Determinants of the Commuter Rail Mode Share in Ireland: A Spatial Analysis

By Edgar L. W. Morgenroth

Economic and Social Research Institute

Abstract: Ireland in common with most other developed countries faces a challenge to reduce green house gases (GHG). The major sources of GHG emissions are industry, agriculture and transport. Transport emissions have grown more rapidly than other emissions and are projected to continue growing. It is therefore not surprising that there is significant policy focus on reducing emissions from transport. A key aspect of this is to achieve modal shift away from cars towards more sustainable modes including public transport. In terms of modal shares Ireland differs from other EU countries in that rail has a relatively low share particularly for commuting journeys. Therefore, in order to design appropriate policy responses to increase the mode share for rail it is important to consider the factors that might constrain this share. To this effect a model of rail modal share for small spatial units (3401 electoral districts) is estimated using a cross section of data for 2011 and 2006. Explanatory variables include accessibility, measured as minimum drive time to the nearest railway station, measures of service quality and a number of socioeconomic variables. Access to a railway service is found to be an important determinant. Measures of service quality and in particular frequency explain a larger portion of the variation in the data than accessibility, which suggests that an increase in service frequency could significantly increase rail mode share. High levels of car ownership reduce the rail share, which concords with the finding in the literature.

Key Words: Rail mode share, Service quality, Tobit, Multi Level Model

JEL Code: C21, R15, R14

Paper prepared for presentation at the Irish Transport Research Network (ITRN) Annual Conference 5th to 6th of September 2013

at Trinity College Dublin

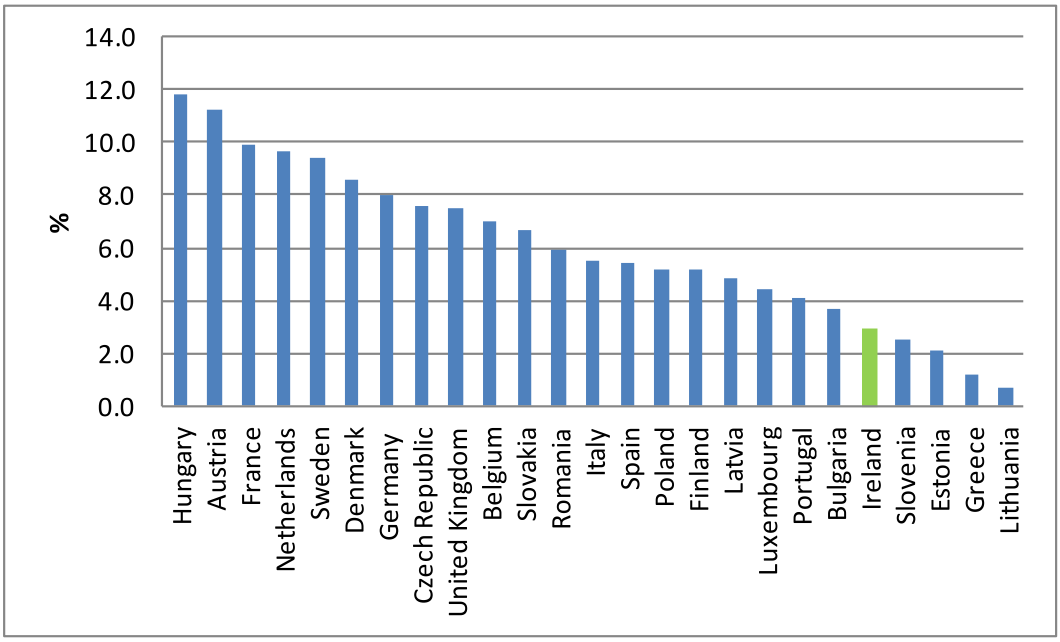
Preliminary draft - not for citation without the permission of the author.

1. Introduction

Ireland in common with most other developed countries faces a challenge to reduce green house gases (GHG). The major sources of GHG emissions are industry, agriculture and transport. Transport emissions have grown more rapidly than other emissions and are projected to continue growing (see Curtis et al, 2013). It is therefore not surprising that there is significant policy focus on reducing emissions from transport. A key aspect of this is to achieve modal shift away from cars towards more sustainable modes including public transport, which requires an understanding of the factors driving modal choice.

In terms of modal shares Ireland differs from other EU countries in that rail has a relatively low share particularly for commuting journeys. Therefore, in order to design appropriate policy responses to increase the mode share for rail it is important to consider the factors that might constrain this share.

Figure 1. Modal split of passenger transport



Source: Eurostat

Previous research on mode shares has largely focused on modelling individual choices[1]. While individual level analysis yields many important insights, it tends consider spatial aspects only in a relatively crude way, usually by including an urban/rural dummy or region dummies. Furthermore, the impact of area wide ('environmental' in the broadest sense) variables is also not considered. These factors appear to be important as the mode shares

[1] There are also some papers applying a macro approach where passenger demand is estimated at the country level (e.g. FitzRoy and Smith, 1995).

have interesting spatial patterns (see Map 1), which appear to derive from location specific factors.

This paper departs from the usual approach of modelling individual choice by modelling the rail mode share across small spatial units using data for Ireland. This allows for a careful modelling of locational factors, particularly the relationship between proximity to a railway station and rail usage. Furthermore, the analysis can also account for the service quality variation across railway stations by including a number of service quality variables, such as frequency, speed and the number of stops, which could be used to influence the demand for rail services. Apart from providing an alternative perspective on the determinants of mode share, modelling spatial data also provides important insights for the land-use-transport interaction and can be used to assess the impact of changes to rail services on transport patterns in specific spatial units.

