Developing a methodology for ex post evaluation of the wider impact of the restoration of rail services to previously disconnected or isolated regions on employment and property prices and accessibility to jobs and essential services.

John Stephen Dixon

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Institute for Transport Studies

December 2018

"I won't be going again on the slow train. On the main line and the goods siding the grass grows high at Dog Dyke, Tumby Woodside and Trouble House Halt. The sleepers sleep at Audlem and Ambergate. No passenger waits on Chittening platform or Cheslyn Hay."

The Slow Train - Michael Flanders and Donald Swann (1963)

Acknowledgements

The author wishes to express his sincere gratitude, respect and heartfelt thanks to his supervisors, Professor Richard Batley and Mr. Dan Johnson for their invaluable guidance, patience, and encouragement throughout his study period. Without their generous support and help, this thesis would not have attained its present shape.

The author is especially indebted to the Institute for Transport Studies at the University of Leeds for providing the ideal background to further his study, and to fellow students who supplied encouragement along the way. Special thanks must also be conveyed to Transport Scotland and the people of the Scottish Borders region for permitting the author to collect the initial data which acted as a catalyst for this research. Also to his father Jack who instilled in him the belief that the pursuit of education was a lifelong activity.

Finally, the author is honoured to dedicate this work to his wife, Anne and daughters Elizabeth and Sarah who encouraged him to take on the task, and supported him unequivocally during his study period through the darkest hours.

ABSTRACT

As well as improving access to jobs and essential services, the re-establishment of rail links to larger, more remote areas may produce wider economic and social impacts both within the region and beyond. Recent developments have highlighted the need for improved ex post evaluation of such impacts, particularly in formerly disconnected or isolated regions. The main thrust of this research was through investigating ex post situations both spatially and temporally to determine cause-effect relationships. This required developing a methodological approach which would match those objectives and adapted pre-existing methods to develop a methodology for appraisal particularly relevant to remote, rural or disconnected regions.

Using three case studies - the Robin Hood line (1998), the Stirling-Alloa line (2008), and Borders Rail (2015) as representing different stages of recent rail investment in previously disconnected regions, and applying mainly secondary data sources, a counterfactual was developed which allowed a meaningful comparison between areas subject to treatment i.e. rail intervention, and those not treated i.e. either unaffected or minimally affected by the intervention and to establish any differences between findings in urban studies. Treatment groups were based on distance thresholds where the control group was selected from remaining locations in the region. There appeared to be some benefit in application of clustering and propensity matching to effect a more balanced comparison between similar locations in the treatment and control groups.

An important consideration was the accessibility characteristic which conventionally has been distance to the nearest rail station. However, two additional measures were utilised here: a distance to station ratio (which measured the percentage improvement in distance to station following the rail intervention) and a job accessibility index which assessed the improvement in access to jobs based on skills matching and the cost of commuting.

The job accessibility index was developed to take into account the limitations in travel in more remote communities where services are less frequent and commuting distances often greater than in the urban situation. The cost of travel was recognised as a key factor affecting accessibility and generalised cost allowed the cost of commuting to be calculated using local values of speed and cost of transport. Job accessibility was based on comparing the percentage skills share at each location, matched to actual jobs at all neighbouring destination locations. The job accessibility index allowed a measure of accessibility based on the original job market, but could also be used to assess the effect on accessibility of a slump in employment by considering the current job market.

Without job skills matching, job accessibility appeared to be overestimated as the seemingly high attraction of job opportunities may not always synchronise with the skills set in that location. It yielded good results when used as an accessibility characteristic in the hedonic models, being a more complex measure than distance from rail station as it encompassed the whole regional employment picture relative to each location.

Previous research had suggested some correspondence between rail access improvements and increased property price and employment levels. Four different approaches were examined here to assess causality: a descriptive comparison approach, a DID (difference-in-difference) model, a fixed effects hedonic model and a GWR (Geographically Weighted Regression) model. These incorporated other factors such as changes in local and property characteristics over the period spanning each intervention.

The descriptive approach looked at individual variables in isolation pre- and post-intervention broken into treatment and control to assess any impacts but ignored the combined effect of other explanatory factors. The output indicated a discernible effect of treatment in some cases, and was useful in corroborating variables to carry forward to the model. For property impacts, the difference-in-difference approach produced contrasting findings for the case study regions. For job impacts, there was a positive effect on employment density of being closer to a rail station and of improvement in job accessibility, but for Borders Rail this was not statistically significant which may be due to the limited amount of data available at this stage. The fixed effects model showed that for property impacts the distance to rail station and distance ratio and improvement in job accessibility were all significant factors.

A modified spatial-temporal version of Geographically Weighted Regression estimated local parameters through time by examining changes in coefficients for two separate years spanning the intervention for each case study. The property model showed variation across each region in the negative relationship between price and distance to the nearest station for both the established case study regions. For the jobs model, the relationship between employment density and distance to the nearest station showed that the distance from the rail network was critical in terms of the job market.

The findings suggest some causality linking rail investment to house price changes and employment density, dependent on the scale of the rail intervention and the regional context. Improvements in rail transport infrastructure could produce economic benefits affected by the proximity to new stations and relate to the effect on property prices. Although improved job accessibility allowing increased commuting, spatial, temporal and economic barriers may still prevent more economically vulnerable neighbourhoods within each region from receiving the full benefit of the intervention.

In conclusion, there are implications for practice in terms of making a case for new rail infrastructure, application in a WebTag style appraisal or evaluation, and new information on spatial patterns of employment and property prices. In addition, consideration is given to expansion of the methodology to other types of transport intervention as well as application in an urban context.

CONTENTS

List of figures

List of Tables

1 Chapter One: Introduction

1.1 Background and rationale

This chapter introduces the rationale for this research, presents the aims, objectives and research questions, and provides a schematic overview of the structure of the thesis.

1.1.1 The reduction in the rail network

The Beeching cuts in the 1960s brought about a reduction of route network and a restructuring of the railways in Great Britain, according to a plan outlined in two reports, *The Reshaping of British Railways* (British Railways Board, 1963) and *The Development of the Major Railway Trunk Routes* (British Railways Board, 1965). One third of the rail network was shut down (Figure 1-1) including many branch lines, with the objective of abating the large losses being incurred from increasing road transport competition.

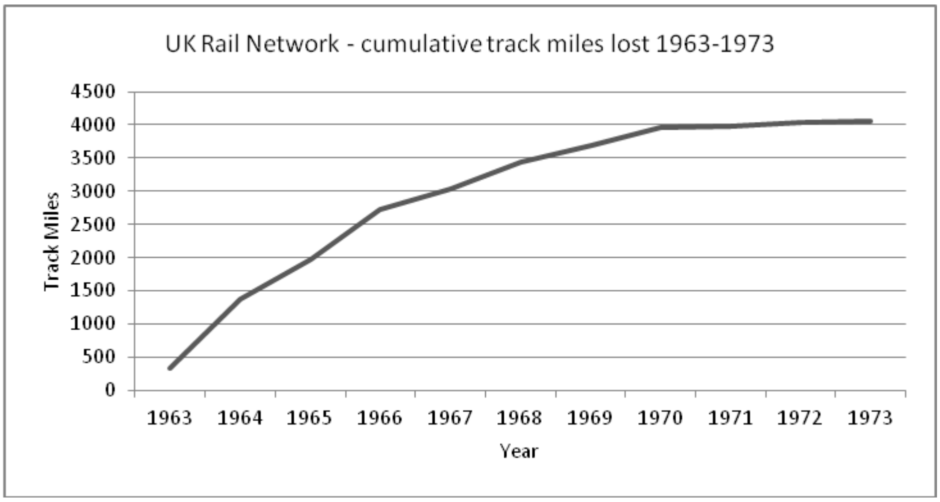


Figure 1-1 UK Rail Network - cumulative track miles lost 1963-1973

(Source: Gourvish, 1974)

The Reshaping of British Railways identified 2,363 stations and 5,000 miles of railway line as suitable for closure, 55% of stations and 30% of route miles. The objective was to stem the large losses being incurred during a period of increasing competition from road transport and reduce the rail subsidies necessary to keep the network running. Beeching's analysis showed that the least-used 1,762 stations had annual passenger receipts of less than £2,500 each, that over half of the 4,300 stations open to passengers in 1960 had receipts of less than £10,000, that the least-used 50% of stations contributed only 2% of passenger revenue, and that one third of route miles carried just 1% of passengers.

His second report, *The Development of the Major Railway Trunk Routes*, identified a small number of major routes for significant investment, concluding that of the 7,500 miles of trunk railway only 3,000 miles "should be selected for future development" and invested in. Traffic would be routed along nine lines; traffic to the West Midlands, the North West and Scotland would be routed through the West Coast Main Line to Carlisle and Glasgow; traffic to the North East would be concentrated through the East Coast Main Line, to be closed north of Newcastle; and traffic to Wales and the West would follow the Great Western Main Line to Swansea and Plymouth.

Many of the recommended closures sparked protests from communities that would lose their train service, some of which (especially rural communities) had no alternative public transport. Some stations and lines were saved, but the majority were closed as planned. A few routes have since reopened, and some short sections have been preserved as heritage railways, whilst others have been incorporated into the National Cycle Network or used for road schemes; others are now lost to construction, simply reverted to farm land, or remain derelict. After the early 1970s, with a few exceptions, proposals to close other lines were met with great public opposition and quietly shelved. In the early 1980s, the possibility of more Beeching-style cuts was raised again in the Serpell Report (Department of Transport, 1982) which set out a number of options. One option involved there being no railways west of Bristol and none in Scotland apart from the central belt, and included the closure of the Midland Main Line, and even the East Coast Main Line between Berwick-upon-Tweed and Edinburgh. The report met with fierce resistance from many quarters and was quickly abandoned.

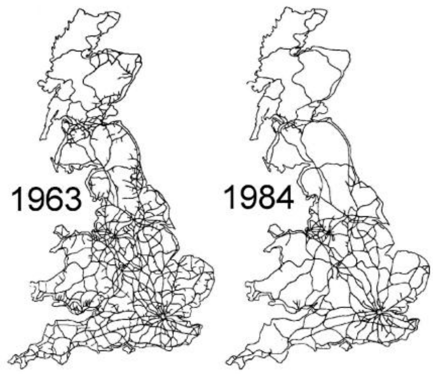


Figure 1-2 Britain’s railways before and after the Beeching cuts.
(Source: The National Council on Inland Transport, 1985).

There has been much debate since Beeching regarding the severity of the action as the closures took place at a time when car travel was in the ascendancy, and rail was deemed an outmoded method of travel, to be confined to main arterial routes. Decisions to close were based largely on passenger numbers and financial considerations, with little interest in the impact on communities, particularly those in more isolated locations (Campaign for Better Transport, 2012).

1.1.2 Transport trends since 1960

Since 1952 (the earliest date for which comparable figures are available), the UK experienced a growth of car use, which increased its modal share, accompanied by a decline in the use of buses and slow growth in rail use. However, since the 1990s, rail has started increasing its modal share at the expense of cars, increasing from 5% to 10% of passenger-kilometres travelled). In 1952, 27% of distance travelled was by car or taxi; with 42% being by bus or coach and 18% by rail. Transport modal share from 1952-2016 (Figure 1-3) shows the initial rise in car use which then peaked and started a gentle decline as rail use has increased in the past twenty years.

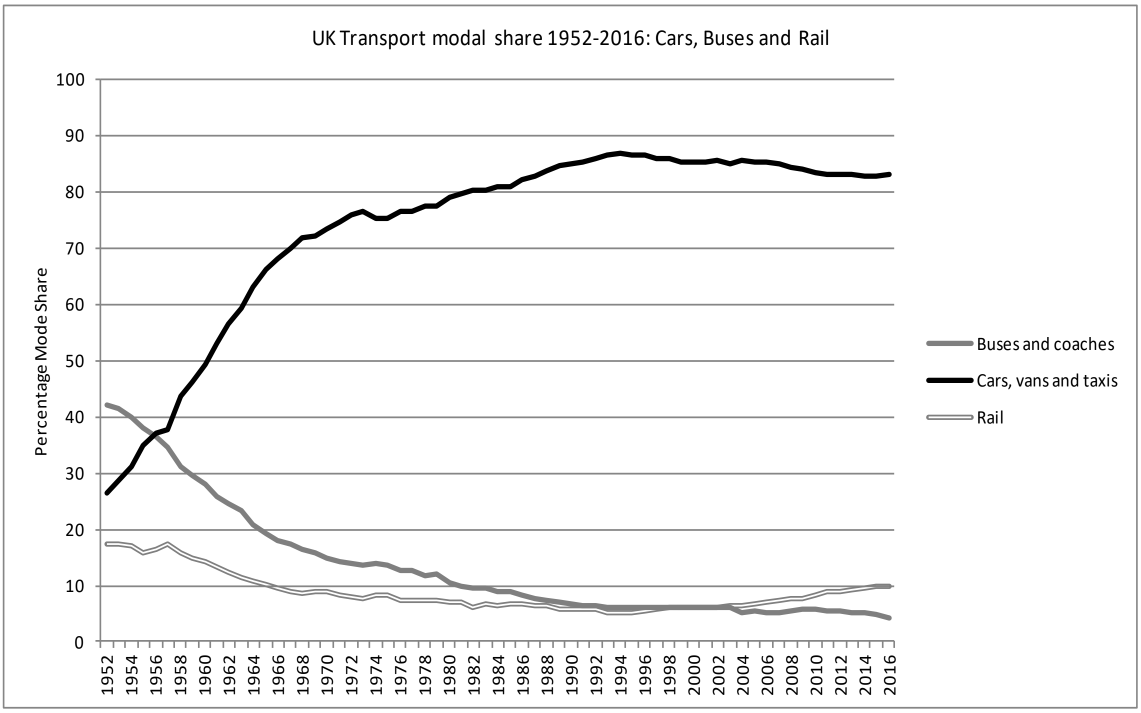


Figure 1-3 UK Transport modal share 1952-2016 (based on passenger kms)
(*Source: DfT Statistics: Passenger transport: by mode, annual from 1952., 2018*)

By 2015 83% of distance travelled was by car or taxi; with 5% being by bus and 10% by rail. In terms of passenger-kilometres, slightly over 662 billion were made by cars, motorcycles vans and taxis, 78 billion by rail, and 39 billion by bus.

Although the decline in rail use led to a reduction in the length of the rail network, the length of the road network has not increased in proportion to the increase in road use. Whereas the rail network has halved from 19,471 miles in 1950 to 10,014 miles today, the major road network has only increased from 44,710 miles to 50,265 miles in a similar period (DfT, 2018). In the years since Beeching, there has been a rapid growth in car ownership (UK Census, 1991, 2001, 2011), but the road building and maintenance programme has not kept pace with demand, resulting in frequent congestion of roads and delayed journeys as shown in *Road congestion and reliability statistics, 2012-2016* (DfT, 2016b).

Traffic congestion is one of the most serious transport problems facing the UK (Department for Transport, 2008), and bottleneck roads are in serious danger of becoming so congested that it may damage the economy (Eddington, 2006). More recently, there has been a great upsurge in the popularity of rail travel as shown by the following chart (Figure 1-4) which shows an increase in passenger numbers since 1995. There have been record levels of passengers on the railways, and rail travel has become increasingly popular, especially for commuting.

Figure 1-4 Rail Passengers in Great Britain from 1829-2016
(Source: Association of Train Companies Billion Passenger, 2008)

Figure 1-4 tracks the early era of small companies, the amalgamation into the "Big Four", nationalisation and finally the current era of privatisation. Figure 1-5 focuses on the period since 1946 and highlights the post-war slump, the Beeching era and the gradual steady rise since 1980 to the present day.

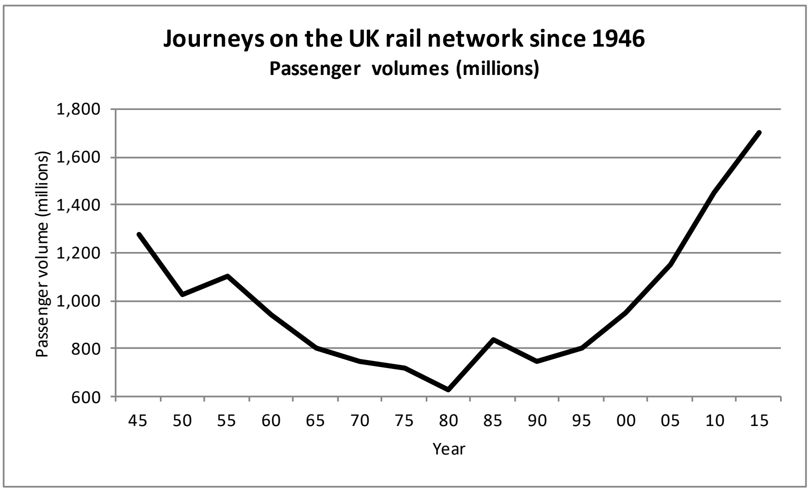


Figure 1-5 Journeys on UK Rail Network since 1946
(Source: Office of Rail Regulation, 2016)

1.1.3 The current transport situation

Since 1952, growth in car traffic has risen dramatically, reflecting increasing household car availability (Figure 1-6), enabling people to make longer trips, and leading to an increase in the distance travelled by individuals on average. In 2014, 90% of passenger journeys were by road which was a similar proportion to 1952, when records began.

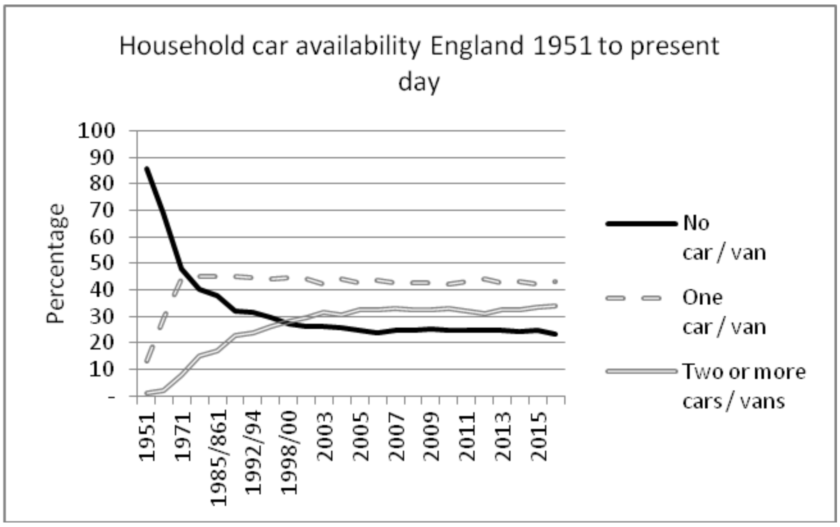


Figure 1-6 Household car availability England 1951 to present day
(Source: National Travel Survey, 2014)

However, the distance travelled and the vehicles used have changed considerably over the last 60 years. Distance travelled by car or vans has increased by over 1000%. 84% of all personal trips - which include commuting, shopping, education and leisure travel - were made by car, as a driver or passenger (DfT, 2016a). The long term trend of growth in traffic has mainly been a result of growth in car ownership. Strong growth in household access to a car between the 1950s and 1970s was followed by growth in households with two or more cars in recent decades (Department for Transport, 2016).

There still appears to be some scope for further growth in ownership; according to the National Travel Survey (Department for Transport, 2014), 48% of those in the lowest income quintile households are without access to a car.

For all the main reasons for travelling, most journeys are made by road, in particular by car. In fact, two-thirds of commuting/business journeys are made by car. The continuing reliance on the car as the chief mode of transport coupled with the decline in bus services has had particular effects in rural and semi-urban areas distant from the main centres of population. These may be regions which either have never possessed alternative modes of transport such as rail, or may have once had a rail link which disappeared in the 60s with the demise of the rail network.

In addressing transport constraints in communities disconnected either through remoteness of location or social exclusion through age, disability, child care or unemployment, in *Making the Connections* (Social Exclusion Unit, 2003) found that poor transport provision and accessibility to services may reinforce social and economic exclusion by preventing people from accessing key local services or activities, such as jobs, learning, healthcare, shopping or leisure. There was a reliance on the car for those who could afford it, whereas the disadvantaged must rely on limited public transport mode choice using the bus as the main mode of transport where rail was not an option.

Ten years later, the House of Commons Environmental Audit Committee (2013) suggested that budget restrictions in central and local government had caused further deterioration in accessibility. In particular, travel times to key services, particularly hospitals, were steadily increasing over time. Funding cuts and the centralisation agenda had particularly affected rural areas, which had suffered a substantial decline in facilities and services (Action with Communities in Rural England, 2014).

The Consumer Council (2013) identified difficulties travelling to and from facilities, and highlighted the need to tackle co-ordination, cost and flexibility of transport services. Transport costs were often comparatively high, and developments including housing, hospitals, business and out of town retail centres, were frequently located in areas not easily accessible to people without a car. Accessibility difficulties were further exacerbated by the increasing centralisation of essential services and shopping facilities (The Consumer Council, 2013). Those without their own transport were often the least able to afford the high costs of public transport, and lower income groups generally paid a higher proportion of their income on travel costs, particularly in rural areas. Here transport was frequently cited as the overriding problem, where access to services equated to access to transport (Scottish Executive Poverty Inclusion Working

Group, 2001). In Scotland, the scattered and often sparse population made progress towards tackling poverty and social exclusion more problematic, in contrast to urban areas where more typically there were concentrations of deprivation.

In The Role of Transport on Social Exclusion in Urban Scotland Literature Review, Gaffron et al. (2001) suggested that assessment of the impact of transport on exclusion required identification of the types of activity from which an individual is excluded, determining the factors responsible for this exclusion, and the extent to which transport is implicated. There was a need to classify areas of activity where transport could have most effect. These included labour markets and employment, education, health, housing markets, social networks and public utilities. It was also necessary to identify those transport-related mechanisms of exclusion creating this effect, whether physical, temporal, economic, spatial or psychological.

Transport policy could be seen as an element of supply-side employment policy (Green and Owen, 2006). New public transport services and free bus passes in deprived areas could enable people to seek work over a wider area, and also help in accessing other services and participating in social life. Good information about services was as important as good services themselves (Lucas et al., 2004).

1.1.4 Isolation of communities

"There is evidence that people in rural and sparsely settled districts do not necessarily make more or fewer trips than people in urban areas, but the distances travelled and times taken are significantly higher" (Nutley, 2003,p.59).

Past recessions have had very unequal effects across the UK (Green and Owen, 2006), and despite extensive literature on structural economic change, there is less indication of the impact on local neighbourhoods and the communities that live in them, and "more is known about the effects on businesses and services than on local people" (Audit Commission, 2009, p.51). The housing market, local services and community relations are all affected, and the impact is felt by small neighbourhoods and the communities living in them, with some neighbourhoods more vulnerable than others.

Invariably these regions experience a degree of isolation with characteristics clearly distinct from cities and busy inter-urban networks, and experience limited choices in finding employment and accessing goods and services. Such regions are typically characterised by low population densities and low incomes, and may have limited choice options on route, travel mode or departure times. Where alternatives do exist, there is often a high cost compared to the preferred route, mode or travel time.

There has been a general tendency to concentrate vital services such as hospitals and other facilities such as shops and banks into larger centralised entities, generally in cities or larger conurbations. This has led to closures of local hospitals, schools, branches of banks and town centre shops (*House of Commons Environmental Audit Committee*, 2013). Trends in globalisation and centralisation have strengthened the position of larger regional centres and cities to the detriment of smaller and rural communities (Collits, 2000).

Small towns have suffered as a consequence of the relocation of both public and private sector services, such as banking, to larger centres of population. Even during periods of growth, shops and private sector services were withdrawing from more marginalised communities, leaving residents with inaccessible or costly shops and services (Speak and Graham, 2000). Together with improved transport and mobility, this has encouraged inhabitants of smaller centres to look to regional cities for their services. People in the hinterland are bypassing local businesses and accessing services in the regional centre. This has led to further loss of business and represents an economic leakage from the local economy reducing the ability of smaller towns to service their populations.

As a result, there has been greater reliance on car and public transport (generally buses) as a transport mode (Stokes, 2002). Historically, there has been no particular body responsible for ensuring that people can access key services and employment locations. Consequently, services have been developed with insufficient attention to accessibility, leading to greater isolation for those in rural locations (Action with Communities in Rural England, 2014).

The debate around centralisation centres on encouraging people away from the car and onto public transport. This would only come about if public transport were more accessible and economically viable. One solution to bring this about could include provision of rail access where there is either currently no rail network or railway stations are too distant (Campaign for Better Transport, 2012).

However, in making these larger centres more accessible, although beneficial in many respects, it may produce a detrimental effect for others - "the two-way road effect". An improved, faster link to a larger conurbation may provide greater access to jobs and essential services, but local businesses may suffer because local customers now have better choices available to them (DfT, 2013). There is no current measure as to how much this is affected by the location concerned, and whether any conclusions can be drawn from this in a general situation.

The JRF study of Transport and Poverty (Titheridge et al., 2014) and the review of the links between transport and poverty by Lucas et al. (2016) highlighted where the transport system may cause difficulties in accessing employment for those in low income neighbourhoods. Broadly, these problems can be grouped as:

- Limitations in the provision of transport services.
- Resource constraints on using transport.
- Travel times and their interaction with care responsibilities.

1.2 Rationale

1.2.1 The feasibility of rail interventions

This growing popularity of the railways as a mode of travel, and the feasibility and benefits of reopening lines has been discussed for some time, but more recently with greater intensity. The Campaign for Better Transport (2012) in proposing the case for rail network expansion, recommended new connections serving communities not on the network, in addition to a programme to reopen or provide new stations on existing lines, or reinstating missing links. With its focus on rural locations, The Campaign to Protect Rural England (2015) also outlined the social benefits generated by a rail reopening, and suggested the need to consider the social and economic context for the re-opening of lines, the likely areas of impact and significant factors that would need to be considered in an investment appraisal.

Increasing road congestion and environmental concerns offered the scope to selectively expand the existing network (The Independent Rail Consultancy Group, 2004), and employing former rail routes provided a lower cost way of doing this. As well as impacts on existing businesses and residents there were potential benefits to the visitor economy, combined with the increased potential for commuting, with its positive and negative consequences. This could potentially impact on local income and skills levels and productivity, as well as improve access to health and education facilities.

Expanding the rail network through utilisation of existing track beds, would cause the least interference in terms of land use (Independent Rail Consultancy Group, 2004). The additional benefits that this would bring in terms of business, jobs and social inclusion have added to that push for reinstatement.

A modest number of the railway closures have been reversed, and a small but significant number of closed stations have reopened, with passenger services restored on lines where they had been previously removed. Many of these were situated in urban metropolitan counties, where Passenger Transport Executives have a role in promoting local passenger rail use.

The Robin Hood line in Nottinghamshire connecting Worksop and Nottingham via Mansfield is a notable example. As a consequence of *The Reshaping of British Railways*, all stations on the line were closed in October 1964, leaving Mansfield as the largest town in Britain without a rail link. Following the re-opening in 1995-1998, it is now estimated that up to 3500 people travel on the line each day, with over a million using it each year. In Wales, the Ebbw Vale Railway was closed to passenger traffic on 30 April 1962, prior to the Beeching Axe, and after a gap of 46 years in 2008, passenger services were restored to the line between a new station at Ebbw Vale Parkway and Cardiff Central.

In Scotland there have been a series of recent rail station and network re-openings. These include Larkhall-Milngavie (SYSTRA, 2015), Airdrie-Bathgate (Glen, 2010), and Laurencekirk Rail Station (Canning et al., 2015). The latest high profile reopening in 2015 is the 35 mile Borders Rail link from Edinburgh to Tweedbank (Johnston and Causley, 2014). The closure of this line in 1969 left the Scottish Borders region without any rail links.

These recent rail interventions have demonstrated the potential for the re-opening of disused former railway lines by reintroducing either heavy or light rail, which may bring associated wider economic impacts on the community and local business (Independent Rail Consultancy Group, 2004), and there have been a number of appraisals - Lewes-Uckfield (Network Rail, 2008), Skipton-Colne (SELRAP, 2014), March and Wisbech (Mott MacDonald, 2014) and Tavistock-Bere (Devon County Council, 2014).

1.2.2 Types of intervention

Recent rail restorations can be broadly categorised either on a line or station basis:

- New lines
- Previously closed lines that have been re-opened
- Extensions to current lines - these can be new or part restored lines
- Extensions where there has been a doubling of the track
- New stations on existing lines
- Re-opened stations on existing lines

Although all types of rail intervention were of interest, the focus for this study is on previously closed lines that have reopened in the context of certain regional characteristics. Each region has its own particular characteristics which vary spatially and geographically. These include socio-demographic characteristics, distribution and provision of services, the local economy, and the current transport situation.

In order to assess causality and impacts, the timing of the intervention is important as this will affect the availability of data and also require a reasonable interval for impacts (especially wider economic impacts) to be realised. As this research will investigate the effects of commuting related to accessibility to jobs and property prices, an essential attribute is the existence of a link or possible link to a larger city or urban centre. This may be a very large conurbation like London or Manchester, or an important centre for trade and employment.

In addition, as isolation is important to this study, it was necessary to categorise into types of isolation: urban but with poor transport connections; mixed urban/rural; industrial decline; rural - inland or rural - coastal; sparse population.

1.3 The focus of this research

This study seeks to appraise the case for new rail interventions by projecting benefits for the local community through accessibility to jobs and services. It also seeks to measure the wider implications of rail interventions in linking remoter locations to larger towns and cities. Although some research has been carried out, there are several areas where there are either gaps in the knowledge or further avenues need to be explored, especially in developing methodology.

- Monitoring the before and after effects of changes in rail transport infrastructure.

Rail infrastructure investments can have significant impact on national and regional development. Nevertheless, the overwhelming majority of studies investigating the socio-economic impact of such infrastructure are conducted at the early stages of the project usually as a part of a feasibility study, and outcomes are rarely controlled or enriched by later studies (Geurs et al., 2009). The complexity of modern economies, and the rapidly changing economic and competitive environment, suggest a need for approaches with a broad perspective, with a shift of focus from short-term performance to critical long-term issues such as wider economic benefits, reduction of inequality between regions, and development opportunities (DfT, 2005). Research into ex post evaluation also has implications for future ex ante evaluations allowing development of further before and after survey methodology structures (Stopher and Greaves, 2006).

- Measuring the wider economic impacts of rail transport mode including the local economy (property prices) and employment/unemployment levels.

One aspect of this research is to assess the impact of changes in accessibility on the local economy through tracking movements in property prices over the period spanning the intervention. House price movements can act as a proxy for the state of the local economy. Rising house prices generally reflect higher consumer spending and higher economic growth, whereas a sharp drop in house prices adversely affects consumer confidence, construction and leads to lower economic growth. (Falls in house price both in 1990 and 2007 corresponded with an economic downturn.)

A rise in house prices creates an increase in wealth for householders. As a consequence, householders will generally be more confident about spending and borrowing, as they can potentially sell their house in an emergency. A rise in house prices also enables homeowners to take out a bigger mortgage, and households could use this to spend on other items and thus create a significant increase in consumer spending. The danger is that this could result in gentrification with the consequence that low income renters may be forced out.

The other aspect is to measure impacts of improvements in accessibility on employment through job accessibility. More recently there is recognition that rural areas should be considered differently as there is often a lack of alternatives and choices for travel, employment and suppliers. With limited job opportunities in rural areas, labour markets are often viewed as thin (Findeis and Jenson, 1998; Vera-Toscano et al., 2004), the ultimate case being a thin labour market where only one employer exists for labour. Unemployment is more pronounced and job search costs are often the cause of market failure in remote labour markets (The Scottish Government, 2009). DfT (2013) suggested that study of the employed and employment movement and impacts on business would be extremely worthwhile as currently very limited analysis had been done, and there needed to be greater emphasis on developing new analytical methods.

(A more detailed theoretical basis for transport investments producing economic benefits is set out in Chapter 4.)

- Measuring specific effects - particularly commuting - resulting from linking remoter areas with smaller economies to larger urban areas.

Laird et al. (2013) specifically focused on extending methods of measuring wider economic benefits, and considered the two-way road effect, where a rural highway improves the accessibility from the region to its market, while at the same time also facilitates access from the economic core of the country to the region. The remote region is therefore subject to new competitive forces and some substitution of activities may occur. While this is not a main consideration here, the change in commuting to jobs is analysed in detail.

As an initial phase, a survey of stakeholders in the Scottish Borders raised several issues which provided a focus for measures to consider in ex post evaluation of rail projects for previously disconnected regions. Consequently, the main focus of this research has involved consideration of changes in accessibility and the wider economy. Specifically, this related to two types of community previously disconnected from the rail network:

- More isolated or rural regions at a distance from urban centres of population.
- Urban and semi-rural areas subject to poor transport infrastructure.

It considers the effect of change in rail access on local residents in terms of improvements in the local economy, manifest in movement in property prices and accessibility to suitable jobs to match the skills availability in the area.

- Investigating the potential for extrapolating accessibility issues and wider impact benefits to other similar situations.

Although there is diversity surrounding each intervention, there are certain elements which have a commonality. Through determining the contributory factors in each situation, the prospective transferability of those elements to other situations offers a worthwhile and rewarding study, which would demonstrate potential for restoration of rail links in some regions (Campaign for Better Transport, 2012). To date there has been a limited amount of research in this area, and there is scope to fill some gaps in the knowledge.

1.4 Aims and objectives

1.4.1 Aims

The aim of this research is to develop a quantitative, scientific methodological approach to the measurement of wider economic and accessibility benefits through reference to three recent rail restorations which reconnected previously isolated communities to the rail network, where the interventions had different aims and were at different evolutionary stages of implementation. Through consideration of appropriate measures and outcomes, the methodology will adopt and compare innovative approaches to evaluate and appraise the socio-economic impact of such rail transport projects on the local community and seek to establish any causal effect of the rail intervention. The study will focus on the key outcomes of employment, property price and accessibility to jobs and services. This ex post evaluation will inform ex ante appraisal guidance for assessing the potential benefits of future new rail interventions.

1.4.2 Objectives

This can be broken down into more detailed objectives as follows:

- To gain initial insight into the current situation regarding disconnected and isolated communities.
- Through reference to appropriate case studies, to categorise relevant characteristics of the affected regions - such as the location and characteristics of the population, distance to public transport, alternative travel modes and the nature and frequency of the rail service.
- Ex post evaluation of each case study region based on impacts to date.
- To develop innovative approaches to evaluation methodology involving the application of existing statistical methods in a different context.
- To associate measures of accessibility change with measures of employment outcomes to produce suitable comparators (e.g. an accessibility index).
- By investigating ex post situations both spatially and temporally, to determine cause-effect relationships and relationships between measures and drivers.
- In particular, to assess the economic impact and potential impact on:
    - employment
    - property prices
    - accessibility to jobs and essential services
- To investigate the diversity and temporal and spatial barriers within each case study region.
- Considering evidence of an “opening date effect” by comparing impacts with those from other studies.

1.4.3 Hypotheses

In particular, the study will address the following hypotheses:

- Improvements in rail transport infrastructure to previously disconnected communities produce measurable net property value impacts.
- These impacts will be affected by the proximity of neighbourhoods to new stations.
- The re-establishment of a rail link between previously disconnected regions and a larger conurbation results in measurable net employment impacts.
- Even with new rail improvements, spatial, temporal and economic barriers may prevent more economically vulnerable neighbourhoods from receiving the full impact.

1.4.4 Academic contribution

In methodologies which are still emerging in this research field, new ways of assessment are continually being proposed. The main academic contributions of this study are to increase the knowledge about ex post appraisal methodology by applying and comparing previously published methods which either have not been considered together or in the context of sparsely populated or remote regions. In particular the contribution includes:

- Re-contextualisation of existing techniques and models i.e. applying those techniques in a new context and demonstrating the applicability of models in a new situation.
- Relating these to remote and sparsely populated areas subject to smaller impacts through identification of regional, property and employment characteristics.
- Development of a counterfactual through adapting methods such as Cluster Analysis and Propensity Matching to address selection into treatment issues.
- Analysing the efficacy of alternative accessibility measures.
- Developing a job accessibility index applicable to more remote regions as a standalone model or being incorporated into the hedonic models addressing remote regions in terms of:
    - The decay effect.
    - Skills mismatch for local residents.
    - Commuting feasibility and thresholds.
- Addressing spatial and temporal impacts and heterogeneity through Geographically Weighted Regression (GWR) to examine relationships at different points in space.
- Estimating the "opening date effect".

1.5 Overview of the thesis structure

This chapter has offered a brief overview of the background, context, rationale, overall structure and aims and objectives of the thesis. It has:

- Provided background information about transport trends, rail network changes and the state of disadvantaged communities.
- Outlined the development of interest in rail interventions, particularly in previously rail-disconnected remote regions.
- Indicated what this study sets out to achieve, and how.

The remaining chapters are organised as follows.

Chapter 2 describes the stakeholder survey and findings carried out prior to the opening of Borders Rail. Although not part of the eventual methodology, it was a key component of this study in fundamentally determining the direction of this research, and being instrumental in clarifying the required methodological approach.

Chapter 3 contextualises the study and potential methodology with reference to relevant literature, in particular assessing advantages and disadvantages of evaluation and methodological approaches.

Chapter 4 outlines the research methodology and methods applied. It describes and justifies the use of various approaches and reflects on the strengths and weaknesses of the methodology applied in this research. These include statistical modelling approaches, derivation of a job accessibility index and its application to modelling property prices and employment. It also considers the 'selection into treatment' problem and suggests a suitable approach using propensity testing.

Chapters 5 to 7 apply the methodology described with reference to three different case study regions, which represent different stages in the restructure process. The unique nature of each region is described and the output and findings for the individual region are highlighted. Statistically significant relationships are identified, with possible explanations for these discussed.

Chapter 8 then provides a synthesis in drawing together the core findings of the research methods as applied to each case study region. It compares the differences and similarities from each case study analysis, and discusses them in the context of the research questions. It offers a critical discussion of the findings and potential alternative explanations for its findings, as well as discussing the limitations. It builds upon, and is compared to other literature findings in relevant research fields.

Chapter 9 then summarises these into key findings and recommendations to take forward and assesses how the findings relate to and respond to the research questions, and contribute to the current state of knowledge on the topic of ex ante evaluation methodology. Figure 1-7 provides a schematic overview of the thesis.

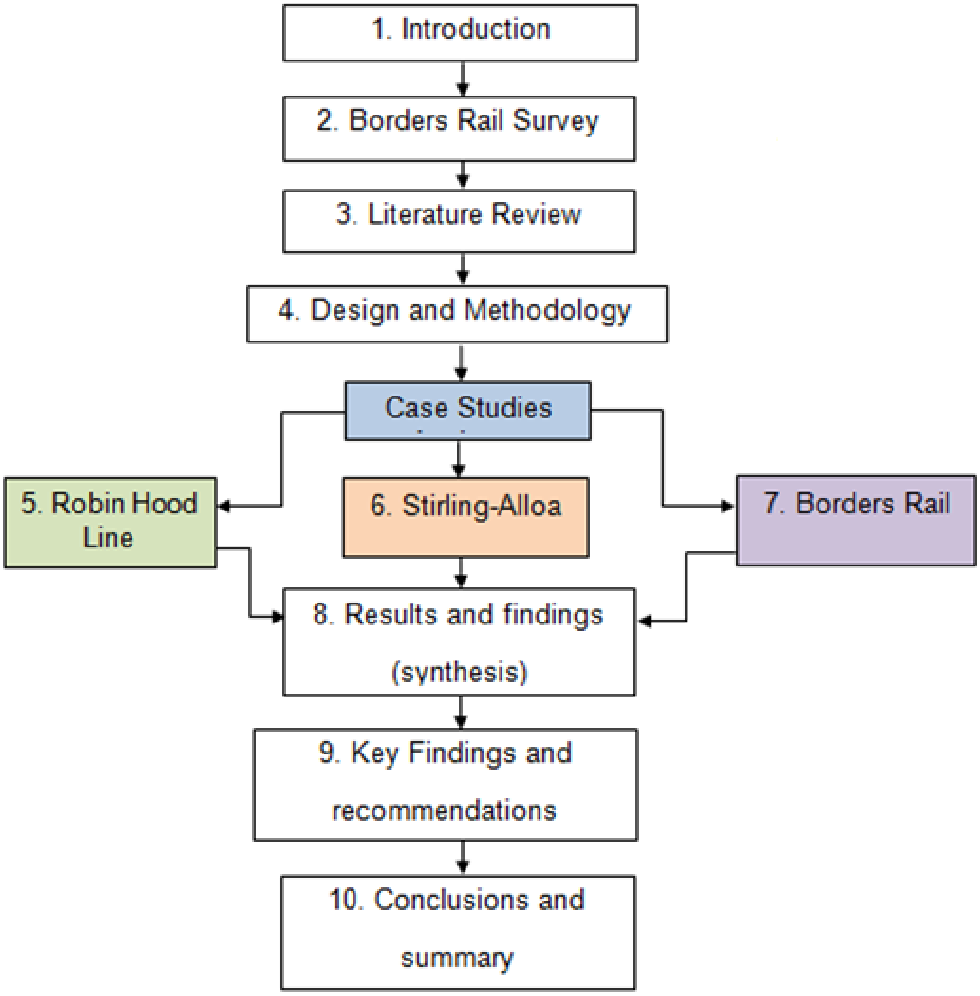


Figure 1-7 Schematic Overview of the Thesis

2 Chapter Two: Borders Rail Survey

2.1 Overview

The overall motivation for this research study was to evaluate the accessibility issues and wider impacts resulting from restoration of rail links in regions previously disconnected from the rail network. The original inspiration was the imminent reopening of Borders Rail which involved restoration of part of the former Waverley line to reconnect Edinburgh to the Scottish Borders at Tweedbank.

In order to establish the aims and objectives of this study, a survey was carried out of stakeholder groups in the Scottish Borders region prior to the opening of Borders Rail in September 2015. Although this survey was not an integral part of the eventual methodology, it was a key component of this study in fundamentally determining the direction of this research, and being instrumental in clarifying the required methodological approach.

The survey sought to gain some insight into the attitudes and views of various interest groups in the community. It was primarily a qualitative study canvassing current attitudes and opinions, but by making use of supplementary questions allowed other specific concerns to be raised. By focussing on stakeholder groups, the aim was to determine whether responses were defined by group interest and objectives, or reflected those of the whole population, and to distinguish any differences in attitude between organisations closer to the rail corridor from those further away.

Interviews were carried out with a sample of local stakeholders having a vested interest in the impact of the intervention, and selected from a cross-section of the community representing various interests - education, police, housing associations, charities, local interest and community groups. The interviews took place on location in the period leading up to and including the official opening. Where a face to face interview was not possible, a representative of the group completed an online questionnaire, although the limitations over personal contact were recognised.

A generic survey questionnaire - constructed using BOS online surveys - provided a template for the interviews and allowed for additional follow up questions. Each interview lasted between 45 minutes and one and a half hours, largely dependent on the availability of the respondent and what they felt able to contribute. The themes of the questionnaire included views on the current state of transport, attitudes to all forms of transport, perceived accessibility to essential services and shops, accessibility for the elderly and disabled and expectations of the new rail link in terms of access to

Edinburgh, tourism, jobs and education. The survey data was collated and analysed to estimate the overall view and variance of opinion, and to highlight any apparent difference in views expressed both between groups and also those based near the rail corridor and further away.

The limitations of this survey were its small sample size and time restrictions that did not allow a more representative spread over a wider range of interests. Nevertheless, it did provide a platform for assessing attitudes where group interest and objectives may have been the dominant factor, and also allowed exploration of other concerns that would not necessarily have been voiced.

2.2 Objectives of the survey

The key objectives of this survey were:

- To discover pre-existing attitudes to public transport, the car and walking and cycling.
- To observe how each group perceived the existing transport situation as regards accessibility to essential and leisure pursuits, including access for the elderly and disabled.
- To canvass expectations on the impact of the introduction of Borders Rail particularly in terms of jobs, access to Edinburgh, tourism and the effect on local retailers – the 'two-way' effect. Thus to assess the positive and negative benefits of Borders Rail to the local economy by estimating additional effects brought about by the new rail link.
- To distinguish attitudes and expectations by group interests, location and the section of the population served based on age, gender and disability.
- To distinguish attitudes and expectations by comparing the views of those based closer to the rail corridor to those further away.
- To allow for a revisiting of all the issues raised subsequent to the line re-opening for a before and after comparison.

2.3 Survey Design and implementation

2.3.1 Selection of stakeholder groups

For the purpose of this study, 'stakeholder groups' equated to any representative group potentially affected by the new rail link. There were two categories:

- Primary – those directly affected either positively or negatively e.g. business groups.
- Secondary – those indirectly affected either positively or negatively e.g. local interest groups.

Groups were selected on the basis of what bound them together (e.g. issues, identity, interaction, location) and on the following overall interests:

- Social change - improved social aspects e.g. activities.
- Environment - conservation of resources, environmental benefits.
- Physical health - healthy activity e.g. cycling and walking.
- Health - access to community health centres, adult day care etc.
- Safety and security - improved safety for the community as a whole.
- Economics - economic prospects for low-income people, for example.

2.3.2 Sample location

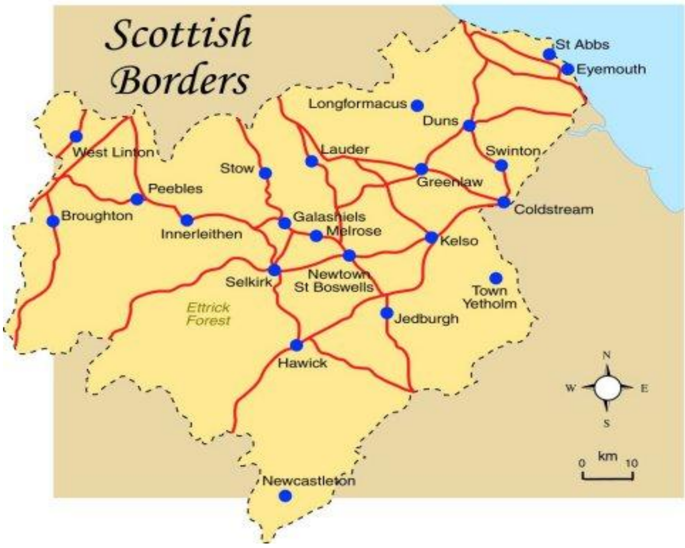


Figure 2-1 Map of Scottish Borders Region
(Source: The Gazetteer for Scotland: 2018)

The sampling strategy was designed to capture the views of stakeholder groups on the rail scheme by enlisting a cross-sectional range across the community based in different areas within the Borders region. A small sample of stakeholder groups was targeted both in the central Borders area (including Galashiels and Melrose), but also outside the proposed rail corridor (Peebles and Selkirk) to distinguish different responses based on relative proximity and ease of access to Edinburgh. Figure 2-1 is a map of the region highlighting the main towns.

Peebles and Selkirk are further away from the new Borders Rail link but have a similar profile to Galashiels, i.e. without a rail service having previously had one. Peebles is closer to Edinburgh but distant enough from the new rail link making it impractical as a feasible transport mode to either Edinburgh or Galashiels. Very little impact would be expected as preferred access to Edinburgh would still entail car or bus transport modes. In Selkirk, the rail link may have more impact potentially, being further away from Edinburgh, and use could be made of park-and-ride facilities.

2.3.3 Survey Construction

An online questionnaire was constructed using BOS online software consisting of a generic version applicable to all stakeholders, but also incorporating a special section for business groups with questions more pertinent to their interests. Responses were provided by one representative who spoke for the entire group. In line with the objectives, the questionnaire was constructed to:

- Canvass current attitudes and opinions on transport and those relating to the forthcoming transport infrastructure changes.
- Seek opinion on the positive and negative benefits that Borders Rail will bring to the local community through the eyes of its stakeholders.
- Allow other avenues and specific concerns to emerge, particularly through the use of supplementary questions in the interview mode.
- Allow for a revisiting of all issues after the opening of the line for a before and after comparison to be made.

It consisted of the following sections:

- The organisation - the type of organisation and interests, its location, and how long established.
- The current transport situation - modes of transport used by members, and views on the current state of transport in the region.
- Accessibility issues - perceptions of access to essential services and shopping and recreational facilities and current access provision for elderly disabled etc.
- Attitudes towards travel and transport - particularly public transport, private transport and walking and cycling.
- Forthcoming changes to the transport Infrastructure - awareness and likely impact of Borders Rail and the perceived difference e.g. travel to Edinburgh.

2.3.4 Survey Implementation

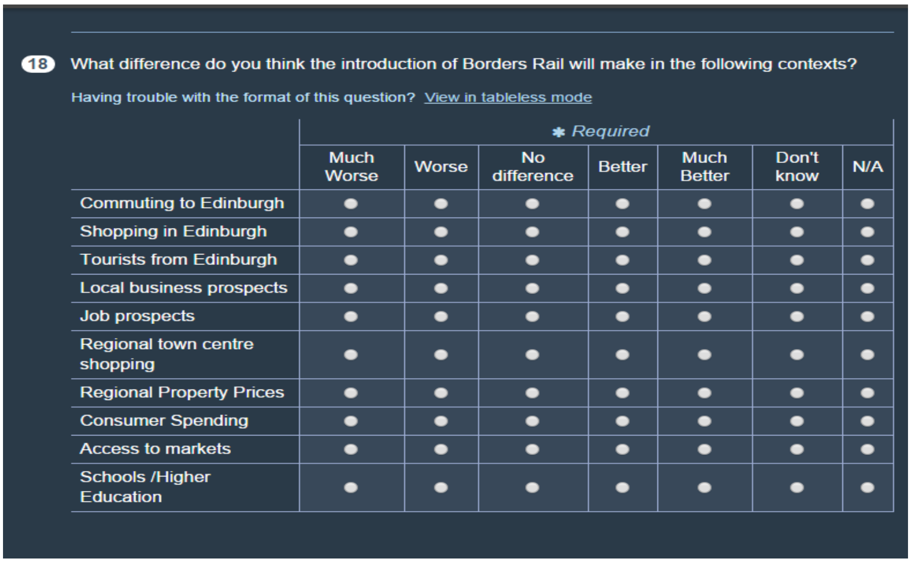
A pilot study was completed by mid-July 2015 where a provisional copy of the online questionnaire was sent to several volunteer respondents who provided feedback on its construction and wording. These suggestions were assimilated before a final version was published ready for implementation. The recruitment of recipients and fieldwork took place from mid-July to mid-August.

- A list of 50 suitable potential contacts was compiled to provide a representative sample drawn from a range of business, educational, community and local interest groups.
- Stakeholders were contacted individually by email and where necessary this was followed up by a telephone call to arrange a suitable date.
- The survey was administered either through:
    - A scripted face to face interview using a prepared questionnaire script. These were carried out on location during w/c 31st August 2015 and w/c 7th September 2015.
    - An online questionnaire (depending of the sample uptake and availability of contacts in arranging a face to face interview.) which went live on 31st August 2015.

Out of 50 organisations contacted, there were 10 face to face scripted interviews and 10 completed online questionnaires, a response rate of 40%. A list of respondents can be found in Appendix 10.8.1 with a key to the coding in Appendix 10.8.2. The following points were clarified before the interview began:

- Who we were
- Why we were carrying out the survey
- Why they were being consulted and how their responses would be used
- The information we would collect and how it would be used - information gathered would be treated and used responsibly and reported honestly
- How we would feed back to the community

The interviews and online questionnaire entry were concluded by late September 2015 followed by collation and analysis which was completed by mid-October 2015. The standard survey questionnaire consisted of a mixture of open and closed questions and can be found in Appendix 10.8.3 and a typical screen shot is illustrated in Figure 2-2.


18 What difference do you think the introduction of Borders Rail will make in the following contexts?

Having trouble with the format of this question? View in tableless mode

✱ Required

	Much Worse	Worse	No difference	Better	Much Better	Don't know	N/A
Commuting to Edinburgh	○	○	○	○	○	○	○
Shopping in Edinburgh	○	○	○	○	○	○	○
Tourists from Edinburgh	○	○	○	○	○	○	○
Local business prospects	○	○	○	○	○	○	○
Job prospects	○	○	○	○	○	○	○
Regional town centre shopping	○	○	○	○	○	○	○
Regional Property Prices	○	○	○	○	○	○	○
Consumer Spending	○	○	○	○	○	○	○
Access to markets	○	○	○	○	○	○	○
Schools /Higher Education	○	○	○	○	○	○	○

Figure 2-2 Screen shot - online questionnaire

2.3.5 Data Collection and Analysis

Data from the scripted interviews and online survey were amalgamated and all standard questionnaire responses transferred onto an Excel spreadsheet. For all supplementary and text-based information:

- There was an initial process of coding and thematic analysis.
- Responses were reviewed line by line and each emerging concept assigned a code.
- Identically coded sections were compared to check they represented the same concept.
- Through this iterative process, emerging themes were identified and interpreted.
- Data was analysed through the use of BOS and Excel, and statistical analysis of responses estimated the average (mean) view expressed on each question, and the variation (variance) in responses. The latter demonstrated the degree of unanimity or divergence expressed on each topic.
- In addition, because the survey aimed to contrast attitudes based on distance from the proposed rail corridor, the analysis also compared views on the current transport situation, accessibility through public transport and the expectations from the introduction of Borders Rail, broken down by group interest to detect any variation in opinion depending on the objectives of that group.

2.4 Survey Output

2.4.1 Demographic of stakeholder groups

Most groups had their headquarters within the Scottish Borders region, but the areas they serviced were not restricted to the region. They predominantly represented local residents, but also covered a range of activities across the community, distributed fairly evenly by age and gender, with the exception of young children.

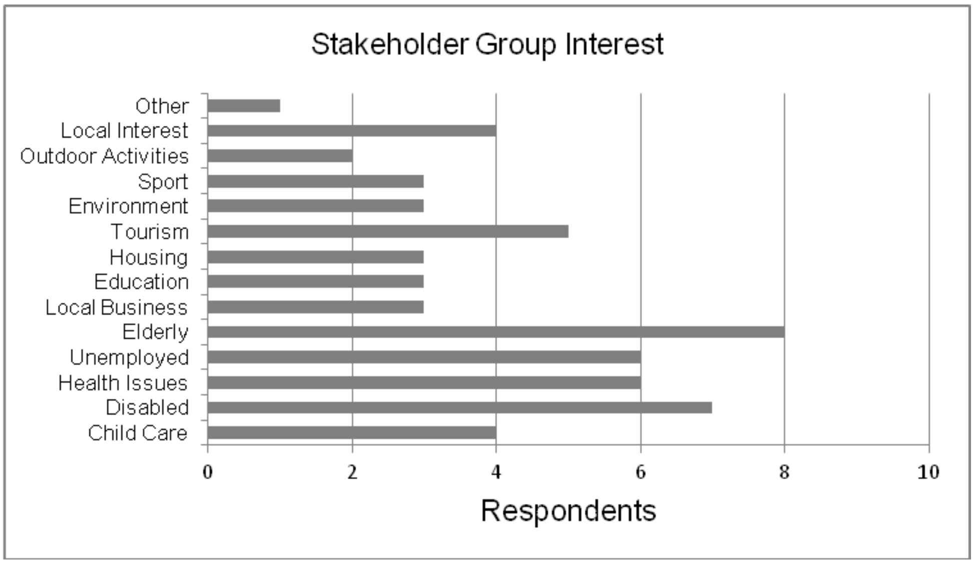


Figure 2-3 Stakeholder Interest

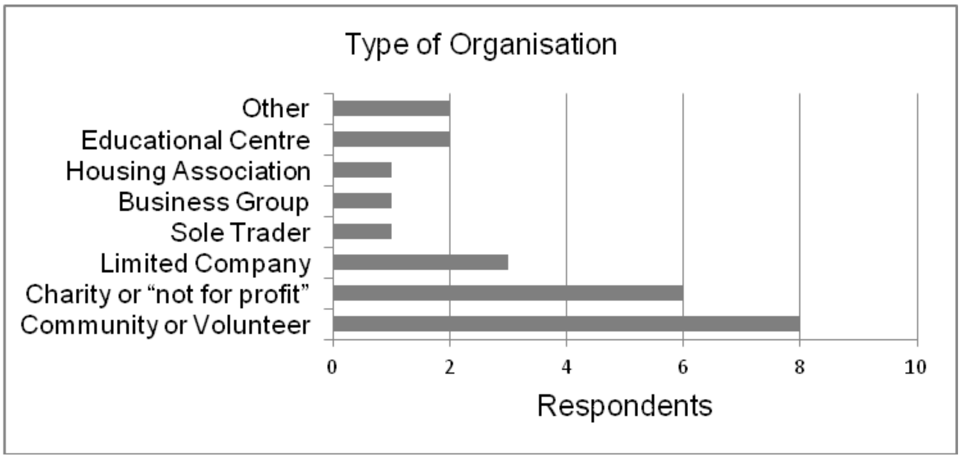


Figure 2-4 Type of organisation

The majority of groups questioned were either volunteer or charity groups with interest in the community, but for a wider spectrum of views there were contributions from the business, housing and educational sectors. Their sphere of interest covered a wide range, but there was good representation of those covering the elderly, unemployed and disabled.

2.4.2 Current attitudes to modes of transport in the region

A summary of the results from the survey can be found in Appendix 10.8.4 where for each question a ranking system of 1 to 5 was used with 1 representing the lowest and 5 the highest ranking. Despite contributing to environmental impact, there was a general dissatisfaction with public transport locally. Public transport was not satisfying, unreliable, offered a below average level of choice and provided a limited service with poor connectivity throughout the region. Despite negative effects on health, fitness and climate change, of the existing travel options in the region, the car was ranked more highly than public transport and seen as satisfying, affordable, reliable, convenient, time efficient and of good value (Figure 2-5), aided by a road network largely considered of good quality. Although car sharing had increased, it should be promoted more, and if buses and trains were less expensive and more available it would be feasible to restrict car use.

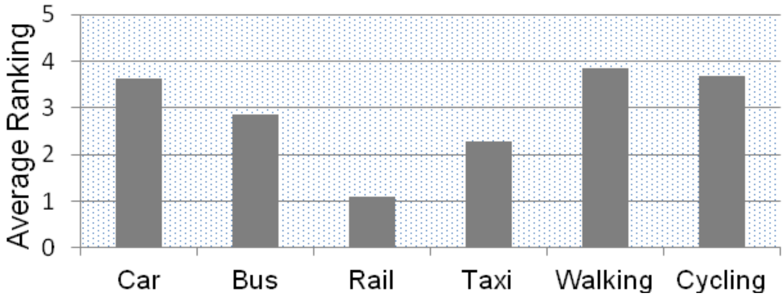


Figure 2-5 Assessment of current travel options in the region

Cycling and walking also ranked very highly, with good recreational facilities and information on walking and cycling routes. Although there was also excellent provision for walkers and cyclists through paths and cycle tracks, still more was needed. Cycle lanes needed better maintenance and wider availability, especially on roads such as the A7 and A68 where competing with heavy vehicles was difficult. Also there should be more facilities for carrying cycles on public transport which is currently subject to a two bike limit on trains.

In terms of distance from the rail corridor there were similar attitudes to public transport which was more highly rated in the town centres, but thought very limited in the smaller villages. There were variations by location, and public transport was viewed more favourably in Peebles than in Selkirk and Hawick. Although a satisfactory bus service operated North-South, East-West connections were poor. Figure 2-6 shows how the sample reflected different distance from the new rail corridor.

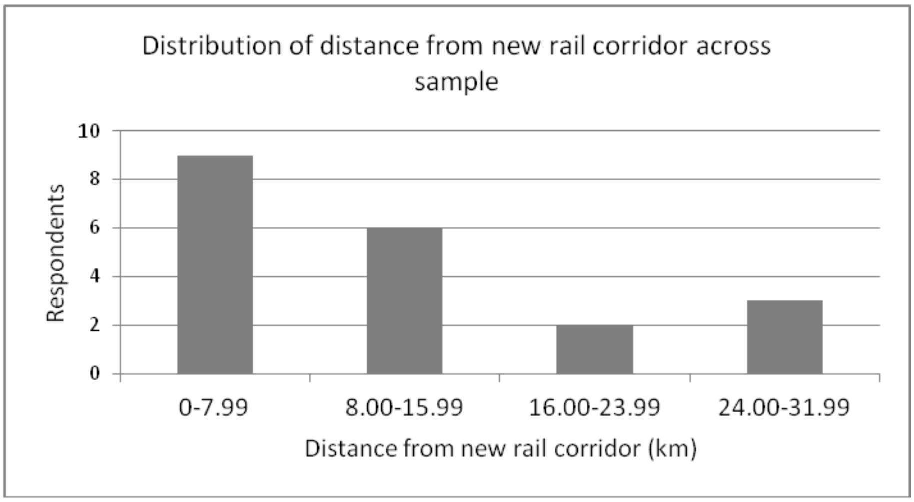


Figure 2-6 Distribution of distance from new rail corridor across sample

2.4.3 Current state of transport and accessibility concerns

There was a pressing need for progress both in transport connections and scheduling of services, which was essential for the new rail link to be more effective in improving accessibility both for essential services and recreational pursuits. There was no apparent joined up transport, and the travel experience differed greatly when travelling north/south rather than east/west. Buses were infrequent and offered a limited service to the villages and more remote areas, but were generally much better in serving the A68/A7 corridor and larger town centres. There was criticism of the level of choice and service frequency, and connections both to the hospital and throughout the region.

Generally, accessibility by public transport to local shops, supermarkets and other facilities was satisfactory, and with no difficulty in accessing leisure and shopping facilities using private transport. However, a recurring theme was the conflicting accessibility experience between towns and villages, particularly to most essential services with the exception of schools. There was limited bus service to villages and more remote areas and this service stopped very early and consequently cars and taxis were often the only option for an evening out, and in villages, fares were much more expensive.

Residents of towns nearer to Edinburgh like Peebles had no difficulty reaching the capital, but elsewhere access to Edinburgh involved a long bus ride. Some currently accessed Edinburgh for work or shopping, but for financial reasons this did not apply to the whole socio-demographic spread. The elderly used their bus pass to get to Edinburgh and may continue to do so after the rail link restoration, despite the longer journey, because of the relative costs of public transport.

Accessibility to essential services and shops depended largely on the place of residence, and was particularly good in Peebles where there is an excellent service to Edinburgh. Those living in towns could reach all services relatively easily, in contrast to those more distant from the main centres of population and satisfaction levels dropped in nearby villages.

> "Out of Peebles and Selkirk public transport is good but in small villages not so good."

> "Out of town it is again a matter of location, and if you are living in villages nearby the levels of satisfaction will go down by at least one notch."

At the southern part of the Borders in Selkirk and Hawick, access by public transport to local shops, supermarket and the towns was only satisfactory, and Hawick was too far away from Edinburgh to provide a viable commuting or shopping journey.

Figure 2-7 summarises information from Question 10 of the survey where respondents were asked to rate current accessibility on a scale of 1 (very difficult) to 5 (very easy). Apart from the hospital, accessibility to essential services in towns using public transport was thought satisfactory to good, as most services were based near to the various centres of population. Although, primary and secondary schools had special transport laid on, this was not true of nurseries which could be located in less accessible places. Access to the hospital was often difficult as there was just one regional hospital at St. Boswells. For example, those living in Hawick would need three buses to reach it, and this was true of some other services e.g. Hawick was also the nearest custody centre.

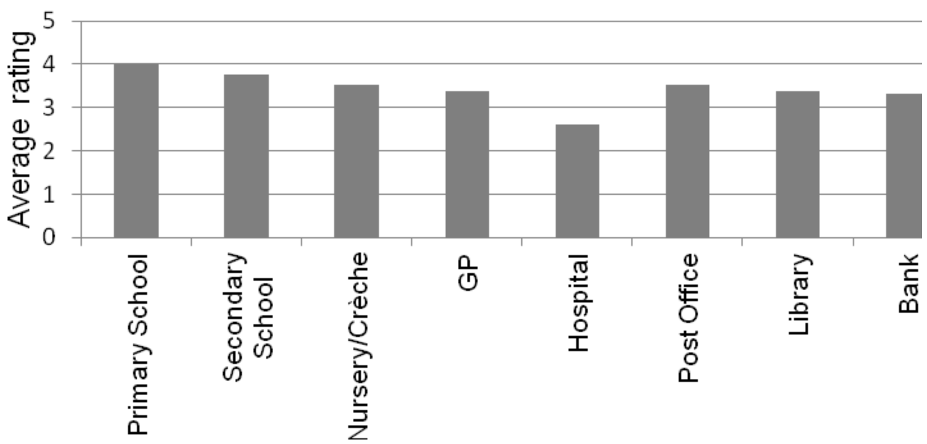


Figure 2-7 Accessibility to essential services by public transport

Although travel could involve long distances, accessibility by car was generally better than other modes as parking locally was either free or inexpensive, and there was no problem of access wherever you lived in the region. Again from Question 18 on the survey (Figure 2-8), respondents were asked to assess the expected impact of Borders Rail on a scale of 1 (much worse) to 5 (much better). There was general positivity about the potential impact of the new rail link providing a welcome boost for the region especially with the attendant publicity regarding the re-opening. All were well informed of the Borders Rail link which would benefit their organisation, especially if properly planned and other initiatives were instigated. The rail link would provide a good resource for the region, even were they or their group never to make use of it.

Figure 2-8 Expected impact of Borders Rail

2.4.4 Wider Economic Benefits

The feedback shows that there was general consensus that the new link would lead to an increase in property prices in the rail corridor, and house prices were expected to rise in Galashiels and Midlothian. It was commented that there would be a decay effect with distance, so that Peebles, Selkirk and Hawick would be largely unaffected.

> "Hawick is too far away from Edinburgh for it to be a viable commuting or shopping journey."

> "In Peebles people use the bus and with Park & Ride in Edinburgh handy and attractive there is no advantage in having access to rail for commuting and shopping"

> "House prices will change - along the rail line slowly at first - with a decay effect as one gets further away from the rail corridor"

Peebles was a more prosperous town than Galashiels, and currently house prices were already high in Peebles and depressed in Galashiels. There was also general agreement that more opportunities would now be accessible through Edinburgh, and there would be improvements in commuting and shopping with a boost to tourism coming from Edinburgh.

Although there was expected to be reduced benefit for areas distant from the rail corridor, they would gain through the wider benefits, and the intervention could potentially benefit all organisations over time. The exception would be the Peebles areas as spatially it would not be practical to use the rail link regularly, but in other areas it would have a very positive influence.

> "It will not make much difference in the Peebles areas as the rail will be too far away to be of use, but in other areas it could have a very positive influence."

Local business prospects would be better, particularly for those based in the central Borders or involved with the tourism sector. However, although more jobs would be available in Edinburgh and the Midlothian area and be more accessible, there were reservations over a significant difference to job prospects in the Scottish Borders. Town centre shopping was already depressed in Galashiels where even charity shops were closing and more pop-up shops were appearing. Although supermarkets had cornered the food market, customers would go to Edinburgh for luxury goods and clothes. There was also concern in Tweedbank being the terminus for the line as potential benefits may be limited as the station is unmanned with few commercial outlets nearby, whereas Melrose might have been a better option with its more extensive facilities and historic context.

Although the rail link should make little impact on schools, it may improve the situation in further and higher education by attracting students from outside of the region. Rail access would have limited impact on the current student population who travelled from all over the Borders region where there may still be no rail connection. The cost of rail travel may impose a constraint on those with limited incomes and locals might not pay for rail travel when bus passes were much cheaper.

It was thought there would be little impact on town centre shopping and consumer spending to areas outside the rail corridor, particularly towns like Hawick and those further south which, because of their location, look more towards Carlisle than Edinburgh. In Peebles, people utilise the bus and park and ride facilities in Edinburgh, which are both handy and attractive, and there is no advantage in access to rail for commuting and shopping. An extension of the railway to Peebles would be very welcome, but this may not be feasible given that the upper Tweed railway track is now dedicated to multi-use paths. Depending on how far away the centres were from

Galashiels and the rail corridor there would be potential for commuting to Edinburgh and access to the influx of tourists from the capital.

2.5 Summary of Findings

The findings highlighted the issues that most concerned the various stakeholder groups in the region, and pointed to themes which would be explored further through research into various secondary source datasets. These included issues of accessibility, social inclusion and wider economic benefits, and how they can be measured effectively with the possibility of transferability to other situations subjected to rail infrastructure changes.

The Borders region had suffered more than most rural locations through disconnection from the rail network, and the consensus was that all forms of public transport needed to improve to boost overall standards in the region. A continuing complaint was the wide variation in local public transport levels of service, with accessibility to services and shopping facilities depending greatly on where you lived. Those residing outside the main town centres suffered much more from social exclusion and general accessibility, relying heavily on the car for transport. This affected all groups and applied both from the perspective of the central Borders and the more outlying areas. Accessibility was most critical involving public transport to the hospital, which was very inaccessible for those not living in the central Borders area, often requiring several buses unless private transport was available.

The car was the most popular means of transport, and although there was an increase in car sharing, it should receive more promotion. There was general dissatisfaction with public transport which offered very limited choice and level of service, and a great variation in travel experience along the North-South and East-West corridors. Buses were infrequent and offered a reduced service to villages and more remote areas, which restricted travel options in the evening. There were different accessibility experiences between the towns and villages, particularly in reaching essential services. There was also restricted connectivity throughout the region, especially in accessing the hospital, and an urgent need for “joined up” transport with better scheduling of connections.

In terms of social inclusion for the disadvantaged, elderly people could utilise their bus pass to travel to Edinburgh and may continue to do so even after the rail link was in place because of the relative cost of transport balanced against the travel time factor. Easier commuting via the rail link should provide improved access to job opportunities and greater access to shopping in Edinburgh. However, train travel costs may impose a barrier on those without the ability to pay, as limited finances may constrain those seeking employment from accessing Edinburgh and Midlothian.

Expectations for the impact of Borders Rail were very positive overall, even outside the immediate rail corridor, and the restoration of the rail link would add a much needed extra dimension to regional public transport. However, the opportunities that the rail link brought should be accompanied by promoting other activities to benefit the region, as an improvement to towns like Galashiels would still need additional initiatives to boost the local economy. Local business prospects were expected to improve, particularly in the Central Borders and through the tourism sector. However, there were doubts about the impact on local job prospects and town centre shopping. The rail link should make no difference to schools, but improve the situation in attracting students from outside of the region into further and higher education. There was also a general consensus that there would be an increase in property prices in the rail corridor and house prices were expected to rise in Galashiels and Midlothian.

It was thought that benefits would be a reduced or those further away from the rail corridor, and that centres such as Peebles, Selkirk and Hawick would be little affected. In particular, for the Peebles area the rail link was thought too far away to be practical, whereas in other areas it may still have a positive influence. Outside of the rail corridor little impact was expected on town centre shopping and consumer spending, and the link would provide little benefit to towns further south like Hawick which are not Edinburgh-centric and look more to Carlisle than Edinburgh.

The Borders was a great area for cycling and walking which were very popular, and could provide a direction for revitalising the area in conjunction with the rail link and an improved bus service to the more remote locations. Although there was better provision for cycles, there should be an expansion of cycle lanes and limited cycle carriage on public transport could prove an obstacle.

Although there was a general similarity in attitude across all groups, there were differences when comparing those closer to the proposed rail corridor to those further away. The wider benefits would be mainly confined to the central Borders region and Midlothian, although some peripheral benefits might filter to the fringes of the region. The link to Edinburgh should bring more prosperity through increased tourism, and improve access to jobs in the capital through better commuting times, which may also translate into business opportunities and jobs locally.

2.6 Limitations of this approach

The main limitations of this approach were the relatively small sample which targeted only certain stakeholder groups, generally geographically based in one or two centres of population (Appendix 10.8.1). It also tacitly assumed responses from one person were representative of the group as a whole. When identifying community stakeholders it was important to be aware that they represent sections of the community. It should be evident that working with only established community leaders is not always an adequate substitute for engaging the wider community. It also pre-empted several questions as to whether the range of views forthcoming from the representative of the group reflect the interests of that group, and do certain groups e.g. disability hold similar views throughout? Use of a generic questionnaire offered a simplified approach to the “two-way” effect as in some cases feedback may be given by those without special knowledge of the local business scene, and sought to determine the reasons for these expectations, some involving the rail link and access to Edinburgh.

2.7 Issues taken forward for the research study

By highlighting the issues that most concerned stakeholder groups in the region, the survey findings supported the formulation of aims and objectives for the research study by providing a suitable context and direction for the methodological approach. These issues (Table 2-1) were particularly applicable in the context of remote or disconnected regions, and included accessibility to jobs together with wider economic benefits. In particular, the following avenues were worth exploring further based on the accompanying relevant respondent observations:

Socio-demographics and employment

- Analysis of regional socio-demographics to estimate those most affected.

"Borders region has been neglected - low wages, more elderly people and poor transport but nobody complains!"

- Current and projected employment levels in the region.

"There would be no difference to job prospects locally, but there would be more jobs available in Edinburgh"

"Central Borders and the biggest impact will be on tourism will be most affected"

Transport modes

- Usage of the rail and bus to and from Edinburgh and purpose of travel.

"Access to Edinburgh for shopping is poor to satisfactory because of the long bus ride."

"Locally ASDA and Tesco corner the food market, but customers will go to Edinburgh for luxury goods and clothes."

- Monitoring the level of transport mode usage in the region

"Bus and taxi access is good in main centres but poor to satisfactory outside"

Accessibility

- Analysis of bus timetables and frequency of service throughout the region particularly looking at out of town destinations and finishing times.

"Bus service is poor in villages but good on the A68/A7 corridor and in larger towns"

"Buses present the biggest problem. Service is satisfactory if you live in towns, but not as good as it should be in the villages. "

"Service to the villages stops very early in the evenings forcing people into cars and taxis for their nights out in town."

- Analysis of distance travelled to essential services and employment from various parts of the region and the means of getting there and back.

"Accessibility depends greatly where people live and most in Galashiels could access all services reasonably easily, but it is a different situation as you move further away."

"Centralisation of services at St. Boswells means that from Hawick you need three buses to get to hospital"

- Calculation of the transport costs and times involved in reaching services.

"With the bus pass being cheaper will people would pay out £10/£15 for rail travel."

"There is a good Bus service between Duns and Edinburgh but not East/West e.g. to go to Melrose from Earlston must go to Galashiels and then backtrack"

- Analysis of transport links to assess connectivity.

"Joined up transport required - need for better connections and scheduling

"The links are poor and there is no joined up transport; there is a difference travelling North/South compared to East/West."

Housing

- Analysis of regional property price movements to identify those areas experiencing a rise and attributing causality both to the rail link and other factors.

"Peebles is more prosperous than Galashiels and house prices are already high in Peebles and depressed in Galashiels, but Midlothian house prices should rise"

"Property prices will not alter as it needs more than a train line to make a difference."

Spatiality

- Comparison between the various parts of the region e.g. Scottish Borders.

"There are variations in public transport by location which is favourable to the Peebles area but unfavourable to areas like Selkirk, Hawick and Earlston."

- Analysis of effective distance from the rail corridor.

"There would be more impact if the train continued to Hawick"

"House prices will change - along the rail line slow at first - with a decay effect as one gets further away from the rail corridor"

Table 2-1 Issues taken forward from the Borders stakeholder survey

Accessibility to jobs and services	Wider economic benefits
Socio-demographics of the region to estimate those most affected	Movement in employment levels
Analysis of bus timetables and frequency of service	Spatial comparison across region
Analysis of transport links and routes	Monitoring of rail usage from main conurbation and purpose of travel
Analysis of distances travelled to the essential services and jobs and the means of getting there and back	Analysis of distance from rail corridor
Calculation of the transport costs and travel times involved	Analysis of property prices across the region
Transport mode usage	Travel mode transfer from car to public transport

3 Chapter Three: Literature review

3.1 Introduction

This review supports the aims of this research in developing a suitable methodology for the evaluation of the wider economic and accessibility benefits, covering literature of particular relevance to the main study objectives. The focus is on ex post evaluation of key outcomes of accessibility to jobs and services, property price and employment, and methodological approaches to evaluation and the establishment of causal effect of the rail intervention. In reviewing the extensive literature covering ex post evaluation of transport infrastructure impacts, it has been noticeable that the primary focus has been on city and urban rather than rural environments.

Having defined the various concepts with reference both to current guidance and recent literature, through a critical investigation of the success and worthiness of previous evaluation studies, and taking into account a selection of suitable methods that may be utilised, an appropriate approach to evaluation has emerged and is detailed in the conclusion.

The current situation as regards the background to disconnected and isolated communities and the feasibility of rail investment has been previously described in the thesis introduction chapter and so this review goes on to cover:

- The nature of wider economic impacts: the impacts of transport interventions on improved accessibility and the value of property and employment opportunities.
- Evaluation design - techniques and guidance: an overview of evaluation techniques comparing previously published methods. Various evaluation approaches are highlighted through reference both to current guidance and recent literature relating specifically to ex ante (based on forecasts rather than actual results) and ex post (based on actual results rather than forecasts).
- Previous evaluation studies: a review of previous evaluation studies, particularly those addressing rail interventions, towards more robust ex post evaluation.
- Methodological approaches: the development of evaluation methodology involving suitable measures for accessibility to jobs and essential services, and indicators of wider economic impacts on employment and housing. Looking at methods that have been applied, along with others potentially suitable to incorporate into an appropriate methodological structure addressing the objectives of this study, including methodologies which are still emerging in this field, and new ways of assessment.

- Spatial effects and accessibility: investigation of the evaluation of spatial, temporal and economic barriers in determining cause-effect relationships, and how they differ at different points in space and time. A review of existing statistical methods of relevance in the context of this study, particularly those for assessing regional heterogeneity and diversity, and how these benefits may be affected by the proximity of new stations and timing of an intervention.

3.2 The nature of wider economic impacts

Banister and Berechman (2000) suggested that transport investment may afford both transport and non-transport benefits. Transport benefits were generally measured in terms of user benefits (e.g. reductions in travel time, increases in accessibility), but there is no clear methodology available to measure non-transport benefits, which can be experienced at macro-economic output level (e.g. productivity increases), at meso-level (e.g. through employment changes), and at local level (e.g. through land value and property market effects).

A key concern of this study is the wider economic impacts of transport interventions brought about by improved accessibility, and how they may affect the local economy in terms of property values and employment opportunities. An initial definition of wider benefits and their application in remote and rural environments, leads on to the linkages with accessibility, employment and housing.

Wider Economic Benefits (WEB) are defined as effects on the economy caused by the existence of market imperfections in transport-using industries. For Vickerman (2007) these were all economic benefits not captured in the direct user benefits of the type which are normally analysed in a well-constructed transport cost benefit analysis after allowing for environmental and other directly imposed external costs.

CBA (Cost Benefit Analysis) is a systematic process comparing the expected benefits of a project with its expected costs. In CBA, as long as the markets are perfectly competitive, the user benefit equals the total benefit of the investment (Jara-Diaz, 1986), and adding spill over effects leads to double counting (Mohring, 1993). (Perfect competition involves a large number of small firms, identical products sold by all firms, freedom of entry in and out of the industry, and perfect knowledge of prices and technology.)

However, increased attention has been given to market imperfections where the full benefits of a transport investment may not be captured by CBA (SACTRA, 1999). Independently of the market structure in the transport sector, market imperfections produce utility effects which do not cancel out and may therefore cause an under-estimation of the user benefit of a project (Venables and Gasiorek, 1998; SACTRA, 1999). Laird et al. (2014) found that, although CBA provides a robust framework, methods must be developed to capture wider economy impacts when dealing with transformational projects. Several case studies (Venables, 2007; Vickerman, 2008; Graham and Dender, 2011) have tried to estimate wider economic benefits of transport schemes concluding that there may be some additionality to the direct benefits, but these are contextual and therefore not relevant to all transport projects.

Laird et al. (2005) suggested that the '"state of the art" was limited on links between transport and the wider economy, and Banister and Thurstain-Goodwin (2011) found that traditional evaluation methods had not been very successful in accounting for non-transport benefits, arguing that different impacts could be measured at different levels relating more to labour market effects or land and property market effects.

Vickerman (2007) concluded that in seeking a robust method for measuring economic benefits it was necessary to clarify objectives, as different methodologies may produce very diverse answers not necessarily reflecting inconsistency in results, but rather incompatibility in method. A substantial literature on the economic and social effects of transport infrastructure investments adopted a variety of scientific methods (Mohring Jr. and Williamson, 1969; SACTRA, 1999; Oosterhaven and Elhorst, 2003; Vickerman, 2007; Lakshmanan, 2010; Tavasszy et al., 2002). There was general consensus that spatially detailed models were best suited to evaluate the ripple effects in the economy of large infrastructure investments, and the most satisfactory way of modelling the economic impacts of new transport infrastructure on the economic structures in a given region (Knaap et al. 2001; Oosterhaven and Knaap 2003; Tavasszy et al. 2002; Simmonds and Feldman 2011)

3.3 How transport improvements may contribute to the economy

The level of transport investment together with the amount of expenditure on transport operations can have wider effects on the economy resulting in reduced household expenditure on other goods and services. In 2005, the UK Department for Transport (DfT, 2005) detailed how to estimate four wider economic benefits: agglomeration economies, increased competition due to increased transport, increased output in imperfectly competitive markets, and welfare benefits from improved labour supply. It dismissed the second of these impacts as likely to be negligible but found the others likely to be significant, especially agglomeration economies.

Accessibility

The direct effects of transport investment are to reduce transport time and costs through reducing travel times, decreasing the operating costs of transport and enhancing access to destinations within the network. Transport investment may also mitigate any economic disadvantages, for example where projects reduce congestion. These incremental benefits of transport investments may be measured through conventional cost-benefit analysis.

Agglomeration and Productivity

Other indirect consequences of transport investments include effects on productivity and the spatial pattern of economic development. Eventually, transport investments contribute to economic development by stimulating a variety of inter-connected economy-wide processes, yielding spatial and regional effects that augment overall productivity. Agglomeration benefits, as set out in Graham (2006), arise when firms and people locate near one another together in cities and industrial clusters Agglomeration is assumed to stem from greater business interaction, more efficient labour market interaction and more efficient input and output markets due to reduced freight costs. When caused by improved transport infrastructure, agglomeration can be argued on the basis that the infrastructure increases accessibility of an area thereby making it available for a greater number of employees and businesses, increasing the economic concentration in the affected areas. This can lead to greater productivity as costs of production decrease so more is produced, more cheaply, and also market share and competitiveness increase. To calculate the agglomeration effect of a reduction in distance, such as travel time, it is possible to use the difference in effective density and the elasticity. The latter has been the most influential in the development of guidelines for calculating agglomeration effects by practitioners e.g. WebTAG.

Employment

There is improved access to better paid jobs and a larger labour pool for employers. The consequence of people getting better jobs may be reduced benefits, increased tax take, and higher wages. Job creation is often held up as a major impact of transport investment as better transport may make it easier for people to get to work. On the supply side, individuals' labour force participation decisions are based on comparing the costs of working (including commuting costs), against the wages earned from a job. By reducing the cost in time and money of getting to work, a transport investment is likely to increase the returns to working. On the demand side, induced investment can generate new employment opportunities. If new jobs are created in one place then the value of output produced by each new job is the wage, and this is set against the value of what workers would have done in the absence of the jobs created.

Property

If house prices rise, then the wealth effect is likely to cause an increase in consumer spending. This will cause higher aggregate demand, and it is likely to cause an increase in real GDP and a higher rate of economic growth. Rising house prices, generally encourage consumer spending and lead to higher economic growth. A sharp drop in house prices adversely affects consumer confidence, construction and leads to lower economic growth (falling house prices can contribute to economic recession). Rising house prices can also redistribute wealth within an economy – increasing the wealth of homeowners (primarily older people), but reducing effective living standards for those who do not own a house (often the young)

Land value uplift measures the difference between the price of land in its new and former uses and represents the private gain to land owners. It provides a convenient way of estimating the economic value of a development which is dependent on a transport intervention. It should only ever be used in the appraisals of dependent developments. Land value uplift will capture any impacts which are capitalised into land values. It could potentially capture any of the following impacts: user benefits, land market distortions and other wider economic impacts, such as agglomeration economies that occur within that development.

3.4 The particular case of rural and rural/urban locations

Although most literature on wider benefits has had an urban focus, there is more recent recognition that rural areas should be considered differently. In isolated locations, poor transport infrastructure provides a major obstacle to significant increases in economic activity, where insufficient transport capacity and accessibility may limit prospects for employment. Also with limited job opportunities, labour markets are often viewed as thin (Findeis and Jenson, 1998; Vera-Toscano et al., 2004), the ultimate “thin” case being of only one employer. Job search costs are often the cause of market failure in remote labour markets (Laird and Mackie, 2009).

Laird et al. (2013) specifically focused on extending methods of measuring wider economic benefits to remote rural areas which often experienced a lack of choices for travel employment and suppliers. In identifying the importance of wider economic benefits for transport schemes in such areas, they considered the “two-way road effect”, where as well as improving the accessibility from the region to its market, there is also better access from the economic core of the country to the region. In a study of four road projects in rural areas in Scotland, Laird and Mackie (2014) concluded that wider economic benefits are relevant to transport project appraisal, and their omission could potentially bias the appraisal.

3.5 Transport and Social Exclusion

Social exclusion occurs when individuals or areas suffer from a combination of related problems such as unemployment, limited skills, low incomes, poor housing and access to education, high crime environments, bad health and family breakdown (OECD, 2002). Social exclusion and distributional effects have achieved increasing recognition within the transport literature, stressing the broader consequences associated with the inability to access or participate in key activities (Church et al., 2000), where lack of transport has been identified as a contributory factor in the social exclusion of disadvantaged and vulnerable population groups (Social Exclusion Unit, 2003) but does not always result in social exclusion (Currie and Stanley, 2010).

There is limited quantitative primary research on the role of travel in social exclusion. “Little theoretical work has been undertaken on social capital and transport, apart from the recognition that it does play a role" (Stanley and Vella-Brodick, 2009, p.90). Traditional transport appraisal methods do not sufficiently capture the social dimensions of accessibility over particular areas for specific population groups (Lucas, et al., 2015). Markovich and Lucas (2012) suggested that the social impacts of transport largely comprise negative impacts, with the majority reflecting disadvantages affecting the most socially excluded members of society. There should more emphasis on analysing the full range of distributional effects for each impact and on the longer-term temporal implications of transport-related social impacts.

3.6 Transport and Employment

Improvements in transport infrastructures may reduce commuting costs and bring increasing employment and lower wages (Gibbons and Machin, 2003). Venables (2007) described how this labour market effect interacts with the productivity increase; employees decided where to live based on the cost of commuting and the trade-off that wages were higher in cities. UK guidelines for calculating wider economic impacts WEBTAG (Department for Transport, 2012) provide a methodology to estimate the effects of transport improvements on labour supply where the additional value of increased labour supply to the economy and the resulting tax revenue is calculated using a general elasticity of labour supply and generalised transport costs.

There is some evidence of the importance of context. For instance, in rural or more remote regions there may be heterogeneity between areas requiring an area specific elasticity of labour supply. Here, a transport improvement changes conditions for commuters, enabling easier access to places with jobs (Gibbons and Machin, 2005) but local labour markets may be non-competitive and wage effects may differ between low skilled and high skilled workers. In addition, the housing supply may be inelastic with property price effects depending on the type of property available.

However, although the labour supply effect may be dominant in the case of rail projects due to improvements in accessibility, there may be more to employment effects such as the labour demand dimension. Employment also increases due to economic growth and if transport costs reduce, firms may produce more i.e. employment can also be seen as an economic indicator like GVA.

3.7 Transport and Housing

Previously, research had concentrated mainly on traditional transport impacts through demand pattern changes and travel mode switches, but improved quality of data from HM Land Registry and similar sources, has provided opportunities to address the land and property value uplift effects of transport, particularly in the residential sector.

Residential property is traded infrequently and by nature is very heterogeneous with considerable price variations between different areas. In normal market conditions there are distinct challenges in tracking transaction prices for comparable residences, and this is even more difficult during times of depression. The housing market experiences deep economic cycles, and prices have crashed twice in the last 30 years. The 1990s housing market recession lasted seven years, and there is some evidence that this had lasting impacts on the worst affected neighbourhoods (Forrest et al., 1997). Overall, poorer areas in wealthier regions suffered most, whereas poor areas in poorer regions were least affected (Dorling, 1993).

The development of transport infrastructure causes changes in accessibility and has economic effects enjoyed by those located at its proximity, measured in particular by its effect on property prices. Over much of the past 25 years, the cycles of house price and consumption growth have been closely synchronised (Attanasio et al., 2011) as an increase in house prices raises households' wealth and desired level of expenditure, and house price growth increases the collateral available to homeowners, thus facilitating higher consumption.

Changes in transport infrastructure typically improved accessibility effecting an increase in residential property prices. These raised property prices reflected the increased desirability of an area due to better connectivity and accessibility and this translated into economic benefits. Commuters were thought to respond to transport improvements by moving to areas more distant from their work in search of lower house prices or improved quality of life. There was a resultant increase in house prices and land values at more distant locations, subsequently impacting on housing supply and construction activity.

The classical argument on the relationship between transport and land-use benefits concluded that "changes in land values as may result from transportation improvements involve transfer of income among members of the population, not additional benefits (or losses) that must in some fashion be added to those arising directly from the improvement" (Mohring, 1976, pp. 119).

3.8 Railway stations and property value

Rail investment provided a practical solution to the rising congestion from vehicular traffic and urban sprawl, thus sustaining a more compact urban structure (Goldberg 1981). Studies in property values mainly consider three groups of determinants: physical characteristics, accessibility and environmental amenities (Debrezion et al., 2007; Banister and Thurstain-Goodwin., 2005).

Physical characteristics were identified as inherent factors representing the qualitative and quantitative features of property. Railway stations possessed two impact characteristics: accessibility and environmental amenity (the externalities associated with the neighbourhood), however, earlier studies mostly considered accessibility solely. Whereas the impact of railway stations on property value diminished with distance from the station, negative environmental externalities could affect properties along the railway line (Debrezion et al., 2007).

3.9 Evaluation techniques and guidance

Following on from consideration of wider impacts, the next step was to investigate evaluation through reference to recommended and applied methodology and techniques. In highlighting relevant evaluation approaches, there would be reference both to current guidance and recent literature, examining advantages and limitations in the context of this study. Consideration of the wider economic and accessibility benefits of a transport intervention can be found in various guidance documents addressing project appraisal and evaluation.

3.9.1 The Green Book and Magenta Book

The Green Book (HM Treasury, 2018) provides a broad framework for evaluation, detailing its application in particular contexts. Evaluation comprises a robust analysis conducted like an appraisal applying almost identical procedures. It should incorporate a quantified comparative assessment through one or more counterfactuals e.g. a ‘control group’ to whom the activity did not apply, to compare what would have happened had the intervention not been implemented.

The Magenta Book (HM Treasury, 2011) focuses on the design and management of evaluations and highlights the practical issues to be considered, providing guidance on data collection and case studies. Importantly, it questions the impact of an intervention through specific outcomes for different groups of people, comparing what was expected at the outset with what would have happened in the absence of the scheme (counterfactual). A good evaluation is not possible if the estimated counterfactual is unreliable, making it uncertain whether the outcomes would have happened anyway. It suggests that good impact evaluation recognises that most outcomes are affected by a range of factors, and not solely the intervention, and will successfully isolate the effect of the intervention from all other potential influences

3.9.2 WebTAG

WebTAG contains the Department for Transport’s guidance on the conduct of transport studies and provides guideline documents (e.g. TAG Unit A2.1 Wider Economic Impacts) which offer appropriate information on the role of transport modelling and appraisal and identify the likely significance of specific wider impacts for particular transport proposals (Figure 3-1) . The Department’s appraisal process is based on the principles of the HM Treasury Green Book guidance.

