Title	Exploring the drivers of light rail ridership : an empirical route level analysis of selected Australian, North American and European systems
Author(s)	Currie, Graham; Ahern, Aoife; Delbosc, Alexa
Publication date	2011
Publication information	Transportation, 38 (3): 545-560
Publisher	Springer
Link to online version	http://dx.doi.org/10.1007/s11116-010-9314-9
Item record/more information	http://hdl.handle.net/10197/2879
Publisher's statement	The final publication is available at springerlink.com
Publisher's version (DOI)	http://dx.doi.org/10.1007/s11116-010-9314-9

DOI 10.1007/s11116-010-9314-9

Exploring the drivers of light rail ridership: an empirical route level analysis of selected Australian, North American and European systems

G. Currie • A. Ahern • A. Delbosc

Abstract This paper explores the relative influence of factors affecting light rail ridership on 57 light rail routes in Australia, Europe and North America through an empirical examination of route level data. Previous research suggests a wide range of possible ridership drivers but is mixed in clarifying major influences. A multiple-regression analysis of route level ridership (boardings per route km) and catchment residential and employ-ment density, car ownership, service level, speed, stop spacing, share of accessible stops, share of segregated right of away and integrated fares was undertaken. This established a statistically significant model (99% level, R^2 = 0.76) with five significant variables including service level, routes being in Europe, speed, integrated ticketing and employment density. In general these findings support selected results from previous research. A sec-ondary analysis of service effectiveness measures (boardings/vehicle km, i.e. the relative ridership performance for a given level of service), established a statistically significant model (99% level, R^2 = 0.67) with 6 significant explanatory variables including being in Europe, speed, employment density, integrated ticketing, track segregation and service level. The latter implies that a higher frequency results in higher service effectiveness. Overall the research findings stress the importance of providing a high level of service as a major driver of light rail ridership. The 'European Factor' is also an important though intriguing influence but its cause remains unclear and requires further research to elaborate its nature.

Keywords Light rail · Patronage levels · Service levels

G. Currie (&) · A. Delbosc
Department of Civil Engineering, Institute of Transport Studies, Monash University, Building 60, Clayton, VIC 3800, Australia
e-mail: graham.currie@monash.edu

A. Delbosc
e-mail: alexa.delbosc@monash.edu

A. Ahern
School of Architecture, Landscape and Civil Engineering, University College Dublin, Newstead Building, Belfield, Dublin 4, Ireland
e-mail: Aoife.ahern@ucd.ie

Introduction

The introduction of new light rail has been proposed as a means of improving the attractiveness of urban transport systems struggling under the combined impacts of traffic congestion, overcrowding, ageing infrastructure and pollution. Transport authorities evaluating the feasibility of light rail face a difficult challenge in assessing the benefits of increased public transport ridership against the high fixed costs of implementation. Factors driving ridership are not always clear and there is much evidence that light rail ridership expectations are not always met (Edwards and Mackett 1996; Mackett and Babalik-Sutcliffe 2003). The cultural and socio-economic context of the cities where light rail is implemented is thought to affect ridership as much as the specifics of service design and operation.

This paper explores the relative influence of factors affecting light rail ridership through an empirical examination of route level data using a multiple-regression analysis of potential explanatory factors. Analysis focuses on 57 light rail routes in Australia, Europe and North America. These systems contrast in a number of ways including system age, operating speed, the nature of system right of way, service and ticket integration, scale of network etc. An aim of the research is to explore the relative value of these features in terms of ridership impacts.

The next section presents a review of previous research literature focussing on factors driving ridership in light rail systems. Section 3 describes the research methodology and Sect. 4 presents a brief description of the light rail routes examined including an outline of existing route level ridership performance and associated explanatory data. Section 5 outlines the findings of the analysis. The paper concludes by summarising key findings and discussing the implications of these. Suggestions for future research are identified.

Research context: light rail ridership drivers

A large range of research evidence based on behavioural research studies has now identified passenger responses to transport system changes for a range of service design aspects (e.g. Balcombe 2004; Evans 2004; McCollom and Pratt 2004; Pratt and Evans 2004). The focus of this paper and hence this literature review is aggregate analysis of system wide or route performance. Aggregate studies necessarily abstract from the detailed perspectives of behavioural research at the individual level. Nevertheless they have value in exploring factors influencing system-wide ridership drivers.