Previous research on mode choice in Ireland has followed the approach of modelling individual choices. For example Commins and Nolan (2010) used Census microdata to analyse the relationship between car ownership and commuting transport mode, and found that public transport availability and population density had a significant impact on public transport being chosen.

A study by Ahern and Tapley (2008) conducted a combined stated choice and revealed preference survey of bus and rail users for two routes (Dublin-Sligo and Dublin-Galway) in order to obtain information on preferences and perceptions that are relevant to travellers choice between bus and rail. They found that cost and travel time were the most important considerations in choosing the mode of transport, while the quality measures frequency and reliability were found to be less important. The latter finding may however be influenced by the relatively high frequency on both routes and the fact that the two routes are inter-urban routes where commuting trips are less likely. Anandarajah et al. (2004) also used a stated preference analysis to analyse the impact of rail transport quality. In particular they found that passengers were willing to pay more for a better service, with quality improvements including speed, reliability and frequency.

A particular focus of the analysis in this paper is on rail quality measures. These have been considered in a studied for other countries. For example Crampton (2002) found that the physical or technical performance of rail transport on its own does not explain usage patterns but that the pricing structure and general passenger accessibility (in particular light rail corridor population density, stop density, peak hour service length) as well as car restriction measures are the key factors in driving usage patterns. However, Ben-Akiva and Morikawa (2002) show that reliability is an important determinant in mode choice in an analysis of the choice between bus and rail travel. This is likely to be particularly important for commuting and business travel. Mandel et.al (1994) found that rail travel is more price and speed sensitive than car travel. Specifically they found that a one percent reduction in travel time would increase rail demand by at least one percent while a one percent reduction in car travel time would increase demand by just 0.14%. They also estimated the impact of a one percent increase in frequency, which would be raise demand by approximately 0.1%. However, in a study of the determinants of rail passenger demand across Europe, FitzRoy and Smith (1995) found that a 1% increase in frequency results in a 0.5% increase in rail passengers. The differences in the impact of frequency changes might be explained by the possibility that this sensitivity varies between different service levels. For instance Wardman (1994) found the frequency elasticity to be lower for longer distances but to be higher at higher service intervals.

2. Model/Hypothesis and Data

Here the aim is to model the mode share of rail in commuting journeys at the small area level. The focus on the small area level allows the modelling of distance effects. Furthermore, it is hypothesised that the level of service has an important impact as do socio-economic variables.

Thus the basic model can be written as:

$$S_i = \alpha + \beta_D D_i + \beta_Q Q_S + \beta_E E_i + e_i$$

Where *S* denotes the rail mode share in spatial unit *i*, ***D*** denotes a vector of distance/geographic variables for spatial unit *i*, ***Q*** denotes a vector of variables that measure the quality of the service for each railway station and **E** is a vector of socio-economic variables which is observed for each ED.

The data for the analysis comes from the Census of Population which held every five years in Ireland, with the most recent being 2011. The Census contains questions regarding transport mode for commuting purposes, and a range of other socio economic variables. At the spatially disaggregated level, data is made available down to the electoral district (ED) level through the Small Area Population Statistics (SAPS). There are just over 3400 electoral districts in Ireland that range in area from 5 hectares to just under 13,000 hectares, in population from 39 persons to just over 16,000 persons. Given the small size of some EDs these are amalgamated yielding a sample of 3401 Eds. The analysis is carried out for the recent 2011 Census and the previous 2006 Census.

The dependent variables in the analysis is the rail mode share is calculated from the SAPS as the share of those commuting to work, school and college using rail, DART or LUAS as their main mode of transport[2]. The rail mode share averages at just over 1% but for 1360 EDs (40%) the share is zero. For those EDs with some rail commuters the average is just over 1.7%. There are some EDs with a rail share of more than 20%. The spatial distribution of the rail share is shown in Map 1. While there is an obvious relationship with the proximity to a rail line, that relationship seems to be quite weak particularly in more rural areas.

[2] DART is the Dublin Area Rapid Transit rail service, which serves the coastal route between Howth and Greystones that was introduced in 1984. LUAS is the name of two tram lines within Dublin, which came into operation in 2004.

As the type of service differs between conventional ‘main line’ rail services and the DART and LUAS services, dummy variables for EDs that have access to the high frequency DART (Dublin Area Rapid Transit) or the Luas tram lines are added to take account of any specific effect associated with these services, which is important as the main focus of this paper is on the so called mainline rail services.

Rather than use distance as a measure of proximity to a railway station, the time it takes to drive to the nearest station (measured from the ED centroid) is used, as this reflects the quality of the roads infrastructure by assuming different average speeds for different types of roads and also adjusts for topographical and other factors that increase the actual distance to a station[3]. To account for possible nonlinear effects of the drive time a squared term is also included. The drive time is also used to define rail station catchments, by allocating EDs to the railway station with the shortest drive time. The catchments are shown in Map 3. The largest, Cork, covers a population of almost 240,000 while 13 stations have no catchment as for surrounding EDs other stations are closer. Using the drive time it is also straightforward to identify the proportion of the population that is within time bands. Table 1 shows that for 2011 almost 47% of the population resided within 15 minutes drive time from a railway station and just 8.9% resided more than one hour from a station.

Table 1 Percentage of the Population by Drive Time from the nearest Railway Station, 2006 and 2011

	2006	2011
Under 15 minutes	45.9%	46.9%
15 -30 minutes	27.4%	27.3%
30-45 minutes	12.0%	11.4%
45 -60 minutes	5.7%	5.6%
60 minutes and more	9.0%	8.9%
Total	100%	100%

The rail service variables, frequency (number of trains in one direction), speed (average speed along the route), and number of stop (maximum number of stops along a route and minimum number of stops was taken from Irish Rail time tables (Iarnrod Eireann). As the frequency might impact on the mode share in a non-linear way the squared frequency is also included in the analysis. The spatial distribution of service frequency is shown in Map 3. Comparing this with Map 1 indicates a strong correlation between mode share and service frequency (r=0.48).