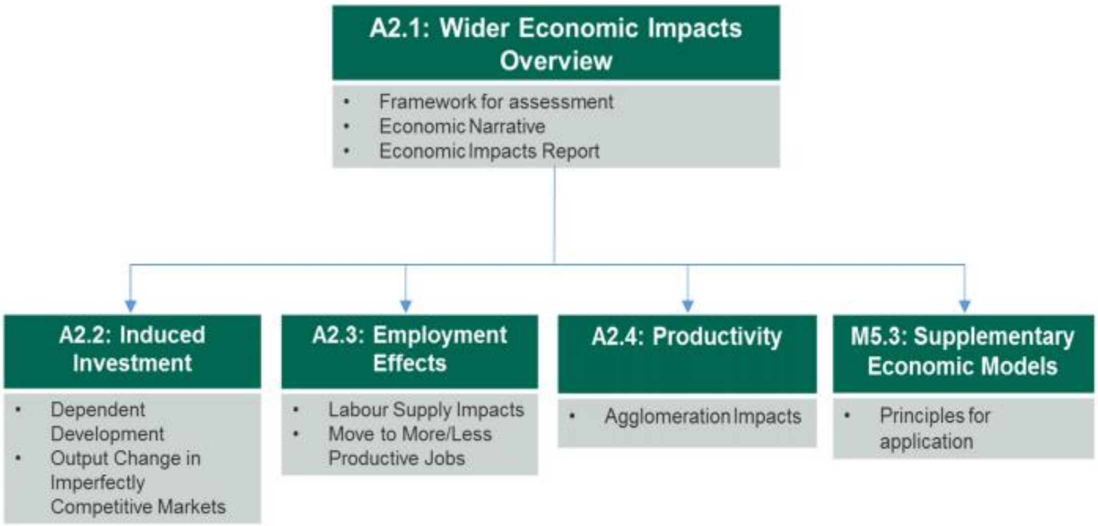


Figure 3-1 WebTAG units

Cost benefit analysis is the preferred approach because it captures a broad range of impacts, such as economic, environmental and social, thereby demonstrating the effect of a transport investment on welfare. Units A1 and A2 provide methods to estimate social welfare benefits associated with boosting the economy. The central estimates for a scheme's impacts on GDP and social welfare can be estimated by summing those impacts in the relevant columns, excluding benefits associated with dependent developments to avoid double-counting. However, this table doesn't include social and environmental impacts which may contribute to both welfare and GDP.

The development of methods to evaluate wider impacts is fairly recent, and there is some conjecture over use of estimates. Indeed, as indicated in *Value for Money Framework* DfT (2015), no wider impacts are part of the initial BCR [1](Benefit Cost Ratio), and some are also not part of the wider adjusted BCR. This is detailed in Box 4.4 of that document (Figure 3-2). These include concerns of this study, accessibility and access to services which are categorised as non-monetised impacts, and property value which is not included at all as mentioned previously.

[1] BCR indicates how much benefit is obtained for each unit of cost

Box 4.4: Typical impacts of a transport proposal

Established Monetised Impacts	Evolving Monetised Impacts	Indicative Monetised Impacts	Non-monetised Impacts
Included in initial and adjusted metrics	*Included in adjusted metric*	*Considered after metric using switching values approach*	
Journey time savings Vehicle operating costs Accidents Physical activity Journey quality Noise Air quality Greenhouse gases Indirect tax	Reliability Static clustering Output in imperfectly competitive markets Labour supply	Moves to more/less productive jobs Dynamic clustering Induced investment Supplementary Economy Modelling*	Security Severance Accessibility Townscape Historic environment Landscape** Biodiversity Water environment Affordability Access to services Option and non-use values

**These are a class of models rather than a specific economic impact*

*** A widely-used methodology for monetisation exists, but this is not included in WebTAG guidance because of concerns about its robustness. Detailed guidance is found in the Supplementary Guidance on Landscape.*

Figure 3-2 Typical impacts of a transport proposal

It is argued that network benefits and environmental impacts are not well represented in WebTAG (Campaign for Better Transport, 2012) and the method of assessment may lead to a bias towards road schemes rather than rail. They argue that If rural roads were assessed on the same basis as rural railways, most of them would be closed too, and the important factor is how a scheme supports transport, planning and economic development strategies; many re-opening schemes support local economic development and so score well in this.

3.9.3 STAG

Scottish Transport Appraisal Guidance (STAG) is based on WebTAG, but incorporates elements more relevant to Scotland's remote and unique topography (Transport Scotland, 2008). In the context of the STAG Appraisal, the Economic Activity and Local Impacts (EALI) measures and Wider Economic Benefit (WEB) measures, where appropriate, need to be considered within the Economy criteria, as well as the impacts of the project in terms of Safety, Integration, Accessibility and Social Inclusion and the

Environment. At present these impacts are not monetised and must be considered separately.

STAG has been modified recently following several post-implementation appraisals of rail interventions conducted by Transport Scotland. Using feedback from the recent pilot ex post evaluation studies at Laurencekirk (Canning et al., 2015), Airdrie (Glen, 2010) and Larkhall (SYSTRA, 2015). Transport Scotland produced a Guidance for the evaluation of rail projects which recommended implementing a comprehensive evaluation when a project has had sufficient time to establish itself, typically three to five years after project completion, but preferably 5-10 years after completion to include an in-depth assessment of wider economic benefits and accessibility. Accessibility analysis would ascertain where the project has made different opportunities and facilities more accessible, and the evaluation should analyse other changes such as the economy, population, property prices, and land use developments as well as changes to other transport modes.

3.9.4 The TIEP Report

In recognising that transport investments can deliver economic benefits over and above conventionally measured user benefits, Venables et al. (2014), in the TIEP report, suggested assessment required an appropriate analytical framework to capture those effects. As techniques were often insufficiently context-specific, they needed a clear narrative about likely economic impacts to inform the modelling and quantification work undertaken and the analytical work and empirical evidence employed. Econometric modelling methods had opened up possibilities for improving prediction and needed to be developed further, including techniques for predicting land use change and its effects. The effectiveness of modelling techniques should be assessed by systematic comparison of ex ante estimates made using various techniques with actual ex post outcomes. In response to the recommendations of the TIEP report, DfT (2014) proposed a major update and restructuring of their guidance to improve analysis of wider economic impacts, and capture the full range of impacts. This included greater flexibility in applying new modelling and valuation approaches to supplement standard appraisal methods, including methods and tools not explicitly defined in WebTAG that may better explain context-specific evidence or dynamic economic impacts and land-use change.

3.10 Ex post evaluation

Ex post evaluation is often perceived as the weak link in the assessment process for transport infrastructure initiatives. Assessment methods have relied on ex ante appraisal, predicting how a scheme might perform rather than being based directly on the outcomes of past decisions (Worsley, 2014). The intervention is seldom the only change taking place within a particular region, and other developments have probably occurred in the years since the project became operational, making it very difficult to determine its net contribution of the intervention. A strong evaluation can distinguish the effect of a project from all other potential influences to produce a good estimate of the counterfactual.

Ex post evaluation centres on impact – which should be evident after a certain period of time, and sustainability – whether the effect is continually produced. It was important to check the causal relationships of the project (Department for Business Innovation and Skills, 2011), and a successful evaluation must establish causality through the counterfactual, distinguishing different impacts depending on gender, ethnicity, or social layers. Weaker models may only suggest association and correlation, and not establish a causal link between the intervention and the outcome.

The problem with evaluation is that whereas the impact of an intervention can only truly be assessed by comparing actual and counterfactual outcomes, the counterfactual is not observed. The so-called “fundamental problem of causal inference” (Holland, 1986) is that causal effects can never be directly observed, because we can never observe both potential outcomes for any individual. Although there has previously been little application of ex post evaluation in the transport sector, its use is increasing (Laird et al., 2012), with its primary role in informing and improving the ex-ante appraisal process (Worsley, 2014).

Various methods have attempted to address the fundamental question of the missing counterfactual. Each method carried assumptions about the nature of potential selection bias in program targeting and participation, which were crucial to developing an appropriate model to determine impacts. These methods include randomised evaluations (Imbens and Angrist, 1994; Kremer, 2003), matching methods, specifically propensity score matching (PSM) (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2002), and double-difference (DID) methods (Abadie, 2005).

The use of control and comparison sites may determine whether any changes associated with a project are a direct consequence of the intervention. However, establishing control sites may be difficult to achieve due to identifying areas which have not been exposed to similar interventions and neighbouring areas may be affected due to spill over effects from the project. Geographic impacts should include these effects where a policy implemented in one area has a positive or negative impact on neighbouring areas, and the “two-way road effect” where negative impacts may be felt in one location due to the competitive impact at other connected points.

3.11 The SMS Scale

Sherman et al. (1997) found that the actual process of conducting evaluations was complex as determining cause and effect often proved difficult, and consequently, they developed a 5-point scale called the Maryland Scientific Method Scale (SMS) to evaluate the methodological quality of studies where confidence in the results is highest at level 5 and level 3 should be the minimum level required to achieve reasonably accurate results. Criteria for each level were:

- Either a cross-sectional comparison of treated and untreated groups or a before-after comparison of treated group, without an untreated comparison group.
- As above - with the addition of control variables or matching techniques to account for cross-sectional differences between groups.
- Comparison of outcomes in the treated group before and after an intervention, with a comparison group providing a counterfactual. Justification must be given that the comparator group is similar to the treatment group with evidence of comparability of groups. Techniques such as regression and propensity score matching may be used to adjust for differences between groups.
- Quasi-randomness in treatment so that it can be credibly held that treatment and control groups differ only in their exposure to the random allocation of treatment.
- Reserved for research designs that involve explicit randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.

3.12 Previous evaluation studies

Ranking the evaluations on the strength of research methods by using an adjusted version of the SMS scale, What Works-Evidence Review (What Works Centre for Local Economic Growth, 2015) retained only the more robust impact evaluations, and shortlisted all impact evaluations with a potential score of three or above. They found only 29 studies of 2,300 transport evaluations but found no evaluations scoring five. There was "a lack of evidence" in that "We found no high quality evaluations that provide evidence on the impact of rail infrastructure on employment, and only a limited number of evaluations showing that road projects have a positive effect" (What Works Centre for Local Economic Growth, 2015 p5).

Seven studies considered the effect of proximity to new rail stations on residential property prices. Although they found evidence of a positive effect, this depended on the distance to the project and may vary temporally. There was weak evidence of an announcement effect – i.e. increase in property prices post-announcement but pre-completion of the project. No high quality evaluations offered evidence of the impact of rail infrastructure on employment, and there was insufficient evidence to generalise about the spatial distribution of effects.

They recommended establishing causality through comparing outcomes between treatment and control groups, using area wide averages with new schemes omitted as a very basic control group. A key issue was the ‘selection into treatment’ problem, and many studies addressed this using variations on difference-in-difference or panel fixed effects methods. Here, the control group was constructed to be similar to the treatment group either by matching on observed characteristics or using control variables taking a before-and-after difference, so eliminating all fixed unobservable differences between the groups. Despite this, there were also likely to be time varying unobservable differences, and these methods cannot account for these underlying factors, and they may bias the evaluation, as the unobservable characteristics determine both treatment and outcomes.

(Mott Macdonald, 2010) found a low number of ex post evaluations outlining the economic impact of rail investment for particular case studies, mostly affecting the relationship between accessibility and property prices. Increased rail investment appeared to centralise economic activity to the better connected areas depending on the underlying economic conditions of those areas. Rural and peripheral neighbourhoods with access to good local services and reasonable options for commuting to large employment centres could experience growth or stabilisation in population and survive as commuting centres.

3.13 Treatment and Control Groups

Tischer and Shea (1997) saw the introduction of a control group as a necessary feature when studying the effects of a transportation modification, otherwise changes in behaviour could also be attributed to extraneous factors operating in the environment.

For selection of treatment and control groups, the literature suggested matching of zones of comparable accessibility subject or not subject to rail infrastructure changes (What Works Centre for Local Economic Growth, 2015). Initial selection of groups has been variously based on several criteria including changes in distance from the rail link due to the intervention (Gibbons and Machin, 2005), but because treatment and control groups may not be fully comparable in terms of observables, techniques such as propensity score matching (Rosenbaum and Rubin, 1983; Steiner and Cook,2013) and clustering techniques (Saxena et al., 2017) could adjust for differences between the groups to reflect similarities and dissimilarities. There has been considerable debate as to the appropriate statistical technique for analysing changes in the non-equivalent control group design. Selection of treatment group based on group differences can result in under-identification of a change model (Linn and Werts, 1977). However, change score analysis can be applied if it is reasonable to assume that any group differences are stationary over time (Kenny, 1975).

3.13.1 Propensity score matching

There is potential for bias in the selection into treatment process because the difference in outcome between two groups of units may depend on characteristics that determined whether or not a unit received a given treatment instead of being due to the effect of the treatment itself. Propensity score-matching (PSM) methods attempt to reduce the bias due to 'confounding variables' and provide a natural weighting scheme that produces unbiased estimates of the treatment impact. The treatment case is matched with one or more control cases based on propensity scores (Rosenbaum and Rubin, 1983) which can reduce selection bias, and hence strengthen causal arguments in quasi-experimental and observational studies.

Matching mimics randomisation by creating a sample of units receiving the treatment comparable on all observed covariates to a sample of units not receiving treatment and pairing treatment and comparison units that are similar in observable characteristics. To predict the probability of a unit being assigned to the treatment group, as many observed covariates as possible should be included in a propensity score model. Although matching is straightforward for a small number of characteristics, if there are too many variables, it may be difficult to decide which dimensions to use to match units.

Propensity score analysis was originally developed on cross-sectional data, but as research data have become increasingly complicated, including longitudinal data, multilevel data, and complex survey samples, there are methodological challenges to its development and use. Some covariates may be influenced by the treatment, and some may not have any association with the outcome, so including them will increase the variance of the estimated treatment effect while not reducing the selection bias (Brookhart et al., 2006). Sensitivity analysis that assesses the potential impact of unobserved confounders on the treatment effect is a useful alternative and should always complement propensity score analysis (Steiner and Cook, 2013).

3.13.2 Clustering Methods

Clustering techniques provide another perspective for identifying heterogeneity within case study regions, revealing other factors for grouping not yet considered. (A cluster is formed when several data points lie in a small interval.) Cluster analysis comprises a set of statistical methods for discovering new group structures when exploring datasets based on the information found in the data describing the objects or their relationships (Everitt et al., 2011). Objects in a group should be related to one other and unrelated to objects in other groups, so that the greater the homogeneity within a group and the greater the difference between groups, the more distinct the clustering. Hierarchical clustering successively merges the pair of most similar clusters together to form a new larger cluster until all points are merged into one cluster. The order and height of merges is normally visualized in a tree-like diagram called a dendrogram (Figure 3-3).

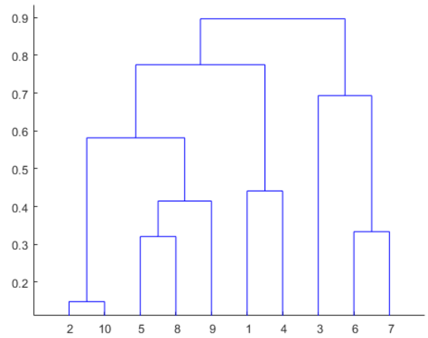


Figure 3-3 Example of a dendrogram

There are several ways to calculate the distance between pairs of clusters containing more than one linkage, and each type of linkage produces a different type of clustering solution on the same data. Single linkage is usually very useful in identifying “outliers” which have a value much greater than or much less than other data in the set (Gower and Ross, 1969). K-means offers an alternative method of clustering, defining dissimilarity using Euclidean distance from the cluster centre, where points are assigned to clusters to minimize the overall distance between points and the cluster centroids of the assigned clusters (MacQueen, 1967; Hartigan, 1975; Steinley, 2006). However, K-means suffers from the familiar problem of locally optimal solutions i.e. a data point is optimal (either maximal or minimal) within a neighbouring set of candidate solutions. The final partition depends on the initial configuration, hence the importance of the choice of starting partitions.

3.14 Accessibility

Underlying definitions of accessibility vary widely, and may be concerned purely with spatial separation of one point from another, or from all other points (de Lannoy and Oudheusden, 1978). Accessibility can also imply relative nearness, either in the sense of a direct linkage or expenditure defined in terms of the travel cost of observed or expected trips (Savigear, 1967). The accessibility of an area has also been defined as “the average opportunity for residents in an area to take part in a particular activity or set of activities” (Wachs and Kumagai, 1973), or “the potential of opportunities for interaction” (Hansen, 1959) i.e. the opportunity for an individual at a given location to take part in a particular activity. It has often been more convenient to compare the accessibility of one location with another in terms of a “Public Transport Needs Gap” (Currie, 2003) by measuring the geographical distribution of transport need, and assessing the distribution and quality of public transport services.

There was evidence that accessibility to a range of activities is valued more by poorer people than by richer, who trade off good accessibility against other factors such as a pleasant environment (Dunphy, 1973; Shindler and Ferrari, 1967). Koenig (1980) discovered a strong correlation between non-work related trip rates of non-working people and their accessibility to shops and services. However, despite access to work appearing the most important accessibility factor affecting residential location, it has been suggested that access to activities such as primary school or shopping centres may be as important. Access to some activities may be an important limiting factor in an individual's choice of residential location, but good accessibility may be sacrificed to gain other attributes.

3.14.1 Accessibility barriers

More recent approaches to accessibility planning have aimed to identify spatial barriers in accessing jobs and key services faced by different groups, particularly those most disadvantaged (DfT, 2005). This is relevant to the rural context where key locations are often less accessible and temporal barriers refer to either a mismatch between when services are available and people can access them, or where the required travel times exceed some acceptable maximum threshold. Each population group required access to specific activities or opportunities, determined through housing supply, and the location and timing of local services and facilities which allow participation to the activity.

The cost of travel is recognised as a key accessibility barrier where the emphasis is on affordability (Social Exclusion Unit, 2003). Travel costs are more significant for some groups than others, with low paid employment only feasible where fares are at a level to make the employment option attractive (Dowler, 2002). In the UK, the steady rise in public transport fares relative to the costs of motoring has tilted the balance in favour of car accessibility.

3.14.2 Measuring accessibility

Whereas basic accessibility metrics measure travel times to specific destinations, other accessibility metrics may consider access to many different types of destinations, using different transportation modes and at various times of day. A primary distinction is made between people measures i.e. travel patterns, attitudes and the needs of specific groups, and place measures i.e. transport usage characteristics of people residing in different types of area such as deprived urban areas and rural areas (Church et al., 2000). However, generalisation about transport needs and accessibility preferences is problematic as the population within any specific area can be quite heterogeneous.

Handy and Niemeier (1997) argued there was no best method for measuring accessibility because different approaches were required for different situations and purposes. In assigning a value to accessibility i.e. its monetary “worth” for the journey–to-work trip, they seemed to underestimate the impacts of accessibility change on social inclusion as it made no allowance for competition for jobs within the employment market (Van Wee et al., 2001). A major difficulty is evaluating the value and quality of each reachable opportunity (Halden, 2011) which could be weighted according to some form of relevance, reflecting choice in obtaining a service or reaching an opportunity.

3.14.3 Approaches to developing an accessibility index

As transport system characteristics, such as level of service attributes, may differ across the day, an accessibility index was often defined to be specific to the time of day. A weighted average would ensure that average peak-period measures receive a higher weight to reflect increased facility use, and a lower weight for off-peak times which has the advantage that variations in traffic flow are recognised in peak-period fluctuations.

Wilson (1971) suggested that the type of disaggregation, a definition of origins and destinations, and a measurement of attraction and impedance should be considered. Accessibility measures could be expressed either in units of opportunity (e.g. number of jobs accessible to a group) or units of transport deterrence (usually time or generalised cost, e.g. average time to reach hospital). For most published accessibility indicators, the units of measurement reflect either the potential opportunities being reached at destinations or a measure of separation of the means of reaching them.

The deterrent effect of travel varies according to the trip purpose and socio-economic group (Handy, 1993). Travel time was most easily quantifiable measure of separation (Department for Transport, 2004), and travel time thresholds for access to education, healthcare, work and shopping could be specified. Generalised cost allowed measures for different trip purposes to be combined since the same units are used for all trip purposes.

There were a wide range of accessibility indices (Koenig, 1980) often defined and interpreted for specific activities such as shopping (Handy, 1993), health (Wachs and Kumagi, 1973), and job commutes. In providing a more comprehensive measure of accessibility encompassing all trip purposes, a weighted average approach to aggregate accessibility measures over a range of transport modes offers an alternative option (Levinson and Kumar, 1994). However, it may result in relatively high indices to areas having few fast modes, and low indices to areas with having a wider range of modes but higher average travel times.

Two basic types of indicator based on this classification are opportunity measures (based on catchment contours) and value measures. Contours do not make implicit assumptions about a person's perception of transport and land use (Geurs and Ritsema van Eck, 2001) and represent more than just a social view of need as they fit reasonably accurately to the observed behaviour for any group. However, they are limited in assuming that all opportunities (e.g. jobs) within the fixed threshold time are equally desirable, regardless of the time spent reaching them (Vickerman, 1974; Ben-Akiva and Lerman, 1979).

Space-time measures define accessibility in terms of the opportunities available within specific time windows, allowing for the time needed to access opportunities, and the minimum time required for participation at the destination (Geurs and Ritsema van Eck, 2001; Ashiru et. al., 2003). However, because of the large data requirements or data being unavailable or insufficient to establish time budget constraints, only small areas can be studied at any one time (Geurs and Ritsema van Eck, 2001).

Infrastructure-based measures consider travel speeds by different modes, operating costs, and the service level of transport infrastructure, such as the "average travel speed on the public transport network". Gravity-based measures reflect travel behaviour so the potential attractiveness is an inverse function of the origin-destination distance (Hansen, 1959). Opportunities are treated differently along a continuum of time and distance using a generic distance decay function as a proxy for the disutility experienced by transport users with increasing travel time, cost or effort (Geurs and van Wee, 2004).

3.15 Job Accessibility

As commuting usually represents the most regular form of travel, employment is considered the most likely single destination type for an accessibility measure (Horner and Mefford, 2005), and one important task of a transport system is connecting workers to jobs (Grengs, 2010). Travel time to activities has been found significant in house location choice, and the jobs-housing concept has long been the centre of transport and land use studies (Zondag and Pieters, 2005; Ma and Banister, 2005). Job accessibility has been variously defined as the "potential of job opportunities for interaction" (Hansen, 1959), the "ease of reaching work places" (Cervero, 1996), or "the amount and diversity of places that can be reached within a given travel time and/or cost" (Bertolini et al., 2005).

Job accessibility comprises three sub-systems: transport, jobs, and workers (Cheng and Bertolini, 2013) (Figure 3-4).

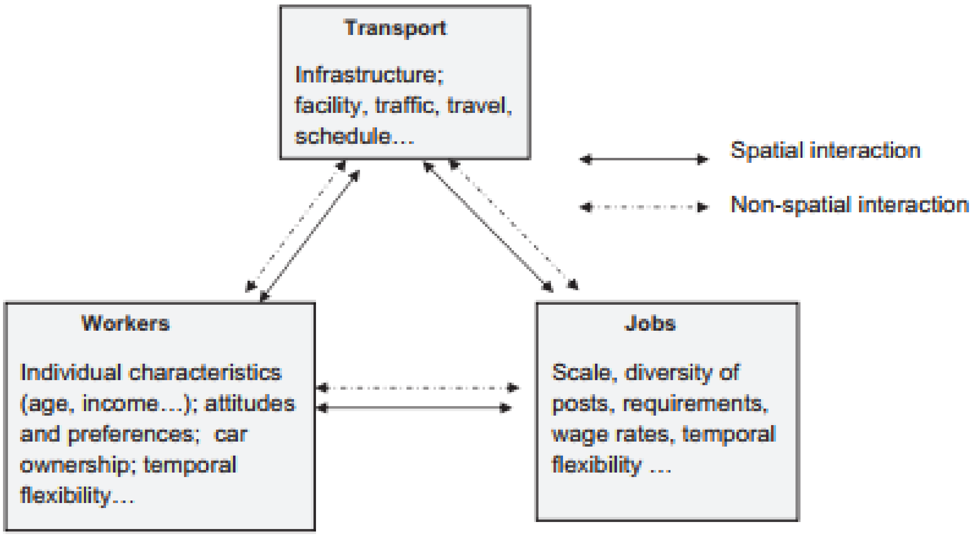


Figure 3-4 Conceptual framework of job accessibility (Cheng and Bertolini, 2013)

Non-spatial elements of the transport sub-system, such as service schedules, contribute largely to variations in mobility provision. For the workers sub-system, individual employee characteristics (e.g. income), attitudes and preferences, flexibility in working hours (e.g. part time) etc. influence job and travel demands (e.g. transport modes, travel frequency). For the jobs sub-system, employers' pay scale and range of jobs impact on demand for workers and provision of transport services to the workplace. Job accessibility depended on the spatial and non-spatial interaction between the sub-systems (Wegener and Fürst, 1999). Spatial interaction between workers and jobs affected spatial dimensions of accessibility such as competition between workers.

3.15.1 Approaches to job accessibility

In measuring longitudinal shifts in job accessibility, Cervero et al. (1998) suggest gravity-based measures or isochronic measures where opportunity is often expressed in terms of number of jobs or workers. This can be represented graphically by isochrones defined for a particular mode of travel by a threshold of travel distance, travel time or cost. A 'basic' gravity-like measure of job accessibility (Hansen, 1959) was measured as in Equation 1:

$$A_i = \sum_j O_j F(C_{ij}) \qquad (1)$$

Equation 1 Gravity measure of job accessibility

A_i - The accessibility index for zone i, standardized as standard deviations from the mean score

O_j - The number of job opportunities in zone j

$F(C_{ij})$ - The impedance for travelling from zone i to zone j for cost of travel C_{ij}

Hansen suggested that if opportunity demand was not equally distributed in space, and/or there was a restriction on the number of opportunities, the method may lead to inaccurate results. Gravity-based measurement generally represented job accessibility taking distance decay into account providing an accurate estimate for accessibility comparison. However, calibration of a distance decay function and parameter had proved difficult, requiring historical or empirical travel survey data (Reggiani et al., 2011).

Key methodological issues proposed in the job accessibility measurement literature were job proximity, job availability, and local job competition. Job proximity influences the ability to seek and hold jobs and can be based on geographical distance, travel time or travel cost. It can also be subject to frontier effects through imposition of an artificial boundary of the catchment area for reachable jobs, though workers may apply for jobs and face competition from outside their residential region.

Job availability (Ihlanfeldt and Sjoquist, 1998) refers to a qualitative match between the skill requirements of the job vacancies and individual skills of the job seekers. For local job competition, a direct measure of vacancies would ideally be used rather than all existing jobs (Ihlanfeldt and Sjoquist, 1998), only counting as competitors the jobless and employed workers actually seeking a new job. Data availability issues may restrict consideration to occupied jobs and active workers instead of vacancies and actual job seekers (Korsu and Wenglenski, 2010).

3.15.2 Spatial barriers

The spatial barrier represents the degree of spatial separation between employers and the residential locations of workers using the measure of distance, time or generalised cost in a physical space. Functions for quantifying the distance decay effect include the inverse power function and the negative exponential function. The former is argued more suitable for analysing short distance interaction at urban or regional level, and the latter for longer distance interactions at national level (Fotheringham and O'Kelly, 1989). Distance decay functions and parameters should be estimated for different modes and household characteristics (Geurs and Ritsema van Eck, 2003). Olsson (1980) and Taylor (1971) considered five different functional specifications for the decay function:

(a) Exponential $e^{-\beta_1 T_{ij}}$

(b) Exponential normal $e^{-\beta_2 {T_{ij}}^2}$

(c) Exponential square root $e^{-\beta_3 \sqrt{T_{ij}}}$

(d) Log normal $e^{-\beta_4 (\log T_{ij})^2}$

(e) Power ${T_{ij}}^{-\gamma}$

T_{ij} - Travel time between location i and location j and can be replaced by cost C_{ij} or distance D_{ij}.

The distance decay parameter (β) is a quantitative conception of human interaction with space in order to access resources (Skov-Petersen, 2001), dependent on spatial structure and interaction behaviour. The perfect value of β is hard to estimate as human behaviour is difficult to model (Fotheringham, 1981) and depends on the condition of the transport system, traffic congestion and other similar factors. Even in perfect conditions, unwillingness to travel also affects the value of β (Harris, 2001).

3.15.3 Job proximity

In the literature, different models are used to measure the proximity of jobs and workers. In the discrete approach, all jobs within a particular distance are reachable, while those located further away are excluded from the worker's local labour market. For Korsu and Wenglenski (2010), jobs (X_j) located less than 60 minutes away from a residence are reachable (Equation 2).

$$\text{Accessibility } A(i) = \sum X_j \left(T_{ij} \leq \tau\right) \tag{2}$$

Equation 2 Accessibility formula based on reachability

$(T_{ij} \leq \tau)$ is an indicator function which equals 1 if travel time between locations i and j is under τ minutes and equals 0 otherwise.

In continuous models with decay function (Bania et al., 2008; Allard and Danziger, 2002; Cervero et al., 1998; Sanchez et al., 2004), jobs are weighted to be inversely correlated with distance with proximity (Equation 3) measured using straight-line distance (D_{ij}) or travel time (T_{ij}).

$$A(i) = \sum X_j e^{-\lambda D_{ij}} \text{ or } \sum X_j e^{-\lambda T_{ij}} \qquad (3)$$

Equation 3 Standard job accessibility formats

Where λ is a decay parameter which calibrates the extent to which each additional km or minute adversely affects job search.

3.15.4 Travel time and travel cost

The demand for transport is explained by standard micro economic theories (Hensher and Brewer, 2001) and behavioural models (Leventhal and Brooks-Gunn, 2000) where a rational passenger will weigh transport against the cost of other goods then choose the transport mode giving the lowest generalised costs for a specific travelling distance. There is a price and alternative use for time, implying that time spent on board transport could be allocated to other activities and has an alternative usage that can be given a monetary value (Becker, 1965; Bruzelius, 1979).

$$GC = P + U(M) \quad (4)$$

Equation 4 Generalised cost formula

where in Equation 4, GC is the generalised cost, P are the monetary costs and U(M) the non-monetary (time) costs of a journey for transport mode M. The total journey cost for passengers making journeys with different transport modes over different distances is derived by relating average speed and time costs of respective transport modes. The total costs are defined as the generalised journey costs in the equation (Balcombe et al., 2004) and are the sum of all the monetary and non-monetary costs.

3.15.5 Job availability and job competition (spatial mismatch)

Even were a job reachable, it may not necessarily be “available” to every worker, since individual characteristics determine the actual matching of jobs and workers. Using aggregated data on both supply and the demand sides of the market, and comparing the stock of workers living in any given zone with the stock of jobs that are reachable, the job availability of any zone equates to the pool of jobs that are reachable from the zone (Ong and Miller, 2005; Johnson, 2006; Bania et al., 2008). Other recent papers use census micro data allowing for a one-dimensional sub-setting of the local labour market. The job availability for a zone equals the pool of jobs within a subset that is reachable according to the proximity measure. This makes the implicit assumption that any job of a given socioeconomic status (Korsu and Wenglenski, 2010) or education level (Mata et al., 2009) is potentially identically available to any worker of the same socio-economic status.

Accessibility to reachable jobs available to any worker will depend on the number of competitors claiming to form a match (Weibull, 1976; Ihlanfeldt, 1993; Harris, 2001; Van Wee et al., 2001; Kawabata and Shen, 2007). Job accessibility is not measured correctly if job competition is ignored (Geurs and van Wee, 2004; Shen, 1998). It is essential to identify the reachable and available jobs for any worker resident in a zone and measure the number of actual labour market competitors for each job. Then, job accessibility is defined as the ratio of weighted reachable jobs to the number of labour market competitors for these jobs as in (Kwok and Yeh, 2004; Sanchez et al., 2004; Shen, 1998). Sanchez et al. (2004) developed a gravity-based accessibility model, incorporating competition effects as well as the distance decay effect using a negative exponential function to represent the travel friction effect.

Bunel and Tovar (2014) argue that different local job accessibility models can lead to significantly different empirical depictions of job accessibility. The empirical differences are spatially differentiated, and they found that failing to fully estimate job availability may overestimate job accessibility levels of poorer areas.

Wang et al. (2003) argued that job accessibility represented the summation of only those jobs spatially and socially accessible to them, depending heavily on the social match between workers and jobs. This match can be interpreted as skills, occupational (Cervero et al., 1995), educational degree (Geurs and Ritsema van Eck, 2003), income (Wang et al., 2003) or even gender (Mata et al., 2010). Cervero et al. (1995) introduced ‘occupational match’ into the measurement of job accessibility, where only matched jobs can be taken by specific groups of workers. Workers and jobs should be segmented according to classification, and diversity accounted for in measuring job accessibility. However, little measurement of job accessibility has appropriately incorporated this diversity element.

It has been widely acknowledged that accessibility is key to understanding variations in unemployment and job search success rates. The spatial mismatch hypothesis, Kain

(1968), proposed increasing physical isolation and inaccessibility of inner-city residents from suburban employment opportunities was the root cause of chronic joblessness and persistent poverty.

Many papers have questioned the empirical reality of Kain's hypothesis and the relative importance of its determinants: individual characteristics, job access, employment decentralisation and residential segregation. There were mixed conclusions, probably stemming from methodological difficulties when assessing local job accessibility (Kain, 1993; Ihlanfeldt and Sjoquist,1998), but using improved measures, subsequent papers did validate the spatial mismatch hypothesis, highlighting the adverse effect of poor accessibility on employment outcomes (Ong and Miller, 2005; Johnson, 2006).

3.16 Hedonic evaluation approach

The economic theory supporting the hedonic hypothesis dates (Griliches, 1961; Lancaster, 1966) was later formulated by Rosen (1974). The classic theory is that "goods are valued for their utility-bearing attributes of characteristics" (Rosen, 1974, p.34) implying that all individual characteristics of each commodity contribute to the price of that commodity. In the hedonic approach the economic value is estimated through implicit pricing of the characteristics based on market values; the more closely this relates to usage of a natural resource, the more suitable this approach for evaluation.

Housing properties are particularly appealing in this context as their values are strongly influenced by location and neighbourhood characteristics. A house has several attributes e.g. number of rooms, bathrooms and car parking availability, all of which impact differently on the price of the property. However, the hedonic approach is restrictive, describing a market perfectly transparent on the supply side, and homogeneous and perfectly competitive on the demand side. It requires perfect transparency of prices and characteristics, and prices must adapt immediately to changes in demand for environmental goods.

The estimation of the function of the prices of private goods is generally obtained applying statistical methods of linear multiple regressions as in Equation 5:

$$\mathrm{Y}= \alpha_0 + \sum \alpha_v x_v + \epsilon \qquad (5)$$

Equation 5 Hedonic price function

Where Y is the dependent variable e.g. property price, ϵ is the error, and x_v (v = 1…p) are the explanatory variables or individual characteristics. Each is associated with a parameter α_v representing the implicit price, if the above conditions are satisfied according to the specification of the variable.

3.16.1 Determinants of property price

Following the argument of Rosen (1974), Visser et al. (2008) categorised attributes into the physical characteristics of the house, and the physical, socio-economic and functional characteristics of the residential environment. The most important factors were the physical housing characteristics including size, volume, surface area, building age and housing type, and several studies suggest these factors explain more than half of the variation in house prices. Physical characteristics of the residential environment concern mainly 'green' (e.g. parks) and 'blue' (e.g. waterways) amenities (Garrod, 1992; Garrod and Willes, 1994; Powe et al., 1995; Daly et al., 2003).

There were studies on the effect of neighbourhood characteristics on house prices including the impact of crime (Dubin and Goodman, 1982; Lynch and Rasmussen, 2001) and neighbourhood composition (Cervero and Duncan, 2004; Rosiers et al., 2001; Noonan, 2005). Powe et al. (1995) found that positive socio-economic factors e.g. high levels of ownership and low levels of unemployment were more noticeable in areas with the highest level of environmental attributes. However, high correlation between socio-economic and physical or environmental variables made interpretation of the estimated coefficients of these variables more difficult.

Initially, as improved proximity to transport facilities increased the speed of travel to the Central Business District (CBD), this was assumed to be the most important factor in determining the value of the property. Accessibility of services was also thought important with improved accessibility leading to increased property values (Miller et al., 1982). However, higher income households with higher education preferred to live in high quality dwellings located further away from the CBD (Kain and Quigley, 1970).

Accessibility is not the only reason for the choice of a house: property characteristics (surface area, level of equipment, type of dwelling) and local characteristics (environment quality, school proximity, rail proximity) may all contribute. So, Tse and Ganesan (1997) adopted a hedonic price model and applied various determinants of house price include structural attributes such as size, age and floor level, and amenities such as accessibility to various means of transport. There are also a number of house-specific explanatory variables, and other factors identified are proximity to transportation facilities (Cervero and Duncan, 2001) and good schools (Gibbons and Machin, 2008).

3.16.2 Transport impact on property prices

For many years it has been considered that development of public transport infrastructure may lead to changes in accessibility which produce economic benefits to those located close by, particularly measured in house price movements. Early studies on residential property value showed this increased in response to transport cost savings and confirmed that the distance from the nearest station had a statistically significant effect on the property value of the land (Boyce et al., 1972; Dewees, 1976; Damm et al., 1980; Grass, 1992). Previous research also suggests a causal link between property prices and accessibility (Henneberry, 1998), and hedonic price methods have been widely considered as an approach to identify the impact of transport investment on land value.

Despite some indication of negligible or negative impact of rail investments, the majority of studies reported a positive effect with negative impacts at locations very close to stations or railway lines where noise, pollution and crime levels were higher. It is argued that railway stations have a higher effect on commercial properties (Weinstein and Clower, 1999; Cervero and Duncan, 2001) as being focal gathering points they attract commercial activities which increases commercial property values. Most empirical literature on property value considers a wider impact area of railway stations for residential properties, but only immediately adjacent areas for commercial properties (Debrezion, 2007). The reason is that the direct proximity effect dominates for commercial properties only when they are within walking distance, otherwise the station is of little use and the attractiveness drops off quickly.

Results show that commercial property values are positively and significantly associated with the accessibility benefits of transport nodes. The distance-band coefficients form a typical distance decay curve for both modes with no detectable disamenity donut effect immediately around the nodes. Only the links of light rail are negatively associated with property values, as hypothesized. When the sample is subdivided by type of commercial property, the magnitude and distance extent of impacts are surprisingly consistent, with light rail stations having stronger impacts than highway exits on all three classes of commercial property: industrial, office, and retail and service.

More recently, Efthymiou and Antoniou (2013) using both hedonic regression and spatial econometrics, found that proximity to transport infrastructure can have either a positive or negative impact depending on the type of system, which was borne out by Mohammad et al. (2013) where a meta-analysis of the impact of rail projects on land and property values in the United States found a large variation in estimates. Context was important as similar transport investments impact differently in locations with different economic conditions. Residential property prices might be depressed in the immediate vicinity of the transport investment or station, and most studies consider a series of key thresholds.

Most research has concentrated on urban rail systems, as offering the best opportunities to test for property market effects. Hedonic pricing has generally been the preferred approach to identify and isolate the property market effects. Gibbons and Machin (2005) studied the effects of rail access on house price in London and the South East England, observing what happened before and after a transport intervention. Adopting an empirical approach to determine how consumers valued rail access, they studied the effect of the intervention on house prices where the distance to the nearest station altered for some households but not others. Rail access appeared significantly valued by households compared to that of other local amenities and services.

There is variability in the relationship between transport accessibility and land value over space (Landis et al., 1994; Cervero and Duncan, 2002a, 2002b). In hedonic modelling this has been met by applying hedonic price models to either submarkets or to different types of properties (Adair et al., 2000), but these approaches require the land area to be subdivided using some arbitrary “rule” to give boundaries for the separate hedonic models. Immediate locations are expected to produce higher effects than locations further away, and high population movement leads to the development of retail activities (Bowes and Ihlanfeldt, 2001).

Blainey and Preston (2010) employed control groups to ascertain whether changes in the economic and social circumstances of areas with new stations could be directly linked to the intervention. They established the impact of the stations on employment and property prices in the areas, and found that new stations may have additional impacts not currently considered during project appraisal. However, quantifying these effects for inclusion in the appraisal process was problematic e.g. it was not obvious what value should be placed on changes in population and employment. Billings (2011) estimated one hedonic model for both treatment and control area types with an extra variable to represent treatment.

A measure of location relative to employment is often included in hedonic housing price models and is most often represented by distance to the CBD (Central Business District). However, this does not consider the decentralization of employment in urban and rural areas. Ottensmann et al. (2008) tested the performance of alternative measures of location and found that distance and time to multiple employment centres are preferable to simple distance to the centre.

Many examples in the literature applying hedonic methods to house prices address mainly an urban environment and adopt different accessibility characteristics (Forrest et al., 1995; Henneberry, 1998). Visser et al. (2008) linked a variety of environmental attributes to the properties involved in transaction data. These incorporated physical characteristics of the residential environment; the socio-economic status of the neighbourhood; location characteristics of the dwelling relating to proximity of a railway station; the accessibility of jobs and the local housing market. They found that neighbourhood characteristics have different effects between urban and rural housing markets and under different housing market conditions. Bowes and Ihlanfeldt (2001) measured accessibility both by distance to a railway station and the quality of services provided at the station. Considering both the nearest and the most frequently chosen railway station, the latter proved a more effective measure, but there was dissimilarity between the results of the two models, dependent on the urbanisation level of the metropolitan area.

One alternative to hedonic pricing is the comparison approach, where the relationship between property price and transport accessibility is determined by isolating the latter from other factors through comparisons of property price. However, this method may not identify the more complicated, multi-dimensional features underpinning property values, and there may be other factors at play. Cervero and Landis (1993) found that commercial property adjacent to some stations was subject to higher rents compared to more distant locations. However, they could not fully control for external variations between control and station areas, and using similar methods, Du and Mulley (2007b) were unable to identify any significant change in property prices as a result of the extending the Tyne and Wear Metro.

3.16.3 The proximity and announcement effect

Often changes will take place in property values in advance of completion of the transport investment in the expectation of improvements in the transport infrastructure. Some hedonic analyses find very high positive valuations for close proximity to railway stations and significant announcement effects of rail projects, while others find new lines have negative or no effects. (The announcement effect refers to the fact that behaviour can be changed merely by announcing a future policy change). Damm et al. (1980) examined movement in property values in anticipation of the construction of a heavy rail transit system in Washington and found increasing distance to a metro station was associated with lower property values with a substantial opening date effect.

However, Tse (2002) and Bae et al. (2003) found an increase in residential prices, but only prior to the line opening. Other studies found a weak effect on property prices and a weak announcement effect regardless of the distance to the station (Gatzlaff and Smith, 1993; Forrest et al., 1995). There was a general consensus that a higher level and quality of facilities in stations has a greater impact on the surrounding properties (Debrezion et al., 2007). Mayor et al. (2008) used a hedonic house price model to estimate the value of transport networks to residents in the Dublin area and found that this depended on the availability of alternative transport options in the area and how far from the property the network is located.

3.16.4 Problems with the hedonic approach

The hedonic approach estimates value based on actual choices, drawing on property data which is typically very reliable. In addition, property markets are relatively efficient in responding to information, and hence can be good indicators of value. However, results depend heavily on model specification, and the scope of measurable benefits is mainly limited to matters related to property price. There is difficulty in deciding which variables to incorporate into the model specification for interpretation, particularly between competing models on the strength of model fit where models may include different variables and dissimilar functional forms. The quality of the measures used in the independent ‘explanatory’ variables is of key importance as the use of a proxy measure may result in inaccurate coefficients in the regression analyses. The model assumes that market prices adjust immediately to changes in attributes, but in reality there could be an associated lag, especially in areas where house transactions are infrequent as in rural or sparsely populated regions.

Additionally, the functional form must satisfy the assumptions of multiple regression (Forrest and Glen, 1995; Weinberger, 2001) and may not allow for easy interpretation. These forms include linear, semi-log, double-log, or log-linear format, each of which implies a different relationship between the house price and its determinants (Li and Brown, 1980). The results of different studies are not conclusive, but it is generally acknowledged that the equation should not be nonlinear (Freeman, 1993). Cropper et al. (1988) argue that the simplest functional forms (linear, semi-log, double-log) are suited when certain explanatory variables are not observed or are replaced by a proxy.

More critically, in the presence of spatial autocorrelation or spatial heterogeneity, hedonic models will have some unexplained variance caused by interdependence between observations arising from their relative location in space. Houses located close to one another will have similar, unobservable attributes, which may not be included in the hedonic pricing model and may share local amenities and neighbourhood characteristics. The residuals from the regression model are likely to be spatially correlated (Basu and Thibodeau, 1998). If the housing transactions are spatially clustered in some way, the estimated regression coefficients will be biased and therefore not suitable for predicting property prices.

3.17 Fixed and random effects modelling

Cross-sectional data is recognised as having limited use in addressing causality. Any inference of a causal relationship from a cross-sectional parameter is limited by the possibility of unobserved variable bias (Duncan, 1972; Holland, 1986), endogeneity bias (Hausman, 1978; Berry, 1984; Finkel, 1995) and indeterminacy over the sequencing of the causal mechanism.

3.17.1 Panel Data

With panel data, or “cross-sectional time series”, the same variables are measured but at different time intervals. It is a special case of longitudinal data, where information not necessarily on the same variables is collected over time. This time dimension allows greater scope for investigating causality and accounts for individual heterogeneity and control for unobserved variables or variables that change over time but not across entities.

However, suitable statistical models are generally more complex and difficult to estimate than those for cross-sectional data, and observations for the same unit over time are unlikely to be independent of each another. Panel designs present data collection issues, and being by nature observational do not remedy problems of endogeneity and omitted variable bias. Although they represent an improvement on cross-sectional data for understanding dynamic processes, the fundamental limitation is clearly identifying causal relationships.

In setting up a panel data structure, data is not always available for each year of the study, which reduces the representativeness of the sample and distorts inferences about the population. However, there are established techniques to compensate for this; list-wise deletion which excludes the entire record if any single value is missing but reduces the sample size; pair-wise deletion which deletes a case where it is missing a variable required for a particular analysis, but includes that case in analyses for which all required variables are present. Another technique, imputation preserves all cases by replacing missing data with an estimated value based on other available information.

One form of imputation is "last observation carried forward", where a dataset is sorted according to any of a number of variables and the technique then finds the first missing value using the cell value immediately prior to the missing data to impute the missing value, repeating the process until all such values have been imputed. A missing value can also be replaced by the mean of that variable for all other cases, which has the advantage of not changing the sample mean for that variable. Finally, regression imputation uses regression to predict observed values of a variable based on other variables, which are then used to impute values in cases where that variable is missing.

3.17.2 Fixed and random effects models

Fixed effects (FE) models analyse the impact of variables that vary over time and explore the relationship between predictor and outcome variables. Each entity has its own individual characteristics that may or may not influence the predictor variables. This helps in controlling for unobserved heterogeneity where this is constant over time. Another important assumption of the fixed effects model is that those time-invariant characteristics are unique to the entity and should not be correlated with other individual characteristics of the entity (Kohler and Kreuter, 2012).

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it} \qquad (6)$$

Equation 6 Fixed effects model format

α_i (i=1….n) - The intercept for each entity (n entity-specific intercepts).

Y_{it} - The dependent variable where i = entity and t = time.

X_{it} – An independent variable, β_1 is the coefficient for that variable, u_{it} is the error term

In random effects (RE) models the unobserved variation across entities is assumed to be random and uncorrelated with the predictor or independent variables included in the model (Greene, 2008). This assumption allows time-invariant variables to act as explanatory variables. Random effects are more efficient but more restrictive in the assumptions. A random effects model is applicable where differences across entities influence the dependent variable. However, some variables may not be available causing omitted variable bias in the model. An advantage of random effects is that time invariant variables (e.g. gender) can be included. In the fixed effects model these variables are absorbed by the intercept.

$$Y_{it} = \beta X_{it} + \alpha + u_{it} + \varepsilon_{it} \qquad (7)$$

Equation 7 Random effect model format

α - The intercept.

Y_{it} - The dependent variable where i = entity and t = time.

X_{it} – An independent variable, β is the coefficient for that variable.

u_{it} - The between entity error, and ε_{it} the within entity error.

The model implies the intercept α is the same for all entities and represents a value drawn from the same distribution. Selection between fixed or random effects models can be determined using a Hausman test, where the null hypothesis is that the preferred model is random effects vs. the alternative the fixed effects (Greene, 2008), which basically tests whether the unique errors (u_{it}) are correlated with the regressors.

3.18 Difference-in-difference models (DID)

By studying the differential effect of a treatment on a 'treatment group' versus a 'control group' in a natural experiment, Difference-in-difference (DID) imitates experimental research design using observational study data (Card and Krueger, 1994; Abadie, 2005). It estimates the effect of a treatment (i.e. an explanatory variable) on an outcome (i.e. a dependent variable) by comparing the average change over time in the outcome variable for the treatment group, compared to the average change over time for the control group. Difference-in-difference employs panel data to measure the differences between treatment and control groups of changes in the outcome variable over time, using differencing to cancel out individual fixed effects. Difference-in-difference requires data from both groups measured at two or more different time periods, (at least one time period before 'treatment' and at least one time period after 'treatment').

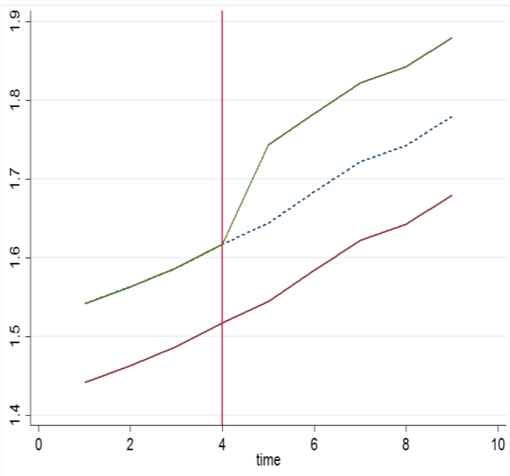


Figure 3-5 DID illustration

The key assumption is that the outcome in treatment and control groups would follow the same time trend in the absence of the treatment. In the example chart Figure 3-5, the bottom line represents the control group and the top line the treatment group which up to time of the intervention (4) are increasing at a similar rate. Post intervention, there is a much greater increase in the treatment group. Difference-in-difference modelling produces the dotted line which estimates what might have happened to the treatment group without the intervention, based on the increase experienced by the control group.

This method may still be subject to certain biases depending on selection into treatment as treatment and control groups may not follow parallel trends. Card and Krueger, in analysing the effect of a minimum wage increase in New Jersey, found it very likely that employment in each state was not only correlated within the state, but also serially correlated. Conventional standard errors often severely understate the standard deviation of the estimators (Bertrand et al., 2004).

In the Card and Krueger case, the equivalent regression model was:

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \delta(NJ_s * d_t) + \varepsilon_{ist} \quad (8)$$

Equation 8 Card and Kruger regression model

NJ - A dummy which is equal to 1 if the observation is from New Jersey.

d - A dummy which is equal to 1 if the observation is post intervention.

ε_{ist} - The error term.

The assumption of a common trend is difficult to verify, and although pre-treatment data may indicate that trends are the same, other policies may be changing at the same time. Including ‘leads’ within the model offers one way to analyse pre-trends, and lags can be included to analyse whether the treatment effect changes over time after treatment.

$$Y_{ist} = \gamma_s + \lambda_t + \sum_{\tau=-q}^{-1} \delta_t D_{s\tau} + \sum_{\tau=0}^{m} \delta_t D_{s\tau} + X_{ist} + \varepsilon_{ist} \quad (9)$$

Equation 9 DID model with leads

- Treatment occurs in year 0.
- Includes q leads or anticipatory effects.
- Includes m leads or post treatment effects.

3.19 Spatial effects and accessibility

Determination of cause-effect relationships has led to investigation of the evaluation of spatial, temporal and economic barriers, and how they differ at different points in space and time. This has led to consideration of existing statistical methods of relevance in the context of this study, particularly those for assessing regional heterogeneity and diversity, and how these benefits may be affected by the proximity of new stations and timing of an intervention.

Although hedonic price methods have been a popular approach to identify transport investment impacts on land value, (Weinberger, 2001: Cervero and Duncan, 2002), the assumption that observations in the regression are independent of each other is likely to be incorrect if there is spatial correlation with the data. Spatial data must be considered in appropriately identifying the relationship between transport infrastructure and land value. More efforts are now being made to integrate and model spatiotemporal data (Cressie, 1993); of particular interest has been the research on panel data models (Holly et al., 2006) and spatiotemporal autoregressive (STAR) models (Pace et al., 1998).

3.19.1 Spatial effects and heterogeneity

Recent developments in hedonic price modelling have used spatial econometrics to model the effects of location as a determinant of property prices because of consistent evidence that property values exhibit a systematic pattern in their spatial distribution, and consequently are said to be spatially auto-correlated. Spatial autocorrelation can be due to the price being affected by those of neighbouring houses, the omission of relevant spatially correlated variables, or the misspecification of the functional form (Wilhelmsson, 2002). Neighbourhoods often develop at the same time, have similar structural characteristics, and may share a number of amenities locally (Basu and Thibodeau, 1998). Sales prices are often influenced by property professionals (Bowen and Prestegaard, 2001), and so local housing market conditions may influence each transaction, and homogeneous neighbourhoods will serve as a proxy for other variables including similar income levels and occupational status of homeowners (Gelfand et al., 1998).

Generally, hedonic price models examine cross-sectional data, but where a longer time series of observations is necessary, indicator variables are commonly used for each designated period (Pace et al., 1998; Thibodeau, 2003). This technique removes some heterogeneity resulting from pooling over time, but may generate an undesirable number of dummy variables. Dorantes et al. (2011) reviewed the theoretical background relating to spatial hedonic models and how they could evaluate the effect of new transport infrastructure. They discovered the presence of submarkets defined by geographic boundaries, implying that economic benefits differed across municipalities. Ahlfeldt (2013) predicted a considerable degree of heterogeneity in terms of the magnitude and spatial extent of price effects around new stations.

3.20 Geographically Weighted Regression

In addition, the technique of Geographically Weighted Regression (GWR) developed by Fotheringham et al. (2002) and Brunsdon et al. (1998) has brought new insight into the understanding of spatial dynamics in econometric models by extending the traditional cross-sectional regression model (Equation 10) into a model where local variations in parameter values can be estimated by taking into account the coordinates of the variable.

$$Y_i = \beta_0 + \sum_k \beta_k \beta_{ik} + \varepsilon_i \qquad (10)$$

Equation 10 Traditional cross-sectional regression model

Brunsdon et al. (1996; 1998) proposed GWR for exploring spatial non-stationarity of a regression relationship for spatial data. Spatial non-stationarity is a condition in which a simple "global" model cannot explain the relationships between some sets of variables, and a model must vary over space to reflect the structure within the data, by producing a separate regression model for each subset of areas within the region. This was achieved by locally fitting a spatially varying coefficient regression model at the location (u_i, v_i) of the form:

$$Y_i(u_i, v_i) = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)\beta_{ik} + \varepsilon_i \qquad (11)$$

Equation 11 GWR Model

(Here there is not one set of coefficients as in a global model, but the coefficients will fluctuate for different locations within the region.)

This is then fitted using a weighted least squares method to estimate the parameters at each location (u_i, v_i), and a predicted value of Y. Weighting gives more influence to observations nearer the location than those further away to generate estimates of parameters for each data point. The weighting process uses a spatial kernel, also called a window or bandwidth, which moves over the study region, and fits the best results for each subarea. The kernel size defines the rate at which the influence of the coefficients decreases as the distance increases.

GWR explores how the relationship between the dependent variable and independent variables vary geographically, so it searches for geographical differences rather than fitting a single "global" model to the entire study region. It recognises the spatial variability of land values by estimating a hedonic model as the global model, extending it to produce a local model at each point of the data. One advantage of the GWR model is its facility to examine the spatial variability of independent variables included as explanatory variables. Some independent variables might be non-significant in the global regression model, but vary significantly over the geographical area and revealed as significant local parameters by GWR modelling.

Despite substantial related research in the United States, considerably fewer studies have been carried out on spatial variability in house prices and accessibility in the

United Kingdom, mainly concentrated in London. However, Du and Mulley (2012) focussed on Tyne and Wear and allowed for estimation of the importance of transport accessibility in determining house prices using GWR methodology with property price as the dependent variable. This was explained by independent variables designed to standardise for household features and spatially defined factors, including the transport accessibility of the house location. Results from GWR (the local model) showed significant spatially varying relationships between property prices and the variables concerned and allowed the impact of accessibility on house prices to be identified. The study offered a new methodology in the transport field that takes account of the spatial nature of the data required in this process.

Banister and Thurstain-Goodwin (2005) identified property value increase resulting from transport infrastructure investment, and found combining hedonic price methods with geographically weighted regression, allowed isolation of small-scale localised changes in the property market. Using data on actual property transactions, this approach also measured how contribution varied through time before and after the opening of the scheme, enabling changing market conditions to be included. Du and Mulley (2007b) studied the relationship between transport accessibility and increases in land value and focused on residential land, through the form of house prices, using the Tyne and Wear Region of the UK as a case study. Using GWR they examined the relationship between transport accessibility and land value and identified that this varies over space.

GWR has been applied many times in cross-sectional settings (Huang et al., 2009; Wrenn and Sam, 2012; Yu, 2010; Wu et al., 2013). Crespo et al. (2007) extended the concept of ‘closeness’ so that data points close in both space and time dimensions have greater influence in estimating local parameters. The name ‘geographically and temporally weighted regression’ (GTWR) appropriately describes the procedure used, which is basically an extension of the cross-sectional GWR weighting function. They incorporated temporal data into the GWR model and developed a spatio-temporal version to forecast and interpolate local parameters through time. The hedonic price model was calibrated using GWR for a period of nineteen years and only the subset of data points for that year was included in the model, so the bandwidths for each year T-1, T, T+1 etc. are calculated independently one another (Figure 3-6).

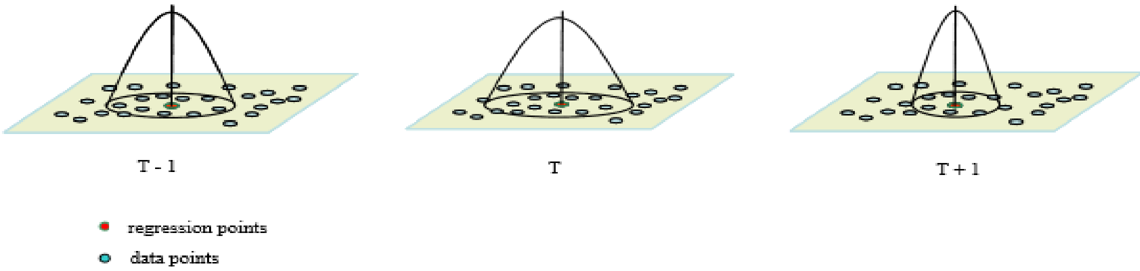


Figure 3-6 GWTR temporal illustration

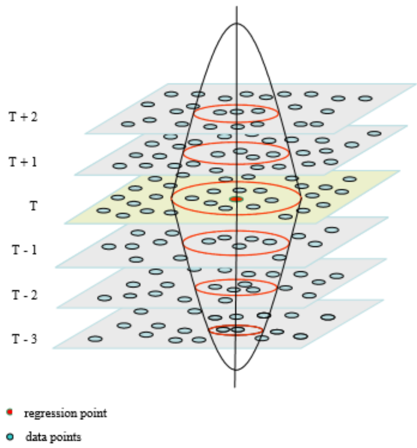


Figure 3-7 GWTR spatio-temporal illustration

In the spatio-temporal approach, the model was also calibrated for each year, incorporating data points from past and future years into the model. Hence, data points were now both spatially and temporally weighted to produce a spatio-temporal time-decay bandwidth, whose size may vary over time. Data points located temporally closer to the regression point were weighted more heavily than those further away.

(Figure 3-7) illustrates a possible time-decay spatio-temporal bandwidth. The regression-year is T, and as can been observed, spatial bandwidths become smaller as the data points are located further from the regression-year T. Wheeler and Tiefelsdorf (2005) raised concerns about the potential correlations among local regression coefficients in cross-sectional GWR..

3.21 Conclusion

Although most of the literature was based on an urban environment, the remit was to develop a methodology and evaluation methods in the context of a rural, remote or mixed urban/rural environment. This conclusion is divided into:

- The main agreements and disagreements in the literature.
- Gaps or areas for future research.
- The overall perspective to carry forward into development of methodology.

3.21.1 The main agreements and disagreements in the literature

It was widely agreed that disregarding wider economic benefits could potentially bias a transport project appraisal, but traditional evaluation methods had not been very successful in accounting for non-transport benefits. These were not restricted to urban areas, and rural and peripheral neighbourhoods should be considered differently as they offered limited job opportunities, and poor transport infrastructure provided a major obstacle to increase in economic activity. Good impact evaluation recognised that most outcomes were affected by a range of factors as well as the intervention, and would successfully isolate its effect from all other potential influences.

It was important to verify the causal relationships of a project, and without a good counterfactual there would be no certainty that the outcomes would have happened anyway, suggesting only association and correlation without establishing a causal link between the intervention and the outcome. Ideally, a successful impact evaluation should preferably establish causality through comparing a treatment group with one or more control groups to whom the activity did not apply.

It is recognised that this kind of approach works well in medicine, where the control can be truly independent, but in the context of this thesis, care must be taken as the selected control locations may be affected positively by the intervention if they are close enough to the new stations, but outside the selected treatment group area. They may also be affected negatively by the intervention for example by the reallocation of property demand towards the new stations. Identification of a control which experiences zero effect as a result of the intervention may prove difficult in this context, although the variation from that may well be small.

For selection into treatment the literature suggested matching zones of comparable accessibility subject or not subject to rail infrastructure changes. There had been considerable debate as to appropriate statistical techniques, and selection of treatment groups based on group differences could result in under-identification of a change model. Propensity score-matching methods could strengthen causal arguments by producing unbiased estimates of the treatment impact, but they had potential for bias because the difference in outcome between groups may depend on characteristics that determined whether or not a unit received a given treatment rather than being due to the effect of the treatment itself.

Changes in transport infrastructure typically improved accessibility effecting an increase in residential property prices. However, the economic benefits of transport infrastructure spending, particularly as a mechanism for generating local economic growth, were not as obvious as they might seem at face value. There was some evidence of the importance of context as improvements in transport infrastructure may improve conditions for commuters living in the vicinity, enabling easier access to places with jobs. Access to work appeared to be the most important accessibility factor affecting residential locations, and employment was considered the most appropriate single destination type for an accessibility measure. Travel time to activities was found significant in house location choice as transport connected workers to jobs; however, non-spatial elements of the transport sub-system, such as service schedules, were often neglected in the literature. Gravity-based measurement taking distance decay into account was often employed to represent job accessibility, but calibration of a distance decay function had proved problematic. There may be significantly different empirical depictions of job accessibility if competition for jobs and workers were ignored. The empirical differences were spatially differentiated, and job availability must be considered to prevent overestimation of job accessibility levels in poorer areas.

In measuring job accessibility, key methodological issues proposed in the literature were job proximity and availability and local job competition. The cost of travel was recognised as a prime factor affecting accessibility, and was more significant for some groups than others, with low paid employment only feasible where fares were sufficiently low to make the employment option attractive. Every job would not necessarily be available to every worker, since individual characteristics determined the actual matching of jobs and workers.

The development of public transport infrastructure was thought to produce economic benefits to those located in close proximity with its impact on house price movements. The hedonic modelling approach was suitable for evaluation the more closely it related to usage of a natural resource. The most important factors for modelling were physical housing characteristics e.g. housing type and neighbourhood composition, however, high correlation between socio-economic and physical or environmental variables made it difficult to interpret estimated coefficients.

Results from hedonic studies depended heavily on model specification, and the scope of measurable benefits was mainly limited to matters related to property price. The functional form may not allow for easy interpretation, and more critically, in the presence of spatial autocorrelation or spatial heterogeneity, hedonic models would have some unexplained variance caused by interdependence between observations arising from their relative location in space. There was a large variation in findings predicting both a positive or negative impact. This often depended on the availability of alternative transport options and proximity of the property to the network. A measure of location relative to employment was also often included in hedonic housing price models, but did not consider the decentralization of employment in urban and rural areas.

Due to endogeneity bias, cross-sectional data was recognised as having limited use in addressing causality, but a time dimension allowed greater scope for investigating causality and assessing individual heterogeneity. However, suitable statistical models were generally more complex and observations for the same unit over time were unlikely to be independent of each another. Difference-in-Difference methods assumed that the outcome in treatment and control groups would follow the same time trend in the absence of the treatment, but this method would still be subject to certain biases depending on selection into treatment as the groups may not follow parallel trends, and the assumption of a common trend was difficult to verify.

Geographically weighted regression (GWR) allowed examination of the spatial variability of explanatory variables where some non-significant independent variables in the global regression model may vary considerably over the geographical area and be revealed as significant in GWR modelling. However, there were concerns about the potential correlations among local regression coefficients.

3.21.2 Gaps or areas for further research

Context is important as similar transport investments may impact differently in locations with different economic conditions. Most research on rail impacts has addressed urban environments, and more should be done to address more remote areas where there is often a lack of alternatives and choices. There is currently limited development of methodology to capture the wider economy impacts of transformational projects, especially using econometric and modelling techniques for providing evidence of land use change and its effects, and new modelling and valuation approaches to supplement standard appraisal methods are required.

Arguably this can be done with LUTI and SCGE models which cover a much wider scope than the analysis in this thesis. The LUTI model offers the means to estimate wider impacts in line with existing DfT guidance, but is very complex and requires specialist expertise to develop. SCGE models help to show how the impacts of a transport intervention are dependent on other factors within the economy changing in order for the change in output to be realised. The data requirements for an operational SCGE model are very substantial, both in terms of the exogenous inputs to the model and in terms of the coefficients in the model (McCartney et al., 2013).

However, this thesis goes much deeper within the areas it covers, so for those particularly interested in the impacts measured here, or wish to be informed about them in more detail than that provided by a LUTI or SCGE model, the thesis offers an alternative or supplementary method. For instance, development of regression and matching techniques to adjust for differences between groups could provide another perspective for exploring heterogeneity within regions, revealing other factors for grouping not yet considered.

Where a project has made different opportunities and facilities more accessible, there should be furthering of accessibility analysis, taking into account other changes such as the economy, population, property prices, land use developments as well as changes to other transport modes, and the role of travel in social exclusion. These accessibility measures should be sensitive to changes in the transport system and land-use system, including the spatial distribution and temporal constraints of opportunities, and the demand for those opportunities. Indicators should recognise and reflect choice, or provide a constraint on the set of opportunities to be reached, and may be specific to the time of day. In considering social match in spatial competition, workers and jobs should be disaggregated according to classification, and diversity accounted for in measuring job accessibility. To date little measurement of job accessibility has appropriately incorporated this diversity element.

In appropriately identifying the relationship between transport infrastructure and land value, more efforts should be made to integrate spatiotemporal data to model more accurately the effects of location as a determinant of property prices. Hedonic pricing methods, combined with geographically weighted regression, offered isolation of small-scale localised changes in the property market.

3.21.3 The overall perspective to carry forward into the methodology

In considering the potential effects of transport infrastructure on the local economy, property prices, social issues and employment, the two main areas of concern to be addressed will be accessibility to jobs and essential services, and wider economic impacts specifically relating to property and employment. The critical approach to previous evaluation studies addressing rail interventions has shown the need to develop a robust method of ex post evaluation by establishing causality through the application of a meaningful counterfactual. The comparative assessment would include a control group not subject to the intervention against a treatment group which has been affected by the intervention. Recognising that the selection of treatment and control groups has been problematic, alternative methods such as propensity score matching and clustering techniques will be applied where comparable areas will be selected in terms of socio-economic characteristics.

Cause-effect in the evaluation of wider economic impacts is much more tenuous due to the difficulty in deciding which factors other than the rail intervention have contributed to the effect, and many wider impacts may take a long time to materialise. There are specific methods for definition and evaluation of impacts applicable to this research study, which may be modified to match the specific historical and geographical background addressed by this research. These include the development of econometric models for employment and property price, and the derivation of appropriate measures for accessibility and wider economic impacts to be incorporated into an evaluation framework.

Through study of accessibility, access to jobs and services can be considered alongside the socio-demographic profile. Accessibility will focus on consideration of employment as job accessibility which needs to allow for job availability and spatial mismatch between available jobs and skill levels. It will also give special consideration to rural and mixed rural/urban environments accounting for their greater distances, less populated settlements and often limited transport access.

In the context of this study, the hedonic approach appears to offer the most appropriate method for evaluation in addressing wider economic effects in the property market. Movements in property prices are very appealing in this context, as their values are strongly influenced by location characteristics. In addition, impacts can be analysed using both a panel data approach for a difference-in-difference comparison and fixed effects modelling. Due to the limitations in the hedonic approach in assessing spatial effects and heterogeneity, geographically weighted regression offers a potential methodology in its facility to allow for local heterogeneities.

4 Chapter Four: Methodological approach

4.1 Overview

Following consideration of the various evaluation approaches referenced in the literature review, a methodological structure has been derived to address the specific aims and objectives and context of this study. This structure establishes a mechanism for assessment of cause-effect relationships through investigation of datasets and comparison of suitable metrics and measurement of inputs, outputs, outcomes and impacts across three case study regions. In terms of evaluation there is a narrow focus, specifically on changes in accessibility to jobs and services and the consequent effect on property prices and employment levels and how this manifests itself in regional heterogeneity. It offers a practical evaluation framework which may have some application in future ex ante appraisals.

A standard approach is adopted here as proposed in HM Treasury (2011) and Hills (2010) for The Tavistock Institute by clarifying the context in which the evaluation is being undertaken, and defining the purpose and nature of the intervention being evaluated. A logic model has been mapped to consider the factors to be measured, and highlight evidence requirements. This logic model *(Appendix 10.4)* imposes a structure for the evaluation strategy in identifying outcomes and impacts attributable to the intervention. The model subdivides outcomes and impacts over a period of time to describe resulting short-term (1 to 3 years), and long-term impact (4 to 10 years) identifying the following elements in evaluating the intervention:

- Issues addressed and context within which the intervention has taken place.
- Inputs, i.e. the resources required.
- Initial outputs
- Outcomes (i.e. short and medium-term effects) e.g. improved access to the city).
- Anticipated impacts (i.e. long-term results such as better access to hospitals, property price increase, agglomeration effects etc.).

4.2 Theoretical Framework

Introduction

This theoretical framework focuses on specific variables and interpreting the data to be gathered. It provides a basis for the examining the hypotheses and choice of research methods and how they vary under different circumstances. The focus is on producing practical results in order to achieve the research objectives whilst minimising theoretical problems as far as possible. This research attempts to determine a balance between complexity and accuracy in order to evaluate post-intervention impacts.

Impact pathways between causes and effects in different markets

This framework restricts itself to impact pathways of cause and effect for the transport, labour and property markets. Wider economic impacts are illustrated in the right hand part of the flow chart (Figure 4-1) and manifest as a consequence of transport's impact on economic geography.

- **Transport market**

Better transport increases proximity making economic agents closer together and may also trigger economic activity as households respond to new opportunities (Laird and Venables, 2017). Typical impacts to be considered include:

- Change in travel mode: Has there been a movement towards rail travel and away from other travel modes and how is that influenced by distance from the station?
- Distance to work: Are people travelling further to work because of access to rail and how does that relate to distance from the station?

- **Labour market**

Job creation is often held up as a major impact of transport investment, with two distinct mechanisms being suggested. On the supply side, better transport may make it easier for people to get to work. On the demand side, induced investment may create new employment opportunities. Typical impacts to be considered include:

- Employment movement: Has the employment situation improved nearer the line and what other factors may have influenced that?
- Accessibility to jobs and essential services improved and how does this vary across the region?

- **Property market**

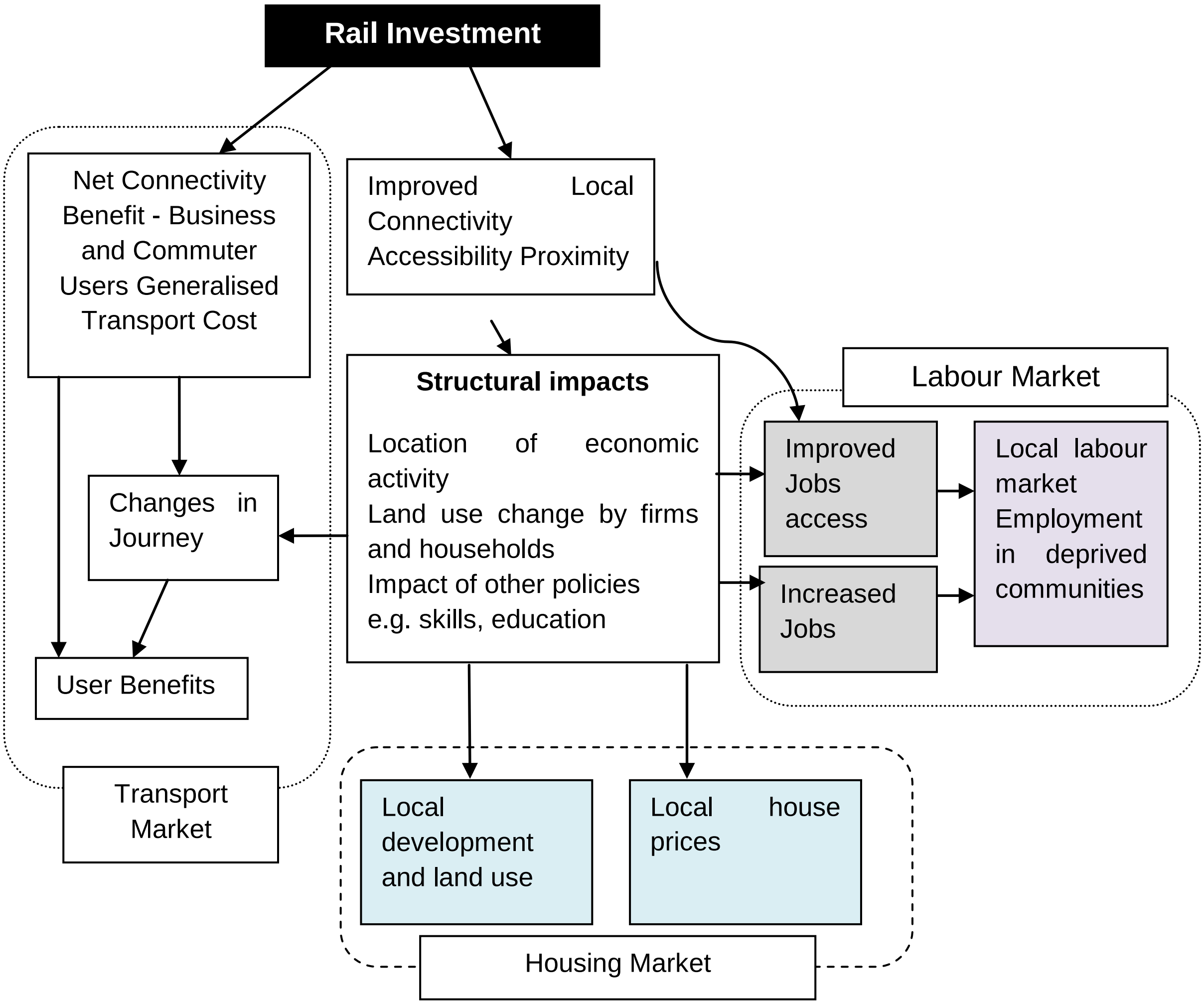
The improvement may make affected locations more attractive destinations for investment. The user-benefits experienced by residents and workers may have further value by changing the "attractiveness" of affected places and impact on the property market. However, the land market is an imperfect place to measure investments (Mohring, 1993) and the use of land value uplift is fraught with difficulties associated with addressing displacement effects and netting out the land value uplift from construction costs and the effects of speculation. Typical impacts to be considered include:

- House price changes: Has the rail investment led to an increase in house prices nearer the line and what other factors may have contributed?

Modelling approaches

There is no single best approach to capturing all the location change and quantity effects induced by a transport improvement (Venables et al., 2014). Different methods are applicable in different contexts, and should be informed by the narrative of what the most important impacts of the project are likely to be. Computer simulation models are a valuable tool, although hampered by the difficulty of understanding what key elements of the model exactly drives their results. They need to be complemented by fuller use of "bottom-up" local information and econometric approaches. The latter are not yet at a stage where they can give sufficiently project specific results, although rapid progress is being made. Nevertheless, they provide a way of assessing the plausibility of results obtained by other means (Venables et al., 2014).

Figure 4-1 Theoretical Framework



4.3 Causality concerns

To reach a meaningful conclusion on wider economic impacts, there must be sufficient elapsed time across the intervention for a comparative analysis (HM Treasury, 2011). To establish causality it is important that a counterfactual is created to allow before and after comparison between areas affected or unaffected by the intervention (treatment and control groups). This has dictated the direction of this methodology, which concentrates on effects that may become apparent over shorter period of time to allow a comparison over all three case study regions.

Given the relative remoteness of each case study and proximity to much larger urban centres, it is important to control for potential contributory external factors in addition to the rail intervention, to be confident that outcomes relate directly to the transport investment. Hence, a more detailed econometric analysis through modelling is adopted here, requiring careful selection of dependent and independent variables, and developing suitable accessibility characteristics. By comparing the situation prior to and following the intervention, assessment is made of relevant outcomes such as property price movements and job accessibility.

Importantly, accessibility reflects changes in access to jobs and appropriate occupations, and when other influences are taken into account, change in accessibility to jobs through reduced commuting journey times provides an impact. In the context of these case studies, reconnecting to the city for commuting and shopping is critical with its resultant impact on employment and the local economy through property prices. Similarly, increases in property prices when related to proximity to the new infrastructure represent an impact of the intervention, if separated from potential confounding factors such as changes in the global economy.

Spatial and temporal measures are also key to this process to avoid concentrating solely on the immediate area of economic impact. As the effect of an intervention will differ spatially and over time, the methodology is structured to reveal any heterogeneity across locations within each case study region where local variations become apparent through study of neighbourhood characteristics of different parts of the region.

4.4 The methodological structure

The methodology generates findings for comparison and synthesis across three case study regions, and encompasses an innovative approach to appraisal methodology through the application of existing statistical methods in the particular context of remote of previously rail-disconnect regions. Following a survey of stakeholders in the Scottish

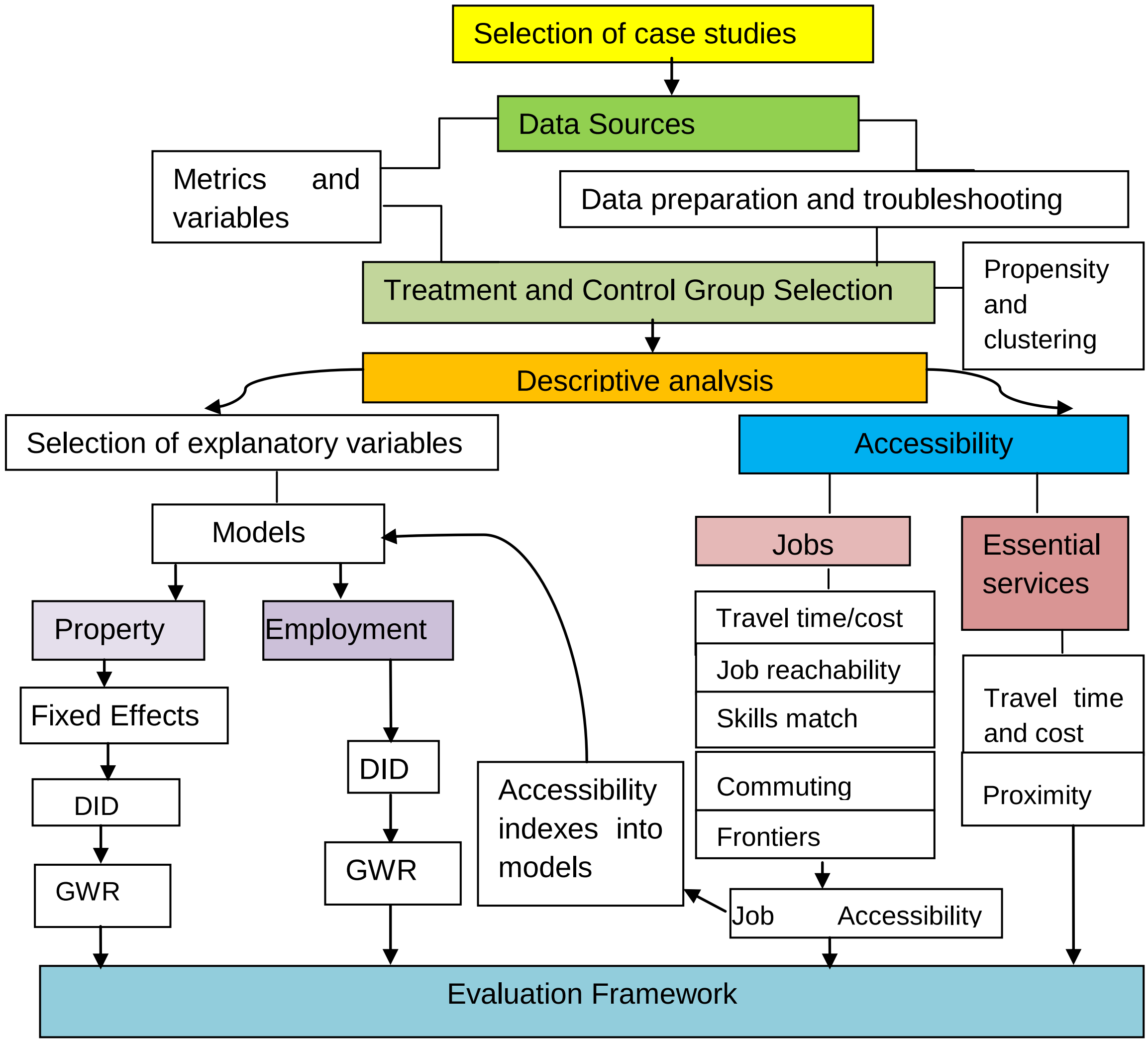
Borders prior to the reopening of the new Borders Rail line between Edinburgh and Tweedbank, several important issues emerged regarding this specific context, and provided a direction for the measures which are considered in this research study.

The methodology is specifically designed to establish causal rail infrastructure impacts in remote and disconnected regions by concentrating on:

- Treatment and control group selection to differentiate effects and establish causality.
- Construction of accessibility indicators to measure temporal movement appropriate to the remote context.
- Applying those indicators as accessibility characteristics in econometric models of property price and employment.
- Modelling impacts on jobs and property before and after the intervention.
- Exploration of heterogeneity in more remote regions through techniques such as GWR and clustering.

A simplified schematic overview of the methodology is illustrated in Figure 4-2.

Figure 4-2 Schematic overview of the methodology



The methodological structure consists of several sequential phases which build upon each other:

Case studies

Three suitable case studies (Robin Hood Line, Stirling-Alloa Link and Borders Rail) are selected to allow comparison of different types of region at various stages in intervention (4.6 *Selection of case studies).*

Data sourcing and analysis

Relevant data is then collated for each case study region, mainly though established secondary sources *(*4.7 *Data Sources*)*,* covering a period of time spanning each intervention, to allow comparison before and after the rail opening. Using a range of data sources including the UK census and Land Registry, these are collated into a disaggregated database of property prices, and an aggregated database for econometric modelling and assessment of temporal and spatial impacts. The data collected is then used to provide context by investigating the current situation and contextualising relevant characteristics through a geographical, socio-demographic and economic overview of each case study region, including the current transport situation and state of the property market (10.3 Data sources*)*.

Treatment and Control Groups

To appraise the impact of the rail intervention and establish causality through a meaningful counterfactual, representative treatment and control groups are set up to provide comparison pre- and post-intervention (What Works Centre for Local Economic Growth, 2015; Tischer and Shea, 1978). Selection into control and treatment groups is pivotal to the comparison process *(*4.9 *Selection of treatment and control groups*), in providing causality evidence through the application of accessibility indexes to property price and employment models. The treatment groups are constructed based on those areas subject to change in rail access across the intervention period. To cater for the economic diversity within groups due to different levels of deprivation and property prices, a further enhancement is the application of propensity score matching (Rosenbaum and Rubin, 1983) and clustering (Saxena et al., 2017). This enables a fairer comparison between the groups, highlighting the spatial effects within each region and unearthing underlying groupings of area characteristics not patently evident.

Accessibility measures

An important objective is to measure specific effects of re-linking larger urban areas to more remote areas with smaller economies. As accessibility to jobs and essential services is key to this study, suitable accessibility indexes for jobs and services are developed, specifically appropriate to more sparsely populated areas. These indexes allow comparison of accessibility both before and after the intervention *(*4.10 *Accessibility to jobs)* and have a two-fold purpose, representing both stand-alone indicators and an accessibility characteristic explanatory variable for the property price and employment econometric models. As job accessibility is seen as a major factor affecting wider impacts, a job accessibility index is calibrated to reflect each case study region. This takes into account proximity to jobs, available opportunities, skills matching and commuting practicability, and so allows comparison of job accessibility on both a time and cost basis and matches the jobs available with local skills. This is further analysed by treatment and control groups, and includes a spatial analysis of accessibility across the region (4.10 *Accessibility to jobs*).

Essential services

A similar but narrower development creates an essential services accessibility index which measures access to hospitals, schools etc. which are only available in a limited number of locations *(*4.11 *Accessibility to essential services).*

Property price modelling

In following another key objective of this study, to appraise the impact on residential property values four approaches are applied. Firstly, a basic descriptive approach compares property prices movements and other individual characteristics broken down by treatment and control group for a period spanning the rail intervention. Secondly, difference-in difference modelling allows some fixed effects to be nullified, and provides a comparison between control and treatment groups to predict what may have happened in the absence of the intervention.

Thirdly, for a more robust evaluation of causality, a fixed effects hedonic model is also applied to model property prices across the period spanning the intervention temporally and spatially *(*4.12 *Impact on residential property values)* and further analysed by treatment and control group. Various model variations are tested for feasibility, and explanatory variables consolidated with the addition of spatial and temporal dummy variables (*0).* Finally, as the property price models assume a homogenous effect across the whole of each region, localised heterogeneity and diversity are analysed comparing two OLS cross-sectional models over separate years spanning the intervention using geographically weighted regression (*4.14 Exploring heterogeneity*).

Employment modelling

Another key objective was to assess the impact of the rail intervention on jobs and employment (*4.13 Impact on jobs and employment*).

Firstly, a basic descriptive approach compares employment levels and other individual characteristics broken down by treatment and control group for a period spanning the rail intervention. Secondly, a model of employment density[2] has been structured through selection of explanatory variables broken down by treatment and control group (*4.13.2 Employment modelling*) for difference-in-difference applied as for the property models. However, there was insufficient data on employment to populate panel data for a fixed effects model.

---

[2] Employment density is a measure of the number of jobs in each location divided by the location area so that in sparser areas this reflects the relative size of the job market and hence is more applicable to remote regions.

Analysing heterogeneity

As the property price and employment models assume a homogenous effect across the whole of each region, they are further expanded to investigate localised heterogeneity and diversity into a local model using geographically weighted regression (GWR) (*4.14 Exploring heterogeneity)*

The final evaluation framework

As a final stage, and evaluation framework is assembled and all outcomes generated for each case study region are then carried forward for synthesis and cross case study comparison to determine factors common to all the case studies.

4.5 Academic contribution

The methodology makes an academic contribution to the existing literature in the following aspects.

- It specifically addresses remote, rural and disconnected regions whereas most studies are US-based and apply in a city or more intensely urban environment.

- It tests the application of range of methods of selection to treatment and control groups to address the remote context of different case studies, adapting methods such as cluster analysis and propensity matching to the particular situation. In so doing it investigates the sensitivity of predicted impacts to group selection adopting a Maryland Scientific Method Scale (SMS) Level 3 approach (Sherman et al., 1997). This implies a comparison of outcomes in a treated group before and after an intervention with a comparison group providing a counterfactual.

- Adapting previous index models it develops a job accessibility index which can stand alone as an accessibility model, or be incorporated into econometric models and is feasible in remote regions in terms of commuting and commuting practicability, and allowance for skills mismatch and job competition. It is tailored to this particular remote regional application in allowing for the 'thin' labour market, the infrequency of transport and the length of commute.

- The context of this study considers regions where there has been great industrial decline and consequently job skills may not match those required in the city.

- It addresses spatial and temporal impacts and heterogeneity peculiar to remote regions, in particular through application of Geographically Weighted Regression (GWR) to produce a local model which allows for different relationships to exist at different points in space and time. This is extended to examine cross-sectional GWR across time periods to examine temporal as well as spatial differences.

4.6 Selection of case studies

After consideration of a range of potential case studies, three were finally selected.

- Robin Hood Line: Mansfield to Nottingham corridor
- Borders Rail: Edinburgh to Galashiels corridor
- Stirling to Alloa: Stirling to Fife corridor

From their attributes summed up in Table 4-1, these examples encapsulated three different types of region, and could be categorised as previously disconnected with a line re-opening, and subject to differing degrees of remoteness (*Appendix 10.1 Rural and urban areas classification*).

The potential problems with each case study region are outlined in Table 4-2, and accounted for in later comparative analysis.

Table 4-1 Case study attributes

Location	Robin Hood Line	Stirling - Alloa	Borders Rail
When opened	1993-1998	2008	2015
Elapsed time	20-25	10	3
Large urban centre	Nottingham	Stirling (Glasgow)	Edinburgh
Regional Characteristics	Urban with significant rural- a mixture of rural and medium-sized towns subject to industrial decline with the demise of the mining industry.	Urban with significant rural - a mixture of rural and small urban areas with links to Stirling.	Mostly rural - rural with pockets of industry. Some black spots such as Galashiels and Hawick suffering industrial decline.
Degree of isolation	Mansfield was previously the largest town (population 100,000) with no railway station.	Previously up to 15 miles to nearest rail link.	Previously 30 miles from rail link and disconnected for 50 years.
Major town	Mansfield	Alloa	Galashiels

Table 4-2 Potential problems with case studies

Robin Hood Line	The region represents isolation, more due to industrial decline rather than a rural context. This equates to isolation rather than remoteness, and so there could be some overspill in population and jobs from other parts of the East Midlands.
Stirling-Alloa	The region represents isolation rather than remoteness and involves just a short rail extension into the region bringing a potential link to additional areas specifically for commuting and shopping, rather than regenerating a whole region such as the Borders.
Borders Rail	Comparing pre and post intervention data, a thorough analysis of the wider benefits cannot be carried out for some years. So unlike the other case studies, insufficient time has elapsed to fully assess the long term outcomes and impacts.

4.7 Data Sources and impacts

The relevant data for each case study region covered an appropriate period spanning the intervention, to allow an assessment of impacts through comparison of the situation before and after each rail intervention. Various secondary data sources were accessed for the relevant intervention periods, and falling into five distinct categories:

- Socio-demographic
- Transport Mode
- Accessibility
- Housing
- Employment

A table of data sources (Appendix *10.3 Data sources*) indicates which data is available annually or monthly, and which from the UK census (10 yearly).