Table 1 presents a summary of the proposed drivers of light rail ridership based on mainly aggregate empirical studies in the research literature.

The density of urban development has long been identified as a major driver of rider-ship: ‘‘nearly every study that has focussed on transit ridership has provided evidence that density is the primary determinant of transit ridership’’ (Johnson 2003, p. 32). Although urban density is often cited as a ridership driver in light rail based research it is often not identified as a primary driver. Kain and Liu (1999) examined the factors determining the high ridership of light rail systems in Houston and San Diego. While stating that factors like urban density and employment levels play a role in determining patronage levels, they concluded that the most important factors to determine patronage are high service levels (measured in vehicle kilometres on a route) and cheap fares. Babalik-Sutcliffe (2002) suggests that it is a mix of supportive factors rather than any single factor which attracts high light rail patronage. Although she suggests that supportive urban development, including density and design features, is important the findings also emphasise a supportive

Table 1 Light rail ridership drivers—previous aggregate research

Identified ridership driver	Research source
High density residential development	(Hass-Klau and Crampton 2002)
	(Johnson 2003)
	(Seskin and Cervero 1996)
	(Kain et al. 2004)
	(Babalik-Sutcliffe 2002)
	(Kain and Liu 1999)
Public transport network effect	(Babalik-Sutcliffe 2002)
	(Mackett and Babalik-Sutcliffe 2003)
	(Denant Boemont and Mills 1999)
High service levels	(Kain and Liu 1999)
	(Mackett and Babalik-Sutcliffe 2003)
Low car ownership	(Babalik-Sutcliffe 2002)
	(Mackett and Babalik-Sutcliffe 2003)
Low fares	(Kain and Liu 1999)
	(Mackett and Babalik-Sutcliffe 2003)
Modal integration	(Babalik-Sutcliffe 2002)
	(Kain et al. 2004)
Ticket integration	(Crampton 2002)
	(Hass-Klau and Crampton 2002)
	(Mackett and Babalik-Sutcliffe 2003)
Pedestrianisation	(Hass-Klau and Crampton 2002)
Reliable service	(Mackett and Babalik-Sutcliffe 2003)
Strong economic conditions	(Babalik-Sutcliffe 2002)
High employment	(Kain and Liu 1999)
High speed	(Hass-Klau and Crampton 2002)
	(Crampton 2002)
Stop distance	(Hass-Klau and Crampton 2002)
	(Crampton 2002)
Strong policy support	(Knowles 2007)
Light rail network density	(Hass-Klau and Crampton 2002)
	(Crampton 2002)
Easy station access	(Kain et al. 2004)

public transport policy environment including integration of public transport modes and services. Similar findings resulted from the work of Kain et al. (2004) in a study of light rail use in the United States. They found residential population density near stations and integration with other modes to be significant factors in high ridership although they also noted the importance of easy accessibility to stations for a wider range of user types. A study of system-wide light rail data from 24 cities (Crampton 2002; Hass-Klau and Crampton 2002) found that population density was the second of two influential explanatory variables explaining a light rail performance index (including ridership) which were significant at the 99% confidence level (the first being integrated ticketing). A series of other factors were also significant (but at a 95% confidence level).

High service levels, measured in terms of frequency and span of hours covered, has often been cited as an important driver of light rail patronage. It was considered a principal driver in the US light rail research by Kain and Liu (1999). Fitzroy and Smith (1998) in their study of the European Freiburg public transport system state that high service levels are important for achieving high patronage levels. Route level demand forecasting models developed by Stopher (1992) also found that service quantity of buses, measured as the number of buses per hour, was the single largest most significant factor in an empirical analysis of US bus routes.

Interestingly Hass-Klau and Crampton (2002) in their study of 24 light rail systems did not find service levels (measured as peak headway or service span or hours run at peak frequencies) to be a significant influence on their light rail performance index.

A number of researchers have cited the importance of an integrated public transport network as a key driver of high light rail patronage (FitzRoy and Smith 1998; Denant Boemont and Mills 1999; Babalik-Sutcliffe 2002). Patronage drivers in this case involve service and fare integration as well as the wider 'network effects' these can generate. Service integration occurs where transfers between modes/services involve a short walk and wait plus well-coordinated and closely scheduled arrival and departure times.