As the number of stops along a line might influence the decision to use a rail service either because frequent stops allow passengers allows for passengers to access multiple destination while more stops reduce the overall speed and may impact on the degree of crowding on the service for which we have no data. Given that the number of stops differs

[3] The drive times were calculated using MapPoint and the MPM Mile Charter utility, using the default speed settings.

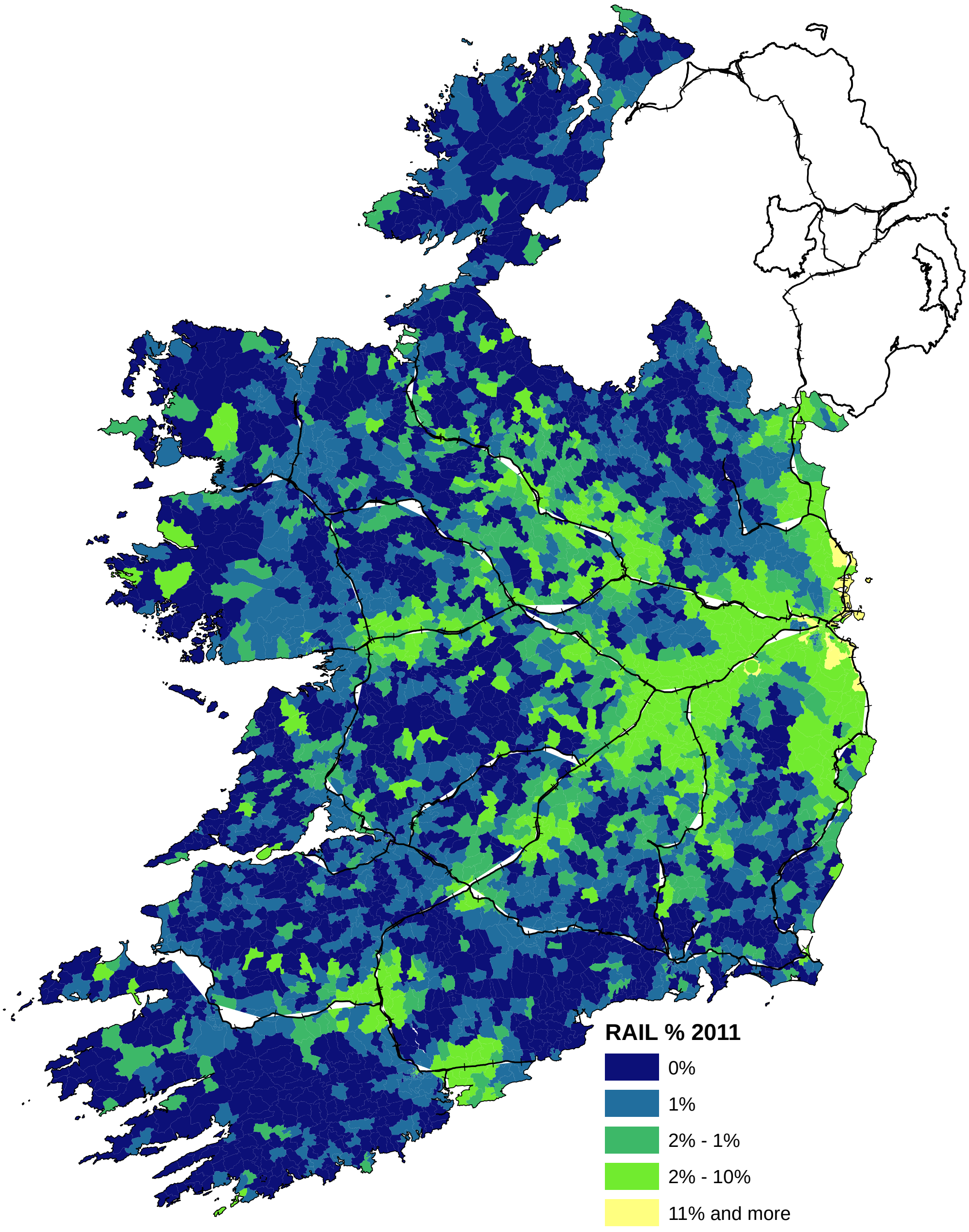
between different trains calling to the same station we include variables for the maximum and minimum number of stops. For 2011 details of whether a train station has car parking spaces are available.

A number of additional socio-economic variables are also included in the analysis. As it is often argued that a higher population density increases rail usage population density is included[4]. While the dependent variable relates to commutes, a high unemployment rate might proxy lower income areas. Likewise areas with a higher proportion of professionals are likely to be more prosperous. Thus these two variables are added to account for income differences across EDs.