Socio-demographic

Population density and population levels were provided via the UK Census 1991/2001/2011, National Records of Scotland, Scottish Neighbourhood Statistics, and GROS mid-year population estimates. The UK Census also provided age and gender profiles, and although the census occurs only every 10 years, this neatly spanned the rail intervention period for each case study region to allow a before and after comparison. Deprivation levels are published every 3 to 4 years and extracted from the Index of Multiple Deprivation for England and Wales, and for Scotland from the Scottish Index of Multiple Deprivation (Scottish Government, 2016).

Educational and qualification levels for each region are aggregated at LSOA or data zone level (Source: UK Census, 2001/2011) and includes:

- People aged 16-19 not in full time Education, employment or training (%).

- Proportion of 17- 21 year olds entering higher education (%).
- SQA Pupil Performance on SQA at Stage 4.

This data has been collated to estimate the following comparative impacts:

- Changes in population and socio-demographic mix.
- Regional demographic mix to national.
- Pockets of deprivation and proportion of the region's data zones for the 20% most deprived.
- Level of educational attainment in the region in particular residents with no qualifications and ‘Level 4 and above’ qualifications.

Transport mode -The UK Census provided data on car ownership by area as a percentage of total households (Source: UK Census, 1991/2001/2011), and also the distribution of travel to work patterns and method of travel to work across the region (Source: UK Census, 1991/2001/2011).

Station statistics were available through the ORR Portal, traffic flow data though Transport Scotland, and analysis of bus routes and times from individual timetables and schedules. This data has been collated to estimate the following comparative impacts:

- Changes in commute time.
- Increase in commuting to larger urban centres.
- Increased use of public transport for commuting purposes, and modal shift from car to public transport, and subsequent reduction in transport cost.
- Changes in public transport network coverage and analysis of transport links post rail intervention.
- The proportion of households with no cars or vans.

Accessibility - Both the DfT and Scottish Neighbourhood Statistics (SNS) sites provide access to services statistics, and distance to work and drive times, which are published at the same time as the index of deprivation. They include:

- Drive time to GP, petrol station, retail centre and post office.
- Drive time to primary school and secondary school.
- Public transport travel time to GP, retail centre and post office.

The DfT publishes annual accessibility statistics, covering seven essential local services, from employment to primary education to food stores. These statistics are available at reasonably detailed levels geographically, allowing trends to be observed and comparisons made down to local authority level. The principal metric is the 'average minimum travel time' in minutes and by transport mode. This data has been collated to estimate changes in accessibility to essential services such as employment, education, leisure, shopping and hospital broken down by socio-demographic group.

Housing - In England, house price transaction data is available through the Land Registry, and in Scotland this can be accessed through the commercial property internet site, www.rightmove.com which uses data produced by the Registers of Scotland. The Land Registry is unable to provide house price information prior to 1995. All individual house sale transactions have been collated for each region by address and post code over the period spanning the intervention allowing a disaggregate analysis of the mix of housing types. For each transaction there is information on the selling price and basic description of the property e.g. number of bedrooms.

Other sources included:

- Housing Mix: Mix of Housing Types in Study Area, and Scotland (Source: SNS).
- Housing Tenure (Source: UK Census, 1991/2001/2011).
- Housing completions and new developments index.

Scottish Neighbourhood Statistics (SNS) supply details of new starts and completions, and census data showing the property type e.g. detached, semi-detached, etc. is available from the 2001 and 2011 census at data zone level allowing estimation of the following impacts:

- House price and house sales changes - analysis of property prices across the region, identifying areas experiencing a rise in the index of median house prices.
- Housing completions and new developments index of housing completions.

Employment

Employment data, including employment by occupation, was available through the UK Census, and more frequently published sources such as BRES, SNS and ONS. Further employment structure and jobs by occupation and sector are extracted from the Annual Population Survey and Employment Structure via BRES. Nomis provides data on the labour market, earnings and job seekers allowance claimants. SNS also supplies information on unemployment, including shift in employment levels, and Key Benefits and Job Seekers Allowance claimants. ONS annual survey of hours and earnings provides earnings by residence, and earnings by workplace. This information was collated to gain some background knowledge but because it did not sufficiently cover the periods of intervention is not detailed here.

The data was used to determine:

- Changes in total employment by regional and local areas by industry.
- Change in accessibility of jobs.
- Change in average journey times for commuters, attributable to improved accessibility.
- Impact on local employment opportunities - % percentage of the working population claiming key benefits and Jobseeker's Allowance since 2002.
- Change in unemployment levels for selected occupational groups.

4.8 Data preparation

A disaggregate property price database was set up by sourcing and downloading the individual property transactions mentioned above. OS Post Code northings and eastings reference data were cross-referenced with house price data to allow accurate measurement of straight line distance to nearby locations. Distances were measured using the "as the crow flies" calculation (Equation 12) where Northing 1=" N1" Easting 1="E1" Northing 2="N2" Easting 2="E2".

$$\text{Distance in kilometres} = \sqrt{(N1 - N2)^2 + (E1 - E2)^2}/1000 \quad (12)$$

Equation 12 "As the crow flies" distance calculation

Location data were added on all essential services in each case study region, allowing straight line distance calculation via post codes sourced from readily available data on the internet which supplied northings and eastings for each establishment. This has facilitated analysis pre- and post-intervention into treatment and control groups. Data from the property price database was combined with socio-demographic, employment, and accessibility data into an aggregated database for each case study region. The database was structured (Figure 4-3) to allow behaviour across locations (e.g. LSOA units) to be observed on an annual basis and be employed for modelling effects.

Data Zone	Year	Average Property Price	Distance to nearest station	Population	Age Spread	…
S01000833	2001	£164,416	10.14	575		
S01000833	2002	£175,000	10.14	642		
…	…	…	…	…	…	…
S01000833	2016	£177,580	2.08	1059	…	…

Figure 4-3 Aggregated database format

This aggregated database allowed characteristics only available at data zone level and above to be included in addition to price and distance elements. For this database, data was not always available for each year of the study period, (for example, deprivation levels are not published annually). Established techniques were applied to compensate for this, so that the overall framework of the datasets is modelled effectively (Peugh and Enders, 2004). After considering five different techniques, imputation was selected as it preserved all entries by replacing missing data with an estimated value based on other available information. The form of imputation was "last observation carried forward", where the dataset is sorted and the cell value immediately prior to the missing data was used to impute the missing value. This was applied to data organised in order of years as in (Figure 4-3) above.

4.9 Selection of treatment and control groups

4.9.1 Change in distance

To examine issues of causality and heterogeneity, the region is divided into various pairings of treatment and control group using the following criteria:

- Areas subject to a change in distance to a rail link due to the intervention. Here the treatment group comprises those locations where rail is now accessible up to a prescribed threshold distance e.g. 2 km, with the control group selected from the remaining locations outside that threshold where there had been no difference in accessibility to rail.

  By using a range of threshold distances from the intervention, the size of the treatment group could be varied to greater thresholds to allow for differences in rural and urban experiences to produce a set of 'base' group configurations.

- Each 'base' group is then modified to create additional groups which match treatment and control locations having similar socio-demographics and other characteristics through application of propensity matching and clustering techniques.

This echoes methods used by Gibbons and Machin (2005). (There a postcode unit was assigned to the treatment group if it experienced a fall in distance to a railway station with the opening of new stations and the new distance was less than 2 km, otherwise, the postcode unit was assigned to the control group.). Although this offers a potential "selection for treatment" method, it is adjusted to take into account factors relating to this particular study:

- The 2km threshold is arbitrarily based on previous findings in the literature and applies to densely populated urban regions and cities.
- In isolated and disconnected regions there may be still some slight impact outside the distance threshold and so allocation of these to a control group may be biased.

For each case study region, a change in distance threshold with 2km contours up to a maximum 10km radiating from the nearest station allows variation in the assignment to treatment and control groups.

4.9.2 Application of propensity score matching

To strengthen the causal argument by reducing selection bias, each of the previous base groupings is then modified through a selection procedure where locations in comparative groups are matched using propensity score matching (Rosenbaum and Rubin, 1983). Each treatment case is matched with one or more control cases having a similar set of characteristics. The matching technique uses nearest neighbour 1-to-1 matching which matches a treated unit to the nearest comparable control. The comparators used for matching are area size and the following deprivation markers from the Scottish and UK publications on deprivation:

- Health
- Education
- Crime
- Housing

Area size is used as a comparator because although data zones and LSOAs are constructed to be of similar population size, in more remote regions these zones vary tremendously in area and so area size reflects the degree of remoteness. These comparators thus allow comparison of locations in the treatment and control group pairing with similar socio-economic profiles and level of remoteness, but avoid the potential bias of using other comparators which are also included later as explanatory variables for modelling. For each case study region, histograms compare the selection "shape" before and after matching, and jitter plots help visualise the quality of the matching, where each circle on the jitter plot represents a case’s propensity score. The two matched sets of locations are then realigned respectively into the treatment and control group pairings to produce another group pairing.

4.9.3 Application of clustering

Clustering techniques (Saxena et. al., 2017) provide another alternative amendment to selection into treatment. The revealed criteria for grouping are also incorporated into treatment and control groups. Again the comparators are area size and deprivation as representative of key elements in defining a location, namely level of affluence and level of security. Two methods of clustering are applied, hierarchical and K-means to corroborate the clusters.

Hierarchical clustering produces a hierarchy of cluster solutions for each case study dataset. For each region, average linkage is adopted (where the distance between clusters is defined to be average of all the distances between all pairs of points), and for each case study a dendrogram tree structure is used to provide visual evidence of the clusters present in the data (Gower and Ross, 1969).

K-means defines dissimilarity using a distance measure from the cluster centre, and assigns observations to a fixed number of clusters (MacQueen, 1967; Hartigan, 1975; Steinley, 2006). This varies for each case study region and is determined by plotting the percentage of variance explained by the clusters against the number of clusters. Where the marginal gain drops, giving an angle in the graph the “elbow criterion”, the number of clusters is chosen at this point, and so an elbow chart is generated for each case study. This is corroborated by selecting the value to give the largest average silhouette width (which measures how similar an object is to its own cluster compared to other clusters). Applying this method to the case study regions yielded similar numbers of distinct clusters, which suggests that cluster groups offer a more reasonable comparison of impacts.

4.10 Accessibility to jobs

A key objective of this study is to measure changes in accessibility to jobs with the re-introduction of rail services. Consequently, after considering the advantages and disadvantages of various approaches proposed in the current literature, an index is developed to measure changes in access to employment and essential services across different travel modes and appropriate to previously remote and disconnected regions. It compares the ease with which a given population segment within each case study region can access job opportunities and services using different modes of transport.

Following on from (4.9 *Selection of treatment and control groups)* an initial descriptive analysis compares effects across the intervention period between the various treatment and control group selections. This entails an empirical comparative analysis using the characteristics studied previously which are applicable particularly to accessibility to jobs and services:

- Car ownership - comparing access to a car between groups.

- Changes in travel to work patterns as indicated by method used and distance travelled to work and analysis of origin and destination.
- Rail passenger usage to study uptake of rail travel in the region.
- Accessibility to essential services e.g. Hospitals, Schools, Shopping outlets, Rail stations to compare movement in accessibility over the period for each group.
- Travel time, destination and origin indicators by mode of travel.

The method used here to derive a suitable accessibility index utilised a gravity-based measure combining an attraction or "opportunity" measure reflecting the "opportunities" (or jobs available), and a spatial barrier represented by a decay measure (Allard and Danziger, 2002; Cervero et al., 1997; Sanchez et. al., 2004). In including these elements, an index for each mode of travel is weighted to be inversely correlated with proximity between locations measured in terms of travel time or cost. The measure is designed either to be stand-alone, representing accessibility at a specific moment in time, or as an explanatory variable, capable of incorporation into an econometric model as an accessibility characteristic.

4.10.1 Methodological issues

Taking into account several methodological issues cited in the job accessibility literature, the index incorporates job reachability and availability and local job competition (Bunel and Tovar, 2014), and is developed to reflect these elements in the context of regions restored to the rail network. Job reachability takes into account available transport, travel time and the cost of travel. It also considers the transport mode and timing and location of services to determine if services exist between any two locations. This is particular relevant to rural and urban/rural environments where public transport is often infrequent or involves long journeys.

Job suitability considers a qualitative match between the skill requirements of job vacancies in different locations with the individual skills of the job seekers or residents in an origin location. (This is particularly valid in the context of this study as job skills requirements in the city may not match those in rural areas or where there has been great industrial decline).

Local job competition uses a direct measure of vacancies instead of all existing jobs, but counts as competitors all employed or jobless workers as job seekers. (Because of data availability issues, actual job seekers were not feasible (Korsu and Wenglenski, 2010). Threshold effects constrains reachable jobs within case study regional boundaries (workers may apply for jobs outside their residential region), but again dataset size issues prevent this.

4.10.2 Development of the index

A simplified approach is based on a matrix of distances between locations and applying standardised cost and travel parameters. As the purpose of this accessibility measure is to estimate the impact of improvements in rail infrastructure on jobs, it is important to capture the proximity between each origin and destination location by considering time and cost as deterrents.

Distance does not offer an appropriate proximity measure for comparative evaluation as over a rail intervention period it will generally undergo very little change. Here job proximity is measured for each mode of travel using alternatively generalised travel time and travel cost, applying the value of time (VOT) and standards for transport speed and other costs accessed through WebTAG and other sources (Wardman et al., 2013), and based on a distance measure between origin and destination location. The significant difference in accessibility between travel modes is affected by the speed of transport mode, waiting and walking time, and the unit cost of travel.

There are important characteristics specifically relating to more remote areas. Firstly, infrequency of public transport, where proximity measures will not necessarily reflect accessibility for commuting to jobs, and travel may only be feasible for a limited range of activities. Secondly, public transport may serve only a limited number of destinations, and not for all travel modes. Finally, the normal thresholds for travelling to jobs may require redrawing where the introduction of rail has made commuting easier. These are weighed against the potential job pool in each destination location, matched to the skills available in the origin location. Within the regional context of this study, the practicability of commuting may still remain unaltered even after the intervention.

4.10.3 Generic Job Accessibility Index

A generic accessibility index acts as a basis for considering the above issue which:

- Applies to different travel modes.
- Is measurable at different time intervals to detect impacts.
- Combines an attraction with an impedance function.
- Either the number of jobs available at other locations or adjusted to account for skills matching provides the attraction element.
- Adopts a negative exponential impedance function based on proximity between locations in terms of generalised travel time or travel cost.
- Assesses commuting practicability is also on the feasibility of commuting to jobs.

Based on the above, the generic format for the job availability index (Equation 13) for origin location i, travel mode m, in year T is:

$$A_m(i,T) = \sum_j O_j(T)\, f\left(p_{ij}\right) / \sum_j O_j(T) \quad j = 1 \dots N \qquad (13)$$

Equation 13 Generic job accessibility index

p_{ij} - Proximity which can be represented by generalised travel time or travel cost.

$O_j(T)$ - The opportunity measure for origin location at destination location j (based on the number of vacancies or other factors) in year T.

$f(p_{ij})$ - The decay measure depending on the proximity p_{ij} between i and j.

Two 'opportunity' measures were considered:

- the total number of vacancies available at each destination location
- the number of vacancies matching the skills at the origin location

The index is 'normalised' through division by $\sum_j O_j(T)$, the total number of vacancies across the region. The calculation of the index for different years (T) allows this to be monitored at different intervals across the period of the intervention. The impedance function (decay measure) is $f\left(p_{ij}\right) = e^{-\beta p_{ij}}$ where β represents the decay constant[3], so the index is rewritten as in Equation 14:

$$A_m(i,T) = \sum_{j=1}^{n} O_j(T)\, e^{-\beta p_{ij}} / \sum_{j=1}^{n} O_j(T)) \qquad (14)$$

Equation 14 Generic job index with impedance function

The job accessibility index is then expanded to cover the key methodological issues: job proximity measures - travel time and travel cost, skills matching and local job competition, and practicability of travel mode.

4.10.4 Developing the proximity metric

As distance between locations is important in calculating generalised time and cost, prior to its application, three different distance measures were considered:

---

[3] The β parameter is estimated empirically (*4.10.12 Calibrating the index and costs for each case study)* through observation of travel patterns and varies by context for different case studies as well as different travel purposes.

1. Euclidean distance: if X = (a, b) and Y = (c, d) is $\sqrt{(a-b)^2+(c-d)^2}$

2. Manhattan (or rectangular) distance: the distance that would be travelled to get between data points if a grid-like path is followed: the sum of the differences of corresponding components. If X= (a, b) and Y =(c, d) this is |a-b|+|c-d|.

3. Network distance

More sophisticated and precise measures such as travel times create computational difficulties, hence, verification is required that simpler Euclidean distances are a good proxy for network and other distances at a regional level. For each case study region travel distances have been estimated between a selection of LSOA zones and datazones in that region. This has been validated for randomly selected journeys, by comparing calculated travel times with those available via on-line mapping providers (Google Maps). Pearson correlation coefficients calculated to assess the strength of the associations between the three distance measures show that the association between all three measures is very strong with correlations above 0.97 (Table 4-3) supporting the notion that Euclidean distances are a good approximation of the two other more specific distances for the regional scale (Figure 4-4). Therefore to keep the index as simplified as possible Euclidean-based distances are applied.

Table 4-3 Correlation of distance measures

	Euclidean	Manhattan	Network
Euclidean	1.000		
Manhattan	0.980	1.000	
Network	0.979	0.960	1.000

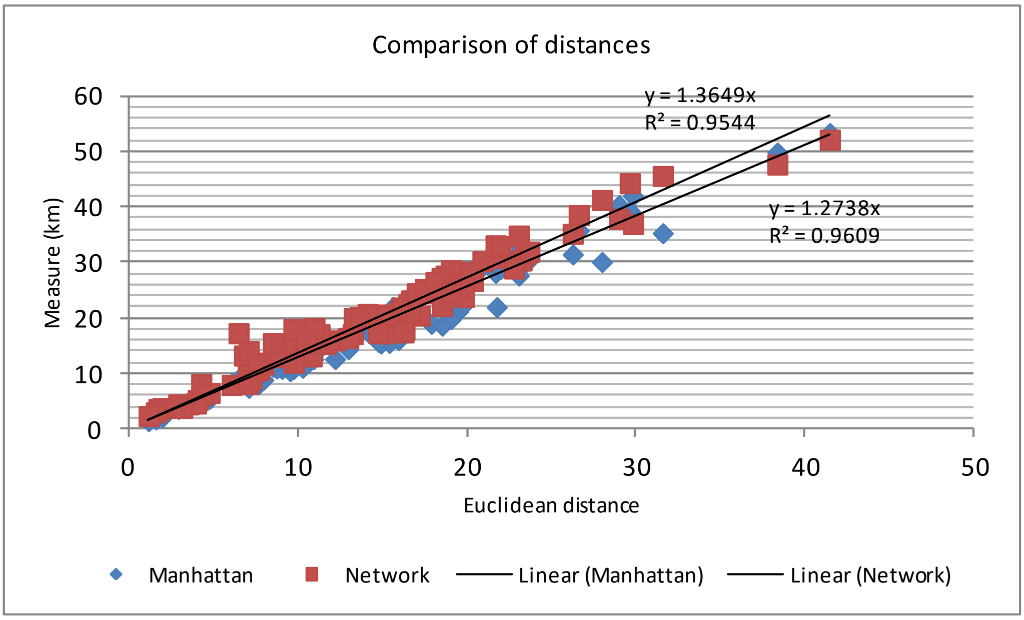


Figure 4-4 Comparison of distance measures

4.10.5 Job reachability

Job reachability is incorporated into the job accessibility index by considering the affordability, timing and location of services for each transport mode. This requires an assessment of travel time and cost, and also the feasibility of transport links between any two given locations, and defines the area within which jobs can be reached by any given employee, so that jobs more distant from the employee's residential location are less reachable than those closer.

4.10.6 Generalised travel time

In assessing travelling time between locations, key elements were the relative speeds of different transport modes and a rational measurement of the actual travel distance. Travel times were measured by applying the average speed of the transport mode to the Euclidean distance (*4.10.4 Developing the proximity metric)* between each zone's geographic centroid.

In reality, many journeys are multi-modal involving various forms of transport, but for the purpose of this simplified accessibility index, the core stage of the journey is assumed single mode to allow comparison of accessibility for various transport modes. Total travel time between locations is calculated using a combination of distance and transport speed with accessibility broken down by travel mode. A typical journey comprises four basic stages:

1. Origin to nearest stop/station (t_{WO})
2. Waiting for transport at stop/station (t_{WAIT})
3. Transport travel time (t_{RIDE})
4. Nearest stop/station to destination (t_{WD})

Origin to nearest stop/station (t_{WO}) represents the time taken to reach a bus stop or station.

- For car this is assumed to be zero.
- For bus, a default distance of 500 m is adopted at walking pace to the nearest bus stop, so t_{WO} = 0.5 km.
- For rail, except for locations within walking distance of a rail station, there will always be an element of travel at the beginning of each journey. This may vary largely, so the distance to the nearest station is assumed covered at the bus travel time and speed.

Average waiting time (t_{WAIT}) for car mode is assumed zero. For public transport, waiting time will depend on the service level (frequency) for the particular transport mode. In rural and disconnected regions, public transport service frequency may be as low as 1 per day up to a maximum of 1 per hour, and for commuting purposes, anything below 1 service per hour is probably not feasible.

Scheduled waiting time for mode m in year T, $W_m(T)$ is normally estimated as half the headway i.e. the interval between services. (For rural areas, a standard waiting time $W_m(T)$ of 15 minutes has been assumed, as were services only every 4 hours, travellers will not be waiting two hours before the transport arrives).

Transport travel time (t_{RIDE}) represents the time spent travelling in the core stage of the journey using the main transport mode based on the relevant distance for each mode. For car travel this would be the distance between start and end of the journey. For bus this would be the distance between start and end points minus 1 km (which is the total combined distance to and from bus stop assuming 0.5 km walk at each end). For rail this would be the distance between the nearest origin and destination railway stations.

Nearest stop/station to destination represents the time taken from the nearest bus stop or rail station to the destination.

- For car this is assumed to be zero.
- For bus a default distance of 500 m has been adopted at walking pace from the nearest bus stop so t_{WO} = 0.5 km.
- For rail, there will generally be some additional travel at the end of the journey. As this may vary largely, the distance from the nearest station at the destination has been assessed at bus travel time and bus speed.

Average speed of a transport mode $S_m(T)$ for transport mode m (bus, car, rail, walk etc.) in year T is based on scheduled speed and compared with empirical experiences. This would vary depending on traffic congestion between any two locations, but generalised averages are calculated for each case study region.

Total travel time between locations i and j for mode m in year T is taken as:

$TT_m(i,j,T) = t_{WO} + t_{WAIT} + t_{RIDE} + t_{WD}$ and calculated using $S_m(T)$

Table 4-4 Travel time calculation by mode for each stage of the journey

Mode	Bus (m=1)	Car (m=2)	Rail (m=3)
Origin to nearest stop/station	$\frac{0{,}5}{S_0(T)}$	0	$\frac{dns_i}{S_1(T)}$
Waiting for transport at stop/station	$W_m(T)$	0	$W_m(T)$
Travel time in transport	$(d_{ij}-1)/S_m(T)$	$d_{ij}/S_m(T)$	$ds_{ij}/S_m(T)$
Nearest stop/station to destination	$\frac{0{,}5}{S_0(T)}$	0	$\frac{dns_j}{S_1(T)}$
Total travelling time $TT_m(i,j,T)$	$W_m(T) + (d_{ij}-1)/S_m(T) +$	$d_{ij}/S_m(T)$	$W_m(T) + \frac{dns_i + dns_j}{S_1(T)} +$

Table 4-4 summarises the travel time calculations for all three modes.

- d_{ij} - Generalised distance between i and j.
- $S_m(T)$ - The average speed for travel mode m in year T.
- dns_i - Distance to nearest station from origin dns_j = distance to nearest station from destination.
- ds_{ij} - Distance between nearest stations to i and j.

These are then applied to the impedance function and weighed against 'attraction' of job opportunities, to update job accessibility for mode m at location i in year T to:

$$A_m(i,T) = \sum_{j=1}^{n} O_j(T) f(TT_m(i,j,T)) \,/\, \sum_{j=1}^{n} O_j(T) \qquad (15)$$

Equation 15 Job Accessibility Index based on generalised time

The impedance function $f\ (TT_m(i,j,T)$ equates to $e^{-\beta TT_m(i,j,T)}$ using the travelling times calculated for each mode as in Table 4-4 to the following for each mode.

4.10.7 Generalised travel cost

Another key proximity factor is the relative cost of different travel modes. Total costs can be defined as the generalised journey costs shown in Equation 16 (Balcombe et al., 2004).

$$GC = P + U(m) \qquad (16)$$

Equation 16 Generalised cost formula

where *P* is the sum of monetary costs and $U(m)$ represents the non-monetary (time) costs of a journey for transport mode m at time T. $U(m)$ can be calculated as the product of the total generalised travelling time of the journey ($TT_m(i, j, T)$) - from the previous calculation of generalised travel time (Table 4-4), and the opportunity cost of the traveller's time value of time (*VOT*), so that:

$$GC = P + U(m) = VOT * TT_m(i,j,T) \quad (17)$$

Equation 17 Generalised cost with value of time

Certain assumptions are made here. Travel cost per mile will not be constant, and average speed may vary depending on congestion and length of trip. When comparing car to public transport, transport costs for a car could incorporate fuel costs, insurance, depreciation etc., and public transport costs may relate only to the ticket fare for the distance travelled, but would be further complicated by considering concessionary travel. The cost of travel could also include other factors such as car ownership and percentage of household budget for transport costs, but in the interests of simplification they are not included in this index. For transport mode m in year T, assuming a unit transport cost $C_m(T)$ and using the distance calculations used in Table 4-4, the monetary cost and non-monetary part of the journey are combined to give a generalised travel cost $TC_m(i,j,T)$ as shown in Table 4-5.

Table 4-5 Generalised cost calculation by transport mode

Bus	$C_1(T) * (d_{ij} - 1) + VOT * TT_1(i,j,T)$
Car	$C_2(T) * d_{ij} + VOT * TT_2(i,j,T)$
Rail	$C_3(T) * ds_{ij} + C_1(T) * (dns_i + dns_j) + VOT * TT_3(i,j,T)$

The job accessibility using mode m for location i in year T can be updated to:

$$A_m(i,T) = \sum_{j=1}^{n} O_j(T) \frac{f\left(TC_m(i,j,T)\right)}{\sum_{j=1}^{n} O_j(T)} \quad (18)$$

Equation 18 Job accessibility index based on generalised cost

4.10.8 Commuting practicability

In comparing accessibility for various travel modes (Korsu and Wenglenski, 2010) consideration must be given as to whether:

- Each travel mode is available between origin and destination.
- There are feasible multi-mode combinations between origin and destination.

In remote or disconnected regions, public transport often runs infrequently, and so the timing of services for different purposes becomes crucial. The availability of travel between any origin-destination pair and infrequency of public transport services impact critically on the accessibility measure. For comparative purposes, the index provides a measure of 'commuting practicability' by combining the availability of travel mode with feasible multi-mode journeys. By considering all possible routes between any two locations, only combinations both available and feasible will count towards commuting practicability.

Post-intervention, various transport mode combinations may then become both available and feasible for commuting. Consequently, a commuting practicability variable $\varphi_m(\mathrm{i}, \mathrm{j}, \mathrm{T})$ is added into the attraction element for each origin destination pair (i, j) at time T which would equal 1 if both availability and feasibility between origin and destination equal 1 and otherwise equal zero. So that even were there job vacancies at a destination, with $\varphi_{\mathrm{m}}(\mathrm{i}, \mathrm{j}, \mathrm{T})$= 0 they would not contribute to the attraction value. The generic job accessibility index then becomes as in Equation 19:

$$A_m(i,T) = \sum_{j=1}^{n} \varphi_m(i,j,T)\, O_j(T)\, f(p_{ij}) \,/ \sum_{j=1}^{n} O_j(T) \qquad (19)$$

Equation 19 Allowance for practicability in generic index

4.10.9 Thresholds

Threshold effects arise when reachable jobs are constrained to those living within the regional boundary. Travel thresholds or “frontiers" are an arbitrary measure representing the limits that workers are willing to travel which could be further in rural and remote areas. According to the National Travel Survey (NTS, 2016), between 2011 and 2014, miles travelled per head was 80% more in the smallest settlements and rural areas than in the Greater London Built-up Area.

The threshold distance for which potential accessibility value reaches zero is defined as the maximum travel distance observed for all commuters for each region, aggregated over all transport modes. This differs by case study region and transport mode, where people will travel further for a faster travel mode. However, because the negative exponential function is short tailed, long distances will have limited effects on estimation of accessibility, and truncation does not lead to an important loss of information.

The shaded areas in Table 4-6 represent trips where total travelling time exceeds a 75 minute threshold which is based on aggregating UK Census Travel to Work information over the case study regions. This extends the potential model to allow that commuters not only have access to opportunities in the area where they reside, but also in the area where they work and post-intervention they may be prepared to travel greater distances.

Table 4-6 Table of travel times

Service Frequency	Waiting time	Time to transport	Time from transport	Transport travel time				
				15m	30m	45m	60m	75m
				Total Travelling Time				
10 m	5	10	5	35	50	65	80	95
12 m	6	10	5	36	51	66	81	96
15 m	8	10	5	38	53	68	83	98
20 m	10	10	5	40	55	70	85	100
30 m	15	10	5	45	60	75	90	105
Hourly	30	10	5	60	75	90	105	120
2 hourly	60	10	5	90	105	120	135	150

4.10.10 Job suitability - occupational matching

Reachable jobs may not be appropriate for every worker as individual characteristics determine the matching of jobs and workers, and job accessibility depends on the number of matching competitors. For local job competition, the workforce faces competition not only from workers living within its own residential region, but also from those outside its boundaries. Current jobs by occupation of residents at the origin location reflect its skills profile, and job vacancies by occupation reflect 'opportunities' at each destination location.

Job suitability refers to the possibility of a qualitative match between the skill requirements of the job offers and the individual skills of the job seekers (Ihlanfeldt and Sjoquist, 1998). An element of occupational matching reflects the relationship between job accessibility and employment for each case study region. Building upon the work of Wachs and Kumagai (1973), this index incorporates theory from Cervero et al. (1998) where conditions like occupational mismatches are explicitly accounted for in the job accessibility index.

Occupational or skills matching applies a weighting effect so that the closer the available jobs in each location j match the skills profile in the origin location, the greater the attraction. The attraction function thus depends on the origin location i and would differ for each origin. Amending for occupational matching the attraction function, based on Cervero et al. (1998) becomes as in Equation 20:

$$O_j(T) = \sum_k r_{ikT} E_{jkT} \quad (20)$$

Equation 20 Occupational matching attraction function

r_{ikT} - The percentage of employed residents in location i working in occupational class k in year T.

k = 1 (executive, professional, managerial), 2 (sales, administration, clerical), 3 (services), 4 (technical) etc.

E_{jkT} - Vacancies in location j in occupational class k in year T.

The job-accessibility index for origin location i in year T can then be refined using $O_j(T)$ in Equation 20 to:

$$A_m(i,T) = \sum_j \sum_k r_{ikT} E_{jkT}\, f(p_{ij}) / \sum_j \sum_k E_{jkT} \quad (21)$$

Equation 21 Job Accessibility Index with skills matching

An 'occupational match' accessibility index adds an important dimension to the analysis as the opportunities at location j are not considered equally available to the origin location. For any origin location i, proximity to jobs in destination location j contributes positively to the accessibility index based on the percentage of employed residents in location i matching the occupational opportunities in location j. Subtracting the standard accessibility index from this 'occupational match' index provides a 'matching effect' which indicates the relevance of occupational matching in the calculation of job accessibility.

As the estimated occupational opportunities at the origin location should never exceed the number of matching vacancies at the destination location, the calculation is modified so that the closer the match, the higher the contribution to the attraction factor. This produces the amended Job accessibility index as in Equation 22.

$$A_m(i,T) = \sum_j \sum_k \delta_{ij}(k)\, f(p_{ij}) / \sum_j \sum_k E_{jkT} \quad (22)$$

Equation 22 Amended attraction function for skills matching

$$\delta_{ij}(k) \quad = E_{jkT} \qquad if\ r_{ikT} \sum_k E_{jkT} > E_{jkT}$$

$$= r_{ikT} \sum_k E_{jkT} \quad \text{otherwise}$$

The attraction function can be further disaggregated to provide a separate accessibility index for each occupation class level. In this case the attraction function would only include jobs for a particular occupation class e.g. professional. Expanding the attraction function to $O_j(T)$ this is calculated as in Equation 23.

$$
\begin{aligned}
O_j(T) &= E_{jkT} && if\ r_{ikT}\sum_k E_{jkT} > E_{jkT} \\
&= r_{ikT}\sum_k E_{jkT} && if\ r_{ikT}\sum_k E_{jkT} \leq E_{jkT}
\end{aligned}
\tag{23}
$$

Equation 23 Attraction function for individual occupation class

4.10.11 Skills matching basis

Two alternative methods were used as the job opportunities attraction element in skills matching:

In method 1, the job opportunities were fixed on the number of jobs available in the base (pre-intervention) year and therefore was relative to the jobs profile that existed at the time. This is suggested in the literature (Gibbons et al., 2012), who advocate using original employment levels in a revised accessibility measure. This reflected a change in nearness to the rail link rather than movements in the job market. Hence, this method was adopted as more suitable for econometric modelling as it reflected the position leading up to the intervention.

In method 2, the job opportunities element was not fixed and was based on the number of jobs available in the current (post-intervention) year. Although this offered a realistic measure of the current state of job accessibility, it would not be applicable as accessibility characteristic in the econometric models as it represented a future position. Hence this method was not thought applicable to the econometric models but acted as an evaluation of the current post-intervention relative to job availability and infrastructure changes.

4.10.12 Calibrating the index and costs for each case study

A calibration process will estimate the β parameter to reflect behaviour in each case study region for different measures of proximity (Reggiani et al., 2011). The process defines two elements of the model specification: the travel impedance and the set of potential destinations, which differ for each case study region. In order to specify travel impedance for each region, two steps are required:

1. Setting the parameters of this function (e.g. the β constant).
2. Setting generalised values such as transport speed and cost

The exponential function decay parameter (β) has been determined for each case study region for both time and cost proximity measures by performing non-linear regression analysis on the non-transformed data. Using distance travelled to work data, non-linear regression analysis for several distance exponents has found the best fit. Since the probability of travelling more than any specified distance to work equals 1 at zero distance (i.e. people will always travel some distance to work - apart from home workers), the constant term is not included in the regression model which took the form $Q = e^{-\alpha D \beta}$ where Q is the probability for travel more than a specified distance D.

For each case study region, average speed is based on scheduled speed and compared with empirical experiences where available. In order to come up with a price comparison for the individual modes of public transport, a price per km is calculated is based on train, bus and car modes using the following methods:

- Train: connections for ten popular routes were analysed in terms of distance and price. The average price available was used to calculate price per km, and earlier years costs were extrapolated from the train fares price index.
- Bus: Ten popular connections were analysed in terms of distance and current price. The average price available was used to calculate price per km. The cost for earlier years was extrapolated using the bus fares price index.
- Car: Rather than use estimates of actual costs of motoring, a perceived cost was calculated omitting depreciation as being an "invisible" cost, but including fuel cost plus an allowance for overhead and maintenance costs which are assumed the same for all case study regions. Historical UK fuel prices for years 1991 to 2017 were factored up to allow for changes in everyday running costs based on an annual average of 16000 km at 56 km per gallon.
- Walk: Total walk time adopts the standard walking speed of 4.8 Km per hour or 80 metres per minute as suggested by Wu and Hine (2003).

For each case study region, the job accessibility index is calculated separately based on a travel time and travel cost basis using the calibrated values applicable. There is further subdivision dependent on the application of matching comparing the effect of skills matching based on an impedance of travel time and cost, before and after the intervention. For each case study region, there is a separate comparison based on the treatment group and control groups generated in (4.9 *Selection of treatment and control groups*). The variations in the job accessibility index are displayed in map format allowing more detailed spatial analysis to be carried out for each case study region.

4.10.13 Sensitivity Analysis

Considering the assumptions made, a sensitivity analysis to test the robustness of the index has determined where compromise in construction of the index may relate to sources of uncertainty in the input parameters. The primary test focuses on the most sensitive parameters to increase understanding of the relationships between input and output variables in the index. Core parameters tested were:

- Decay coefficient β
- Transport speed $S_m(T)$
- Headway $W_m(T)$
- Unit cost of travel $C_m(T)$

The testing method adopted is one-at-a-time (OAT/OFAT) where changing one input variable, whilst keeping others at their baseline values, sensitivity is measured by monitoring changes in the output (Czitrom, 2012). However, the OAT approach does not allow for the simultaneous variation of input variables, so it cannot detect the presence of interactions between input variables. Using the standard impedance function: $f\left(p_{ij}\right) = e^{-\beta p_{ij}}$, each core parameter has been varied separately whilst keeping all the other parameters constant (K_1 and K_2 are constants)

- Decay coefficient β - the function tested (all modes) was $e^{-K_1\beta}$
- Transport speed $S_m(T)$- the function tested was:
- Bus/Rail $e^{-(K_1 + K_2/S_m(T))}$
- Car $e^{-K_1/S_m(T))}$
- Headway $W_m(T)$ - the function tested (all modes) was: $e^{-(K_1 W_m(T) + K_2)}$
- Unit cost of travel $C_m(T)$ - the function tested (all modes) was: $e^{-(C_m(T) * K_1 + K_2)}$

The results of applying this method are shown in Table 4-7.

Table 4-7 Sensitivity analysis for impedance function

Sensitivity of impedance function to 1% increase					
Parameter	Mode	Average	Maximum	Minimum	SD
Decay coefficient	All	-0.9468%	-0.9094%	-0.9851%	0.0259%
Transport speed	Bus/Rail	0.0001%	0.0004%	0.0000%	0.0001%
	Car	0.0005%	0.0030%	0.0000%	0.0010%
Headway	Bus/Rail	-0.1281%	-0.1231%	-0.1333%	0.0035%
Unit cost of travel	All	-0.1281%	-0.1231%	-0.1333%	0.0035%

This indicates the percentage change in impedance function for a 1% increase in each core parameter and shows that variations in the transport speed are not as critical as for the other parameters. The most sensitive parameter is by far the decay coefficient, and headway and unit travel cost are also significant.

4.11 Accessibility to essential services

A suitable transport accessibility index for essential services is developed similarly to that for job accessibility. Gravity-based measures are again used to estimate the accessibility of opportunities, but represent the count of each service available at each location, which may be zero in some locations. The location's accessibility is calculated by adding together all the opportunities available in all other locations, weighted by a function of the difficulty of reaching that location (Hansen, 1959).

Data requirements are the size and placement of the services, and the travel time or distance between locations in the case study regions. In this case, arbitrary weights could have been applied to prioritise one service against another, but because of the difficulty in assigning these weightings, a separate index for each service has been developed having the following features:

- Consideration of both travel time and distance based on travel mode.
- Measurement of differences over time to highlight areas most impacted due to economic and rail infrastructure movements, and also changes in travel to essential services.
- Measurement of change in accessibility to services over years.

Again, the index reflects the type of opportunity, time period and travel mode. A generic calculation for the availability of essential services index for origin location i, service X, travel mode m, in year T is shown in Equation 24:

$$A_m(i,X,T) = \sum_{j=1}^{n} O_j(X,T)\, f\left(p_{ij}\right) / \sum_j O_j(X,T) \qquad (24)$$

Equation 24 Generic index for essential services

$O_j(X,T)$ - The opportunity measure for service X at destination location j at year T (= 1 if that service e.g. hospital exists at location j in year T, = 0 otherwise).

$f\left(p_{ij}\right)$ - The decay measure or function which depends on the proximity p_{ij} between i and j for mode m and time T, and again an exponential format was used as before.

This measure evaluates the combined effect of land-use and transport elements, and incorporates assumptions on a person's perceptions of transport by using a distance decay function similar to the job index. The measures are appropriate as social indicators for analysing the level of access to social and economic opportunities for different socio-economic groups, and may be incorporated into land use and employment models.

So as per job accessibility, the essential services accessibility of location i for opportunity x (e.g. hospital) at time T for transport mode m can be measured as:

$$A_m(i,x,T) = \sum_{j=1}^{n} O_j(x,T)\, e^{-\beta p_{ij}(m,T)} / \sum_j O_j(x,T) \qquad (25)$$

Equation 25 Accessibility index with impedance function

Using the above calculation, the essential service accessibility for each case study region is calculated for rail station, primary school, secondary school, nursery, and health services for each travel mode. This is further sub-divided into treatment group, and control group classifications as in (4.9 *Selection of treatment and control groups*).

4.12 Impact on residential property values

Four methods or modelling approaches are applied which systematically go into increasing detail to address property impacts. Initially, the descriptive approach highlights changes in relevant individual characteristics pre- and post-intervention using the treatment and control groups described as in (4.9 *Selection of treatment and control groups*).

To address the question of causality, a difference-in difference model compares before and after situations based on different accessibility characteristics through nullifying fixed effects and then examining changes between the treatment and control groups. A hedonic aggregated fixed effects model is developed using various explanatory variables including property characteristics and the local economy using the data assembled as in 4.7 *Data Sources*. This is applied using alternative accessibility characteristics including distance to rail link, change in distance from the rail link and the job accessibility index (4.10).

Finally, there is further investigation into heterogeneity and spatial diversity within each case study region using geographically weighted regression (GWR) through reference to a locally varying model (*4.14 Exploring heterogeneity*).

4.12.1 Initial descriptive analysis

A descriptive analysis considers characteristics particularly relating to property and housing. For each case study region, this is broken up into treatment and control groups as specified in (4.9 *Selection of treatment and control groups*) for the respective period of intervention. These are:

- Accommodation Type analysis - comparing property sales by house type to detect any movement in house type profile e.g. detached, terrace etc. over the intervention period by treatment and control groups.
- House Price analysis – to compare changes in property prices between the two groups for a period spanning the intervention by postcode sector for all areas adjacent to the rail link by year and house type.

The descriptive approach examines the relationship between property price and transport accessibility by isolating transport accessibility from other factors through comparisons of property price between years prior to the intervention and post intervention divided into treatment and control groups and shows the range of values and distribution of house prices for each group.

Although the rail intervention covers a specific period, there may be some lead or lag in terms of impact on property prices for the following reasons:

- There may be some increase in prices prior to the rail opening in anticipation - the "announcement effect".
- There may be a period of time before the benefits of the rail link become apparent to the housing market.

For each case study, in order to detect evidence of this phenomenon, and to define a “post intervention” cut off point, property price movement is also analysed either side of the intervention in close proximity to the rail line in the overlapping years prior to and post rail intervention.

4.12.2 Hedonic modelling

There is a large enough sample of housing market transactions for each case study region collated from Land Registry information to allow econometric modelling to separate out the implicit price of the attributes though an hedonic property price model. This model assesses the impact of transport infrastructure on land value through changes in property prices, whilst controlling for specific property characteristics, on the basis that if all the features of the property are accounted for, then the price of the property should reflect the value of the land on which the property sits. House prices are effectively standardised for internal and environmental factors so that the relevance of accessibility and location can be determined.

Two hedonic model approaches are specified using panel data, a difference-in-difference model where a comparison is made between two years either side of the intervention, and a fixed effects model using cross-sectional time series data for individual years spanning the intervention. The semi log functional form is chosen as it provided the best model fit (Bowes and Ihlanfeldt, 2001).

For the fixed effects model, in addition to transport and environmental characteristics, time dummies are included to allow for inflation and trends in the housing market and locality dummies for area-specific variations as suggested in Conniffe and Duffy (1999). These dummies attempt to capture much of the unexplained variation in house prices.

4.12.3 The panel data approach

For panel data the same variables are measured at different times providing greater scope for questions of causality and allowing for individual heterogeneity and control of variables not observed or measured. The panel data structure makes it possible to assess how various variables have changed over the intervention period, and here the movement in house prices can be measured against changes in property and location characteristics and general time effects.

The variables included in the model specification have been decided by choosing between competing models on the basis of model fit, where models included different variables with different functional forms. An investigation of collinearity and possible inclusion of too many variables has involved consideration of various models and model structures, testing, graphical output and resultant findings.

For each case study region, a correlation and regression analysis of the explanatory variables has confirmed a definite relationship between all three pairs of explanatory variables resulting in one of each pair being omitted from the model. There was significant correlation between:

- Income deprivation and Employment rate
- Employment Rate and Income deprivation - Employment deprivation - Health - Education
- Employment deprivation and Health
- Education and Health and Housing

Consequently, the final version of the fixed effects model has been adjusted to include the dependent and explanatory variables listed in Table 4-8 into the following categories: property, accessibility, neighbourhood and location characteristics and status variables. The accessibility variable includes distance to nearest station, change in distance ratio to station and job accessibility index.

Table 4-8 Model explanatory variables

Characteristic	Variable Name	Description	Measure
Property	Terrace	Number of terraced Properties	1 = terraced 0 = otherwise
	Detached	Number of detached Properties	1 = detached 0 = otherwise
	Semi	Number of semi-detached properties	1 = Semi 0 = otherwise
Accessibility	Distance ratio	Distance ratio	%
	Nearest Station	Distance to nearest Station	km
	Job Accessibility index	Job Accessibility Index	Between 0 and 1
	Pop Density	Population density	population per sq km
Neighbourhood environment	Full Time	Number of people in full time employment	employees
	No Cars	Number of households without a car	households
	1 car	Number of households with 1 car	households
	Level 1	Population at basic level (Level 1) education	residents
	Level 4	Population at highest level (Level 4) education	residents
	% employed	% of population employed	% number employed /Work population
Location and time dummy variables	Location	Location specific dummy variable	1 = Location i 0 = otherwise
	Time	Time specific dummy variable	1 = Year t 0 = otherwise
	Post	Post intervention year	0 = pre 1 = post

4.12.4 Difference-in-difference model

Difference-in-difference (DID) estimates the effect of a treatment (i.e. an explanatory variable) on an outcome (i.e. a dependent variable) by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group. It requires measuring data from treatment and control groups from two or more different time periods, before and after 'treatment'. DID allows estimation of the effect of the intervention by comparing average change over time in property price variable for the treatment group and the control group. The aim here is to study the effect of the rail intervention across a cross-section data zones over at least two time intervals using data measured at different time periods. Such an approach highlights differences across various data zones and indicates the degree of heterogeneity.

The difference-in-difference model was run using the different variables of accessibility mentioned in the previous section. The difference-in-difference regression model (Equation 26) was derived from differencing the standard fixed effects model shown later (Equation 27) where suffix 0 refers to pre-intervention and suffix 1 refers to post-intervention.

$\Delta A = A_{i1} - A_{i0}$ Change in accessibility to nearest station

$\Delta X = X_{i1} - X_{i0}$ Change in property and location characteristics

$\Delta g = g_1 - g_0$ Change due to time effects

$$(ln\, p_{i1} - ln\, p_{i0}) = \Delta A\beta + \Delta X\gamma + \Delta g + (\epsilon_{i1} - \epsilon_{i0}) \qquad (26)$$

Equation 26 Difference-in-difference model

The difference-in-difference model compares two separate years spanning the rail intervention by "differencing" the fixed effects model used previously, effectively cancelling out any fixed location effects. Using panel data for each case study region, a comparison has been made between pairs of selected years to detect and measure any level of impact. By differencing across different periods, most of data for the intervening years is ignored, and so consequently trends and detailed movements in house price are overlooked. As with the fixed effects model specified later, three accessibility characteristics are applied - change in distance to nearest station, change in distance to station ratio and change in job accessibility index.

4.12.5 Standard fixed effects model

Fixed-effects (FE) models analyse the impact of variables that vary over time and explore the relationship between predictor and outcome variables. Each entity has its own individual characteristics that may or may not influence the predictor variables. An important assumption of the FE model is that time-invariant characteristics are unique to the individual and should not be correlated with other individual characteristics

(Kohler and Kreuter, 2012). This model applies across the entire case study region, and house prices are effectively standardised for internal and environmental factors so that the relevance of accessibility and location can be determined.

The fixed effects model using log price as the dependent variable was preferable to a random effects model as corroborated by running various comparison tests including the Hausman test. Using samples of data from each case study region, it also shows the presence of time-fixed effects (F test and Lagrange Multiplier Test), cross-sectional dependence and homoskedacity (Breusch-Pagan test), and serial correlation (Breusch-Godfrey/Wooldridge and Augmented Dickey-Fuller Test). There was also analysis to detect the presence of spatial autocorrelation or spatial heterogeneity when unexplained variance may be caused by interdependence between observations as a result of their relative location in space.

Where i denotes the location index (or the one by which data is grouped) and t the time index, the format of the standard fixed effects model is:

$$\ln p_{it} = \alpha_i + A_{it}\beta_1 + X_{it}\gamma + f_i + g_t + \epsilon_{it} \qquad (27)$$

Equation 27 Standard fixed effects model

- p_{it} Price of a property at location i in year t (dependent variable).
- A_{it} Accessibility matrix e.g. distance to nearest station, accessibility index.
- X_{it} Vector of other property and location characteristics e.g. housing types.
- f_i Represents location-specific unobserved components fixed over time.
- g_t Represents general time effects.
- ϵ_{it} Represents the error term.

Three accessibility characteristics are applied to A_{it} :

- *Distance to the nearest station* (DNS) is the distance to the nearest station which will change between pre- and post- intervention dates for those locations subject to rail structure investment, but remain unaltered otherwise. This is a basic measure applied in the literature (Gibbons and Machin, 2005), but does not account for the time or cost in reaching the station.

- *Distance to station ratio* indicates the improvement in accessibility, expressed as a ratio, where there has been a change over time in distance to the nearest station due to new rail infrastructure. For example, where the location was previously 20 km from a railway station and is now 5 km, the improvement is 15 km and the distance ratio is thus 15/20 or 0.75. This again ignores time or cost in reaching the station, but indicates a relative improvement in access to the rail network. The ratio is $(d_{io} - d_{it})/d_{io}$ where

  d_{io} - The initial distance from the station (year 0)

  d_{it} - The distance to the nearest station (year t)

For no change this will be 0 and will vary between 0 and 1, approaching 1 for large changes in distance.

- *Job accessibility*. Accessibility has long been accepted as a central driver of property values, often represented simplistically, as in the distance from a single Central Business District (CBD) (Alonso, 1964). A deeper understanding of the relationship between accessibility and property values determines how households value travel time and diversity in job destinations, and predict how the housing market will respond to changes in transport infrastructure. The accessibility index is covered in (4.10 *Accessibility to jobs*) which takes into account the decay effect as well as commuting possibilities.

The last two factors represent:

- *Location specific fixed effects over time* (f_i) expressed as $\mu_2 D_{2i} + \mu_3 D_{3i} + \ldots$ where D_{ji}= 1 if the observation belongs to location j and 0 otherwise. The number of dummies equals the number of locations minus 1 to avoid the dummy variable trap.
- *Time effects* (g_t) expressed as $\lambda_1 Dum_{Year1} + \ldots + \lambda_{13} Dum_{Year(N-1)}$ where Dum_t takes a value of 1 for an observation in year t and 0 otherwise. Year N is taken as the base year to avoid the dummy variable trap.

Assuming constant slope coefficients, but with an intercept that varies over locations as well as time the finalised model structure used in Equation 28 was:

$$\ln p_{it} = \propto_i + (A_{iT}\beta_1 + X_{it}\gamma + (\mu_2 D_{2i} + ..) + (\lambda_1 Dm_{1993} + ..) + \epsilon_{it} \quad (28)$$

Equation 28 Finalised fixed effects model

In terms of model runs there is a comparison of output based on the treatment and control groups mentioned earlier.

4.13 Impact on jobs and employment

Another key objective is to examine the effects of rail infrastructure changes on jobs and employment. Four methods or modelling approaches are applied which systematically go into increasing detail to address property impacts.

Initially, the descriptive approach highlights changes in relevant individual characteristics pre- and post-intervention using the treatment and control groups described as in (4.9 *Selection of treatment and control groups*).

As before this led to the development of an econometric model to assess impacts on employment levels. This initial model assesses the impact of transport infrastructure on employment by concentrating on changes to employment characteristics. The model takes into account various explanatory variables including house characteristics, the local economy, and accessibility characteristics. As before a difference-in difference model using panel data format has been applied and results analysed by treatment and control group. A definitive model then emerged through consolidation of explanatory variables and spatial and temporal dummy variables which led to further considerations of heterogeneity and spatial diversity by the application of geographically weighted regression (GWR) techniques (*4.14 Exploring heterogeneity).*

4.13.1 Initial descriptive analysis

A descriptive analysis considers characteristics particularly relating to jobs and employment. For each case study region, this is broken up into treatment and control groups as specified in (4.9 *Selection of treatment and control groups*) for the respective period of intervention. This has entailed an empirical comparative analysis using the characteristics particularly relevant to jobs and employment. These are:

- Education levels and qualifications
- Job Seekers Allowance by LSOA[4]

---

[4] Jobseekers allowance provides an alternative to income or employment deprivation, and continuous data is available for all years between 2004 and 2017.

- Economic activity
- Job profile by industry and occupation

The comparison method examines the relationship between employment and transport accessibility by isolating transport accessibility from other factors. Simplified comparison methods are used to evaluate impacts where changes in employment levels in the treatment group are compared with the employment density changes in the control group, but this does not identify the more complicated, multi-dimensional features that underpin employment movements.

4.13.2 Employment modelling – Difference-in-Difference

For employment modelling, owing to limitations of data, it was decided not to run a fixed effects model as it was not possible to effect a meaningful comparison of output across the three case study regions. Whereas for the property modelling there was annual data on property prices from 1995 onwards, although there was more detailed employment data from 2004 onwards, prior to that date, information was restricted to the census years only and the UK census was the only available source for comparison purposes.

A definitive model has been developed using:

- panel data: representing a cross-sectional time series spanning the intervention
- Difference-in-difference where a comparison is made between pairs of years either side of the intervention.

Employment density is defined as the total number of filled jobs in a location divided by the land area of that location. The total number of jobs is a residence-based measure of jobs and comprises employees, self-employment jobs, government-supported trainees and HM Forces. Employment density has been used here rather than employment level; it offers a better representation of the 'thin' labour market in much of these case study regions, and avoids the potential correlation with population levels.

The relationship between employment density and transport accessibility is modelled by standardising for a number of attributes in a global multiple regression model with the dependent variable of employment density as applied in Gibbons and Machin (2003). This relationship is assumed to hold everywhere in the case study region, but in reality often varies in the presence of spatial effects, which become evident when mapping the residuals of the multiple regressions where distinct spatial patterns can be identified.

The global model relates the dependent variable of log employment density to a group of variables:

$$ln(y_{it}/a_i) = \beta_0 + A_{it}\beta_1 + X_{it}\beta_2 + f_i + g_t + \epsilon_{it} \qquad (29)$$

Equation 29 Employment global model

y_{it} - Average employment in location i in year t (annual figure).

a_i - The land area of the location.

y_{it}/a_i - Employment density.

A_{it} - A vector of station distances or other accessibility characteristics.

X_{it} - A vector of other zonal characteristics – specifically population density and distance to the central business district (CBD).

The explanatory variables used in this model are divided into the following categories as detailed in Table 4-9.

Table 4-9 Characteristics used in DID employment model

Characteristic	Variable Name	Description	Measure
Property Characteristics	Terrace	Number of terraced properties	1 = terraced 0 = otherwise
	Detached	Number of detached properties	1 = detached 0 = otherwise
	Semi	Number of semi-detached properties	1 = semi 0 = otherwise
Accessibility	Dist Ratio	Distance ratio	%
	Nearest Station	Distance to nearest Station	km
	Job Accessibility	Job Accessibility Index	
Neighbourhood environment characteristics	Pop Density	Population Density	population per sq. km
	Full Time	Number of people in full time employment	employees
	No Cars	Number of households without a car	households
	1 car	Number of households with 1 car	households
	No qualifications	Population at basic level No qualifications	residents
	Level 1	Population at basic level (Level 1) education	residents
	Level 4	Population at highest level (Level 4) education	residents

The accessibility variables include distance to the nearest station, change in distance ratio to nearest station and accessibility indices as specified in 4.10 *Accessibility to jobs.* Care is taken when including the job accessibility index in the employment model to avoid double counting, and allow for the possibility that job accessibility may be endogenous to employment. However, employment density measures the level of employment for residents in each origin location, whereas the accessibility index is based on employment at locations other than the origin. This is in conjunction with the job vacancies available there, either by split by occupation or in total. Using the same methodology as in (*4.12.3 The panel data approach*) the final version of the model is adjusted to include the explanatory variables in Table 4-9.

Using Equation 29 over two separate years defined the DID model which followed a parallel logic to the property model. In this case, DID allows estimation of the change of employment density over time for the treatment and control group. The difference-in-difference model is run alternatively applying the different variables of accessibility mentioned previously. The difference-in-difference regression model (Equation 30) has by differencing the employment global model in Equation 29:

$\Delta A = A_{i1} - A_{io}$ Change in accessibility to nearest station

$\Delta X = X_{i1} - X_{io}$ Change in property and location characteristics

$\Delta g = g_1 - g_o$ Change due to time effects

$$ln\,(y_{i1}/a_i) - ln(y_{i0}/a_i) = \Delta A\beta + \Delta X\gamma + \Delta g + (\epsilon_{i1} - \epsilon_{i0}) \quad (30)$$

Equation 30 Difference-in-difference model - employment

As with the DID property price model, difference-in-difference compares two separate years spanning the rail intervention by "differencing" the employment density global model, effectively cancelling out any fixed location effects. Using panel data for each case study region, a comparison is made between pairs of selected years to measure levels of impact. By differencing across different periods, most of data for the intervening years is ignored, and so consequently trends and detailed movements in employment density are overlooked.

4.14 Exploring heterogeneity

Although the hedonic models offer a basis for explaining variations in property price and employment across on a regional basis, they give no indication of spatial variability, and the global relationship is assumed to hold everywhere across each case study region. However, in the analysis of property prices and employment, this is often contravened due to spatial effects that are evident when mapping the residuals of the multiple regressions.

4.14.1 Property prices

Therefore to explore questions of heterogeneity, a global hedonic model based on the previous models is calibrated into a more disaggregate local model to investigate the spatial variability in house prices and accessibility. Geographically weighted regression (GWR) explores how the relationship between the dependent and independent variables fluctuates geographically, and so searches for geographical differences rather than fitting a single "global" model to the entire case study region. It recognises the spatial variability of land values by first estimating a cross-sectional OLS model representing the global model, then extending it to produce a local model at each location in the region. For each case study an analysis is carried out using GWR with log of property price as the dependent variable to determine the relative impacts locally across the region. As in Brunsdon et al. (1998), this involves fitting a spatially varying coefficient regression model of the form shown in Equation 31:

$$y_i = \sum_{j=1}^{p} \beta_j(u_i, v_i)x_{ij} + \epsilon_i \quad \mathrm{i} = 1, 2, \cdots, \mathrm{n} \qquad (31)$$

Equation 31 GWR regression model

- $y_i; x_{i1} \dots x_{ip}$ - The set of observations of the response y and explanatory variables $x_1 \dots x_p$ at each location (u_i, v_i) in each case study region.
- $\beta_j(u_i, v_i)$ (j = 1, 2, · ·, p) - Regression coefficients at each location (u_i, v_i).
- ϵ_i (i = 1, 2, · · · , n) - Error terms.

Examination of the spatial variability of independent variables tests to see if some explanatory variables, though not significant in the global model, may vary significantly over the geographical area and become significant local parameters through GWR modelling. The procedure which is cross-sectional and indicates spatial variation across the region for one specific year only is as follows:

- The transport accessibility characteristics which were applied in the global model are alternatively substituted to generate separate sets of local models.
- The values of the regression coefficients are estimated and monitored using a "*moving window"* approach which sweeps over the study area and fits a local model to each location.
- The optimal search window size is calculated and the number of neighbouring locations within the search window fixed. This varies in area from location to location within each region - covering less where sample points are close together and greater where points are sparse. This is ideal for analysing census data (because census zones are of a variable size: smaller where population density is higher and vice versa).
- Estimates of the parameters are made for each data point with coordinates, and then mapped.

However, as this research study is concerned with before and after comparison the process is extended using ‘geographically and temporally weighted regression’ (GTWR) which is basically an extension of the cross-sectional GWR weighting procedure described above for multiple years. In a cross-sectional approach, the local model is calibrated using GWR for two separate years spanning the intervention, and for each year only the subset of data points for that year is included in the model. Hence a comparison can be made between coefficients before and after the intervention to observe any change in heterogeneity due to the intervention. This method has previously not been applied in the context of mainly rural and disconnected regions, and will appropriately reflect changes in local parameters over a period of time spanning each intervention.

4.14.2 Jobs and employment

As for property prices, the global model is calibrated into a more disaggregate local model. An analysis is carried out for each case study region using geographically weighted regression (GWR) with log of employment density as the dependent variable to assess the relative impacts across the region. The methodology follows that outlined in (*4.14.1 Property prices*). The variables selected for the local model mirror those adopted for the global model. Using the global model for employment density considered earlier, the transport accessibility characteristics were substituted in turn to generate separate global models as the basis for the GWR model. The GTWR (geographically and temporally weighted regression) method is again applied as in (*4.14.1 Property prices*) and will appropriately reflect changes in local parameters of the employment model over a period of time spanning each intervention.

4.15 The final evaluation framework

As a final stage, all outcomes generated for each case study region are then carried forward for synthesis and cross case study comparison to determine factors common to all the case studies.

5 Chapter Five: The Robin Hood Line Case Study

A significant proportion of the literature examines economic and social impact of rail intervention almost exclusively with reference to city or urban environments (Gibbons and Machin, 2005; Oosterhaven and Elhorst, 2003; Vickerman, 2007; Lakshmanan, 2010; Tavasszy et al., 2002). However, more recently, there has been an attempt to consider the rural or semi-rural perspective. Laird et al. (2013) specifically focused on extending methods of measuring wider economic benefits to remote rural areas often offering a lack of alternatives and choices for travel employment and suppliers. Such analyses need to encompass the complexity of rural and semi-rural issues, particularly in a diverse setting such as Nottinghamshire as served by the Robin Hood line.

This chapter addresses the impact of rail links using the case study of the Robin Hood Line, which was reopened twenty years ago and links the city of Nottingham with more remote towns which were disconnected under Beeching. The region mainly encompasses the demographically diverse county of Nottinghamshire, with its mixture of more remote areas suffering industrial decline on the one hand, and increased prosperity in the city of Nottingham on the other. This chapter applies the methodology specified in Chapter 4 contextualised through an overview of the geography of the region and its socio-demographic and economic profile and property market. Through division of the region into treatment and control groups, econometric modelling estimates the impact on property price and employment density using three alternative accessibility indicators. The sensitivity of selection into treatment criteria is analysed using different group configurations with a view to differentiating between urban and rural applications.

5.1 Robin Hood Line and its region

At the outset it is necessary to describe the unique situation of the Robin Hood line as part of the growing rail network. Nottinghamshire sits on extensive coal measures situated largely in the north of the county. The centre and south west of the county, around Sherwood Forest, features undulating hills with ancient oak woodland (Figure 5-1). Outside the city of Nottingham, the area is dominated by Mansfield as the main central town - a largely urban area with links to Nottingham and Sheffield. Historically, the region has been influenced heavily by its industrial past, with coal mining and textiles, but over the past thirty years has suffered considerable industrial decline, particularly in the mining industry,

Figure 5-1 Features of the Robin Hood Line region

(Source: World Guides.com, 2016)

As a consequence of the Beeching report (British Railways Board, 1963), all stations on the line were closed in October 1964 leaving Mansfield as the largest town in the country without a rail link. In the late 1980s, a consortium of local authorities sought to restore passenger services between Nottingham and Worksop. Subsequently, the section between Nottingham and Newstead re-opened in 1993, extending to Mansfield Woodhouse by November 1995, with the final extension to Worksop opened in 1998 (Figure 5-2).

Figure 5-2 The Robin Hood Line route
(*Source: East Midlands Trains, 2016)*

The Robin Hood Line links this previously rail-isolated region to Nottingham which involved reopening a substantial section of line and creating ten new rail stations. It has been in place for some time, hence provides an opportunity to observe wider impacts both economically and socially, particularly through changes in commuting patterns, and local economy impacts in housing and job accessibility. The region consists of communities cut off through rail station closure rather than being geographically remote, and the populations of the LSOAs are generally larger and the area covered smaller than for the other case studies. As a case study, the rail infrastructure has been in place much longer (20-23 years), and it presents a much more complex picture owing to numerous confounding factors. These include the surrounding motorway network, and a number of stations on the periphery of the region that existed prior to the intervention. There is greater potential for overspill from other parts of the East Midlands, and the city of Nottingham presents a much larger conurbation which, since the introduction of the rail link, now has its own tramway system.

5.2 Definition of case study region

For the purpose of this analysis, the Robin Hood Line region is illustrated in Figure 5-3 where the region studied is shown in white with a light grey boundary. This incorporates all LSOAs adjacent to the Robin Hood Line between Nottingham and Worksop and incorporates all of Nottinghamshire and some parts of the adjoining counties extending eastward to Newark and including the south of Nottingham city. This allowed a mixture of areas close to the rail intervention and those further away for further comparison of impacts into treatment and control groups.

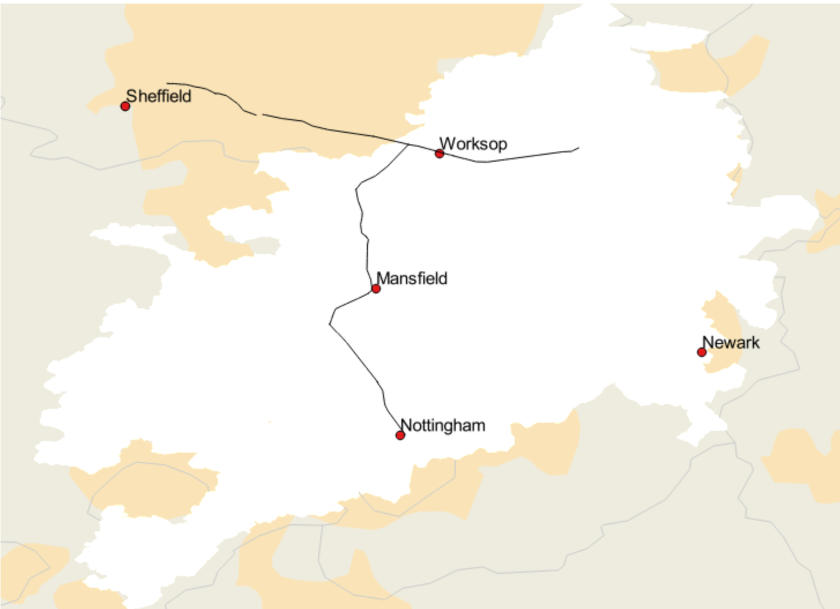


Figure 5-3 Robin Hood Line Case Study regional boundary

5.3 Population profile

The industrial decline and poor transport connections and subsequent restoration of rail links should be evident in a study of population movement. Indeed, over the past century the population of Mansfield had declined in parallel with its traditional industrial base. However, much has been done to diversify the economic base and replace jobs lost. Health is the largest sector, accounting for 14.6% of all employment, but there is still a strong manufacturing base, with 13.8% of employment being within this sector compared with 13.3% in the East Midlands and 8.4% in England. The retail sector is the third largest employer with 11.1%.

Recent mid-year population forecasts reveal that since 2011 the population had risen from 104,600 to 107,400 in 2016 (ONS, 2018), and most towns neighbouring the Robin Hood line have experienced an increase in population over the census periods (2001-2011), more noticeably Mansfield and Worksop . On the other hand, the age profile of Nottinghamshire in 2011 was slightly older than the national average, with 19% of the population aged 65+ compared with 17% in England, and this trend is predicted to continue over the next fifteen years with the number of 65-84 year olds increasing by over 30%.

The diversity across the region is apparent in the employment figures. While in 2005 the UK had 4.7% unemployment, the East Midlands 4.4%, and the Nottingham commuter belt area 2.4%. However, Mansfield has 20.2% of its working age population seeking key out-of-work benefits (NOMIS, 2016). Compared to the national level, there is a below-average proportion of residents with high skills (NVQ4 or above). Educational attainment of the workforce in the county is lower than the average for England at all levels, but particularly for NVQ4+.

5.4 The property market

Property prices are often seen as representing a weather vane for economic prosperity. The latest figures from the Land Registry House Price Index (Land Registry, 2017) show the average cost of a home in Nottingham in November 2016 is £91,446, up from £83,602 in November 2013. Meanwhile in the wider county, prices of properties were up 4.7 per cent to £126,255. As seen in Figure 5-4 Nottinghamshire house prices (1995-2016), in Nottinghamshire, the market has gradually recovered to pre-crisis levels over recent years with the exception being areas like Rushcliffe where prices started to increase earlier in the property cycle.

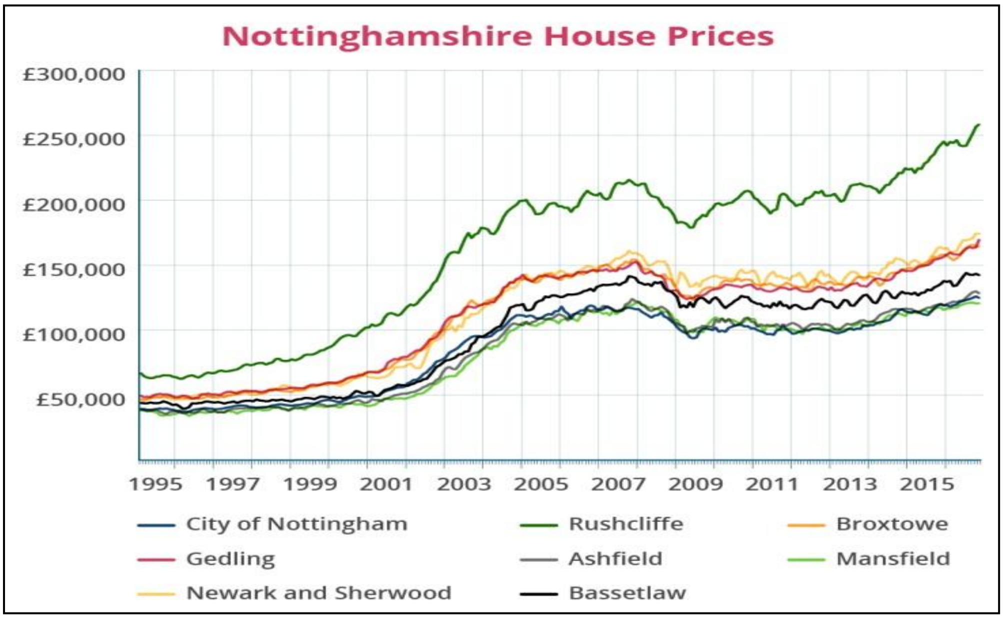


Figure 5-4 Nottinghamshire house prices (1995-2016)

(Source: Nationwide, 2016)

5.5 Transport provision

Examination of the current transport situation provides further information about the amount of disconnection in the region. Although over 90% of households in Nottinghamshire can access an hourly bus service within 10 minutes walking distance during the day (Nottinghamshire County Council, 2015), this drops to just over 70% for evenings and Sundays, and in the rural districts can be fewer than 50% of households.

Correspondingly 21% of households in Nottinghamshire (excluding the city of Nottingham) have no car, and this is more critical when broken down by population groups such as all single person households (45%), elderly people living alone (58%) and lone parent families with dependent children (33%) (Nottinghamshire County Council, 2015). Car ownership levels are lowest in urban districts such as Mansfield and Ashfield where there are higher levels of deprivation. Rural areas of Nottinghamshire have some of the highest levels of car ownership. However, residents without a car may experience difficulties in accessing services by public transport as provision is poorest in these areas.

Regional rail services link the towns and cities of the East Midlands, plus Central and Northern England. The Robin Hood Line train service runs from Nottingham to Worksop via Mansfield, Monday to Saturday and from Nottingham to Mansfield on Sundays. The Nottingham Express Transit (NET) tram system opened in 2004 with a second phase in 2015. The network consists of two lines crossing the city; Line 1 runs between Chilwell and Hucknall, and Line 2 between Clifton and Phoenix Park. There are between 4 and 8 trams per hour, depending on the day and time of day (Source: Nottingham Express Transit). There are indications that the introduction of the tram has had some impact on usage of those stations close to Nottingham such as Hucknall.

Just under 90% of households are within 15 minutes travel time by public transport from a GP surgery/health centre, and 98% of households are within 30 minutes travelling time (Nottinghamshire County Council, 2015). Access is poorer in rural areas such as Newark and Sherwood where 60-70% of households are within 15 minutes travel time and public transport frequency is lower. 77% of households in Nottinghamshire (excluding Nottingham City) are within 30 minutes travelling time of a hospital by public transport. Access is again poorer in rural areas with 43% of households within 30 minutes travel and public transport frequency is lower.

5.6 Treatment and control groups

Having placed the Robin Hood Line case study region in context, as a first part of the process outlined earlier, it is now necessary to divide the region into treatment and control groups in order to appraise the impact of the rail intervention and establish causality through creation of a meaningful counterfactual. A key objective of this research is to measure specific effects attributable to the rail intervention, and selection into control and treatment groups is a pivotal part of the comparison process in helping to provide causality evidence through application in property price and employment models.

The groups are constructed based on those areas subject to change in rail access across the intervention period. To cater for the economic diversity within groups due to different levels of deprivation and property prices, a further enhancement is the application of propensity score matching (Rosenbaum and Rubin, 1983) and clustering (Saxena et al., 2017) enabling a fairer comparison between the groups, highlighting the spatial effects within each region and unearthing underlying groupings of area characteristics that are not patently evident.

5.6.1 Selection process

As a first step in the process, the region is divided into various treatment and control group pairings using the methodology outlined in *4.9 Selection of treatment and control groups.* The purpose of considering multiple treatment group configurations is to distinguish between urban and rural methods of assessing the influence of new rail infrastructure.

Most of the literature adopts a 2 km threshold around the rail station as providing the most impact, but this is predominantly based on an urban scenario, and the analysis will examine its validity in a rural or more remote setting where perhaps a wider threshold is necessary. One treatment/control group configuration is adopted as the standard and used throughout for comparison purposes. Later in the chapter, sensitivity analysis will investigate the effect of adopting different configurations to indicate any difference from the urban situation.

5.6.2 Change in distance

For the base groupings, allocation of treatment and control groups involves a change in distance to the rail network combined with successive 2 km contours radiating from the nearest station up to a maximum of 10 km, which allows five variations in treatment group specification. For each configuration, the control group is represented by all LSOA zones not in the treatment group.

Figure 5-5 shows the different treatment groups based on the alternative methods of selection. This selection method particularly addresses the regional context relevant to this study, where impact may be experienced at a greater distance than in an urban situation. These groups will be applied later against accessibility indices and the property and employment models for a comparative analysis.

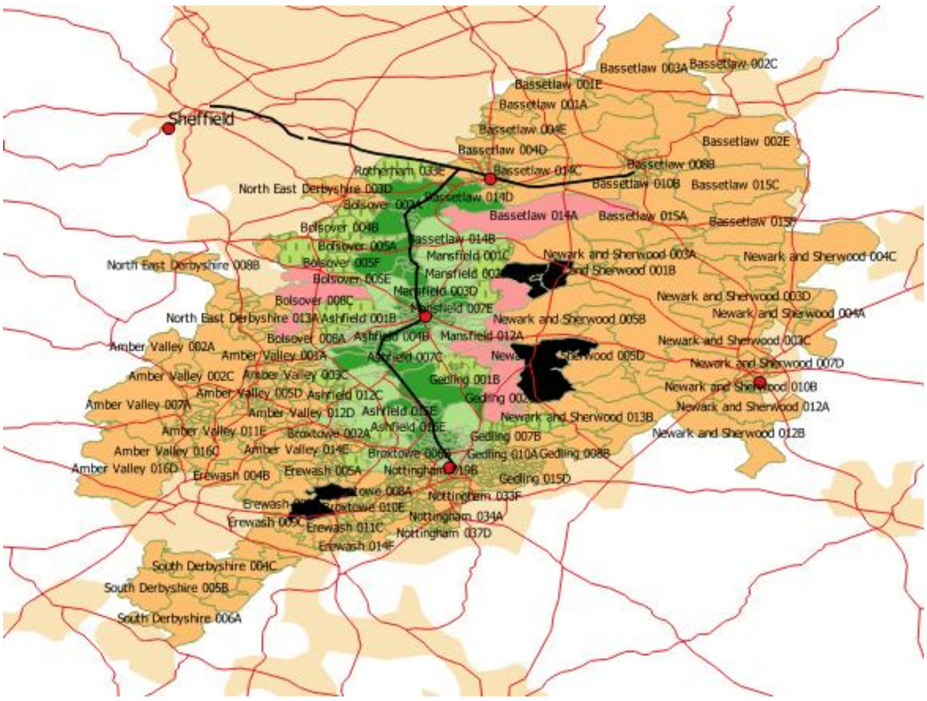


Figure 5-5 Robin Hood Line - variation in treatment and control group allocation

	< 2km
	2km-4km
	4km-6km
	6km-8km
	8km-10km
	> 10km

The shading used in Figure 5-5 is shown here where each colour represents an extension to the treatment group areas.

5.6.3 Propensity score matching

It is important to recognise the limitations of using distance as the sole basis for grouping as more affluent districts in the treatment group may be being compared to more depressed areas in the control group. To allow for this, propensity matching is applied in combination with distance using nearest neighbour 1-to-1 matching. The matching comparators are the area size and deprivation markers taken from the UK publications on deprivation, health, education, crime and housing. These are chosen to match those areas in the treatment group with those in the control group having similar levels of deprivation thus avoiding comparators applied later in hedonic modelling. For the Robin Hood Line, only changes in distance to the nearest station up to 4 km are used as a base standard. Consideration of higher distances requires a larger base standard treatment group to be matched against a much larger selection of potential locations to create a matching control group, which would have involved collating a much more extensive dataset. As shown in Table 5-1, matching works very well based on a 2 km distance threshold because it allows selection from a larger cohort of non-treatment locations. However, it should be noted that it excludes consideration for treatment those LSOAs further than 2km away from rail line where they may have also experienced benefits. The "summary of balance for all data" section shows that before matching the mean percentage of health was 0.49 less, housing 2.27 less, education 14 less, and crime 0.65 less in the treatment group than in the control group[5].

Table 5-1 Summary of propensity matching output: 2 km threshold - Robin Hood Line

Summary of balance for all data							
	Means Treatment	Means Control	SD Control	Mean Diff	eQQ Median	eQQ Mean	eQQ Max
Distance	0.425	0.210	0.169	0.215	0.236	0.215	0.346
Health	0.871	0.382	0.554	0.489	0.500	0.491	0.850
Housing	18.599	16.320	9.129	2.28	2.930	2.828	10.340
Education	43.198	29.142	19.488	14.05	13.960	14.12	23.930
Crime	0.947	0.297	0.627	0.649	0.650	0.653	0.830

[5] **eQQ** The quantile-quantile determines if two data sets come from populations with a common distribution and compares the quantiles of the first data set against the quantiles of the second data set.

Summary of balance for matched data							
Distance	0.425	0.366	0.182	0.059	0.074	0.058	0.126
Health	0.871	0.749	0.509	0.121	0.130	0.126	0.470
Housing	18.599	16.919	9.476	1.679	1.760	1.953	5.490
Education	43.198	37.405	21.2861	5.792	5.150	5.831	13.010
Crime	0.947	0.825	0.537	0.121	0.140	0.137	0.580

Percent Balance Improvement				
	Mean difference	eQQ Median	eQQ Mean	eQQ Max
Distance	72.679	68.631	72.640	63.321
Health	75.190	74.000	74.195	44.705
Housing	26.300	39.937	30.921	46.905
Education	58.787	63.109	58.717	45.633
Crime	81.346	78.465	79.014	30.120

Sample sizes		
	Control	Treatment
All	347	127
Matched	127	127
Unmatched	220	0

After matching, the treatment and control groups are now much similar in terms of health, education and crime deprivation markers, and the rightmost columns in these summary data show the median, mean, and maximum quartile between the treated and control data; smaller QQ values indicate better matching. The mean differences in percentage between treated and control areas reduced to 0.12 for health, 1.68 for housing, 5.7 for education and 0.12 for crime. Jitter plots and histograms help visualize the quality of the matching. In the jitter plot Figure 5-6, the absence of cases in the uppermost stratification indicates that there are no unmatched treatment units. The middle stratifications show the close match between the treatment and the matched control units, and the final stratification shows the unmatched control units. Both the numerical and visual data indicate that the matching was successful.

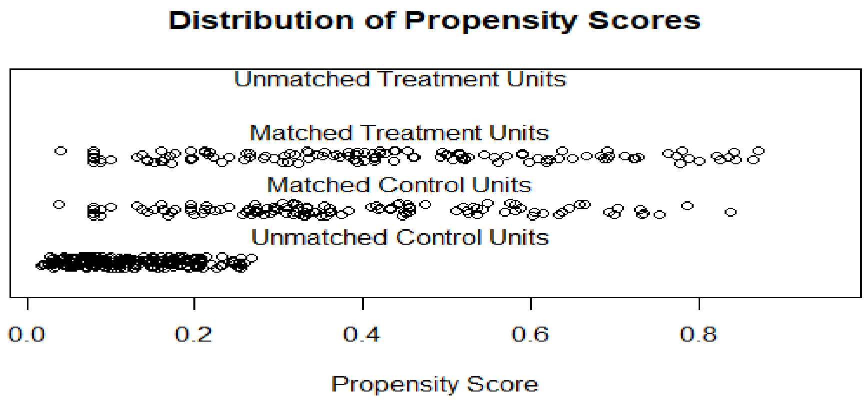


Figure 5-6 Distribution of propensity scores: 2 km threshold Robin Hood Line

Figure 5-7 shows the histograms before and after matching showing those before matching (on the left) differ to a great degree. However, the histograms after matching on the right are very similar. So both the numerical and visual data indicate that the matching was successful.

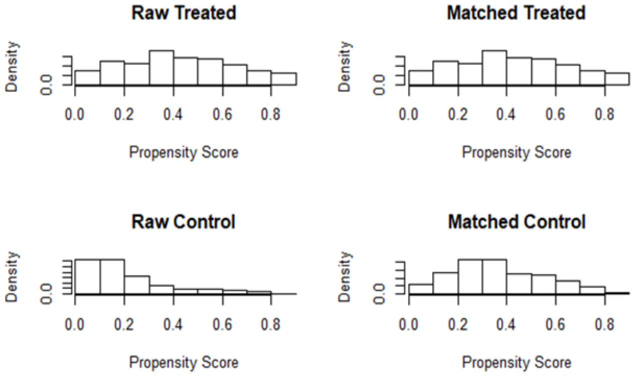


Figure 5-7 Histograms showing shapes before and after matching: 2 km threshold Robin Hood Line

Figure 5-8 illustrates the selection of treatment and control groups after the application of propensity testing based on a 2 km threshold. An example of matching treatment and control areas for a 4 km distance threshold are shown in Appendix 10.5.2, but this was much less effective and so only the 2 km threshold was used.

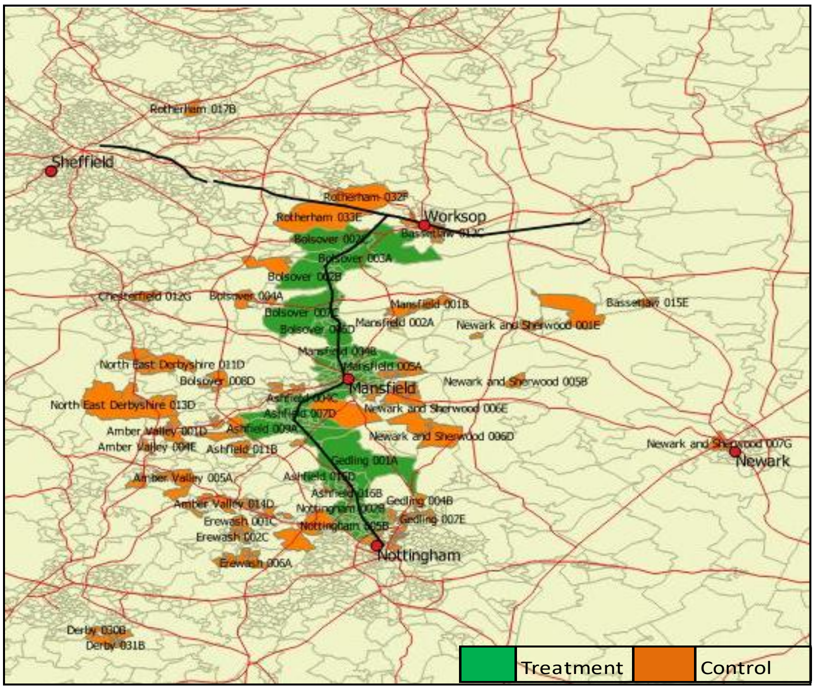


Figure 5-8 Treatment and control groups with propensity matching: 2 km

5.6.4 Application of clustering

There are limitations in using propensity matching because of the limited pool available for the control group. So as an alternative to propensity matching, clustering techniques produce a pairing of control and treatment groups which do not have to match numerically. This produces a further group pairing with again deprivation levels used as comparators using the dominant cluster and the 2 km "base" group, based on deprivation indicators only which allowed LSOAs within and outside the 2km threshold to be identified for each cluster. The membership of the cluster groups is as shown in Table 5-2 below:

Table 5-2 Cluster group memberships: Robin Hood Line

Cluster Group	Membership
1	176
2	58
3	131
4	48
5	61

A dendrogram provides visual evidence of the number of clusters (*Appendix 10.5.3*) indicating 5 distinctive clusters as shown again by the green ovals with an outlier as a red oval. The elbow chart in Appendix 10.5.3 also suggests the optimum number of clusters is 5. Examples of cluster analysis shown in Appendix 10.5.3 also suggest cluster groups to be incorporated into treatment and control groups. This confirms cluster group 1 to be the dominant cluster, and following initial division into the base grouping using 2 km distance as a basis for allocation, clustering is then applied by mapping matching clusters in the two groups.

5.6.5 Treatment and control group combinations

Subsequently, the different methods of allocation to treatment group outlined in sections 5.6.2 to 5.6.4 are taken forward, and out of eight possible combinations of treatment and control group pairings only five were used for further analysis (Table 5-3):

Table 5-3 Treatment/Control Group combinations: Robin Hood Line

Treatment Group ID	Comments	Distance contour (km)	Status
TG1	Base groups	2	
TG2		4	
TG3		6	
		8	Not used
		10	Not used
TG4	TG1 with added propensity	2	
	TG2 with added propensity	4	Not used
TG5	TG1 with added clustering	2	

Treatment/Control group combinations TG1 to TG3 represent selection to treatment based on increasing contour distance from the station where there has been a change in rail access. TG4 expands treatment group TG1 with the application of propensity matching, and TG5 is based on treatment group TG1 with the application of clustering using the dominant cluster for matching.

The other combinations were discarded as not suitable for sensitivity analysis as they provide an insufficient number of LSOA zones in the control group, without much greater expansion of the dataset. One specific treatment/control group combination (TG4) was adopted for the general analysis. This was chosen as it was based on a 2 km distance from a rail station, a threshold often taken as standard in the literature but with the added benefit of matching similarly deprived locations. Variations brought about by using the different treatment group selections are considered later in the sensitivity analysis.

5.7 Accessibility characteristics

The previous section covered the division of the region into treatment and control groups, defining five group treatment/group combinations to be used for investigations of causality in later sections of the chapter. In this section consideration is now given to four different measures for defining accessibility in the region.

These are firstly, distance from the nearest rail station, secondly the distance ratio (which expresses in percentage terms the change in distance to the rail network relative to the original distance), and thirdly a job accessibility index relating primarily to proximity to jobs. In addition, a fourth measure, an essential services index measuring proximity to hospitals, schools etc. provides a more complete picture of accessibility in the region. All measures are designed to be stand-alone and also incorporated into econometric models as accessibility characteristics.

5.7.1 Initial descriptive analysis

Consideration of each of these indicators requires prior knowledge of the background to jobs and services in the Robin Hood Line region, in particular:

- Car ownership
- Travel to work patterns by method and distance travelled
- Rail passenger usage to study uptake of rail travel in the region
- Accessibility to essential services

Using the standard treatment and control group (TG4) helps to contextualise the background and how it has affected "treated" areas. We can see this in the case of car ownership between 1991 and 2011 (Table 5-4) where proportionately more people now have access to a car in the treatment group which shows an overall drop of 9.93% in households having no access to a car against 8.90% for the control group.

At the other extreme, more households in the treatment group now have 2 cars or vans (7.04%), marginally higher than the treatment group (6.72%), which among other reasons may indicate either a greater affluence in the neighbourhood, or the necessity of the car for quick access to the station.

Table 5-4 % Changes in car ownership over the intervention period: Robin Hood Line

Group	No cars or vans	1 car or van	2 cars or vans	3 or more cars or vans
Treatment	-9.93%	0.33%	7.05%	2.54%
Control	-8.90%	-0.97%	6.72%	3.15%

This greater car ownership is evident in the means of travel to work figures (Table 5-5) comparing the immediate period of the intervention (1991-2001) and the remaining intervention period (2001-2011). Whereas between 1991 to 2001, there is an increase of 1.11% in rail share for the treatment group compared to 0.21% for the control group, between 2001 to 2011 there is a small decrease (-0.10%) for the treatment group but a similar increase (0.17%) for the control group. This may be explained by examining the tram figures (*Appendix 10.5.8*10.5.8) as the tram network was introduced in 2004 and takes a 2.32% share of transport mode.

Most of the impact of the tram would be expected in those treatment group locations closer to Nottingham as highlighted in *Appendix 10.5.1* (*Station Usage 1997-2015*). An increase in car driver travel resonates with the car ownership figures, but is less for the treatment group (10.61%) than the control group (11.38%) which may indicate greater use of rail and tram in the treatment group locations.

Table 5-5 % Changes in methods of travel to work over the intervention period: Robin Hood Line

Group	Year	Home	Tram	Train	Bus	Car driver	Car Passenger
Treatment	91-01	4.07%	0.01%	1.11%	-3.27%	4.03%	-1.40%
	01-11	-4.02%	2.32%	-0.10%	-2.01%	6.58%	-0.96%
Control	91-01	3.87%	0.02%	0.21%	-1.97%	4.60%	-1.14%
	01-11	-3.83%	0.41%	0.17%	-1.15%	6.78%	-1.24%

This is also accompanied by a decrease in bus mode over the whole period, which is more marked in the treatment group (-5.28%) than the control group (-3.12%), again suggesting a greater influence from rail and tram alternatives.

Table 5-6 % Changes in distance travelled to work 2001-2011: Robin Hood Line

Area	Year	Work mainly from home	Less than 10km	10km to less than 30km	30km and over	Other
Treatment	2001-2011	0.62%	-2.56%	3.09%	1.81%	-2.96%
Control	2001-2011	0.92%	-2.04%	2.74%	1.61%	-3.23%

However, the distance travelled to work over the period 2001 to 2011 indicates more commuting to distances of over 10 km, and is greater in the treatment group (4.90%) than the control group (4.35%) (Table 5-6). When compared to similar pre-existing stations nearby, rail station usage on the Robin Hood Line has grown more rapidly over the initial period which corroborates the census findings on rail mode usage. Appendix 10.5.1 shows station usage for all stations along the Robin Hood line which indicates a steady uptake, but decreasing use of 'competitor' stations near to Nottingham city since the introduction of the tram service there.