Fare integration occurs where there is no requirement to incur additional fare cost, or to have to go through a fare payment transaction whilst transferring. 'Network effects' have been closely related to both service and fare integration. These are 'synergies' where cohesiveness, cooperation and efficient interconnectivity within networks enable the achievement of goals or performance outcomes which exceed the sum of individual factors (Capineri and Kamann 1998). In public transport patronage terms they occur when pas-sengers can easily access a large network rather than a single route thus considerably enhancing the opportunities for access to a wider range of destinations.

Integrated fares were identified as the single most influential of two variables in the Hass-Klau and Crampton (2002) study of 24 light rail systems. A 0.64 correlation with an index of light rail system wide performance (including ridership) was identified.

A range of other factors have been suggested which might also influence light rail ridership. Cheap fares were cited two reports (FitzRoy and Smith 1998; Kain and Liu 1999). A number of researchers cite the importance of a strong policy context as a basis for high light rail ridership (e.g. Knowles 2007). Several researchers have suggested that high car ownership can act to reduce light rail usage (Babalik-Sutcliffe 2002; Mackett and Babalik-Sutcliffe 2003). Hass-Klau and Crampton (2002) suggested that pedestrian zone length in cities, average speed, stop distance and the density of the light rail network were also related to their index of light rail performance (based on ridership per route km). Correlation analysis suggested better performance (ridership) at slower speeds and short stop distances (Crampton 2002) whereas pedestrian zone scale and density of the network were positively related with ridership variables.

The results of the Hass-Klau and Crampton (2002) and Crampton (2002) analyses are worthy of further examination relative to this study since a multiple regression analysis of light rail performance factors was also undertaken. System-wide (rather than route level) data from 24 light rail systems from around the world (but with 75% from Europe) was compiled. The dependent variable was a performance index based on the ranking of systems. This was based on per capita light rail ridership, total transit ridership per capita, light rail ridership per route km, the annual growth rate of light rail ridership, the growth rate of total transit ridership and light rail passenger kms per track km. The authors note the weakness of their small sample and hence suggest caution in use of their analysis. They report an R^2 of 0.60 with major explanatory variables including travel card use (ticket integration), pedestrianisation, population density and low fares.

Overall many factors are considered to have influenced ridership on light rail systems however on balance what factors are important appear varied between studies. Clarifying these issues is a major aim of the research methodology developed for the project.

Methodology

The method aims to measure the strength of the relationship between light rail route level ridership and a series of possible explanatory variables.

Variables

Table 2 shows the variables used to explore ridership drivers and the major sources of the data collated. The selection of variables was based on previous research and also on the availability of data. Only those variables that could be reasonably measured across countries were included and this was a challenging and time consuming part of the project. An unfortunate limitation is that we were unable to generate a ‘‘cost’’ variable that was applicable across such different contexts and appropriate to the years that ridership data was available.

Whenever possible, route level data was collated for this analysis rather than the system wide data used in previous studies (e.g. Hass-Klau and Crampton 2002). This increased the scale of the data collection task but was chosen to increase the number of available points and also to enable a wider exploration of explanatory factors e.g. by exploring variations in ridership within specific light rail systems. In general route data is consistent with the base year 2006–2008 however it was not possible to exactly match data for a specific year in each case and best nearest case for matching data was used.

Whenever possible, car ownership, urban density and employment density was measured in the 800 m catchment around a route; Australia, Toronto, and Dublin were measured in this way. However this route-catchment data was not available for those routes in the UK, France and the United States. For these locations data for the entire city was used and interpolated to estimate catchment density.

The statistical model

A linear regression modelling approach was adopted using the following model:

$$Y_i = b_0 + b_1X_{i1} + b_2X_{i2} + \cdots + b_nX_n + e_i$$

where Y_i = dependent variable i, X_i = independent variables predicting Y_i, b = regres-sion coefficients to be estimated, e = error.

A step-wise inclusion criteria was used where variables were included in the model based on their level of statistical significance (significance probability of 95% for inclusion and removal if significance dropped below 90%).