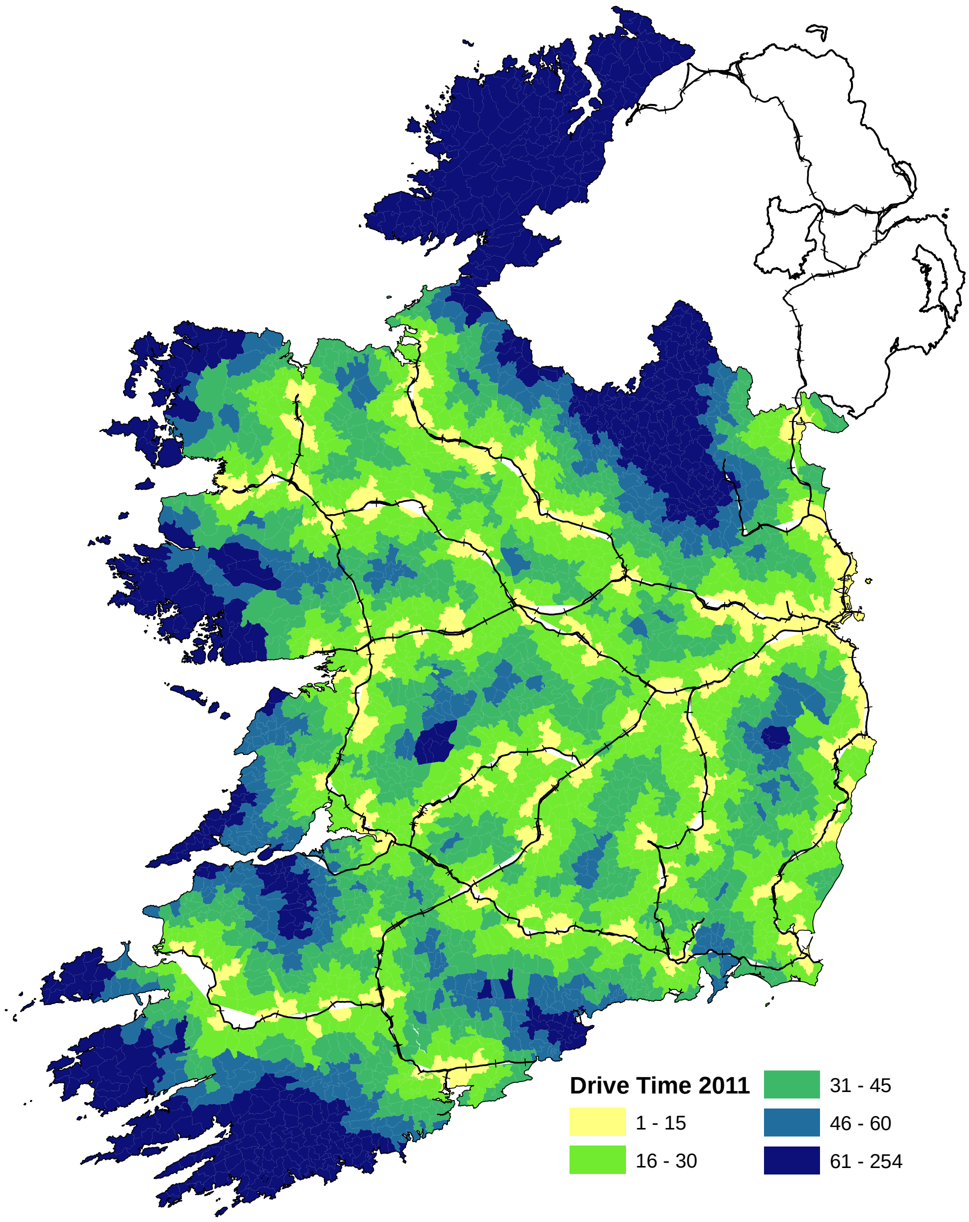
During the last decade a significant number of immigrants came to Ireland. In so far as these are more concentrated in some EDs the share of the population accounted for by immigrants could impact on public transport usage either because they are less likely to own a car or perhaps because their attitudes to public transport might differ to that of the Irish population (for example many New EU Member States have a tradition of good public transport services at low prices. To account for the possibility that mode choices differ in more rural areas the percentage of the workforce employed in agriculture, forestry and fishing is included in the analysis. Long distance commuters are likely to have different commuting behaviour than those travelling only a short distance, not least because modes like walking and cycling are not feasible modes. Thus, the percentage of commuters that are long distance commuters is included in the model. This variable is defined as the percentage of commuters with a travel time in excess of one hour. Research has shown that car ownership negatively impacts on public transport usage and consequently the average number of cars per household is included in the model. All of these additional socio-economic variables are drawn from the SAPS. As there may be specific agglomeration effects in the major urban areas the rive time to the centres of the four largest cities, Dublin, Cork, Limerick and Galway are also included in the analysis.

[4] For example Glick (1992) argues that light rail use intensities are significantly determined by land use policies and particularly population densities in close proximity to rail lines.

