	.423	2.364
	p_jobmanpr	-.317	.044	-.099	-7.288	.000	.234	4.269
	p_jobblu	-1.212	.037	-.532	-32.784	.000	.164	6.079
	p_worker_man	.381	.041	.074	9.348	.000	.684	1.463
	p_wocon	-.937	.071	-.135	-13.181	.000	.414	2.414
	p_woret	1.966	.084	.194	23.420	.000	.632	1.581
	p_worker_manpr	.036	.028	.016	1.305	.192	.294	3.400
	p_worker_manblu	-.726	.028	-.311	-26.132	.000	.306	3.273

a. Dependent Variable: ln_edertt1000

APPENDIX 65 WAGE EQUATION: LOG LINEAR REGRESSION ON EFFECTIVE JOB DENSITY PREMIUM

Syntax:

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_ejdtt1000
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 p_worker_man p_wocon p_woret p_worker_manpr p_worker_manblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_ejdtt1000	2.1888	.32272	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771
p_worker_man	.0945	.05223	11771
p_wocon	.1095	.03845	11771
p_woret	.1009	.02633	11771
p_worker_manpr	.3172	.11622	11771
p_worker_manblu	.3585	.11452	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.756[a]	.572	.571	.21130

a. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu
b. Dependent Variable: ln_ejdtt1000

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	700.856	12	58.405	1308.176	.000[b]
	Residual	524.946	11758	.045		
	Total	1225.802	11770			

a. Dependent Variable: ln_ejdtt1000
b. Predictors: (Constant), p_worker_manblu, p_jobret, jwr_1.5, p_jobcon, p_woret, p_worker_man, p_jobman, jwr_0_0.75, p_jobmanpr, p_wocon, p_worker_manpr, p_jobblu

Coefficients[a]

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.	Collinearity Statistics

		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.583	.037		69.296	.000		
	jwr_0_0.75	-.053	.006	-.077	-9.280	.000	.525	1.906
	jwr_1.5	.254	.007	.297	34.089	.000	.480	2.085
	p_jobman	1.011	.039	.277	25.625	.000	.312	3.206
	p_jobcon	.061	.045	.011	1.333	.182	.552	1.810
	p_jobret	.533	.039	.126	13.546	.000	.423	2.364
	p_jobmanpr	.160	.048	.041	3.315	.001	.234	4.269
	p_jobblu	-1.104	.041	-.401	-26.958	.000	.164	6.079
	p_worker_man	.980	.045	.159	21.721	.000	.684	1.463
	p_wocon	-.753	.079	-.090	-9.564	.000	.414	2.414
	p_woret	1.815	.093	.148	19.512	.000	.632	1.581
	p_worker_manpr	.033	.031	.012	1.058	.290	.294	3.400
	p_worker_manblu	-1.219	.031	-.433	-39.622	.000	.306	3.273

a. Dependent Variable: ln_ejdtt1000

APPENDIX 66 WAGE EQUATION: LOG LINEAR REGRESSION ON TOTAL TRIP ATTRACTION PREMIUM

Syntax:

```
COMPUTE ln_alltrip_att=ln(alltrip_att).
EXECUTE.
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_alltrip_att
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Descriptive Statistics

	Mean	Std. Deviation	N
ln_alltrip_att	6.4494	1.41132	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.622[a]	.386	.386	1.10589

a. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr

b. Dependent Variable: ln_alltrip_att

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9057.684	7	1293.955	1058.020	.000[b]
	Residual	14386.115	11763	1.223		
	Total	23443.799	11770			

a. Dependent Variable: ln_alltrip_att

b. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	9.035	.144		62.797	.000		
	jwr_0_0.75	-.939	.029	-.314	-32.352	.000	.554	1.804

	jwr_1.5	.503	.037	.134	13.545	.000	.530	1.887
	p_jobman	4.778	.172	.299	27.780	.000	.450	2.224
	p_jobcon	-.804	.214	-.033	-3.750	.000	.681	1.469
	p_jobret	4.989	.200	.269	24.916	.000	.448	2.232
	p_jobmanpr	-3.071	.235	-.182	-13.069	.000	.270	3.703
	p_jobblu	-5.903	.209	-.490	-28.268	.000	.173	5.770

a. Dependent Variable: ln_alltrip_att

APPENDIX 67 WAGE EQUATION: LOG LINEAR REGRESSION ON THE PROPORTION OF TRAIN TRIP ATTRACTION PREMIUM

Syntax:

```
DATASET ACTIVATE DataSet1.
SAVE OUTFILE='C:\Users\14356234\Desktop\NEW BACKUP 10 JULI 2015\2015_DRAFT\DRAFT 1\REPORT\FULL '+
 'DISSERTATION_SITI\bahan pelengkap\data_wagetradeoff_FN.sav'
 /COMPRESSED.
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_ptrain_w22
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_jobman p_jobcon p_jobret p_jobmanpr p_jobblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_ptrain_w22	-4.398639349466825	1.006848708655876	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_jobman	.0782	.08838	11771
p_jobcon	.1072	.05764	11771
p_jobret	.1090	.07604	11771
p_jobmanpr	.3073	.08349	11771
p_jobblu	.3277	.11726	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.443[a]	.196	.195	.903100670700898

a. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr

b. Dependent Variable: ln_ptrain_w22

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2337.976	7	333.997	409.515	.000[b]
	Residual	9593.795	11763	.816		
	Total	11931.771	11770			

a. Dependent Variable: ln_ptrain_w22

b. Predictors: (Constant), p_jobblu, jwr_1.5, p_jobret, p_jobcon, jwr_0_0.75, p_jobman, p_jobmanpr

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-5.179	.117		-44.077	.000		
	jwr_0_0.75	.103	.024	.048	4.329	.000	.554	1.804
	jwr_1.5	.476	.030	.178	15.678	.000	.530	1.887
	p_jobman	-.634	.140	-.056	-4.515	.000	.450	2.224
	p_jobcon	-.324	.175	-.019	-1.854	.064	.681	1.469
	p_jobret	1.356	.164	.102	8.293	.000	.448	2.232
	p_jobmanpr	3.263	.192	.271	17.009	.000	.270	3.703

APPENDIX 68 WAGE EQUATION: LOG LINEAR REGRESSION ON THE TOTAL TRIP PRODUCTION PREMIUM

Syntax:

```
DATASET ACTIVATE DataSet1.
SAVE OUTFILE='C:\Users\14356234\Desktop\NEW BACKUP 10 JULI 2015\2015_DRAFT\DRAFT 1\REPORT\FULL '+
 'DISSERTATION_SITI\bahan pelengkap\data_wagetradeoff_FN.sav'
 /COMPRESSED.
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_alltrip_prod
 /METHOD=ENTER jwr_0_0.75 jwr_1.5
 p_worker_man p_wocon p_woret p_worker_manpr p_worker_manblu
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_alltrip_prod	6.8775	1.98847	11771
jwr_0_0.75	.6657	.47176	11771
jwr_1.5	.1714	.37691	11771
p_worker_man	.0945	.05223	11771
p_wocon	.1095	.03845	11771
p_woret	.1009	.02633	11771
p_worker_manpr	.3172	.11622	11771
p_worker_manblu	.3585	.11452	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.605[a]	.366	.366	1.58365

a. Predictors: (Constant), p_worker_manblu, jwr_1.5, p_woret, p_worker_man, p_wocon, jwr_0_0.75, p_worker_manpr
b. Dependent Variable: ln_alltrip_prod

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17037.655	7	2433.951	970.498	.000[b]
	Residual	29500.895	11763	2.508		
	Total	46538.549	11770			

a. Dependent Variable: ln_alltrip_prod
b. Predictors: (Constant), p_worker_manblu, jwr_1.5, p_woret, p_worker_man, p_wocon, jwr_0_0.75, p_worker_manpr

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	8.095	.168		48.104	.000		
	jwr_0_0.75	.323	.042	.077	7.719	.000	.547	1.827
	jwr_1.5	-1.345	.054	-.255	-24.714	.000	.506	1.974
	p_worker_man	-1.280	.332	-.034	-3.861	.000	.710	1.408
	p_wocon	4.814	.476	.093	10.117	.000	.637	1.570
	p_woret	19.519	.626	.258	31.163	.000	.783	1.277
	p_worker_manpr	-3.067	.214	-.179	-14.346	.000	.345	2.898
	p_worker_manblu	-7.264	.217	-.418	-33.491	.000	.345	2.896

a. Dependent Variable: ln_alltrip_prod

APPENDIX 69 WAGE EQUATION: LOG LINEAR REGRESSION ON THE PROPORTION OF TRAIN TRIP PRODUCTION PREMIUM

Syntax:

```
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
```

```
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT ln_ptrain_wo
/METHOD=ENTER jwr_0_0.75 jwr_1.5
p_worker_man p_wocon p_woret p_worker_manpr p_worker_manblu
/SCATTERPLOT=(*ZRESID ,*ZPRED)
/RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Model:

Descriptive Statistics

	Mean	Std. Deviation	N
ln_ptrain_wo	-2.989081587782398	.664430688716655	10952
jwr_0_0.75	.7027	.45709	10952
jwr_1.5	.1263	.33218	10952
p_worker_man	.0925	.02848	10952
p_wocon	.1129	.03348	10952
p_woret	.1028	.01750	10952
p_worker_manpr	.3181	.11337	10952
p_worker_manblu	.3511	.10362	10952

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.365[a]	.133	.133	.618810854385207

a. Predictors: (Constant), p_worker_manblu, jwr_0_0.75, p_woret, p_wocon, jwr_1.5, p_worker_man, p_worker_manpr
b. Dependent Variable: ln_ptrain_wo