The main dependent variable was ridership per route km however a secondary analysis was also undertaken using boardings per vehicle km as the dependent variable. The latter explored ridership influences after accounting for variations in relative service levels. Service levels were found to be important in explaining ridership per route km and varied

Table 2 Explanatory variables collated—light rail ridership drivers

Variable/measure	Method and source		
	Australia	North America	Europe
Boardings per year			
Used to calculate boardings/route and vehicle km	2007 data Data provided by operators	Toronto—TTC[a] US—FTIS[b]	Based on SYPTE[c] and website data[d]
Vehicle Kilometres			
Used to calculate boardings/vehicle km	Melbourne provided by operator; others estimated from published timetables	As above for Boardings p.a.	Based on SYPTE[c] and website data[d]
Residential density			
Residents per square metre	ABS[e]	Toronto—SC[f], US—census[g]	Dublin: CSO[h] UK: census[i] Others: based on SYPTE[c] Rouen—estimated from Wikipedia
Employment density			
Jobs per square metre	ABS[e]	Toronto—SC[f], US—FTIS[b]	Dublin: CSO[h] UK: census[i] Others: based on SYPTE[c] Selected European centres using Data from INSEE[j]
Car ownership			
Cars per 1,000 residents	ABS[e]	Toronto—TT Survey[k] US—census[g]	Dublin: CSO[h] UK: census[i] France—CERTU[l]
Service level			
Vehicle trips per annum	Based on an analysis of published timetables for 2007	Toronto—TTC[a] US—FTIS[b]	Based on SYPTE[c] and website data[d]
Speed			
Average travel time divided by route length (kph)	As above for service level	As above for service level	As above for service level
Stop spacing			
Route length divided by number of stops minus 1	As above for service level	As above for service level	As above for service level
Share accessible stops			
Proportion of stops that are wheelchair accessible	As above for service level	As above for service level	As above for service level
Share segregated right of way			
Proportion of track out of mixed traffic	Data provided by VicRoads, and an analysis of Google Maps	Toronto: based on route inspection; others : visual inspection of Google Maps	Visual inspection of Google Maps Dublin: Data provided by RPA UK systems: web site data[d]
Integrated fares			
No fare on transfer	Operator website	Operator website	Operator website

Table 2 Explanatory variables collated—light rail ridership drivers

Variable/measure	Method and source		
	Australia	North America	Europe
Route length			
Used to calculate service level, speed, stop spacing and ROW	Melbourne provided by operator; others Google Earth	Toronto—TTC data Google Earth	Mix of web site data and Google Earth Route Inspection UK/ Dublin—web site data[d]

a Toronto Transit Commission 2008 data (www.ttc.ca) (last accessed Nov 2009)

b 2006 data from Florida Transit Information System, see http://www.ftis.org/ (last accessed Nov 2009)

c A study of European Light Rail Performance for South Yorkshire Passenger Transport Executive undertaken by (Egis Semaly Ltd and Faber Maunsell 2003). Data is thought to related to the calendar year 2003

d UK/Dublin website data at www.tramlink.co.uk, www.centro.org.uk, www.railway-technology.com, www.supertram.com http://www.rpa.ie/en/Pages/default.aspx (last accessed Nov 2009)

e GIS Analysis of (Australian Bureau of Statistics 2006)

f Based on 2006 data and GIS analysis of (Statistics Canada 2007)

g Major statistical area, 2000, (U. S. Census Bureau 2000) http://www.census.gov/ (last accessed Nov 2009)

h GIS analysis of Central Statistics Office, Ireland, Census for 2006 at http://www.cso.ie/ (last accessed Nov 2009)

i GIS analysis of UK Census data for 2001, https://www.census.ac.uk/Default.aspx (last accessed Nov 2009)

j INSEE—National Institute of Statistics and Economic Studies—France, http://www.insee.fr/en/default.asp (last accessed Nov 2009)

k Transport Tomorrow Survey, (University of Toronto 2006)

l Center for Studies on Networks, Transport, Urban Planning and Public Works, France, http://www.certu.fr/spip.php?page=sommaire&lang=en (last accessed Nov 2009)

widely between systems. The secondary analysis explored ridership influences by removing this effect.