Finally, comparing distance to a range of essential services after the intervention by treatment and control group (Table 5-7), the average distance to nursery and primary schools is approximately the same for each group, but for secondary schools and hospitals is generally higher in the control group[6]. In addition, there is a much greater spread in distances within the control group when comparing the minimum and maximum values for individual LSOAs.

Table 5-7 Distance (km) to essential services over the intervention period (2001-2011): Robin Hood Line

	Nursery School			Primary School			Secondary School			Hospital		
Groups	Average	Min	Max	Average	Min	Max	Average	Min	Max	Average	Min	Max
Control	8.80	0.05	31.67	0.153	0.00	1.63	1.057	0.00	8.65	6.049	0.00	22.19
Treatment	8.10	0.78	22.53	0.144	0.00	0.52	0.751	0.00	4.50	1.967	0.00	5.54

This initial descriptive analysis of the accessibility background leads to consideration of three accessibility measures to be utilised in the econometric models.

5.7.2 Distance to nearest station

A basic measure used widely in the literature is the distance to the nearest available rail station on the network, and is calculated here using the Euclidean method, but factored up to allow for road network configurations not being in a straight line. For most locations in the Robin Hood Line region distance to the nearest station will decrease following the intervention as shown in Figure 5-9 which indicates a prevalent shift, so that post-intervention most LSOA areas in the case study region are now much closer to a rail station. Hence, as an indicator of accessibility to the rail network, it provides a basis for making comparisons of impacts for use in economic modelling, but has limitations in taking no account of cost or time.

[6] Distances were calculated using the straight line method based on OS coordinates for each type of service

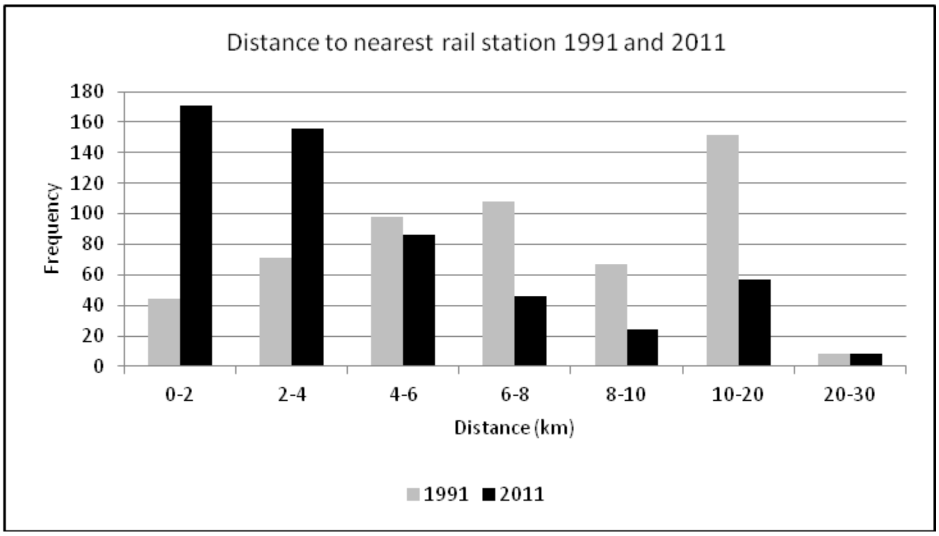


Figure 5-9 Distance to nearest station comparison (1991-2011): Robin Hood Line

5.7.3 Distance ratio

The second measure used here describes the percentage difference that the rail intervention has made in being comparatively closer to a rail station. It indicates the relative improvement in accessibility, expressed as a ratio, where there has been a change over time in distance to the nearest station due to new rail infrastructure. For example, a location previously 20 km from a rail station and now 5 km, has an improvement of 15 km and the distance ratio is thus 15/20 or 0.75. This ignores time or cost in reaching the station, but highlights a relative improvement in access to the rail network.

Here the ratio is $(d_{i2001} - d_{i2011})/d_{i2001}$ where

d_{i2001} -The initial distance to the nearest rail station in 2001

d_{i2011} - The distance to the nearest rail station in 2011

So for no change its value is zero but it will vary between 0 and 1, approaching 1 for large changes in distance.

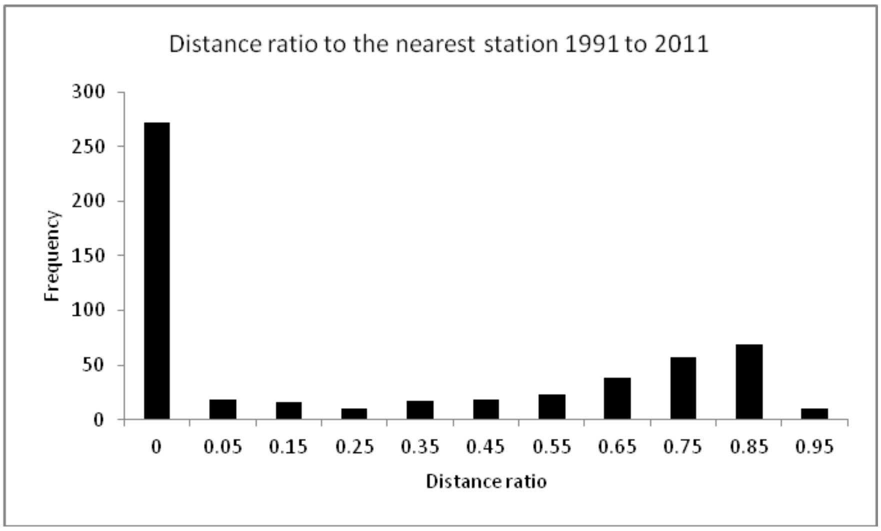


Figure 5-10 Distance ratio to nearest station (1991-2011): Robin Hood Line

Figure 5-10 shows a sizeable shift so that in 2011 most LSOA areas are now much nearer a rail station, and hence as an indicator of accessibility to rail mode it is a potential comparator for use in econometric modelling.

5.8 Job Accessibility index

The third measure is a more comprehensive indicator of accessibility than the previous two. A job accessibility index is calibrated to be applicable to the Robin Hood Line case study based on considerations in *4.10 Accessibility to jobs.* Alternatively applying generalised travel time (*4.10.6 Generalised travel time)* and generalised travel cost (*4.10.7 Generalised travel cost)* as proximity measures, a generic job accessibility index is derived.

The time and cost values and decay parameter for the region are based on transport data estimates using the methodology outlined in *4.10.12 Calibrating the index and costs for each case study.* The index also allows for commuting practicability (*4.10.8 Commuting practicability)* and those constraints which impose a boundary in access to jobs *(4.10.9 Thresholds),* and this is critical in the case of Robin Hood Line representing a more remote and disconnected region.

The index is alternatively measured based on consideration of all jobs available or only matching those that are suitable for the skills available in the region (*4.10.10 Job suitability - occupational matching)* and by comparison between the two methods, a comparator for the effect of job matching can be generated. Accessibility change is measured against original employment levels (1991) to match the models and approach undertaken here. Before accessibility can be calculated, consideration must be given to calibrating the parameters which define the measure for this particular region.

5.8.1 Setting the standard parameters

Allowing for differences between regions, a price comparison for individual modes of transport, requires a price per km, speed and headway for train, bus and car mode for the Robin Hood Line region, and these are derived using the methods prescribed in *4.10.12 Calibrating the index and costs for each case study*. From consideration of the above, standard values from Table 5-8 are applied in estimating the standardised time and cost of a journey for each transport mode.

Table 5-8 Cost and speed values used in index calculation: Robin Hood Line

Transport Mode	Year	Transport Speed	Service Frequency	Headway	Cost of Travel km
Bus	1991	42	1	30	£0.06
Bus	2001	42	1	30	£0.10
Bus	2011	42	1	30	£0.16
Car	1991	50	0	0	£0.15
Car	2001	50	0	0	£0.29
Car	2011	50	0	0	£0.48
Rail	1991	67	0	20000	£0.10
Rail	2001	67	1	30	£0.16
Rail	2011	67	1	30	£0.25
Walk	All	4.8	0	0	£0.00

As for the standard values, a decay parameter (β) is determined which produces the best fit for travel behaviour in the Robin Hood Line region (*4.10.12 Calibrating the index and costs for each case study)*. Through analysis of census distance to work statistics by LSOA for 2001 and 2011, using 2001 as the base year, Table 5-9 summarises cumulative distances travelled to work over the whole region, showing the variation between LSOAs.

The mean is the average over all LSOAs in the region for each distance, min is the lowest LSOA and max the highest LSOA value. This indicates that on average 39.85% of the working population travel at least 10 km to work, although there is a great variation across the region where this can be as high as 78.58% or as low as 20.71% with a median point of 7.5 km.

Table 5-9 Decay effect of travel distance: Robin Hood Line

Distance	Mean	Min	Max	SD
> 2 km	80.64%	60.43%	95.28%	8.01%
> 5 km	61.25%	34.17%	87.75%	12.54%
> 10 km	39.85%	20.71%	78.58%	11.01%
> 20 km	22.91%	14.38%	50.80%	5.82%
> 30 km	17.71%	12.80%	40.26%	3.88%
> 40 km	16.06%	11.86%	32.33%	3.34%
> 60 km	14.84%	11.06%	31.70%	3.02%
> 100 km	12.21%	8.43%	30.34%	2.70%

From these calculations a chart is constructed which maps the impedance value against distance travelled (Figure 5-11) and by fitting an exponential decay function graphically, an overall value of β can be estimated. By taking into account the mean rate of decay by LSOA a mean value is determined and this is used in future calculations.

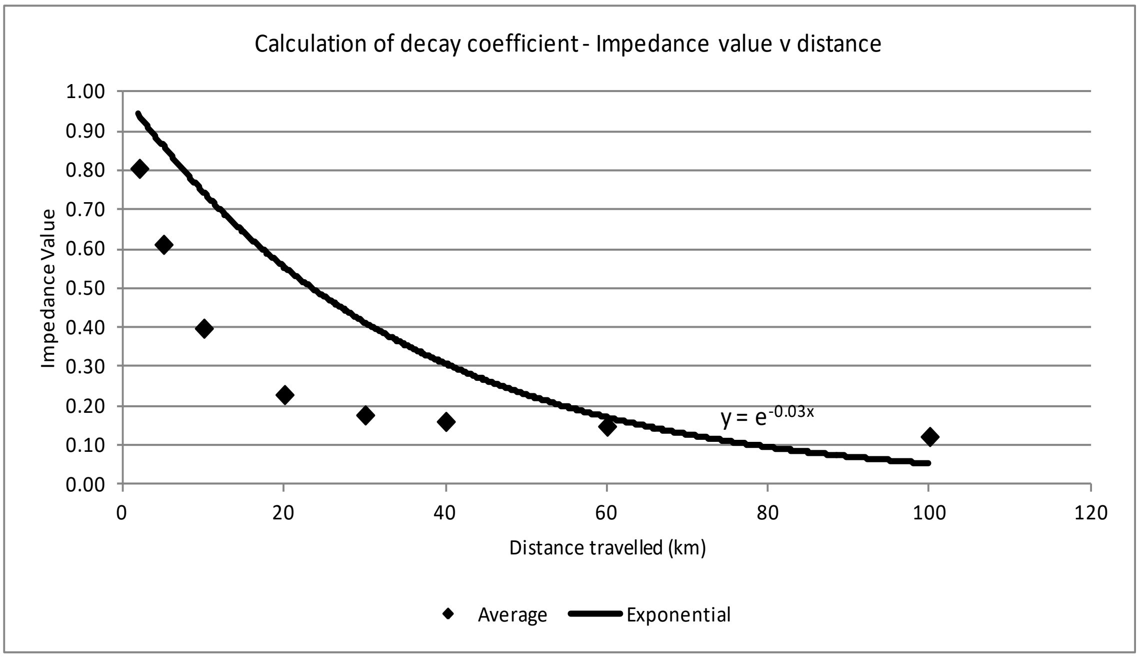


Figure 5-11 Decay coefficient calculation: Robin Hood Line

This initial value of β is based on distance only, but adjusted when calculating job accessibility where time or cost provides the basis for proximity and is dependent on the particular mode of transport as indicated in *4.10.12 Calibrating the index and costs for each case study*. This is reflected in the average speed and cost of that mode in the region.

A further consideration is the imposition of a threshold beyond which job opportunities are not included. As the negative exponential function is short tailed, long distances have limited effect on accessibility estimation, and thus function truncation does not result in any important loss of information in this case, and there are no long commutes to consider that are not within the specified case study region. Because no spatial interactions are observed beyond a certain distance threshold for a given purpose, and because of the asymptotic nature of the exponential function (i.e. every opportunity contributes to the potential value calculation), a maximum distance is defined beyond which opportunities are not included in the calculation. The threshold distance is defined using commuting data where the potential accessibility value reaches zero at the maximum travel distance observed in the region for all commuters (100 km).

5.8.2 Job accessibility comparison

Having determined the specific cost and decay parameters for the Robin Hood Line, the next stage was to compare the job accessibility index alternatively based on travel time and travel cost before and after the intervention. This was broken down further by applying either job skills matching or no matching after allowing for commuting thresholds. The index values are regionally standardised and the comparison measures the relative movement in values over the intervention period. The analysis breaks this down by treatment and control group using the specified treatment/control group combination (TG4).

There are two comparisons that job accessibility reflects:

1. infrastructure changes only by using the job situation prior to the intervention as the basis for job opportunity and skills matching
2. infrastructure changes and movement of jobs by factoring in the prevailing job situation in post-intervention years

In studying the values of the index to use in econometric modelling it was preferable to use original employment levels (Method 1) as an accessibility characteristic as this reflects the change in proximity to jobs rather than any movement in jobs over the period. It is also a measure of the immediate effect of the rail intervention experienced by the existing job market at the moment of intervention.

Infrastructure changes only

In this comparison, the index was calculated pre- and post-intervention by taking into account just rail infrastructure changes and using the job situation pre-intervention years. Using both a travel time and travel cost basis and without skills matching, accessibility to rail mode increased post-intervention for both the treatment group and control group (Table 5-11), but was considerably greater for the former, and the change on a cost basis was valued higher. This suggested that there were more benefits from improvements using cost as a basis.

Table 5-10 Method 1: Change in accessibility to rail treatment v control (1991-2011): Robin Hood Line

		with skills matching			without matching			Matching effect: Ratio matching to no matching	
Group	Method	pre	post	% change	pre	post	% change	pre	post
Treatment	Cost	0.026	0.029	12%	0.140	0.235	68%	5.385	8.103
	Time	0.097	0.102	5%	0.531	0.819	54%	5.474	8.029
Control	Cost	0.030	0.026	-13%	0.161	0.201	25%	5.367	7.731
	Time	0.108	0.098	-9%	0.586	0.774	32%	5.426	7.898

However, with the application of skills matching, accessibility to rail mode increased post-intervention for the treatment group but not the control group, and the change on a cost basis was again valued higher. For all other modes, there was negligible change on both a time and cost basis.

Pre- and post-intervention differences show an impact due partially to the change in proximity brought by rail. All results are aggregated job accessibility indexes for all locations within the Robin Hood Line region and suggest that job skills matching is required so that accessibility takes into account the skill sets in the region. There is evidence of this when dividing the unmatched jobs accessibility index from the 'skills match' index to provide a 'matching effect' which indicates the relevance of occupational matching in the calculation of job accessibility. This ratio of matching to non-matching values of the index shows a factor of approximately 5.5 pre-intervention rising to approximately 8 post-intervention. The wider difference when the index is based on travel cost rather than travel time, suggests that the index is more sensitive to cost, which may imply that it is more of an impediment.

Infrastructure changes and movement of jobs

In this comparison, the index is calculated pre- and post-intervention by taking into account not just the rail infrastructure changes but also job situation in pre- and post-intervention years. Using both a travel time and cost basis and without skills matching, accessibility to rail increased post-intervention for both the treatment group and control group (Table 5-11) but considerably more for the latter, and the change on a cost basis was valued much higher for the treatment group.

Table 5-11 Method 2: Change in accessibility to rail treatment v control (1991-2011): Robin Hood Line

		with skills matching			without matching			Ratio matching to no matching	
Group	Method	pre	post	% change	pre	post	% change	pre	post
Treatment	Cost	0.026	0.022	-15%	0.140	0.193	38%	5.385	8.773
	Time	0.097	0.088	-9%	0.531	0.784	48%	5.474	8.909
Control	Cost	0.030	0.019	-37%	0.161	0.162	1%	5.367	8.526
	Time	0.108	0.084	-22%	0.586	0.742	27%	5.426	8.833

With the application of skills matching, accessibility to rail mode decreased post-intervention for both groups, but this was substantially more for the control group, and the change on a cost basis was again valued higher. For all other modes, there was negligible change on both a time and cost basis.

Appendix 10.5.6 maps job accessibility by rail mode based on travel cost and compares predictions for pre- and post intervention values using alternatively no matching i.e. all job opportunities and skills matching i.e. only those opportunities matching. Pre- and post-intervention differences show an impact due partially to the change in proximity brought by rail but with interference from the slump in the job market up to 2011. All results are aggregated job accessibility indexes for all locations within the Robin Hood Line region, and this again suggests that without job skills matching there may be again an overestimation of accessibility due to the seemingly high attraction of job opportunities. The application of job skills matching highlights that the reduction in levels nearer to the line because of fewer jobs available which did not match the skills profile of residents in those locations.

Analysis of the reduction in the range of accessibility index values across the intervention period over all data zones is summarised in Table 5-12 which compares the variation in index value for all individual locations pre- and post-intervention. The reduction in this variation indicates that accessibility values are less variable spatially following the intervention. This decreased variability is more apparent when skills matching is taken into account where there is greater narrowing of the index range. This is less noticeable on a cost basis which may imply that there is still spatial diversity within the region due to cost constraints. Without matching the difference in accessibility across the region is predicted to be only slightly less than it was before the intervention.

Table 5-12 Spatial analysis showing reduction in range of job accessibility post-intervention: Robin Hood Line

	Matching		Without matching	
Parameter	Cost	Time	Cost	Time
Infrastructure changes and movement of jobs				
Range	55.6%	62.2%	21.1%	33.6%
Infrastructure changes only				
Range	39.5%	52.0%	9.8%	30.5%

5.9 Essential services accessibility index

Although the job accessibility index represents one of three main accessibility characteristics, to present a complete picture, an appropriate accessibility index for essential services is also derived based on *4.11 Accessibility to essential services.* This index measures access to five essential services and reflects the particular characteristics of the region, estimating how rail developments have impacted on levels of accessibility through analysis by treatment and control group.

An analysis of changes in the rail station accessibility index is summarised in Appendix 10.5.7, based on both a time-based and cost-based method. This indicates that regardless of a cost or time basis, across all modes, there is an increase in accessibility in the treatment group compared to the control group. On a time basis, average accessibility pre-intervention is similar for both treatment and control group for all services apart from nursery schools. Following the intervention, this increases more across all services for the treatment group (2.6%) compared to the control group (1.6%). This also holds on a cost basis with a greater increase for the treatment group (31.6%) compared to the control group (15.7%).

Spatial analysis reveals variations in the accessibility index across the region for individual LSOA areas post intervention. The spread of accessibility as shown by the minimum, maximum and standard deviation is greater in the control group for nursery schools, secondary schools and hospitals. Comparing rail accessibility to the hospital from treatment group zones (those close to the line) and control group zones shows that those nearer to the line have greater accessibility to the hospital using the rail mode.

5.10 Impact on residential property values

Having dealt with the provision of treatment and control groups and selection of accessibility measures, it is now important that these are carried forward to the application of property econometric models. This section covers assessment of the impact of the rail intervention on property prices, by incorporating the three defined accessibility measures as accessibility characteristics in econometric property price models using three modelling approaches: difference-in-difference, fixed effects and GWR (Geographically Weighted Regression).

5.10.1 Initial descriptive analysis

Before consideration of each of these model approaches, some background relevant to property price movements in the Robin Hood Line region will validate the variables applicable to the models. Following the methodology outlined in *4.12 Impact on residential property values* the situation before and after rail intervention is compared using the standard treatment/control group configuration (TG4) to compare changes in property type and house prices.

The type of property e.g. terraced, detached will influence property prices as areas with a higher number of detached properties will generally reflect a higher average house price. Table 5-13 illustrates how this impacts for both treatment and control groups across the intervention period showing a decrease in the percentage of terraced property from 31.1% to 25.9% for the treatment group and 28.5% to 24.2% for the control group between 1991 and 2011.

This represents a 5.2% drop for the treatment group against a 4.3% drop for the control group. Using the same measure, detached properties show a 2.97% increase for treatment group and 2.4% increase for the control group, and semi-detached properties a 1.75% increase for treatment, and only 0.59% increase for control. The treatment group now has a much higher percentage of flats (9.32%).

Table 5-13 Accommodation profile - treatment v control (1991-2011): Robin Hood Line (*Source: UK Census, 2011)*

Group	Year	Detached	Semi-detached	Terraced	Flat etc.
Treatment	1991	18.06%	39.96%	31.11%	8.71%
	2001	21.15%	42.30%	26.63%	8.78%
	2011	21.03%	41.71%	25.89%	9.32%
Control	1991	17.87%	43.20%	28.50%	7.19%
	2001	19.82%	44.50%	24.96%	7.64%
	2011	20.27%	43.79%	24.23%	8.71%

In terms of average house price, the increase between 1995 and 2016 is similar for both treatment and control groups, especially in the period 1995 to 2002 (the immediate period spanning the intervention) with the control group marginally higher (Figure 5-12). This is also echoed for minimum prices, but the maximum price is higher in the treatment group dropping below that of the control group after 2001.

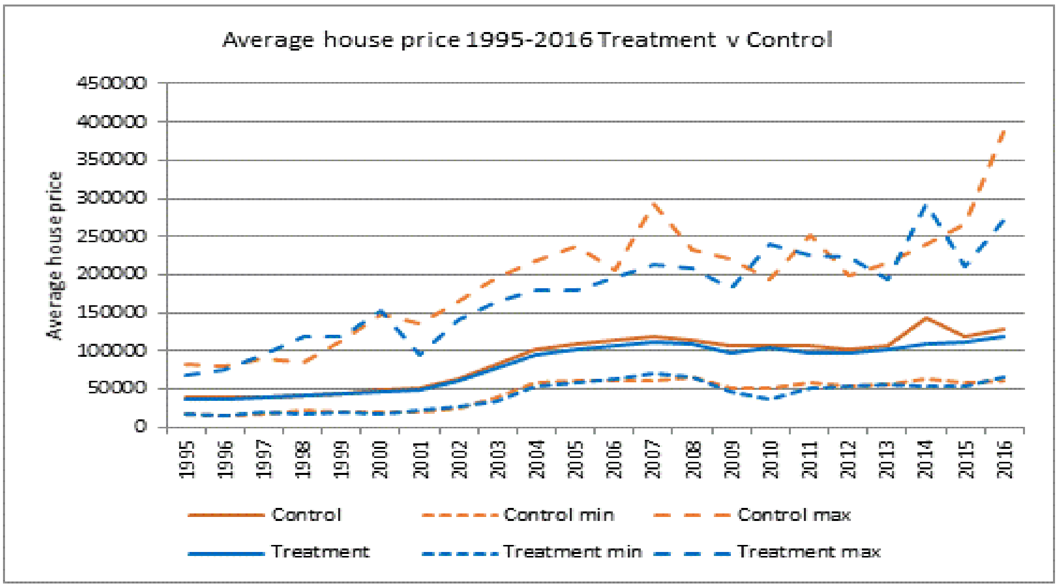


Figure 5-12 Average house prices 1995 to 2016 - treatment v control: Robin Hood Line (*Source: Land Registry, 2017)*

The regional respective on property prices (Figure 5-13) shows a clear distribution shape with a small number of very costly properties. The distribution is strongly positively skewed - suggesting a large upper tail - i.e. a central group of typically-priced houses together with a long tail of relatively expensive ones. The summary statistics are mean (£123324), median (£111616), standard deviation (£53392), and skewness (2.137).

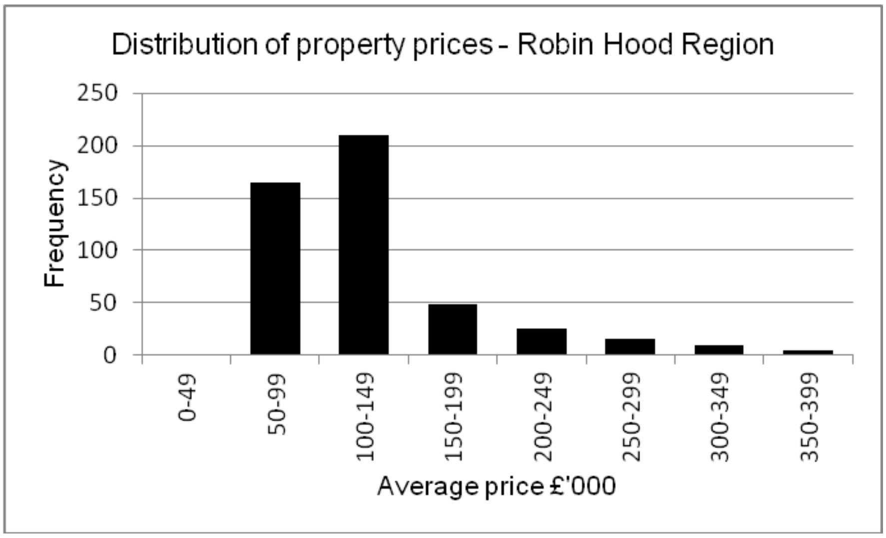


Figure 5-13 Distribution of property prices (1991-2011): Robin Hood Line
(*Source: Land Registry, 2017)*

5.10.2 Difference-in-difference model

The difference-in-difference model compares two separate years pre- and post-intervention to predict what would have happened had there been no intervention using the methodology in *1.1* but applying two different variables of accessibility which reflect distance from the nearest station and job accessibility. It should be noted that the DID models only takes property prices in the start and end year (1991 and 2011) and therefore have less statistical power than fixed effects models which consider every year.

Distance to station

Firstly, with distance to nearest rail station as the accessibility characteristic variable, Table 5-14 compares 1991 to 2011 showing a statistically significant negative coefficient when considering all groups together. This indicates that being closer to a rail station post intervention effects an increase in property prices, and shows an R^2 value of 0.1533. When a treatment variable replaces the distance to station, this suggests a statistically significant effect of treatment and an R^2 value of 0.1474.

Education Level 4 also shows as significant in reflecting the price of property because in this region the percentage of those educated at this level has risen dramatically over the twenty year period relative to other factors. However, care must be taken in interpreting these findings as causality due to changes in recording criteria over that time.

Table 5-14 Property DID results for distance to station: all areas and treatment variable (1991-2011): Robin Hood Line

	All areas				All areas with treatment variable			
	Coefficient	SE	t value		Coefficient	SE	t value	
(Intercept)	0.9495	0.0492	19.320	***	0.9524	0.0501	18.991	***
Distance to station	-0.0144	0.0049	-2.908	**				
No car households	-0.0005	0.0005	-0.961		-0.0006	0.0005	-1.081	
1 car household	0.0004	0.0006	0.682		0.0004	0.0006	0.657	
% employed	0.3964	0.2621	1.512		0.4605	0.2610	1.764	.
Education Level 1	0.0009	0.0006	1.527		0.0008	0.0006	1.472	
Education Level 4	-0.0019	0.0003	-6.287	***	-0.0019	0.0003	-6.255	***
No qualifications	0.0001	0.0002	0.386		0.0001	0.0002	0.458	
Population density	-0.0036	0.0035	-1.037		-0.0034	0.0035	-0.986	
Treated					0.1005	0.0436	2.305	*
R^2	0.1533				0.1474			
Number of observations: 948								

Distance Ratio

Using distance ratio as the accessibility characteristic, Table 5-15 shows a positive coefficient when considering all groups together, indicating that being proportionately closer to a rail station post intervention has a positive effect on prices.

Table 5-15 Property DID results for distance ratio: all areas (1991-2011): Robin Hood Line

	Coefficient	SE	t value	
(Intercept)	0.9357	0.0504	18.555	***
Distance Ratio)	0.1567	0.0516	3.036	**
No car households	-0.0005	0.0005	-1.071	
1 car household	0.0003	0.0006	0.499	
% employed	0.4373	0.2601	1.682	.
Education Level 1	0.0008	0.0006	1.415	
Education Level 4	-0.0019	0.0003	-6.202	***
No qualifications	0.0001	0.0002	0.506	
Population density	-0.0038	0.0035	-1.086	
R^2	0.1547	Number of observations: 948		

Again Education Level 4 appears significant due to a dramatic rise in the percentage of those educated at this level and, as before, care must be taken in interpreting these as causality due to changes in recording criteria.

Job Accessibility Index

Using the job accessibility index based on fixed labour and on a cost basis and skills matching as the accessibility characteristic, Table 5-16 shows a positive coefficient when considering all groups together. However, this does not imply a statistically significant impact on property prices as this is not reflected in the t-value of 1.045.

Table 5-16 Property DID results for job accessibility: all areas (1991-2011): Robin Hood Line

	All areas			
	Coefficient	Std. Error	t value	
(Intercept)	0.9891	0.0474	20.884	***
Job Accessibility Index	0.4862	0.4654	1.045	
No car households	-0.0005	0.0005	-0.970	
1 car household	0.0005	0.0006	0.844	
% employed	0.4845	0.2625	1.846	.
Education Level 1	0.0009	0.0006	1.559	
Education Level 4	-0.0020	0.0003	-6.582	***
No qualifications	0.0001	0.0002	0.350	
Population density	-0.0037	0.0035	-1.055	
R^2	0.1393			
Number of observations: 948				

As previously, Education Level 4 appears significant and again care must be taken in interpreting these as causality due to changes in recording criteria.

5.10.3 Fixed effects model

The standard fixed effects model was applied to the case study datasets for the Robin Hood Line. Aggregate data for the Robin Hood Line was collated from individual house transactions spanning the years 1991 to 2016 for each zone (LSOA) into an aggregated property price database. This acted as the base data for a standard model using the three different variables of accessibility. Eventually, only data from the years 1993 to 2006 was used for the following reasons:

- This data spanned the time period containing the intervention without being too heavily dependent on later years' data.

- Over that period property prices increased steadily up to the house market slump making it easier to distinguish impacts when time trend was taken into account.
- Prior to 1995 there were no details of individual house transactions.

Distance to the nearest station

The model run results with distance to the nearest station as the accessibility characteristic indicate that this is a significant factor in influencing the price of property. With log of property price as the dependent variable there is an R^2 value of 0.629 (Table 5-17). Against a background of fixed location and time effects, the distance to station has a coefficient of -0.027, indicating that distance to the nearest station has a negative effect on property price as distance increases. For a property in at £100,000 level this means a reduction of £2729 for each 1 km distance from the station.

Table 5-17 Model output with distance to station as accessibility characteristic (1991-2011): Robin Hood Line

	Coefficient	Std. Error	t-value	
Distance to nearest station	-0.027660	0.003409	-8.115	***
Terrace	-0.000107	0.000455	-0.236	
Detached	0.000481	0.000298	1.616	
Semi	0.002031	0.000435	4.664	***
No car households	-0.001958	0.000439	-4.454	***
1 car household	-0.001068	0.000497	-2.148	*
% employed	2.148806	0.286705	7.495	***
Level 1 education	-0.000163	0.000504	-0.323	
Level 4 education	0.002285	0.000204	11.212	***
No qualifications	-0.001688	0.000353	-4.782	***
Population Density	0.001915	0.003883	0.493	
R^2	0.629580			
Fixed effects groups: locations (474) time (16)				
Number of observations: 12324				

Other notable factors include the employment rate, where a higher percentage of employment is reflected in the price of property for that particular LSOA location; also where there is a greater amount of no car or one car ownership, the property price is relatively lower. This also holds for educational levels where there is a higher percentage of residents with no qualifications, property prices are lower. This is all against a background of fixed location and time effects.

Distance to station ratio

Running the model with distance to station ratio as the accessibility characteristic also corroborates the previous findings, and may be more relevant as they reflect a percentage improvement in accessibility (Table 5-18). This shows an R^2 value of 0.635, and the distance ratio has a coefficient of 0.327 which indicates that an increase in distance ratio i.e. being proportionally nearer the station, has a positive effect on property price as the relative change in distance increases; so being 10% nearer the station the property price on a £100,000 house will increase by £3326.This again shows that the employment rate has a positive effect on property price, whereas having no car or lack of qualifications have a negative effect on property price, with higher levels of education having a positive effect.

Table 5-18 Model output with distance ratio as accessibility characteristic (1991-2011): Robin Hood Line

	Coefficient	Std. Error	t-value	
Distance to station ratio	0.327300	0.035948	9.105	***
Terrace	-0.000002	0.000452	-0.005	
Detached	0.000516	0.000294	1.755	.
Semi	0.001931	0.000433	4.463	***
No car households	-0.001890	0.000436	-4.334	***
1 car household	-0.001288	0.000493	-2.611	**
% employed	2.183600	0.284360	7.679	***
Level 1 education	-0.000146	0.000500	-0.291	
Level 4 education	0.002275	0.000202	11.25	***
No qualifications	-0.001582	0.000350	-4.513	***
Population Density	0.001789	0.003854	0.464	
R^2	0.635130			
Fixed effects groups: locations (474) time (16)				
Number of observations: 12324				

Job accessibility index

Finally, running the model with cost-based job accessibility index (Table 5-19) as an alternative accessibility characteristic, taking into account distribution of jobs and the cost of reaching those jobs as a deterrent yields an R^2 value of 0.608. The job accessibility has a coefficient of 0.498, indicating an increase in this job accessibility index has a positive effect on property price, but of a lower order than predicted by the previous characteristics. For example, a 10% increase in job accessibility would generate a 5.12% increase in property price from £100,000 to £105,115. Again, where fewer households have no access to a car or a higher percentage of residents with no qualifications, the property price is relatively lower.

Table 5-19 Model output with job accessibility as accessibility characteristic (1991-2011): Robin Hood Line

	Coefficient	Std. Error	t-value	
Job accessibility	0.498846	0.286201	1.743	***
Terrace	-0.000130	0.000468	-0.278	
Detached	0.000786	0.000309	2.545	
Semi	0.002616	0.000449	5.826	***
No car households	-0.002483	0.000455	-5.462	***
1 car household	-0.001174	0.000511	-2.296	*
% employed	2.301550	0.295879	7.779	***
Level 1 education	-0.000278	0.000519	-0.536	
Level 4 education	0.002287	0.000218	10.492	***
No qualifications	-0.001710	0.000364	-4.703	***
Population Density	0.001389	0.003999	0.347	
R^2	0.607610			
Fixed effects groups: locations (474) time (16)				
Number of observations: 12324				

5.10.4 GWR Model

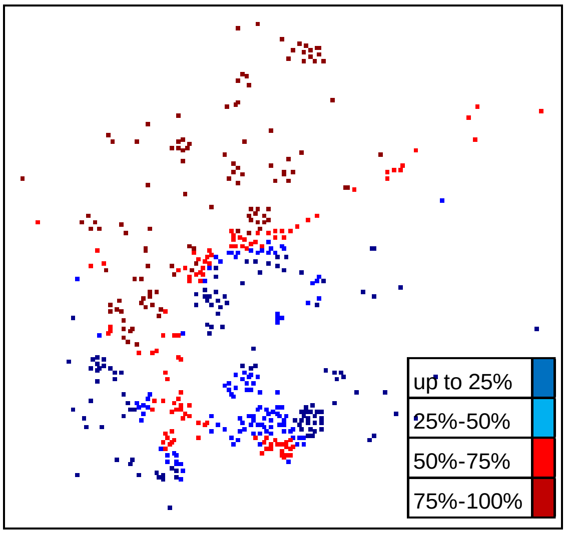


Figure 5-14 Spatial distribution of coefficients

Although the previous models suggest a relationship between accessibility characteristics and property price, they do not fully reflect any spatial variability which may be apparent when dealing with regions such as the Robin Hood Line region. As a third approach, Geographically Weighted Regression (GWR) models changes in spatial diversity over a period of time spanning the rail intervention. Using the methodology from *4.14.1 Property prices*, for the GWR analysis, data is collated for years 1991, 2001 and 2011 spanning the intervention. The data frame contains census information collected for 545 LSOA zones in the Robin Hood Line area. The analysis uses house price data representing a sample of houses transacted in the region between 1995 and 2016.

As it is a cross-sectional model, for comparison purposes it is run for each of the three years to compare coefficients in changes at a local level both spatially (across the LSOA zones) and temporally with focus on 2011 difference-in-difference employment results. GWR produces a set of parameter estimates and model statistics for each sample. As for previous models, distance to nearest station and the job accessibility index provide an alternative accessibility characteristic for both the GWR and global (OLS) models. In some locations the relationship between the log of property price and distance to station or job accessibility is negative, and in others positive. Coefficient ranges for the other variables suggest some interesting spatial patterning (Figure 5-14).

Table 5-20 summarises the parameter estimates fitted by GWR for 1991, 2001 and 2011, including the median, upper quartiles, lower quartiles, range, minimum and maximum comparing this with the OLS global model. Using distance to nearest station, for 2001 the coefficients ranged from a minimum value of -0.077 (1 km change in distance from the station resulting in a drop in average house price of £7710) to +0.014 (1 km change in distance resulting in an increase in average house price of £1337).

Table 5-20 Spatial variation of coefficients for different accessibility characteristics (1991-2011): Robin Hood Line

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	SE	t value	R^2
Distance nearest station										
1991	-0.008	0.0003	0.003	0.011	0.022	0.010	0.003	0.003	1.101	0.715
2001	-0.077	-0.061	-0.022	-0.002	0.014	0.059	-0.014	0.009	-1.522	0.738
2011	-0.007	0.0038	0.010	0.018	0.060	0.014	0.005	0.004	1.363	0.681
Job Accessibility										
1991	-7.717	-6.097	-2.150	-0.173	1.377	5.924	-1.416	0.931	-1.522	0.683
2001	-11.530	-6.542	-2.617	3.536	6.574	10.078	-0.479	1.767	-0.271	0.733
2011	-17.830	-8.701	-5.827	0.652	2.204	9.353	-1.848	2.121	-0.871	0.680

The t-value applies to the global model only, and shows that the coefficients for the global model do not appear to be significant at 95% level, as evidenced by the spread of values in the GWR analysis. Hence, in comparing GWR model results with a basic cross-sectional OLS model (represented by the global model), the analysis detects a significant advantage in using GWR with an improvement in fit of a GWR model over a global model using an F test to check significance. These GWR results do not affect the validity of those for the fixed effects model as the global model is based on one year only e.g. 2001 rather than a span of years e.g. 1995-2016 as in the fixed effects model.

Also, as all the interquartile values were greater than their corresponding 2 standard errors of the global estimates for distance from the nearest station it is clearly evident that there is non-stationarity of relationships over the case study region. However, it is important to test whether it is significant or not (Shi et al., 2006). If the relationships do not vary significantly, there is no advantage to using the GWR model over the simplified OLS model (Shi et al., 2006). A Monte Carlo test on the parameter estimates suggested that all the parameter estimates were highly significant at 0.1% of significance level which confirms the existence of significant local non-stationary relationships. Therefore, applying a global model which assumes stationarity under such conditions, leads to under-prediction of the model.

5.11 Impact on jobs and employment

In this section, a similar methodology is used to assess the impact of the rail intervention on jobs and employment. This section covers assessment of the impact of the rail intervention on employment density, by incorporating the defined accessibility measures as accessibility characteristics in econometric property price models using three modelling approaches: difference-in-difference and GWR (Geographically Weighted Regression).

5.11.1 Initial descriptive analysis

Consideration of each of these model approaches requires some prior knowledge of the background relevant to jobs and employment in the Robin Hood Line region. This follows the methodology outlined in (*4.13 Impact on jobs and employment)* - an analysis evaluates the situation before and after rail intervention with reference to the standard treatment/control group (TG4) comparing changes in:

- Education levels and qualifications
- Economic activity
- Job profile by industry and occupation
- Population Levels

In terms of employment levels there was an increase in unemployment in the period 2001 to 2011, but it is noticeable that full time employment decreased more sharply with unemployment showing a higher increase in the control group (Table 5-21). In addition, the percentage of part-time workers has doubled in the control group area.

Table 5-21 Changes in economic activity by treatment and control group (1991-*2011): Robin Hood Line*

Group	Year	Full-time	Part-time	Self-employed	Unemployed	Full-time student
Treatment	91-01	0.18%	1.09%	0.81%	-4.36%	2.28%
	01-11	-3.66%	1.49%	0.71%	0.68%	0.78%
Control	91-01	-0.44%	2.15%	0.13%	-4.09%	2.26%
	01-11	-5.19%	2.27%	1.24%	0.97%	0.71%

Educational levels and qualifications represent potential for employment and Table 5-22 highlights changes in educational standards by treatment and control group over the initial rail intervention period (1991-2001) and later (2001-2011).There is a similarity between the two groups showing much fewer having no qualifications over the initial intervention period combined with a sizeable increase in level 4 and qualifications over the entire period.

Table 5-22 Changes in educational standards by treatment and control group (1991-2011): Robin Hood Line

Group	Year	No qualifications	Level 1	Level 2	Level 3	Other	Level 4 and above
Treatment	1991-2001	-53.54%	15.45%	15.82%	5.67%	7.28%	9.31%
	2001-2011	-7.52%	-2.61%	-0.81%	6.29%	-1.92%	6.57%
Control	1991-2001	-53.20%	14.61%	14.53%	5.56%	7.36%	11.14%
	2001-2011	-6.83%	-2.40%	-0.54%	6.24%	-2.66%	6.19%

The job accessibility index incorporates occupation share as an important element when comparing matching of job opportunities between different locations. From Table 5-23, both treatment and control group show an overall increase in residents with professional or associate professional and technical occupations, which is higher in the treatment group (10.82%) than the control group (9.89%).

Managers directors	1	Professional	2	Assoc. professional	3
Administrative / secretarial	4	Skilled trades	5	Caring leisure	6
Sales and customer service	7	Machine operatives	8	Elementary	9

Table 5-23 Changes in jobs by occupation by treatment and control group (1991-2011): Robin Hood Line

Group	Years	1	2	3	4	5	6	7	8	9
Treatment	91-01	0.12%	1.71%	4.45%	-2.02%	-10.01%	-1.77%	0.85%	0.91%	5.76%
	01-11	-2.69%	5.14%	-0.48%	-0.86%	-1.23%	3.52%	0.96%	-3.44%	-0.90%
Control	91-01	-0.02%	1.73%	4.22%	-1.65%	-9.33%	-1.54%	1.15%	-0.09%	5.52%
	01-11	-2.88%	4.69%	-0.75%	-0.30%	-0.58%	3.40%	1.27%	-2.98%	-1.88%

In addition, Table 5-24 indicates a sizeable move away from the traditional industries e.g. manufacturing and mining together with a move towards hotels and catering and the construction industry which is similar both groups.

Table 5-24 % changes in jobs by industry by treatment and control group (1991-2011): Robin Hood Line

	Treatment		Control	
	1991-2001	2001-2011	1991-2001	2001-2011
Agriculture, hunting and forestry	0.20%	-0.50%	0.11%	-0.50%
Mining and quarrying	-1.36%	-0.41%	-2.92%	-0.40%
Manufacturing	-4.63%	-8.38%	-4.87%	-8.36%
Electricity, gas and water supply	-7.03%	0.91%	-5.72%	0.86%
Construction	0.98%	0.49%	0.50%	1.00%
Wholesale and retail trade, repair of motor vehicles	-1.65%	0.03%	-1.13%	0.29%
Hotels and catering	4.50%	0.67%	4.84%	0.45%
Transport, storage and communication	1.59%	-0.25%	1.30%	-0.71%
Financial intermediation	-4.05%	-0.58%	-3.96%	-0.38%
Real estate, renting and business activities	8.26%	-7.54%	8.66%	-7.63%
Public administration and defence	4.38%	0.74%	4.44%	0.54%
Education	5.62%	1.89%	6.22%	1.99%
Health and social work	11.39%	2.92%	10.83%	2.64%
Other	-17.70%	5.71%	-17.81%	5.77%

To place this in context, analysis of change in population level by control and treatment group between 1991 and 2011 (Figure 5-15) reveals an increase in both groups over the period, which is more noticeable for the treatment group between 2001 and 2011.

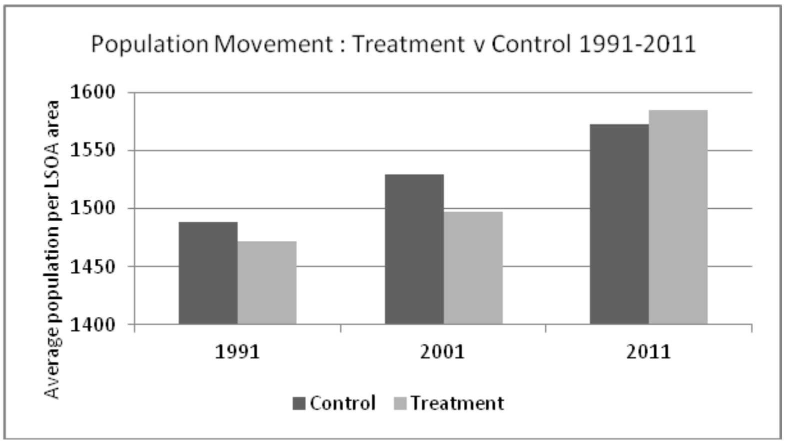


Figure 5-15 Population levels 2002-2014 by treatment and control group (1991-2011): Robin Hood Line

Figure 5-16 is a snapshot showing the distribution of employment density for the entire region in 2001, which has a median of 6 jobs per square km but a high proportion of areas with density less than 2. At the other extreme, there are very few locations where the density exceeds 30 per square km. This typifies the whole region with its mixture of locations with little or no employment or potential for jobs (a thin market), and an urban cohort with a limited level of employment.

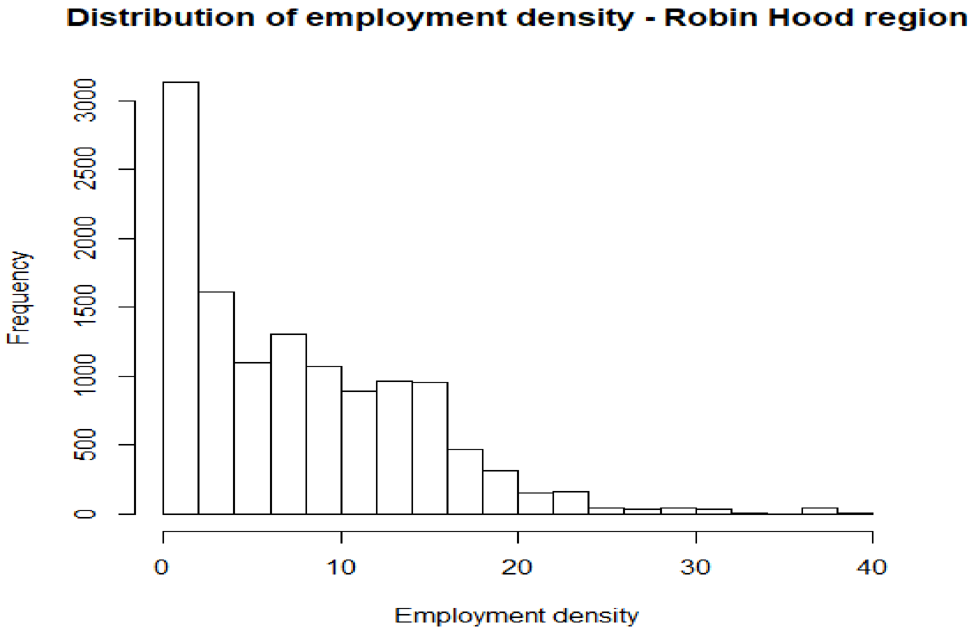


Figure 5-16 Distribution of employment density 2001: Robin Hood Line

5.12 Difference-in-Difference model

As for the property model earlier, the difference-in-difference model compares two separate years pre- and post-intervention using the methodology in *4.13.2 Employment modelling – Difference-in-Difference* but applying different variables of accessibility.

Distance to nearest rail station

Firstly, with distance to nearest rail station as the accessibility characteristic, Table 5-25 shows a significant negative coefficient when considering all groups together. This indicates that if difference in distance to the rail station post intervention is negative i.e. the worker is nearer a station, this will have a significant positive effect on employment density and vice versa (R^2 value of 0.6122).

When a treatment variable replaced the distance variable, the effect on employment density of treatment i.e. being nearer a station, is positive and statistically significant and has an R^2 value of 0.6093.

Table 5-25 Employment DID results for distance to station: all areas and with treatment variable (1991-2011): Robin Hood Line

	All areas				All areas with treatment variable			
	Coefficient	Std. Error	t value		Coefficient	Std. Error	t value	
(Intercept)	-0.0883	0.0362	-2.438	*	-0.0833	0.0369	-2.258	*
Distance to nearest station	-0.0101	0.0036	-2.790	**				
Terraced	0.0015	0.0004	3.972	***	0.0015	0.0004	3.905	***
Detached	0.0020	0.0003	7.648	***	0.0020	0.0003	7.642	***
Semi	0.0013	0.0004	3.274	**	0.0013	0.0004	3.228	**
No car households	0.0011	0.0004	2.952	**	0.0011	0.0004	2.796	**
1 car household	0.0017	0.0004	3.890	***	0.0017	0.0004	3.892	***
Education Level 1	0.0021	0.0004	5.537	***	0.0022	0.0004	5.589	***
Education Level 4	0.0004	0.0002	1.980	*	0.0004	0.0002	1.934	.
No qualifications	-0.0003	0.0002	-1.923	.	-0.0003	0.0002	-1.854	.
Population density	0.0044	0.0026	1.739	.	0.0045	0.0026	1.734	.
Treated					0.0677	0.0324	2.094	*
R^2	0.6122				0.6093			

Distance ratio

Table 5-26 Employment DID results for distance ratio: all areas (1991-2011): Robin Hood Line

	Coefficient	Std. Error	t value	
(Intercept)	-0.0854	0.0373	-2.291	*
Distance Ratio (DSR)	0.0802	0.0384	2.086	*
Terraced	0.0015	0.0004	3.938	***
Detached	0.0020	0.0003	7.750	***
Semi	0.0013	0.0004	3.304	**
No car households	0.0011	0.0004	2.840	**
1 car household	0.0017	0.0004	3.817	***
Education Level 1	0.0022	0.0004	5.579	***
Education Level 4	0.0004	0.0002	1.901	.
No qualifications	-0.0003	0.0002	-1.872	.
Population density	0.0042	0.0026	1.642	
R^2	0.6093			

Using distance ratio as the accessibility characteristic, Table 5-26 shows a significant positive coefficient considering all groups together, which indicates that being proportionately closer to a rail station post intervention has a significant positive effect on employment density.

Job accessibility

Using job accessibility index based on cost as the accessibility characteristic with labour constant, Table 5-27 shows a significant positive coefficient considering all groups, indicating that improvement in job accessibility post intervention has a positive effect on employment density.

Table 5-27 Employment DID results for job accessibility: (1991-2011): Robin Hood Line

	All areas			
	Coefficient	SE	t value	
(Intercept)	-0.0618	0.0344	-1.795	.
Job Accessibility Index	0.7625	0.3431	2.223	*
Terraced	0.0015	0.0004	3.904	***
Detached	0.0020	0.0003	7.592	***
Semi	0.0013	0.0004	3.211	**
No car households	0.0011	0.0004	2.875	**
1 car household	0.0018	0.0004	4.023	***
Education Level 1	0.0021	0.0004	5.500	***
Education Level 4	0.0004	0.0002	1.886	.
No qualifications	-0.0003	0.0002	-1.809	.
Population density	0.0045	0.0026	1.768	.
R^2	0.6098			

5.12.1 GWR Model

In contrast to a single set of constant values over the study region generated by an OLS model, GWR produces a set of parameter estimates and model statistics for each sample. Using the methodology outlined in *4.14.2 Jobs and employment* for the GWR analysis, data was collated for the years 1991, 2001 and 2011 spanning the intervention, and the GWR model is run for those three years (Table 5-28). The data frame contains census information collected from 545 LSOA zones in the Robin Hood Line area. Using a global OLS model for employment density, the accessibility characteristics used previously were substituted in turn to generate two separate global models as the basis for the GWR model.

There was an indication that the relationship between log of employment density and distance to the nearest station was in most cases negative - reflecting the global model. It was also skewed in that the minimum coefficient value was -0.878 with a median of -0.057 and a maximum of 0.1394. There was also a narrow spread in terms of population density (0.269 to 0.380) with a median of 0.322.

For distance to nearest station, the coefficients ranged from a minimum value of -0.878 (1 km increase in distance from the station resulting in a drop in average employment density of 87.8%) to +0.1394 (1 km increase in distance from the station resulting in an increase in average employment density of 13.94%). For half of the LSOA zones in the dataset, as distance to station rose by 1 km, average employment density decreased between 9.6% and 1.9% (the inter-quartile range between the 1st quarter and the 3rd quarter).

As with the property model, the range of GWR local parameter estimates is examined with confidence intervals around the OLS global estimate of the equivalent parameter. The analysis compares GWR model results with the cross-sectional OLS model (which represents the global model) to detect any significant advantage in using GWR by reporting an improvement in fit of a GWR model over a global model using an F test to check significance.

Table 5-28 Comparison of GWR coefficients for different accessibility characteristics (1991-2011): Robin Hood Line

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	SE	t value	R^2	
Distance nearest station											
1991	-0.4455	-0.0947	-0.023	0.0159	0.328	0.1106	-0.054	0.0117	-4.582	0.77	***
2001	-0.3308	-0.2136	-0.146	-0.0973	-0.009	0.1163	-0.123	0.0128	-9.643	0.82	***
2011	-0.3636	-0.2186	-0.163	-0.1122	-0.001	0.1064	-0.13	0.0122	-10.69	0.83	***
Job Accessibility											
1991	-10.96	2.357	7.572	12.57	30.03	10.213	18.37	2.5548	7.19	0.78	***
2001	-9.604	1.421	7.777	18.89	45.07	17.469	32.23	4.3489	7.411	0.80	***
2011	-10.02	0.5572	10.74	21.85	39.6	21.293	28.963	6.0228	4.809	0.79	***

Table 5-28 compares the GWR model coefficients over three different years when applying different accessibility characteristics to the model. It is clearly evident that all the interquartile values were greater than the corresponding 2 standard errors of the global estimates for distance from the nearest station indicating the presence of non-stationarity of relationships over the study region. A Monte Carlo test on the parameter estimates suggested that all the parameter estimates were highly significant at 0.1% of significance level, confirming the existence of significant local non-stationary relationships over study area. If we apply the global model which assumes stationarity under such conditions, there is certainly misspecification of the true situation, and under-prediction of the model.

5.13 Sensitivity analysis

Analysing the impact of the rail intervention through application of a job accessibility index and property and employment models using the standard combination (TG4) illustrated that some causality could be attributed, but it was neither practical nor feasible at that stage to investigate every possible group combination. This section revisits that analysis by examining the effect of applying alternative selections into treatment. The accessibility indices and property and job difference-in difference models are re-run using each treatment group/control group combination to investigate the sensitivity of results.

Job Accessibility

The job accessibility index for rail is calculated to see how group selection impacts on the value of the accessibility measure. The variations using different group combinations for accessibility, calculated on a cost and time basis and with or without skills matching are shown in Table and represent the estimated increase in accessibility between 1991 and 2001 expressed as a percentage. Here, Table looks at how the predicted changes in job accessibility will vary if alternative treatment groups are chosen.

The predicted increase in job accessibility is fairly consistent, regardless of the selection of treatment group -2.63% variation with skills matching and 0.31% variation with no matching on a cost basis and similar figures on a time basis. However, there is a much wider variation for the respective control groups -7.63% variation with skills matching and 1.02% with no matching on a cost basis and similar figures on a time basis. This reflects the reducing size of the control group with expansion of the treatment group to cover greater distance.

Table 5-29 Job Accessibility variations for different treatment groups (1991-2011): Robin Hood Line

Group Combination ID	Group	Time basis		Cost basis	
		Matching	No Matching	Matching	No Matching
TG1	Control	-1.14%	17.09%	-0.49%	3.13%
	Treatment	0.45%	28.72%	0.36%	9.50%
TG2	Control	-1.26%	15.13%	-0.60%	2.00%
	Treatment	-0.07%	26.53%	0.14%	8.37%
TG3	Control	-1.51%	13.31%	-0.68%	1.37%
	Treatment	-0.01%	26.47%	0.11%	7.94%
TG4	Control	-1.06%	18.84%	-0.46%	3.95%
	Treatment	0.45%	28.72%	0.36%	9.50%
TG5	Control	-1.66%	16.25%	-0.61%	3.27%
	Treatment	-0.03%	26.08%	0.27%	8.99%

Property value

All treatment and control group combinations are run against the difference-in-difference model to compare the effect of alternative "selection into treatment" processes. For the difference-in-difference model, as the distance from the nearest station there is a consistently positive effect of treatment, which is noticeably less for groups TG4 and TG5 where matching is applied. For job accessibility the effect of treatment is positive where matching is applied.

Employment

All treatment and control group combinations are run against the difference-in-difference model to compare the effect of alternative "selection into treatment" processes. For the difference-in-difference model, with distance from the nearest station, there is a consistently negative effect of increased distance and positive effect of increased job accessibility.

5.14 Conclusion and Summary

The findings indicate that rail accessibility has had some impact on property prices and employment density, and that for the Robin Hood Line, providing new rail infrastructure has provided extra benefit in increases in jobs and house prices. However, because of the complexity of the region with its many confounding factors of the length of time that has elapsed since the intervention, this exact impact is difficult to evaluate. Analysis of the models shows that in more remote parts of the region, there are alternatives to distance to rail station as the sole accessibility measure, namely the distance to station ratio and job accessibility which give consistent and significant results in the model runs. The inclusion of the job accessibility index to represent the opportunities reachable in each location is perhaps more relevant to this case study where there are pockets of higher employment in addition to the job market in Nottingham, and works better in the employment model.

The region was divided into treatment and control groups to monitor effects of infrastructure changes using distance thresholds from the intervention to create basic groups. Propensity matching of those treatment and control groups with similar levels of deprivation worked well here for the 2 km and 4 km distance contours, as did clustering techniques using the dominant cluster groups for matching. Going beyond the 4 km contour required matching a larger base treatment group against a much larger selection of potential locations in the control group.

A comparative analysis of treatment and control groups shows that in the treatment group, proportionately more households have access to a car, which may indicate a greater affluence in the neighbourhood, or the necessity of quick access to the station for commuting. An increase in car travel is less for the treatment than the control group which may reflect greater use of rail and tram in the former. Despite the rapid growth of rail travel over the early intervention period, there is evidence of the impact of the tram in treatment groups nearer to Nottingham with decreased use of "competitor" stations nearer Nottingham centre.

Job accessibility

There are two comparisons that job accessibility reflects:

1. Infrastructure changes only - using the job situation prior to the intervention as the basis for job opportunity and skills matching for use in modelling.
2. Infrastructure changes and movement of jobs - factoring in the prevailing job situation in post-intervention years.

Rail mode shows a marked increase in accessibility post-intervention for the treatment group, compared to a reduction when taking into account all modes. All results suggest that without job skills matching there may be an overestimation of accessibility due to the seemingly high attraction factor of job opportunities. The wider difference when the index is based on travel cost rather than travel time, suggests that cost is more of an impediment. Hence a job accessibility measure based on travel cost and allowing for skills matching was taken forward as the best option for application in the econometric models. In retrospect, it would have been useful to estimate the models with both accessibility measures to compare the effects of sorting of labour and also a comparison of the impact of the skills matching accessibility measure versus no matching.

The job accessibility measure is fairly consistent for different treatment groups, but with a much wider spread of values for the respective control groups, reflecting the reducing size of the control group with expansion of the treatment group contour to cover greater distances, and suggesting a limitation on this method where the dataset is comparatively small.

Impact on residential property values

Descriptive analysis shows a similar increase in average property value between 1995 and 2016 for both treatment and control groups, especially in the period 1995 to 2002 (the immediate period spanning the intervention). However, the maximum price is higher in the treatment group dropping below that of the control group after 2001.

The difference-in-difference model indicates that a reduction in distance to the station has effected a significant increase in house prices especially for locations in the treatment group. Similarly, an increase in job accessibility and distance ratio suggest an increasing effect on house prices over the intervention period.

The fixed effects model indicates an impact on property prices of rail accessibility with distance to the nearest station having a negative effect as distance increases, similarly, distance to station ratio results indicate that an increase in distance ratio i.e. being proportionally nearer the station after the intervention, has a positive effect on property price as the relative change in distance increases. Also, an increase in the job accessibility index has a significant positive effect on property price.

Geographically Weighted Regression (GWR) indicates that the relationship between prices, distance to station and job accessibility is negative in some locations and in others positive. The model results show a significant advantage in using GWR over the cross-sectional OLS model, and indicate the presence of non-stationarity of relationships over the study area and the benefit of using GWR in this case study.

Impact on jobs and employment

Evidence from the work on accessibility demonstrates that accessibility varies spatially, having a varied effect on employment across the region. Initial comparison shows a higher drop in unemployment in the treatment group 1991 to 2001 and a higher increase in the control group 2001 to 2011. The difference-in-difference model indicates that a reduction in distance to the station has resulted in increased employment density over the intervention period. Similarly, an increase in job accessibility and distance ratio have also effected an increase in employment density. There is a consistently positive effect, regardless of different selections into treatment and whichever accessibility characteristic is used there is a pattern that suggests a 2 km threshold with either clustering or propensity matching offers the most suitable configuration.

The effect of employment density significantly increases marginally with a higher population density suggesting a difference between the urban environment and more remote places where there are fewer large centres of population, but with a higher density of jobs.

For GWR, the relationship between employment density and distance to the nearest station was in most cases negative - reflecting the OLS model and the presence of non-stationarity of relationships is clearly evident over the study area. Comparing 2001 with 2011 produces a different set of coefficients of the same order highlighting a spatial variation following the intervention. GWR modelling is valid in this type of situation and brings out the diversity between data zones across the region not evident in fixed effects modelling.

6 Chapter Six: The Stirling-Alloa case study

Following on from the previous chapter where the Robin Hood Line provided the case study, this chapter considers the impact of the reopening of the Stirling to Alloa Line which reopened ten years ago and runs through the Central Scotland corridor linking Alloa and Stirling and on to Glasgow and Edinburgh. The region mainly encompasses the county of Clackmannanshire with its mixture of rural and more remote communities. Through consideration of particular features of the region, the specific context of Stirling-Alloa can be compared to the other case study regions.

This chapter applies the methodology specifically outlined in Chapter 4 using the Stirling-Alloa Line as a case study example. This is contextualised through an overview of the geography of the region and its socio-demographic and economic profile and property market. Through division of the region into treatment and control groups, the impact on property price and employment density is economically modelled using three alternative accessibility indicators.

6.1 Stirling-Alloa Line and its region

The Stirling-Alloa line serves a previously isolated region classified as "urban with significant rural" now re-linked to the national rail network and the Glasgow-Stirling-Edinburgh axis. The rail intervention reopened a relatively small section of line, and has been in place for only 10 years and so offers limited opportunity to observe long term wider economic impacts, but a sufficient period to observe some employment and house price movements. Stirling-Alloa differs from the Robin Hood Line for the following reasons:

- The topography of the region makes it effectively a cul-de-sac.
- Data zones often cover large areas having a lower than average population.
- Just one new rail station has been created, so it effectively represents a single branch line from the main network.
- There are pre-existing neighbouring stations (e.g. Larbert and Dunblane).
- It is a previously isolated region now having much better links to a larger conurbation (Stirling) and to the much larger cities of Edinburgh and Glasgow.
- Edinburgh now has installed a tramway system since the introduction of the rail link.
- A "two way road effect" may lead to increased commuting outside the region added to some movement in retail and tourism.

Of the three case studies, Stirling-Alloa represents the most simplified picture owing to its lack of external confounding factors, and its mix of smaller communities previously remote from the rail network in a geographically cut-off location. Rather than having the potential to regenerate a whole region, it represents a short extension which makes the rail network more accessible to a limited number of locations in the vicinity of Alloa.

The Ochil Hills bound the northern part of an area dominated by Alloa and its road and links to Stirling in the west (Figure 6-1). Alloa lies on the north bank of the Firth of Forth 8.9 km east of Stirling with a resident population of 14,130 at the time of the 2011 UK census. The economy of the town once relied heavily on trade through its port with mainland Europe, but is now dependent on retail and leisure after the closure of the main industries of the town.

Figure 6-1 Topography of the Stirling-Alloa region (*Source: Google Maps, 2018)*

The restored line, between Stirling station and Longannet power station, reopens a 21 km stretch of the former Stirling-Dunfermline Railway which closed in 1968. Reopening of the route in 2008 reconnected Alloa to the national rail network for the first time in 40 years (Figure 6-2). The new line operates an hourly direct passenger service between Alloa, Stirling and Glasgow Queen Street stations with a 10 minute reduction in journey time, allowing passengers to change at Stirling for onward travel to Edinburgh. Passenger

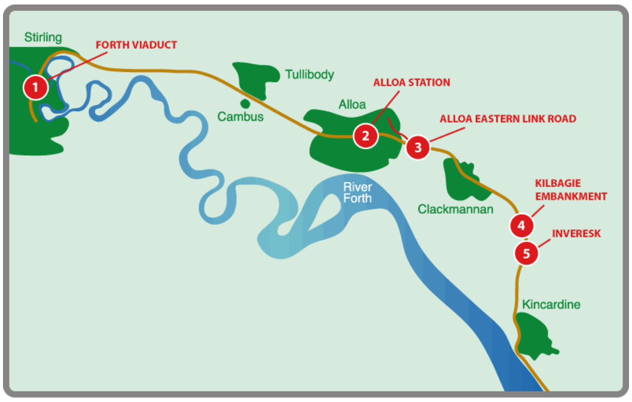


Figure 6-2 Stirling-Alloa rail link
(*Source: Transport Scotland, 2008)*

trains run on the new line only as far as Alloa, but there is potential to extend the line via Kincardine to Dunfermline and over the Forth Bridge to Edinburgh.

6.2 Definition of case study region

For the purpose of this analysis, the Stirling-Alloa region is illustrated in Figure 6-3 where the region studied is shown in white with a light grey boundary. This incorporates all datazones adjacent to the new section of the line and extends westwards to Stirling, eastward to Dollar and south to Kincardine. This allowed a mixture of areas close to the rail intervention and those further away to allow comparison of impacts later.

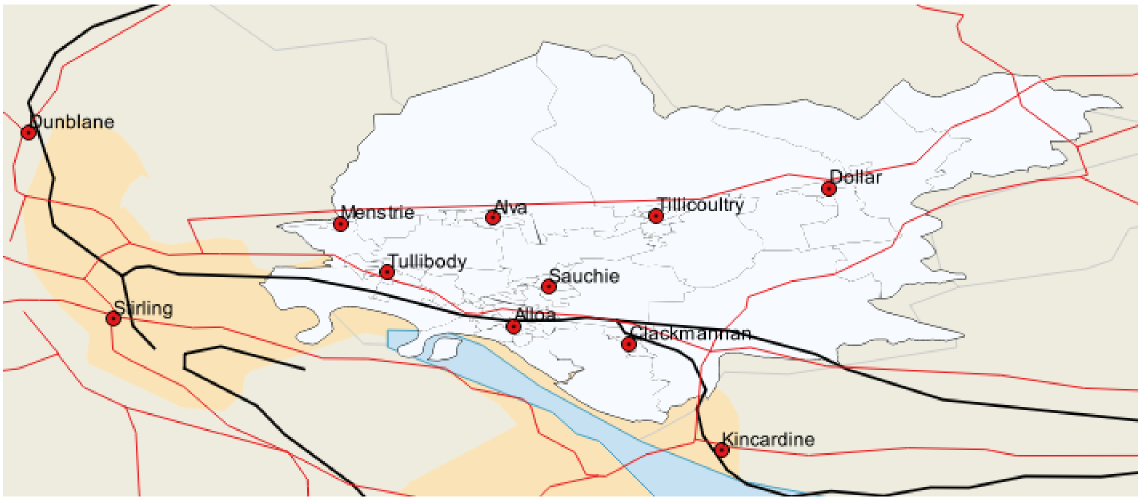


Figure 6-3 Stirling-Alloa Rail Case Study regional boundary

6.3 Population profile

Compared to the previous case study, the region (mainly Clackmannanshire) represents a much smaller location with a fairly static population. The 2014 population for Clackmannanshire was 51,190; down 0.2% from 51,280 in 2013 and accounted for just 1.0 per cent of the total population of Scotland. However, like the previous case study, it is skewed in terms of age group with 16.0% of the population aged between 16 and 29 years which is less than Scotland (18.3%) while those aged 60 and over make up 24.6%, again slightly more than Scotland (24.0%).

The region has declined economically over the years, and there are pockets of deprivation in the region with ten data zones classified as the 15% most deprived areas in Scotland. Alloa South and East remains the most deprived area with 5 data zones found in the 5% most deprived areas in Scotland. However, there are signs of improvement in the local economy with the region's annual GVA growth rate of 5.3% in 2013 (compared to 0.7% in 2012) well ahead of the rest of Scotland (2.9%).

Employment is dominated by the production sector (39%) and education and health sectors (22%) with regional employment levels (2015) at 71.6% compared to Scotland at 73.1%. Unemployment at 6.1% compares to Scotland at 6%. The workforce is slightly lower skilled than nationally with 17% of employees working in professional occupations compared to the Scotland and UK averages of 20%. This is reflected in educational attainment where the proportion of 16-64 year olds in Clackmannanshire with high level qualifications (36% at SCQF 7-12) is below the Scotland rate (41%) and those with no qualifications (11%) is higher than the Scotland rate (9%). Amongst 16-24 year-olds, 17% have higher-level qualifications (SCQF 6), again below the Scottish rate (27%).

6.4 The property market

Property prices can reflect the state of the local economy and, as seen in Figure 6-4, the market has slowly recovered to pre-2007 levels following a low in 2013. This follows a similar pattern to the national picture where the steep rise from 2000 peaks in 2007 followed by a gradual decline but picking up again in 2013. This period of decline coincides with the re-opening and early days of the Stirling-Alloa line, and needs consideration as a confounding factor in assessing causality. The latest figures from Rightmove show that Clackmannanshire, with an overall average price of £137,256, is similar in terms of sold prices to nearby Falkirk (County) (£139,090), but cheaper than Stirling (County) (£195,089) and Fife (£156,409). The most expensive area within Clackmannanshire is Dollar (£254,549) and the cheapest is Sauchie (£102,546).

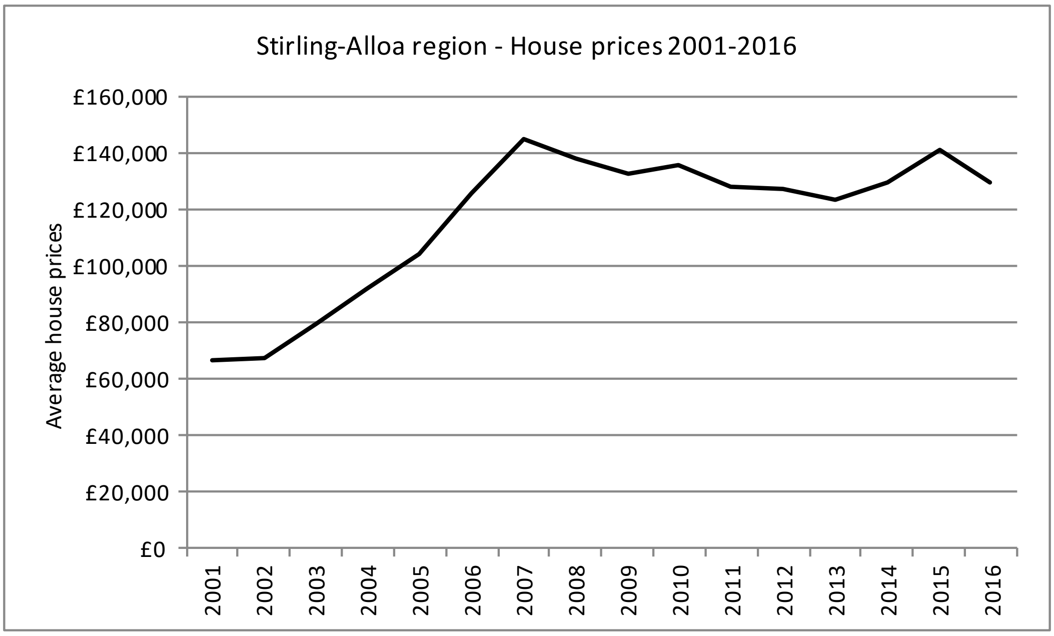


Figure 6-4 Stirling-Alloa house prices 1995-2016

(Source: Rightmove, 2017)

Transport provision

Analysis of the current transport situation in the Stirling-Alloa region provides further information about the variable levels of transport disconnection. The region is generally well served with buses, with frequent services between the various townships and Stirling. Over 80% of households have access to an hourly bus service within 10 minutes walking distance during the day, but this is not replicated evenly across the region where in evenings and Sundays this drops to just over 70%, and in rural locations just under 40% of households.

This disparity is also evident with access to the regional hospital - Forth Valley Royal. Direct services connect with the hospital throughout Monday to Saturday. From Alloa and Clackmannan, the direct service is hourly and takes up to 25 minutes, however, from areas such as Alva and Dollar, the service is only two hourly, and can take up to 55 minutes. In addition to the bus service connection to the Forth Valley Hospital, there is also a rail connection between Alloa and Larbert which links with the bus service at Larbert to the hospital. This runs approximately every hour for the whole week with a combined journey of approximately 35 to 40 minutes.

Similarly to the previous case study, 25% of households in the region do not possess a car, but this figure varies substantially when broken down by population groups from all owner occupiers (11%) to social sector (44%) (Scottish Government, 2015). Car ownership levels are lowest in districts where there are higher levels of deprivation, such as Alloa South and East, whilst rural areas of Clackmannanshire such as Dollar and Muckhart have some of the highest levels.

Following the reopening of the Stirling-Alloa line there is an hourly train from Alloa to Stirling and Glasgow running seven days a week, with an additional peak hour commuter service to Edinburgh running six days a week. The journey time to Stirling is 12–14 minutes. Scheduled passenger services operate only between Alloa and Stirling and onwards to Glasgow and Edinburgh. Previously rail commuters to Glasgow would require a multi-mode journey of bus and train to get to work.

Just under 90% of households are within 30 minutes of a GP surgery or health centre by public transport and 57% of households within 10 minutes travelling time. 84% of households are within 15 minutes travel time by public transport to a retail centre. 35% of households in the region are at least 10 km from the nearest hospital.

6.5 Treatment and control groups

As for the previous case study, the region is divided into treatment groups (likely to be more affected by the intervention) and control groups (less likely to be affected) in order to appraise the impact of the rail intervention and establish causality through creation of a meaningful counterfactual. The groups are constructed based on those locations subject to change in rail access across the intervention period following the methodology in *4.9 Selection of treatment and control groups.* As the region contains considerably fewer data zones than the Robin Hood Line, the groups themselves are consequently much smaller.

6.5.1 Change in distance

For the base groupings, selection into the treatment includes those data zones subject to a change in distance to the rail network for different contour boundaries i.e. 2 km contours radiating from Alloa up to a maximum of 10 km, affording five variations in base group specification.

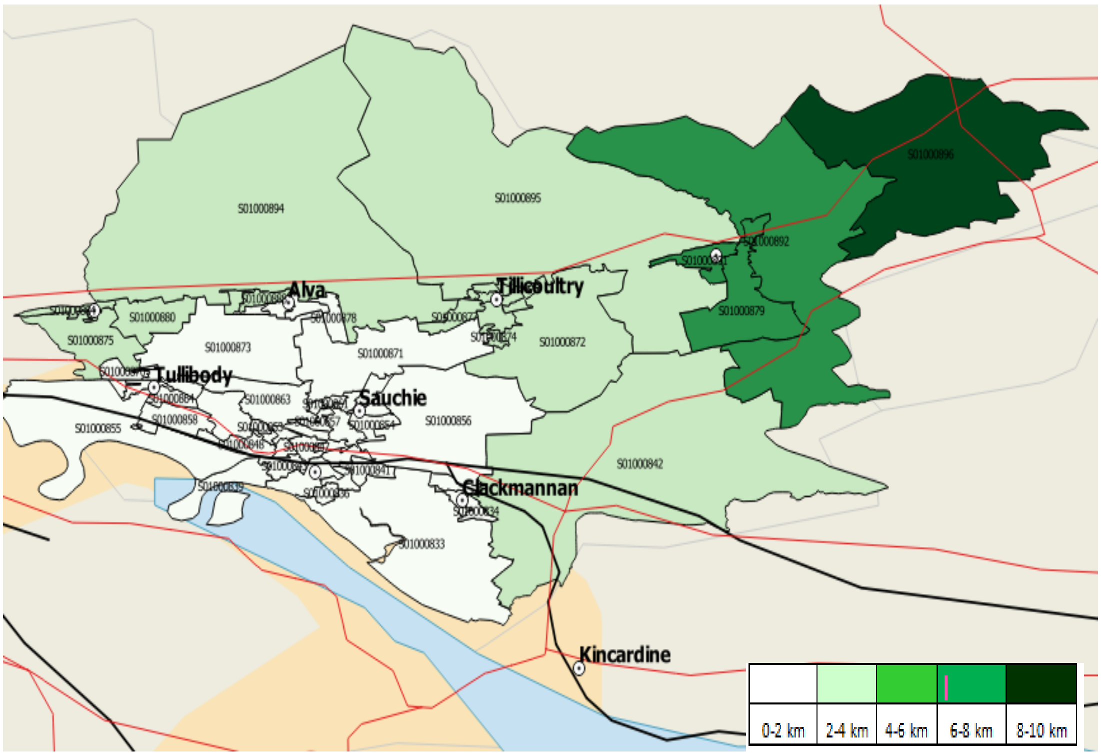


Figure 6-5 Variation in treatment and control group allocation: Stirling-Alloa

Figure 6-5 displays the different treatment groups based on alternative methods of selection where it is apparent that going to a 10 km contour would encompass the whole region as the treatment group, so it has not been feasible for this case study to consider distances beyond 6 km without greatly expanding the dataset.

6.5.2 Propensity score matching

As for the Robin Hood Line, the selection of the basic groups is expanded using propensity scoring matching techniques with nearest neighbour 1-to-1 matching. The comparators used for matching are data zone area size and the following deprivation levels taken from the Scottish Index of Multiple Deprivation (Scottish Government, 2016), health, education, crime, and housing.

In this case study, propensity matching is only applied in conjunction with a base group of 4 km for changes in distance to the nearest station for the reasons of dataset size mentioned previously. Matching treatment and control areas works very well using a 2 km distance threshold as the base as highlighted in the mean values for each comparator (Table 6-1) which indicate a marked high percentage balance improvement. However, matching is ineffective at the 4 km threshold and so is not considered for potential grouping.

Table 6-1 Summary of propensity matching output: Stirling-Alloa

Summary of balance for all data							
	Means Treated	Means Control	SD Control	Mean Diff	eQQ Med	eQQ Mean	eQQ Max
distance	0.543	0.2116	0.211	0.3314	0.3611	0.3326	0.5118
health	0.4864	-0.1291	0.5394	0.6155	0.65	0.6364	1.03
housing	13.9776	11.2841	5.8021	2.6935	2.93	2.886	4.03
educ	0.4644	-0.2412	0.976	0.7056	0.59	0.7385	1.85
crime	590.6	316.3148	238.278	274.2852	241	273.76	1506

Summary of balance for matched data							
distance	0.543	0.3763	0.2075	0.1666	0.174	0.1666	0.2778
health	0.4864	0.2528	0.3624	0.2336	0.19	0.2336	0.48
housing	13.9776	13.0748	4.6666	0.9028	0.88	1.0844	3.34
educ	0.4644	0.2415	0.6714	0.2229	0.22	0.2261	0.88
crime	590.6	449.04	252.031	141.56	70	151.96	1506

Percent Balance Improvement				
	Mean Diff	eQQ Med	eQQ Mean	eQQ Max
distance	49.7127	51.8189	49.9005	45.7299
health	62.0455	70.7692	63.2935	53.3981
housing	66.4826	69.9659	62.4255	17.1216
educ	68.4041	62.7119	69.3809	52.4324
crime	48.3895	70.9544	44.4915	0

In the "Summary of balance for all data" section before matching, the mean deprivation score for health is 0.60 more, housing 2.7 more, education 0.70 more, and crime 274 more in the treatment area than in the control area. After matching, in the "Summary of balance for matched data" the mean differences in deprivation score between treated and control areas reduce to 0.23 for health, 0.90 for housing, 0.22 for education, and 141.6 for crime. After matching, the treated and control areas are much more similar in terms of health, housing and education deprivation markers in particular. The rightmost columns in Table 6-1 also show the median, mean, and maximum quartile between the treated and control data; larger QQ values indicate better matching.

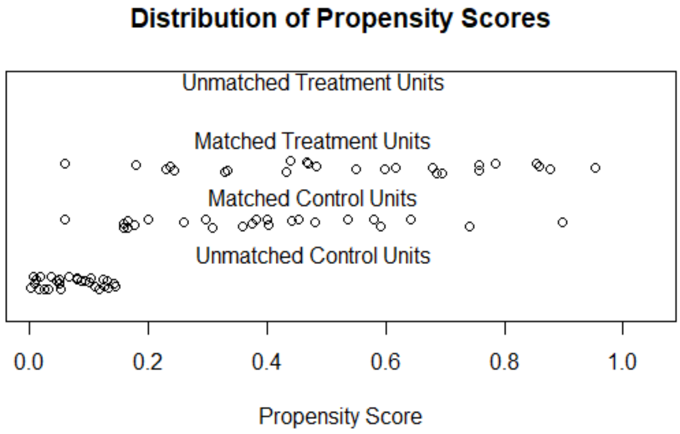


Figure 6-6 Distribution of propensity scores: Stirling-Alloa

Jitter plots and histograms help visualize the quality of the matching. In the jitter plot (Figure 6-6) the absence of cases in the uppermost stratification indicates that there are no unmatched treatment units. The middle stratifications show the close match between the treatment and the matched control units, and the final stratification shows the unmatched control units. Both the numerical and visual data indicate that the matching was successful.

Figure 6-8 compares the histograms before and after matching where those before matching (on the left). Before matching these are seen to differ to a great degree. However, the histograms after matching on the right are much closer in shape.

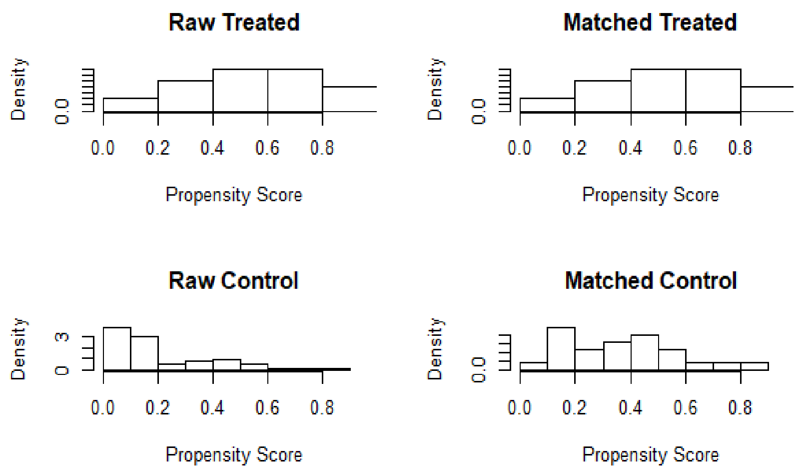


Figure 6-8 Histograms showing shapes before and after matching

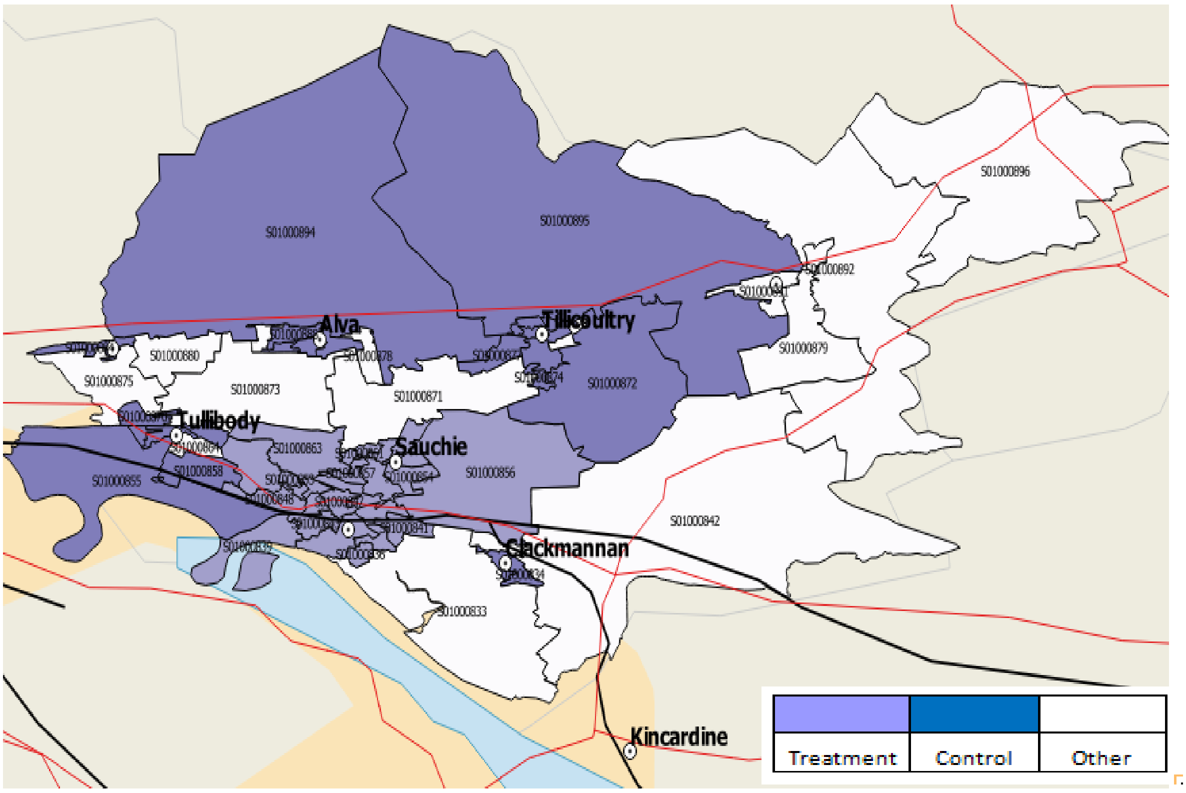


Figure 6-7 Treatment and control groups with propensity matching: Stirling-Alloa

Figure 6-7 illustrates the selection of treatment and control groups after the application of propensity testing based on a 2 km threshold.

6.5.3 Application of clustering

As an alternative to propensity matching, clustering techniques are again applied and unlike propensity testing do not require exact matching. This produces a further pairing of control and treatment group where again deprivation levels are used as comparators using a combination of clusters and the 2 km "base" group. A dendrogram provides visual evidence of the number of clusters (*Appendix 10.6.2*).

This indicates 3 distinct clusters as shown by the ovals which corresponds with the elbow chart (*Appendix 10.6.2*) and represents fewer clusters than the Robin Hood Line case study reflecting the smaller less diverse regional sample. Appendix 10.6.2 suggests three cluster groups to be considered in allocation of the treatment and control groups.

The membership of the cluster groups is shown in Table 6-2 and is mapped against the base groups at 2 km threshold in Figure 6-9 and Figure 6-10. This shows cluster group 1 to be marginally the dominant cluster, and following initial division into base groups using 2 km distance from the station as a basis for allocation, clustering is then applied by mapping matching clusters in the two groups.

Table 6-2 Cluster group memberships: Stirling-Alloa

Cluster Group	Membership
1	32
2	31
3	16

Figure 6-9 Clusters within treatment group: Stirling-Alloa

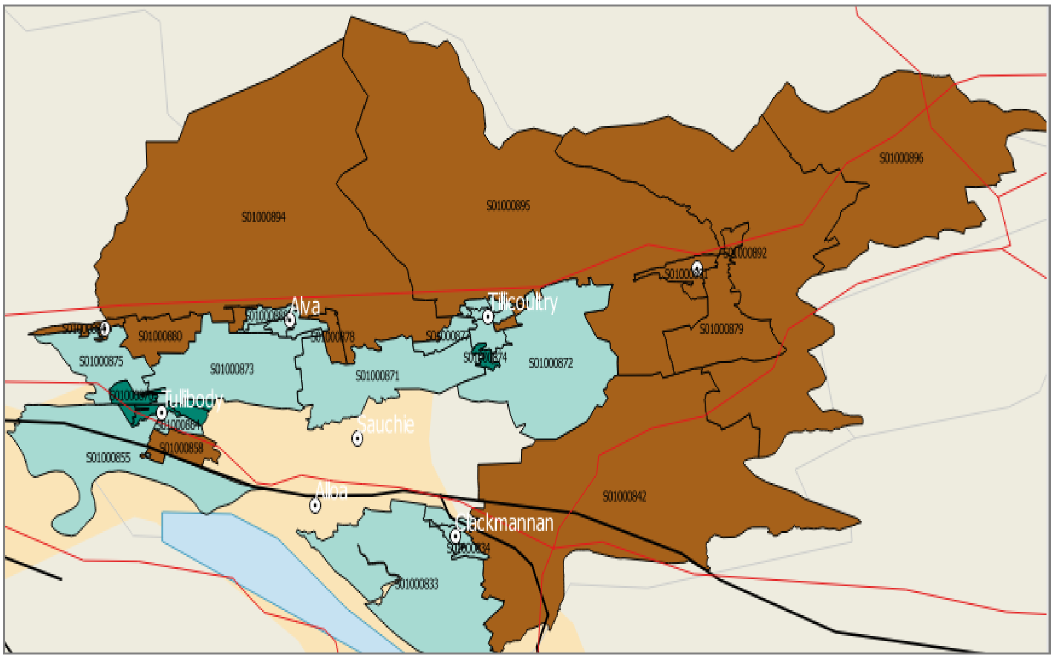


Figure 6-10 Clusters within control group: Stirling-Alloa

6.5.4 Treatment and control group combinations

Taking forward the different methods of allocation to treatment outlined in previous sections, the possible combinations of treatment and control group pairings used for further analysis are summarised in Table 6-3:

Table 6-3 Treatment/Control Group combinations: Stirling-Alloa

Treatment Group ID	Comments	Distance contour (km)	Status
TG1		2	
TG2		4	
TG3	Base groups	6	
		8	Not used
		10	Not used
TG4	TG1 with added propensity	2	
	TG2 with added propensity	4	Not used
TG5	TG1 with added Clustering	2	

Treatment/Control group combinations TG1 to TG3 represent selection to treatment based on increasing contour distance from the station where there has been a change in rail access. TG4 expands treatment groups TG1 with the addition of propensity matching, and TG5 is based on treatment group TG1 with the addition of clustering matched on the dominant cluster.