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	643.766	7	91.967	240.167	.000[b]
	Residual	4190.752	10944	.383		
	Total	4834.518	10951			

a. Dependent Variable: ln_ptrain_wo
b. Predictors: (Constant), p_worker_manblu, jwr_0_0.75, p_woret, p_wocon, jwr_1.5, p_worker_man, p_worker_manpr

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	.753	.171		4.418	.000		
	jwr_0_0.75	.287	.017	.198	16.741	.000	.568	1.760
	jwr_1.5	.363	.023	.182	15.786	.000	.598	1.672
	p_worker_man	-.378	.369	-.016	-1.025	.305	.317	3.156
	p_wocon	.126	.238	.006	.529	.597	.552	1.811
	p_woret	-1.235	.463	-.033	-2.669	.008	.533	1.876
	p_worker_manpr	-5.531	.205	-.944	-26.941	.000	.065	15.494
	p_worker_manblu	-5.932	.213	-.925	-27.884	.000	.072	13.899

a. Dependent Variable: ln_ptrain_wo

APPENDIX 70 WAGE EQUATION: LOG LINEAR REGRESSION ON THE TRADE OFF BETWEEN THE HOURLY WAGES (MANAGERS/PROFESSIONALS OCCUPATION) AND TRAVEL TIME PARK AND RIDE FOR EACH SUBURB CATEGORIES BASED ON JOB-HOUSING BALANCE CRITERIAM

Syntax:

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT ln_wagemanpr
/METHOD=ENTER avettpnr jwr_0_0.75 jwr_1.5 avett_cat1 avett_cat3.
```

Model:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.550[a]	.302	.302	.127360677591469

a. Predictors: (Constant), avett_cat3, avettpnr, avett_cat1, jwr_0_0.75, jwr_1.5

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	82.632	5	16.526	1018.838	.000[b]
	Residual	190.837	11765	.016		
	Total	273.469	11770			

a. Dependent Variable: ln_wagemanpr
b. Predictors: (Constant), avett_cat3, avettpnr, avett_cat1, jwr_0_0.75, jwr_1.5

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.875	.017		223.130	.000
	avettpnr	-.005	.000	-.344	-26.229	.000
	jwr_0_0.75	-.199	.022	-.616	-9.094	.000
	jwr_1.5	-.286	.035	-.707	-8.217	.000
	avett_cat1	.002	.000	.474	7.242	.000
	avett_cat3	.005	.000	1.052	12.525	.000

a. Dependent Variable: ln_wagemanpr

APPENDIX 71 WAGE EQUATION OF TRAVEL TIME SAVING (MANAGERS/PROFESSIONAL OCCUPATION) BASED ON JWR CATEGORY

Syntax:

```
COMPUTE avett_cat1=avettpnr*jwr_0_0.75.
EXECUTE.
COMPUTE avett_cat2=avettpnr*jwr_0.75_1.5.
EXECUTE.
COMPUTE avett_cat3=avettpnr*jwr_1.5.
EXECUTE.
REGRESSION
 /DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT ln_wagemanpr
 /METHOD=ENTER avett_cat1 avett_cat2 avett_cat3
 /SCATTERPLOT=(*ZRESID ,*ZPRED)
 /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

Descriptive Statistics

	Mean	Std. Deviation	N
ln_wagemanpr	3.444065764402756	.152428252984216	11771
avett_cat1	53.8609	38.83715	11771
avett_cat2	14.1110	32.54125	11771
avett_cat3	13.4109	29.64188	11771

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.544[a]	.296	.296	.127932986086102

a. Predictors: (Constant), avett_cat3, avett_cat2, avett_cat1
b. Dependent Variable: ln_wagemanpr

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	80.880	3	26.960	1647.229	.000[b]

	Residual	192.589	11767	.016			
	Total	273.469	11770				

a. Dependent Variable: ln_wagemanpr

b. Predictors: (Constant), avett_cat3, avett_cat2, avett_cat1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1.000000	(Constant)	3.731881	.010022		372.372124	.000		
	avett_cat1	-.004011	.000124	-1.021846	-32.420274	.000	.060245	16.598910
	avett_cat2	-.003580	.000117	-.764285	-30.537938	.000	.095549	10.465841
	avett_cat3	-.001587	.000132	-.308656	-12.019219	.000	.090753	11.018923

a. Dependent Variable: ln_wagemanpr

APPENDIX 72 RELATED PUBLICATIONS IN CONFERENCE PROCEEDINGS

Nurlaela, S., & Curtis, C. (2012). Modeling household residential location choice and travel behavior and its relationship with public transport accessibility. *Procedia Social and Behavioral Sciences, 54*, 56-64. doi: 10.1016/j.sbspro.2012.09.725

Nurlaela, S., & Pamungkas, A. (2014). Assessing the impact of accessibility on property value capitalisation post-Perth - Mandurah railway opening. *Procedia Social and Behavioral Sciences, 135*, 96-105. doi: 10.1016/j.sbspro.2014.07.331