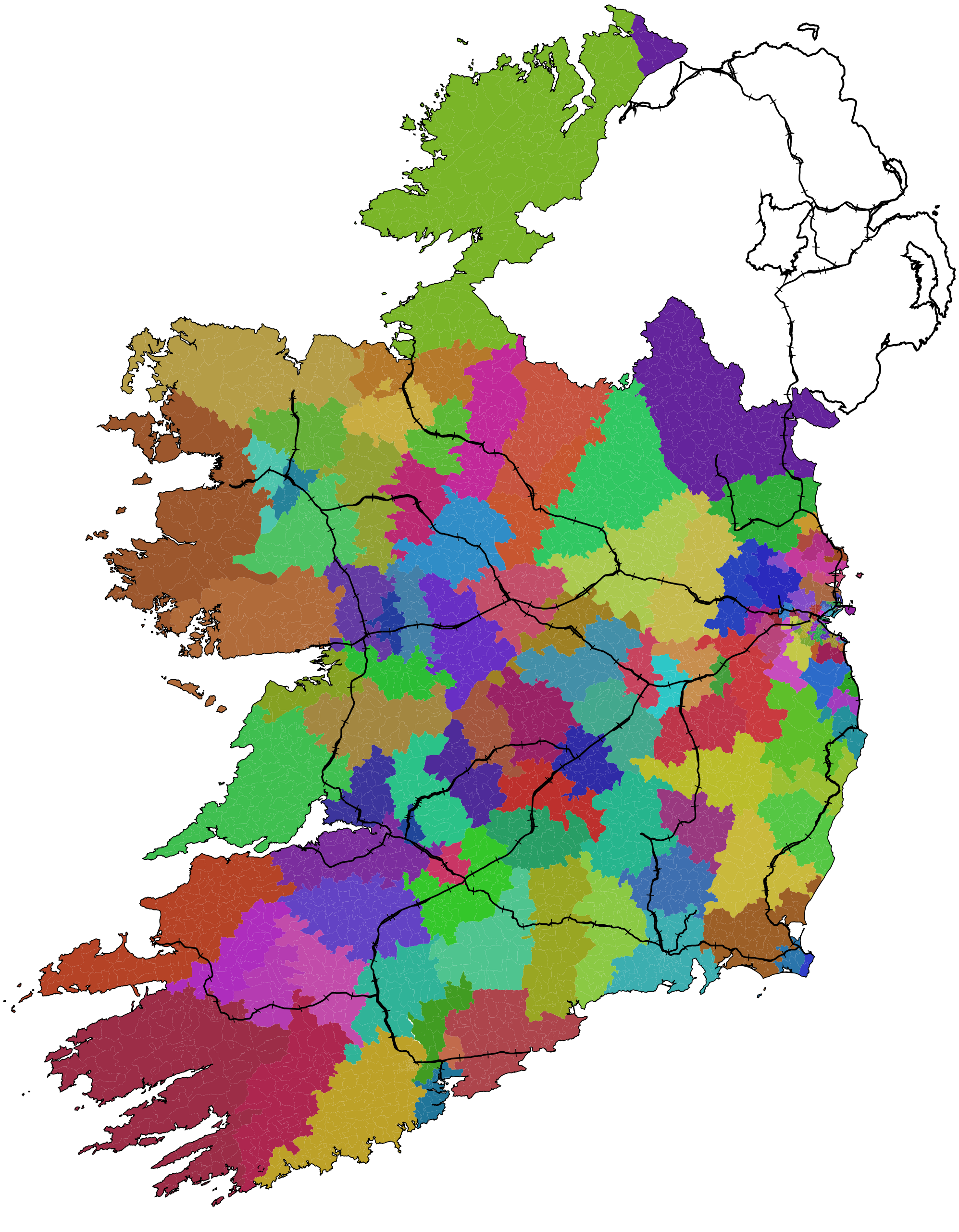
Map 1. Rail Mode Share for Commuting to School, College and Work (2011)



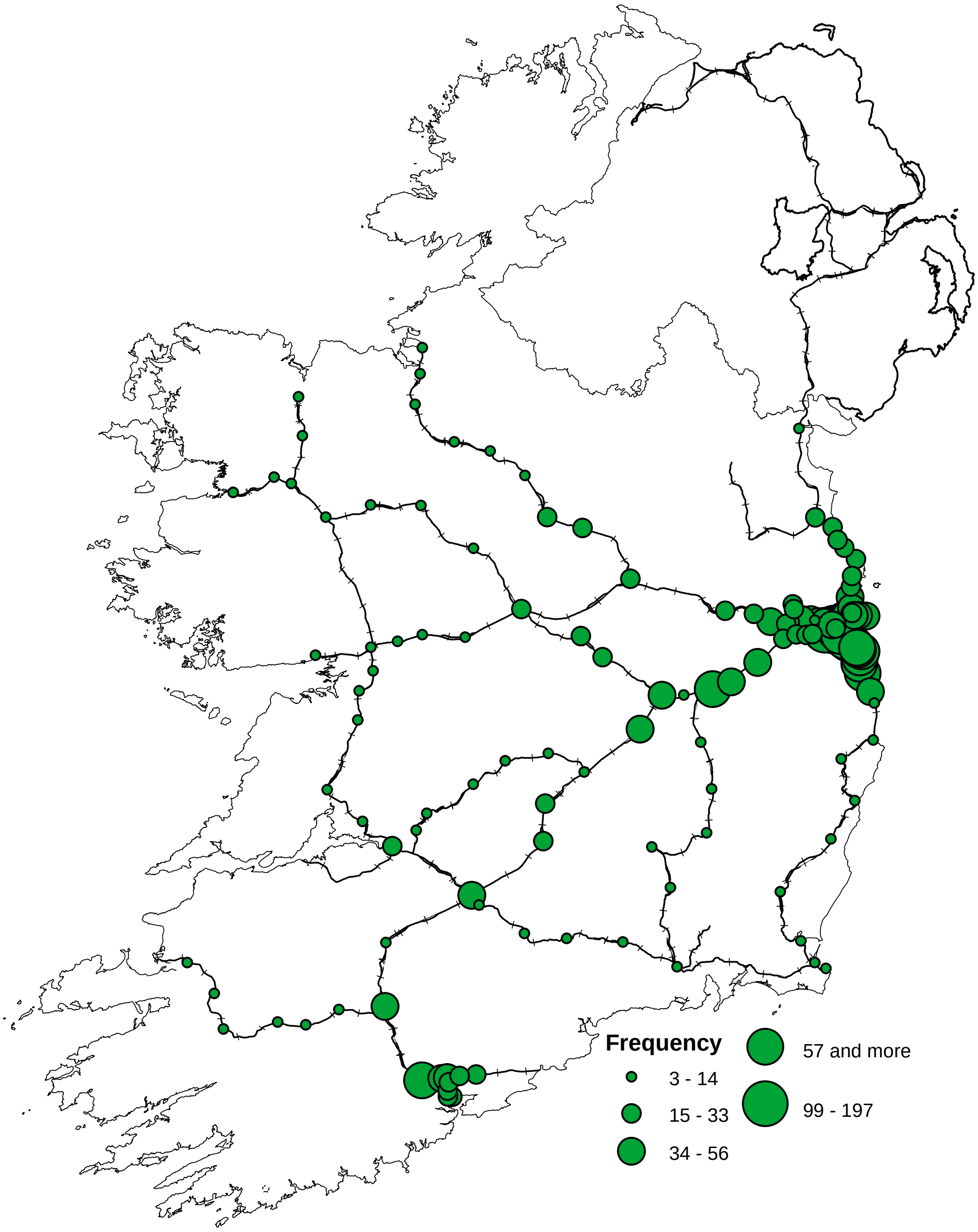
Map 2. Drive Time to Railway Stations (minutes), 2011