Three possible combinations were not suitable for sensitivity analysis because of the limited number of data zones (79) in the Stirling-Alloa dataset which produce groups too small to analyse, or provide an insufficient number of data zones in the control group without much greater expansion of the dataset. Hence only five combinations TG1 to TG5 are considered in the sensitivity analysis.

6.6 Accessibility characteristics

Following the division of the region into treatment and control groups, this section considers the three different measures for defining accessibility in the Stirling-Alloa region.

6.6.1 Initial descriptive analysis

Consideration of these indicators requires some prior knowledge of the background relevant to jobs and services in the Robin Hood Line region, in particular:

- Access to a car (car ownership)
- Travel to work patterns by mode and distance travelled
- Uptake of rail travel in Alloa through study of rail passenger usage
- Accessibility to essential services

Using the standard treatment and control group configuration (TG4) helps to contextualise the background and how it has affected "treated" areas. In the case of car ownership, (Table 6-4) over the intervention period (2001 to 2011) the treatment group shows an overall drop of 3.5% in households having no access to a car which is similar to the control group (3.40%).

At the other extreme, more households in the treatment group can now access two or more cars or vans (6.3%), which is higher than in the control group (5.0%). So, in the treatment group, proportionately more households now have access to a car, which may indicate a comparative improvement in economic status, or the benefit of using a car for quick access to the station.

Table 6-4 Changes in car ownership over the intervention period (2001-2011): Stirling-Alloa

Group	No cars or vans	1 car or van	2 cars or vans	3 cars or vans	4 or more cars or vans
Treatment	-3.5%	-2.7%	4.2%	1.5%	0.6%
Control	-3.4%	-1.6%	3.3%	1.3%	0.4%

This greater car ownership is reflected in means of travel to work (Table 6-5) where the increase in car driver travel is similar for both groups and mirrors the car ownership figures. The treatment group shows an increase of 1.58% in rail mode share against 0.54% for the control group. The decrease in bus mode is more noticeable in the control group (-4.81%) than the treatment group (-3.68%), which again could also suggest some multi-mode journeys in reaching the rail station.

Table 6-5 Changes in methods of travel to work over the intervention period (2001-2011): Stirling-Alloa

Group	Home	Tram	Train	Bus	Car Driver	Car Passenger
Treatment	4.35%	0.01%	1.58%	-3.68%	6.08%	-3.89%
Control	4.65%	-0.03%	0.54%	-4.81%	6.16%	-3.65%

As only a single average distance to work figure is available for 2001 and there is no information on the range of distances travelled, a comparison of distance travelled is not feasible. Over the intervention period the average distance to work has reduced by 4.56 km in the treatment group but increased by 2.49km in the control group. Table 6-6 shows distance travelled to work in 2011 by group and indicates more working from home in the control group, and more commuting to distances of under 10 km in the treatment group (56.46%) than the control group (48.13%).

This probably reflects the geographical nature of the region where there is a concentration of population and employment around Alloa (i.e. typically treatment group) and residents in more remote places (i.e. those typically in the control group) must travel further to work and are more likely to work from home.

Table 6-6 Distance travelled to work (2001-2011): Stirling-Alloa

Group	Work from home	Less than 2 km	2 km to Less than 5 km	5 km to Less than 10 km	10 km to Less than 20 km	20 km to Less than 30 km	30 km to Less than 40 km	40 km to Less than 60 km	60 km and over
Treatment	8.36%	20.77%	17.95%	17.74%	13.44%	2.98%	3.41%	3.44%	1.47%
Control	9.16%	8.59%	16.67%	22.86%	18.89%	3.28%	3.58%	4.75%	1.37%

When compared to similar pre-existing stations nearby, usage of Alloa station has grown more quickly over the initial period as shown in Appendix 10.6.1 for both Alloa and neighbouring stations by full fare, reduced fare and season ticket categories. Table 6-7 compares uptake since the rail re-opening between 2008 and 2016 of Alloa to Bridge of Allan, an existing station at a similar distance from Stirling. Although Bridge of Allan is on the main Glasgow to Perth line, there is a greater increase in both full and season tickets from Alloa, which may suggest an increase in regular commuting to Stirling and Glasgow.

Table 6-7 Comparison of rail usage (2008-2015): Stirling-Alloa

Full		Season Tickets	
Alloa	Bridge of Allan	Alloa	Bridge of Allan
127.6%	116.6%	140.6%	121.9%

Finally, comparing accessibility to a range of essential services over the period for treatment and control groups (Table 6-8), the average distance to all essential services is seen to be greater for the control group, but there is also a much greater spread in accessibility within the control group when comparing the minimum and maximum values for individual data zones.

Table 6-8 Distance to essential services over the intervention period (2001-2011): Stirling-Alloa

Groups	Nursery School			Primary School			Secondary School			Hospital		
	Average	Min	Max	Average	Min	Max	Average	Min	Max	Average	Min	Max
Treatment	0.83	0.27	1.52	0.47	0.16	0.99	1.40	0.37	2.60	9.20	7.92	10.74
Control	3.64	1.22	7.74	0.74	0.20	4.78	2.45	0.41	6.83	10.01	6.34	13.61

6.6.2 Distance to nearest station

A basic measure predominantly used in the literature is the distance to the nearest available rail station on the network calculated, here using the Euclidean method, then factored up to allow for the road network. For most locations in the Stirling-Alloa region this decreased following the intervention when Alloa station became the nearest accessible station to the rail network. This is more apparent in this case study as, unlike the Robin Hood Line, there are no confounding factors of nearby pre-existing rail stations prior to the intervention.

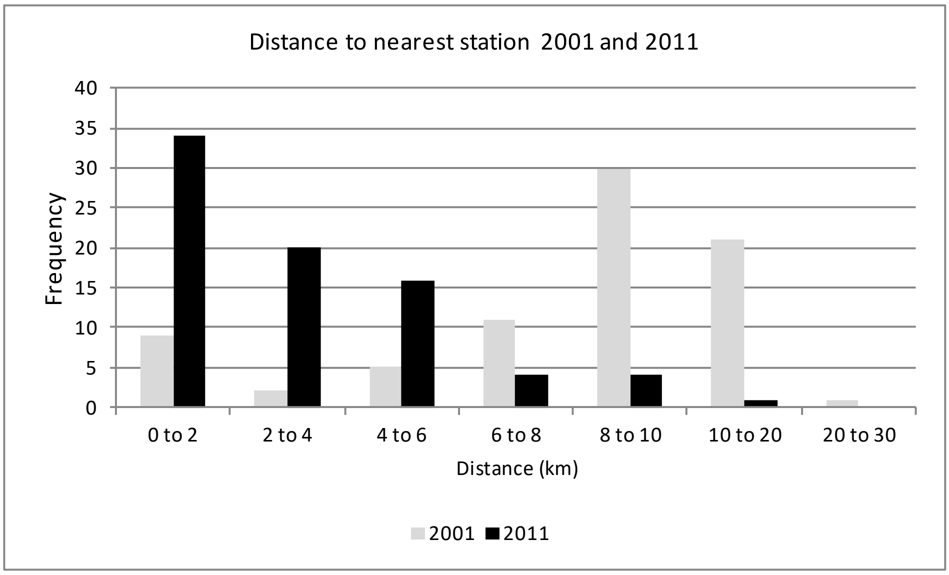


Figure 6-11 Distance to nearest station comparison (2001-2011): Stirling-Alloa

This is illustrated in Figure 6-11 which shows a sizeable shift so that in 2011 most data zones are now much nearer a rail station, and as an indicator of accessibility to rail mode it is a potential comparator for use in economic modelling.

6.6.3 Distance ratio

As a consequence of the above, the distance ratio which describes the percentage difference that the rail intervention has made in being comparatively closer to a rail station expressed as a ratio, has much more potential for this case study.

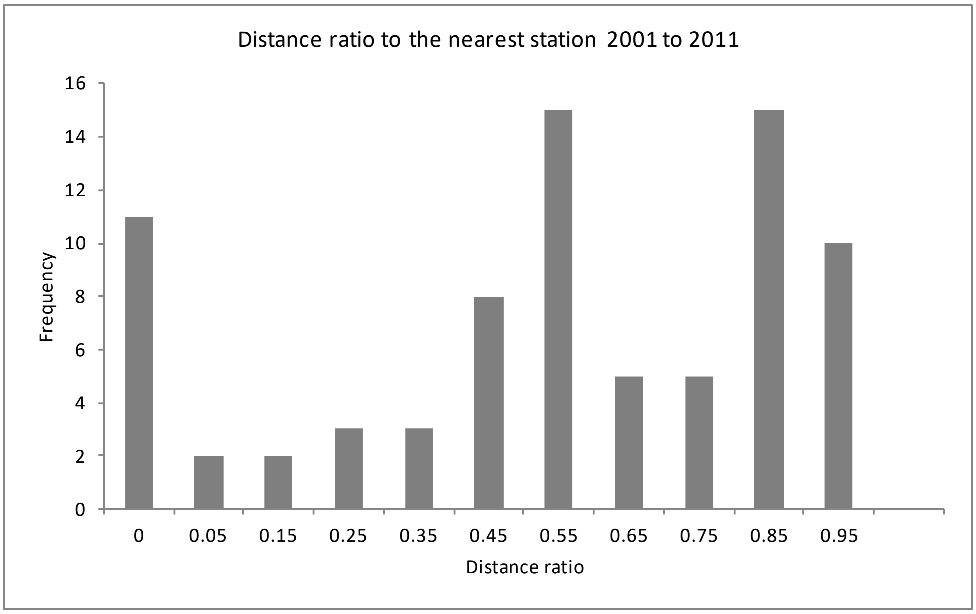


Figure 6-12 Distance ratio to nearest station comparison (2001-2011): Stirling-Alloa

Figure 6-12 indicates the predominant shift so that in 2011 most data zones are now much nearer a rail station, suggesting its potential applicability as an indicator of accessibility to rail mode for use in econometric modelling.

6.7 Job accessibility index

To measure job accessibility, this index is developed as applicable to Stirling-Alloa Line, based on considerations in *4.10 Accessibility to jobs.* This index is calculated with generalised travel time and generalised travel cost as proximity alternative measures in a generic job accessibility index. The time and cost values and decay parameter for the region are calibrated based on regional transport data estimates using the methodology outlined in *4.10.12 Calibrating the index and costs for each case study.* The index allows for commuting practicability and imposes a boundary in access to jobs, which is critical in the case of Stirling-Alloa representing a more remote and disconnected region. The index is alternatively measured based on consideration of all jobs available or only matching those suitable for the skills available in the region (*4.10.10 Job suitability - occupational matching).* Prior to calculation of accessibility, consideration must be given to the elements defining the measure.

6.7.1 Setting the standard parameters

As for the Robin Hood Line, a price comparison for individual modes of transport was derived using the methods prescribed in *4.10.12 Calibrating the index and costs for each case study.* Journey cost and times for Stirling-Alloa are calculated for estimated unit cost and speed values (Table 6-9).

Table 6-9 Cost and speed values used in index calculation: Stirling-Alloa

Transport Mode	Year	Transport Speed	Service Frequency	Headway	Cost of Travel km
Bus	2001	40	1	30	£0.10
	2011	40	1	30	£0.16
Car	2001	50	0	0	£0.29
	2011	50	0	0	£0.48
Rail	2001	65	1	30	£0.16
	2011	65	1	30	£0.25

As for the standard values, a decay parameter (β) is determined which produces the best fit for travel behaviour in the Stirling-Alloa region (*4.10.12 Calibrating the index and costs for each case study)*. Through analysis of distance to work statistics by data zone for 2001 and 2011, using 2001 as the base year (UK Census), Table 6-10 indicates that on average 49.36% of the working population travel at least 10 km to work, and this tails off gradually up to 40km (78.95%).

Distance	Average
> 2 km	87.58%
> 5 km	72.60%
> 10 km	50.64%
> 20 km	32.03%
> 30 km	26.18%
> 40 km	21.05%
> 60 km	13,89%
> 100 km	11.70%

Table 6-10 Decay effect of travel distance (2001-2011): Stirling-Alloa

From these calculations the impedance value is mapped against distance travelled (Figure 6-13) and by fitting an exponential decay function graphically, an overall value of β estimated. As seen in Figure 6-13 this has an R^2 value of 0.6094, and value of β of -0.028. This value is then used for calculating job accessibility.

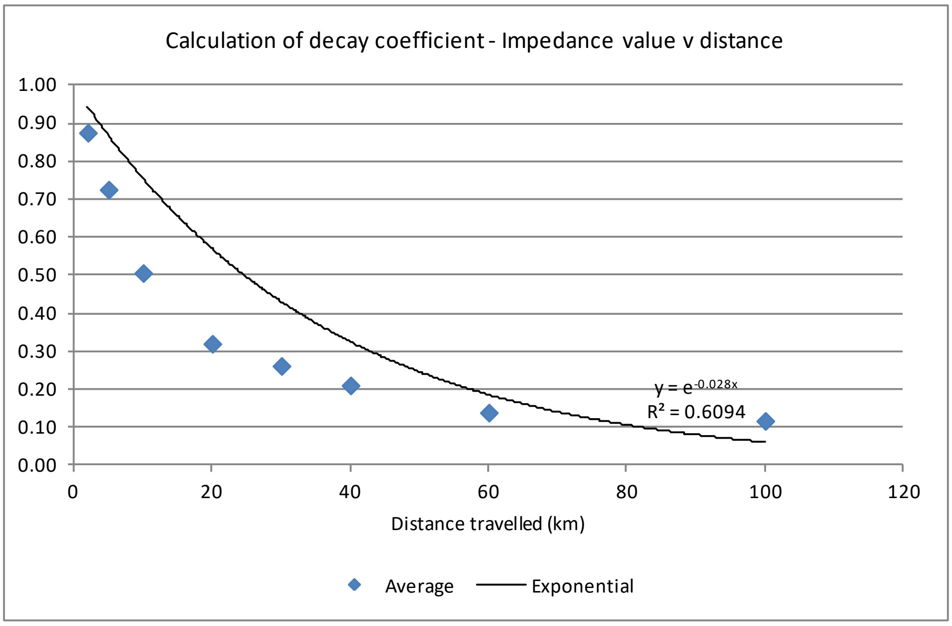


Figure 6-13 Calculation of decay coefficient (2001-2011): Stirling-Alloa

This initial value of β is based on distance only, and is adjusted when calculating job accessibility, where time or cost provides the basis for proximity and is dependent on the particular mode of transport, as indicated in *4.10.12 Calibrating the index and costs for each case study*. This is reflected in the average speed and cost of that mode in the region.

A further consideration is the imposition of a threshold beyond which job opportunities are excluded. As the negative exponential function is short tailed, long distances have limited effect on accessibility estimation, so function truncation may not result in any important loss of information. Because every opportunity contributes to the potential value calculation, a maximum distance can be defined (*4.10.9 Thresholds*).

However, in the special case of Stirling-Alloa, exceptions are made here for Stirling and Glasgow, which although outside the case study region, provide a powerful pull in terms of job opportunities, more relevant since the reintroduction of rail travel from Alloa. Employment data for Stirling and Glasgow have been incorporated into the accessibility calculations, and the threshold distance defined using commuting data where the maximum travel distance observed in the region (100 km) may be greater following the rail intervention.

6.7.2 Job accessibility comparison

Having determined the specific cost and decay parameters for the Stirling-Alloa, the next stage was to compare the job accessibility index alternatively based on travel time and travel cost before and after the intervention, broken down further by applying either job skills matching or no matching after allowing for commuting thresholds. There are two comparisons that job accessibility reflects:

1. Infrastructure changes only by using the job situation prior to the intervention as the basis for job opportunity and skills matching.

2. Infrastructure changes and movement of jobs by factoring in the prevailing job situation in post-intervention years.

Infrastructure changes only

Using both a travel time and travel cost basis and without skills matching, accessibility to rail mode increased post-intervention for both the treatment group and control group (Table 5-11), but was marginally greater for the former on a cost basis and lower on a time basis. This suggested that there were more benefits from improvements using cost as a basis.

Table 6-11 Method 1: Change in accessibility rail (2001-2011): Stirling-Alloa

		with skills matching			without matching			Matching Effect (Ratio matching to no matching)	
Group	Method	pre	post	% change	pre	post	% change	pre	post
Treatment	Cost	0.022	0.039	80%	0.212	0.380	80%	9.802	9.774
	Time	0.075	0.088	17%	0.740	0.867	17%	9.811	9.836
Control	Cost	0.019	0.033	71%	0.192	0.327	70%	9.916	9.874
	Time	0.070	0.085	23%	0.692	0.850	23%	9.919	9.942

Applying skills matching accessibility to rail mode suprisingly produced similar percentage increase as that estimated with no matching, which may suggest that because the region studied is small and compact there is a similar skills set across the region. Pre- and post-intervention differences show an impact due partially to the change in proximity brought by rail. All results are aggregated job accessibility indexes for all locations within the Stirling-Alloa region, and suggest that in this case, job skills matching provides a similar relative change in accessibility.

There is evidence of this when dividing the unmatched jobs accessibility index from this "skills match" index to provide a 'matching effect' which indicates the relevance of occupational matching in the calculation of job accessibility. This ratio of matching to non-matching values of the index is very similar across the board. The wider difference when the index is based on travel cost rather than travel time, suggests that the index is more sensitive to cost, which may imply that it is more of an impediment.

Infrastructure changes and movement of jobs

In this comparison, using both a travel time and cost basis and without skills matching, accessibility to rail increased only marginally post-intervention for both the treatment group and control group (Table 6-12) with no significant difference in time and cost bases.

However, with the application of skills matching, accessibility to rail mode increased marginally more post-intervention for both groups with little difference whether a cost or tme basis was used. For all other modes, there was negligible change on both a time basis and cost basis.

Table 6-12 Method 2: Change in accessibility rail (2001-2011): Stirling-Alloa

		with skills matching			without matching			Matching Effect: Ratio of matching to no matching	
Group	Method	pre	post	% change	pre	post	% change	pre	post
Treatment	Cost	0.038	0.041	7%	0.380	0.383	1%	9.774	9.251
	Time	0.088	0.093	6%	0.867	0.867	0%	9.833	9.292
Control	Cost	0.033	0.035	6%	0.326	0.329	1%	9.864	9.366
	Time	0.085	0.090	6%	0.849	0.851	0%	9.937	9.409

Appendix 10.6.3 maps job accessibility based on travel cost and compares predictions for pre- and post intervention values using alternatively no matching (i.e. all job opportunities) and skills matching (i.e. only those opportunities matching). Pre- and post-intervention differences show very little impact due to change in proximity brought by rail, despite interference from the slump in the job market up to 2011. All results are aggregated job accessibility indexes for all locations within the Robin Hood Line region, and this again suggests that without job skills matching there may be again an overestimation of accessibility due to the seemingly high attraction of job opportunities. The application of job skills matching highlights the reduction in levels nearer to the line because of fewer jobs available which did not match the skills profile of residents in those locations.

6.8 Essential services accessibility index

An appropriate accessibility index for essential services is also derived based on *4.11 Accessibility to essential services.* This index measures access to five essential services and reflects the particular characteristics of the region, estimating how rail developments have impacted on levels of accessibility through analysis by treatment and control group.

Analysis of changes in the rail station accessibility index can be found in Appendix 10.6.4 alternatively using time and cost as the basis, reflecting average accessibility across the case study region for treatment group TG4. On a time basis, average accessibility pre-intervention is similar for both treatment and control group for all services apart from nursery schools. Following the intervention, this increases more across all services for the treatment group (2.6%) compared to the control group

(1.6%).This also holds on a cost basis (Appendix 10.6.4) with a greater increase for the treatment group (31.6%) compared to the control group (15.7%). Spatial analysis reveals variations in the accessibility index across the region for individual data zones post intervention. The spread of accessibility as shown by the minimum, maximum and standard deviation is greater in the control group for nursery schools, secondary schools and hospitals.

6.9 Impact on residential property values

The next stage in the process is to carry forward the treatment group configurations and accessibility characteristics applicable to Stirling-Alloa into the property econometric models using three modelling approaches: fixed effects, difference-in-difference and GWR (Geographically Weighted Regression). Stirling-Alloa differs from the Robin Hood Line in being basically a branch line from Stirling and it may be problematic in discerning any house price movements due to the rail intervention, and benefit may be from greater commuting potential to the job market in Glasgow.

6.9.1 Background to the models

Before consideration of each of these model approaches, some background relevant to property price movements in the Robin Hood Line region will validate the variables applicable to the models. Following the methodology outlined in *4.12 Impact on residential property values* the situation before and after rail intervention is compared using the standard treatment/control group configuration (TG4) to compare changes in property type and house price.

The type of property e.g. terraced, detached in each data zone is one factor that will influence property prices as areas with a higher number of detached properties will generally reflect a higher average house price. The property type profile for treatment and control groups for the period 2001 to 2011 (Table 6-13) indicates a decreased percentage of terraced properties (2.22% for the treatment group and 1.36% for the control group). This is against a 3.75% increase for treatment and 2.67% increase for control for detached properties, and a 1.75% increase for treatment and 0.59% increase for control for semi-detached properties.

Table 6-13 Accommodation profile treatment v control (2001-2011): Stirling-Alloa
(*Source: UK Census, 2001/2011)*

Group	Year	Detached	Semi-detached	Terraced	Flat etc.
Control	2001	15.40%	27.28%	24.89%	32.43%
	2011	18.17%	27.03%	23.53%	31.28%
Treatment	2001	12.72%	28.10%	25.66%	33.53%
	2011	16.47%	27.65%	23.44%	32.44%

Figure 6-14 charts the increase in average house prices between 2001 and 2016, with a dip between 2010 and 2013. Between 2008 and 2010 (spanning the immediate period either side of the intervention), there is a marginally higher increase for the treatment group echoed in the movement of minimum prices, but the maximum price is higher in the treatment group dropping below that of the control group after 2011.

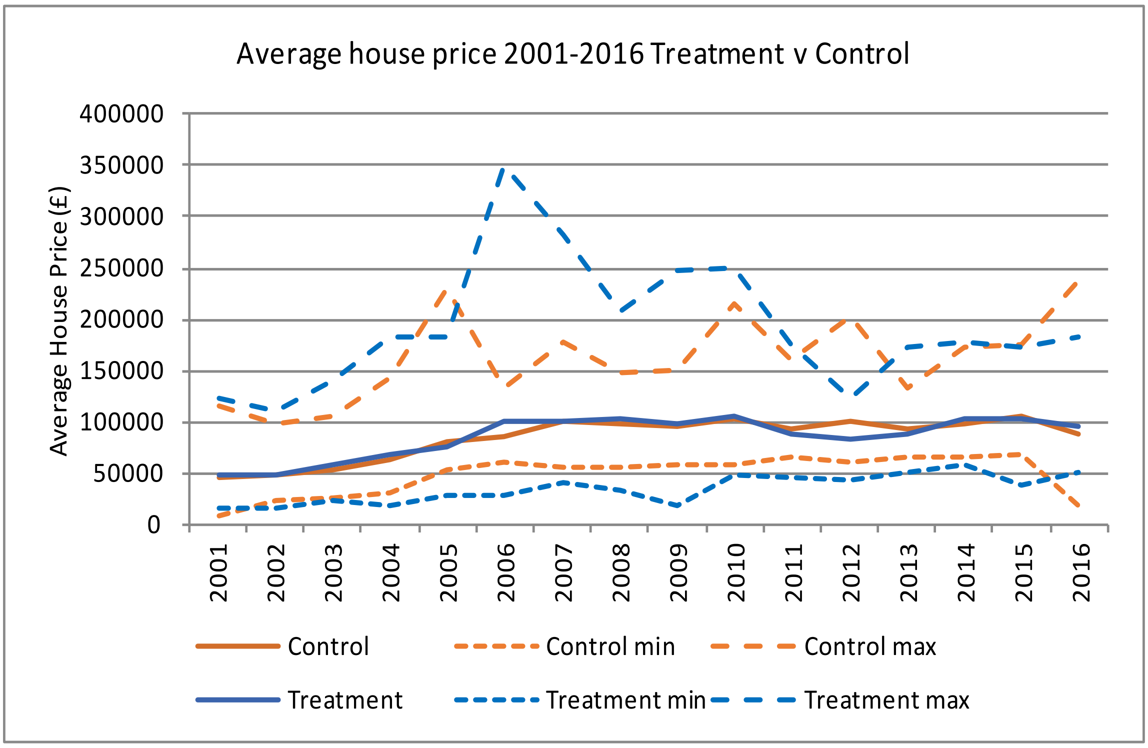


Figure 6-14 Average house prices 2001-2016 - treatment v control: Stirling-Alloa
(*Source: Rightmove, 2017)*

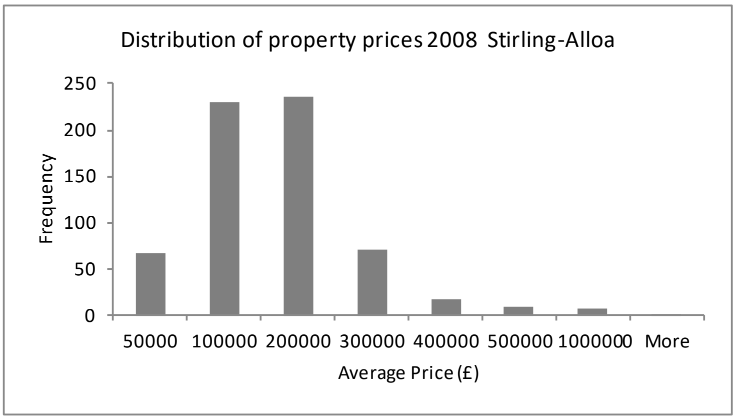


Figure 6-15 Distribution of property prices

(*Source: Rightmove, 2017)*

The regional perspective is illustrated in Figure 6-15 showing the distribution of property prices in the region in 2008 prior to the rail intervention where a clear distribution shape indicates a small number of very costly properties. The distribution is strongly positively skewed with a central group of typically-priced houses together with a long tail of relatively expensive ones. The summary statistics are mean (£140273), median (3107334), standard deviation (£116386), and skewness (3.737).

6.9.2 Difference-in-difference model

The difference-in-difference model compares two separate years pre- and post-intervention to predict what would have happened had there been no intervention and applying three different variables of accessibility.

Table 6-14 DID model output - distance from station (2001-2011): Stirling-Alloa

	All areas				All areas with treatment variable			
	Coefficient	SE	t value		Coefficient	SE	t value	
(Intercept)	0.6116	0.259	2.356	*	0.3591	0.1592	2.256	*
Distance to nearest station	0.0327	0.027	1.174					
Terraced	0.0041	0.006	0.627		0.0049	0.0066	0.744	
Detached	-0.0019	0.003	-0.548		-0.0021	0.0035	-0.608	
Semi	0.0025	0.004	0.588		0.0028	0.0043	0.652	
No car households	-0.0024	0.004	-0.577		-0.0032	0.0042	-0.771	
1 car household	0.0009	0.004	0.204		-0.0003	0.0041	-0.080	
% employed	-0.7192	1.153	-0.624		-0.5845	1.1730	-0.498	
Education Level 1	0.0038	0.003	1.447		0.0035	0.0027	1.318	
Education Level 4	-0.0030	0.003	-1.008		-0.0020	0.0029	-0.691	
No qualifications	-0.0004	0.003	-0.155		0.0000	0.0026	0.019	
Population density	-0.0040	0.013	-0.317		-0.0042	0.0129	-0.324	
Treated					-0.0028	0.0939	-0.030	
R^2	0.1548				0.1184			

Distance to nearest station

Firstly, with distance to nearest rail station as the accessibility variable, Table 6-14 compares 2001 to 2011 using distance to station as the accessibility variable. This has an R^2 value of 0.155, and for treated locations a reduction in distance to the station has not increased house prices over that period.

Distance ratio

Secondly, with distance ratio as the accessibility variable, Table 6-15 compares 2001 to 2011. This has an R^2 value of 0.1257 and indicates that a percentage reduction in distance to the station again does not produce an increase in house prices over that period.

Table 6-15 Output from DID model using distance ratio (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t value	
(Intercept)	0.4744	0.2688	1.765	.
Distance Ratio (DSR)	-0.1500	0.2894	-0.518	
Terraced	0.0047	0.0066	0.720	
Detached	-0.0021	0.0034	-0.603	
Semi	0.0027	0.0043	0.637	
No car households	-0.0028	0.0042	-0.684	
1 car household	-0.0001	0.0041	-0.026	
% employed	-0.6469	1.1733	-0.551	
Education Level 1	0.0037	0.0027	1.381	
Education Level 4	-0.0023	0.0029	-0.778	
No qualifications	-0.0001	0.0026	-0.039	
Population density	-0.0036	0.0129	-0.283	
R^2	0.1257			

Job Accessibility Index

Using job accessibility with cost basis and skills matching and ignoring movement of labour gives an R^2 value of 0.1323 (Table 6-16). Although an increase in job accessibility is reflected in a positive coefficient, it cannot be seen to indicate a positive impact on house prices as the t-value of 0.718 suggests it is not significant.

Table 6-16 Output from DID model using job accessibility (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t value	
(Intercept)	0.1574	0.3138	0.502	
Job Accessibility Index (JAI)	50.1100	69.8000	0.718	
Terraced	0.0047	0.0065	0.723	
Detached	-0.0018	0.0034	-0.528	
Semi	0.0035	0.0044	0.802	
No car households	-0.0041	0.0043	-0.963	
1 car household	0.0000	0.0041	0.011	
% employed	-0.5038	1.1680	-0.431	
Education Level 1	0.0040	0.0027	1.466	
Education Level 4	-0.0016	0.0029	-0.559	
No qualifications	0.0005	0.0026	0.203	
Population density	-0.0054	0.0129	-0.422	
R^2	0.1323			

6.9.3 Fixed effects model

The Standard Fixed effects model was applied using the case study datasets for the Stirling-Alloa. The aggregate data for the Stirling-Alloa Line spanning the years 2001 to 2016 for each data zone was collated from individual house transactions into an aggregated property price database and input into a standard model using the three different variables of accessibility mentioned previously.

Distance to the nearest station

Firstly, using distance to the nearest station as the accessibility characteristic, and log of property price as the dependent variable, there is an R^2 value of 0.627 (Table 6-17).

Against a background of fixed location and time effects, the distance to station variable has a coefficient of -0.0534, indicating that distance to the nearest station has a negative effect on property price as distance increases. For a property at the £100,000 level there is a reduction of £5200 for each 1 km distance from the station.

Table 6-17 Model output with distance to nearest station as accessibility characteristic (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t-value	
Distance to nearest station	-0.0534	0.0180	-2.971	**
Terraced	0.0073	0.0078	0.935	
Detached	-0.0085	0.0039	-2.206	*
Semi	-0.0052	0.0061	-0.864	
No car household	-0.0074	0.0049	-1.493	
1 car household	0.0084	0.0046	1.835	.
% employed	-3.9646	1.7961	-2.207	*
Level 1 education	0.0011	0.0039	0.287	
Level 4 education	0.0053	0.0020	2.665	**
No qualifications	-0.0022	0.0037	-0.594	
Population density	0.0132	0.0247	0.534	
$R^{2:}$ 0.62778				
Fixed effects - locations (79) time (16) Number of observations: 1264				

Other relevant factors include car ownership (where lack of car access has a negative effect on property price), and education levels (where lack of qualifications also reflects negatively on price, but higher education levels have a positive effect).

Distance to station ratio

The model run results with distance to station ratio as the accessibility characteristic also corroborate the previous findings and indicate that a percentage improvement in accessibility is also a relevant measure (Table 6-18).

This measure representing the improvement in accessibility through a change in distance to the nearest rail station due to new rail infrastructure, gives an R^2 value of 0.626 which is similar to distance to the station. The distance ratio has a coefficient of 0.537 which indicates that an increase in distance ratio i.e. being proportionally nearer the station after the intervention, has a positive effect on property price as the relative change in distance increases.

For example, if 10% nearer the station a property at the £100,000 level will increase in price by £5520. This again shows that having no car or lack of qualifications have a negative effect on property price, with higher levels of education having a positive effect.

Table 6-18 Model output with distance ratio as accessibility characteristic (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t-value	
Distance ratio	0.5371	0.1837	2.924	**
Terraced	0.0073	0.0078	0.926	
Detached	-0.0090	0.0038	-2.343	*
Semi	-0.0057	0.0060	-0.935	
No car household	-0.0076	0.0049	-1.544	
1 car household	0.0091	0.0046	1.981	.
% employed	-3.8485	1.8221	-2.112	*
Level 1 education	-0.0021	0.0037	-0.573	
Level 4 education	0.0006	0.0040	0.158	
No qualifications	0.0056	0.0020	2.828	**
Population density	0.0116	0.0247	0.471	
R^2: 0.62644				
Fixed effects - locations (79) time (16) Number of observations: 1264				

Job accessibility index

Finally running the model with a cost-based job accessibility index as an alternative accessibility characteristic takes into account distribution of jobs and the cost of reaching those jobs as a deterrent. In this model, the percentage employed is omitted as a contributory factor as it is already incorporated into the job accessibility index and may cause endogeneity.

Again using log of property price as the dependent variable gives an R^2 value of 0.57 (Table 6-19), and yields a coefficient of 61.60, which indicates an increase in the job index has a positive effect on property price. For example, an increase in accessibility of 1% would generate an increase in property price from £100,000 to £105,350 or 5.35%.

Table 6-19 Model output with job accessibility as accessibility characteristic (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t-value	
Job accessibility	61.6000	17.8800	3.440	***
Terraced	0.0002	0.0080	0.028	
Detached	-0.0114	0.0040	-2.879	**
Semi	-0.0045	0.0065	-0.697	
No car household	-0.0043	0.0053	-0.808	
1 car household	0.0074	0.0049	1.529	
No qualifications	-0.0048	0.0039	-1.229	
Level 1 education	0.0071	0.0039	1.826	.
Level 4 education	0.0055	0.0024	2.317	*
Population Density	0.0029	0.0261	0.112	
Fixed effects - locations (79) time (16)				
R^2: 0.5702				

Where there is greater percentage of households with no access to a car, the property price is relatively lower. Also where a higher percentage of residents have no qualifications, property prices will again be lower.

6.9.4 GWR Model

As an alternative approach, Geographically Weighted Regression (GWR) is used to model changes in spatial diversity over a period of time spanning the rail intervention. The data frame contains census information collected for 79 data zones in the Stirling-Alloa Line area with a grid reference marking the population weighted centre for each zone. The analysis uses house price data representing a sample of houses transacted in the region between 2001 and 2016. In contrast to a single set of constant values over the study area generated by an OLS model, GWR produces a set of parameter estimates and model statistics for each sample.

As for previous models, the following distance to nearest station and job accessibility provide an alternative accessibility characteristic for both the GWR and global (OLS) models. The models are run for each of 2001 and 2011 to compare coefficients both spatially (across the data zones) and temporally for the two separate years with focus on 2011 difference-in-difference employment results. Detailed runs are shown in Appendix 10.6.6. Table 6-20 provides a summary of the parameter estimates fitted by GWR for 2001 and 2011, including the median, upper quartiles, lower quartiles, range, minimum and maximum and compares this with the OLS global model. This compares the GWR model coefficients over the two separate different years and applying different accessibility characteristics to the model.

By varying the search window across the case study area, it is apparent that there is a negative relationship between price and distance to the nearest station which varies across the region. Using distance to nearest station, the coefficients ranged from a minimum value of -0.0251 to a maximum of -0.0101. So a 1 km increase in distance from the station results in a drop in average house price which varies between £2500 (2.5%) and £1000 (1%) for a £100,000 house. The analysis compares the GWR model results with the cross-sectional OLS model (which represents the global model) to detect any significant advantage in using GWR which involves reporting an improvement in fit of a GWR model over a global model using an F test to check significance. It is clearly evident that all of the interquartile values were less than their corresponding 2 standard errors of the global estimates (Table 6-20) indicating the lack of non-stationarity of relationships over the study area.

Table 6-20 Spatial variation of coefficients for different accessibility characteristics (2001-2011): Stirling-Alloa

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	SE	t value	R^2
Distance nearest station										
2001	-0.025	-0.023	-0.021	-0.016	-0.010	0.006	-0.017	0.017	-0.980	0.485
2011	0.015	0.020	0.023	0.025	0.033	0.005	0.026	0.016	1.638	0.681
Job Accessibility										
2001	3.064	8.203	16.490	23.540	30.750	15.337	4.894	21.327	0.229	0.478
2011	-16.190	-12.850	-11.110	-10.510	-9.657	2.340	-15.670	10.800	-1.451	0.678

6.10 Impact on jobs and employment

A similar methodology is used to assess the impact of the rail intervention on jobs and employment by including the different defined accessibility measures as accessibility characteristics in econometric models of property price. Unlike the property model, two modelling approaches are considered: difference-in-difference and GWR (Geographically Weighted Regression).

6.10.1 Initial descriptive analysis

Before embarking on the models, some prior knowledge of the background relevant to jobs and employment in the Stirling-Alloa region is required to validate variables that may be applicable. Following the methodology outlined in (*4.13 Impact on jobs and employment*), an analysis evaluates the situation before and after rail intervention with reference to the standard treatment/control group (TG4), comparing changes in education levels and qualifications, economic activity, job profile by industry and occupation, and population levels.

Educational levels and qualifications represent potential for employment, and Table 6-21 highlights changes in educational standards over the rail intervention period by treatment and control group which show a similarity between the two groups, but a greater reduction in those having no qualifications and more possessing level 3 and above qualifications in the treatment group.

Table 6-21 Changes in educational standards (2001-2011): Stirling-Alloa

(*Source: UK Census, 2011)*

Group	No Qualifications	Level 1	Level 2	Level 3	Level 4 and above
Treatment	-7.43%	-1.22%	-0.12%	3.46%	5.30%
Control	-6.61%	-1.06%	-0.71%	2.73%	5.65%

In terms of employment, Table 6-22 highlights changes in employment levels by group over the intervention period showing a greater increase in unemployment in the control group, with net employment (full, part-time and self-employed) decreasing less sharply in the treatment group (-2.37%) than in the control group (-3.70%), In addition, the increase in percentage of part-time workers in the treatment group is more than double that in the control group.

Table 6-23 Jobs by occupation - treatment v control (2001-2011): Stirling-Alloa

(Source: UK Census, 2011)

	Treatment	Control
Managers directors etc.	-1.00%	-0.70%
Professional	0.43%	-0.15%
Associate professional and technical	-0.49%	0.78%
Administrative secretarial	0.09%	-0.65%
Skilled trades	0.98%	3.39%
Caring leisure and other services	-1.17%	-0.33%
Sales and customer service	-0.30%	-0.28%
Process plant and machine operatives	1.12%	-1.60%
Elementary occupations	0.34%	-0.46%

Occupation is an important element in the job accessibility index when comparing matching of job opportunities between different locations, particularly in deriving a skills matched job accessibility index. The treatment group shows an overall increase in the percentage of residents with professional occupations against a decrease for the control group (Table 6-23), but the reverse is true for the associate professional and technical. Although skilled trades have increased for both groups, this is more marked for the control group (3.39%).

Table 6-24 highlights changes in types of industry over the period, indicating a move away from traditional industries of manufacturing and mining which is slightly higher in the treatment group (-5.05%) than the control group (-4.28%). There has been a similar move towards the public sector, e.g. public administration, health and social work and education, and this is generally more marked in the treatment group.

Table 6-24 Jobs by industry - treatment v control (2001-2011): Stirling-Alloa
(*Source: UK Census, 2011)*

	Treatment	Control
Agriculture hunting and forestry	-0.23%	-0.42%
Fishing	-0.07%	-0.01%
Mining and quarrying	-0.25%	-0.54%
Manufacturing	-5.05%	-4.28%
Electricity Gas and Water Supply	0.88%	1.13%
Construction	0.48%	-0.51%
Wholesale and retail trade repairs	1.34%	0.62%
Hotels and restaurants	0.57%	0.56%
Transport storage and communications	-1.07%	-0.64%
Financial intermediaries	0.07%	-0.66%
Real estate renting and business activities	-8.48%	-7.97%
Public administration and defence	4.98%	3.60%
Education	0.32%	1.85%
Health and social work	3.28%	2.84%
Other	3.21%	4.43%

To place this in context, analysis of change in population level by control and treatment group between 2001 and 2014 (Figure 6-16) reveals a gradual decline for the control group over the period, but a more noticeable increase for the treatment group between 2002 up to 2008, followed by a slow decline post 2008. There is a similar pattern in the period following the intervention, with the control group levels being marginally higher throughout. This does not suggest any substantial changes between the groups that may explain some of the differences exhibited elsewhere.

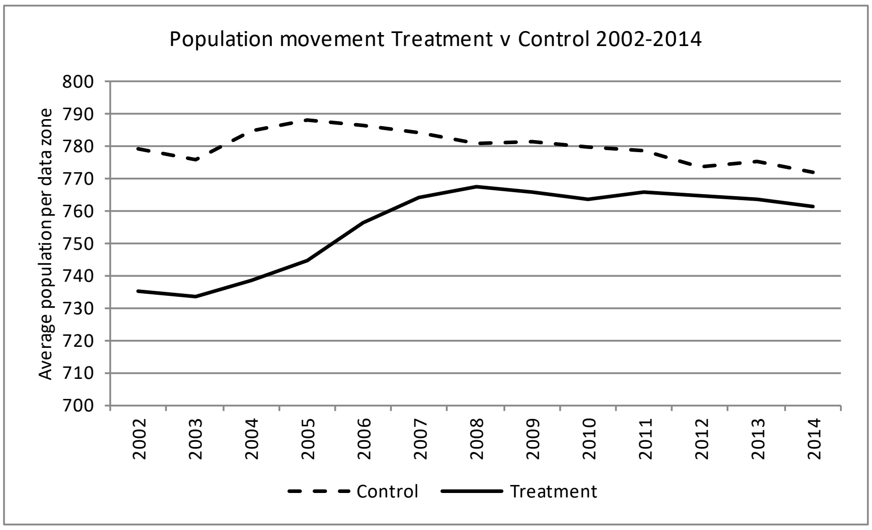


Figure 6-16 Population levels 2002-2014 (2001-2014): Stirling-Alloa

Figure 6-17 shows the distribution of employment density for the entire region, where there is a median of 6 jobs per unit area but a high proportion with density less than 2. At the other extreme, there are very few locations where the density exceeds 24 per unit area. This typifies the whole region with a sizeable number of rural locations with little or no employment or potential for jobs (the "thin" market) and semi-urban locations with some employment on a limited scale.

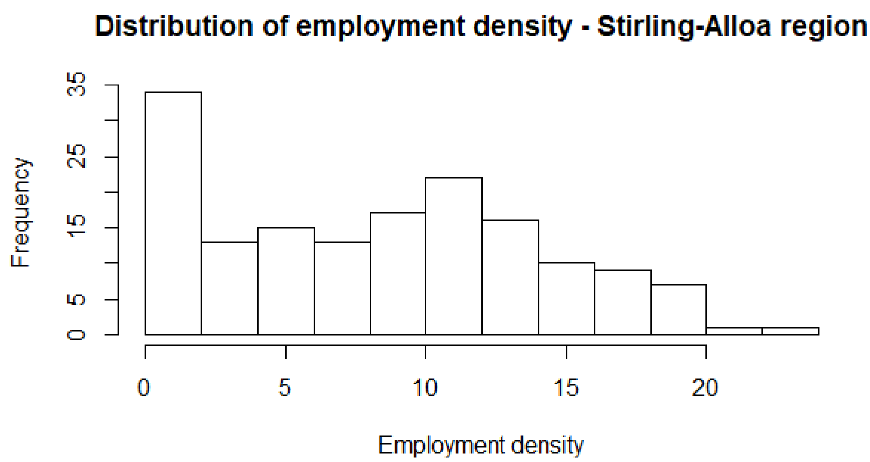


Figure 6-17 Distribution of employment density 2001 Stirling-Alloa

(Source: UK Census, 2001)

6.10.2 Difference-in-Difference model

The difference-in-difference model compares two separate years pre- and post-intervention to predict what would have happened had there been no intervention using the methodology in *4.13.2 Employment modelling – Difference-in-Difference* but applying only two different variables of accessibility which reflect distance from the nearest station and job accessibility.

Distance to nearest rail station

Table 6-25 compares 2001 to 2011 and indicates that a reduction in distance to the station has increased employment density over that period. It also predicts a positive effect of treatment for the standard treatment group (TG4).

Table 6-25 DID job model using distance to station (2001-2011): Stirling-Alloa

	All areas				All areas with treatment variable			
	Coefficient	Std. Error	t value		Coefficient	Std. Error	t value	
(Intercept)	-0.0155	0.0637	-0.244		-0.0145	0.0307	-0.471	
Distance to nearest station	-0.0058	0.0069	-0.837					
Terraced	-0.0015	0.0016	-0.927		-0.0013	0.0014	-0.909	
Detached	0.0005	0.0008	0.564		0.0003	0.0007	0.363	
Semi	0.0013	0.0011	1.208		0.0012	0.0009	1.279	
No car households	-0.0018	0.0010	-1.800	.	-0.0020	0.0009	-2.326	*
1 car household	0.0006	0.0010	0.595		0.0008	0.0009	0.856	
Education Level 1	0.0030	0.0007	4.562	***	0.0029	0.0006	5.066	***
Education Level 4	0.0009	0.0007	1.281		0.0010	0.0006	1.672	
No qualifications	0.0014	0.0006	2.131	*	0.0013	0.0006	2.316	*
Population density	0.0037	0.0031	1.184		0.0034	0.0028	1.209	
Treated					0.0647	0.0204	3.180	**
R^2	0.8502				0.8829			

Distance Ratio

Table 6-26 compares 2001 to 2011 for distance ratio as the accessibility variable, and indicates that being proportionally closer to a station improved employment density over that period.

Table 6-26 DID job model using distance ratio (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t value	
(Intercept)	-0.0972	0.0606	-1.603	
Distance Ratio (DSR)	0.1588	0.0661	2.402	*
Terraced	-0.0014	0.0015	-0.929	
Detached	0.0005	0.0008	0.617	
Semi	0.0013	0.0010	1.304	
No car households	-0.0021	0.0009	-2.242	*
1 car household	0.0006	0.0010	0.615	
Education Level 1	0.0029	0.0006	4.644	***
Education Level 4	0.0011	0.0007	1.596	
No qualifications	0.0014	0.0006	2.426	*
Population density	0.0032	0.0029	1.099	
R^2	0.8698			

Job accessibility

Table 6-27 compares 2001 to 2011 for job accessibility based on cost with skills matching and labour fixed as the accessibility variable, and indicates that increased job accessibility has led to an increase in employment density over that period.

Table 6-27 DID job model using job accessibility (2001-2011): Stirling-Alloa

	Coefficient	Std. Error	t value	
(Intercept)	0.0107	0.0368	0.2900	
Job Accessibility Index	8.3402	8.3695	0.9970	
Terraced	-0.0018	0.0016	-1.1280	
Detached	0.0008	0.0009	0.8610	
Semi	0.0014	0.0011	1.3130	
No car households	-0.0017	0.0010	-1.7170	.
1 car household	0.0009	0.0010	0.9020	
Education Level 1	0.0029	0.0007	4.3870	***
Education Level 4	0.0007	0.0007	1.0150	
No qualifications	0.0012	0.0006	1.9150	.
Population density	0.0035	0.0031	1.1140	
R^2	0.8515			

6.10.3 Geographically Weighted regression (GWR)

As an alternative approach, Geographically Weighted Regression (GWR) models changes in spatial diversity over a period of time spanning the rail intervention. Using the model based on employment density as the dependent variable considered earlier, the same transport accessibility characteristics, distance to the nearest station, and job accessibility index are substituted in turn to generate the OLS global models and GWR model.

The models are run for years 2001 and 2011 to compare coefficients both spatially (across data zones) and temporally. Detailed runs are shown in Appendix 10.6.6. Table 6-28 provides a summary of the parameter estimates fitted by GWR for 2011, including the median, upper quartiles, lower quartiles, range, minimum and maximum.

Going across the 79 sample points, employment density decreases with increasing distance to the nearest station which reflects the global model. For distance to nearest station, the coefficients ranged between -0.2569 and -0.1247. This means that a 1 km increase in distance from the station results in a percentage drop in employment density which can vary between 87.8 and 13.94.

The distribution of values is symmetrical as the median is -0.2569 approximately halfway between the maximum and minimum values. This shows that over the intervention period, the distance from the nearest station has had an increasing effect on employment density i.e. distance from the rail network is now more critical in terms of the job market.

For distance to the nearest rail station, the coefficients range between - 0.3854 and - 0.1247 which means that a 1 km increase in distance from the station results in a drop in average employment density which varies between 87.8% and 13.94%.

As with the property model, the range of GWR local parameter estimates is examined with confidence intervals around the OLS global estimate of the equivalent parameter. The analysis compares GWR model results with the cross-sectional OLS model (which represents the global model) to detect any significant advantage in using GWR by reporting an improvement in fit of a GWR model over a global model using an F test to check significance.

It is clearly evident that all of the interquartile values were less than their corresponding 2 standard errors of the global estimates (Table 6-28) indicating the lack of non-stationarity of relationships over the study area.

Table 6-28 Comparison of GWR coefficients for different accessibility characteristics (2001-2011): Stirling-Alloa

Distance nearest station										
	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	SE	t value	R^2
2001	-0.245	-0.213	-0.169	-0.122	-0.012	0.092	-0.046	0.041	-1.117	0.737
2011	-0.385	-0.319	-0.257	-0.161	-0.125	0.159	-0.176	0.062	-2.843	0.751
Job Accessibility										
2001	47.000	67.820	72.290	78.570	81.110	10.750	61.142	31.146	1.963	0.746
2011	64.690	70.930	77.630	82.130	87.320	11.200	71.266	71.266	28.332	0.745

6.11 Sensitivity analysis

Analysing the impact of the rail intervention through application of a job accessibility index and property and employment models has used the standard treatment/control group configuration treatment/control group combination (TG4). This section examines the effect of applying alternative selections into treatment. The accessibility indices, property and job models are re-run using each treatment group/control group combination to investigate how sensitive the results are to different criteria for selection into treatment.

Job Accessibility

The job accessibility index for rail is calculated to see how group selection impacts on the value of the accessibility measure. Variations in the job accessibility for all combinations of treatment and control groups are shown in Table 6-29. The figures represent the estimated increase in accessibility between 2001 and 2011 expressed as a percentage.

Table 6-29 Job Accessibility variations for different treatment groups

Group Combination ID	Group	Cost basis		Time basis	
		Matching	No Matching	Matching	No Matching
TG1	Control	6.83%	0.67%	6.53%	0.09%
	Treatment	6.75%	0.97%	5.91%	0.07%
TG2	Control	7.72%	0.61%	7.51%	0.20%
	Treatment	6.19%	0.90%	5.40%	-0.01%
TG3	Control	9.11%	0.35%	8.98%	0.19%
	Treatment	6.23%	0.89%	5.51%	0.05%
TG4	Control	6.41%	0.98%	5.87%	0.22%
	Treatment	6.75%	0.97%	5.91%	0.07%
TG5	Control	8.55%	0.57%	8.30%	0.21%
	Treatment	8.80%	1.19%	7.92%	0.24%

Property value

All treatment and control group combinations are run against the difference-in-difference model to compare the effect of alternative "selection into treatment" processes. For the difference-in-difference model, with both distance from the nearest station and job accessibility, there is a consistently positive effect of treatment except for TG3.

Jobs and employment

All treatment and control group combinations are run against the difference-in-difference model to compare the effect of alternative "selection into treatment" processes. For the difference-in-difference model, with distance from the nearest station, there is a consistently negative effect of increased distance from the station and positive effect of job accessibility increase.

6.12 Conclusion and Summary

Of the three case studies, Stirling-Alloa offers the most simplified picture with its lack of external confounding factors, and mix of smaller communities previously remote from the rail network and in a geographically cut-off location. However, having been in place for just 10 years, it offers limited opportunity to observe long term wider economic impacts and represents a minor addition to the rail network involving a relatively small branch line extension.

The findings indicate that rail accessibility shows some impact on property prices and employment density, but the estimate of this impact depends upon the accessibility measure adopted. In this case, because of the short time elapsed since the intervention, the limited size of region studied, the dataset, and the extent of the intervention (one station), the potential benefits are relatively small apart from a greater accessibility to jobs in Glasgow and Stirling. Analysis of the models shows that in a more remote or non-urban region, distance to station should not be considered the sole accessibility measure, and the distance to station ratio and job accessibility provide viable alternative measures, giving consistent results in the model runs. In particular, the inclusion of the job accessibility index represents the opportunities reachable in each location reflecting the feasibility of commute and the attraction of other locations and works better in the employment model.

Counterfactual

The region was divided into treatment and control groups to monitor the effects of infrastructure changes. Alternate treatment group configurations allowed comparison between selection into treatment for urban and rural situations. However, the extension above 6 km distance contours from Alloa station was not feasible because the limited dataset of 79 datazones resulted in groups too small to analyse. This suggests that as a case study region it is perhaps too small for use in this methodology and would benefit from an extension to include other neighbouring datazones. Propensity matching with similar levels of deprivation worked well here for the 2 km distance contours, as did clustering techniques using the dominant cluster groups for matching, but going beyond the 2 km contour required a larger base treatment group to be matched against a much larger selection of potential locations in the control group and matching at 4 km was not satisfactory.

A comparative analysis of treatment and control groups across the intervention period shows proportionately more households having access to a car and increased rail mode share in the treatment group, and average distance to work has reduced in the treatment group but is now greater in the control group. The concentration of population and employment around Alloa (i.e. typically treatment group) and residents in more remote places (i.e. typically control group) must travel further to work and are more likely to work from home. Rail usage has grown steadily since the re-opening with evidence of regular commuting to Stirling and Glasgow.

Job accessibility

There are two comparisons that job accessibility reflects:

1. infrastructure changes only by using the job situation prior to the intervention
2. infrastructure changes and movement of jobs using the prevailing job levels

Rail mode shows a marked increase in accessibility post-intervention for the treatment group. All results suggest that, without job skills matching, accessibility may be overestimated due to the high attraction value of all job opportunities. A job accessibility measure based on travel cost and allowing for skills matching offered the best option to take forward for application in the models, assuming travel cost rather than journey time will be the prime deterrent factor in more remote regions such as Stirling-Alloa, added to the fact that cost will include an element of time in its derivation. The job accessibility measure is fairly consistent for different treatment groups, but with a much wider spread of values for the respective control groups, reflecting the reducing size of the control group with the expansion of the treatment group contour to cover greater distances, and suggests a limitation on this method where the dataset is comparatively small. Job opportunities are normally excluded beyond a defined threshold, but in this case exceptions are made for Stirling and Glasgow, outside the case study region but providing a powerful pull in terms of job opportunities.

Impact on residential property values

In the Stirling-Alloa region there are a small number of very costly properties with a central group of typically-priced houses together with a long tail of relatively expensive ones, and the percentage of terraced properties has decreased across the period compared to an increase for detached properties. A comparative analysis showed that between 2008 and 2010 (spanning the immediate period either side of the intervention), there was a marginally higher increase for the treatment group.

The difference-in-difference model does not suggest that a reduction in distance to the station has had any significant impact on house prices, and in fact has had a negative effect in the treatment group which represents those data zones closest to Alloa station. This is also reflected in results for distance ratio and job accessibility. This effect is consistent for different alternative treatment groups.

However, the fixed effects model indicates an impact of improved rail accessibility on property prices with distance to the nearest station having a negative effect as distance increases, and similarly distance to station ratio results indicate that an increase in distance ratio i.e. being proportionally nearer the station after the intervention, has a positive effect on property price as the relative change in distance increases. Also an increase in the job accessibility index has a positive effect on property price. For property impacts, the fixed effects model predicts a consistently positive effect of decrease in distance to station and increase in distance ratio on property price especially for treatment group TG4, but this appears more marked in the control groups.

Geographically Weighted Regression (GWR) indicates that the relationship between prices, distance to station and job accessibility is negative in some locations and in others positive. There is a cluster of higher negative coefficient values in the centre of the region, and lower negative coefficient values on the periphery. Comparing model coefficients over the two separate years and applying different accessibility characteristics to the model did not produce the same effects in 2011 as in 2001. However, the model results show a significant advantage in using GWR over the cross-sectional OLS model in highlighting that distance is more important to property price in the central areas around Alloa than in the more remote fringe regions. Analysis also indicates the lack of the presence of non-stationarity of relationships over the study area.

Impact on jobs and employment

The work on accessibility demonstrates that it varies spatially and has a varied effect on employment across the region. Initial comparison shows that net employment decreased less sharply in the treatment group following the intervention. The distribution of employment density typifies the whole region with a mixture of locations with little or no employment or potential for jobs and semi-urban cohort with a limited level of employment.

The difference-in-difference model indicates that a reduction in distance to the station has effected an improvement in employment density over the intervention period. Similarly, an increase in job accessibility and distance ratio have impacted positively on employment density over the intervention period. There is a consistently positive effect regardless of treatment group configuration, and whichever accessibility characteristic is used there is a pattern that suggests a 2 km threshold with either clustering or propensity matching offers the most suitable configuration.

For GWR, the relationship between employment density and distance to the nearest station was in most cases negative - reflecting the OLS model and the presence of non-stationarity of relationships is clearly evident over the study area. This varies across the region with local clustering in the centre and diminishing effects in the peripheral areas. Comparing 2001 with 2011 produces a different set of coefficients of the same order allowing spatial and temporal comparison.

7 Chapter Seven: The Borders Rail case study

The previous chapters took the Robin Hood Line and Stirling Alloa as case studies, both of which have been established for some time. In contrast, this chapter considers the impact of reopening the Borders Rail line from Edinburgh to Tweedbank in the Scottish Borders which represents a relatively recent intervention (2015), and so as a case study offers reduced potential for assessing impacts due to the limited historical data available.

Despite these limitations, as Borders Rail provided the catalyst and motivation for this research, relevant elements are addressed here. Again this case study is contextualised through an overview of the geography of the region and its socio-demographic and economic profile and property market. Through division of the region into treatment and control groups there is econometric modelling of the impact on property price and employment density using three alternative accessibility indicators. The sensitivity of selection into treatment criteria is analysed using different group configurations with a view to differentiating between urban and rural applications.

7.1 Borders Rail and its region

To put this in context, it is necessary to describe the specific situation of the Borders Rail line as a sizeable addition to the rail network in Scotland. Through re-opening the Borders Rail Line - a substantial section between Tweedbank and Edinburgh - the Scottish Borders and to a lesser extent Midlothian - previously isolated regions - are now re-linked to the rail network and the Edinburgh-Glasgow axis for the first time in nearly 50 years. The case study region is best considered as two separate areas - the Scottish Borders to the south of the region and Midlothian to the north of the region and much closer to Edinburgh

Although as a case study Borders Rail provides an ideal context for addressing the aims and objectives of this research, the short period of time (3 years) since its introduction is insufficient to come to a definitive conclusion on wider economic impacts. However, it will be seen that commuting patterns, job accessibility and house price movements show early indications of an economic effect. Like the other case studies, it represents a previously isolated region which now has much improved links to a large conurbation (Edinburgh), but differs from the other case study regions in the following respects:

- Its regional geography makes it effectively a cul-de-sac beyond Hawick.
- There are many geographically remote communities represented by data zones which often cover a larger area than those considered previously, and consequently have a lower average population density.
- The intervention has created seven new rail stations which link to stations nearer Edinburgh already in place.
- Apart from the Midlothian area there are no pre-existing rail stations, so until recently rail has not provided a realistic choice of travel mode.
- Although Edinburgh installed a tramway system in 2014 just prior to the rail link, this does not overlap with the rail locations nearer the city.
- There has been little development of the road network, hence fewer potential confounding factors.
- The region provides a further contrast in offering an attractive tourist location.
- There may be signs of a “two way road effect” through an increase in commuting out of the region, combined with some retail movement and increased tourism.

Scotland is a relatively sparsely-populated country with most population concentrated in the industrial Central Lowlands. The Scottish Borders is one of its more remote regions, lying in the eastern part of the Southern Uplands. The region is hilly and largely rural, with the River Tweed flowing through it. Borders Rail follows most of the alignment of the northern part of the Waverley Route, a former double-track line in southern Scotland and northern England that ran between Edinburgh and Carlisle, and provided direct rail services between Edinburgh, the textile towns of the Borders, Carlisle and North Yorkshire and onward to London St Pancras. As part of the Beeching cuts that line was controversially closed in 1969, leaving the Borders region without any access to the National Rail network, Hawick being the most remote UK town from the rail network.

The new line opened in September 2015 and involves over 30 miles of new railway infrastructure, and the development of seven new rail stations, four in Midlothian and three in the Borders (Figure 7-2). There is a journey time of less than one hour between Edinburgh Waverley and Tweedbank, with 12 intermediate stops servicing at shorter distances.

7.2 Definition of case study region

For the purpose of this analysis, the Borders Rail region is illustrated in Figure 7-1 where the region studied is shown in white with a light grey boundary. This incorporates all datazones adjacent to the new section of the line and extends westwards to Peebles, eastward to Coldstream and 20 miles south of Hawick. This allowed a mixture of areas close to the rail intervention and those further away to allow comparison of impacts later.

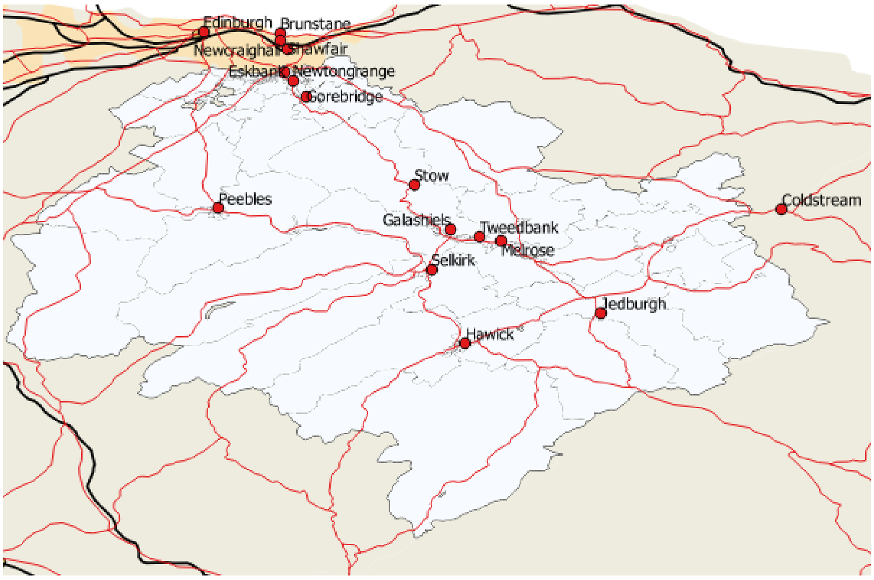


Figure 7-1 Borders Rail Case Study regional boundary

7.3 Population profile

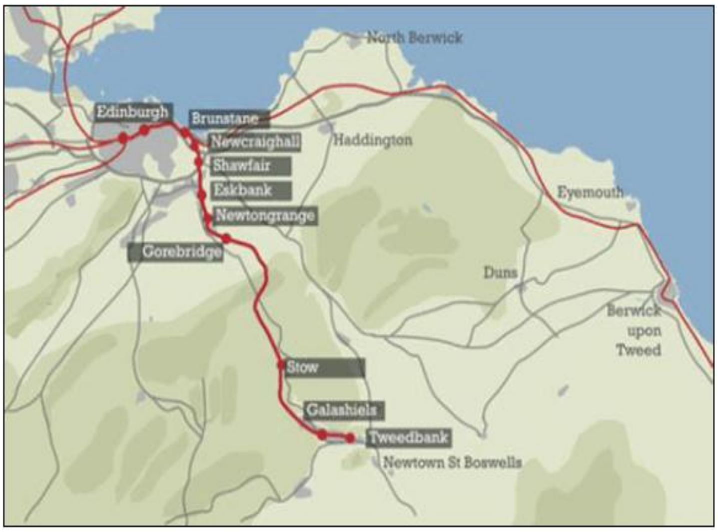


Figure 7-2 Borders Railway route
(Source: Transport Scotland, 2014)

The region is similar to Stirling-Alloa in being more sparsely populated with a few larger townships scattered around. At the 2011 UK Census, the population of the region was 114,000, up 6.78% from the previous census. 30% of the population live in small hamlets or settlements of fewer than 500 people. The largest town is Hawick with a population of 14,029 (UK Census, 2011), followed by Galashiels with 12,604. The only other towns with a population of over 5,000 are Peebles, Kelso and Selkirk. The Scottish Borders has a smaller proportion of its population of working age (62%) compared to the rest of Scotland (66%). However, although the proportion of children under 16 is around the Scottish average at 17%, there is a higher aging population with the proportion of those aged 65 and over well above average at 20.9% compared with 16.8% in Scotland.

The remoteness of the region is evident in the access deprivation which is higher in the Scottish Borders compared to Scotland as a whole. Of the 130 data zones in Scottish Borders 29% are within the 15% most access deprived areas of Scotland, reflecting the rural geography of the Scottish Borders compared to Scotland (Scottish Government, 2016). This is also reflected in other deprivation statistics which show employment deprivation is lower than average, but increasing rapidly. The most deprived data zone is in Hawick - on the fringes of the region.

Overall, unemployment is lower compared to the rest of Scotland, with a higher proportion of employment in agriculture and manufacturing, but a lower proportion in banking, finance and insurance which continues to affect the GVA for the Scottish Borders. Between 2002 and 2011 Gross Value Added (GVA) for the Scottish Borders increased by 33% from £1,119M to £1,513M, which was less than the 44% increase for Scotland.

Like the rest of the UK, the proportion claiming Jobseeker's Allowance (JSA) has increased over the past 10 years, particularly for the 18-24 age group. Due to the economic recession, the rate of unemployment doubled between 2007 and 2010 from 3.1% to 6.2%. Since then, the rate of unemployment has decreased slightly to 5.9 per cent in 2012. However, there is a well-educated and skilled workforce with a lower proportion of people of working age having low or no qualifications compared to Scotland as a whole. The Scottish Government's labour market publication (based on the ONS Annual Population Survey) indicates that qualification levels of the working age population in the Scottish Borders are consistent with those for Scotland, and attainment in secondary schools is equivalent or better than Scotland as a whole.

7.4 The property market

Property prices can be a reflection of the state of the local economy and, by 2017, the market in the Borders has slowly recovered to exceed pre-2007 levels (Figure 7-3).

This follows a similar pattern to the national picture where a drop in 2007-2010 is followed by a steady increase over a period covering the rail intervention in 2015.

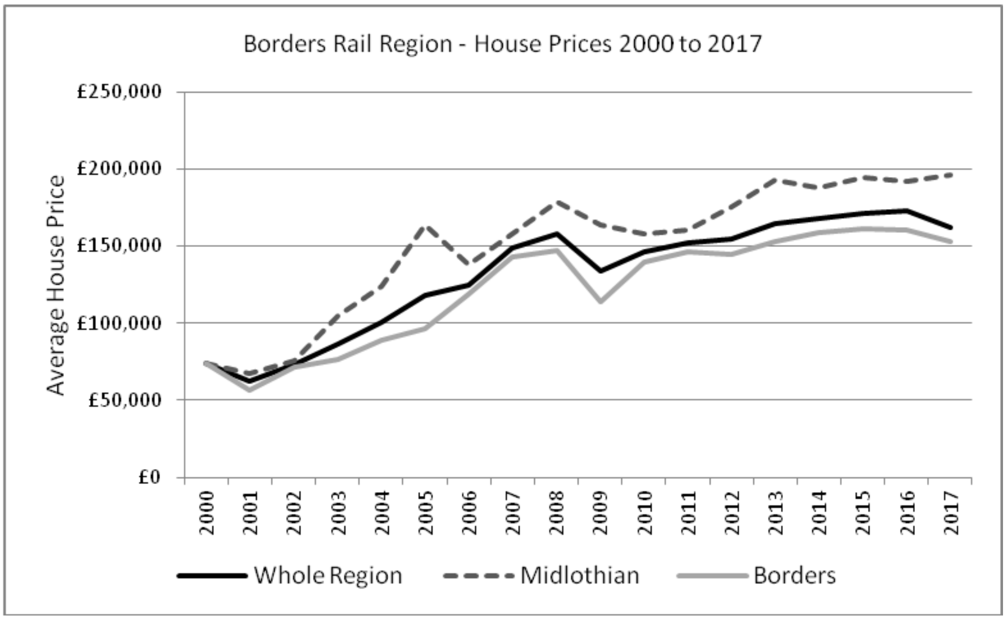


Figure 7-3 House prices 2000-2017: Borders Rail (*Source: Rightmove, 2018)*

Although, there was a marked decrease in the number of new housing builds for 2011-2012 within the Scottish Borders (137%) compared to a Scottish decrease of just 2%, this was offset by a 23% increase in 2012-2013. It is evident that average house prices in Midlothian were consistently higher than those in the Scottish Borders, and from 2014 to 2017 a gradual increase in Midlothian as against a levelling out and slight decline in the Scottish Borders. This may point to a greater benefit for Midlothian in being nearer to Edinburgh and more commutable in terms of time and cost.

7.5 Transport provision

Analysis of the current transport situation provides further information concerning the variable levels of transport disconnection between the Scottish Borders and Midlothian. Until September 2015, the Scottish Borders had no working railway stations although there were some close to Midlothian. The area is served by buses which connect the main population centres with express bus services linking the main towns with rail stations at Edinburgh and Carlisle.

The Scottish Borders has a lower proportion of people journeying to work by public or active transport according to the Scottish Household Survey (Scottish Government, 2015), and consistently has a negative commuting to work inflow of about 10%, so more people go outside of the region than come into it for work. An average of 1.25

cars in each household compared to 1.04 for Scotland (UK Census, 2011) reflects the predominantly rural nature of the Scottish Borders. A lower proportion of households have no car, and the region has approximately 1,900 miles of adopted routes available.

Since the re-opening, rail passenger services run half-hourly on weekdays until 20:00, then hourly until 23:54 on Sundays with the majority of services between Tweedbank and Edinburgh having a journey time of less than one hour. The timetable also allows charter train promoters to run special excursions. As a consequence of its low population density, long journeys are necessary to reach educational, medical, shopping and leisure facilities, with the Scottish Government classifying a third of the region as being "remote rural" in nature.

7.6 Treatment and control groups

As with the previous case studies, the region is again divided into treatment and control groups to appraise the impact of the rail intervention and establish causality through creation of a meaningful counterfactual. The groups are based on those locations subject to change in rail access across the intervention period. Unlike Stirling-Alloa, the region contains considerably more data zones and consequently groups are larger and statistically more manageable. Division into treatment and control group combinations follows the methodology in *4.9 Selection of treatment and control groups* adopting the same criteria as for the other case studies,

7.6.1 Change in distance

For the "base" groupings, allocation of groups involves selection into treatment data zones experiencing a change in distance to the rail network for different contour boundaries i.e. 2 km contours radiating from the nearest rail stations up to a maximum of 10 km, affording five variations in base group specification. Figure 7-4 shows the different base treatment groups using alternative methods of selection, and unlike Stirling-Alloa, the 10 km contour presents a feasible alternative without greatly expanding the dataset.

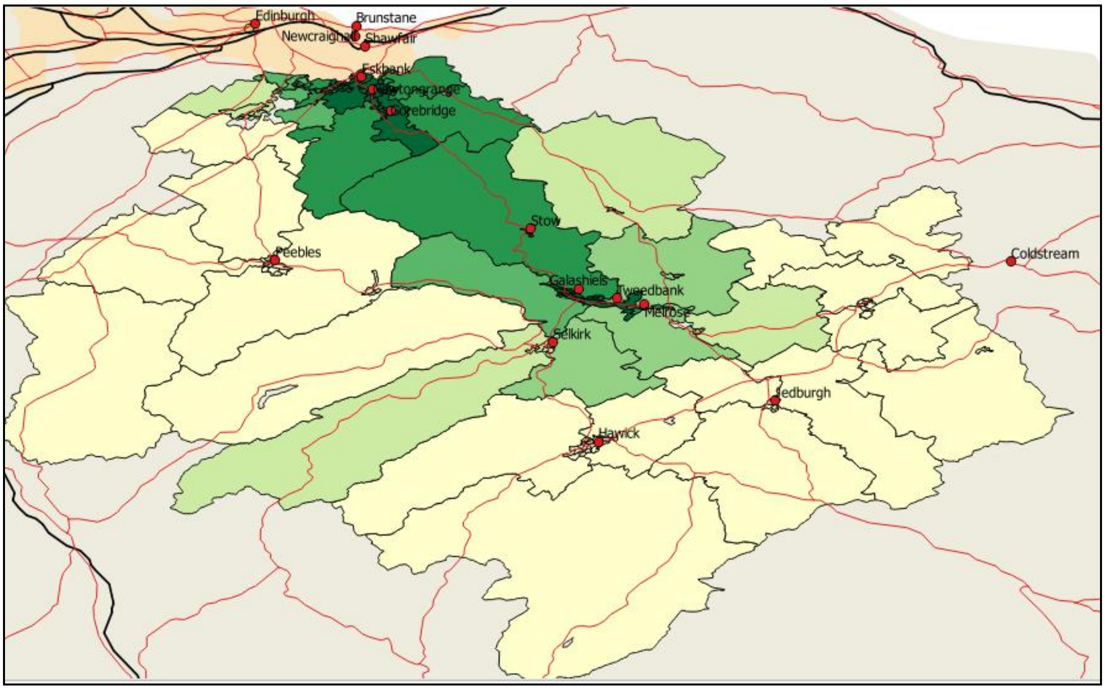


Figure 7-4 Variation in treatment and control Group allocation: Borders Rail

0-2 km | 2-4 km | 4-6 km | 6-8 km | 8-10 km | > 10 km

7.6.2 Propensity score matching

It is important to recognise the limitations of using distance as the sole basis and the need to compare similar zones in each group. This selection of the basic groups is expanded using propensity scoring matching techniques with nearest neighbour 1-to-1 matching. The comparators used for matching are zone area size and deprivation levels taken from the Scottish Index of Multiple Deprivation (Scottish Government, 2016), health, education, crime and housing. In this case study, propensity matching is applied in conjunction with base groups up to 4 km for changes in distance to the nearest station. Extension to higher distances would require matching a larger base standard treatment group against a much larger selection of potential locations to create a control group. This would involve extracting and collating a much more extensive dataset. Matching treatment and control areas works well when adopting a 2 km distance threshold as the base. This is highlighted in the mean values for each comparator which indicate a marked high percentage balance improvement section (Table 7-1).

Table 7-1 Summary of propensity matching output at 2 km threshold: Borders Rail

Summary of balance for all data							
	Means Treated	Means Control	SD Control	Mean Diff	eQQ Med	eQQ Mean	eQQ Max
distance	0.3671	0.2933	0.1056	0.0739	0.0526	0.0743	0.2947
health	0.1392	-0.0842	0.5529	0.2233	0.28	0.2815	0.78
housing	15.0211	13.8503	6.0112	1.1707	1.47	2.1005	8.9
education	0.0163	-0.1304	0.7009	0.1467	0.26	0.2532	0.5
crime	405.3509	269.8699	210.1728	135.481	145	135.4737	474

Summary of balance for matched data							
	Means Treated Means	Control	SD Control	Mean Diff	eQQ Med	eQQ Mean	eQQ Max
distance	0.3671	0.349	0.1234	0.0181	0.0049	0.019	0.2161
health	0.1392	0.0793	0.5322	0.0599	0.16	0.2045	0.85
housing	15.0211	14.5307	5.3926	0.4904	2.08	2.1321	3.92
education	0.0163	-0.0479	0.7763	0.0642	0.16	0.1642	0.5
crime	405.3509	366.5263	248.5586	38.8246	29	45.2105	474

Percent balance improvement				
	Mean Diff.	eQQ Med	eQQ Mean	eQQ Max
distance	75.5009	90.6421	74.4303	26.6782
health	73.177	42.8571	27.3618	-8.9744
housing	58.1157	-41.4966	-1.5034	55.9551
education	56.2263	38.4615	35.1314	0
crime	71.3432	80	66.6278	0

Sample sizes		
	Control	Treated
All	123	57
Matched	57	57
Unmatched	66	0
Discarded	0	0

For the 2 km threshold, in the "Summary of balance for all data" section (Table 7-1), before matching, the mean percentage of health was 0.22 more, housing 1.27 more, education 0.15 more, and crime 135 more in the treatment area than in the control area. After matching, ("Summary of balance for matched data"), the mean differences in percentage between treated and control areas reduces to 0.06 for health, 0.49 for housing, 0.02 for education, and 38.8 for crime. The treated and control areas are now much more similar in terms of health and crime deprivation markers in particular. The rightmost columns in Table 7-1 show the median, mean, and maximum quartile (QQ) between the treated and control data; larger QQ values indicate better matching. The jitter plot for Borders Rail region (Figure 7-5), shows an absence of cases in the uppermost stratification indicating that there are no unmatched treatment units.

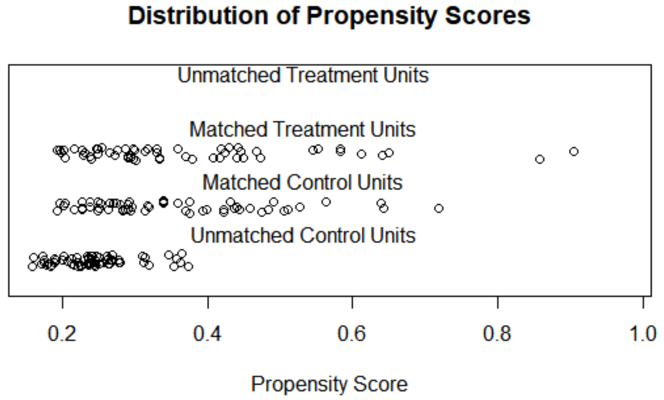


Figure 7-5 Distribution of propensity scores: Borders Rail

The middle stratifications show the close match between the treatment units and the matched control units, and the final stratification shows the unmatched control units. Both the numerical and visual data indicate that the matching was successful. shows the histograms before and after matching where those before matching (on the left) are seen to differ to a great degree. However, after matching those on the right are much closer in shape.

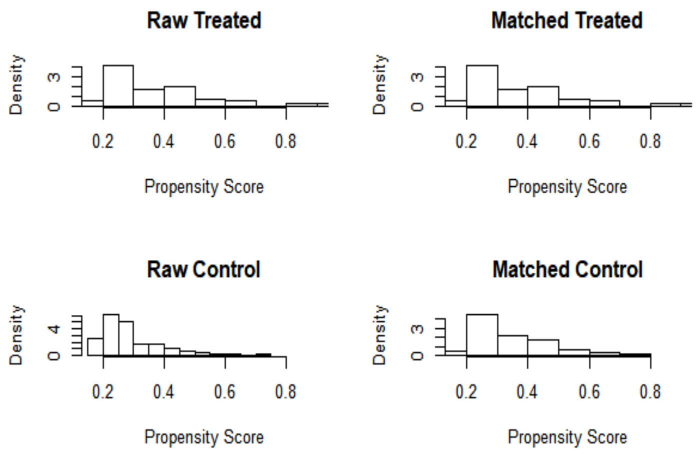


Figure 7-6 Histograms showing shapes before and after matching: Borders Rail

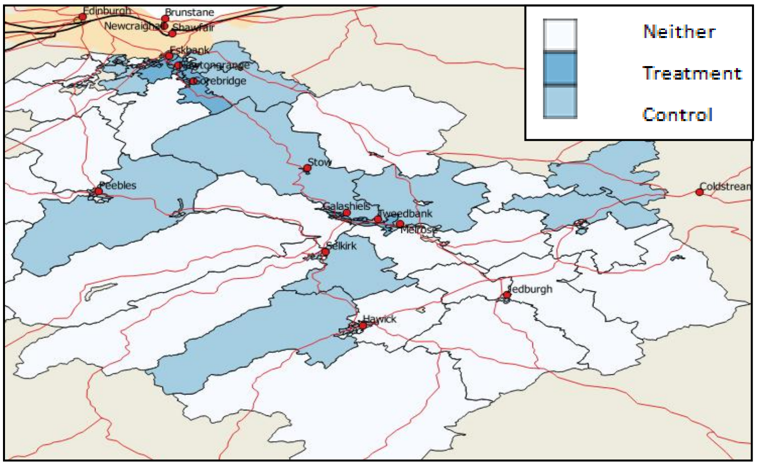


Figure 7-7 Treatment and control groups with propensity matching: Borders Rail

Figure 7-7 illustrates the selection of treatment and control groups after the application of propensity testing based on a 2 km threshold. Although matching is less effective at the 4 km threshold, there is still some improvement shown, especially in education (Table 7-2), however, there are relatively few unmatched data zones because the number in the treatment group has increased to 81 requiring 81 matches out of an available pool of only 99 in the base control group.

Table 7-2 Summary of propensity matching output 4 km threshold: Borders Rail

Percent Balance Improvement		Sample sizes		
distance	29.6052		Control	Treated
health	19.8929	All	99	81
housing	34.6127	Matched	81	81
education	40.9722	Unmatched	18	0
crime	25.9185	Discarded	0	0

7.6.3 Application of clustering

As an alternative to propensity matching, clustering techniques are applied to produce a set of alternative treatment and control groups. A dendrogram (Appendix 10.7.3) indicates 4 distinct clusters as shown by the ovals. The elbow chart shown in Appendix 10.7.2 concurs that the optimum number of clusters is 4. This represents fewer clusters than the Robin Hood Line case study, but more than Stirling-Alloa, which reflects the size of the regional sample of data zones. The four cluster groups are considered in allocation of treatment and control groups. The membership of the cluster groups is as shown in Table 7-3 and mapped against the base groups at 2 km threshold in Figure 7-8 and Figure 7-9. This shows cluster group 1 to be the dominant cluster, and following initial division into base groups using 2 km distance as a basis for allocation, clustering is applied by mapping matching clusters in the two groups.

Table 7-3 Cluster group memberships: Borders Rail

Cluster Group	Membership
1	79
2	52
3	16
4	33

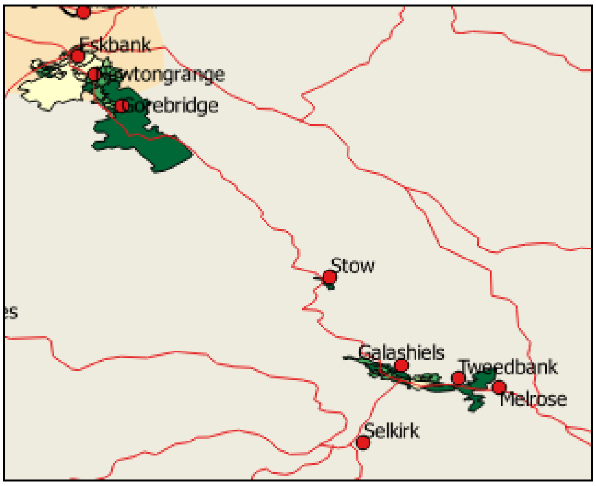


Figure 7-8 Clusters within treatment group: Borders Rail

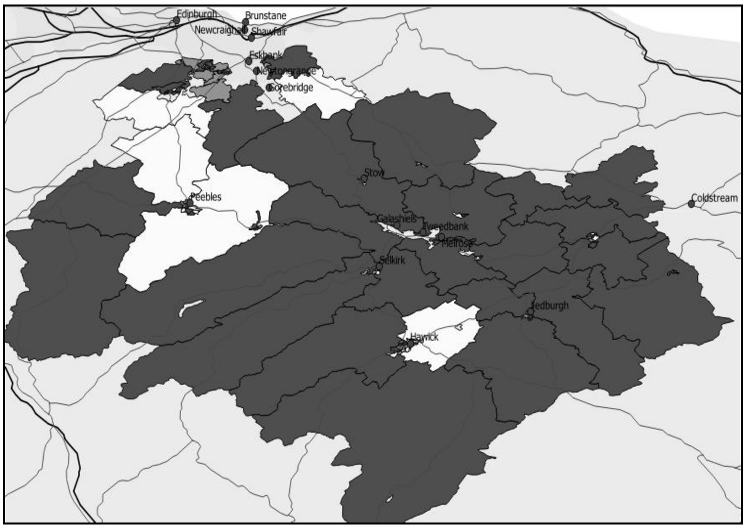


Figure 7-9 Clusters within control group

7.6.4 Treatment and control group combinations

The different methods of allocation to treatment are summarised in Table 7-4 to give a total of eight possible treatment and control group pairings:

Table 7-4 Treatment/Control Group combinations: Borders Rail

Treatment Group ID	Comments	Distance threshold (km)	Status
TG1	Base	2	
TG2		4	
TG3		6	
		8	Not used
		10	Not used
TG4	Base + Propensity	2	
	Base + Propensity	4	Not used
TG5	Base + Clustering	2	

For this case study, all combinations are feasible because although the number of data zones in the Borders Rail region is small (180), there are a sufficient number of data zones in the control group without expansion of the dataset. However, so that comparison could be made between the case studies, only the same five combinations that were relevant for the other case studies were retained. Treatment /Control group combination TG1 to TG3 represent allocation based on various threshold distances from the station where there has been a change in access. TG4 expands TG1 with the addition of propensity matching, and TG5 is based on TG1 with the addition of clustering matched on the dominant cluster.

One specific treatment/control group combination is adopted throughout the analysis for comparison purposes. This is group TG4 in Table 7-4, which represents locations up to a 2 km distance from a rail station experiencing a post-intervention change and incorporating propensity matching.

7.7 Accessibility characteristics

Following division into treatment and control groups, consideration is now given to the different measures for defining accessibility in the Borders Rail region.

7.7.1 Initial descriptive analysis

Study of accessibility issues and the calculation of suitable measures require prior knowledge of the background relevant to jobs and services in the Borders Rail region, in particular:

- Car ownership
- Travel to work patterns by mode and distance travelled
- Rail passenger usage to study uptake of rail travel in the region
- Accessibility to essential services

The standard treatment and control group configuration (TG4) is used to contextualise the background and how it has affected "treated" areas, but owing to the short period since the Borders Rail reopening, data for comparison purposes before and after the intervention is very limited, and so, unlike the other case studies, merely reflects the situation prior to 2015. Prior to the intervention period in 2011, the treatment group (TG4) had a similar percentage with no access to a car (26.6%) as the control group (23.5%), but much more than those in neither group (Table 7-5). This differential is also reflected in a lower percentage of households with access to at least one car suggesting that application of propensity matching has produced two similar groupings for comparative purposes.

Table 7-5 Car ownership prior to the intervention period (2011): Borders Rail
(Source*: UK Census, 2011)*

Group	No cars or vans	1 car or van	2 cars or vans	3 cars or vans	4 or more cars or vans
Treatment	26.60%	43.65%	23.78%	4.61%	1.36%
Control	23.51%	44.77%	24.93%	5.04%	1.76%
Neither	17.65%	45.23%	28.41%	6.46%	2.25%

For travel to work prior to the intervention period (Table 7-6), the treatment and control groups have a similar profile, but more travel by bus in the treatment group, and more work from home in the control group. This reflects the better accessibility to buses in the treatment group locations e.g. Galashiels. The majority using a car either as driver or passenger for both groups reflects the reliance on cars across the region.

Table 7-6 Method of travel to work prior to the intervention period (2011): Borders Rail (*Source: UK Census, 2011)*

Group	Work from home	Tram	Train	Bus	Taxi	Driving a car or van	Passenger in a car or van	Motorcycle	Bicycle	On foot
Neither	14.75%	0.02%	0.36%	6.71%	0.21%	60.65%	5.06%	0.33%	0.97%	10.26%
Treatment	9.08%	0.01%	0.31%	13.34%	0.35%	59.61%	6.20%	0.49%	1.17%	8.88%
Control	13.17%	0.03%	0.35%	8.28%	0.26%	57.63%	5.37%	0.40%	1.05%	12.87%

Table 7-7 shows distance travelled to work, with more either working from home or travelling less than 2 km in the control group (31.82%) than the treatment group (22.16%), but more commuting to distances of under 10 km in the treatment group (45.85%) than the control group (40.95%). This reflects the geographical nature of the region where there is a concentration of population and employment around Galashiels (i.e. typically treatment group), and residents in more remote places (i.e. typically control group) must travel further to work or work from home.

Table 7-7 Distance travelled to work prior to intervention period (2011): Borders Rail *(Source: UK Census, 2011)*

Group	Work from home	less than 2km	2km to less than 5km	5km to less than 10km	10km to less than 20km	20km to less than 30km	30km to less than 40km	40km to less than 60km	60km and over	Other
Neither	14.75%	15.66%	9.44%	12.67%	19.24%	6.74%	3.96%	3.24%	2.33%	11.97%
Treatment	9.08%	13.08%	16.21%	16.56%	24.27%	3.90%	2.25%	2.45%	1.39%	10.80%
Control	13.17%	18.65%	8.72%	13.58%	16.84%	6.96%	5.55%	2.55%	2.20%	11.79%

Despite limited data over the short period since the rail reopening, the rail usage figures indicate the uptake of rail use over that time showing that usage of Borders Rail stations has grown more quickly compared to similar pre-existing stations. Appendix 10.7.1 shows station usage for both Borders Rail stations and neighbouring pre-existing stations by full fare, reduced fare and season ticket categories. This is shown graphically in Figure 7-10 highlighting uptake since the rail re-opening between 2015 and 2017, and comparing Borders Rail with other stations.

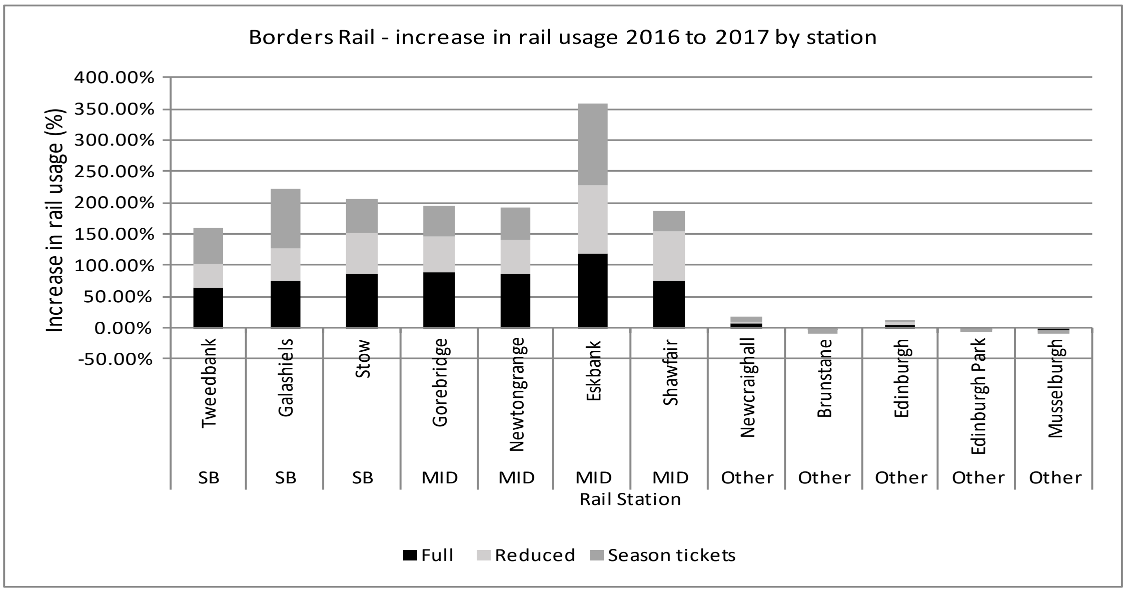


Figure 7-10 Comparison of rail usage between 2015 and 2017: Borders Rail
(Source: ORR Portal, 2018)

In Figure 7-10 the stations are shown in route order from Tweedbank to Edinburgh. "SB" are new stations in the Scottish Borders section of the line, "MID" new stations in the Midlothian section of the line and "Other" are pre-existing stations. This indicates a noticeably positive uptake in rail usage since the re-opening, with an increasing number of season tickets, especially from Galashiels, which represents the most populated centre being served. This contrasts with low figures for Newcraighall, Brunstane and Musselburgh which were previously the nearest available stations for Midlothian.

7.7.2 Distance to nearest station

For most locations in the Borders Rail region distance to the nearest station will decrease following the intervention with seven new stations offering closer access to the rail network.

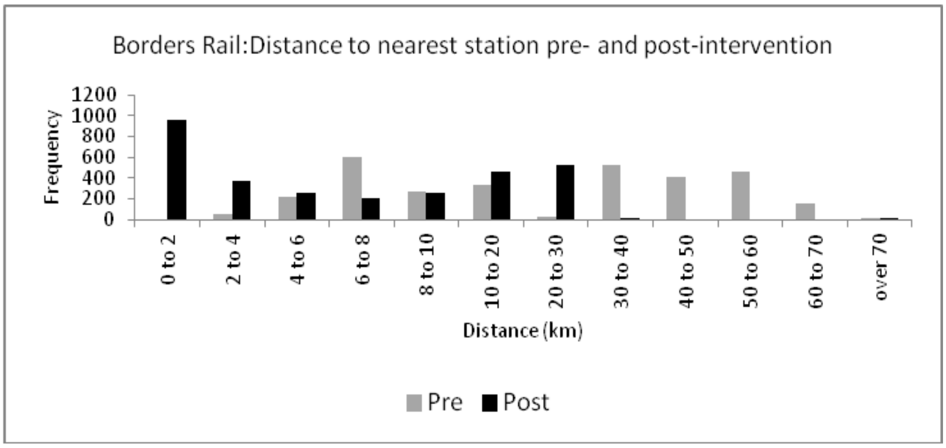


Figure 7-11 Distance to nearest station comparison (2011-2017) Borders Rail

This is more apparent for the Borders Rail case study as shown in Figure 7-11 which shows a predominant shift, with most data zones now much nearer a rail station in 2017. This is not surprising as most locations in the Scottish Borders were some distance from a rail station before the intervention. So as an indicator of accessibility to rail mode it is a potential comparator for use in econometric modelling, if limited in that it takes no account of journey cost or time. If this is broken down further into Scottish Borders and Midlothian (Figure 7-12), it is noticeable that Midlothian datazones were much nearer existing stations (such as Newcraighall and Brunstane) before the intervention than those in the Scottish Borders.

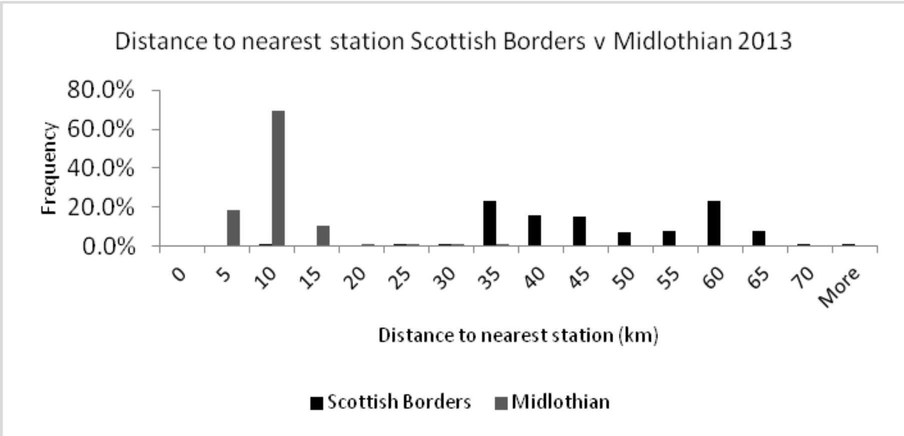


Figure 7-12 Distance to station comparison (2011-2017) Scottish Borders and Midlothian: Borders Rail

7.7.3 Distance ratio

This measure which was applied in both the Robin Hood and Stirling-Alloa case studies indicates the relative improvement in accessibility where there has been a change over time in distance to the nearest station due to new rail infrastructure. As an indicator of accessibility to rail mode, it offers a potential comparator for use in econometric modelling as it represents the percentage improvement in distance access to a station. Figure 7-11 maps distance ratio frequency by post code and shows a major shift which is skewed towards larger values of the ratio, so that in 2017 most data zones are now much nearer a rail station.

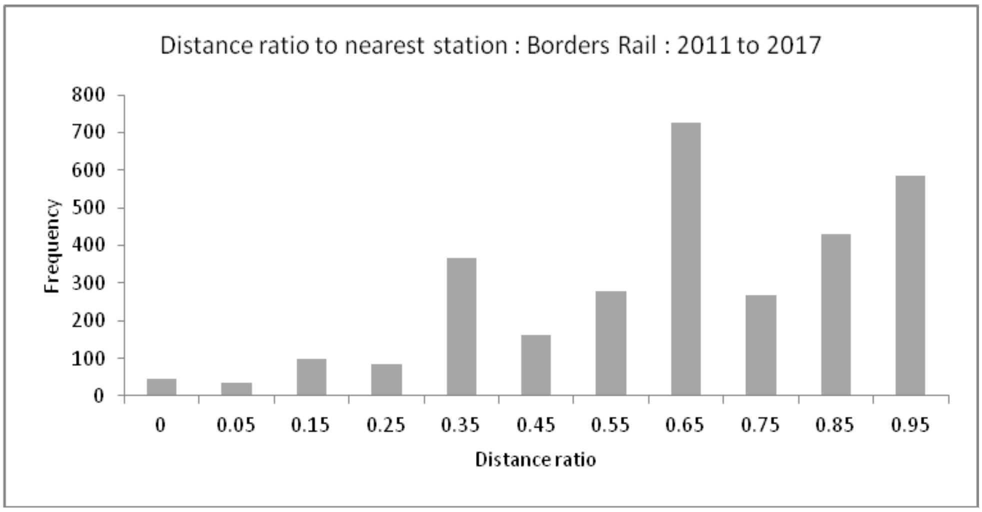


Figure 7-13 Distance ratio to nearest station (2011-2017): Borders Rail

Breaking this down further into the Scottish Borders and Midlothian separately (Figure 7-14), it is noticeable that for Scottish Borders the main movement is in the 0.55-0.65 range with a further peak at 0.95, whereas for Midlothian this is in the 0.75-0.85 range.

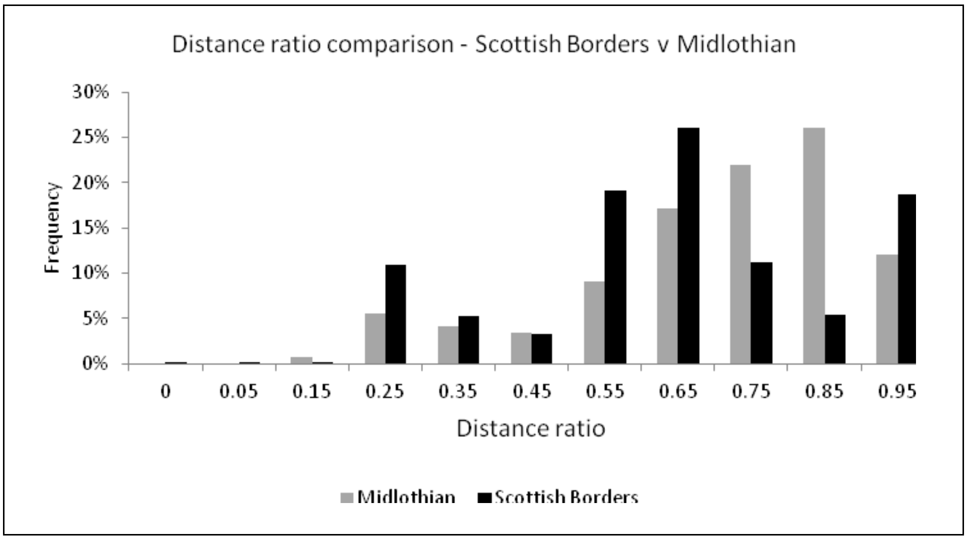


Figure 7-14 Distance ratio comparison Scottish Borders and Midlothian (2011-2017): Borders Rail

This anomaly can be explained, as the distance ratio is a relative measure based on the percentage difference in distance to the nearest station. Hence, the ratio can be larger for a datazone in Midlothian, even though the original distance to the nearest station and actual change in distance is much smaller than in the Scottish Borders (Figure 7-12).

7.8 Job Accessibility index

The third alternative measure of accessibility is a job accessibility index which is modified to be applicable to the Borders Rail Line case study region based on considerations in *4.10 Accessibility to jobs.* It is varied by alternatively applying generalised travel time and travel cost as proximity measures in a generic job accessibility index. The time and cost values and decay parameter for Borders Rail are based on transport data estimates, with allowance for commuting practicability and accounting for constraints which impose a boundary in access to jobs. As for the other case studies, the alternative elements which define the measure for the Borders Rail region are calibrated prior to calculation of accessibility.

7.8.1 Setting the standard parameters

As each case study region is different, a price comparison for individual transport modes requires a price per km, speed and headway for train, bus and car mode for the Borders Rail region, derived using methods prescribed in *4.10.12 Calibrating the index and costs for each case study.* As a first element, the unit cost and speed values are estimated so that the cost of a journey for each transport mode can be calculated (Table 7-8).

Table 7-8 Cost and speed values used in index calculation: Borders Rail

Transport Mode	Transport Speed	Service Frequency	Headway	Cost of Travel
Bus	35	2	30	£0.23
Car	50.03	0	0	£0.48
Rail	67	2	30	£0.18
Walk	4.8	0	0	£0.00

Distance	Average
Work mainly at or from home	88.14%
Less than 2km or work from home	71.57%
2km to less than 5km	60.10%
5km to less than 10km	45.62%
10km to less than 20km	25.38%
20km to less than 30km	19.73%
30km to less than 40km	16.16%
40km to less than 60km	13.55%
60km and over	11.63%

Table 7-9 Decay effect of travel distance: Borders Rail

As for the standard values, a decay parameter (β) is determined producing the best fit for travel behaviour in the Borders Rail region (*4.10.12 Calibrating the index and costs for each case study)*.

Through analysis of UK census distance to work statistics by data zone for 2011, Table 7-9 summarises cumulative average distances travelled over the whole region, indicating that on average 45.62% of the working population travel at least 10 km to work, but this tails off gradually up to 40km (16.16%). From these calculations a chart is constructed mapping the impedance value against distance travelled (Figure 7-15), and by fitting an exponential decay function graphically, an overall value of β is estimated. This has an R^2 value of 0.5184, and value of β of -0.028.

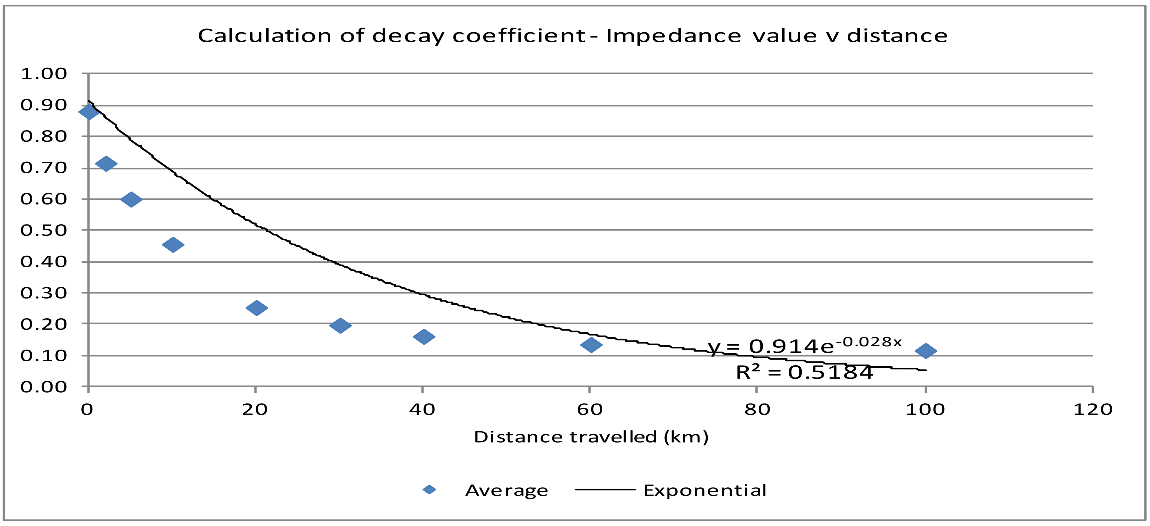


Figure 7-15 Calculation of decay coefficient: Borders Rail

This initial value of β is based on distance only, but adjusted when calculating job accessibility, to allow for a time or cost basis of proximity. It will also be modified based on the mode of transport used. This is reflected in the average speed and cost of that mode in the region as indicated in *4.10.12 Calibrating the index and costs for each case study*.

A further consideration is the imposition of a threshold beyond which job opportunities are not included. However, for Borders Rail, exceptions are made for Edinburgh, which although outside the case study region, provides a powerful pull in terms of job opportunities, made more relevant with the reintroduction of rail travel from the Borders. Employment data for Edinburgh has thus been incorporated into the accessibility calculations, and the threshold distance is defined using commuting data where the maximum travel distance observed in the region for all commuters (100 km) may be greater following the rail intervention.

7.8.2 Job accessibility comparison

Having determined the specific cost and decay parameters for Borders Rail, the next stage was to compare the job accessibility index values alternatively based on travel time and travel cost before and after the intervention, broken down further by applying either job skills matching or no matching after allowing for commuting thresholds. Unlike the other two case studies, because Borders Rail has only been open for three years, only one comparison is made i.e. using the job situation prior to the intervention as the basis for job opportunity and skills matching. Comparing the average job index by travel mode for two separate years spanning the intervention 2011 and 2017, accessibility by rail in the treatment group improves over the control group (Table 7-10). The control group behaves similarly to the treatment group, but on a much smaller scale than the treatment group.

Table 7-10 Method 1: Change in accessibility rail (2011-2017): Borders Rail

		with skills matching			without matching			Ratio matching to no matching	
Group	Method	pre	post	% change	pre	post	% change	pre	post
Rail									
Treatment	Cost	0.007	0.017	159%	0.1	0.2	155%	11.09	10.89
	Time	0.022	0.054	149%	0.2	0.6	144%	11.09	10.86
Control	Cost	0.003	0.007	112%	0.2	0.2	0%	57.09	26.91
	Time	0.012	0.023	97%	0.4	0.4	0%	35.91	18.19

All results are aggregated job accessibility indexes for all locations within Borders Rail region, and suggest that without job skills matching there may be an overestimation of accessibility due to the seemingly high attraction factor of job opportunities. The wider difference when the index is based on travel cost suggests that cost is more of an impediment. Pre- and post-intervention differences show an impact due partially to the change in proximity brought by rail. Appendix 10.7.4 maps job accessibility based on travel cost and compares predictions for pre- and post intervention values using alternatively no matching and skills matching.

7.9 Essential services accessibility index

Although the job accessibility index represents one of the three main accessibility characteristics, to complete the picture, an appropriate accessibility index for essential services applicable to the Borders Rail case study region is also derived based on *4.11 Accessibility to essential services.* This index measures access to five essential services and reflects the particular characteristics of the region, estimating how rail developments have impacted on levels of accessibility through analysis by treatment and control group. An analysis of changes in the rail station accessibility index is shown in Appendix 10.7.5 with time as the basis, reflecting average accessibility across the case study region using treatment group TG4. On a time basis, average accessibility pre-intervention is higher for the treatment group, and less for the control group for all services. Following the intervention, there is a greater increase across all services for the treatment group (2.6%) as against the control group (1.6%).This also holds on a cost basis (*Appendix 10.7.5*) with a greater increase across all services for the treatment group (31.6%) as against the control group (15.7%). It is notable that the cost-based index is more sensitive to change than the time-based. Spatial analysis highlights variations in the services accessibility index across the case study region for individual data zones post intervention, alternatively applying a time-based and cost-based index. Appendix 10.7.5 summarises statistics for rail mode where travel time and travel cost have been the basis. The spread of accessibility is highlighted by the minimum, maximum and standard deviation, and is greater in the control group for nursery schools, secondary schools and hospitals.

7.10 Impact on residential property values

The treatment group configurations and accessibility characteristics applicable to Borders Rail are carried forward into the next stage in the process - the application of property and employment econometric models. Although this represents a sizeable rail improvement, Borders Rail differs from the other case study regions in being a more recent considerable intervention, and it may be difficult to detect any lasting house price impacts due to the rail intervention because of the limited amount of data available. As with Stirling-Alloa the immediate benefit at this stage will be of greater commuting potential to the job market in Edinburgh.

7.10.1 Initial descriptive analysis

Before applying these model approaches, knowledge of the background relevant to property price movements in the Borders Rail region may clarify variables applicable to the models. Following the methodology outlined in *4.12 Impact on residential property values*, an analysis evaluates the situation before and after rail intervention with reference to the standard treatment/control group (TG4), comparing changes in property types and house prices. TG4 incorporates a 2 km contour threshold with propensity matching, and there is effectively a third group, included here for comparison purposes, which is neither treatment nor control but comprises those datazones originally in the control group but rejected after application of propensity matching. The property type profile for treatment and control groups in 2011, prior to the intervention (Table 7-11), indicates a similar distribution across different types of property for the treatment group and the control group, whereas for data zones in neither of these two groups, although terraced property share is again similar, there is a substantially greater share of detached and semi-detached properties, and a much lower proportion of flats and apartments. This also reaffirms the validity of the matching process.

Table 7-11 Accommodation profile prior to intervention (2011-2017): Borders Rail

Group	Detached	Semi-detached	Terraced	Flat
Neither	31.15%	29.00%	22.61%	17.24%
Treatment	21.16%	25.13%	22.55%	31.16%
Control	25.60%	23.48%	21.20%	29.72%

A sharp drop in average house prices between 2007 and 2010 in the control group was followed by a similar drop in the treatment group areas between 2008 and 2013 (Figure 7-16). For two years either side of the intervention, for zones in neither group, there was a dropping off and levelling in prices. However, in the same period, there was an increase in house prices for both treatment and control groups, and this was steadier and more evident for the treatment group. This suggests an impact in those locations nearer the new rail link which was manifest prior to the reopening and may reflect a lead time effect in expectancy of improved accessibility.

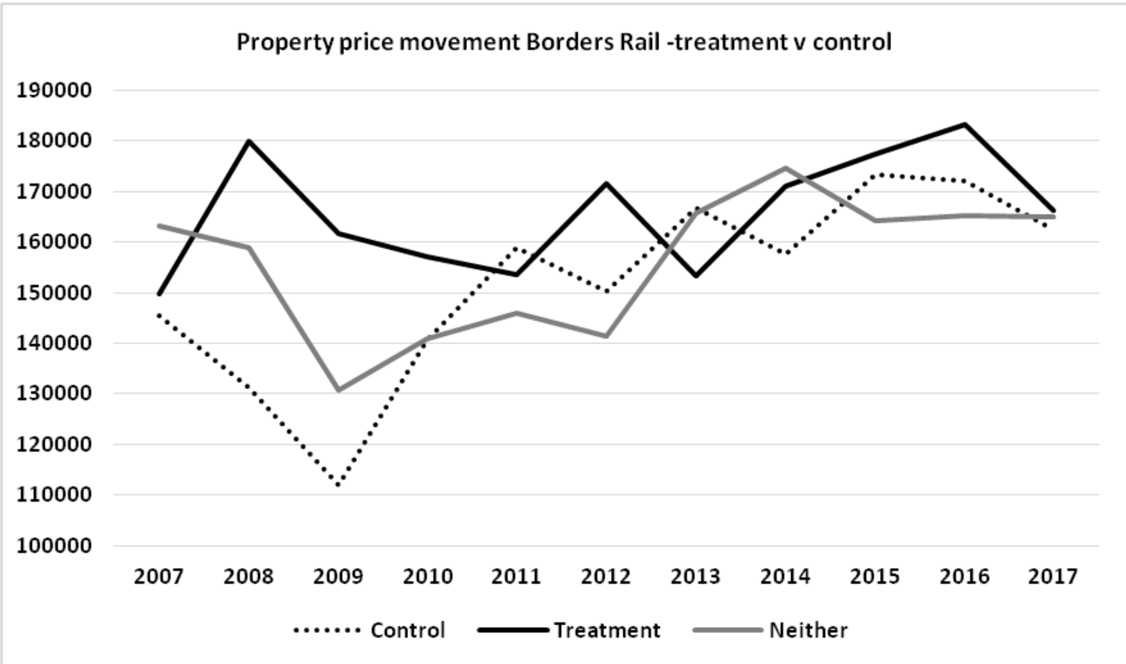


Figure 7-16 Average house prices 2007 to 2017 - treatment v control

(*Source: Rightmove, 2018)*

The regional perspective for Borders Rail is illustrated in Figure 7-17 which shows the distribution of property prices in the region in 2013 prior to the rail intervention, indicating a clear distribution shape with a small number of very costly properties. This distribution is strongly positively skewed - suggesting a large upper tail - i.e. a central group of typically-priced houses together with a long tail of relatively expensive ones. The summary statistics are mean (140582), median (127500), standard deviation (118564) and skewness (3.100).

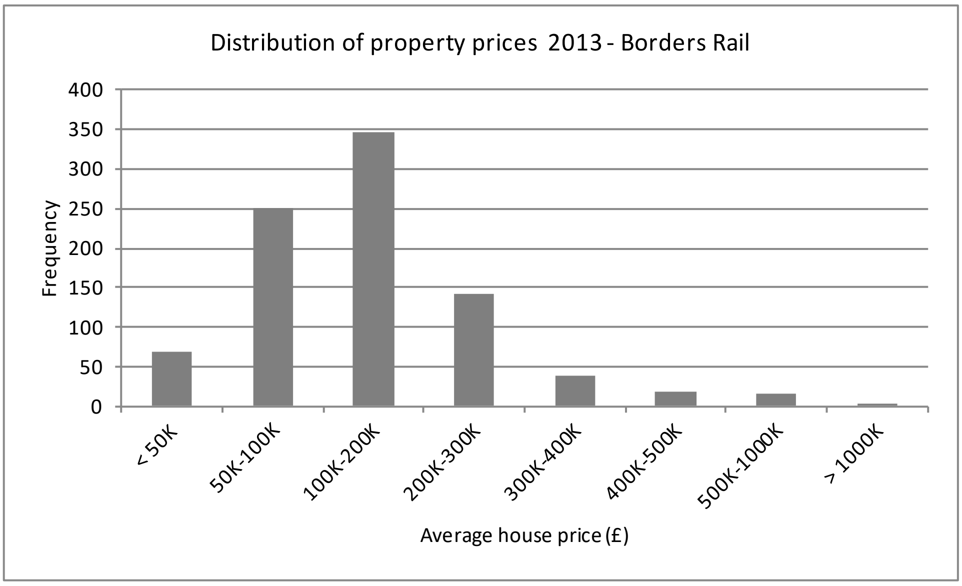


Figure 7-17 Distribution of property prices 2013 - Borders Rail

(*Source: Rightmove, 2018)*

7.11 Difference-in-difference model

Table 7-12 tabulates the model outputs for the difference-in-difference model applying two different variables of accessibility which reflect distance from the nearest station and job accessibility index. Firstly, with distance to nearest rail station as the accessibility variable,

Table 7-12 compares 2011 to 2017 for distance to rail station. This fails to indicate that a reduction in distance to the station made any difference in house prices over that period, but this is not statistically significant. If no distance characteristic is used, although there is a positive coefficient for treatment, again it is not statistically significant.

Table 7-12 Output from Property DID Model distance to station (2011-2017): Borders Rail

	Coefficient	SE	t value
Distance to nearest station			
(Intercept)	0.1971	0.1573	1.2530
Distance to nearest station	0.0074	0.0045	1.6440
Population density	0.0003	0.0386	0.0070
R^2	0.0535		
(Intercept)	-0.0511	0.0598	-0.8540
Treated	0.0035	0.1008	0.0350
Population density	0.0034	0.0397	0.0850
R^2	0.0002		

Using distance ratio (Table 7-14) again fails to indicate that a decrease in distance ratio had impacted positively on house prices over that period.

Table 7-13 Output from Property DID Model distance ratio (2011-2017): Borders Rail

	Coefficient	SE	t value
Distance Ratio			
(Intercept)	0.0456	0.1650	0.2760
Distance Ratio (DSR)	-0.1291	0.2135	-0.6050
Population density	0.0042	0.0395	0.1060
R^2	0.0077		

Using job accessibility on a cost basis with skills matching and ignoring movement of labour (Table 7-14) may suggest from the coefficient that an increase in job accessibility is related to an increase in house prices but again this is not statistically significant.

Table 7-14 Output from Property DID Model job accessibility (2011-2017): Borders Rail

	Coefficient	SE	t value
Job Accessibility			
(Intercept)	-0.0680	0.1139	-0.5970
Job Accessibility Index (JAI)	0.1034	0.5899	0.1750
Population density	0.0028	0.0398	0.0700
R^2	0.0008		

7.12 Fixed effects model

A fixed effects model is then applied to the Borders Rail dataset using aggregate data spanning the years 2013 to 2017 for each data zone with a record for each data zone and year. Property price data is collated from individual house transactions into an aggregated property price database, to run the model using the three different variables of accessibility.

For this model, a fourth accessibility characteristic was also considered - the time to reach Edinburgh - which was based on distance to Edinburgh divided by the fastest public transport mode using the average speeds calculated earlier. So prior to 2016 bus was the fastest at 35 km per hour, but post intervention, rail was the fastest at 65 km per hour.

Whereas the distance to Edinburgh had not changed over the intervention period, there was a change in proximity in terms of time. This additional consideration would also explore any movement in house prices detected due to the perceived greater ease of reaching Edinburgh, and to observe differences in Midlothian (nearer Edinburgh) and the Scottish Borders (further away).

Table 7-15 indicates the model summary with distance to the nearest station as the accessibility characteristic and log of property price as the dependent variable. The distance to rail station variable has a coefficient of -0.00016, suggesting distance to the nearest station having a negative effect on property price as distance increases. Because of the limited time span the result cannot be seen as significant.

Table 7-15 Model output with distance to nearest station as accessibility characteristic (2011-2017): Borders Rail

	Coefficient	Std. Error	t-value
Distance to nearest station	-0.00016	0.00073	-0.22300
Population density	-0.02607	0.02323	-1.12220
R^2	0.00359		

Fixed effects - locations (180) time (5)

Table 7-16 shows model results with distance ratio as the accessibility characteristic with an R^2 value of 0.0036 which is a slightly better fit than that based on distance to nearest station. The distance ratio has a coefficient of 0.00614 which indicates that an increase in distance ratio i.e. being proportionally nearer the station after the intervention, has a very small positive effect on property price as the relative change in distance increases, although this is not significant for the reasons given above.

Table 7-16 Model output with distance ratio as accessibility characteristic (2011-2017): Borders Rail

	Coefficient	Std. Error	t-value
Distance Ratio	0.00614	0.03422	0.17940
Population density	-0.02619	0.02325	-1.12650
R^2	0.00355		

Fixed effects - locations (180) time (5)

Applying cost-based job accessibility index yields a coefficient of 0.06881 which indicates that an increase in the job accessibility index has a positive effect on property price, but again this is not significant.

Table 7-17 Property model output using Job accessibility with skills matching (2011-2017): Borders Rail

	Coefficient	Std. Error	t-value
Job accessibility	0.06881	2.05918	0.03340
Population density	-0.02608	0.02326	-1.12100
R^2	0.00346		

Finally, if time to Edinburgh is added to the model (Table 7-18) there is a coefficient of –0.00510 which indicates that increasing travel time from Edinburgh has a negative effect on property price, but again this is not significant.

Table 7-18 Model output distance ratio using time to Edinburgh added (2011-2017): Borders Rail

	Coefficient	Std. Error	t-value
Time to Edinburgh	-0.00510	0.03601	-0.14170
Population density	-0.02601	0.02323	-1.11970
R^2	0.00351		

7.13 GWR Model

As an alternative approach, Geographically Weighted Regression (GWR) models changes in spatial diversity over the period of the rail intervention. The data frame contains census information collected for 180 data zones in the Borders Rail region with a grid reference marking the population weighted centre for each zone and uses house price data representing a sample of houses transacted in the region in 2013 and 2017. In contrast to a single set of constant values over the study area generated by an OLS model, GWR produces a set of parameter estimates and model statistics for each sample.

Because of the limited dataset, the models are run only for 2017 to compare coefficients spatially across the data zones. Detailed runs are shown in Appendix 10.7.6, and Table 7-19 provides a summary of the parameter estimates fitted by GWR, including the median, upper quartiles, lower quartiles, range, minimum and maximum and compares this with the OLS global model. This compares the GWR model coefficients applying different accessibility characteristics to the model. It is apparent that there is a negative relationship between price and distance to the nearest station which varies across the region. Using distance to nearest station, the coefficients ranged from a minimum value of -0.00904 to a maximum of -0.00025. So a 1 km increase in distance from the station results in a drop in average house price which varies between £2500 (2.5%) and £1000 (1%) for a £100,000 property.

The analysis compares GWR model results with the cross-sectional OLS model (which represents the global model) to detect any significant advantage in using GWR which involves reporting an improvement in fit of a GWR model over a global model using an F test to check significance. It is clearly evident that all of the interquartile values were less than the corresponding 2 standard errors of the global estimates (Table 7-19) indicating the lack of non-stationarity of relationships over the study area.

Table 7-19 Spatial variation of coefficients for different accessibility characteristics: (2011-2017): Borders Rail

Distance to rail station											
Minimum	25% quartile	Median	75% quartile	Maximum	IQ Range	Global (OLS)	SE	2SE	t value	R-squared	F stat
-0.0090	-0.0055	-0.0041	-0.0022	-0.0003	0.0033	-0.0050	0.0051	0.0103	-0.97	0.6653	14.46 on 11 and 80 DF
Job accessibility											
Minimum	25% quartile	Median	75% quartile	Maximum	IQ range	Global	SE	2SE	t value	R-squared	F stat
11.52	17.57	22.68	24.72	31.41	7.15	13.99	17.81	35.62	0.79	0.562	7.917 on 12 and 74 DF

7.14 Impact on jobs and employment

Following the methodology outlined in (*4.13 Impact on jobs and employment*) an initial comparative analysis, particularly relevant to jobs and employment, is applied to the data and characteristics of the Borders Rail line, and analysed by treatment and control groups, again using treatment group combination TG4 for illustration purposes into education levels and qualifications, Job Seekers Allowance by LSOA, economic activity and job profile by industry and occupation.

Educational levels and qualifications represent potential for employment, and Table 7-20 highlights variations in educational standards prior to the rail intervention by treatment and control group. There is a similarity between the two groups and also those datazones in neither group, except that there are fewer in the treatment group at level 4 and above.

Table 7-20 Educational standards prior to the rail intervention (2011): Borders Rail (*Source: UK Census, 2011)*

Group	No qualifications	Level 1	Level 2	Level 3	Level 4 and above
Neither	26.65%	23.35%	13.99%	8.76%	27.25%
Treatment	27.16%	25.74%	14.29%	9.06%	23.75%
Control	28.10%	23.27%	13.98%	8.56%	26.09%

Table 7-21 highlights employment levels prior to the rail intervention by treatment and control group. Unemployment is similar in both groups, with fewer self-employed in the treatment group.

Table 7-22 Economic activity prior to rail intervention (2011): Borders Rail (*Source: UK Census, 2011)*

Group	Full-time	Part-time	Self-employed	Full-time students	Unemployed
Neither	53.32%	20.98%	17.16%	3.00%	5.54%
Treatment	56.52%	21.01%	11.50%	4.35%	6.62%
Control	54.04%	21.50%	15.19%	3.09%	6.18%

Occupation is important to the job accessibility index when comparing job opportunity matches between different locations, particularly in deriving a skills matched job index

Table 7-23 Jobs by occupation in 2011 - treatment v control (*Source: UK Census, 2011)*

Group	Managers	Professional	Associate professional and technical	Administrative and secretarial	Skilled trades	Caring leisure and other service	Sales and customer service	Process plant and machine operatives	Elementary occupations
Neither	9.88%	15.88%	11.37%	10.41%	16.43%	10.46%	7.22%	7.27%	11.08%
Treatment	8.75%	15.32%	11.57%	12.33%	13.39%	11.27%	9.58%	6.92%	10.87%
Control	9.50%	15.48%	10.73%	10.41%	16.05%	10.70%	8.04%	7.38%	11.70%

There are slightly less percentages of residents in the treatment group with managerial and professional occupations and skilled trades in the treatment group (Table 7-23) in contrast to a higher proportion in administrative, care and leisure services and sales occupations.

Figure 7-19 shows the distribution of employment density for the entire Borders Rail case study region. There is a median of 6 jobs per unit area but a high proportion with density less than 2, and at the other extreme very few locations where the density exceeds 24 per unit area.

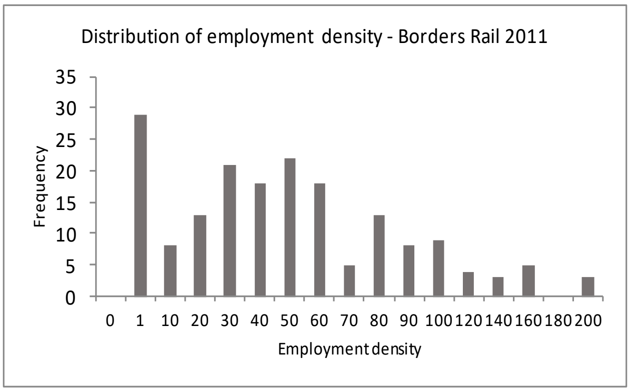


Figure 7-18 Distribution of employment density (2011): Borders Rail (Source: UK Census, 2011)

This places the region in context where there are a number of locations with very little employment or potential for jobs, and a semi-urban group of locations having a limited level of employment (the "thin" market). Hence, it provides a background appropriate to this study as a disconnected region, in contrast to the intensely urban nature of those areas previously studied in the literature.

Table 7-24 indicates less manufacturing and agriculture in the treatment group, but otherwise a broadly similar split of industries across all groups.

Table 7-24 Jobs by industry - treatment v control (2011): Borders Rail

Group	Neither	Treatment	Control
Agriculture forestry and fishing	6.20%	1.66%	4.71%
Mining and quarrying	0.24%	0.27%	0.24%
Manufacturing	8.04%	5.81%	8.50%
Electricity gas and water supply	0.48%	0.70%	0.43%
Construction	8.74%	8.69%	8.99%
Wholesale and retail trade repairs	13.81%	15.80%	14.72%
Transport and storage	3.84%	4.53%	3.81%
Accommodation and food service activities	5.12%	5.05%	5.66%
Information and communication	1.88%	2.08%	1.87%
Financial and insurance activities	4.11%	6.74%	4.21%
Real estate activities	1.68%	1.33%	1.60%
Professional scientific and technical activities	5.40%	4.80%	5.20%
Administrative and support service activities	4.03%	4.22%	3.92%
Public administration and defence	6.78%	7.64%	5.97%
Education	8.14%	8.13%	8.07%
Human health and social work activities	15.41%	16.58%	15.73%
Water supply	0.71%	0.90%	0.79%
Arts entertainment and recreation	5.19%	4.99%	5.43%

7.15 Difference-in-Difference model

The difference-in-difference model compares two separate years pre- and post-intervention to predict what would have happened had there been no intervention and applying two different variables of accessibility which reflect distance from the nearest station and the job accessibility index.

Comparing 2013 to 2017 for distance to nearest rail station as the accessibility variable (Table 7-25) suggests a positive effect on employment density given a reduction in distance to the station, but it is not statistically significant. This is also reflected when considering treatment independent of distance characteristic.

Table 7-25 DID job model using distance to station (2011): Borders Rail

	Coefficient	Std.	t value	
(Intercept)	1.18166	0.92756	1.274	
Distance to nearest station	-0.0247	0.02666	-0.927	
Population density	-0.06009	0.22764	-0.264	
R^2	0.0195			
(Intercept)	1.65665	0.33593	4.932	***
Treated	0.98057	0.56616	1.732	.
Population density	-0.09136	0.22286	-0.41	
R^2	0.0607			

The output shown in Table 7-26 compares 2011 to 2017 for distance ratio and job accessibility as the accessibility variable. This suggests that an increase in job accessibility may cause some increase in employment density over that period, but this is not significant at this stage, probably due to the limited amount of data available.

Table 7-26 DID job model using distance ratio and job accessibility (2011-2017): Borders Rail

	Coefficient	Std. Error	t value
Distance Ratio			
(Intercept)	1.4645	0.9551	1.533
Distance Ratio (DSR)	0.7273	1.2358	0.589
Population density	-0.0748	0.2287	-0.327
R^2	0.0091		
Job Accessibility			
(Intercept)	1.52118	0.65501	2.322
Job Accessibility Index (JAI)	2.74871	3.39168	0.81
Population density	-0.08834	0.22888	-0.386
R^2	0.0154		

7.16 Geographically Weighted regression (GWR)

Using the model based on employment density as the dependent variable considered earlier, the same transport accessibility characteristics are substituted in turn to generate the OLS global models and GWR model.

Because of the limited dataset, the models are run only for 2017 to compare coefficients spatially across the data zones. Detailed output is shown in Appendix 10.7.7, and Table 7-27 summarises the parameter estimates for 2017, including the median, upper quartiles, lower quartiles, range, minimum and maximum.

Table 7-27 Comparison of GWR coefficients for different accessibility characteristics (2011-2017): Borders Rail

Distance to rail station									
Min	25% quartile	Median	75% quartile	IQ range	Max	Global (OLS)	SE	t value	R^2
-0.077	-0.065	-0.046	-0.041	0.024	-0.032	-0.053	0.023	-2.307	0.821
Job accessibility									
Min	25% quartile	Median	75% quartile	IQ range	Max	Global	SE	t value	R^2
25.14	153.2	243.8	326	173.2	394.5	260.6	165.5	1.575	0.821

Going across the 180 sample points, employment density decreases with increasing distance to the nearest station which reflects the global model. For distance to nearest station, the coefficients ranged between -0.0676 and -0.0324, so a 1 km increase in distance from the station results in a percentage drop in employment density which can vary between 87.8 and 13.94.

The distribution of values is symmetrical as the median is -0.0459, approximately halfway between the maximum and minimum values. This shows that over the intervention period, improvement in distance to the nearest station has had the effect of increasing employment density. For job accessibility, the coefficients range between 25.14 and 394.5, so a 1 km increase in distance from the station results in a drop in average employment density which varies between 87.8% and 13.94%. The analysis compares GWR model results with the cross-sectional OLS model (which represents the global model) to detect any significant advantage in using GWR by reporting an improvement in fit of a GWR model over a global model using an F test to check significance. It is clearly evident that all the interquartile values were less than their corresponding 2 standard errors of the global estimates for distance from the nearest station (Table 7-27) indicating the lack of non-stationarity of relationships over the study area.

7.17 Sensitivity analysis

Analysing the impact of the rail intervention through application of a job accessibility index and property and employment models using the standard treatment/control group configuration treatment/control group combination (TG4) illustrated limited evidence of attributable causality, but it was neither practical nor feasible at that stage to investigate every possible group combination. This section revisits that analysis by examining the effect of applying alternative selections into treatment. The job accessibility index is re-run using each treatment group/control group combination to investigate how sensitive the results are to different criteria for selection into treatment.

The job accessibility index for rail is calculated to see how group selection impacts on the value of the accessibility measure. Variations in the job accessibility for all combinations of treatment and control groups are shown in Table 7-28 which shows the estimated increase in accessibility between 2011 and 2017 expressed as a percentage.

Table 7-28 Job Accessibility increases for different treatment groups

Group Combination ID	Group	Cost basis		Time basis	
		Matching	No Matching	Matching	No Matching
TG1	Control	196.40%	196.40%	257.10%	257.10%
	Treatment	151.40%	148.00%	126.60%	125.50%
TG2	Control	313.00%	309.80%	401.70%	396.10%
	Treatment	129.40%	127.40%	111.20%	110.90%
TG3	Control	509.10%	494.00%	580.90%	563.50%
	Treatment	123.40%	121.80%	107.50%	107.40%
TG4	Control	183.20%	181.50%	243.80%	241.70%
	Treatment	151.40%	148.00%	126.60%	125.50%
TG5	Control	231.00%	232.50%	297.80%	300.70%
	Treatment	233.40%	225.30%	188.80%	184.40%

The predicted increase in job accessibility is fairly consistent, regardless of the selection of treatment group - 2.63% variation with skills matching and 0.31% variation with no matching on a cost basis and similar figures on a time basis. However, there is a much wider variation for the respective control groups - 7.63% variation with skills matching and 1.02% with no matching on a cost basis and similar figures on a time basis. This reflects the reducing size of the control group with expansion of the treatment group to cover greater distance.

7.18 Conclusion

Of the three case study regions, Borders Rail is definitive in meeting all the criteria of this research in reconnecting remote areas to the rail network and represents a substantial section of line and access to the network via seven new stations. It provides a good commuting link to Edinburgh with it larger job market, and there is also a lack of confounding factors which would make establishing causality more difficult. There is a mix of smaller communities previously remote from the rail network, and in a geographically cut-off location. However, on the negative side, as the rail intervention has been in place for just 3 years, there is very limited opportunity to observe long term wider economic impacts, although there is some early suggestion here of some property impacts.

The findings indicate that rail accessibility shows some impact on property prices and employment density. However, because of the short time since the intervention, any impacts are so far relatively small, apart from a greater accessibility to jobs in Edinburgh with evidence of some movement in house prices in the immediate intervention period in treatment group areas. It is apparent that the distance from Edinburgh may be a further significant factor to consider, and analysis has shown that in more remote or non-urban regions, distance to a major centre (Edinburgh) can also provide a viable alternative measure. In addition, the divide between Midlothian and the Scottish Borders needs to be taken into account. For Borders Rail, more time is needed before a proper assessment of impacts can be carried out.

Counterfactual

The region was divided into treatment and control groups to monitor the effects of infrastructure changes using distance thresholds from the intervention to create basic groups. Unlike Stirling-Alloa, the extension above the 6 km distance contour to the 10 km contour presents a feasible alternative without the need to extend the remit of the dataset, and the case study incorporates 180 datazones which represent a larger pool for division into treatment and control groups.

Propensity matching of those treatment and control groups with similar levels of deprivation worked well here for the 2 km distance contours, as did clustering techniques using the dominant cluster groups for matching. Although matching was less effective at the 4 km threshold, there was still some improvement, but going beyond the 4 km contour required a larger base treatment group to be matched against a much larger selection of potential locations in the control group.

Unlike the other case studies much data is limited to that at the 2011 UK Census. The high percentage using car either as driver or passenger for both groups reflects the reliance on cars across the region. More commute to distances of under 10 km in the control group than the treatment group, which reflects the geographical nature of the region with its concentration of population and employment around Galashiels (i.e. typically treatment group), and those in more remote places (i.e. typically control group) must travel further to work. Usage of Borders Rail stations has grown faster than similar pre-existing stations nearby. A markedly positive uptake in rail usage since the re-opening of Borders Rail includes an increasing number of season tickets especially from Galashiels - the most populated centre being served.

Accessibility characteristics

The region should be considered as made up of two halves - Midlothian to the north - much closer to Edinburgh and the Scottish Borders to the south - more distant from Edinburgh. It is therefore important to consider the pull of the Scottish capital and potentially different impacts in the two sub-regions. For most locations, the seven new stations offering closer access to the rail network, but this will be much greater in the Scottish Borders. However, for Borders Rail it was apparent that care must be taken in using the ratio as it can be larger for a datazone in Midlothian, even though the original distance to the nearest station and actual change in distance is much smaller than in the Scottish Borders.

Job accessibility

Comparing the average job index by travel mode for two separate years spanning the intervention (2011 and 2017), accessibility by rail improves more for the treatment group than the control group, and this improvement is more noticeable for rail mode where there is an increase in accessibility post-intervention. Again, results indicate that without job skills matching the job opportunity attraction can be overestimated. Consequently, an index based on a combination of cost and job skills matching is preferable for this case study, assuming travel cost rather than journey time will be the prime deterrent factor in more remote regions such as Borders Rail. Job opportunities are normally excluded beyond a defined threshold; in this case exceptions are made for the city of Edinburgh, which although outside the case study region, provides a powerful pull in terms of job opportunities. The predicted increase in job accessibility is fairly consistent, regardless of the selection of treatment group on a cost or time basis. However, there is a much wider variation for the respective control groups with application of skills matching.

Impact on residential property values

There is evidence of a "lead time" effect where prices rise in anticipation of the new rail link. For two years either side of the intervention, the increase in average house price is steadier and more marked for the treatment group, whereas for zones in neither group, there was a dropping off and levelling in prices. However, in the same period, there was an increase in house prices for both treatment and control groups, and this was steadier and more evident for the treatment group. This suggests an impact in those locations nearer the new rail link which manifest itself prior to the reopening and may reflect a lead time effect in expectancy of improved accessibility.

However, the difference-in-difference model gave mixed results in trying to establish causality. It could not identify any house prices effect of change in distance to station or distance ratio, but did find some effect of an increase in job accessibility.

Using the limited data available since the intervention, the fixed effects model is applied to the Borders Rail dataset using aggregate data spanning the years 2013 to 2017. For these models, an additional proximity characteristic was also considered - the time to Edinburgh - which would reduce over the period and was thought to influence house prices as house price levels in Midlothian (nearer Edinburgh) are higher than those in the Scottish Borders (further away). When time to Edinburgh is added to the models, it does appear to have an effect and suggests that ease of access to Edinburgh is also important. Results indicate an impact of rail accessibility on property prices. Distance to the nearest station has a negative effect on property price as it increases, and an increase in distance ratio i.e. being proportionally nearer the station after the intervention, has a positive effect on property price as the relative change in distance increases, but this is not significant. Finally, the cost-based job accessibility index with skills matching indicates an increase in the job index has a positive effect on property price.

The negative relationship between price and distance to the nearest rail station varies across the region with a cluster of higher negative coefficient values at the centre of the region, and lower negative coefficient values on the periphery. Comparing the GWR model with the cross-sectional OLS model to detect any significant advantage in using GWR, there is a lack of non-stationarity of relationships over the region.

Impact on jobs and employment

The distribution of employment density typifies the whole Borders Rail region with its mixture of locations with little potential for jobs and small townships and a limited level of employment - the "thin" market. For the difference-in-difference model, using distance to station as the accessibility variable suggests that a reduction in distance to the station or increased job accessibility will contribute to an increase in employment density over the intervention period.

For GWR, the relationship between employment density and distance to the nearest station was in most cases negative - reflecting the OLS model. Unlike the results for property, the presence of non-stationarity of relationships is clearly evident over the study area. Comparing GWR model results with the cross-sectional OLS model it is clearly evident that there is a lack of non-stationarity of relationships over the study region.

8 Chapter Eight: Results and findings

Each case study has focussed on the key outcomes of employment, property price and accessibility to jobs and services and categorised relevant characteristics of each region such as population, distance to public transport, alternative travel modes and local transport provision. Following on, this chapter considers the importance of regional context, especially in the measurement of a counterfactual (treatment and control groups), and in deriving the accessibility characteristics, in particular the job accessibility index which acts as both a model for job accessibility and an accessibility characteristic for inclusion in econometric modelling. A comparison of the various models includes a conclusion as to their findings along with a critique of apparent strengths and weaknesses.

8.1 The importance of regional context

A key finding, through reference to appropriate case studies and assessment of relevant characteristics of the affected regions, is the importance of context in affecting the economic impact resulting from each rail intervention. There are differences in all three case studies, specifically based on the local geography and size of the intervention which have influenced the extent of this impact. All three case studies comprise a mix of medium and small communities previously disconnected from the rail network as a result of Beeching, and a common feature is the reconnection to a large urban centre e.g. Edinburgh, Glasgow, Nottingham (Table 8-1). There are also similarities between each region in the state of the local economy reflected in the population, property and employment profiles, and also in levels of deprivation and age distribution, where there are more elderly and fewer younger people. All regions have experienced some degree of industrial decline which has brought about a redistribution of industry.

Table 8-1 Case study context comparison

Background	Robin Hood Line	Stirling - Alloa	Borders Rail
When opened	1993-1998	2008	2015
Elapsed time	20-25 years	10 years	3 years
Large urban centre	Nottingham	Stirling (Glasgow)	Edinburgh
Major town	Mansfield	Alloa	Galashiels

In contrast, the re-opened rail links have been in place for varying lengths of time, and differ in the scale of line re-opening. There are also differences in the size of the region, the distribution of larger communities and the prevalence of other transport modes. The widest contrast is between the Robin Hood Line - re-opened over twenty years ago - and comprising remote townships experiencing industrial decline, and Borders Rail - re-opened very recently - where 30% of the population reside in settlements of fewer than 500 people. The 1990s housing market recession lasted seven years, and there is some evidence that the "crash" had lasting impacts on the worst affected neighbourhoods (Forrest et al., 1997). This was followed by a similar crash from 2007 to 2011.

The case studies represent rail interventions, not just at different stages of establishment, but on a varying scale in terms of number of new stations opened, length of track, and service levels. This is reflected in the findings which show a variable impact on house prices, jobs and commuting patterns. There are also differences in the size and topography of each region from the smallest in Stirling-Alloa to the largest in the Robin Hood Line. Borders Rail is similarly large but more remote, and in addition, unlike the other case study regions, can be divided into the Scottish Borders, where travel distances are greater, and Midlothian, where there may be a greater benefit in being closer to Edinburgh, which is consequently more commutable in terms of time and cost. This has led to some difficulties over the size of the dataset and having to allow for locations of influence external to the immediate case study region.

The "remoteness" element is reflected in the distribution of communities. Stirling-Alloa and Borders Rail regions generally comprise sparsely populated settlements with consequently fewer job opportunities, which in the case of Borders Rail are very thinly distributed. For the Robin Hood Line there is a mixture of settlement sizes, but some towns, such as Mansfield, are much larger than any in the other case studies. This is apparent in the relative distribution of employment density. All regions comprise a mixture of locations with little or no employment or potential for jobs (a thin market), and urban or semi-urban locations with a limited level of employment. The Robin Hood Line has a high proportion of areas with employment density of less than 2 jobs per square km and very few locations where the density exceeds 30 jobs per square km. Stirling-Alloa and Borders Rail also have a high proportion with density less than 2 jobs per square km but very few locations where the density exceeds 24 jobs per square km.

The level of transport provision offers a further contrast as the Stirling-Alloa and Robin Hood Lines had accessible rail stations in place prior to the intervention which although not necessarily offering a viable alternative for commuting, were near enough to access via other transport modes. There were also alternative modes available especially in Stirling-Alloa, and in the case of the Robin Hood Line, the development of other transport provision development in the tramway system.

8.2 Selection into treatment and control groups

The literature stresses the importance of checking causal relationships of the project and that a successful wider impacts impact evaluation must establish causality through the counterfactual. The What Works-Evidence Review (What Works Centre for Local Economic Growth, 2015) recommended establishing causality through comparing outcomes between treatment and control groups. Many studies have addressed the key issue of "selection into treatment" using variations on difference-in-difference or panel fixed effects methods. In this study, the control group was constructed to be similar to the treatment group through matching on observed characteristics taking a before-and-after difference to eliminate all fixed unobservable differences between groups. However, establishing control sites has often proved problematic due to identifying areas not exposed to similar interventions.

For this study, treatment groups were selected based on distance thresholds with the control group selected from remaining locations in the region. However, as this could lead to treatment and control groups not fully comparable in terms of observables, there was further application of propensity score matching and clustering techniques to adjust for differences between treated and untreated groups to reflect the similarities and dissimilarities.

Table 8-2 Matching statistics by case study region

Matching		Robin Hood Line	Stirling - Alloa	Borders Rail
Propensity Matching	Matched at 2 km	127	25	57
	Improvement %	72.68	49.71	75.50
	Matched at 4 km	224	N/A	81
	Improvement %	11.94	N/A	30.20
Clustering	Clusters	5	3	4
	Largest Cluster Size	37%	41%	44%

For all case study regions, propensity matching on area size and deprivation has worked very well for a 2 km threshold, by producing a similar deprivation profile in each group (Table 8-2). Matching was best for Robin Hood Line and Borders Rail at over

70% improvement over the base group, probably due to the size of the groups which were much larger than those for Stirling-Alloa. Matching at 4 km threshold required larger treatment and control groups and hence matching was more difficult, Hence the improvement over the base group was much lower (only 11.94% for the Robin Hood Line), and the dataset for Stirling-Alloa was too small to extend to 4 km. This limitation was due to the size and remit of the datasets, requiring a larger base standard treatment group to be matched against a much larger selection of potential locations to create a control group.

As an alternative to propensity matching, clustering produced respectively five distinct clusters for the Robin Hood Line, four clusters for Borders Rail and three for Stirling-Alloa (Table 8-2) based on five deprivation characteristics, and in each case the dominant cluster was matched against a 2 km "base" group. The number of clusters reflected the relative size of the regions and diversity across communities, which is more noticeable in the Robin Hood Line region than the others.

For all the case study regions, there appeared to be some benefit in further application of clustering and propensity matching so that a more meaningful comparison could be made between similar locations in the treatment and control groups. The limitation caused by the size of the dataset prevented further analysis at extended thresholds, and eventually only five alternative group combinations (3 base groups and 1 each propensity matching and clustering) were carried through into the modelling.

Application of propensity matching produced a treatment/control group combination containing locations with a similar socio-economic profile. The following comparisons in this chapter by treatment and control group are based on a 2 km catchment for the treatment group which is matched to similar locations in the control group. While the results indicate an effect of treatment, there are no hypothesis tests to establish whether the treatment results are significantly different from the control results, but this would be a suitable and feasible addition to the methodology to be included in future research.

8.3 Accessibility characteristics

A key objective was to associate measures with outcomes and to this end several accessibility measures were alternatively applied based on suggestions from the literature*.* Distance to the Central Business District (CBD) - or in this case major urban centre - had been seen as the most important factor as proximity to transport facilities increased speed of travel to the CBD, which translated into the value of the property.

However, this was only applied in the case of Borders Rail where access to Edinburgh was key to the restoration of the line. Previous research suggests a causal link between property prices and accessibility (Debrezion, 2007) and immediate locations are expected to produce higher effects than locations further away (Bowes and Ihlanfeldt, 2001). Three different measures were used for defining accessibility; distance from the nearest rail station, distance ratio, and a job accessibility index.

Table 8-3 Accessibility Characteristics statistics by case study region

<table>
<tr><th>Accessibility Characteristic</th><th colspan="4">Robin Hood Line</th><th colspan="4">Stirling - Alloa</th><th colspan="4">Borders Rail</th></tr>
<tr><td colspan="13">Distance to nearest station</td></tr>
<tr><td></td><td colspan="2">Treatment</td><td colspan="2">Control</td><td colspan="2">Treatment</td><td colspan="2">Control</td><td colspan="2">Treatment</td><td colspan="2">Control</td></tr>
<tr><td>Mean</td><td>9.34</td><td>1.22</td><td>6.21</td><td>3.92</td><td>8.73</td><td>1.06</td><td>9.1</td><td>4.25</td><td>18.06</td><td>1.03</td><td>40.67</td><td>14.69</td></tr>
<tr><td>Median</td><td>8.65</td><td>1.28</td><td>5.34</td><td>3.29</td><td>8.9</td><td>0.94</td><td>8.94</td><td>4.09</td><td>7.95</td><td>0.96</td><td>45</td><td>19.24</td></tr>
<tr><td colspan="13">Distance Ratio</td></tr>
<tr><td>Mean</td><td colspan="2">0.869</td><td colspan="2">0.369</td><td colspan="2">0.879</td><td colspan="2">0.533</td><td colspan="2">0.943</td><td colspan="2">0.639</td></tr>
<tr><td>Median</td><td colspan="2">0.852</td><td colspan="2">0.385</td><td colspan="2">0.895</td><td colspan="2">0.542</td><td colspan="2">0.88</td><td colspan="2">0.572</td></tr>
</table>

Distance to nearest station

For the Robin Hood Line and Stirling-Alloa regions, most locations were closer to a rail station following the intervention (Table 8-3), although some locations experienced negligible or no change. However, for Borders Rail, all data zones were now much nearer by some considerable margin, and this is more noticeable in the Scottish Borders rather than Midlothian - the latter being much closer to Edinburgh. Although a change in distance to rail station suggested that property prices increase the nearer to a rail station, it was not significant, and this could be either because it is not important in this case or that there is insufficient data. A limitation of distance to station, apart from taking no account of travel time or travel cost does not considering the scale of the change in distance to station following the intervention.

Distance ratio

As an alternative, the distance ratio appeared more meaningful as it expressed a change in the relative nearness of the rail station (Table 8-3). The ratio has a value between 0 and 1 (where 0 indicates no change in distance and 1 implies now being adjacent to a station). This worked well as an accessibility characteristic in the models for all the case study regions. Its application in the models indicated some causality, both in terms of property price movement and change in employment density, especially for the Robin Hood Line region.

However, within the Borders Rail region the main movement in Midlothian is in the 0.75-0.85, whereas for the Scottish Borders there are two peaks - the 0.55-0.65 range and a further peak at 0.95. Hence, care must be taken in using the ratio which can be greater for a datazone in Midlothian, even though the original distance to the nearest station is much less. This discrepancy in the Scottish Borders is due to the distribution of population where a concentration of locations around Galashiels experience a great improvement because they are now very close to the rail link. On the other hand, peripheral areas of the region such as Peebles, although nearer the network, are still a considerable distance from the nearest rail station.

8.4 Job accessibility index

The third accessibility measure, the job accessibility index was designed to be applicable to remote regions acting as both an accessibility model in its own right, and also as an accessibility characteristic to be incorporated into an econometric model. It involved a much more complex formulation than the previously discussed characteristics, and required some calibration for each case study region to address specific issues:

- the decay effect in more remote areas
- commuting feasibility in remote areas
- skills mismatch and job competition effects

8.4.1 The decay effect in more remote areas

In measuring longitudinal shifts in job accessibility for gravity-based measures a distance decay function was calibrated for different modes and household characteristics. For this current study it often proved difficult owing to historical data requirements, especially for the Robin Hood Line which opened more than twenty years ago. For each case study, travel to work statistics were analysed to estimate the most applicable decay function (negative exponential) and decay parameter β which was dependent on the state of the transport system, traffic congestion and unwillingness to travel (Harris, 2001). For this study, calibration of the job accessibility index produced a similar decay parameter (-0.028 to -0.030) for all three regions (Table 8-4) which represented a common feature of all case study regions.

Table 8-4 Beta value for each case study region based on distance

	Robin Hood Line	Stirling-Alloa	Borders Rail
Beta	-0.03	-0.028	-0.03
R^2	0.403	0.6094	0.5184

A threshold was set to define the furthest that commuters were prepared to travel which was based on travel to work statistics from the UK Census and other sources. Nottingham is an integral part of the Robin Hood Line region dataset, but for the other regions exceptions were made to allow for Stirling, Glasgow and Edinburgh which, although outside the case study regions, provided a powerful pull in terms of job opportunities. This avoided constraining the scope of reachable jobs to those within the specific case study region as workers may apply for jobs outside their own region.

Commuting feasibility

Employment is considered the most likely single destination type for an accessibility measure (Horner and Mefford, 2005), and so the feasibility of commuting was seen as a key factor in assessing accessibility for this study. Travel time was found significant for house location choice (Zondag and Pieters, 2005), but the Borders Rail Survey (Chapter 2) had identified temporal barriers where there is a mismatch between the availability of services and when people can access them, or where travel times exceed some acceptable maximum threshold. Korsu and Wenglenski (2010) suggested jobs located less than 60 minutes away from a residence were reachable, but after analysing distance travelled to work and taking into account the rural context and nearby larger conurbations, a 75 minute threshold was adopted for this study.

Affordability i.e. the cost of travel had been recognised as a key factor affecting accessibility (Social Exclusion Unit, 2003). Generalised cost allowed measures for different trip purposes to be combined since the same units are used for all trip purposes and assigned a value to accessibility i.e. its monetary “worth” for the journey–to-work trip. For each case study, average speed and time costs of respective transport modes were used to estimate the total journey cost for passengers making journeys with different transport modes over different distances using generalised journey costs. Samples of times and costs from different local transport estimates for the different regions produced a set of similar values for the assessment of travel time and cost.

8.4.2 Job availability and skills mismatch

Accessibility to reachable jobs available to any worker, should depend on the number of competitors claiming to form a match (Weibull, 1976; Ihlanfeldt, 1993; Harris, 2001; Van Wee et al., 2001; Kawabata and Shen, 2007), and here "occupational match" is used as a measure of job accessibility, where previously little measurement of job accessibility has appropriately incorporated this diversity element,

Although a direct measure of vacancies would ideally be used rather than all existing jobs (Ihlanfeldt and Sjoquist, 1998), data availability issues restricted consideration of job availability to occupied jobs and active workers instead of vacancies and actual job seekers as per (Korsu and Wenglenski, 2010). In the skills matching calculations used in this study, job accessibility for each origin location uses the percentage skills share at that location, not the actual number of jobs or skills available. This is then matched to actual jobs at all the other destination locations (excluding the origin location) to counter the potential for endogeneity in the accessibility calculation. Using the gravity formulation specified in the Chapter 4, two alternative methods were used as a basis for the skills matching element:

- Method 1: The number of jobs available in the base (pre-intervention) year
- Method 2: The number of jobs available in the current (post-intervention) year

Table 8-5 Job Accessibility changes by case study - treatment v control

Case Study Region			Robin Hood Line		Stirling - Alloa		Borders Rail	
			% change		% change		% change	
Method	Group	Base	with skills matching	without matching	with skills matching	without matching	with skills matching	without matching
1	Treatment	Cost	12%	68%	80%	80%	159%	155%
		Time	5%	54%	17%	17%	149%	144%
	Control	Cost	-13%	25%	71%	70%	112%	0%
		Time	-9%	32%	23%	23%	97%	0%
2	Treatment	Cost	-15%	38%	7%	1%	159%	155%
		Time	-9%	48%	6%	0%	149%	144%
	Control	Cost	-37%	1%	6%	1%	112%	0%
		Time	-22%	27%	6%	0%	97%	0%

As expected, these predicted different values for the job accessibility index (Table Table 8-5). Using the second method, the fall in the number of jobs post-intervention between 2001 and 2011 reduced the accessibility benefits gained by improvement in proximity to the rail link resulting in a lower value of the index. Although this offered a realistic measure of the current state of job accessibility, it would not be applicable as an accessibility characteristic in the econometric models as it represented a future position. Hence this method was used to evaluate the current post-intervention relative to job availability, but not used in the econometric models.

On the other hand, using pre-intervention job figures predicted improved accessibility post-intervention relative to the jobs profile that existed at the time (Gibbons et al., 2012), and hence reflected a change in nearness to the rail link rather than movement in the job market. Hence, this method was adopted as most suitable for econometric modelling as it reflected the position leading up to the intervention.

One advantage of using two methods for the job accessibility index is that, as well as being a measure of accessibility based on the original job market, it can also predict the effect of a slump in employment on accessibility by applying more recent job figures. In addition, as seen later, it gives good results when used as an accessibility characteristic in the hedonic models, and is more relevant than distance from station as it takes into account the whole regional employment picture relative to each location.

8.4.3 Results summary

Across all case study regions, there is an increase in job accessibility for rail mode. Whether using a travel time or cost basis with skills matching, the effect of being closer to the intervention is illustrated in a greater improvement in accessibility for the treatment group than the control group, and an increase in accessibility post-intervention for rail mode.

Table 8-5 also indicates that without job skills matching, job accessibility may be overestimated due to the seemingly high attraction of job opportunities which may not synchronise with the skills set in that location. For example, for the Robin Hood line, there is a change in accessibility of 68% since the intervention to all jobs available, but this change is only 12% when skills matching is taken into account indicating that many of the jobs in other locations do not match those in each origin location. Comparing the average job index for two separate years spanning the intervention, regardless of a time or cost basis, accessibility by rail in treatment areas improves over the control areas with a positive shift in minimum and maximum values. The greater narrowing of index range in the treatment group and the wider relative difference when the index is based on travel cost rather than travel time suggests that cost is more of an impediment.

Sensitivity analysis from the job accessibility index exhibits consistency in predicting the effect of a change in accessibility due to the rail intervention for different selections into treatment, and the inclusion of a cost or time constraint and allowance for matching of jobs makes it more pertinent to the rural or semi-rural situation. It is also evident that although job accessibility looks favourable without skills matching, this can lead to an overestimation of opportunities, and a more realistic measure takes into account the skills set at each location.

8.4.4 Essential services accessibility

Analysis of the essential services index indicates that, following the intervention, regardless of a cost or time basis, there is an increase in accessibility in the treatment group compared to the control group (Table 8-6). The exception to this is Borders Rail where in the case of hospitals and schools it has made no difference.

Table 8-6 Essential services changes by case study - treatment v control

Robin Hood Line					
Base	Group	Nursery	Primary	Secondary	Hospital
Time	Control	0.38%	0.49%	0.50%	0.50%
	Treatment	1.40%	1.49%	1.49%	1.37%
Cost	Control	-6.29%	2.54%	1.31%	-3.68%
	Treatment	0.00%	9.65%	9.46%	7.55%
Stirling-Alloa					
Time	Control	1.39%	1.25%	1.39%	1.29%
	Treatment	2.01%	2.13%	2.14%	2.19%
Cost	Control	8.51%	9.48%	9.14%	8.28%
	Treatment	15.17%	14.95%	15.05%	14.77%
Borders Rail					
Time	Control	6.59%	3.95%	13.01%	22.59%
	Treatment	17.58%	5.24%	13.31%	23.47%
Cost	Control	2.83%	1.78%	5.57%	9.30%
	Treatment	6.59%	2.30%	5.56%	9.27%

8.5 Comparison of model approaches

The TIEP report (DfT, 2014) found that techniques were often insufficiently context-specific and needed a clear narrative about likely economic impacts to inform modelling and quantification work, and the analytical work and empirical evidence employed. Here four different approaches were applied to the datasets each looking at the data in context and from a different angle.

8.5.1 Descriptive approach

In this current study, the descriptive approach was used to make a comparison between individual variables subject or not subject to treatment and so assess any impacts. However, this did not take into account the effect of other potential causal factors. It is accepted that this method does not identify the more complicated features underpinning property values as reflected in the literature where (Cervero and Landis, 1993), and Du and Mulley (2007a) were unable to discover any significant changes in property prices using this approach.

Accessibility issues

Travel to work by train shows a percentage improvement in the treatment group over the control group for the Robin Hood Line and Stirling-Alloa for the period spanning the intervention although this did not continue into 2001-2011 for the Robin Hood Line which may affected by the new Nottingham Tram System (Table). Distance to work breakdown was not available for the 1991 census, but for 2001-2011 it showed a decrease in distance travelled less than 10 km to work. There was a reduced percentage of those with no access to a car across all groups for all case study regions. In Table Borders Rail is excluded as only UK census figures prior to its opening (2011) were available so no comparison could be made.

Table 8-7 Comparison of accessibility attributes treatment v control

Case Study		Robin Hood Line				Stirling - Alloa	
Group		Treatment	Control	Treatment	Control	Treatment	Control
	Year	91-01	91-01	01-11	01-11	01-11	01-11
Travel to work	Home	4.07%	3.87%	-4.02%	-3.83%	4.35%	4.65%
	Train	1.11%	0.21%	-0.10%	0.17%	1.58%	0.54%
	Bus	-3.27%	-1.97%	-2.01%	-1.15%	-3.68%	-4.81%
	Car driver	4.03%	4.60%	6.58%	6.78%	6.08%	6.16%
	Tram	0.01%	0.02%	2.32%	0.41%	0.01%	-0.03%
Distance to work	Work mainly from home			0.62%	0.92%	8.36%	9.16%
	Less than 10 km			-2.56%	-2.04%	56.46%	48.13%
	10 km - 30 km			3.09%	2.74%	16.42%	22.17%
	30 km and over			1.81%	1.61%	8.32%	9.70%
Car Ownership	No cars or vans	-5.64%	-5.45%	-4.29%	-3.45%	-3.50%	-3.40%
	1 car/van	0.94%	0.53%	-0.62%	-1.50%	-2.70%	-1.60%
	2 cars/vans	3.71%	3.57%	3.33%	3.15%	4.20%	3.30%
	3 or more cars/vans	0.97%	1.36%	1.57%	1.79%	2.10%	1.70%

Property issues

Table 8-8 Property type changes - treatment v control

	Robin Hood Line				Stirling-Alloa			
Group	Detached	Semi-detached	Terraced	Flat etc.	Detached	Semi-detached	Terraced	Flat etc.
Treatment	3.09%	2.34%	-4.48%	0.07%	3.75%	-0.45%	-2.22%	-1.09%
Control	1.95%	1.30%	-3.54%	0.45%	2.77%	-0.25%	-1.36%	-1.15%

There was movement towards detached and semi-detached properties and away from terraced properties which was similar for both the Robin Hood Line and Stirling-Alloa and this was more noticeable in the treatment group areas compared to the control group areas. In Table Borders Rail is excluded as only UK census figures prior to its opening (2011) were available so again no comparison could be made.

Employment issues

In comparing employment across all the case study regions, it should be noted that the interventions span different years when there were changes in employment due to the recession of 2007 onwards.

Both the Robin Hood Line and Stirling-Alloa regions indicate a percentage increase in full-time employment in the respective treatment group areas compared to a percentage reduction in the control groups (Table).

For the Robin-Hood Line there was a percentage improvement in those with Level 1 qualifications and above, which was marginally greater in the treatment groups. This was not the same for Stirling-Alloa as the figures relate to 2001-2011 and include the recession period. However, the reduction in those with qualifications was less in the treatment group and this was offset by a percentage increase in those with Level 4 and above.

For the Robin Hood Line, both treatment and control group show an overall increase in residents with professional or associate professional and technical occupations, which is higher in the treatment group (10.82%) than the control group (9.89%).For Stirling-Alloa, the treatment group shows an overall increase in the percentage of residents with professional occupations against a decrease for the control group but the reverse is true for the associate professional and technical. Although skilled trades have increased for both groups, this is more marked for the control group (3.39%).

For the Robin Hood Line there is a sizeable move away from the traditional industries e.g. manufacturing and mining alongside a move towards hotels and catering and the construction industry which is echoed in both groups. In Stirling-Alloa, changes in types of industry over the period indicate a move away from traditional industries of manufacturing and mining, which is slightly higher in the treatment group (-5.05%) than the control group (-4.28%). There has been a similar move towards the public sector, e.g. public administration, health and social work and education, and this is generally more marked in the treatment group.

Table 8-9 Employment attributes comparison - treatment v control

	Robin Hood Line		Stirling-Alloa		Borders Rail	
	Treatment	Control	Treatment	Control	Treatment	Control
Full-time	0.18%	-0.44%	0.16%	-8.56%	56.52%	54.04%
Part-time	1.09%	2.15%	0.82%	-13.98%	21.01%	21.50%
Self-employed	0.81%	0.13%	0.35%	0.00%	11.50%	15.19%
Unemployed	-4.36%	-4.09%	-0.73%	0.00%	4.35%	3.09%
Full-time student	2.28%	2.26%	-0.59%	-26.09%	6.62%	6.18%
No qualifications	-53.54%	-53.20%	-7.43%	-6.61%	27.16%	28.10%
Level 1	15.45%	14.61%	-1.22%	-1.06%	25.74%	23.27%
Level 2	15.82%	14.53%	-0.12%	-0.71%	14.29%	13.98%
Level 3	5.67%	5.56%	3.46%	2.73%	9.06%	8.56%
Other	7.28%	7.36%				
Level 4 and above	9.31%	11.14%	5.30%	5.65%	23.75%	26.09%
Managers directors etc.	0.12%	-0.02%	-1.00%	-0.70%	8.75%	9.50%
Professional	1.71%	1.73%	0.43%	-0.15%	15.32%	15.48%
Associate professional and technical	4.45%	4.22%	-0.49%	0.78%	11.57%	10.73%
Administrative and secretarial	-2.02%	-1.65%	0.09%	-0.65%	12.33%	10.41%
Skilled trades	-10.01%	-9.33%	0.98%	3.39%	13.39%	16.05%
Caring leisure and other services	-1.77%	-1.54%	-1.17%	-0.33%	11.27%	10.70%
Sales and customer service	0.85%	1.15%	-0.30%	-0.28%	9.58%	8.04%
Process and machine operatives	0.91%	-0.09%	1.12%	-1.60%	6.92%	7.38%
Elementary occupations	5.76%	5.52%	0.34%	-0.46%	10.87%	11.70%
Agriculture hunting and forestry	0.20%	0.11%	-0.23%	-0.42%	1.66%	4.71%
Mining and quarrying	-1.36%	-2.92%	-0.25%	-0.54%	0.27%	0.24%
Manufacturing	-4.63%	-4.87%	-5.05%	-4.28%	5.81%	8.50%
Electricity Gas and Water Supply	-7.03%	-5.72%	0.88%	1.13%	0.70%	0.43%
Construction	0.98%	0.50%	0.48%	-0.51%	8.69%	8.99%
Wholesale and retail trade repairs	-1.65%	-1.13%	1.34%	0.62%	15.80%	14.72%
Hotels and restaurants	4.50%	4.84%	0.57%	0.56%	5.05%	5.66%
Transport storage and comms	1.59%	1.30%	-1.07%	-0.64%	4.53%	3.81%
Financial intermediaries	-4.05%	-3.96%	0.07%	-0.66%	6.74%	4.21%
Real estate renting and business activities	8.26%	8.66%	-8.48%	-7.97%	1.33%	1.60%
Public administration and defence	4.38%	4.44%	4.98%	3.60%	7.64%	5.97%
Education	5.62%	6.22%	0.32%	1.85%	8.13%	8.07%
Health and social work	11.39%	10.83%	3.28%	2.84%	16.58%	15.73%
Other	-17.70%	-17.81%	3.21%	4.43%	17.07%	17.36%

8.6 Difference-in-Difference

Difference-in-difference (DID) required data from the treatment and control groups measured at two or more different time periods, (at least one time period before "treatment" and one time period after 'treatment'), but the assumption of a common trend which was difficult to verify. Although pre-treatment data may indicate that trends are the same, there may be other policies changing at the same time or even vice versa.

Property impacts

There were contrasting findings for the case study regions reflecting their specific regional context. For the Robin Hood Line, there was a positive impact on house prices of being closer in distance to a rail station post intervention and also of being proportionately closer to a rail station post intervention. The Robin Hood Line has experienced a longer period since its rail intervention and so over the twenty years there has been a substantial increase in house prices attributable to all accessibility characteristics.

Table 8-10 Property Difference-in-Difference results - treatment v control

DID	Robin Hood Line			Stirling-Alloa			Borders Rail		
	Coefficient	t value		Coefficient	t value		Coefficient	t value	
Treated	0.101	2.305	*	-0.003	-0.030		0.004	0.035	
Distance nearest station	-0.014	-2.908	**	0.033	1.174		0.007	1.644	
Distance Ratio	0.157	3.036	**	-0.150	-0.518		-0.129	-0.605	
Job Accessibility Index	0.486	1.045		50.110	0.718		0.103	0.175	

However, for Stirling-Alloa, there appears to be no significant impact in the shorter intervention period studied for changes in distance to station. On the other hand, use of the job accessibility index suggests a positive impact of being proportionately closer to a rail station post intervention and consequently being closer to employment.

For Borders Rail, there are indications, though not significant, of a positive effect of treatment. However, as for Stirling-Alloa, the job accessibility index suggests a positive impact of being proportionately closer to a rail station post intervention, and thus nearer to employment.

Employment impacts

For the Robin Hood Line, there was a positive impact on employment density of being closer to a rail station post intervention and the effect of treatment is positive and statistically significant. There was also a positive impact of being proportionately closer to a rail station and an improvement in job accessibility post intervention. Also for Stirling-Alloa, a reduction in distance to the station and increased job accessibility had a positive impact on employment density over the intervention period. For Borders Rail, a reduction in distance to the station was not statistically significant at this stage possible due to the limited amount of data available.

Table 8-11 Employment Difference-in-Difference results - treatment v control

DID	Robin Hood			Stirling-Alloa			Borders Rail		
	Coefficient	t value		Coefficien t	t value		Coefficien t	t value	
Treated	0.068	2.094	*	0.065	3.180	**	0.981	1.732	
Distance nearest station	-0.010	-2.790	**	-0.006	-0.837		-0.025	-0.927	
Distance Ratio	0.080	2.086	*	0.159	2.402	*	0.727	0.589	
Job Accessibility	0.763	2.223	*	8.340	0.997		2.749	0.810	.
R^2	0.612			0.850			0.020		

8.6.1 Hedonic modelling - Fixed effects

An important assumption of the fixed effects model is that time-invariant characteristics are unique to the individual location (LSOA or Scottish datazone) and should not be correlated with other individual characteristics (Kohler and Kreuter, 2012). Results depend heavily on model specification, and the quality of the measures used in the independent "explanatory" variables is of key importance as the use of a proxy measure may result in inaccurate coefficients in the regression analyses.

For Borders Rail, the fixed effects model 2013 to 2017 may not detect any lasting house price impacts due to the rail intervention because of the limited amount of data available. The immediate benefit would be greater commuting potential to access the job market in Edinburgh. An additional factor was also considered - the time to reach Edinburgh - which would decrease over the intervention period and was thought to influence property prices as levels in Midlothian (nearer Edinburgh) are higher than those in the Scottish Borders (further away).

Property impacts

For both Robin Hood Line and Stirling-Alloa fixed effects models the distance to station and distance ratio are both significant along with other factors including a higher rate of employment, a greater amount of no car or one car ownership, and a higher percentage of residents with no qualifications. For instance, in Stirling-Alloa being 10% nearer a rail station predicts an increase in house price by £5520 on a £100,000 property. For Borders Rail, the distance to station and distance ratio are not significant but car ownership and time to reach Edinburgh are. For the Robin Hood Line, improvement in cost-based job accessibility generates an increase in property price and for Stirling-Alloa there is a consistently positive effect of increase in accessibility.

Table 8-12 Property fixed effects results - treatment v control

	Robin Hood Line			Stirling-Alloa			Borders Rail		
	Coefficient	t value		Coefficient	t value		Coefficient	t value	
Distance to nearest station	-0.0277	-8.11	***	-0.0534	-2.97	**	-0.00016	-0.22	
No car households	-0.0020	-4.45	***	-0.0074	-1.49				
1 car household	-0.0011	-2.15	*	0.0084	1.84	.			
% employed	2.1488	7.49	***	-3.9646	-2.21	*			
No qualifications	-0.0017	-4.78	***	-0.0022	-0.59				
R^2	0.6296			0.62778			0.00359		
Distance to station ratio	0.3273	9.10	***	0.5371	2.92	**	0.00614	0.18	
No car households	-0.0019	-4.33	***	-0.0076	-1.54				
1 car household	-0.0013	-2.61	**	0.0091	1.98	.			
% employed	2.1836	7.68	***	-3.8485	-2.11	*			
No qualifications	-0.0016	-4.51	***	0.0056	2.83	**			
R^2	0.6351			0.62644			0.00355		
Job accessibility	0.4988	1.74	***	61.6	3.44		0.06881	0.03	
Terrace	-0.0001	-0.28		0.0002	0.03	***			
No car households	-0.0025	-5.46	***	-0.0043	-0.81				
1 car household	-0.0012	-2.30	*	0.0074	1.53				
% employed	2.3016	7.78	***						
No qualifications	-0.0017	-4.70	***	-0.0048	-1.23				
R^2	0.6076			0.5702			0.00346		

8.6.2 Geographically weighted regression (GWR)

In the presence of spatial autocorrelation or spatial heterogeneity the hedonic model has some unexplained variance caused by interdependence between observations arising from their relative location in space. The residuals from the regression model are likely to be spatially correlated if the housing transactions are spatially clustered in some way; the estimated regression coefficients will be biased and therefore not suitable for predicting property prices (Basu and Thibodeau, 1998).

The GWR method developed by (Fotheringham et al., 2002; Brunsdon et al., 1998) explores spatial non-stationarity of a regression relationship for spatial data where a simple "global" model cannot explain the relationships between some sets of variables. Crespo et al. (2007) incorporated temporal data into the GWR model and developed a spatio-temporal version to forecast and interpolate local parameters through time. This has been applied in the present study by looking at the changes in coefficients for two separate years spanning the intervention for each case study.

One advantage of the GWR model as used in this study is the facility to examine the spatial variability of independent variables included as explanatory variables. Some independent variables appear non-significant in the global regression model, but vary significantly over the geographical region and are revealed as significant local parameters by the GWR model.

Property model

For the Robin Hood Line and Stirling-Alloa the negative relationship between price and distance to the nearest station varies across the region (Table). For the Robin Hood Line this is between -0.0076 and - 0.0101 in 1991. For Borders Rail, there is a negative relationship between price and distance to the nearest station for both 2011 and 2017 which varies across the region.

Table 8-13 Property GWR results by case study

Case Study	Year	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)
Robin Hood	1991	-0.0076	-0.0227	-0.0214	-0.0164	-0.0101	0.0064	-0.0169
	2001	-0.0067	0.0038	0.0101	0.0178	0.0601	0.0140	0.0051
Stirling Alloa	2001	-0.0251	-0.0227	-0.0214	-0.0164	-0.0101	0.0064	-0.0169
	2011	0.0146	0.0198	0.0232	0.0253	0.0326	0.0055	0.0257
Borders Rail	2011	-0.0090	-0.0055	-0.0041	-0.0022	-0.0003	0.0033	-0.0050
	2017	-0.0113	-0.0063	-0.0056	-0.0034	0.0053	0.0029	0.0002

Employment model

For the Robin Hood Line and Stirling-Alloa, the relationship between employment density and distance to the nearest station is universally negative (Table) reflecting the global cross-sectional model (the exception being Robin Hood Line in 1991 where 40% were positive). This shows that over the intervention period, the distance from the nearest station has had an increasing effect on employment density i.e. distance from the rail network is now more critical in terms of the job market. For Borders Rail, the model is run for 2011 only, and compares coefficients spatially across the data zones. Going across the 180 sample points, employment density also increases with decreasing distance to the nearest station reflecting the global model.

Table 8-14 Employment GWR results by case study

	Year	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)
Robin Hood	1991	-0.4455	-0.0947	-0.0230	0.0159	0.3288	0.1106	-0.0535
	2011	-0.3636	-0.2186	-0.1628	-0.1122	-0.0012	0.1064	-0.1304
Stirling-Alloa	2001	-0.2452	-0.2133	-0.1687	-0.1217	-0.0121	0.0916	-0.0457
	2011	-0.3854	-0.3191	-0.2569	-0.1605	-0.1247	0.1586	-0.1763
Borders Rail	2011	-0.0767	-0.0650	-0.0459	-0.0414	-0.0324	0.0236	-0.0533

Non- stationarity

Non-stationarity can simply be defined as processes that are not stationary and that have statistical properties that are deterministic functions of time. For the property model, the Robin Hood Line exhibited non-stationarity of relationships over the study region, but there was a lack of non-stationarity for the other case study regions. Again for the job model, using distance from the nearest station for the Robin Hood Line there was non-stationarity of relationships over the study region, but again there was a lack of non-stationarity for the other case study regions.

8.6.3 A summary of the models' strengths and weaknesses

Each modelling approach has been of some value in this study, especially when used in conjunction with the treatment and control group configurations derived previously. Each approach represents four different perspectives on the datasets.

The descriptive approach

The descriptive approach as applied here basically compares each contributory variable e.g. property price, accommodation type, educational level etc. pre- and post-intervention broken down by treatment and control group. This is a simplified approach which highlights any noticeable difference in the characteristics between the two groups. The output has indicated a discernible effect of treatment in some cases, and the exercise has been useful in suggesting variables to carry forward to the models. However, unlike the other methods, it looks at each variable in isolation, and does not take into account the effect that they may have on each other. Although it often demonstrated an impact on a particular variable, this was not corroborated by the other models. So rather than provide any evidence of causality, it yields some useful information on which characteristics need to be observed.

Difference-in-difference

Difference-in difference (DID) required the assumption of a common trend which was difficult to verify. Although pre-treatment data may indicate that trends are the same, there may be other policies changing at the same time. Thus it was probably more applicable to the two Scottish case studies where there were less confounding factors and measurement was effected over a shorter period. It did indicate some effect of treatment for the Robin Hood Line both on property prices and jobs, and although it was inconclusive for property price for the other case studies, it did predict an effect of treatment on jobs. Some of this may be due to different time periods since intervention between the Stirling-Alloa and Robin Hood lines. Difference-in-difference neutralised the fixed effects, but as such omitted a lot of relevant information on property prices which was taken up by the fixed effects model. It was also unable to identify spatial differences and spatial correlation across the regions which considered in the GWR model.

Fixed effects

The fixed effects model went into much more detail and considered property prices over many years before and after the intervention, and hence picked up movements not apparent in the difference-in-difference model. As such, it was successful in establishing some causality between the accessibility characteristics and property price and employment density for all the regions. However, in adopting a panel data approach there were problems of missing data or data not available, and there was the usual problem of deciding which explanatory variables to use. In addition, it was often difficult to match data over a long period of time as the criteria for describing individual data had changed as had specification of locations in datazones and LSOA areas. Like the DID model it was unable to identify spatial differences and spatial correlation.

Geographically weighted regression (GWR)

One advantage of the GWR model as used in this study is its facility to examine the spatial variability of independent variables included as explanatory variables which was not possible using the other methods. Some independent variables may vary significantly over the geographical region yet appear non-significant in the other models. So GWR showed that in some locations the relationship between variables is negative, and in others positive, and predicted non-stationarity of relationships in the Robin Hood Line region.

However, there are difficulties in using the model temporally as it is based on a cross-sectional model, where the issue of missing variables is a potential problem. Here the model was run for two periods either side of the intervention, so became effectively imitated a difference-in-difference model.

9 Chapter Nine: Conclusions and Summary

9.1 Introduction

This study has sought to develop a quantitative, scientific, methodological approach to apply in evaluating the impact of restoration of rail services, particularly relevant to previously remote or disconnected regions. With reference to three recent rail restorations that have reconnected previously isolated communities to the rail network, where the interventions had different aims and were at different evolutionary stages of implementation, the methodology has focused on property prices, employment and accessibility to jobs. The role of this final chapter is to pull everything together by summarising the research findings, comparing this with previous research, and going on to discuss the limitations of the study and suggestions for future research possibilities.

9.2 Summary of findings

The main thrust of this research was through investigating ex post situations both spatially and temporally to determine cause-effect relationships and relationships between measures, drivers and barriers. This required developing a methodological approach which would match those objectives and adapted pre-existing methods to develop a methodology for appraisal and establishing cause-effect relationships. The remit was to develop innovative approaches to evaluation methodology with the application of existing models and techniques in a different context, but particularly relevant to remote, rural or disconnected regions.

This section summarises the main points and findings of the research. Through consideration of appropriate measures and outcomes, the methodology has adopted innovative approaches to evaluate and appraise the socio-economic impact of rail transport restorations on the local community, specifically examining how this is reflected in jobs and property prices.

9.2.1 Methodological approach

A full account of the methodology is given in Chapter 4. In brief, the motivation for the research was triggered by a small survey of stakeholders just immediately prior to the Borders Rail re-opening. Three case study regions were then analysed quantitatively from established data sources using a mixture of modelling techniques. The results were subsequently synthesised in Chapter 8 and compared with relevant previous research literature.

Establishing a counterfactual

A counterfactual was developed which allowed a meaningful comparison between areas subject to treatment i.e. rail intervention, and those not treated i.e. either unaffected or minimally affected by the intervention, and to establish any differences between findings in urban studies. Consequently, treatment groups were selected based on distance thresholds where the control group was selected from remaining locations in the region. There appeared to be some benefit in application of clustering and propensity matching to effect a more balanced comparison between similar locations in the treatment and control groups. For all case study regions, propensity matching on area size and deprivation worked particularly well for a 2 km threshold, but was less effective at 4 km.

Accessibility characteristics

Three different measures of accessibility were applied; distance from the nearest rail station, distance ratio, and a job accessibility index to see how each measure translated into an impact or outcome. The change in distance to a rail station was effective in suggesting an impact, but distance ratio, which expressed a change in the relative nearness of the rail station, appeared more meaningful as it would project a more noticeable improvement to transport users.

Job accessibility

As the feasibility of commuting to jobs was seen as a key factor in assessing accessibility for this study, a third measure was developed, a job accessibility index applicable to remote regions. This had a dual purpose in being an accessibility model in its own right, which would reflect changes in accessibility to employment, and also an accessibility characteristic to be incorporated into an econometric model and calibrated to allow for the decay effect, commuting feasibility and skills matching.

The index reflected feedback received from the Borders Rail survey where in remote regions accessibility could be reduced according to when services were available and when people could access them, or where travel times exceeded some acceptable maximum threshold. The latter reflected the furthest that commuters were prepared to travel, with exceptions for those outside the case study regions where there was a powerful pull in terms of job opportunities. After analysing distance travelled to work and taking into account the rural context and nearby larger conurbations, a 75 minute threshold was deemed appropriate for this study. The cost of travel was recognised as a key factor affecting accessibility and generalised cost allowed the cost of commuting to be calculated using local values of speed and cost of transport.

The inclusion of a cost or time constraint and allowance for matching of jobs made it more relevant to the rural or semi-rural situation, and an index based on travel cost was chosen as it appeared that cost was more of an impediment to travel. Job accessibility was based on comparing the percentage skills share at each location, matched to actual jobs at all neighbouring destination locations. The job accessibility index allowed a measure of accessibility based on the original job market, and could also predict the effect of a slump in employment. The jobs basis for the skills matching element used, alternatively, the number of jobs available in the base (pre-intervention) year and the number of jobs available in the current (post-intervention) year.

Without job skills matching, job accessibility appeared to be overestimated as the seemingly high attraction of job opportunities may not always synchronise with the skills set in that location. It yielded good results when used as an accessibility characteristic in the hedonic models, being a more pertinent measure than distance from rail station as it encompassed the whole regional employment picture relative to each location.

Model approaches

From recommendations in the literature, different modelling approaches were applied to the datasets, each examining the data in context and from a different angle.

- The descriptive approach

The descriptive approach was seen not to identify the more complicated features underpinning property values, and the literature was unable to discover any significant changes in property prices using this approach (Cervero and Landis, 1993; Du and Mulley, 2007b).

For this study, comparison between individual variables subject to treatment did not take into account the combined effect of other explanatory factors. The descriptive approach looked at individual variables pre- and post-intervention broken down by treatment and control to assess any impacts, and therefore ignored the combined effect of other explanatory factors. The output indicated a discernible effect of treatment in some cases, and was useful in corroborating variables to carry forward to the models. However, unlike the other methods, by looking at each variable in isolation, it did not take into account their effect on each other.