Map 3. Rail Station Catchments, 2011.



Map 4. Frequency of Trains in Each Train Station



3. Estimation and Results

In the first instance the model is estimated using ordinary least squares (OLS). However, there the nature of the data has important implications for the appropriate estimator. Firstly, the rail quality variables frequency and speed are observed at the station level so they are the same across the catchment. This implies that the residuals across the observations for a particular station catchment are likely to be correlated. This can be accounted for by adjusting the standard errors. Another important issue is that as was noted above, some 40% of EDs had no rail commuters at all. One interpretation of this is that the utility of using rail for all commuters in that ED is at least zero and possibly negative, which implies that the data is censored at zero. This implies that estimates relate to the mean of the underlying dependent variable is not equal to the mean of the observed data. The usual way to deal with this is to apply a Tobit model (Tobin, 1958), which splits the contribution to the likelihood function into two parts, one for the non-censored observations and one for the censored observations. The Tobit estimator is therefore utilised here. The fact that some of the variables are observed at the ED level while others are observed at the station (catchment) level i.e. some of the data is clustered into station catchments also suggests an alternative modelling strategy namely to apply multilevel (hierarchical)modelling. In the simplest form these allow for the clustering by introducing random intercepts (effects) for each level, where the bottom level in the data in this study relates to the ED and the second level relates to the station catchment[5]. Thus, multilevel models allow for the correlation of observations across higher levels, rather than simply adjusting the standard errors for clustering. Thus they can deal with the particular structure of the independent variables used in this study. These models can also adapted to allow for censoring so it is possible to estimate a Multilevel Tobit which is also applied here[6].

Thus the model is estimated using OLS, Tobit, Multilevel and Multilevel Tobit estimators. The results are shown in Tables 1 and 2 for 2011 and 2006 respectively. As can be seen from the log-likelihood, the of fit of the Tobit and particularly the Multilevel Tobit is considerably better than the models that disregard the censoring.

There are some differences regarding the parameters and their significance across models but they provide broadly similar results. Being further away from a station significantly reduces the rail mode share at the ED level. A higher frequency of service results in a higher rail mode share. Speed, the number of stops and the availability of parking at the station appear not to impact significantly on the rail mode share. Both the LUAS and DART services are associated with higher levels of rail mode share which is not surprising as they offer a different service compared to mainline rail services. Of course they also service the area

---

[5] One could also allow for a third level, accounting for rail lines (speed and stops are measured for each line). However particularly for interchange stations such as Dublin Connolly or Heuston which service many routes allocating stations to one line would not be appropriate.

[6] The estimation is carried out using STATA 12. The simple multilevel model is estimated using the xtmixed command, while the Multilevel Tobit is estimated using the gllamm procedure. Gllamm estimates generalised linear mixed models in a maximum likelihood framework using adaptive quadrature.

with the highest population density, which may account for the surprising result that population density does not increase the rail commuting mode share, which stands in contrast to other findings. Income levels, proxied by unemployment and the percentage of professionals significantly impact on rail mode share, with a higher unemployment rate reducing the share and more professionals increasing the share. Interestingly, the share of the population accounted for by foreign nationals impacts positively on rail mode share. This may be due to them not having the same level of car ownership or some preference for public transport. A high level of long distance commuting increases the rail share. This may be due to the fact that rail constitutes a more comfortable mode that allows individuals to work or engage in leisure activities (e.g. read) while travelling. A higher level car ownership results in lower rail mode share as has been found elsewhere in the literature. The results for 2006 are very similar to those for 2011. However, the coefficient for the service frequency was considerably larger in 2006 compared to 2011, suggesting that the impact of this has declined over time. The dummy variable for LUAS is insignificant, which may reflect the fact that at that stage LUAS was still a relatively new service. Also the unemployment rate was insignificant, but this was considerably lower and has a lower variance across EDs in 2006 compared to 2011.

As was noted, a high service frequency of leads to higher rail mode shares but this effect only pertains up to a point as is shown in figure 2, which shows the implied contribution of the service frequency to mode share. This is constructed by taking the estimated parameters and multiplying these by the frequency. Given that there are two parameters, one for the level of frequency that is positive and one for squared frequency that is negative, the estimates imply a curved relationship. The figure shows that the Multilevel estimators result in more curvature which also results in a downward shift of the optimal rail frequency. The Multilevel Tobit results suggest that the mode share maximising frequency is 54 while the OLS results suggest that it would be 91.