- Difference-in-Difference

The difference-in-difference (DID) model is actually a type of fixed effects where differencing gets rid of the individual fixed effects, but as such omits a lot of relevant information on property prices (which was considered in the fixed effects model), and it was also unable to identify spatial differences and spatial correlation across the regions which was taken up with the GWR model.

For property impacts, the difference-in-difference approach produced contrasting findings for the case study regions, suggesting a beneficial effect on property prices of being closer in distance or proportionately closer to a rail station for the Robin Hood Line, and a suggestion of a positive effect for Borders Rail, whereas, for Stirling-Alloa, there appeared to be no significant impact. However, application of the job accessibility index suggested a favourable effect on prices of being proportionately closer to employment for all case studies.

For job impacts, there was a positive effect on employment density of being closer to a rail station, and a positive impact of improvement in job accessibility, but for Borders Rail this was not statistically significant at this stage possibly due to the limited amount of data available.

- Hedonic modelling with property fixed effects

The study applied hedonic price methods which were proposed as a popular approach in identifying transport investment impacts on land value (Weinberger 2001; Cervero and Duncan, 2002). However, results depended heavily on model specification and the quality of explanatory variables as the use of a proxy measure may result in inaccurate coefficients in the regression analyses. An important assumption of the fixed effects model was that time-invariant characteristics were unique to the individual and should not be correlated with other individual characteristics (Kohler and Kreuter, 2012). As the residuals from the model could be spatially correlated, the estimated regression coefficients would be biased and therefore not suitable for predicting property prices (Basu and Thibodeau, 1998).

For property impacts, the distance to station and distance ratio and improvement in job accessibility were all significant along with other factors for both Robin Hood Line and

Stirling-Alloa. For Borders Rail, unlike car ownership and time to reach Edinburgh, the distance to station and distance ratio were not significant. However, there were problems of missing data or data not available, and it was often difficult to match data over a long period of time, and inability to identify spatial differences and spatial correlation.

- Geographically weighted regression (GWR)

Few studies have been carried out on spatial variability in house prices and accessibility in the United Kingdom, mainly concentrated in London. However, Du and Mulley (2012) focussing on Tyne and Wear allowed for estimation of the importance of transport accessibility in determining house prices using GWR methodology (Fotheringham et al. 2002).

As the Borders Survey (Chapter 2) indicated spatial variability in transport isolation, it was decided to apply geographically weighted regression (GWR) as particularly pertinent to the rural or disconnected situation, to provide another perspective for this study. This explored spatial non-stationarity of a regression relationship for spatial data where a simple "global" model could not explain the relationships between some sets of variables.

Crespo et al. (2007) incorporated temporal data into the GWR model, developing a spatial-temporal version to forecast local parameters through time, and a modified version of this method was used here to estimate changes spatially over time. This estimated local parameters through time by examining changes in coefficients for two separate years spanning the intervention for each case study.

The property model showed variation across each region in the negative relationship between price and distance to the nearest station for both the Robin Hood Line and Stirling-Alloa. The Robin Hood Line exhibited non-stationarity of relationships over the study region, but no evidence of non-stationarity for the other case study regions. For Borders Rail, the model was run for 2017 only, but for that year the negative relationship between price and distance to the nearest station varied across the region.

For the jobs model, the relationship between employment density and distance to the nearest station was in most cases negative for both the Robin Hood Line and Stirling-Alloa, showing that the distance from the rail network was now more critical in terms of the job market. Again there was non-stationarity of relationships over the study region for the Robin Hood Line but not for the other case studies.

One advantage of GWR was its facility to examine the spatial variability of dependent and explanatory variables and demonstrate that they varied significantly over each geographical region whereas they often appeared non-significant in the other models. However, as it was based on a cross-sectional model, there were difficulties in using the model on a temporal basis due to missing variable bias. Accessibility impacts were still positive in being closer to a rail link, but in terms of job accessibility, while greater than 75% of the region had benefitted from better job accessibility there was spatial variation with pockets where there had been no impact.

This methodology seeks to capture the wider economy impacts of rail infrastructure projects, using econometric and modelling techniques to provide evidence of land use change and its effects, and new modelling and valuation approaches to supplement standard appraisal methods. Arguably LUTI and SCGE models cover a much wider scope, but they are very complex and require specialist expertise to develop and have substantial data requirements. This thesis goes much deeper within the areas it covers, so for those particularly interested in the impacts measured here in more detail than that provided by a LUTI or SCGE model, it offers an alternative or supplementary method.

9.2.2 Research Hypotheses

The original research hypotheses posed in Chapter 1 are restated and a summary given in each case as to how this has been addressed as a result of applying the methodology.

- *Initial insight into the current situation regarding disconnected and isolated communities and through reference to appropriate case studies, to categorise relevant characteristics of the affected regions.*

Context has been important for this study as most literature is US-based and has concentrated on urban and city rail systems as offering the best opportunities to test for property market and job effects. This research study has particularly focussed on remote and sparsely populated regions which were previously disconnected from the rail network, where impacts are often of a lesser magnitude than the urban setting.

It has compared three case study regions which meet the "disconnected" criteria, considering relevant characteristics of the affected regions, such as the location and population characteristics, accessibility to public transport, alternative travel modes and the nature and frequency of the rail service. The interventions had different aims and were at different evolutionary stages of implementation with specific regional, property and employment characteristics. An initial qualitative survey of Borders Rail just prior to its reopening canvassed opinions of local stakeholders on the expected impact, and this highlighted inherent problems in rail-disconnected regions, which were subsequently taken forward for consideration for the whole study.

The scale of line re-opening was a prime factor in discerning any significant impacts, and not surprisingly, where rail links had been in place for varying lengths of time, wider impacts were more apparent than in more recent interventions. This was evident in the findings which showed a variable impact on house prices, jobs and commuting patterns and this was more noticeable in the Robin Hood Line case study where the extensive restoration had been in place for over 20 years. An improvement in proximity to the rail network was shown to translate into an increase in house prices for all models. However, this was not reflected for Stirling-Alloa where, although the fixed effects models estimated an impact, the difference-in-difference model showed no discernible difference between treatment and control groups.

The difference in impact over the intervention period for Stirling-Alloa and Robin Hood Lines may be explained by context as they represent dissimilar transport investments, and residential property prices might be depressed in the immediate vicinity of the station in the case of Alloa. The level of transport provision offered a further contrast as the Stirling-Alloa and Robin Hood Lines had nearby rail stations in place prior to the intervention, which although not necessarily offering a viable alternative for commuting, were close enough to access via other transport modes. There were also plentiful alternative transport modes available, especially in Stirling-Alloa, and in the case of the Robin Hood Line, the development of other provision with the tramway system which impacted on those locations closer to the centre of Nottingham.

There are two other important factors concerning property prices and employment. Firstly, areas where there had been seemingly little impact on property prices were often locations where there was a stagnant housing market. These comprised a high proportion of terraced properties, which did not necessarily attract new rail commuters who may prefer detached or semi-detached property further away from the station. In addition, some locations had a "thin" job market with limited skill sets, hence the job accessibility did not improve to any great extent post-intervention which reflected on property prices.

- *Improvements in rail transport infrastructure to previously disconnected communities produce measurable net economic benefits*

This research question looked in particular at property impacts as an indicator of a net economic benefit, on the basis that an increase in house prices raised household wealth and increased the collateral available to homeowners, thereby facilitating higher consumption. Property prices act as a proxy for the state of the local economy, and hence any improvement may be considered as a measurable economic benefit. From the case study approach it was apparent that context was important in examining rural and disconnected communities as against purely urban regions. The constituency of the region was more crucial, and different levels of transport investments impacted differently in locations with similar economic conditions. The housing profile for each of the three case study regions has a similar distribution shape with a small number of very costly properties, a central group of typically-priced houses and a long tail of relatively expensive ones.

The current study found that developments in transport infrastructure caused changes in accessibility that could have an effect on property prices, and these were measurable using econometric modelling techniques by considering different measures of accessibility. Property benefits depended to some extent on the closeness of neighbourhoods to new stations, and the impact of railway stations on property value diminished with distance from the station. This is also indicated in the consistently positive effect of increase due to improvement in accessibility to jobs for all three case study regions.

For the Robin Hood Line and Stirling-Alloa case studies there was a reduction in percentage of terraced properties over the intervention period, and this was greater for the treatment group areas in Stirling-Alloa, and may impact on the model results. The property type share influenced average property price, and a higher number of terraced properties would generally result in lower average house prices, which is greater for the treatment group areas in Stirling-Alloa. This may be driven by demand in those areas in movement towards detached and semi-detached properties.

For the long-established Robin Hood Line, the substantial increase in house prices was found in part attributable to improvements in rail infrastructure, as against the more recent intervention, Stirling-Alloa, where there appeared to be contrasting evidence of impact depending on the models applied. For the latter, distances and changes in distance were much less than for the other case studies, and so there may have been less incentive for residents to view the rail link as particularly beneficial. In addition, there was extensive availability of good alternative transport options in the area which would reduce the impact of the rail station, except for those commuting or shopping in Glasgow.

However, although only a recent intervention, there were early indications that the introduction of Borders Rail had had a positive effect on house prices in the treatment areas. The immediately visible benefit was greater commuting potential to the job market in Edinburgh and an additional factor - the improved time to reach Edinburgh - was thought to influence house prices as levels in Midlothian (nearer Edinburgh) are higher than those in the Scottish Borders (further away).

Often changes will take place in property values in advance of completion of the transport investment in the expectation of improvements in the transport infrastructure, the so-called "announcement effect". Limitations in data availability i.e. less detailed information on house prices prior to 1995 prevented detection of such an effect for the Robin Hood Line, but for Stirling-Alloa and Borders Rail there was some evidence of this in the treatment areas. For Stirling-Alloa, there is an indication of a marginally higher increase, echoed in movement of minimum prices in the immediate period spanning either side of the intervention, with the maximum price higher in the treatment group dropping below that of the control group after 2011. For Borders Rail, there was an increase in house prices for both treatment and control groups, suggesting an impact on those locations nearer the new rail link prior to its reopening - reflecting an announcement effect in expectancy of improved accessibility.

There could also be an associated lag, especially in areas where house transactions are infrequent. In the case of Stirling-Alloa - a much smaller intervention than the others - there was no noticeable improvement beyond the immediate intervention period, suggesting that rail interventions have a very limited effect on the property prices where the market was generally stagnant, or there was good availability of alternative transport options in the area. A confounding factor may be that the Stirling-Alloa line reopening coincided with the beginning of the nationwide slump in house prices although this is unlikely to show an effect as it is controlled for within the models.

- *The re-establishment of a rail link between previously disconnected regions and a larger conurbation results in measurable net employment impacts.*

Being closer to a rail station post intervention appears to have an increasing effect on employment density for both the Robin Hood Line and Stirling-Alloa lines. There is also evidence that this effect is produced by increased job accessibility. For Borders Rail, a reduction in distance to the rail station accompanies an increase in employment density, although this is not statistically significant, and cannot be ratified at this stage due to the limited amount of data available. The impact on employment density increases marginally where there is a higher population density. For all case study regions, results suggested that an important influence was improvement in access to higher numbers of suitable jobs in neighbouring locations. For Borders Rail, the improvement in time to Edinburgh was also very significant as employment density was found to decrease with increased time to Edinburgh.

For all case study regions, rail usage had grown more quickly than for pre-existing stations nearby. This included an increasing use of season tickets, especially from main centres such as Galashiels and Mansfield, although a direct causal connection could not be established between season tickets and commuting to jobs, an increase in season tickets did correlate with a greater level of commuting. A confounding factor for the Robin Hood Line was the introduction of trams in Nottingham in 2004 which impacted on treatment areas nearer to Nottingham, and detrimentally affected the use of "competitor" stations near Nottingham.

Following on from the Robin Hood Line intervention (2001 to 2011), although full time employment decreased across the region, this was less in the treatment areas, and this is equally true for Stirling-Alloa. Both regions underwent a change in job profile, as evident in the sizeable move away from the traditional industries, with an overall increase in the percentage of residents with professional occupations in the treatment group i.e. those now nearer the rail network.

For Borders Rail, at the last census there were proportionally less residents with managerial and professional occupations and skilled trades, against a higher proportion in administrative, care and leisure services and sales occupations in the treatment group compared to the control group. It is too early at this stage to gauge a shift in job profile into treatment and control groups. However, post-intervention over the whole region there is a noticeable percentage increase in professional occupations against a decrease in administrative, care and leisure and skilled trades.

There was greater use of rail and tram in the Robin Hood Line treatment group, and more commuting to distances of over 10 km and a greater increase in rail mode for the Stirling-Alloa case study. The geographical nature of the Stirling-Alloa and Borders Rail region means there is a concentration of population and employment around Alloa or Galashiels (typically treatment group) and residents in more remote places (typically control group) must travel further to work and are more likely to work from home. Any causal economic impact would also be reflected in the transport market.

- *Even with new rail improvements, spatial, temporal and economic barriers may prevent more economically vulnerable neighbourhoods from receiving the full benefit*

The specified criteria relate to spatial elements (e.g. accessibility to the city and public transport), temporal elements (e.g. public transport availability) and activities (e.g. employment) accessibility. For the present study, the results from the accessibility-based measures based on employment and essential services relate to cost and time for travel, distance thresholds and the timing of services and measure accessibility based on both the original and more recent job markets. Both measures indicate that although average job accessibility has improved across each region there are locations where, because of problems of connection and timing they have not received the full benefit from the intervention. In addressing spatial and temporal impacts and heterogeneity within the region, there was variability of transport accessibility and land value over space. Across all case study regions, there was an overall increase in rail mode accessibility, and whether using a travel time or cost basis with skills matching, the effect of being closer to the rail network was illustrated in a greater improvement in job accessibility.

However, the greater relative narrowing of index range in the treatment group and the relatively wider difference when the index is based on travel cost rather than travel time suggest that cost is more of a deterrent impediment, and may prevent more economically vulnerable neighbourhoods from receiving the full benefit of the restoration. An important factor in determining the impact of an intervention is job reachability reflecting a threshold beyond which travel to jobs was not feasible because of the time and cost constraints. This will vary across a region taking into account spatial and temporal barriers. Some contributory factors were non-significant when looked at overall, but varied significantly over the geographical area and were revealed as significant local parameters.

Job skills matching, synchronised with the skills set in each location, and the inclusion of a cost or time constraint and allowance for matching of jobs, made it more pertinent to the rural or semi-rural situation - often in economically vulnerable neighbourhoods. Hence, those on a lower income may be at a disadvantage both financially and through lack of suitable skills. It was also evident that job accessibility without skills matching assumes that all job opportunities are appropriate for each location, and in more remote regions where there is a "thin" job market a more realistic measure would take into account the skills set at that location as representing the level of attraction. Applying the different variations of job accessibility index indicated that using cost-based job skills matching worked best as an accessibility comparator.

9.3 Comparison with previous research

The current study addresses the gap in knowledge reflected in a clear shortage of "in depth" academic research into the consideration of changes in accessibility and the wider economy. This specifically relates to more isolated or rural regions at a distance from urban centres of population, or urban and semi-rural areas subject to poor transport infrastructure.

Ex post evaluation of interventions provided evidence of impacts after a certain period of time being based directly on the outcomes of past decisions (Worsley, 2014). The intervention was seldom the only change taking place within a particular region and when any observed and time invariant factors were controlled for, any remaining differences in price could be chiefly attributed to changes in rail accessibility. The main contribution of this study has been to increase the knowledge about ex post appraisal methodology by applying and comparing previously published methods which have rarely been considered together, and certainly not in the context of remote or sparsely populated or regions. This has often involved combining two or more ideas to reveal something new and relevant, for instance, applying methods such as Cluster Analysis and Propensity Matching to address selection into treatment issues. This has included utilising models which addressed the spatial and temporal impacts and heterogeneity within the region, in particular applying Geographically Weighted Regression (GWR) to produce a local model which allowed for different relationships at different points in space.

The development of a job accessibility index applicable to remote regions capable of incorporation into hedonic models has addressed the inherent problems in such regions in terms of skills mismatch and commuting feasibility, as well as different transport impact thresholds. By projecting benefits for the local community through accessibility to jobs and services, the study has particularly explored the wider implications of rail interventions in linking remoter locations to larger towns and cities.

Despite the limited amount of research in this area, the study has unearthed certain common elements which are contributory factors in each intervention, suggesting prospective transferability of those elements to other situations and potential for restoration of rail links in other regions as suggested by (Campaign for Better Transport, 2012).

This research has indicated a number of potential quality of life benefits arising from the rail intervention in more remote areas. Unlike previous research, rather than simply taking trip numbers as a proxy for the success of an intervention, it has suggested wider benefits may depend on the extent of the intervention and the specific regional context, thus bridging a gap in current understanding between the current aggregate approaches which was relatively devoid of context.

Context

Unlike a significant proportion of the literature which is either US-based or addresses almost exclusively a city or urban context (Gibbons and Machin, 2005; Oosterhaven and Elhorst, 2003; Vickerman, 2007; Lakshmanan, 2010; Tavasszy et al., 2002), this research recognises that rural areas are different, presenting a thin labour market where there are limited job opportunities and often a lack alternatives and choices for travel and employment (Findeis and Jenson, 1998; Vera-Toscano et al., 2004). Unemployment is more pronounced and job search costs are often the cause of market failure in remote labour markets (Laird and Mackie, 2009). This research has attempted to assess the rural or semi-rural perspective as suggested by Laird et al. (2013), in developing a quantitative, scientific methodological approach to the measurement of wider economic and accessibility benefits through reference to three recent rail restorations that have reconnected previously isolated communities to the rail network.

The counterfactual

This study has recognised the importance of checking causal relationships of the project (Department for Business Innovation and Skills, 2011), and that a strong evaluation should distinguish the effect of a project from all other potential influences to establish causality through the counterfactual. It has thus taken the approach suggested in What Works-Evidence Review (What Works Centre for Local Economic Growth, 2015) of comparing outcomes between treatment and control groups.

A key issue was the "selection into treatment" problem - establishing control sites by identifying neighbouring areas which have not been exposed to the intervention but may be affected due to spill over effects from the project. Establishing control sites had often proved problematic due to identifying areas not exposed to similar interventions, as indicated in Blainey and Preston (2010) who employed control groups to establish the impact on employment and property prices.

This current study applied a rigorous approach. through matching of zones of comparable accessibility, subject or not subject to rail infrastructure changes (What Works Centre for Local Economic Growth, 2015), and treatment groups were selected based on distance thresholds, where the control group was selected from remaining locations in the region (Gibbons and Machin, 2005). However, it took this further by ensuring that treatment and control groups were comparable in terms of observables, through additional application of propensity score matching and clustering techniques.

Accessibility measures

Previous research suggested a causal link between property prices and accessibility (Henneberry, 1998), and immediate locations were expected to produce higher effects than locations further away (Bowes and Ihlanfeldt, 2001), which was often translated in the literature to refer to distance as the accessibility characteristic. Alternatively, distance to the Central Business District (CBD) was seen as an important factor as proximity to transport facilities increased speed of travel to the CBD, which translated into the value of the property. On that basis, this study has considered distance as an accessibility measure, but also explored two other measures, distance to station ratio and job accessibility index as perhaps more meaningful contributors to property prices and employment levels.

Job accessibility index

There was an extensive literature on job accessibility mainly applicable to urban environments using gravity-based measures (Hansen, 1959) and a distance decay function (Geurs and Ritsema van Eck, 2003; Cervero et al., 1998). This often proved difficult to calibrate because of historical data requirements and the difficulty in modelling human behaviour (Fotheringham, 1981), and was dependent on the condition of the transport system, traffic congestion and unwillingness to travel (Harris, 2001; Reggiani et al., 2011). The present study also found this problematic, especially in a rural situation where workers often applied for jobs outside their own administrative boundaries (Rogers, 1997).

Considering employment as the most likely destination type for an accessibility measure (Horner and Mefford, 2005), the feasibility of commuting was seen as a key factor in assessing accessibility for this study because of the potentially longer commuting distances, accepting the significance of travel time in house location choice (Zondag and Pieters, 2005). For this study, there was often a mismatch between when services are available and people could access them, as intimated in DfT (2005). Travel times were expected not to exceed some acceptable maximum threshold, but taking into account the rural context and adjacent larger conurbations, a 75 minute threshold was shown as more applicable to this study compared to the 60 minutes suggested in Korsu and Wenglenski (2010). This is borne out by analysis of travel to work figures which show that a significant proportion of workers travel at least 100 km to work in all the case study regions.

Affordability was also recognised as a contributory factor affecting accessibility (Social Exclusion Unit, 2003), but the monetary "worth" for the journey-to-work trip (Handy and Niemeier, 1997) made no allowance for competition for jobs within the employment market, and could underestimate the impacts of accessibility change (Van Wee et al., 2001). Accessibility to reachable jobs should depend on those workers claiming to form a match (Weibull, 1976; Ihlanfeldt, 1993; Harris, 2001; Van Wee et al., 2001; Kawabata and Shen, 2007). This study accepted that failing to fully estimate job availability could overestimate the job accessibility levels of poorer areas such as those under consideration (Bunel and Tovar, 2014), and consequently has allowed for job availability in combination with use of average speed and generalised journey costs (Balcombe et al., 2004) in estimating the total journey cost. The current study also took account of "occupational match" as introduced in Cervero et al. (1995). Little measurement of job accessibility has appropriately incorporated this diversity element, but it is especially important in the context of this current study where job skills are often limited, and it is important that consideration is given to the suitability of job opportunities to the local job market.

The present study would have preferred to use a measure of vacancies and job seekers as a basis (Korsu and Wenglenski, 2010) rather than all existing jobs as suggested in Ihlanfeldt and Sjoquist (1998), but in common with the latter, data availability issues restricted job availability to consider occupied jobs and active workers. The present study adopted the use of pre-intervention figures as suggested in Gibbons et al. (2012). This allowed for an estimate of improved accessibility post-intervention relative to the jobs profile that existed pre-intervention, whereas other studies had used a post-intervention basis. However, similar to the methods used there, in the accessibility calculation, skills matching in this study used the percentage skills share at each origin location, which was then matched to actual jobs at all the other destination locations. Although there is still the potential for endogeneity, this offers a more robust measure of accessibility in using pre-intervention job figures and pre and post intervention costs and times to isolate the improvement in accessibility of the transport intervention, rather than of the sorting effect of labour movements.

Property impacts

Although Attanasio et al (2011) suggested that changes in accessibility through transport infrastructure developments could have economic effects enjoyed by those located at its proximity, measured in particular by the effect on property prices, there were contradictory findings in the literature where there was a large variation in estimates (Efthymiou and Antoniou, 2013; Mohammad et al., 2013). However, the majority of studies reported a positive effect with negative impacts at locations very close to stations or railway lines where noise, pollution and crime levels were higher.

Many examples in the literature applying hedonic methods to house prices to assess the impact of transport interventions focused mainly on an urban environment and adopted different accessibility characteristics (Forrest et al., 1995; Henneberry, 1998). They showed different effects between urban and rural housing markets (Visser et al., 2008), and dissimilarity between model results dependent on the urbanisation level of the metropolitan area (Bowes and Ihlanfeldt, 2001). By contrast, the present study has focused on a remote, rural or disconnected context, but concurred with some of the findings that related to the housing market and suggested different impacts for different case study regions. These findings agree with Mayor et al. (2008) that there was a reduced impact depending on the availability of alternative transport options in the area (which were quite extensive in the Stirling-Alloa region).

Another consideration, the "announcement effect" - anticipation of a forthcoming rail opening with an effect on prices - was particularly relevant to Borders Rail as its coming was heralded for several years prior to the actual reopening. Although there was evidence in this case of an opening date effect on property prices in treatment areas, it was not possible to confirm causality, as suggested in the literature (Forrest et al., 1995; Gatzlaff and Smith, 1993).

Employment impacts

There was much discussion in the literature on the importance of context, and that improvements in transport infrastructure may reduce commuting costs and bring increasing employment and lower wages and provide easier access to places with jobs (Gibbons and Machin, 2003). Measuring impacts due to improvements in accessibility on employment through job accessibility was again based on the urban situation. However, more recently there was recognition that rural areas are different with often a lack of alternatives and choices for travel, employment and suppliers (Laird et al., 2013).

9.4 Limitations of the study

There are aspects of this study where findings were limited and leave questions less than securely answered.

Data collection

Some of these relate to data collection and data quality. For example, there was a lack of property transaction data for the period prior to the intervention for the Robin Hood Line, which was not available until 1995, just prior to the intervention. Whilst this research provides evidence to suggest that more commuters are using the train, it therefore remains challenging to establish whether this has affected the price of property around the time of the intervention.

Data definitions

Another data problem was the change in definition of LSOA and Scottish data zones for different UK censuses which made it problematic for matching and hence making comparisons over the intervention period. In addition, definitions of data type e.g. education levels, were not always comparable. For example, Level 4 education was interpreted differently for different years, and so required adjustments to account for this. Also not all data was published annually, so there were consequently gaps in the data which had to be filled in by extrapolation methods.

Scope of dataset

Because of limitations on accumulation of data, the dataset for each region was of finite size and therefore possibly influential locations external to the dataset were either omitted or only partially included. This may have had some minor impact on the models, but certainly limited the scope for selection into treatment groups. As a result, the control group included some locations where there had been some change since the intervention, albeit of a very small magnitude.

Treatment and control groups

Application of propensity matching generated a treatment and control group combination with a similar socio-economic profile. Although the comparisons showed an improvement post-intervention for the treatment group, there were no hypothesis tests to establish whether the treatment and control results were significantly different. This would be a suitable and worthwhile extension to be included in future research.

Explanatory variables

As the results depended heavily on model specification, the quality of the measures used as explanatory variables are of key importance and may result in inaccurate coefficients in the model outputs. Here instrumental variables have not been used to deal with the endogeneity issue, but the use of various statistical tests has filtered out variables that exhibited some correlation. Results from the models suggested that there may be additional variables to take into account, possibly at a more disaggregated level.

For instance, because property attributes contribute greatly to the price of a house, a simple breakdown into types of property does not take into account its age or condition, or whether it has a garage etc. These factors have not been included here due to data sourcing limitations, and they may also have some correlation with the explanatory variables used. Also, the GWR models may also be susceptible to omitted variable bias as they are basically cross sectional, although here two separate years have been run allowing a variant on the difference-in-difference approach.

Assumptions in accessibility index

There was limited sampling of times and costs for each region, and a more detailed analysis of routes and fares may have produced more accurate results. Furthermore, the estimation of time and costs made some assumptions on multi-mode travel to access the rail network which did not take into account the time of day or day of the week, and also the level of congestion and typical delays experienced. The lack of vacancy data meant that jobs occupied rather than vacancies were used in calculations.

The level of aggregation

The level of aggregation used in the data did not allow for further exploration of the datasets in terms of age, gender, disability etc. which would have produced more findings in terms of the distributive aspect of the interventions.

9.5 Future Research and recommendations

This section suggests the potential for future research based on the findings to date, for example, what sort of study should be contemplated now, and what issues would it resolve? From the work carried out in this research and the conclusions drawn there are several recommendations that can be made relating to appraisal of disconnected remote and areas of future work where issues have arisen. These relate to expansion of the more simplified remit shown here to develop a methodology which will go deeper and adopt a more detailed approach to economic aspects.

Distributional effects

The evidence suggests that although the introduction of rail to remote areas is potentially an effective mechanism for preventing the onset of isolation and social exclusion, rail accessibility varies across each region, and so if residents do not have reasonable access to rail then the intervention is ineffective. Hence the study could be expanded to distinguish impacts identified in this research across the wider population. This would require further research into the distributional effects of age, gender, ethnicity, social layers and educational skills. For age-related concerns, it could include the influence of the bus pass as a disincentive to use rail, and there should be further research into the impact and effectiveness of rail cards, and implications for those with disabilities.

The research recommends further consideration of the wider social benefits of the rail intervention as against evaluation simply in terms of quantitative use of the network. Given that the train provides a social environment, it may provide benefits to various social groups who would not necessarily use it, and so more understanding is required as to how rail users are responding to the opportunity it presents.

Tracking travel mode transition from bus and car to rail

This research has found many differing responses to the introduction of rail infrastructure with a suggestion of some success in motivating a modal shift, with those who would have commuted by bus and car in the absence of the intervention most likely to respond to it. Some rail users might be walking or catching the bus, or choose to drive to the rail station.

In order to strengthen understanding of the transition from other modes, a travel diary or panel survey could track the changeover to becoming a rail user - maybe making use of a railcard or season ticket - and how this may affect their usage of bus or car. It would also trace the types of activities undertaken, and address the need for greater consideration in determining the reason for the rail trip.

Incorporation of other factors potentially affecting wider economic impacts

When appraising property and job impacts this study has not examined in detail the societal and income characteristics of the markets, and how cost of rail fares effects personal budgets along with prices, wages and rents. Further expansion needs to consider other elements of appraisal including revenue, cost, congestion and environment.

Regeneration considerations

The Borders Rail survey brought up a recurrent theme relating to re-introduction of a rail link in that, on its own, it was thought to have limited wider impact unless there were more regional initiatives to act as a catalyst to regenerate the region. Suitable research may consider both what that regeneration should comprise and also compare that with what has been done previously in reconnected regions. Such research could also assess the extent to which those in possession of a bus pass could be attracted onto rail despite being required to pay.

In addition, one aspect of regeneration which this research has touched on is the transition in type of industry, job skills and educational levels after the intervention. However, this would benefit from further disaggregation so that underlying movements in local job profiles can be detected and the causal links to changes in transport infrastructure can be determined.

9.6 Implications for practice

Use within or alongside a WebTAG-style appraisal/evaluation

TAG UNIT M5.3 *Supplementary Economic Modelling* provides high-level guidance to inform estimation and review of "Supplementary Economic Models" - as non-standard methods to estimate the economic impact of transport schemes. Where Supplementary Economic Modelling can be justified and credible analysis produced, these should be reported. The view of DfT is that there is no single best approach to capture all the economic impacts of transport improvements. Rather, different methods may be applicable to different contexts depending on the scheme's anticipated impacts.

This methodology can assess how transport schemes impact on the spatial distribution of the economy and should be used to supplement rather than replace conventional appraisal methods set out in TAG Units A1 and A2. It can be used to provide a broader understanding of impacts not captured by standard approaches; as well as an appreciation of a range of potential future scenarios. This methodology could be used by consultants and project managers to inform the scoping, undertaking and reporting of *Supplementary Economic Modelling*, both for individual transport schemes and packages of schemes.

An example of this relates to WebTag unit *A2.3 Employment Effects*. If employment effects are to be analysed, this methodology provides additional information for use in the Economic Narrative for Employment Effects within the context of the transport investment. This includes justification of the expected employment effects:

What evidence is there that transport acts as a barrier to employment? For example, poor connections to centres of employment (including low frequency of public transport, inconvenient timetabling, as well as no physical links); High transport costs relative to income. UK census provides data on household access to private vehicles, which can be used to help identify transport barriers to employment.

What evidence is there that requiring employment have the skills required by firms? Even once transport barriers are removed, in order for individuals to gain employment they need to be able to take advantage of the employment opportunities.

Are the expected employment effects fully captured by user-benefits and how are the employment effects to be quantified and valued?

Thus this methodology may provide additional information to:

- Evaluate user-benefits for schemes impacting the economy spatially.
- Capture a broader range of wider economic impacts than provided in the A2 Units of TAG.
- Obtain context-specific estimates of welfare impacts set out in A2
- Estimate sub-national impacts, such as changes in local employment.

How do the results relate to time savings (or user benefits) in appraisal/evaluation?

The direct effects of transport investment are to reduce transport time and costs through reducing travel times, decreasing the operating costs of transport and enhancing access to destinations within the network. In WebTag A1.3 User Benefits, the calculation of transport user benefits is based on the conventional consumer surplus theory where consumer surplus is defined as the benefit which a consumer enjoys, in excess of the costs which he or she perceives. The surplus associated with making a journey will not be the same for everybody and depends on the benefit each individual derives from making that journey. A value of time savings is required to convert the forecast changes in travel time resulting from an intervention into monetary values that can be used in appraisal. The TAG Data Book contains values of travel time savings for working and non-working time that should be used in most economic appraisals of transport projects.

User benefits in appraisal and evaluation relate to a net connectivity benefit to business and commuters through application of generalised transport cost. This is reflected in the results of this study by consideration of the change in distance to access employment and essential services and converted using cost and speed of transport to indicate an improvement in terms of time and cost savings. This study also takes into account the context of the areas affected and goes into more detail as to how these benefits vary across each region where increased proximity may differ considerable spatially. Through disaggregation and attribution of user benefits, the question of who benefits from the scheme can be assessed.

For instance, on the supply side effects of labour there be local economic impacts over and above those captured by user benefits. The expected employment effects will not be fully captured by user-benefits as they do not take into account the spatial distribution of jobs, and barriers such as transport frequency. Also in consideration of property effects the user benefits may impact on the price of houses and provide an economic benefit which will vary across an affected region.

Contribution to making the case for new or re-opened rail infrastructure

This methodology could be applied in addition to conventional appraisal methods in that it goes into more detail in the specific areas of accessibility, property and jobs than is possible within the normal WebTAG or STAG framework. It does not seek to replace those methods, but rather provide supplementary evidence of the need for rail infrastructure. It does so by reference to similar recent re-openings of rail lines and analysis of the common factors relevant to all contexts. It particularly it will contribute by focusing on:

- Job accessibility: by studying the unique employment profile of a region in relation to job skills, and the impact that a rail intervention might have taking into account the existing transport situation.
- Accessibility to essential services: by forecasting the improvement or otherwise that rail may bring across the socio-economic profile of the region.
- Land value uplift: predicting how an intervention might impact on house prices as an indicator of the state of the local economy.

In particular, by looking at the spatial distribution of the region, to predict the expected variation for different parts of the region to assess whether the extent to which the benefits will be felt in terms of job accessibility, property values and employment.

New information on the spatial pattern of employment and property

In the context of this study, some regions have experienced great industrial decline, and consequently job skills may not match those required in the city where there is a greater range of job types. The study has addressed temporal and spatial and temporal impacts and heterogeneity through application of hedonic modelling and Geographically Weighted Regression (GWR). The latter has allowed analysis of the different relationships which exist at different points in space and time. Both produced results which provide new information on the spatial pattern of employment and property impacts, and indicate that there can be significant variation across the region.

The study has confirmed areas impacted by the intervention where there have been changes to the cost of travel. In some cases, there was a much wider area of coverage than the transport intervention itself and the likely area in which user costs change. The public transport corridors likely to be affected by the rail intervention have been identified, and new railway stations have included the station's wider catchment area. For the hedonic models, the aggregation level has been at LSOA or datazone level, but for GWR, the aggregation level varies using the "moving window" approach so this has produced new information on spatial patterns. The application of clustering techniques and propensity testing have also generated new information matching of similar socio-economic zones.

The appraisal of accessibility focuses on the public transport accessibility aspect of accessing employment and services, and considers the accessibility needs of different groups of people, taking into a wide range of factors, including journey times to reach key destinations and service frequencies e.g. is the service every 10 minutes or more during peak hours?

Application in other types of transport re-openings

There is potential for refining the methodology used in this thesis for accessibility, property price and employment impacts for application in wider situations than just rail re-openings. Although the emphasis here has been on changes in accessibility through a rail intervention, and the case studies have particularly looked at regions subject to such an intervention, the methodology already incorporates bus and car mode for the job accessibility element, and could be extended to include light rail and tram interventions.

So, for example, an ex-post appraisal using this methodology could evaluate a new faster bus service, an improvement in road infrastructure such as a new motorway, or the introduction of a new tram service. The latter would be easier to incorporate as it involves a relatively small network similar to heavy rail and there would be fewer confounding factors, whereas a new motorway would also have competition from the wide number of routes available to the more complex road network.

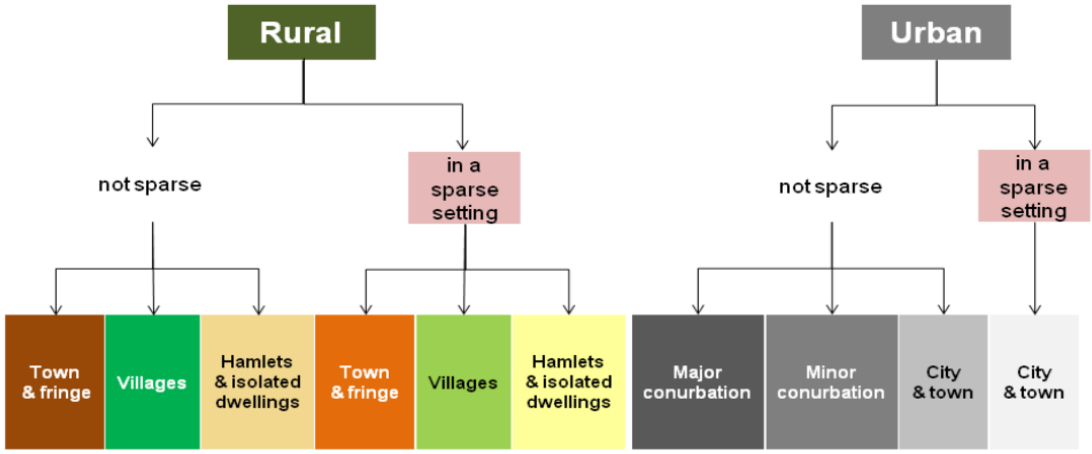
Application in urban contexts

The methodology used in this thesis focuses on remote, rural or disconnected regions and that has been the specific context considered here. However, with a little modification it may also be applicable in an urban or semi-urban scenario. For example, job accessibility reflects the frequency of public transport services and the speed of travel which could be adjusted to allow for the greater frequency of transport in urban areas. It could also be adjusted to reflect the spatial differences in accessibility which would still apply in an urban environment.

10 Appendix

10.1 Rural and urban areas classification

The need for a more consistent approach in distinguishing urban and rural environments led in 2001, led to a review of the definitions of urban and rural by the Department of Transport, Local Government and the Regions (Bibby and Shepherd, 2001). The Rural-Urban Classification distinguishes rural and urban areas, where rural implies falling outside of settlements with more than 10,000 resident population. Census Output Areas - the smallest areas for which data are available from the 2001 and 2011 Censuses - are assigned to one of four urban or six rural categories:



(Source: ONS Defining Rural Areas, 2015 [7])

The local authority categories are:

[7] "in a sparse setting" reflects where the wider area is remotely populated.

10.2 Census aggregation levels

The types of aggregation available were:

Output areas (OA)

Output Areas are the base unit for Census data releases and built up from clusters of adjacent unit postcodes. Due to their smaller size, they allow for finer resolution of data analysis. The OA is the lowest geographical level at which census estimates are provided.

Super Output Areas (LSOA and MSOA)

SOAs were designed to improve reporting of small area statistics and applied to the 2001 and 2011 census. They are constructed from groups of output areas and statistics for lower layer super output areas (LSOA) and middle layer super output areas (MSOA) were originally released in 2004 for England and Wales. LSOAs are built from groups of contiguous Output Areas and have been automatically generated to be as consistent in population size as possible, and typically contain from four to six Output Areas. The minimum population for an LSOA is 1000 with mean of 1500.

Enumeration Districts (ED)

EDs were used for both data collection and output in 1991. In 2001, EDs sometimes straddled 2001 administrative boundaries and deemed unsuitable for data output and used for data collection only. There were no official enumeration districts (ED) created for the 2011 Census. In 2001, England and Wales had 116,895 EDs; the majority were different from 1991 equivalents, with an average size of approximately 200 households.

10.3 Data sources

Measure	Aggregate Level	Period	Description	Source	Frequency
Age Structure	LSOA/DZ	91/01/11	Age distribution	UK Census	10 years
Population by Gender	LSOA/DZ	91/01/11	Gender split and population density	UK Census	10 years
Qualifications and students	LSOA/DZ	91/01/11	Level of education	UK Census	10 years
Car Availability	LSOA/DZ	91/01/11	Cars per household	UK Census	10 years
Method of travel to work	LSOA/DZ	91/01/11	Means of travel to work	UK Census	10 years
Distance Travelled to work	LSOA/DZ	01/11	Distance bands for travel to work	UK Census	10 years
OD by District	District	01/11	Origin destination by district	UK Census	10 years
OD by MSOA	MSOA/IZ	01/11	Origin destination by MSOA	UK Census	10 years
Station Usage	Station	1997-2017	Historic usage of station by year	ORR Portal	Annually
Property Prices	Post Code	01/11	Property transactions	Land Registry Rightmove	Monthly
Rooms	LSOA/DZ	91/01/11	Average rooms per household	UK Census	10 years
Housing Types	LSOA/DZ	91/01/11	Breakdown by housing type	UK Census	10 years
Economic Activity	LSOA/DZ	91/01/11	Economically active and inactive by LSOA	UK Census	10 years
Industry of Employment	LSOA/DZ	91/01/11	Types of industry represented in jobs	UK Census	10 years
Occupation of Employment	LSOA/DZ	91/01/11	Types of occupation represented in jobs	UK Census	10 years
JSA Count	LSOA/DZ	2005-2017	Job Seekers Allowance by occupation	Nomis	Monthly
Job Centre Plus Vacancies	LSOA/DZ	2004-2017	Job Centre Plus Vacancies by Occupation	Nomis	Monthly
Income Support Data	LSOA/DZ	2000-2017	Income Support	Nomis	Monthly

10.4 Logic model

Context

Rail interventions to previously disconnected regions once having a rail link

The rail link now connects previously remote region to a larger conurbation or city

Accessibility to jobs and essential services.

The two-way road effect in relation to jobs, tourism and retail impacts.

Some ex ante evaluation has been done for such cases but very little ex post

Input

- Stakeholders
- DfT - Transport Scotland
- Welsh Government - TfL - EU etc.
- Feasibility
- Design
- Construction
- Funding
- Contractors

Output

- New rail stations
- Connection to rail network
- Link to larger conurbation
- Freight connections
- Passenger connections
- Regular services throughout the day
- Transport Hub to other travel modes

- Employment
- Property
- Accessibility

Outcomes: short term impacts

- Commuting levels
- Accessible jobs
- Employment levels
- Key Benefits and Job Seekers Allowance claimants
- House Price Transactions
- Housing completions and new developments index
- Transport link status
- Transport Patronage Levels
- Vehicle and passenger mileage
- Transport Cost
- Reduction in car use - number or length of car trips
- Travel cost and time to essential services
- Population levels

- Reduced travel time to work
- Improvement in accessibility for commuting to jobs
- Increased commuting to work
- Local job opportunities
- Improved accessibility to and from larger urban areas
- House Price movements
- Movement in GDP/GVA per worker in industry and retail
- Changes in Transport Coverage
- Increased use of public transport and modal shift from car
- Improved accessibility to shopping
- Reduction in transport cost
- Changes in accessibility to essential services by socio-demographic group
- Change in demographic mix

Outcomes: long term impacts

- More jobs taken in city
- Movement in job location
- Employment or residents up Employment on High Street down
- Buoyant house market
- Fewer high street outlets
- Tourist economy improves
- Increased turnover for tourism-centred business
- Agglomeration effects
- Greater productivity
- Improved accessibility and social inclusion for disadvantaged groups

Assumptions There is causality that can be established Access will be available to suitable secondary data and primary data will be accessible through Transport Scotland - valid case studies and a counterfactual can be used

External Factors The state of the economy. Population. Housing supply. House prices. Retail developments and other land use developments. Other transport modes (new services e.g. tram, or bus services withdrawn, changes in the town centre).Opening or closing of major employers.

Risks Data at suitable levels of aggregation may not be available in all cases. Less recent data may be so confounded by other factors that a meaningful causal relationship may not be forthcoming. Time factors and scope

10.5 Robin Hood Line

10.5.1 Station Usage 1997-2015

Full entries

Year	Bulwell	Hucknall	Newstead	Kirkby in Ashfield	Sutton Parkway	Mansfield Town	Mansfield Woodhouse	Shirebrook	Langwith	Creswell	Whitwell
1997/8	30386	37005	8513	19328	10343	28340	14389				
1998/9	28131	43269	9975	23639	11570	34936	14972	7392	5582	5907	3652
1999/0	25281	45329	9677	25730	14912	37550	16603	8890	6042	6702	4197
2000/1	26784	52182	9396	27411	18553	42469	19863	9900	6817	6115	4741
2001/2	25292	51474	8889	27146	20201	48427	17538	10309	7099	6112	3860
2002/3	22062	48595	8445	28663	21147	53442	19430	10839	6979	6420	3727
2003/4	18832	45715	8000	30181	22093	58457	21323	11369	6858	6728	3594
2004/5	15602	42836	7556	31698	23039	63472	23215	11899	6738	7036	3461
2005/6	15495	44481	7378	32916	23826	70774	23213	11650	6345	7327	3821
2006/7	14198	47005	7338	31524	23904	68264	23669	10709	6052	6949	3706
2007/8	10074	47735	7249	28337	23361	65307	24762	9972	5309	6891	3193
2008/9	9286	43092	6762	28362	23913	65435	26191	10500	4825	7244	3282
2009/10	11639	41132	7019	28354	23864	62344	23765	9385	4225	6932	3362
2010/1	11022	42438	7692	28642	25833	65766	24745	10364	4549	7331	4079
2011/2	12420	45197	8634	32782	28695	69996	28006	13640	5332	8847	5237
2012/3	12523	40531	7923	31689	28719	66902	26541	15343	4874	8796	4819
2013/4	11028	37697	7044	27653	27913	61663	25158	13158	4714	7319	4145
2014/5	12327	41433	8696	32454	35623	72658	29787	15838	5600	8690	4502

Reduced entries

Year	Bulwell	Hucknall	Newstead	Kirkby in Ashfield	Sutton Parkway	Mansfield Town	Mansfield Woodhouse	Shirebrook	Langwith	Creswell	Whitwell
1997/8	25799	36522	10140	36560	19328	69060	27731				
1998/9	27952	37361	10185	38872	21677	85835	28210	15747	11417	11677	6853
1999/0	27977	38212	9292	40275	24537	87800	25730	15866	10818	12472	6792
2000/1	29031	43965	9403	40740	26144	92071	28024	16948	12005	12450	6827
2001/2	24611	39401	8825	39656	25249	91807	25235	16226	11035	11601	6190
2002/3	20049	35936	7615	38215	25053	92773	24801	15304	10099	11886	6106
2003/4	15488	32471	6404	36775	24856	93740	24368	14383	9162	12170	6021
2004/5	10926	29006	5194	35334	24660	94706	23934	13461	8226	12455	5937
2005/6	10429	26787	4990	33942	24274	93706	25394	14231	7966	12166	6542
2006/7	8012	26449	5166	33983	22148	90845	26379	14512	6916	12570	5667
2007/8	6257	26096	4150	33697	21635	87736	27196	14579	6586	11505	5226
2008/9	6209	27069	4445	37428	21437	93318	30948	15386	7233	11889	5530
2009/10	6789	31433	7404	51290	32762	119830	41420	21208	8537	15023	6736
2010/1	6010	25440	5370	40464	25443	100020	32649	17121	7245	13775	5263
2011/2	5820	22658	5356	40749	25412	94040	31374	16611	6243	11313	5169
2012/3	6178	18679	5166	38306	25177	89715	31424	17070	5704	10272	4531
2013/4	5135	14906	4513	29680	21742	78199	26707	15740.5	4743	8571	3477
2014/5	6405	17585	4939	34181	27693	90121	31266	17823	5893	9156	4253

Season Tickets

Year	Bulwell	Hucknall	Newstead	Kirkby in Ashfield	Sutton Parkway	Mansfield Town	Mansfield Woodhouse	Shirebrook	Langwith	Creswell	Whitwell
1997/8	5376	12124	4419	6232	7747	9967	9979				
1998/9	4279	11212	4424	7351	8924	9875.5	8864	1459.5	632	344.5	186.5
1999/0	4429	7384.5	3589	7383	9886.5	14347	11232	1842	590.5	632.5	322.5
2000/1	5004.5	9968	3385	11078.5	11307.5	18001	13554	2994.5	105	1476	742.5
2001/2	4765	9538	2916	11318	12455	18873	13219	3816	455	1201	855
2002/3	4870	9200	2942	11723	13301	21076	13350	3419	671	1227	1031
2003/4	4975	8863	2968	12128	14147	23279	13482	3022	887	1253	1207
2004/5	5080	8525	2994	12533	14993	25482	13613	2625	1103	1280	1383
2005/6	4679	7573	2415	12629	14625	21805	15007	3609	759	1909	887
2006/7	4484	9227	1614	10787	13979	18046	15464	3673	570	1933	611
2007/8	4126	7757	1877	8242	13964	16333	13549	2699	270	1322	810
2008/9	4047	8074	2120	9700	10868	15587	14074	1882	578	1009	442
2009/10	5993	11026	1840	9247	11240	14821	12710	1788	144	1364	1005
2010/1	7049	14637	2812	13972	13001	17241	15633	3036	632	2755	1032
2011/2	8178	14823	3385	14479	16109	19593	20790	3083	1196	2270	1813
2012/3	8329	13608	2347	15933	19618	18288	19931	3474	1082	1783	1122
2013/4	7133	13657	2755	16389	17970	17051	17926	2919	1097	1364	1554
2014/5	6744	15699	3334	18823	21213	20650	18293	3708	1197	1334	1508

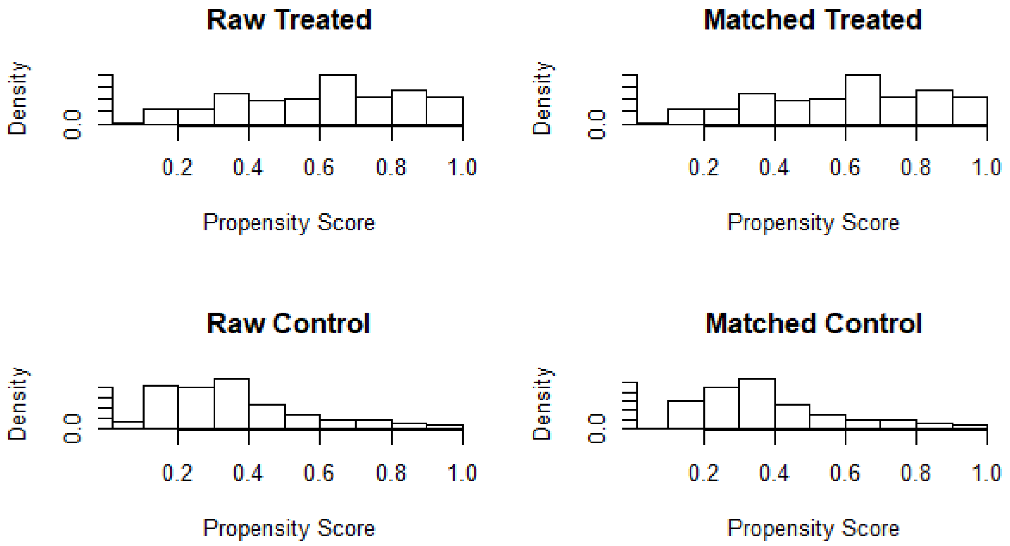
10.5.2 Propensity Testing - 4 km threshold for treatment group

Summary of balance for all data							
	Means Treated	Means Control	SD Control	Mean Diff	eQQ Med	eQQ Mean	eQQ Max
Distance	0.601	0.357	0.1969	0.243	0.2814	0.245	0.361
Health	0.720	0.327	0.560	0.393	0.400	0.398	0.510
Housing	18.625	15.414	9.171	3.213	4.115	3.889	10.340
Education	38.164	28.195	19.051	9.966	9.255	10.221	20.000
Crime	0.798	0.178	0.592	0.620	0.600	0.627	0.850
Summary of balance for matched data							
Distance	0.601	0.386	0.188	0.2152	0.249	0.215	0.340
Health	0.720	0.398	0.534	0.3222	0.340	0.322	0.410
Housing	18.625	15.265	9.281	3.360	4.395	4.041	10.340
Education	38.164	30.112	18.67	8.052	6.660	8.253	18.080
Crime	0.798	0.274	0.541	0.524	0.510	0.524	0.780

Percent balance improvement				
	Mean Difference	eQQ Median	eQQ Mean	eQQ Maximum
Distance	11.608	11.247	12.188	5.679
Health	17.997	15.000	19.135	19.607
Housing	-4.584	-6.804	-3.916	0
Education	19.203	28.039	19.250	9.600
Crime	15.467	15.000	16.314	8.235

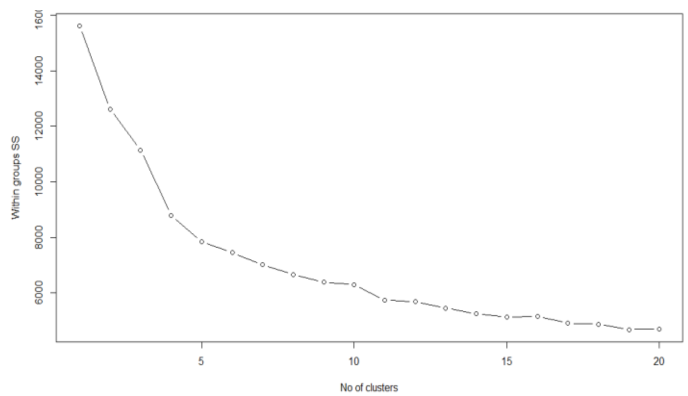
Sample sizes		
	Control	Treated
All	250	224
Matched	224	224
Unmatched	26	0
Discarded	0	0

Histogram

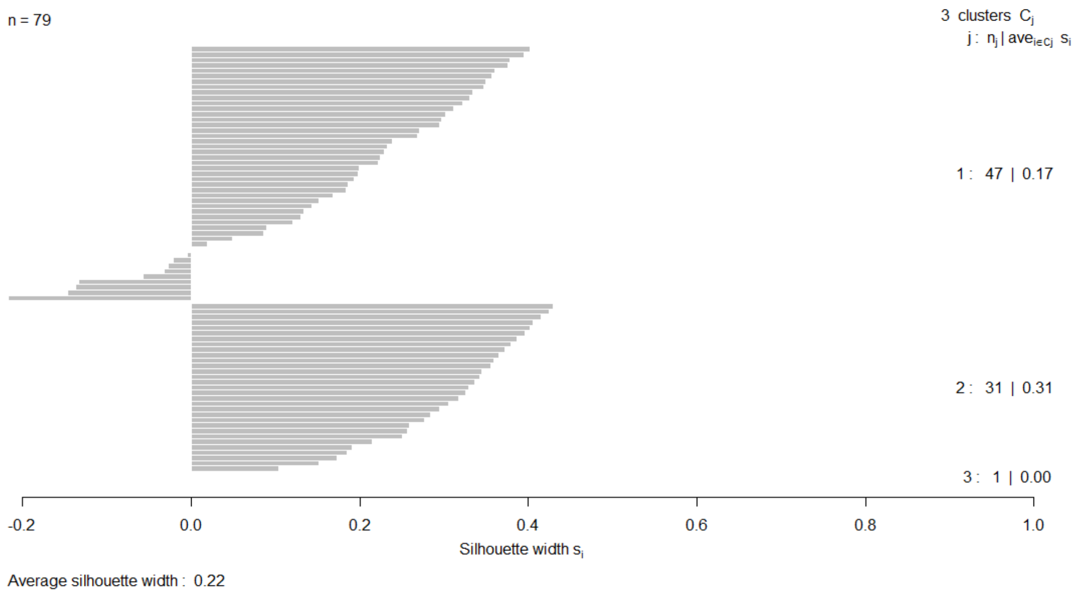


10.5.3 Cluster Analysis

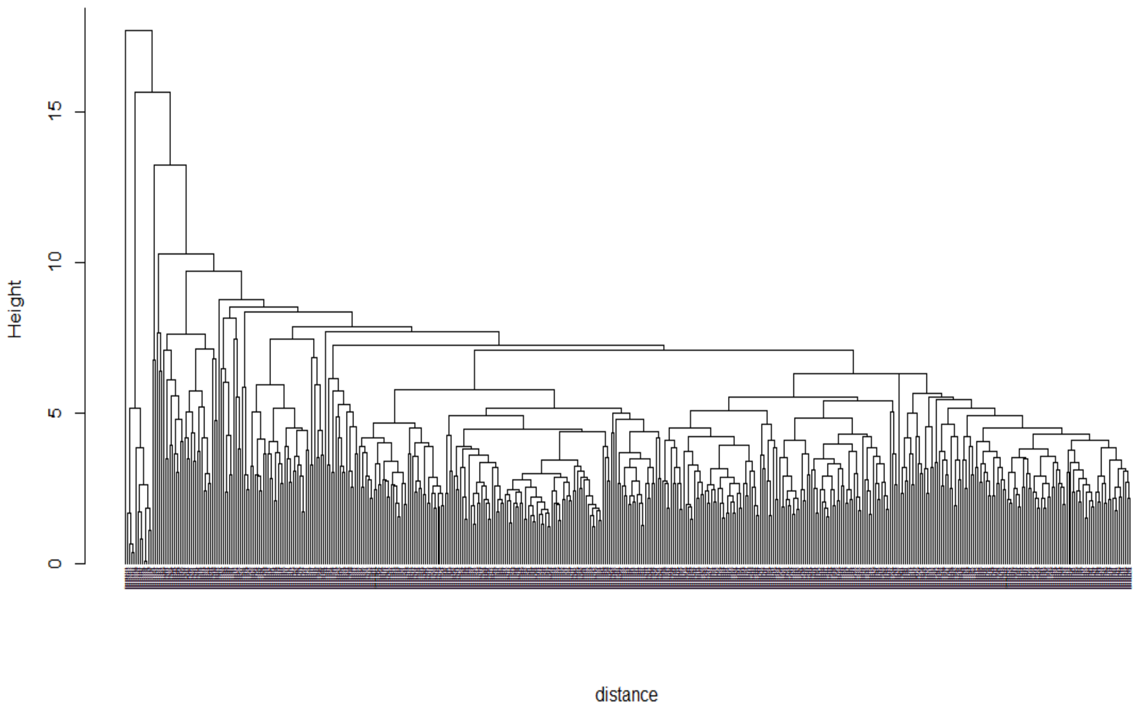
Elbow chart



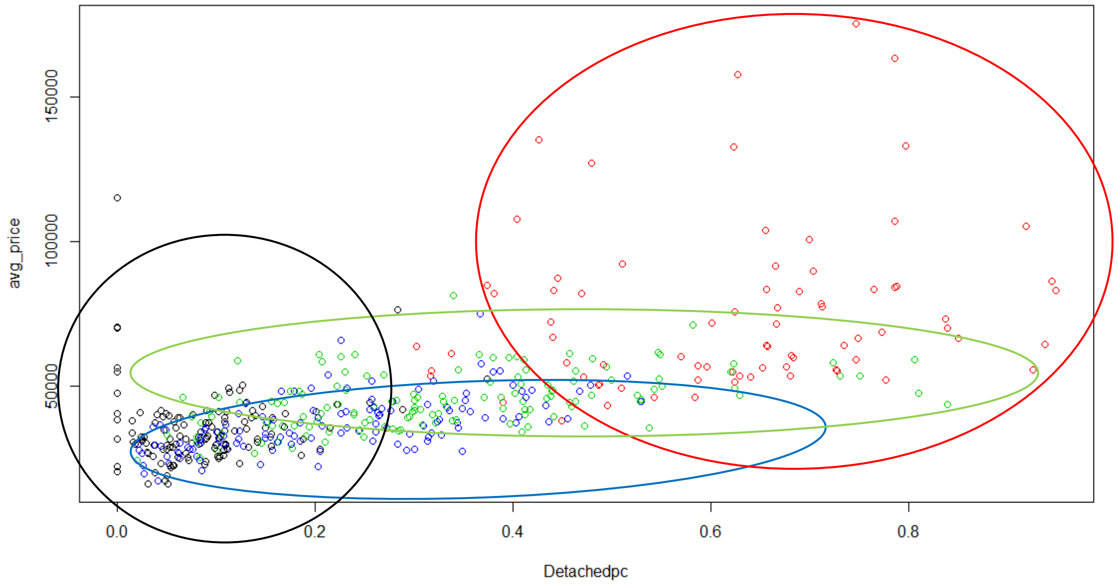
Silhouette diagram



Dendrogram



K-Means Diagram



10.5.4 Overall Statistics

Year	1991			2001			2011		
Measure	Mean	Standard Error	Median	Mean	Standard Error	Median	Mean	Standard Error	Median
Population Density	24.5	1.0	20.4	25.4	1.0	22.2	27.1	1.0	24.4
Distance to nearest station	7.2	0.2	6.5	3.5	0.1	2.7	3.5	0.1	2.7
Average House Price	47928.9	1007.0	42917.0	64326.6	1574.5	56205.3	123324.6	2465.4	111616.7
Detached	165.1	6.0	134.0	193.3	6.7	164.0	208.9	6.9	179.0
Semi	247.7	5.7	235.0	270.7	5.8	258.0	282.5	5.7	275.0
Terrace	148.1	6.0	107.5	136.7	5.6	99.0	143.8	5.3	111.0
Flat	51.2	2.7	31.0	54.5	3.3	28.0	64.7	3.5	37.0
Other	0.3	0.0	0.0	2.3	0.5	0.0	2.1	0.5	0.0
Average rooms	5.1	0.0	5.1	5.3	0.0	5.3	5.5	0.0	5.4
Total population	1446.9	12.2	1443.5	1501.6	8.8	1498.5	1568.2	12.0	1533.5
No qualifications	402.4	5.4	410.5	402.5	5.3	410.0	379.1	5.4	378.5
Level 1	200.7	2.0	199.0	200.8	2.0	199.0	193.1	2.0	191.0
No cars	198.6	4.6	189.0	177.6	4.5	161.5	166.1	4.3	147.0
1 car	265.8	2.6	265.5	284.2	2.4	284.0	292.0	2.7	289.0
Full time	444.7	4.5	435.0	429.2	4.5	425.0	448.2	5.0	436.0

10.5.5 Job accessibility index

Overall

Overall		with skills matching			no matching		
Mode	Method	1991	2001	2011	1991	2001	2011
Bus	Cost	0.039	0.025	0.019	0.218	0.202	0.173
	Time	0.131	0.093	0.080	0.739	0.749	0.716
Car	Cost	0.091	0.049	0.030	0.510	0.390	0.266
	Time	0.164	0.117	0.101	0.935	0.948	0.907
Rail	Cost	0.027	0.024	0.018	0.147	0.193	0.156
	Time	0.101	0.093	0.080	0.551	0.749	0.718
Overall	Cost	0.052	0.033	0.022	0.292	0.262	0.198
	Time	0.132	0.101	0.087	0.742	0.815	0.780

Treatment

Treatment		with skills matching			no matching		
Mode	Method	1991	2001	2011	1991	2001	2011
Bus	Cost	0.042	0.028	0.022	0.238	0.222	0.192
	Time	0.137	0.098	0.085	0.777	0.786	0.753
Car	Cost	0.098	0.054	0.035	0.553	0.435	0.310
	Time	0.165	0.118	0.103	0.941	0.954	0.914
Rail	Cost	0.026	0.029	0.022	0.140	0.235	0.193
	Time	0.097	0.102	0.088	0.531	0.819	0.784
Overall	Cost	0.055	0.037	0.026	0.310	0.297	0.232
	Time	0.133	0.106	0.092	0.750	0.853	0.817

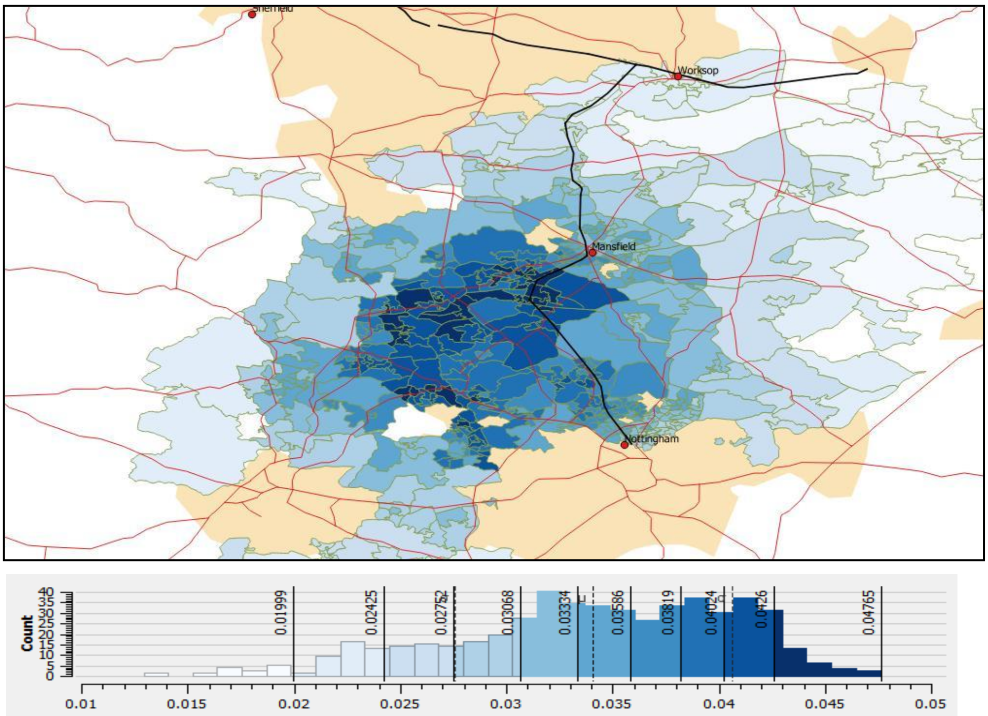
Control

Control		with skills matching			no matching		
Mode	Method	1991	2001	2011	1991	2001	2011
Bus	Cost	0.040	0.026	0.020	0.222	0.207	0.177
	Time	0.134	0.095	0.082	0.748	0.757	0.725
Car	Cost	0.095	0.051	0.031	0.520	0.399	0.274
	Time	0.167	0.119	0.102	0.937	0.950	0.909
Rail	Cost	0.030	0.026	0.019	0.161	0.201	0.162
	Time	0.108	0.098	0.084	0.586	0.774	0.742
Overall	Cost	0.055	0.034	0.023	0.301	0.269	0.204
	Time	0.137	0.104	0.089	0.757	0.827	0.792

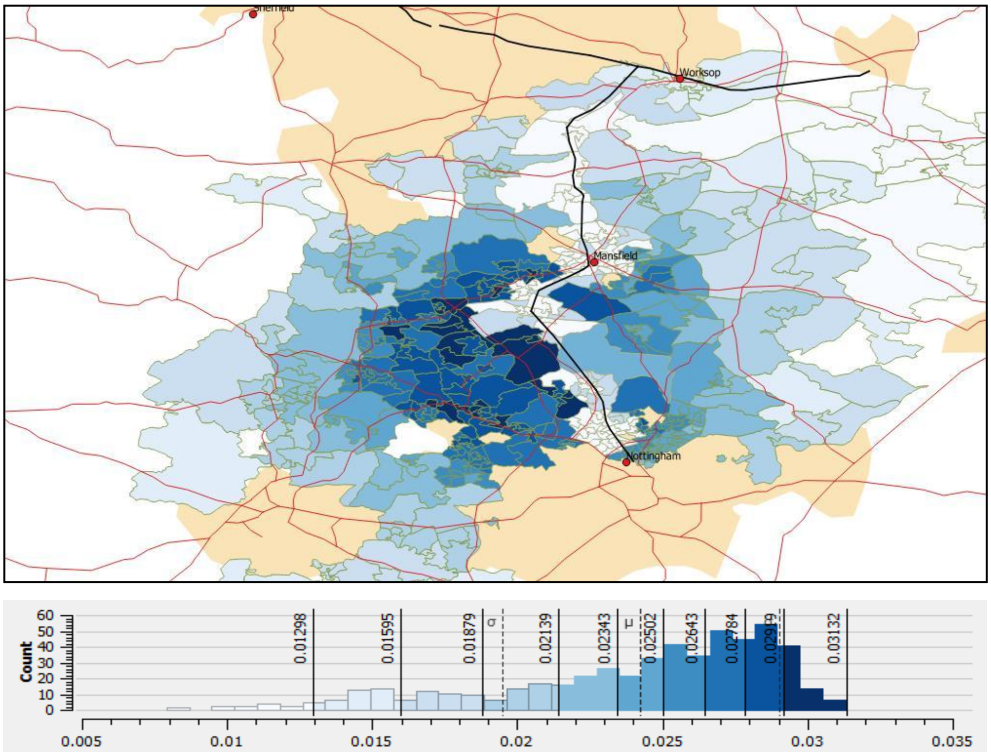
10.5.6 Job accessibility index - Spatial analysis

Cost basis with matching

1991

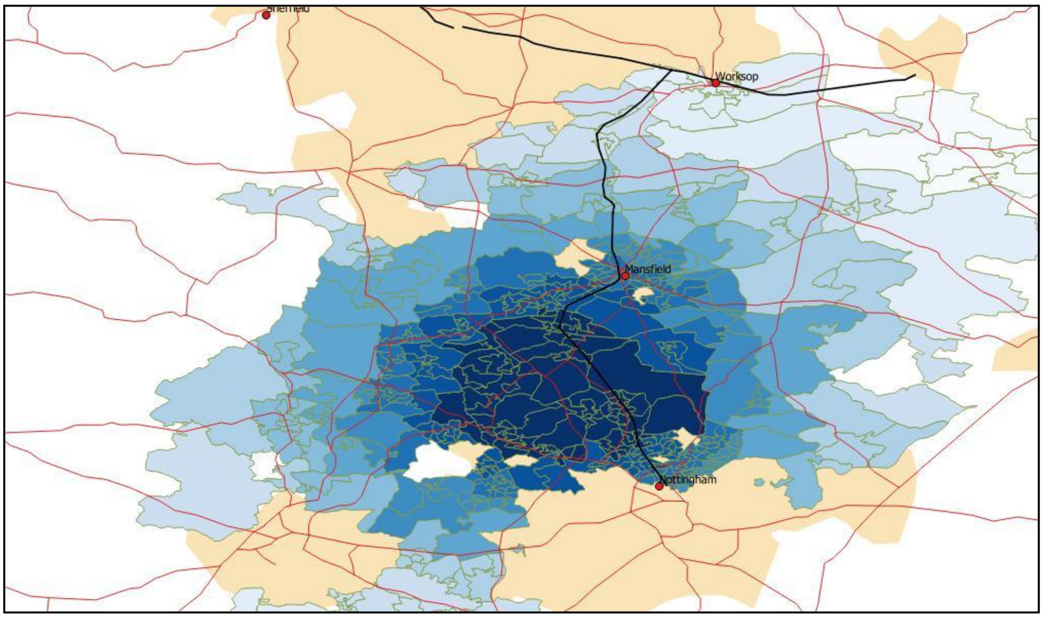


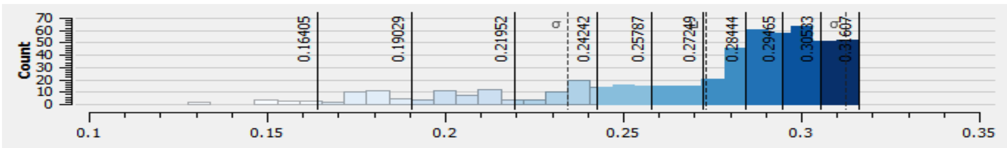
2011



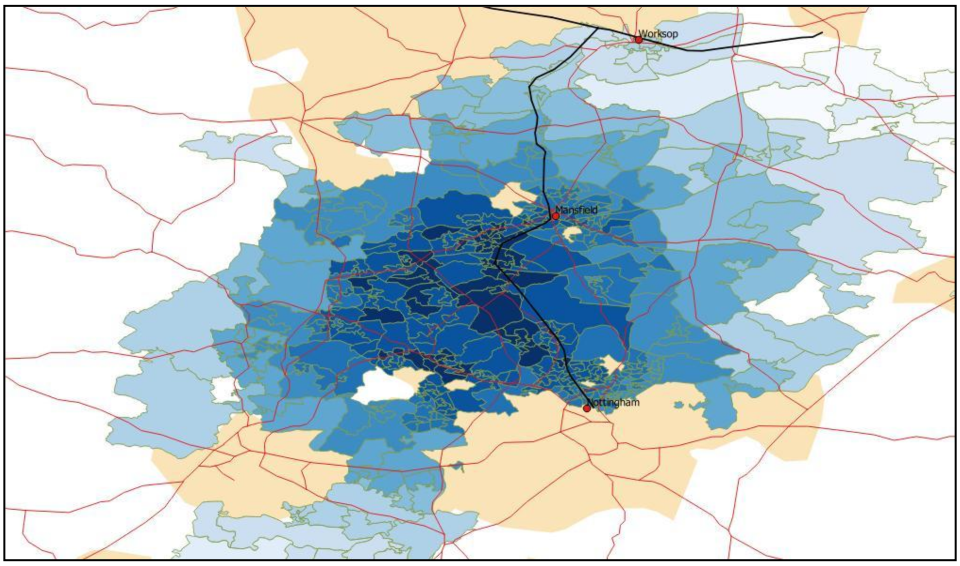
Cost basis without matching

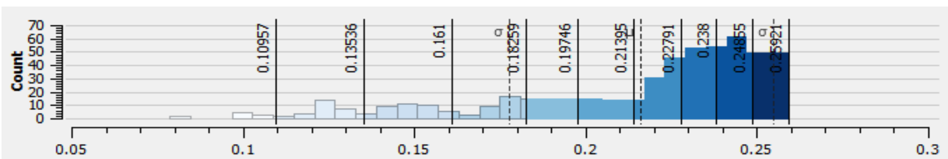
1991





2011





10.5.7 Essential services accessibility index

Accessibility comparison - time basis

Year	Group	Rail station	Nursery School	Primary School	Secondary School	Hospital	Average across all services
1991	Control	0.879	0.789	0.809	0.807	0.795	0.816
	Treatment	0.866	0.786	0.805	0.803	0.801	0.812
2001	Control	0.888	0.792	0.813	0.811	0.799	0.820
	Treatment	0.900	0.797	0.817	0.815	0.812	0.828
2011	Control	0.888	0.792	0.813	0.811	0.799	0.820
	Treatment	0.900	0.797	0.817	0.815	0.812	0.828

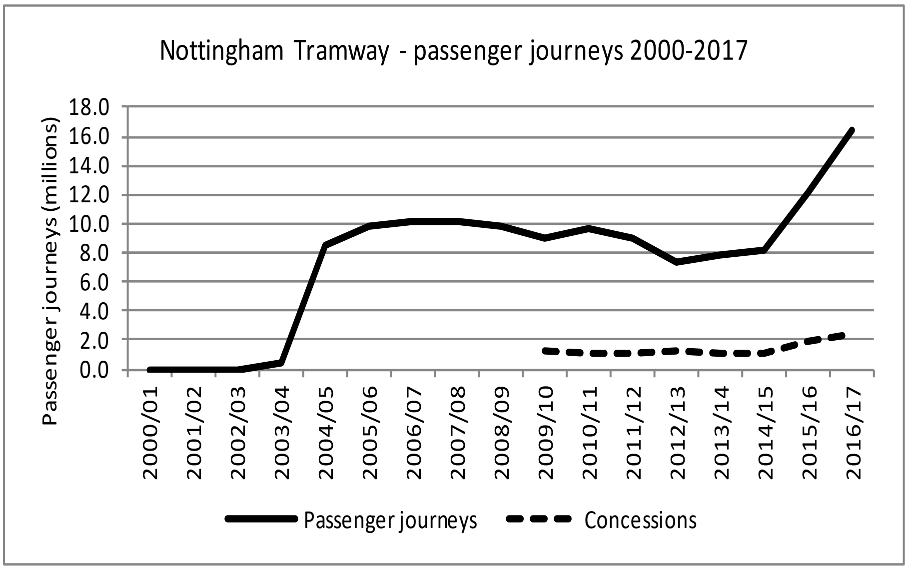
Accessibility comparison - cost basis

Year	Group	Rail station	Nursery School	Primary School	Secondary School	Hospital	Average across all services
1991	Control	0.390	0.175	0.236	0.229	0.190	0.244
	Treatment	0.336	0.169	0.228	0.222	0.212	0.233
2001	Control	0.415	0.164	0.242	0.232	0.183	0.247
	Treatment	0.474	0.169	0.250	0.243	0.228	0.273
2011	Control	0.397	0.146	0.242	0.228	0.168	0.236
	Treatment	0.467	0.148	0.250	0.240	0.221	0.265

10.5.8 Nottingham Tramway usage statistics 2004 to the present

Nottingham Tramway - average trip purpose

Trip Purpose						
Commuting	Business	Education	Shopping	Personal business	Leisure	Other
27.95%	2.23%	11.56%	22.18%	6.89%	26.29%	2.89%



Nottingham Tramway - passenger and concessionary journeys (millions)

10.6 Stirling-Alloa

10.6.1 Station Usage 2008-2015

Full entries

Year	Alloa	Bridge Of Allan	Dunblane	Stirling
0809	38913	47266	75010	308154
0910	48901	52016	72389	315708
1011	50722	49811	70304	330844
1112	50199	51363	70366	319567
1213	47820	52215	69560	313933
1314	47304	52248	72913	325969
1415	49647	55127	75723	342025

Reduced entries

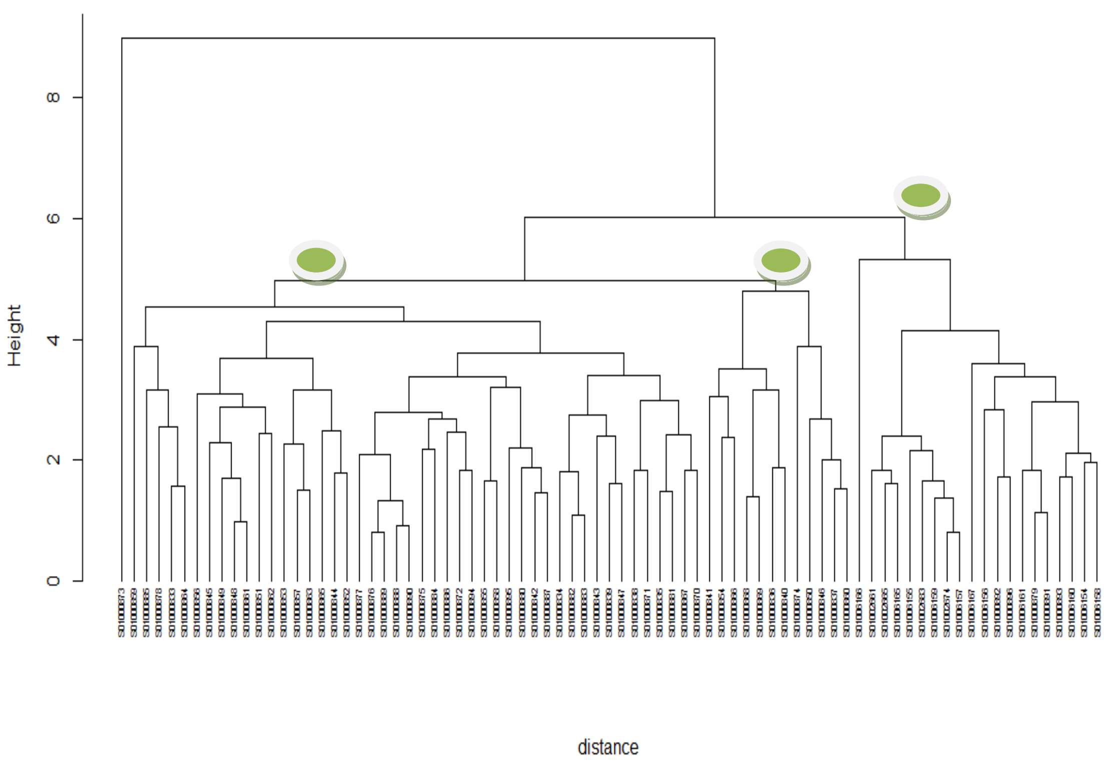
Year	Alloa	Bridge Of Allan	Dunblane	Stirling
0809	101245	36107	133596	578716
0910	110361	38646	133359	589438
1011	110863	38362	130252	610457
1112	114030	40529	130209	617610
1213	107571	40931	131501	614578
1314	104421	42103	128574	608796
1415	112652	47096	138247	668345

Season Tickets

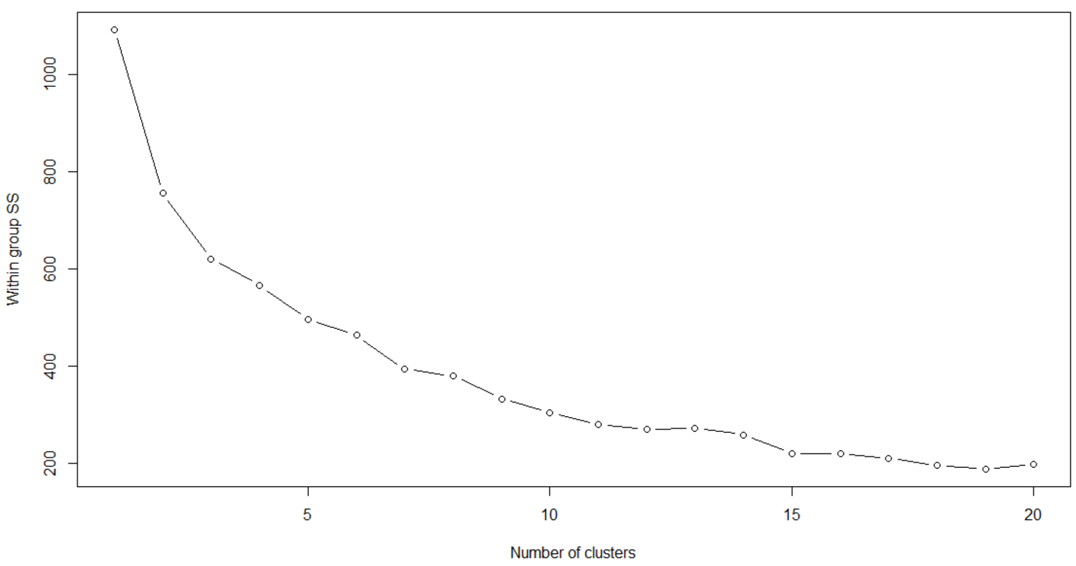
Year	Alloa	Bridge Of Allan	Dunblane	Stirling
0809	27669	28926	49441	179094
0910	35774	26957	42612	172350
1011	35578	27542	48280	192274
1112	36346	29843	46877	193736
1213	35041	30962	50968	190562
1314	40168	35007	52769	195115
1415	38902	35263	53061	197393

10.6.2 Cluster Analysis

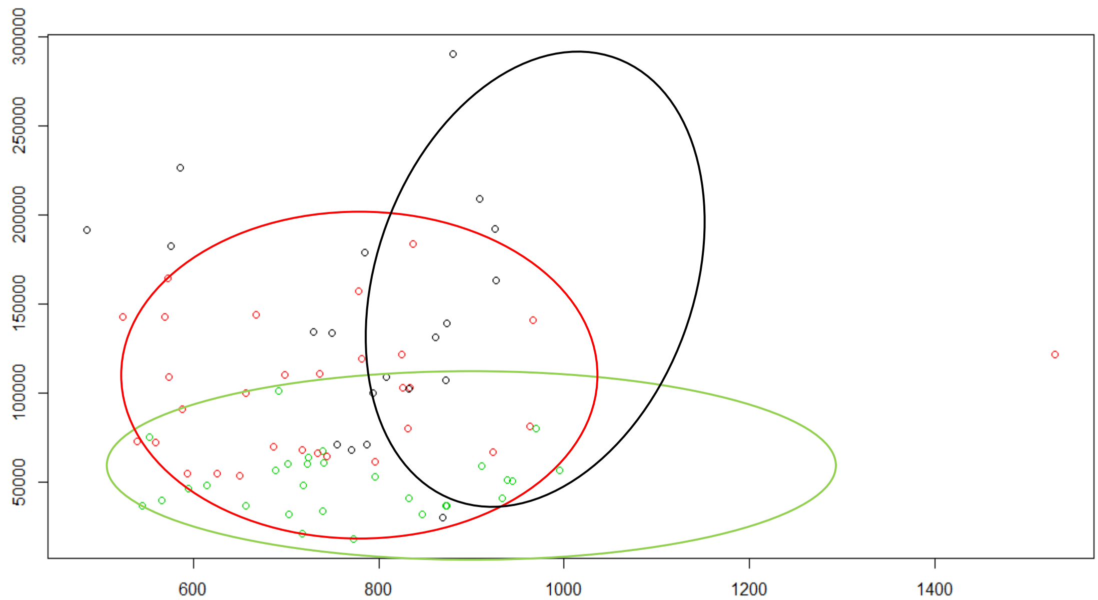
Dendrogram



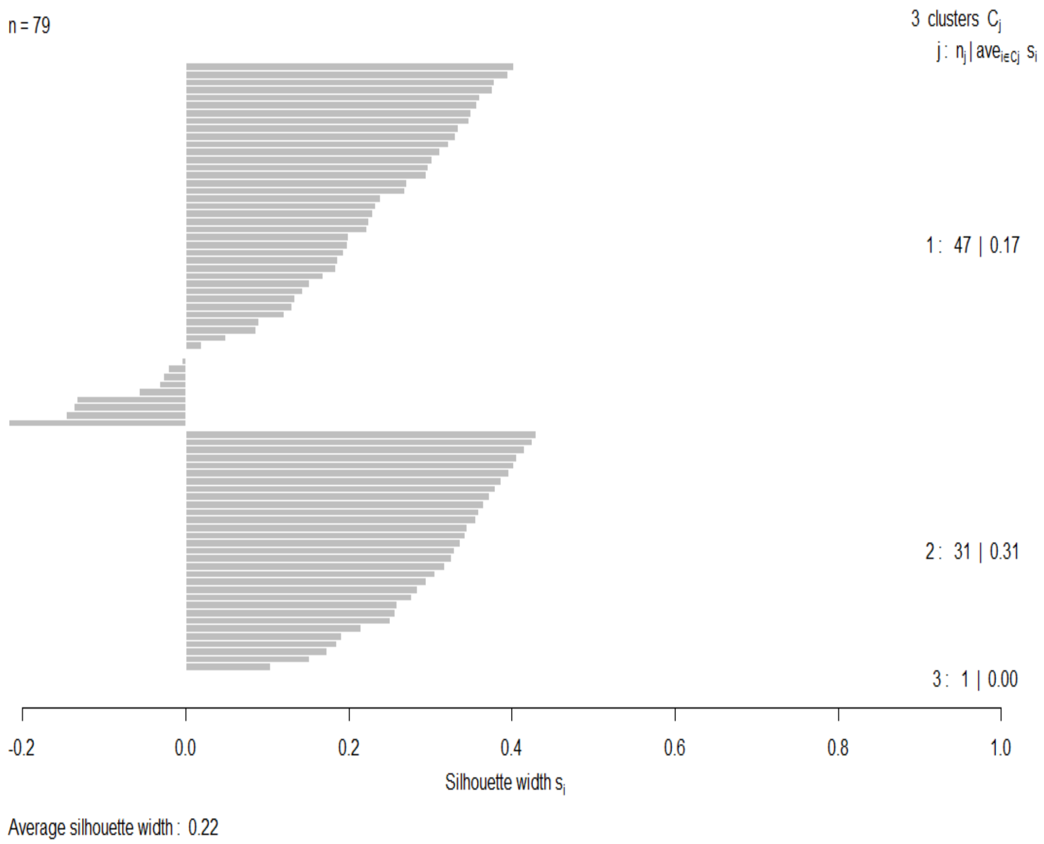
Elbow Chart



K-Means Diagram



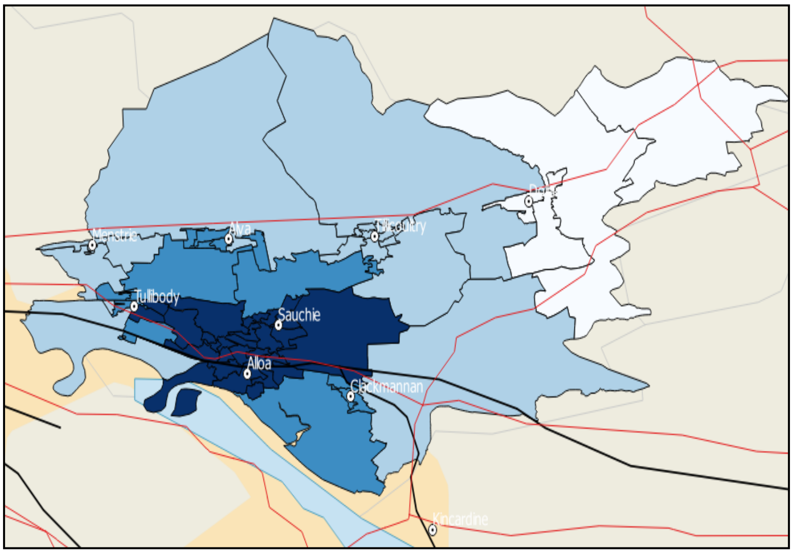
Silhouette Diagram

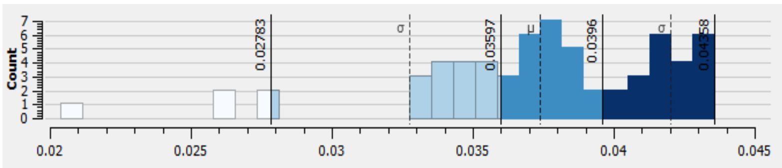


10.6.3 Job accessibility index - Spatial analysis

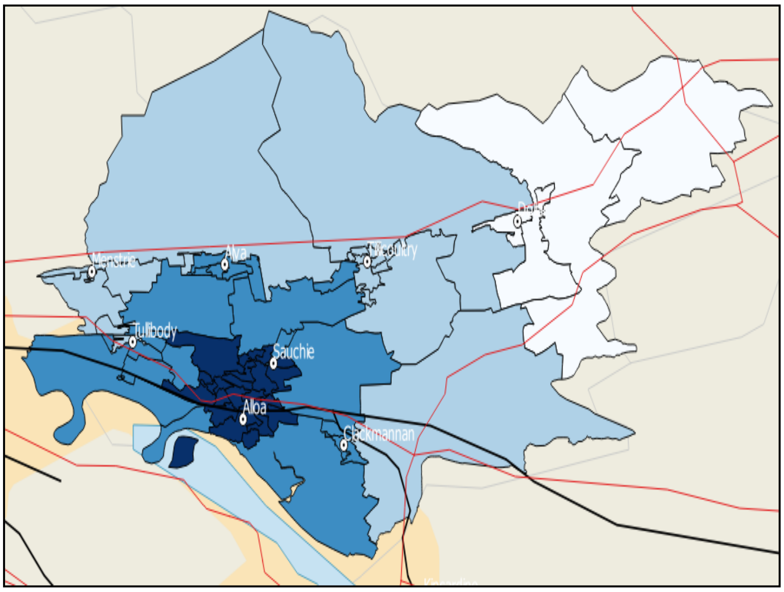
Cost basis with matching

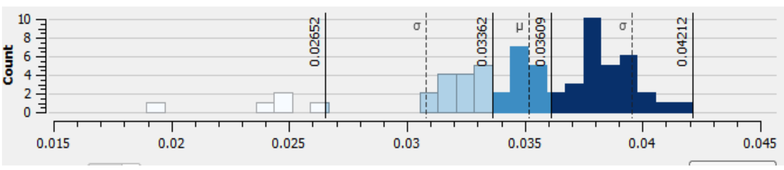
2001





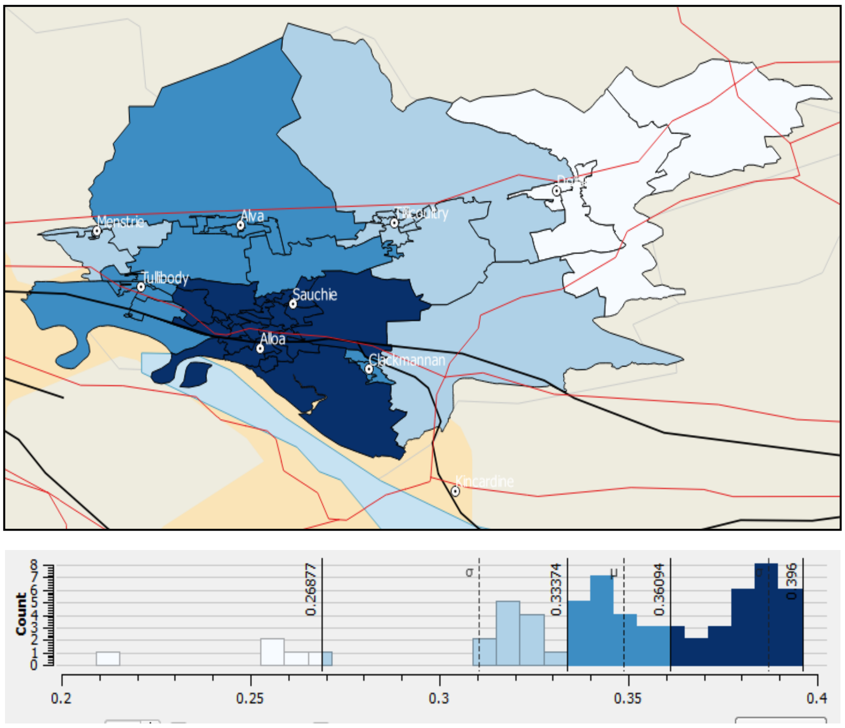
2011



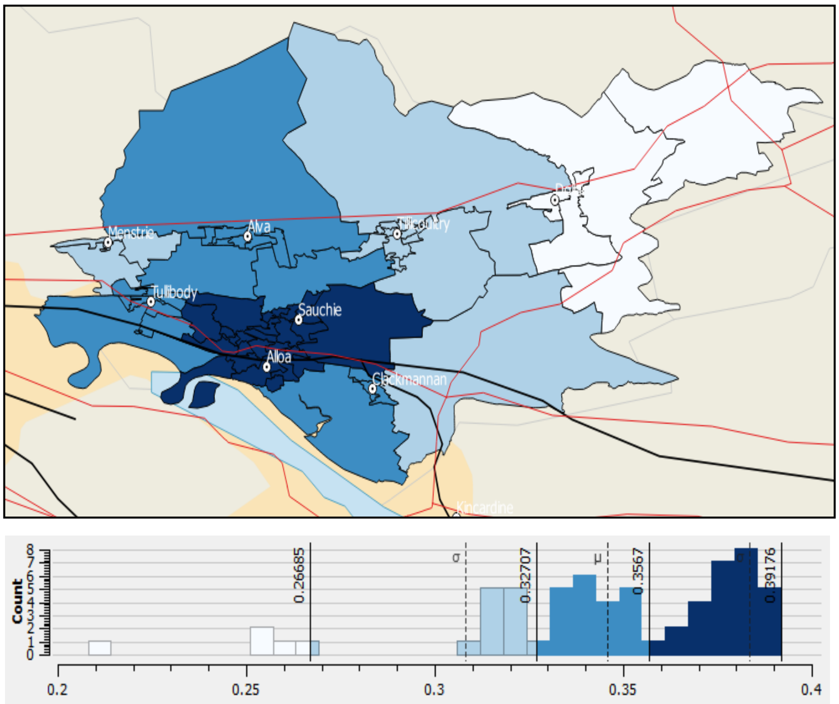


Cost basis without matching

2001



2011



10.6.4 Essential services accessibility index

Time basis comparison

Year	Group	Rail station	Nursery School	Primary School	Secondary School	Hospital	Average across all services
2001	Control	0.859	0.790	0.798	0.793	0.775	0.803
	Treatment	0.860	0.798	0.798	0.796	0.777	0.806
2011	Control	0.883	0.801	0.808	0.804	0.785	0.816
	Treatment	0.899	0.814	0.815	0.813	0.794	0.827