Another way to see the implications of the results is to consider the impact of a change in frequency for a particular station. One interesting example is Dundalk which is on the Dublin-Belfast line. Commuters from Dundalk that would use this line would largely travel towards Dublin. Currently there are 13 services to Dublin while for the next station on the line, Drogheda, there are more than 30 services per day. Morgenroth (2011) in a benchmarking study estimated that the expected service frequency for the Dublin-Belfast train should be 20[7]. Using the results from the Multilevel Tobit it is straightforward to simulate the increase in passengers if the frequency serving Dundalk were increased to 20 trains. Here it is assumed that the frequency only impacts on those EDs which currently have some commuters travelling by rail i.e. the change in frequency is assumed not to impact on the residents in those EDs. In 2011 there were 449 commuters in EDs in the Dundalk catchment using rail. The increased service would result in an additional 314 (+70%) commuters using rail.

[7] Currently there are just 8.

Table 1 Estimation Results for 2011 Data

	OLS	Tobit	Multi-Level	Multi Level Tobit
Drive Time	-0.0281***	-0.0618***	-0.0201***	-0.0460***
Drive Time^2	0.0001***	0.0002***	0.0001***	0.0002***
Frequency	0.0278***	0.0322***	0.0825***	0.0557***
Frequency^2	-0.0002***	-0.0002**	-0.0006***	-0.0005***
Speed	0.0066**	0.0077	-0.0160	-0.0073
Maximum Stops	-0.0495**	-0.0354	0.0401	0.0451
Minimum Stops	0.0379	0.0129	-0.0702	-0.0815
Luas	0.7079	0.8235	1.3975***	1.3118***
Dart	5.1750***	4.8977***	7.6747***	7.5111***
Dublin	0.0000	-0.0029	0.0019	0.0010
Cork	0.0082***	0.0065	0.0016	0.0062
Limerick	-0.0074***	-0.0035	-0.0014	-0.0070
Galway	0.0089***	0.0051	0.0055	0.0113
Population Density	0.0000	-0.0001	0.0000	0.0000
Unemployment rate	-0.0221***	-0.0356***	-0.0166**	-0.0277**
% Foreign	0.0107	0.0077	0.0282***	0.0261***
% Agriculture	0.0224***	-0.0112	0.0210***	-0.0080
% Professional	0.0398***	0.0514***	0.0293***	0.0463***
% Long Distance	0.1109***	0.1674***	0.0664***	0.1091***
HH cars	-1.4326***	-1.9960***	-0.9733***	-1.6325***
Parking	0.0426	0.2250	0.7359	0.6710
Observations	3401	3401	3401	3401
$\bar{R}^2$	0.54			
Log-likelihood	-7082.7	-5410.0	-6599.4	-5222.2

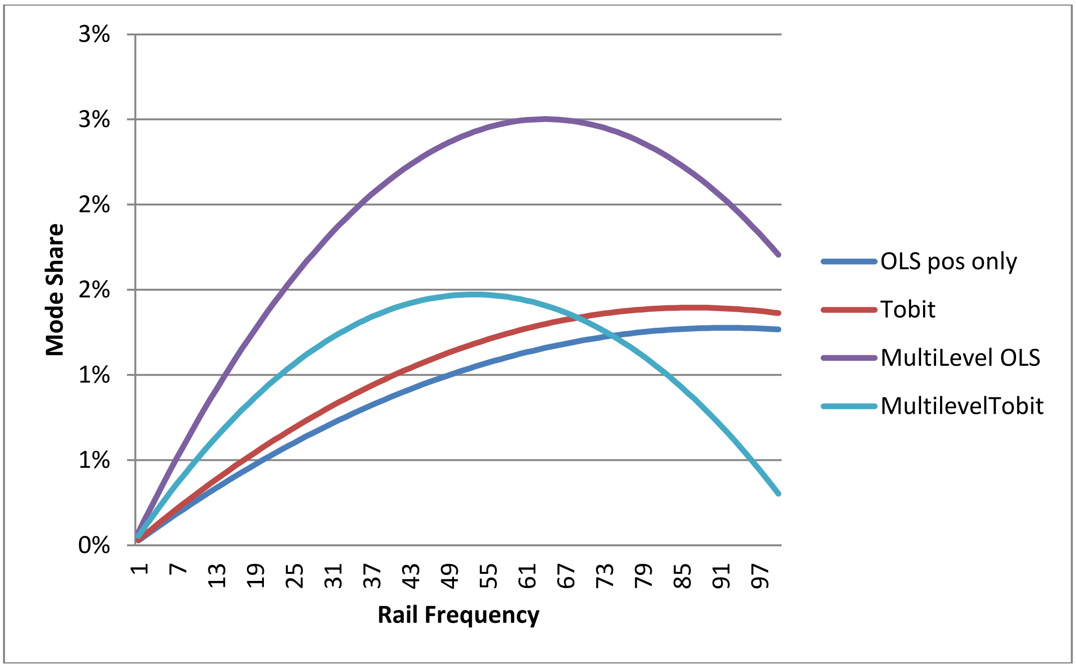
Note: ***, **, * respectively. OLS and Tobit standard errors are adjusted for clustering on rail stations.