Cost basis comparison

Year	Group	Rail station	Nursery School	Primary School	Secondary School	Hospital	Average across all services
2001	Control	0.257	0.188	0.211	0.197	0.145	0.200
	Treatment	0.253	0.211	0.214	0.206	0.149	0.206
2011	Control	0.348	0.204	0.231	0.215	0.157	0.231
	Treatment	0.461	0.243	0.246	0.237	0.171	0.272

10.6.5 GWR Model output - Property

Distance to nearest station

2001

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	2SE	t value
Intercept	10.670	10.790	10.970	11.190	11.420	0.400	11.415	2.718	8.4000
Distance nearest	-0.025	-0.023	-0.021	-0.016	-0.010	0.006	-0.0169	0.034	-0.980
Detached	0.003	0.004	0.0052	0.0057	0.0080	0.0018	0.0034	0.006	1.229
Terraced	0.001	0.001	0.0016	0.0022	0.0030	0.0009	0.0023	0.004	1.118
Semi	0.001	0.0010	0.0012	0.0015	0.0020	0.0005	0.0016	0.004	0.751
% employed	-1.530	-1.251	-0.9690	-0.7143	-0.5999	0.5367	-1.5994	4.042	-0.791
No car household	-0.001	-0.0003	-0.0002	0.0001	0.0003	0.0004	0.0000	0.004	-0.002
1 car household	-0.004	-0.004	-0.0028	-0.0023	-0.0019	0.0013	-0.0028	0.006	-0.912
Level 1 education	0.005	0.0056	0.0057	0.0064	0.0071	0.0008	0.0058	0.007	1.808
Level 4 education	-0.0002	0.0011	0.0015	0.0023	0.0027	0.0013	0.0028	0.004	1.262
No qualification	-0.003	-0.0030	-0.0025	-0.0022	-0.0018	0.0008	-0.0031	0.005	-1.363
Population density	-0.004	-0.0033	-0.0031	-0.0026	-0.0009	0.0008	-0.0038	0.007	-1.102

R^2: 0.4848 F-statistic: 5.731 on 11 and 67 DF

2011

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	SE	t value
Intercept	11.270	11.310	11.320	11.350	11.460	0.040	11.484	0.581	19.750
Distance nearest station	0.017	0.0198	0.0232	0.0253	0.0326	0.0055	0.0257	0.0157	1.638
Detached	0.0027	0.0029	0.0029	0.0030	0.0035	0.0001	0.0033	0.0013	2.514
Terraced	-0.0004	-0.0002	-0.0001	0.0000	0.0003	0.0001	-0.0003	0.0012	-0.239
Semi	-0.0008	-0.0007	-0.0007	-0.0006	-0.0004	0.0001	-0.0006	0.0012	-0.470
% employed	-0.1884	0.0170	0.0612	0.1111	0.2439	0.0941	-0.1566	0.9436	-0.166
No car household	-0.0012	-0.0011	-0.0011	-0.0010	-0.0009	0.0001	-0.0011	0.0013	-0.811
1 car household	0.0000	0.0004	0.0004	0.0005	0.0007	0.0001	0.0006	0.0019	0.325
Level 1 education	-0.0034	-0.0033	-0.0032	-0.0031	-0.0030	0.0002	-0.0033	0.0015	-2.250
Level 4 education	0.0009	0.0014	0.0015	0.0016	0.0019	0.0002	0.0011	0.0009	1.276
No qualifications	0.0009	0.0011	0.0012	0.0012	0.0013	0.0001	0.0010	0.0014	0.717
Population density	0.0012	0.0013	0.0013	0.0014	0.0017	0.0001	0.0015	0.0020	0.728

R^2: 0.6811 F-statistic: 13.01 on 11 and 67 DF

10.6.6 GWR Model output - Employment

Distance to nearest station

2001

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	2SE	t value
Intercept	-0.1433	0.5398	0.9974	1.4720	2.0970	0.9322	-0.1916	1.6244	-0.2360
Distance station	-0.2452	-0.2133	-0.1687	-0.1217	-0.0121	0.0916	-0.0457	0.0818	-1.1170
Detached	-0.0044	-0.0035	-0.0010	0.0026	0.0083	0.0061	-0.0036	0.0094	-0.7670
Terraced	-0.0013	0.0009	0.0022	0.0047	0.0071	0.0038	0.0024	0.0070	0.6780
Semi	-0.0010	0.0001	0.0004	0.0016	0.0028	0.0015	0.0009	0.0072	0.2380
No car household	-0.0002	0.0016	0.0029	0.0033	0.0039	0.0017	0.0000	0.0072	0.0130
1 car household	0.0012	0.0024	0.0038	0.0055	0.0070	0.0031	0.0061	0.0092	1.3160
Level 1 education	0.0042	0.0050	0.0061	0.0068	0.0102	0.0019	0.0086	0.0108	1.5820
Level 4 education	-0.0088	-0.0039	0.0003	0.0020	0.0052	0.0059	-0.0020	0.0072	-0.5400
No qualifications	-0.0101	-0.0088	-0.0079	-0.0071	-0.0050	0.0017	-0.0092	0.0073	-2.5350
Population density	-0.0154	-0.0056	0.0029	0.0134	0.0628	0.0190	0.0500	0.0256	3.9060

R^2: 0.7367 F-statistic: 17.04 on 11 and 67 DF

2011

	Min	25% Quartile	Median	75% quartile	Max	IQ Range	Global (OLS)	2SE	t value
Intercept	-0.4848	-0.0855	0.1565	0.4977	0.8398	0.5832	-0.3492	1.0975	-0.636
Distance station	-0.3854	-0.3191	-0.2569	-0.1605	-0.1247	0.1586	-0.1763	0.1240	-2.843
Detached	-0.0022	-0.0003	0.0007	0.0024	0.0043	0.0027	0.0020	0.0078	0.513
Terraced	-0.0008	0.0005	0.0013	0.0035	0.0052	0.0030	0.0011	0.0073	0.292
Semi	-0.0001	0.0014	0.0029	0.0033	0.0047	0.0019	0.0032	0.0073	0.877
No car household	-0.0012	0.0018	0.0021	0.0024	0.0028	0.0007	0.0011	0.0071	0.317
1 car household	0.0046	0.0069	0.0076	0.0104	0.0168	0.0035	0.0141	0.0097	2.911
Level 1 education	-0.0015	-0.0012	0.0007	0.0011	0.0019	0.0023	-0.0009	0.0087	-0.202
Level 4 education	-0.0060	-0.0026	0.0000	0.0012	0.0033	0.0038	-0.0044	0.0052	-1.719
No qualifications	-0.0083	-0.0077	-0.0066	-0.0064	-0.0058	0.0013	-0.0067	0.0083	-1.618
Population density	0.0303	0.0331	0.0370	0.0422	0.0513	0.0091	0.0443	0.0188	4.718

R^2: 0.7514 F-statistic: 18.41 on 11 and 67 DF

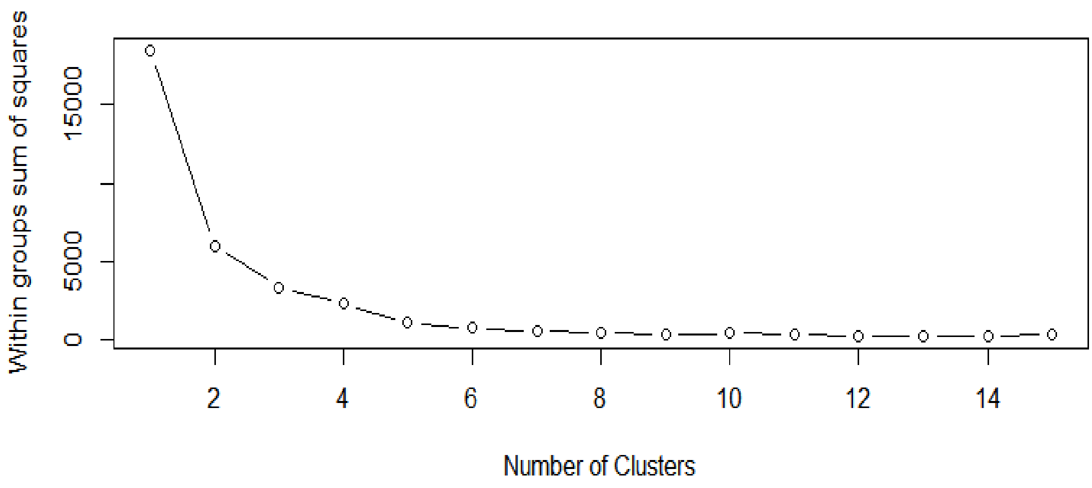
10.7 Borders Rail

10.7.1 Station Usage 2014-2017

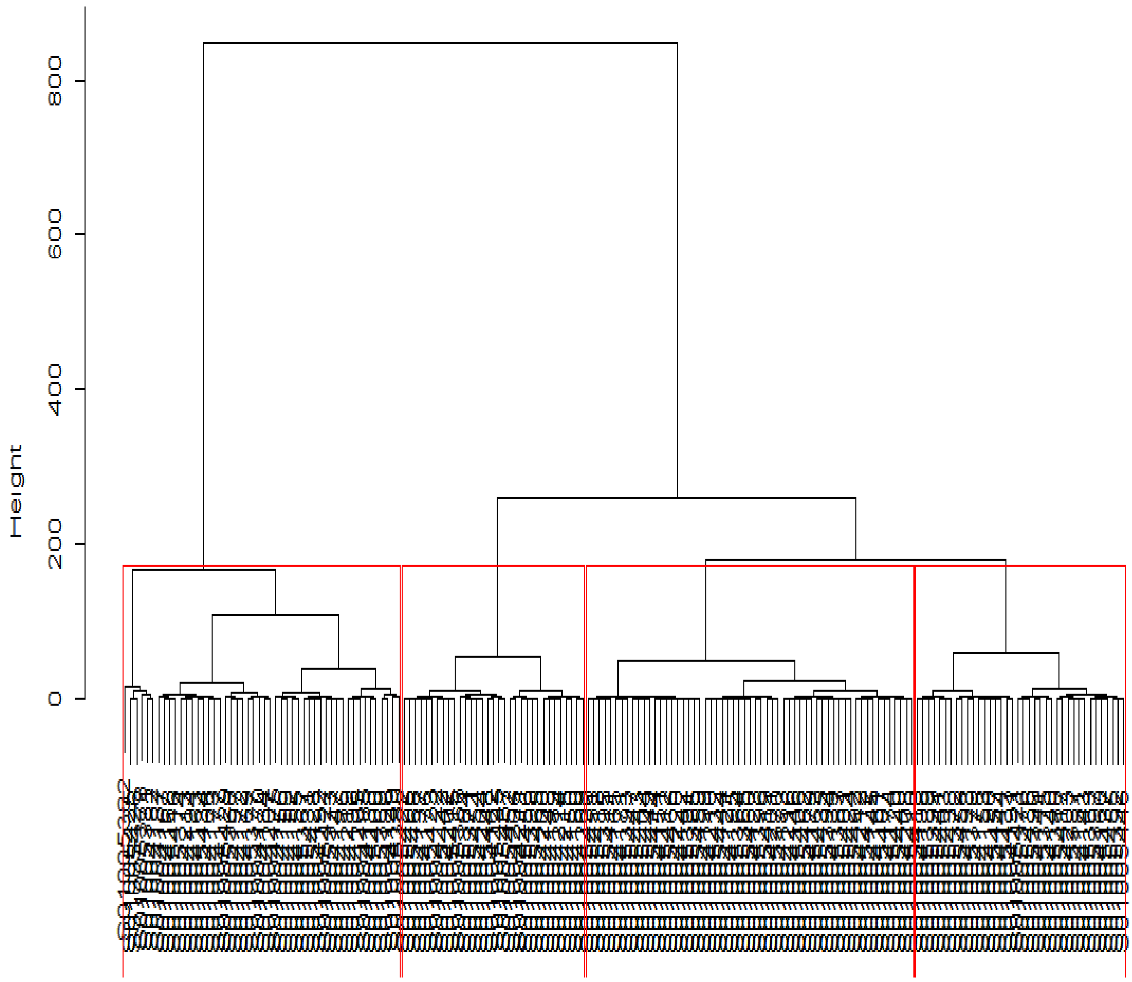
		14/15	15/16				16/17			
Station Name	New station	Total	Full	Reduced	Season tickets	Total	Full	Reduced	Season tickets	Total
Tweedbank	X	-	37,762	103,903	8,636	150,301	62,454	141,989	13,673	218,116
Galashiels	X	-	31,116	67,834	7,930	106,880	54,301	103,356	15,475	173,132
Stow	X	-	6,814	9,396	3,618	19,828	12,703	15,412	5,622	33,737
Gorebridge	X	-	9,258	15,203	5,191	29,652	17,531	23,744	7,834	49,109
Newtongrange	X	-	13,338	20,227	9,634	43,199	24,793	31,509	14,520	70,822
Eskbank	X	-	16,853	41,415	5,881	64,149	36,858	86,910	13,617	137,385
Shawfair		-	2,310	2,913	1,378	6,601	4,054	5,229	1,835	11,118
Newcraighall		121,379	46,773	48,688	16,552	112,013	49,588	49,867	17,930	117,385
Brunstane		82,266	40,873	24,146	17,970	82,989	40,002	24,598	16,437	81,037
Edinburgh		10,553,270	2,553,089	6,355,571	1,953,320	10,861,980	2,635,356	6,672,712	1,983,103	11,291,171
Edinburgh Park		446,764	186,510	118,732	139,488	444,730	184,031	117,344	133,614	434,989
Musselburgh		228,359	118,119	64,940	55,991	239,050	113,716	64,366	53,763	231,845

10.7.2 Cluster Analysis

Elbow Chart



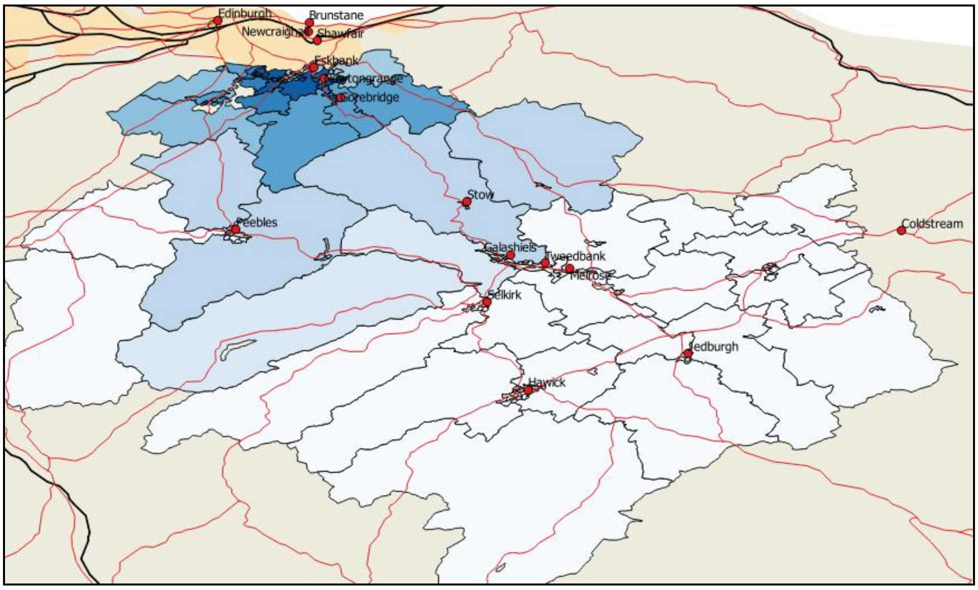
10.7.3 Dendrogram

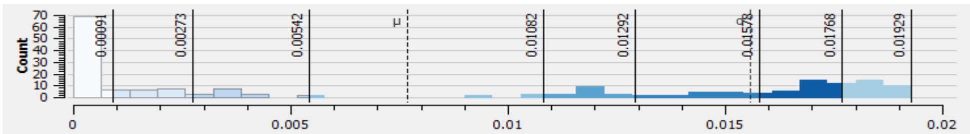


10.7.4 Job accessibility index - Spatial analysis

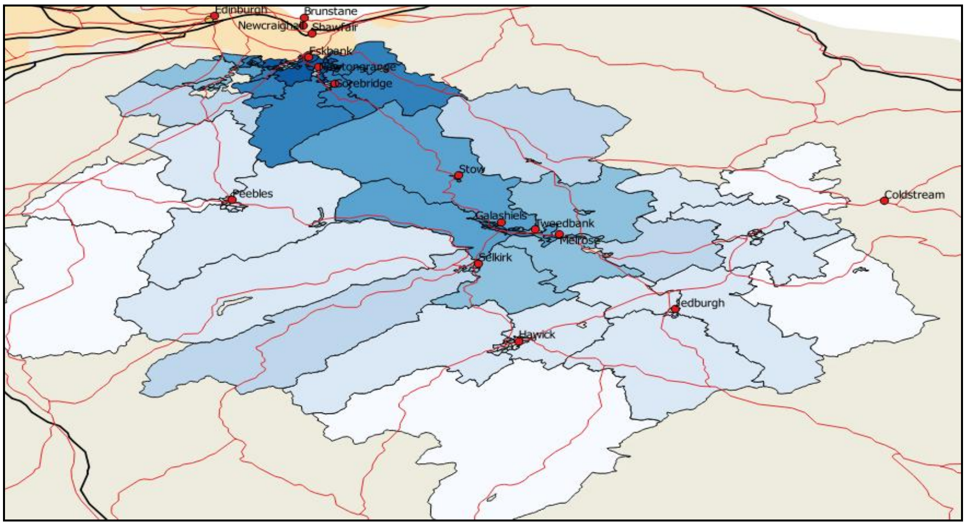
Cost basis with matching

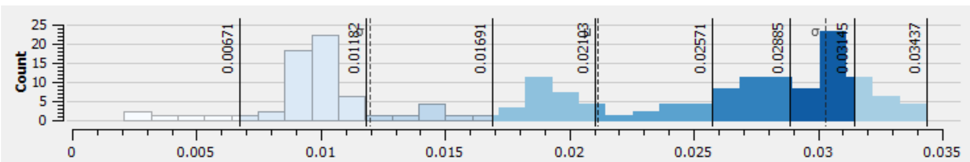
2011





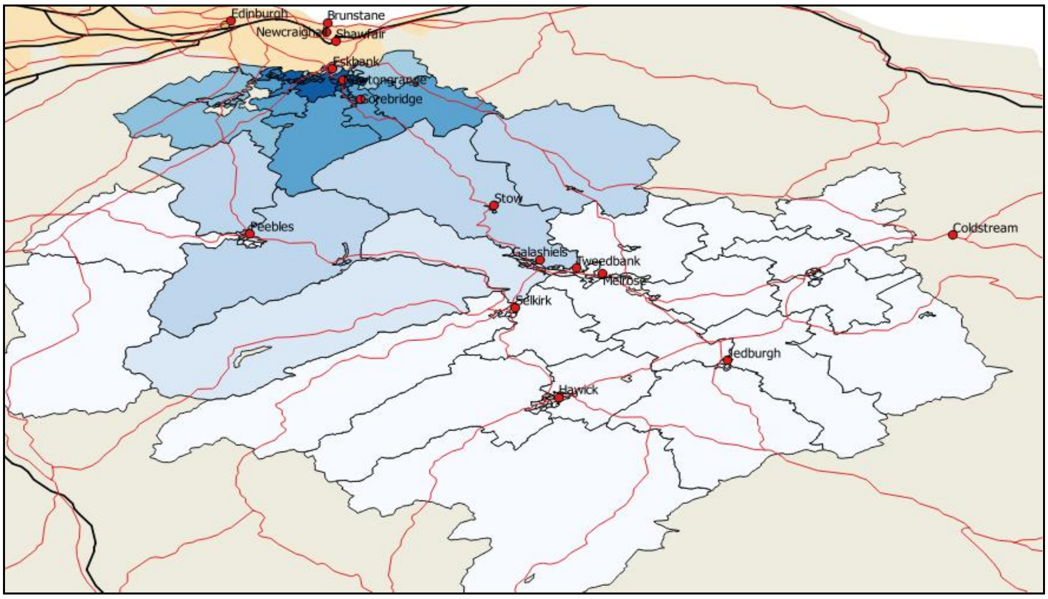
2017

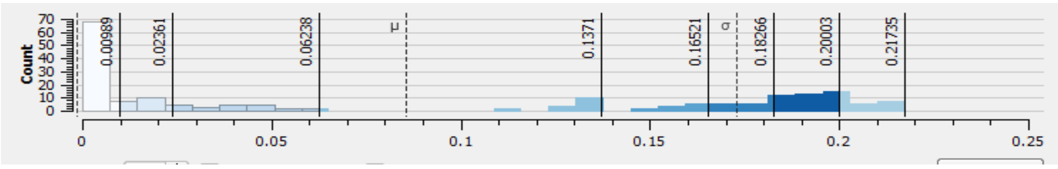




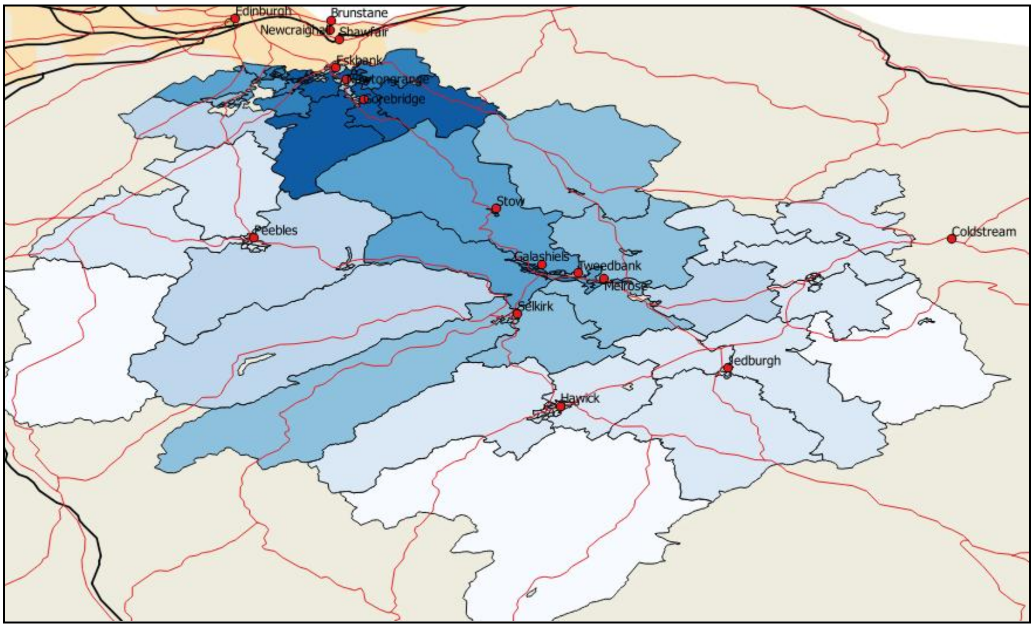
Cost basis without matching

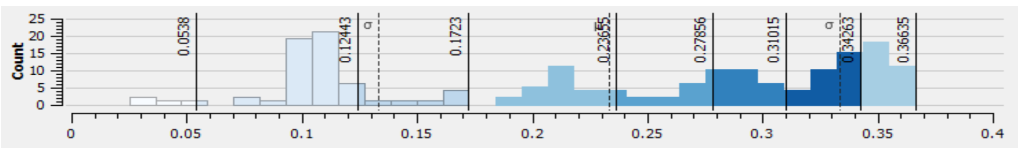
2011





2017





10.7.5 Essential services accessibility index

Time basis comparison

Year	Group	Nursery School	Primary School	Secondary School	Hospital	Average
2011	Treatment	0.835	0.886	0.730	0.602	0.763
	Control	0.603	0.858	0.714	0.571	0.686
2017	Treatment	0.890	0.921	0.825	0.738	0.843
	Control	0.709	0.903	0.809	0.705	0.782

Cost basis comparison

Year	Group	Nursery School	Primary School	Secondary School	Hospital	Average
2011	Treatment	0.847	0.901	0.736	0.602	0.772
	Control	0.607	0.871	0.720	0.572	0.692
2017	Treatment	0.871	0.917	0.777	0.658	0.806
	Control	0.647	0.891	0.760	0.625	0.731

10.7.6 GWR model output - property

Distance to nearest station

2011

	Minimum	1st Q	Median	3rd Q	Maximum	Global
Intercept	11.8800	11.9600	12.2700	12.5200	12.8200	12.3894
Distance to nearest station	-0.0090	-0.0055	-0.0041	-0.0022	-0.0003	-0.0050
Detached	0.0012	0.0014	0.0015	0.0017	0.0024	0.0011
Terraced	-0.0016	-0.0012	-0.0012	-0.0011	-0.0009	-0.0013
Semi	0.0003	0.0005	0.0008	0.0009	0.0011	0.0007
No car households	-0.0026	-0.0023	-0.0018	-0.0018	-0.0012	-0.0024
1 car household	0.0001	0.0003	0.0005	0.0007	0.0010	0.0006
Level 1	-0.0007	-0.0005	-0.0003	-0.0002	0.0001	-0.0003
Level 4	0.0001	0.0001	0.0002	0.0002	0.0003	0.0002
No qualifications	-0.0003	-0.0001	-0.0001	0.0000	0.0002	0.0000
Population density	-0.0106	-0.0099	-0.0098	-0.0098	-0.0092	-0.0096
% employed	-0.6934	-0.3341	0.0125	0.2025	0.5390	-0.1328
2017						
	Minimum	1st Q	Median	3rd Q	Maximum	Global
Intercept	11.3600	11.9300	12.0800	12.3600	12.7000	11.4482
Distance to nearest station	-0.0113	-0.0063	-0.0056	-0.0034	0.0053	0.0002
Detached	0.0002	0.0008	0.0012	0.0014	0.0023	0.0006
Terraced	-0.0020	-0.0014	-0.0012	-0.0007	-0.0003	-0.0018
Semi	0.0003	0.0007	0.0010	0.0012	0.0018	0.0013
No car households	-0.0052	-0.0048	-0.0034	-0.0029	-0.0023	-0.0038
1 car household	0.0017	0.0020	0.0022	0.0023	0.0028	0.0021
Level 1	-0.0002	-0.0002	-0.0001	0.0000	0.0002	-0.0001
Level 4	0.0001	0.0002	0.0003	0.0004	0.0005	0.0003
No qualifications	-0.0004	-0.0002	-0.0001	0.0000	0.0001	-0.0001
Population density	-0.0066	-0.0055	-0.0053	-0.0051	-0.0049	-0.0051
% employed	-1.6120	-1.0800	-0.6717	-0.4394	0.4294	0.2940

10.7.7 GWR model output - employment

Distance to nearest station

	Min	1st Q	Median	3rd Q	Max	Global
Intercept	3.0880	4.2510	4.7240	5.8110	6.3860	4.2914
Distance to nearest station	-0.0768	-0.0650	-0.0459	-0.0414	-0.0324	-0.0533
Detached	-0.0198	-0.0169	-0.0120	-0.0116	-0.0100	-0.0112
Terraced	-0.0129	-0.0121	-0.0101	-0.0088	-0.0038	-0.0078
Semi	-0.0059	-0.0044	-0.0023	-0.0018	0.0008	-0.0049
No car households	-0.0089	-0.0060	-0.0044	-0.0040	-0.0035	-0.0044
1 car household	0.0116	0.0120	0.0125	0.0140	0.0163	0.0121
Level 1	-0.0043	-0.0023	-0.0020	-0.0011	-0.0005	-0.0029
Level 4	-0.0013	-0.0011	-0.0008	-0.0004	0.0004	-0.0004
No qualifications	-0.0013	-0.0001	0.0003	0.0011	0.0017	0.0009
Population density	-0.0611	-0.0189	0.0061	0.0332	0.0441	0.0221

10.8 Borders Rail Stakeholder Survey 2015

10.8.1 Survey Respondents with key characteristics

ID	Sector	Type of organisation	Interest	HQ Location	How long operating (years)	Area cover	Who represent
1	TS	2-3	14	Borders	Over 20	B-E-S	5-6-7
2	CO	2-3	3-4-6-16	Scotland	10-20	B-E-G-S-O	1-2-3-6-7-10
3	CH	2-3-7	1-2-4-5-6-10-16	Borders	Over 20	B	1-2-3-6
4	HA	2-6-7	2-4-5-6-8-16	Borders	Over 20	B	2-3-6-7-10
5	CH	2-3	4-6-8-10-15	Scotland	Over 20	B	1-2-3-5-6-9-10
6	CH	1-3	1-2-4-5-6-9-10-15-16	Borders	Over 20	B	1-2-3-5-6-9-10
7	CH	3-4-5-9	2-3-4-6-16	Borders	Over 20	B	1-2-3-5-6-10
8	ED	5	1-2-3-4-9-10-12-14-15-16	Edinburgh	Over 20	B-M-E-S-U	2-5-6-7-9
9	OT	3-9	2-4-9-11-13-15	Borders	Over 20	B	1-2-5-6-7-8-9
10	TS	10	14-15	Borders	Over 20	B	7-8
11	TS	7	12	Borders	10-20 years	B-M-E-S	1-5-6-7-8-9
12	CH	2-3	2	Midlothian	less than 10	B	7
13	CH		1	Borders	10-20	B-M	6
14	CH	2-3-8-9	5-9-10-11-12-14-15	Scotland	Over 20	S	1-2-3-5-6-7-8-9-10
15	CO	3	2-4	Borders	10-20	B	1-5-6-7-8
16	HA	7	8	Borders N. England	Over 20	B-N	6-7
17	CH	2-3	7-8-16	Borders	less than 10	B	2-3-4-6-10
18	BG	2	9-10-15-16	Borders	Over 20	B	6
19	TS	2-3	14	Borders	Over 20	B	5-7
20	HA	2-6-7	2-4-5-6-8-16	Borders	Over 20	B	2-3-6-7-10

10.8.2 Coding for Respondents

Sectors	
BG	Business Groups
CH	Comm & Vol, Charities & NFP organisations
CO	Conservation/Environment/Heritage
ED	Education/University Centres
HA	Housing Associations
OT	Other
TS	Tourism and Sport

Type of organisation	
1	Business Group
2	Charity or “not for profit”
3	Community or Volunteer
4	Council
5	Educational Centre
6	Housing Association
7	Limited Company
8	Local Interest Group
9	Other
10	Sole Trader

Interest	
1	Child Care
2	Disabled
3	Education
4	Elderly
5	Environment
6	Health Issues
7	Homeless
8	Housing
9	Local Business
10	Local Interest
11	Other
12	Outdoor Activities
13	Security
14	Sport
15	Tourism
16	Unemployed

Who they represent	
1	Children
2	Disabled
3	Health issues
4	Homeless
5	Local Businesses
6	Local Residents
7	Non-Local Residents
8	Seasonal Visitors
9	Students
10	Unemployed

10.8.3 Survey Design as used in Borders Rail Stakeholder Survey

UNIVERSITY OF LEEDS

Midlothian and Scottish Borders Stakeholder Group Survey

Welcome to our survey!

This survey forms part of a Ph.D. research project through the Institute for Transport Studies at the University of Leeds to assess the impact on communities of the opening and re-opening of railway lines. It is designed to gauge attitudes and opinions over a range of travel and transport matters for different groups of stakeholders in the Borders and Midlothian regions in the light of the forthcoming launch of the new Borders Rail Edinburgh to Tweedbank service in September 2015. Please try to answer all sections as this will provide us with a more representative account of the range of views. Perhaps one individual could complete the survey, viewing matters from your organisation's perspective. You can be assured that your responses will be stored securely in line with University of Leeds policy and used for academic purposes only. A report on the findings will be sent to **Transport Scotland** to help them shape provision of better services for the local community, but no individual participants will be identifiable in the report. If you have any questions or would like more information, please contact us via email at **tsjsd@leeds.ac.uk**

Firstly we'd like some general information about your organisation.

1. Which of the following categories most closely describes the structure of your organisation? (Please answer more than one if applicable)

Please select at least 1 answer(s).

- [ ] Community or Volunteer
- [ ] Charity or "not for profit"
- [ ] Limited Company
- [ ] Sole Trader
- [ ] Business Group
- [ ] Housing Association
- [ ] Educational Centre
- [ ] Local Interest Group
- [ ] Other

1.a. If you selected Other, please specify:

2. Which of the following is your organisation most closely interested in? (Please answer more than one if applicable)

Please select at least 1 answer(s).

- [ ] Child Care
- [ ] Disabled
- [ ] Health Issues
- [ ] Unemployed
- [ ] Elderly
- [ ] Local Business
- [ ] Education
- [ ] Housing
- [ ] Tourism
- [ ] Conservation
- [ ] Heritage
- [ ] Environment
- [ ] Sport
- [ ] Outdoor Activities
- [ ] Local Interest
- [ ] Other

2.a. If you selected Other, please specify:

3. Where is the headquarters of your organisation located?

- Scottish Borders
- Midlothian
- Edinburgh
- Scotland
- UK
- Europe
- Globally
- Other

3.a. If you selected Other, please specify:

4. How long has your organisation been operating in this region?

- 0-3 years
- 4-9 years
- 10-20 years
- Over 20 Years
- Don't Know

And now some information about those your organisation represents

5. Which area(s) do you cover? (Please select more than one if applicable)

Please select at least 1 answer(s).

- Scottish Borders
- Midlothian
- Edinburgh
- Scotland
- UK
- Europe
- Global
- Other
- Don't Know

5.a. If you selected Other, please specify:

6. Which of the following groups describes those your organisation represents? (Please pick all that apply)

Please select at least 1 answer(s).

- Local Residents
- Non-Local Residents
- Seasonal Visitors
- Local Businesses
- Students
- Children
- Unemployed
- Disabled
- Health issues
- Other
- Don't Know

6.a. If you selected Other, please specify:

7. If applicable, please indicate the age range and gender of those you represent?

	0-11	12-17	18-24	25-34	35-44	45-54	55-64	65-74	Over 75	Don't Know	N/A
Male											
Female											

We would now like your views on the current usage and state of the transport system in Midlothian and the Borders

8. Please rank by **usage** the modes of transport used regularly by your organisation or those it represents from highest (6) to lowest (1)? Please don't select more than 1 answer(s) per row. Please don't select more than 1 answer(s) in any single column.

	1	2	3	4	5	6	Don't Know
Car	☐	☐	☐	☐	☐	☐	☐
Bus	☐	☐	☐	☐	☐	☐	☐
Rail	☐	☐	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐	☐	☐
Walking	☐	☐	☐	☐	☐	☐	☐
Cycling	☐	☐	☐	☐	☐	☐	☐

9. Overall, how would you assess **current** transport options in the region?

	Very Poor	Poor	Satisfactory	Good	Excellent	Don't Know
Car	○	○	○	○	○	○
Bus	○	○	○	○	○	○
Rail	○	○	○	○	○	○
Taxi	○	○	○	○	○	○
Walking	○	○	○	○	○	○
Cycling	○	○	○	○	○	○

9.a. Please add any general comments about how transport varies across your region

We would also like your opinion on accessibility issues in local transport

10. In this region how easy is it currently to access the following essential services using public or private transport? *Please try to answer this on behalf of your organisation rather than from personal experience.* (1 very difficult to 5 very easy)

	Public Transport							Private Transport						
	1	2	3	4	5	Don't Know	N/A	1	2	3	4	5	Don't Know	N/A
Primary School														
Secondary School														
Nursery/Crèche														
GP														
Hospital														
Post Office														
Library														
Bank														

10.a Please add any general comments about variations in access to essential services across the region

11. In this region how easy is it currently to access the following shopping and recreational facilities using public or private transport? *Please try to answer this on behalf of your organisation rather than from personal experience.* (1 very difficult to 5 very easy)

	Public Transport							Private Transport						
	1	2	3	4	5	Don't Know	N/A	1	2	3	4	5	Don't Know	N/A
Local Shops														
Supermarket														
Town Centre e.g. Galashiels														
City Centre e.g. Edinburgh														
Activity Centre														
Local Park														

11.a Please add any general comments about variations in access to shopping and recreational facilities across the region

12. How would you assess current **public transport** access provision for children, young families, the disabled and the elderly? (1 very poor to 5 very good)

	1	2	3	4	5	Don't Know	N/A
Information about available accessible transport							
Audio-visual information							
Announcements							
Visual information inside the vehicle							
Accessibility getting onto vehicles e.g. wide doors							
Vehicle layout							
Accessibility at stations							
Assistance with connections during journeys							
Timetables & information with disabled & elderly in mind							
Enforcement & provision of disabled parking							

12.a. Please add any additional comments about variations in access provision for children, young families, the disabled and the elderly across the region.

We're also interested to discover current attitudes towards using public transport, the car and walking & cycling.

13. What is your attitude to existing local **public transport**? Please assess your organisation's attitude to each statement.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
A satisfying experience	○	○	○	○	○	○
Affordable and good value	○	○	○	○	○	○
Convenient	○	○	○	○	○	○
Reliable	○	○	○	○	○	○
Time efficient	○	○	○	○	○	○
Good Service throughout the day	○	○	○	○	○	○
Good Level of Choice	○	○	○	○	○	○
Good connections throughout the region and Edinburgh	○	○	○	○	○	○
Good connections to hospitals and schools	○	○	○	○	○	○
Good level of information available	○	○	○	○	○	○
Reduces environmental impact	○	○	○	○	○	○

13.a. Please add any additional comments about attitudes to public transport across the region.

14. This question is about attitudes to **private transport**. Please assess your organisation's view on each statement.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
A satisfying experience						
Affordable and good value						
Convenient						
Reliable						
Time efficient						
The quality and network of local roads is good						
Has negative effects on health & fitness						
Contributes to problems like climate change						
There should be restricted car access in towns						
We should promote a reduction in car use						

14.a. Please add any additional comments about attitudes to private transport across the region.

15. Finally, a question about attitudes to **walking and cycling** in the region**.** Please indicate your level of agreement with each statement, even if you don't cycle or walk very often.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Don't Know
A satisfying experience						
Time efficient						
Reduces environmental impacts						
Benefit health and fitness						
Good facilities for cycles on public transport						
Good level of information about routes & facilities						
Good provision for walkers & cyclists e.g. cycle lanes						

15.a. Please add any additional comments about attitudes to walking and cycling across the region.

In this final section, we're interested in your awareness and attitude towards the forthcoming changes in the transport infrastructure

16. Are you aware that the new Borders Railway from Edinburgh to Galashiels and Tweedbank will begin operation this September?

- No, I have not previously heard of it
- I am aware of it, but know nothing about it
- Yes, but I have limited knowledge of it
- Yes, I am well informed on its progress
- Don't know

17. Do you think the opening of the Borders Railway is likely to have any impact on your organisation or its members?

- Not likely to be of benefit
- Unsure if and how it might benefit it
- May benefit it a little
- May benefit it a great deal
- Don't know

18. What difference do you think the introduction of Borders Rail will make in the following contexts?

	Much Worse	Worse	No difference	Better	Much Better	Don't know	N/A
Commuting to Edinburgh							
Shopping in Edinburgh							
Tourists from Edinburgh							
Local business prospects							
Job prospects							
Regional town centre shopping							
Regional Property Prices							
Consumer Spending							
Access to markets							
Schools /Higher Education							

18.a. Please add any additional comments regarding differences Borders Rail may make across the region.

19. Would you say the new rail link is a good facility to have available - even if you do not intend to or may never use it?

- ○ Definitely Not
- ○ Probably Not
- ○ Possibly
- ○ Probably
- ○ Definitely
- ○ Don't Know

20. Can you suggest any new initiatives or activities to promote the region through the new rail link?

21. Are there any additional comments you would like to make which relate specifically to your organisation?

Thank you for taking time to complete the survey, we greatly value your input. If you would be willing to take part in a repeat survey next year following the opening of the railway line, please fill in your details below and tick the appropriate boxes*. Also if you would like to receive a brief written report on the results of our work, please fill in your details below*. We will send you the report in autumn 2015.

22. I would be willing to:

- ☐ Take part in a repeat survey
- ☐ Receive written feedback on the research

10.8.4 Borders Rail Stakeholder survey results

	Modes of transport ranked by popularity				Assessment of current transport options in the region			
	Mean rank	SD	Lower Q	Upper Q	Mean rank	SD	Lower Q	Upper Q
Car	5.77	0.58	6	6	3.62	0.74	3	4
Bus	4.69	1.14	5	5	2.85	0.66	2	3
Rail	1.15	0.53	1	1	1.08	0.27	1	1
Taxi	2.08	0.28	2	2	2.27	1.05	2	4
Walking	3.23	0.7	3	4	3.85	0.53	4	4
Cycling	3.77	0.7	3	4	3.69	0.72	3	4
	Accessibility to essential services using public transport				**Accessibility to essential services using private transport**			
	Mean rank	SD	Lower Q	Upper Q	Mean rank	SD	Lower Q	Upper Q
Primary School	4	0.68	4	4	4.31	0.61	4	5
Secondary School	3.77	0.8	3	4	4.31	0.61	4	5
Nursery/Crèche	3.54	1.34	3	4	4.23	0.58	4	5
GP	3.38	1	3	4	3.85	0.53	4	4
Hospital	2.62	1.5	2	3	3.77	0.89	4	4
Post Office	3.54	1.15	3	4	4	0.68	4	4
Library	3.38	1.08	3	4	3.85	0.53	4	4
Bank	3.31	1.07	3	4	3.85	0.53	4	4
	Accessibility to leisure and shopping facilities using public transport				**Accessibility to leisure and shopping facilities using private transport**			
	Mean rank	SD	Lower Q	Upper Q	Mean rank	SD	Lower Q	Upper Q
Local Shops	3.62	0.62	3	4	4.08	0.27	4	4
Supermarket	3.23	0.7	3	4	4.08	0.27	4	4
Town Centres	3.38	0.84	3	4	4	0.39	4	4
City Centre e.g. Edinburgh	2.54	1.08	2	3	3.92	0.47	4	4
Activity Centres	3.23	0.7	3	4	3.92	0.27	4	4
Local Park	4	1.36	3	4	4.46	1.08	4	4

Attitudes to walking and cycling in the region				
	Mean rank	SD	Lower Q	Upper Q
A satisfying experience	4.15	0.66	4	5
Time efficient	2.85	0.86	2	4
Reduces environmental impacts	4.54	0.5	4	5
Benefit health and fitness	4.54	0.5	4	5
Good facilities for cycles on public transport	2.08	0.92	1	3
Good level of information about routes and facilities	3.46	0.84	3	4
Good provision for walkers and cyclists	3.31	0.99	2	4

Attitude to existing local public transport				
	Mean rank	SD	Lower Q	Upper Q
A satisfying experience	3.15	0.95	3	4
Affordable and good value	3.23	0.8	3	4
Convenient	2.77	0.97	2	3
Reliable	3.54	0.84	3	4
Time efficient	3.38	1	2	4
Good Service throughout the day	2.77	1.19	2	4
Good Level of Choice	2.54	0.93	2	3
Good connections throughout region and Edinburgh	2.62	1.21	2	3
Good connections to hospitals and schools	2.54	1.5	2	3
Good level of information available	3	1.18	2	4
Reduces environmental impact	3.69	0.91	3	4

Awareness of the new Borders Railway and its likely impact				
	Mean rank	SD	Lower Q	Upper Q
Awareness that the new Borders Railway from Edinburgh to Tweedbank will begin operation this September?	3.77	0.58	3	4
Do you think the opening of the Borders Railway is likely to have any impact on your organisation or its members?	3.08	0.73	3	3

What difference do you think Borders Rail will make in the following contexts?				
	Mean rank	SD	Lower Q	Upper Q
Commuting to Edinburgh	4.38	0.74	4	5
Shopping in Edinburgh	4.23	0.7	4	5
Tourists from Edinburgh	4.31	0.61	4	5
Local business prospects	4.46	0.84	4	5
Job prospects	4.31	0.99	4	5
Regional town centre shopping	3.38	0.84	3	4
Regional Property Prices	4.15	1.17	3	4
Consumer Spending	3.85	0.86	3	4
Access to markets	3.92	0.83	3	4
Schools /Higher Education	3.54	0.63	3	4

REFERENCES

Abadie, A. 2005. Semi-parametric Difference-in-Differences Estimators. *Review of Economic Studies*. **72**, pp.1–19.

Action with Communities in Rural England. 2014. *Transport - Affordable transport solutions for rural communities have never been more essential: ACRE Policy Position Paper 2014.* Action with Communities in Rural England.

Adair, A., McGreal, S., Smyth, A., Cooper, J. and Ryley, T. 2000. House prices and accessibility: The testing of relationships within the Belfast urban area. *Housing Studies*. **15**, pp.699–716.

Ahlfeldt, G. M. 2013. If We Build it, Will They Pay? Predicting Property Price Effects of Transport Innovations. *Environment and Planning A: Economy and Space*. **45**(8), pp.1977–1994.

Allard, S. and Danziger, S. 2002. Proximity and Opportunity: How Residence and Race Affect the Employment of Welfare. *Housing Policy Debate*. **13** (4), pp.675-700.

Alonso, W. 1964. *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge, Massachusetts: Harvard University Press.

Ashiru, O., Polak, J. W. and Noland, R. B. 2003. Space-Time User Benefit and Utility Accessibility Measures for Individual Activity Schedules. *Transportation Research Record*. **1854**(1), pp.62–73.

Attanasio, O., Kugler, A. and Meghir, C. 2011. Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*. **3**(3).

Audit Commission. 2009. *When it comes to the crunch…How councils are responding to the recession*. Audit commission.

Bae, C. H. C., Myung J.J. and Park, H. 2003. The impact of Seoul's subway Line 5 on residential property values. *Transport Policy*. **10**(2), pp.85-94.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M. and White, P., 2004. *The demand for public transport: a practical guide*. TRL Report 593. Crowthorne, UK:TRL.

Bania, N., Leete, L. and Coulton, C. 2008. Job Access, Employment and Earnings: Outcomes for Welfare Leavers in a US Urban Labour Market. *Urban Studies*. **45**(11)

Banister, D. and Berechman, J. 2000. *Transport Investment and Economic Development*. London: University College Press.

Banister, D. and Thurstain-Goodwin, M. 2011. Quantification of the non-transport benefits resulting from rail investment. *Journal of Transport Geography.* **19**(2), pp. 212–223.

Banister. D, and Thurstain-Goodwin. M. 2005. Property Values and Public Transport Investment. *Association for European Transport and contributors 2005.*

Basu, S. and Thibodeau, T. 1998. Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics.* **17**(1), pp 61–85.

Becker, G. 1965. A Theory of the Allocation of Time. *The Economic Journal*. **75**(299), pp. 493-517.

Ben-Akiva, M. and Lerman, S.R. 1979. Disaggregate travel and mobility choice models and measures of accessibility. In: Hensher, D.A., Sopher, P.R. ed(s). *Behavioural Travel Modelling.* Croom Helm: Andover, Hants, pp.654–679.

Berry, W. D. 1984. *Nonrecursive Causal Models*. London: Sage Publications.

Bertolini, L., le Clercq, F. and Kapoen, L. 2005. Sustainable accessibility: a conceptual framework to integrate transport and land use plan-making. *Transport Policy*. **12**(3), pp.207-220.

Bertrand, M., Duflo, E. and Mullainathan, S. 2004. How much should we trust Differences-in-Differences estimates?. *Quarterly Journal of Economics*. **119** (1), pp.249-275.

Bibby, P. and Shepherd, J. 2001. *Developing a New Classification of Urban and Rural Areas for Policy Purposes – the Methodology.* Department of Transport Local Government and the Regions (DTLR).

Billings, S. B. 2011. Estimating the value of a new transit option. *Regional Science and Urban Economics*. **41**, pp. 525–536.

Blainey, S.P. and Preston, J.M. 2010. Gateways to prosperity? The long term impacts of new local railway stations*. Association for European Transport and contributors* 2010.

Bowen, W. and Prestegaard, D. 2001. Theoretical and Empirical Considerations Regarding Space in Hedonic Housing Price Model Applications. *Growth and Change*. **32**(4), pp.466–490.

Bowes, D.R. and Ihlanfeldt, K. 2001. Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics*. **50**, pp.1-25.

Boyce, D., Allen, D., Mudge, R., Slater, P. and Isserman, A. 1972. *The impact of rapid rail on suburban residential property values and land development analysis of the Philadelphia high-speed line*. Philadelphia: Regional Science Department, Wharton School, University of Pennsylvania

British Railways Board.1963. *The Reshaping of British Railways*. British Railways Board.

British Railways Board. 1965. *The Development of the Major Railway Trunk Routes.* British Railways Board.

Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J. and Stürmer, T. 2006. Variable selection for propensity score models. *Am J Epidemiol*. **163**(12), pp. 1149-56.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. 1996. Geographically Weighted Regression: a method for exploring spatial nonstationarity. *Geographical Analysis.* **28,** pp.281–298.

Brunsdon, C., Fotheringham, A. S. and Charlton, M. 1998. Geographically Weighted Regression — modelling spatial nonstationarity. *The Statistician*. **47**, pp.431–443.

Bruzelius, N. 1979. *The value of travel time: theory and measurement*. London:Croom Helm.

Bunel, M. and Tovar, E. 2014. Key Issues in Local Job Accessibility Measurement: Different Models Mean Different Results. *Urban Studies*. **51**(6), pp.1322–1338.

Campaign for Better Transport. 2012. *The case for expanding the rail network*. Campaign for Better Transport.

Campaign to Protect Rural England. 2015. *Reopening Railways: the case for growing the rail network and how it can be achieved.* Greengauge 21.

Canning, S. , Leitham, S. and McCartney, P. 2015. *Evaluation of Laurencekirk Railway Station Final Report.* Transport Scotland.

Card, D. and Krueger, A. 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*. **84** (4).

Cervero, R. 1996. Paradigm Shift: From Automobility to Accessibility Planning. *Institute of Urban and Regional Development. Journal of the American Planning Association*. **65**(1), pp.96–111.

Cervero, R. and 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), pp.13-22.

Cervero, R. and Duncan, M. 2002a. Benefits of Proximity to Rail on Housing Markets: Experiences in Santa Clara County. *Journal of Public Transportation***. 5**(1), pp.1-18.

Cervero, R. and Duncan, M. 2002b. Land Value Impacts of Rail Transit Services in Los Angeles County. *National Association of Realtors, Urban Land Institute*.

Cervero, R. and Duncan, M. 2004. Neighbourhood Composition and Residential Land Prices: Does Exclusion Raise or Lower Values?. *Urban Studies*. **41**(2).

Cervero, R. and Duncan, M. 2001. Transit's value-added: effects of light and commuter rail services on commercial land values. *University of California, Urban Planning Department, mimeo.*

Cervero, R. and 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**, pp.13–22.

Cervero, R., Rood, T. and Appleyard, B. 1995. *Job Accessibility as a Performance Indicator: An Analysis of Trends and Their Social Policy Implications in the San Francisco Bay Area*. UC Berkeley: University of California Transportation Center.

Cervero, R., Rood, T. and Appleyard, B. 1997. Job accessibility as a performance indicator: An analysis of trends and their social policy implications in the San Francisco Bay Area*. Working paper 692, University of California, Berkeley*.

Cervero, R., Rood, T. and Appleyard, B. 1998. Tracking Accessibility: Employment and Housing Opportunities in the San Francisco Bay Area. *Environment and Planning A: Economy and Space*. **31**(7), pp.1259–1278.

Cheng, J. and Bertolini, C. 2013. Measuring urban job accessibility with distance decay, competition and diversity. *Journal of Transport Geography*. **30**, pp.100–109.

Church, A., Frost, M. and Sullivan, K. 2000. Transport and social exclusion in London. *Transport Policy*. **7**, pp.195-205.

Collits, P. 2000. Small Town Decline and Survival: Trends, Success Factors and Policy Issues. In: *Future of Australia's Country Towns Conference at La Trobe University, Bendigo.*

Conniffe, D. and Duffy, D. 1999. Irish House Price Indices Methodological Issues. *The Economic and Social Review*. **30**, pp.403-423.

Crespo, R., Fotheringham, A.S. and Charlton, M. 2007. Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models. In: Demšar, U. (Ed.), *Proceedings of the 9th International Conference on Geocomputation*. Maynooth, Ireland: NCG, NUI.

Cressie, N. 1993. *Statistics for Spatial Data*. New York: John Wiley and Sons.

Cropper, M., Mcconnell, K. and Deck, L. 1988. On the Choice of Functional Form for Hedonic Price Functions. *Review of Economics and Statistics* **70**(4), pp.668-75.

Currie, G. 2003. Public Transport Needs Gap Analysis. *Paper presented at the 2004 TRB Conference, January.*

Currie, G. and Stanley, J. 2010. Investigating Links between Social Capital and Public Transport*. Transport Reviews.* **28**(4), pp.529-547.

Czitrom, V. 2012. One-Factor-at-a-Time versus Designed Experiments. *The American Statistician*. **53**(2).

Daly, C., Duncan, G., McDonough, P. and Williams, D. 2003. Daly et al. respond. *Am J Public Health*. **93**(1), p.13.

Damm, D., Lerman, S. R., Lerner-Lam, E. and Young, J. 1980. Response of Urban Real Estate Values in Anticipation of the Washington Metro*. Journal of Transport Economics and Policy*. **14**, pp.315-336.

de Lannoy, W. and Van Oudheusden W. 1978. The accessibility of nodes in the Belgian road network***. GeoJournal January 1978***. **2**(1), pp.65–70.

Debrezion, G., Pels, E. and Rietveld, P. 2007. The impact of railway stations on residential and commercial property value: a meta-analysis*. Journal of Real Estate Finance and Economics*. **35**, pp.161–180.

Dehejia, R. H. and Wahba, S. 2002. Propensity Score-Matching Methods for Non-experimental Causal Studies. *The Review of Economics and Statistics 2002.* **84**(1), pp.151-161.

Department for Business Innovation and Skills. 2011. *Guidance on evaluating the impact of interventions on business.* Department for Business Innovation and Skills.

Department for Transport. 2004. *Guidance on Accessibility in Local Transport Plans*. Department for Transport

Department for Transport. 2008. *Tackling congestion on our roads - managing demand*. Department for Transport.

Department for Transport. 2014. National Travel Survey: 2014 Personal travel statistics for residents of England during 2014. Department for Transport

Department for Transport. 2012. *The wider impacts sub-objective, TAG Unit 3.5.14*. Department for Transport.

Department for Transport. 2016. *Understanding and Valuing the Impacts of Transport Investment*. Department for Transport.

Department for Transport. 2016. National Travel Survey: 2016. Department for Transport.

Department for Transport. 2018. *Passenger transport: by mode, annual from 1952*. Department for Transport.

Department of Transport.1982. *Serpell Committee Report on the Review of Railway Finances*. Department of Transport.

Devon County Council. 2014. *Tavistock to Bere Alston Rail Appraisal Specification Report*. Devon County Council.

Dewees, D. 1976. The effect of a subway on residential property values in Toronto. *Journal of Urban Economics*. **3**(4), pp.357-369.

DfT. 2005. *Transport, Wider Economic Benefits, and Impacts on GDP*. DfT.

DfT. 2014. *Transport investment and economic performance: Implications for project appraisal*. DfT 2014.

DfT. 2015. *Value for Money Framework Moving Britain Ahead*. DfT.

DfT. 2016a. *National Travel Survey: Household car ownership by region and Rural-Urban Classification: England, 2002/03 to 2016/17*. DfT.

DfT. 2016b. *Road congestion and reliability statistics 2012-2016*. Department for Transport.

Dorantes, L. M., Paez, A. and Vassallo, J. M. 2011. Analysis of House Prices to Assess Economic Impacts of New Public Transport Infrastructure: Madrid Metro Line 12. *Transportation Research Record.* **2245**(1), pp.131–139.

Dorling, D. 1993. The spread of negative equity. *JRF Housing Research Findings, (December, Ref: 101). York: Joseph Rowntree Foundation.*

Dowler, E. 2002. *Poverty Bites: Food Health and Poor Families*. Child Poverty Action Group.

Du, H. and Mulley, C. 2007a. The short-term land value impacts of urban rail transit: Quantitative evidence from Sunderland, UK. *Land Use Policy*. **24**(1), pp.223–233.

Du, H. and Mulley, C. 2007b. Transport accessibility and land value: a case study of Tyne and Wear. *RICS Research paper series.* **7**(3).

Du, H. and Mulley, C. 2012. Understanding spatial variations in the impact of accessibility on land value using geographically weighted regression. *The Journal of Transport and Land use*. **5**(2), pp.46–59.

Dubin, R. and Goodman, A. 1982. Valuation of education and crime neighborhood characteristics through hedonic housing prices. *Population and Environment*. **5**(3), pp.166–181.

Duncan, O.D. 1972. Unmeasured variables in linear panel models. *Sociological Methodology*. **4**, pp.36-82.

Dunphy, R. 1973. Transit Accessibility as a determinant of automobile ownership. In: *52nd Annual Meeting of the Highway Research Board*. Washington District of Columbia, United States

Eddington, R. 2006*. Delivering Choice and Reliability*. Department for Transport.

Efthymiou, D. and Antoniou, C. 2013. How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice*. 52(C), pp.1-22.

Everitt, B., Landau, S., Leese, M. and Stahl, D. 2011. *Cluster Analysis*. 5th ed. Chichester: Wiley.

Findeis, J. and Jensen, L. 1998. Employment Opportunities in Rural Areas: Implications for Poverty in a Changing Policy Environment. *American Journal of Agricultural Economics*. **80**(5), pp. 1000-1007.

Finkel, S. 1995. *Causal analysis with panel data. Sage university papers series. Quantitative applications in the social sciences, No. 07-105. Causal analysis with panel data*. Thousand Oaks, CA, US: Sage Publications, Inc.

Forrest, D., Glen, J. and Ward, R. 1995. The Impact of a Light Rail System on the Structure of House Prices: A Hedonic Longitudinal Study. *Journal of Transport Economics and Policy*. **30**(1), pp. 15-29.

Forrest, R., Kennett, T. and Leather, P. 1997. *Homeowners on new estates in the 1990s*. Bristol: The Policy Press.

Fotheringham, A.S. 1981. Spatial Structure and Distance-Decay Parameters. *Annals of the Association of American Geographers.* **71**(3), pp.425-436.

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

Fotheringham, S and O'Kelly, M. 1989. *Spatial interaction models: formulations and applications*. London: Klewer.

Freeman, A. 1993. *The Measurement of Environmental and Resource Values: Theory and Methods.* Washington, D.C.: Resources for the Future.

Gaffron, P., Hine, J. and Mitchell, F. 2000. *The Role of Transport on Social Exclusion in Urban Scotland Literature Review.* Scottish Executive - Central Research Unit.

Garrod, G. 1992. The amenity value of woodland in Great Britain: A comparison of economic estimates. *Environmental & Resource Economics*. **2**(4), pp.415-434.

Garrod, G. and Willes, K. 1994. An economic estimate of the effect of a waterside location on property values. *Environmental and Resource Economics*. **4**(2), pp. 209–217.

Gatzlaff, D. H. and Smith, M. T. 1993. The impact of the Miami Metrorail on the value of residences near station locations. *Land Economics*. **69**(1), pp.54-66.

Gelfand, A., Ghosh, S., Knight, J. and Sirmans, C. 1998. Spatio-Temporal Modeling of Residential Sales Data. *Journal of Business & Economic Statistics*. **16**(3), pp. 312-321.

Geurs, K.T. and Ritsema van Eck, J.R. 2001. *Accessibility measures: review and applications*. RIVM report 408505 006. Bilthoven: National Institute of Public Health and the Environment.

Geurs, K.T. and van Wee, B. 2004. Accessibility evaluation of land-use and transport strategies: review and research directions*. Journal of Transport Geography*. **12**, pp.127-140.

Geurs, K. T., Boon, W. and van Wee, B. 2009. Social impacts of transport: literature review and state-of-the-practice transport appraisal in the Netherlands and the United Kingdom. *Journal Transport Reviews.* **29**(1).

Geurs, K.T. and Ritsema van Eck, J.R. 2003. Accessibility evaluation of land-use scenarios: the impact of job competition land-use and infrastructure developments for the Netherlands. *Environment and Planning B.* **30**(1), pp.69–87.

Gibbons, S. and Machin, S. 2003. Employment and rail access: an evaluation of the wider benefits of transport improvements. *Department of Transport. 2003.*

Gibbons, S. and Machin, S. 2005. Valuing rail access using transport innovations. *Journal of Urban Economics*. **57**(1), pp.148-169.

Gibbons, S. and Machin, S. 2008. Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy*. **24**(1), pp.99-119.

Gibbons, S., Lyytikäinen, T., Overman, H. and Sanchis-Guarner, R. 2012. New road infrastructure: the effects on firms. *SERC Discussion paper 117.*

Glen, A. 2010. *Airdrie-Bathgate Rail Link: Reconnecting Communities.* Isle of Man: Lily Publications.

Goldberg, M. 1981. *Transportation systems and urban forms: performance measurement and data requirements*. Proceedings of the International Symposium on Surface Transportation Performance. pp. 256–268.

Gourvish, T.R. 1974. *British Railways 1948–73: A Business History.* Cambridge: Cambridge University Press.

Gower, J. and Ross, J. 1969. Minimum Spanning Trees and Single Linkage Cluster Analysis. *Journal of the Royal Statistical Society. Series C (Applied Statistics).* 18(1), pp. 54-64.

Graham, D. J. 2006. *Wider economic benefits of transport improvements: Link between agglomeration and productivity*. Stage 2 Report, Department for Transport, London.

Graham, D.J. and Dender, K. 2011. Estimating the agglomeration benefits of transport investments: some tests for stability. *Transportation, Springer*. **38**(3), pp. 409-426.

Grass, R.G. 1992. The estimation of residential property values around transit station sites in Washington, D.C. *Journal of Economics and Finance*. **16**, pp.139–146.

Green, A. and Owen, D. 2006. *The geography of poor skills and access to work.* York: Joseph Rowntree Foundation.

Greene, W. 2008. Fixed and Random Effects Models for Count Data - *NYU Working Paper No. 2451/26046. Journal of Mixed Methods Research January 1, 2008.*

Grengs, J. 2010. Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography*. **18**(1), pp.42-54.

Griliches, Z. 1961. Hedonic Price Indices for Automobiles, in: Price Indices and Quality Change: *Studies in New Methods of Measurement*. 1961 ed., ed. Z Griliches. Cambridge: Harvard University Press.

Halden, D. 2011. The Use and Abuse of Accessibility Measures in UK Passenger Transport Planning. *Research in Transportation Business and Management*. **2**, pp.12–19

Handy, S. L., 1993. *Regional versus Local Accessibility: Neo-Traditional Development and its Implications for Non-Work Travel*, Built Environment, **18**(4), pp. 253–267.

Handy, S.L. and Niemeier, D. A. 1997. Measuring Accessibility: An Exploration of Issues and Alternatives. *Environment and Planning*. **29**(7), pp.1175-1194.

Hansen, W. G. 1959. How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*. **35**(2), pp.73-76.

Harris, B. 2001. Accessibility: concepts and application. *Bureau of Transportation Statistics, U.S. Dept. of Transportation, Washington, D.C*

Hartigan, J.A. 1975. *Clustering Algorithms*. New York: John Wiley and Sons.

Hausman, J. 1978. Specification Tests in Econometrics. *Econometrica*. **46**(6), pp.1251-71.

Henneberry, D. 1998.Transport investment and house prices. *Journal of Property Valuation and Investment*. **16**(2), pp.144-158.

Hensher, D. and Brewer, A. 2001. *Transport: an economics and management perspective*. Oxford: Oxford University Press, 2001.

Hills, D. 2010. *Logic mapping: hints and tips for better transport evaluations*. London: Tavistock Institute.

HM Treasury. 2011. *The Magenta Book*. HM Treasury*.*

HM Treasury. 2018. *The Green Book - appraisal and evaluation in central government*. HM Treasury.

Holland, P.W. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association.* **81**(396), pp.945-960.

Holly, S., Hashem, M., Yamagata, T. 2006. A Spatio-Temporal Model of House Prices in the U*. CESifo Working Paper No. 1826.*

Horner, M.W. and Mefford, J.N. 2005. GIS-based strategies for measuring worker accessibility to job opportunities: The case of bus transit. *Proceedings of the 84th Annual Meeting of the Transportation Research Board, Washington, DC.*

House of Commons Environmental Audit Committee. 2013. *Transport and accessibility to public services - Third Report*. House of Commons Environmental Audit Committee.

Huang, L., Stinchcomb, D.G., Pickle, L.W., Dill, J. and Berrigan, D. *Identifying Clusters of Active Transportation Using Spatial Scan Statistics*. *Am J Prev Med.* **37**(2), pp.157–166.

Ihlanfeldt, K. 1993. The Spatial Mismatch between Jobs and Residential Locations within Urban Areas. *Cityscape: A Journal of Policy Development and Research*. **1**(1), pp.219–44.

Ihlanfeldt, K. and Sjoquist, D. 1998. The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform. *Housing Policy Debate*. **9**(4), pp. 849-892.

Imbens, G.W. and Angrist, J.D. 1994. Identification and Estimation of Local Average Treatment Effects*. Econometrica*. **62**(2), pp.467-475.

Jara-Diaz, S.R. 1986. On the relation between users' benefits and the economic effects of transportation activities. *Journal of Regional Science*. **26**(2). pp. 379-391.

Johnson R.C. 2006. Landing a job in urban space: The extent and effects of spatial mismatch. *Regional Science and Urban Economics.* **36** (2006) pp.331–372

Johnston, K. and Causley, J. 2014. Take that Doctor Beeching: The case for the rebirth of Borders Rail. In: *STAR Conference 2013*.

Kain, J. 1968. Housing Segregation, Negro Employment, and Metropolitan Decentralization. *The Quarterly Journal of Economics*. **82**(2), pp.175–197.

Kain, J. 1993. The spatial mismatch hypothesis: three decades later. *Housing Policy Debate*. **3**, pp.371-460.

Kain, J. and Quigley, J. 1970. Measuring the value of housing quality. *Journal of the American Statistical Association*. **65**, pp. 532-548.

Kawabata, M. and Shen, Q. 2007. Commuting inequality between cars and public transit: The case of the San Francisco Bay Area, 1990-2000. *Urban Studies*. **44**(9).

Kenny, D. A. 1975. A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. *Psych. Bull*. **82**: pp.345-362.

Knaap, T., Oosterhaven, J., Ruijgrok, C. and Tavasszy, L.A. 2001. On the development of RAEM: the Dutch spatial general equilibrium model and its first application. In: *41th European Regional Science Association Conference, Zagreb, 2001.*

Koenig, J. G. 1980. Indicators of Urban Accessibility: Theory and Application. *Transportation*. **9**, pp.145–172.

Kohler, U. and Kreuter, F. 2012. *Data Analysis Using Stata.* 3rd ed. Stata Press.

Korsu, E. and Wenglenski, S. 2010. Job Accessibility, Residential Segregation and Risk of Long-term Unemployment in the Paris Region. *Urban Studies*. **47**(11), pp.2279-324.

Kremer, M. 2003. Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. *American Economic Review Papers and Proceedings*. **93**(2), pp.102-115.

Kwok, R. and Yeh, A. 2004. The Use of Modal Accessibility Gap as an Indicator for Sustainable Transport Development. *Environment and Planning A: Economy and Space*. **36**(5).

Laird, J.J., Nellthorp, J. and Mackie, P.J. 2005. Network effects and total economic impact in transport appraisal. *Transport Policy*. **12**(6), pp.537–544.

Laird, J.J. and Mackie, P. 2009. *Review of Economic Assessment in Rural Transport Appraisal - Research Findings.* Scottish Government.

Laird, J.J. and Mackie, P.J. 2013. Wider economic benefits of transport schemes in remote rural areas. *Research in Transportation Economics*. **47**, pp.92 - 102 (11).

Laird, J., Kelly, C. and Nellthorp, J. 2013. Ex Post Appraisal: What lessons can be learnt from EU Cohesion funded transport projects. *Paper presented at the European Transport Conference, Glasgow, October 2012.*

Laird, J.J., Nash, C. and Mackie, P.J. 2014. Transformational transport infrastructure: cost benefit analysis challenges. *The Town Planning Review*. **85**(6), pp.709 - 730.

Laird, J.J. and Venables, A.J. 2017. Transport investment and economic performance: A framework for project appraisal, *Transport Policy*, **56**, 2017, pp.1-11.

Lakshmanan, T.R. 2010. The broader economic consequences of transport infrastructure investments. *Journal of Transport Geography*. **19**(1), pp.1-12.

Lancaster, K.J. 1966. A New Approach to Consumer Theory. *The Journal of Political Economy*. **74**(2), pp.132-157.

Land Registry. 2017. *House price index* Downloaded from http://www.landregistry.gov.uk

Landis, J., Guhathakurta, S. and Zhang, M. 1994. Capitalization of Transit Investments into Single-Family Home Prices. *The University of California Transportation Center*.

Leventhal, T. and Brooks-Gunn, J. 2000. The neighborhoods they live in: The effects of neighbourhood residence on child and adolescent outcomes. *Psychological Bulletin*. **126**, pp.309-337.

Levinson, D. and Kumar, A. 1994. Multimodal Trip Distribution: Structure and Application. *Transportation Research Record*. **1466**, pp.124 – 131.

Li, M.M. and Brown, H.J. 1980. Micro-Neighborhood Externalities and Hedonic Housing Prices. *Land Economics.* **56**(2), pp.125-141.

Linn, R. L. and Werts, C. E. 1977. Analysis implications of the choice of a structural model in the nonequivalent control group design. *Psychological Bulletin*. **84**(2), pp. 229-234.

Lucas, K., van Wee, B. and Maat, K. 2015. A method to evaluate equitable accessibility: combining ethical theories and accessibility-based approaches. *Transportation*. **43**(3), pp. 473–490.

Lucas, K., Fuller, S., Psaila, A. and Thrush, D. 2004. *Representing neighbourhood environmental concerns within Local Strategic Partnerships*. York: Joseph Rowntree Foundation.

Lucas, K., Mattioli, G., Verlinghieri, E. and Alvaro Guzman. 2016. Transport poverty and its adverse social consequences. *Proceedings of the Institution of Civil Engineers-Transport.* **169**(6), pp. 353-365.

Lynch, A. and Rasmussen, D. 2001. Measuring the impact of crime on house prices. *Applied Economics.* **33**(15), pp.1981-1989.

Ma, K-R. and Banister, D. 2006. Extended excess commuting: a measure of the jobs-housing imbalance in Seoul. *Urban Studies*. **43**(11), pp.2099–2113.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob*. Vol. 1 (Univ. of Calif. Press, 1967), pp.281-297.

Markovich, J. and Lucas, K. 2012. *The Social and Distributional Impacts of Transport: A Literature Review.* Oxford: Transport Studies Unit School of Geography and the Environment.

Mata, A., Raymond, J-L. and Roig, J.L. 2009. Car ownership and access to jobs in Spain. *Transportation Research Part A: Policy and Practice.* **43**(6), pp. 607-617.

Mata, A., Raymond, J-L. and Roig, J.L. 2010. Job Accessibility and Female Employment Probability: The Cases of Barcelona and Madrid. *Urban Studies*. **47**(4), pp.769–787.

Mayor, K., Lyons, S., Duffy, D. and Tola, R. 2008. *A* Hedonic Analysis of the Value of Rail Transport in the Greater Dublin Area. *Working Paper ESRI November 2008*

McCahill, C. and Ebeling, M. 2015. Tools for measuring accessibility in an equity framework. *Congress for the New Urbanism 23rd Annual Meeting Dallas, Texas April 29 – May 2, 2015.*

McCartney, P., Lowe, S., Worsley, T., Hall, K. and Mackie, P. 2013. *Assessment of Methods for Modelling and Appraisal of the Sub-National, Regional and Local Economy Impacts of Transport*. Report to the Department for Transport.

McKenzie, R. P. 1984. The Measurement of Accessibility to Employment. *TSU Ref. 245, Transport Studies Unit, Oxford University.*

McNicholas, P. D. and Murphy, T.B. 2010. Model-based clustering of longitudinal data. *The Canadian Journal of Statistics*. **38**(1), pp.153-168.

Miller, N.G. (1982) Residential property hedonic pricing models: A review. In: *C.F. Sirmans (Ed.) Urban housing markets and property valuation, Research in Real Estate* 2 pp. 31-56.

Mohammad, S., Graham, D.J., Melo, P. and Anderson, R. 2013. A Meta-Analysis of the Impact of Rail Projects on Land and Property Values. In: *1st European Symposium on Quantitative Methods in Transportation Systems, 4-8th September 2012*.

Mohring Jr, H. and Williamson, H. 1969. Scale and Industrial Reorganisation Economies of Transport Improvements. *Journal of Transport Economics and Policy*. **3(**3), pp. 251-271.

Mohring, H. 1976. Transportation Economics. *Transportation Economics*, Ballinger Publishing Co. The Engineering Economist. **23**(3), pp.205-207.

Mohring, H. 1993. Maximizing, measuring, and not double counting transportation-improvement benefits: A primer on closed-and open-economy cost-benefit analysis. *Transportation Research Part B: Methodological.* **27**(6), pp.413-424.

Mott MacDonald. 2010. *Regional and Local Economic Impacts of Rail Investments* Final Report Welsh Government.

Mott MacDonald. 2014. *Wider Economic Benefits of a Rail Service between March and Wisbech*. Report to Cambridgeshire County Council.

Network Rail. 2008. *Lewes-Uckfield Railway Line Reinstatement Study*. Network Rail.

Noonan, D.S. 2005. Neighbours, Barriers and Urban Environments: Are Things 'Different on the Other Side of the Tracks'? *Urban Studies*. **42**(10).

Nottinghamshire County Council. 2015. *JSNA: The People of Nottinghamshire*. Nottinghamshire County Council.

Nutley, S. 2003. Indicators of Transport and Accessibility Problems in Rural Australia. *Journal of Transport Geography*. **11**, pp.55-71.

OECD. 2002. *Impact of Transport Infrastructure Investment on Regional Development*. Organisation for Economic Co-operation and Development, Paris

UK Census. 2001. *Census aggregate data.* Office for National Statistics via UK Data Service (Edition: May 2011).

UK Census. 2011. *Census aggregate data*. Office for National Statistics via UK Data Service (Edition: June 2016).

Office for National Statistics. 2018. *Annual Population Survey, 2004-2017.* UK Data Service, 13th Edition

Olsson, G. 1980. *Birds in Egg/Eggs in Bird.* London: Pion Limited.

Ong, P.M. and Miller, D. 2005. Spatial and Transportation Mismatch in Los Angeles. *Journal of Planning Education and Research*. **25**(1).

Oosterhaven, J. and Elhorst, J. 2003. Effects of Transport Improvements on Commuting and Residential Choice. In: *ERSA conference papers ersa03 p29, European Regional Science Association*.

Oosterhaven, J. and Knaap, T. 2003. *Spatial economic impacts of transport infrastructure investments. Transport projects, programmes and policies: Evaluation needs and capabilities*. Pearman, A., Mackie, P. & Nellthorp, J. (eds.). London: Ashgate Publishing, pp. 87 - 105.

Ottensmann, J., Payton, S., and Man, J. 2008. Urban Location and Housing Prices within a Hedonic Model. *JRAP*. **38**(1), pp.19-35.

Pace, R.K., Barry, R. and Sirmans, C.F. 1998. Spatial Statistics and Real Estate. The *Journal of Real Estate Finance and Economics.* **17**(1), pp.5-13.

Paez, A. and Scott, D.M. 2004. Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal.* **61**, pp.53–67.

Paez, A. and Scott, D.M. and Morency, C. 2012. Measuring accessibility: positive and normative implementations of various accessibility indicators. *Journal of Transport Geography.* **25**(0).

Peugh, J.L. and Enders, C.K. 2004. Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement. *Review of Educational Research*. **74**(4).

Powe, N.A., Garrod, G.D., Brunsdon, C. F. and Willis, K. G. 1995. Using a geographic information system to estimate an hedonic price model of the benefits of woodland access. *Forestry*. **70**(2), pp.139-149.

Reggiani, A., Bucci, P. and Russo, G. 2011. Accessibility and Impedance Forms: Empirical Applications to the German Commuting Network. *International Regional Science Review*. **34**(2), pp.230-252.

Rogers, C.L. 1997. Job Search and Unemployment Duration: Implications for the Spatial Mismatch Hypothesis. *Journal of Urban Economics*. **42**(1), pp.109-132

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy.* **82**, pp.34-55.

Rosenbaum, P.R. and Rubin, D.B. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. **70**(1), pp.41–55.

Rosiers, F., Lagana, A. and Theriault, M. 2001. Size and proximity effects of primary schools on surrounding house values. *Journal of property research*. **18**(2).

SACTRA. 1999. *Transport and the economy: full report*. SACTRA

Sanchez, T. W., Shen, Q. and Peng, Z.-R. 2004. Transit mobility, jobs access and low income labour participation in US metropolitan areas*. Urban Studies*. **41**(7), pp.1313–1331.

Sarkar, A.K. and Mashiri, M. 2002. Quantification of Accessibility Levels of Rural areas: A case study in the Northern Province of South Africa. *Journal of the Indian Roads Congress, December, New Delhi, India*. pp.37-51.

Savigear, F. 1967. A Quantitative Measure of Accessibility. *Town Planning Review*. **38**, pp. 64–72.

Saxena, A., Prasad, M., Gupta, A., Bharill., N., Patel, O., Tiwari, A., Er, M., Ding. W. and Lin, C-T. 2017. A Review of Clustering Techniques and Developments. *Neurocomputing (2017).*

Scottish Executive Poverty Inclusion Working Group. 2001. *Poverty and Social Exclusion in Rural Scotland.* Scottish Government.

Scottish Government. 2016. The Scottish Index of Multiple Deprivation (2012 and 2016). The Scottish Government.

Scottish Government. 2015. Scottish household survey 2015: annual report. Scottish Government.

SELRAP. 2014. *Skipton-Colne Rail Link Phase 1b Needs Assessment and Appraisal*. SELRAP.

Shen, Q. 1998. Location characteristics of inner-city neighbourhoods and employment accessibility of low-wage workers. *Environment and Planning B*. **25**(3), pp.345-366.

Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. and Bushway, S. 1997. *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: U.S. Department of Justice. National Institute of Justice.

Shi, A., Laurent, E.J., LeBouton, J. and Liu, J. 2006. Local spatial modeling of white-tailed deer distribution. *Ecological Modelling* **190**(1-2): pp.171-189.

Shindler, R. and Ferrari, M.G. 1967. Auto ownership as affected by transportation. *System Traffic Engineering.* October 24-28.

Simmonds, D. C .and Feldman, O. 2011. Alternative approaches to spatial modelling. *Research in Transportation Economics.* **31**(1), pp.2-11.

Skov-Petersen, H. 2001. Estimation of distance-decay parameters - GIS-based indicators of recreational accessibility. *ScanGIS*. 2001. pp.237-258

So, H.M., Tse, R.Y.C., and Ganesan,S. 1997. Estimating the influence of transport on house prices: evidence from Hong Kong. *Journal of Property Valuation and Investment*. **15**(1), pp.40-47.

Social Exclusion Unit. 2003. *Making the Connections: Final report on Transport and Social Exclusion*. London: Office of the Deputy Prime Minister.

Speak, S. and Graham, S. 2000. *Service not included? Social implications of private sector service restructuring in marginalised neighbourhoods*. Bristol: The Policy Press.

Stanley, J. and Vella-Brodrick, D. 2009. The Usefulness of Social Exclusion to Inform Social Policy in Transport. *Transport Policy*. **16**, pp.90-96.

Steiner, P. and Cook, D. 2013. Matching and propensity scores. *Sociological Methods & Research*. **46**(2).

Steinley, D. 2006. K-Means Clustering: A Half-Century Synthesis. *British Journal of Mathematical and Statistical Psychology*. **59**(1), pp.1-34.

Stokes, G. 2002. *Strategies to change people's travel behaviour - where do we start?*. Paper to PTRC conference on Changing Travel Behaviour, London.

Stopher, P. and Greaves, S. 2006. Guidelines for samplers: measuring a change in behaviour from before and after surveys. *Transportation*. **34**(1), pp.1–16.

SYSTRA. 2015. *Larkhall – Milngavie Railway Project Evaluation Study Final Report*. Transport Scotland.

Tavasszy, L.A., Thissen, M., Muskens, J. and Oosterhaven, J. 2002. Pitfalls and solutions in the application of spatial computable general equilibrium models for transport appraisal. In: *42nd Congress of the European Regional Science Association, 2002, Dortmund, Germany.*

Taylor, P.J. 1971. Distance Transformation and Distance Decay Functions. *Geographical Analysis.* **3**(3), pp. 221-238.

The Consumer Council. 2013. *Transport Services funded by the Rural Transport Fund and the Transport Programme for People with Disabilities: Views from passengers, operators and stakeholders.* The Consumer Council.

The Independent Rail Consultancy Group. 2004. *Rail re-opening toolkit: A complete guide to the re-opening of a disused railway*. The Countryside Agency.

Thibodeau, T.G. 2003. Marking Single–Family Property Values to Market. *Real Estate Economics*. **31**(1), pp.1-22.

Tischer, M.L. and Shea C.G. 1978. Attitude-Behaviour Changes in a Before-After Mode Choice Situation, in NA - *Advances in Consumer Research*. Volume 05, eds. Kent Hunt, Ann Abor, MI: Association for Consumer Research: pp.456-464.

Tischer, M.L. and Shea, C.G. 1997. Attitude-Behaviour Changes in a Before-After Mode Choice Situation. *Advances in Consumer Research.* **05**, pp. 456-464.

Titheridge, H., Christie, N., Mackett, R., Hernandez, D. and Ye, R. 2014. Transport Poverty: Concepts, Measurement and Policy Perspectives. In: *the Association of American Geographers' Annual Meeting, Tampa, Florida, USA. 8th-12th April 2014*.

Transport Scotland. 2008. *Scottish Transport Appraisal Guidance (STAG).* The Scottish Government, Edinburgh 2008.

Tse, R. Y. C. 2002. Estimating neighbourhood effects in house prices: Towards a new hedonic model approach. *Urban Studies*. **39**(7), pp.1165-1180.

Van Wee, B., Hagoort, M. and Annema, J.A. 2001. Accessibility Measures with Competition. *Journal of Transport Geography*. **9**, pp.199-208.

Venables, A. and Gasiorek, M. 1999. *The Welfare Implications of Transport Improvements in the Presence of Market Failure: The Incidence of Imperfect Competition in UK Sectors and Regions*. London: DETR.

Venables, A., Laird, J. J. and Overman, H. G. 2014. Transport investment and economic performance: Implications for project appraisal. *UK Department for Transport.*

Venables, A.J. 2007. Evaluating Urban Transport Improvements: Cost-Benefit Analysis in the Presence of Agglomeration and Income Taxation. *Journal of Transport Economics and Policy*. **41**(2), pp.173-188.

Vera-Toscano, E., Phimister, E. and Weersink, A. 2004. Panel estimates of the Canadian rural/urban women's wage gap*. American Journal of Agricultural Economics*. **86**, pp.1138-1151.

Vickerman, R.W. 1974. Accessibility, attraction and potential: A review of some concepts and their use in determining mobility. *Environment and Planning*. **6A,** pp. 675-691.

Vickerman, R. 2007. Recent Evolution of Research into the Wider Economic Benefits of Transport Infrastructure Investments. *OECD Statistics Directorate. OECD/ITF Joint Transport Research Centre Discussion Papers 2007-9*. (Dec 2007).

Visser, P., Van Dam, F. and Hooimeijer, P. 2008. Residential Environment and Spatial Variation in House Prices in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie*. **99**(3), pp.348-360.

Wachs, M. and Kumagi, T. G. 1973. Physical Accessibility as a Social Indicator. *Socioeconomic Planning Science*. **7**, pp. 437 – 456.

Wang, D., Tsui, A., Zhang, Y. and Ma, L. 2003. Employment relationships and firm performance: evidence from an emerging economy. *Tijdschrift voor Economische en Sociale Geografie*. **99**(3), pp.348-360.

Wardman, M., Batley, R., Laird, J.J., Mackie, P., Fowkes, T., Lyons, G., Bates, J. and Eliasson, J. 2013. *Valuation of Travel Time Savings for Business Travellers*. Report prepared for the Department of Transport, Institute for Transport Studies, University of Leeds, Leeds.

Wegener, M. and Fürst, F. 1999. *Land-Use Transport Interaction: State of the Art*. Deliverable D2a of the project TRANSLAND (Integration of Transport and Land use Planning).

Weibull, J. W. 1976. An Axiomatic Approach to the Measurement of Accessibility. *Regional Science and Urban Economics*. **6**, pp. 357 – 379.

Weinberger, R. 2001. Light rail proximity: Benefit or detriment? - The case of Santa Clara County, California*. Proceedings of the 80th Annual Meeting of the Transportation Research Board, Washington, DC.*

Weinstein, B. and Clower, T. 1999. *The Initial Economic impacts of the DART LRT System.* University of North Texas, Center for Economic Development and Research.

Weisbrod, G. 2016. Estimating Wider Economic Impacts in Transport Project Prioritisation Using Ex-Post Analysis. *OECD.*

What Works Centre for Local Economic Growth. 2015. What Works-Evidence Review 7 Transport July 2015. What Works Centre for Local Economic Growth.

Wheeler, D. and Tiefelsdorf, M. 2005. Multi-collinearity and correlation among local regression coefficients in geographically weighted regression. *J Geograph Syst*. (2005) **7**, pp.161–187.

Wilhelmsson, M. 2002. Spatial Models in Real Estate Economics. *Journal Housing, Theory and Society.* **19**(2).

Wilson, A. G. 1971. *A* Family of Spatial Interaction Models, and Associated Developments. *Environment and Planning*. **3**, pp.1–32.

Worsley, T. 2014. Ex-post Assessment of Transport Interventions and Policy Interventions. Roundtable Summary and Conclusions. *OECD ITF Discussion paper 2014 /19.*

Wrenn, D., Sam, A. and Irwin, E. 2012. Searching for the Urban Fringe: Exploring Spatio-Temporal Variations in the Effect of Distance versus Local Interactions on Residential Land Conversion Using a Conditionally-Parametric Discrete-Time Duration Model. In: *2012 Annual Meeting, August 12-14, 2012, Seattle, Washington from Agricultural and Applied Economics Association*.

Wu, K., Liu, B., Huang, B. and Lei, Z. 2013. Incorporating the multi-cross-sectional temporal effect in Geographically Weighted Logit Regression. In: *International Conference on Information Systems and Computing.*Technology (ISCT). London: CRC Press, Taylor & Francis Group, pp. 3–14.

Wu, M., and Hine J. P. 2003. A PTAL approach to measuring changes in bus service accessibility. *Transport Policy*. **10**(4), pp. 307–320 (Transport and Social Exclusion).

Yu, D.L. 2010. Exploring Spatiotemporally Varying Regressed Relationships: The Geographically Weighted Panel Regression Analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. **38**(2).

Zhang, L. and Shi, H. 2004. Local Modeling of Tree Growth by Geographically Weighted Regression. *Forest Science*. **50**(2), pp. 225-24

Zondag, B. and Pieters, M. 2005. Influence of accessibility on residential location choice*. 84th Annual Meeting of the Transportation Research Board, Washington, DC.*