Table 2 Estimation Results for 2006 Data

	OLS	Tobit	Multi-Level	Multi Level Tobit
Drive Time	-0.0307***	-0.0597***	-0.0205***	-0.0421***
Drive Time^2	0.0001***	0.0002***	0.0001***	0.0001***
Frequency	0.0587***	0.0649***	0.1285***	0.1147***
Frequency^2	-0.0003***	-0.0004***	-0.0011***	-0.0010***
Speed	0.0034	0.0104	-0.0240	-0.0096
Maximum Stops	-0.0655**	-0.0583	-0.0870	-0.0811
Minimum Stops	0.0020	-0.0081	-0.1139	-0.0976
Luas	-0.0539	-0.0914	0.7025**	0.5852
Dart	5.0656***	4.7039***	7.6123***	6.7333***
Dublin	0.0027	-0.0019	0.0020	0.0017
Cork	0.0112***	0.0093*	0.0048	0.0126*
Limerick	-0.0080	-0.0033	0.0007	-0.0055
Galway	0.0079**	0.0045	0.0070	0.0129*
Population Density	0.0000	-0.0001**	0.0000	0.0000
Unemployment rate	-0.0036	-0.0065	-0.0146	-0.0251
% Foreign	0.0233	0.0283	0.0400***	0.0456***
% Agriculture	0.0043	-0.0569***	0.0066	-0.0596***
% Professional	0.0397***	0.0521***	0.0238***	0.0380***
% Long Distance	0.1154***	0.1406***	0.0574***	0.0844***
HH cars	-1.1178***	-1.4978***	-0.7889***	-1.1689***
Observations	3401	3401	3401	3401
$\bar{R}^2$	0.54	0.18		
Log-likelihood	-7082.7	-5245.5	-6626.4	-5074.9

Note: ***, **, * respectively. OLS and Tobit standard errors are adjusted for clustering on rail stations.

Figure 2. Estimated Relationship between Frequency and Mode Share.



4. Conclusions

This paper has analysed the rail mode share. In doing so it has taken a different approach to that usually taken in the literature by considering the rail mode share at the micro-spatial level. As part of this the proximity to rail stations was estimated as the drive time to a station rather than distance which is often used as a proxy. Considering drive times allows for differences in the underlying infrastructure. By estimating the drive time to stations, station catchments were identified, for which service quality measures were collected from railway timetables. The modelling approach accounted for the fact that a considerable proportion of EDs had no rail commuters and the fact that some explanatory variables were identical across a number of observations i.e. they were clustered on rail stations.

The results confirm the importance of accessibility. As the drive time increases the rail mode share quickly declines, which explains the relatively narrow rail corridors shown in Map 1. From a policy perspective there are two options to address accessibility issues. Firstly, one could build more railway lines, which would be very costly. Secondly, one could facilitate access for example through the integration with other modes. Anspacher et al. (2005) in a study for San Francisco, show that using feeder buses can significantly increase the demand for rail, particularly in areas of low population density. Furthermore they show that a significant proportion of the population would be willing to use a shuttle. This suggests that if multi-modal journeys are not facilitated then the rail catchment size will be limited.

While other quality variables do not seem to matter, the service frequency was found to have an important bearing on the rail mode share. Thus, simply providing a rail line is not sufficient to create demand for rail services. A lack of frequency renders a service inflexible and results in longer waiting times. As this is a key policy variable i.e. could relatively easily be changed by the service provider, the results suggest that simulations around the optimal use of the rolling stock should be carried out. A simple simulation for Dundalk station suggests that improvements in rail frequency could result in very significant increases in rail use.

While the modelling work has accounted for some of the issues arising out of the nature of the data that is being modelled, further work needs to adjust for the likely presence of spatial autocorrelation which could arise from the if unobserved (missing) variables are spatially correlated. This may well be an issue as for example the presence and quality of bus services is not accounted for with any of the variables.

References

Ananderajah, G., Ahern, A., and N. Tapley (2004) "Intercity Rail Transportation Quality and Suggested Improvements to Increase the Modal Share of Rail in Ireland", Proceedings of the European Transport Conference 2004

Ahern, A., and N. Tapley (2008) "The Use of Stated Preference Techniques to Model Modal Choices on Interurban Trips in Ireland", *Transportation Research A*, Vol. 42(1), pp. 15-27.

Anspacher D., Khattak A., and Y. Yim (2005) "The Demand for Rail Feeder Shuttles", *Journal of Public Transportation*, Vol. 8(1), pp. 1-20

Ben-Akiva M., and T. Morikiwa (2002) "Comparing Ridership Attraction of rail and Bus", *Transport Policy*, Vol. 9, pp.107-116.

Crampton G. (2002) "International Comparison of Urban Light Rail Systems: the roles of integrated ticketing, pedestrianisation, and population density" paper presented at the ERSA Congress held in Dortmund, 27th to 31st of August 2002.

Cummins, N., and A. Nolan (2010) "Car Ownership and Mode of Transport to Work in Ireland" *The Economic and Social Review*, Vol. 41 (1), pp.

Curtis, J., Pentecost, A., Lyons, S., Morgenroth, E., and V. Di Cosmo (2013) *Towards a Green Net National Product in Ireland.* EPA STRIVE Report Wexford: Environmental Protection Agency.

FitzRoy, F., and I. Smith (1995) "The demand for rail transport in European Countries", *Transport Policy*, Vol. 2(3), pp. 153-158.

Glick, F. (1992) Light rail transit and effective land use planning: Portland, Sacramento and San Diego" *Transportation Research Record*, Vol. 1361, pp. 75-80.

Mandel, B., Gaudry, M., and W. Rotherngatter (1997) "A Disaggregate Box-Cox Logit Mode Cheice Model of Intercity Passanger Travel in Germany and its Implications for High-Speed Rail Demand Forecasts", *Annals of Regional Science*, Vol.31, pp.99-120

Morgenroth, E., (2011) "Improving the Dublin Belfast Rail Service to Lower Barriers between North and South" *Journal of Cross Border Studies in Ireland*, Vol. 6, pp. 59-70.

Tobin, J. (1958) "Estimation of relationships for limited dependent variables" *Econometrica, Vol.* 26, pp. 24–36.

Wardman, M. (1994) "Forecasting the Impact of Service Quality Changes on the Demand for Inter-Urban Rail Travel", *Journal of Transport Economics and Policy*, Vol. 238(3), pp. 287-306.