The R value is the multiple correlation of the regression model. The adjusted R^2 is the proportion of the variance in the dependent variable that is explained by the regression model. Beta (b) values represent the statistically standardised relationship between an explanatory variable and its dependent variable. These values show the relative influence of the variables within the model.

A major concern for reliable use of multiple regression methodology is the use of a large data set from which results can be based (Green 1991; Kelley and Maxwell 2003). Hair et al. (2006) provide some simple advice which enables an informed assessment of sample requirements;

- Simple regression can be effective with a sample size of 20, but maintaining power at 0.80 in multiple regression requires a minimum sample of 50 and preferably 100 observations for most research situations.
- The minimum ratio of observations to (explanatory) variables is 5–1, but the preferred ratio is 15 or 20–1, and this should increase when stepwise estimation is used (which is used in this analysis).

The eventual number of data points collated for the research was 57. This only just meets the minimum threshold of requirements suggested by this source hence clearly the statistical reliability of the method requires careful examination in its application.

Statistical reliability tests

A number of statistical tests were undertaken to assess the reliability of the analysis:

- Collinearity tests whether the predictors in the model are so highly correlated as to be interchangeable. Collinearity can inflate error values resulting in an unstable model. A Variance Inflation Factor (VIF) over 10 is cause for concern (Myers 1990).
- Casewise Diagnostics—This examines whether any cases are having an unusual influence on the model, either as a spurious outlier that conflicts with the model or as an unduly large influence on the model. In an ordinary sample only 5% of cases should have standardised residuals outside ± 2.
- Three criteria determine whether a single case is having undue influence on the model: Cook's Distance, leverage and Mahalanobis Distance. A Cook's Distance value greater than 1 is of concern (Cook and Weisberg 1982). Any values over 3 times the average leverage (k ? 1/n) are of concern (Stevens 2002). With a sample size below 50, Mahalanobis distances approaching 11 or over may be cause for concern (Barnett and Lewis 1978).
- Regression assumptions—The Durbin–Watson value tests to see if residual errors are uncorrelated. A value less than 1 or greater than 3 is cause for concern (Durbin and Watson 1951). Plots of standardised residuals can assess if residuals are skewed.

The results of these tests are reviewed as part of assessment of the reliability of modelling results.

Light rail system data

Table 3 shows the 57 routes for which light rail data was collated. Data collection was mainly based on data availability and ease of access. However it was hoped to explore how performance varied by right of way design (Melbourne and Toronto have significant mixed traffic operations whilst the other systems operate on significant segregated rights of way (Currie and Shalaby 2007)).

Table 4 shows summary data from each of the systems analysed. Ridership is considerably higher on European routes. Australian routes have less than half the rate of Boardings/Route Km but they are more effective than North American in terms of Boardings/Vehicle Km.

Residential and employment density are, perhaps surprisingly, considerably lower in Europe than in Australia or North America data. The residential density result for Australia is consistent with Hass-Klau and Crampton (2002), who found Melbourne to have an urban density within 600 m of tram routes which was higher than all European cities excluding Koln. The same source found densities for Dallas (whole system) and San Diego to be above many European Cities (e.g. Dublin and Rouen which are included in this analysis).

Car ownership in the light rail catchment is highest in North America. Australia is second with 18% less car ownership in light rail route catchments than North America closely followed by Europe (25% less than the US).

Table 3 Light rail routes selected for analysis

	Australia	North America	Europe
Number of route data points	24	21	12
Cities/ routes	Melbourne—109, 96, 86, 112, 19, 75, 59, 8, 16, 1, 3, 5, 48, 55, 67, 57, 72, 6,70, 64,78–79, 82	Toronto—501, 502–503, 504–508,505,506, 509–510, 511, 512	Dublin—Red, Green Croydon—Wimbledon, Beckenhamand New Addington lines[a]
	Adelaide (one route operated)	Boston, MA—Green Line	Sheffield—Meadowhall, Halfway and Middlewood lines[a]
	Sydney (one route operated)	Baltimore, MD—(one route operated)	Tyne and Wear—Green and yellow lines[a]
		Charlotte, NC—Lynx Light Rail	Midland Metro—Birmingham to Wolverhampton
		Houston, TX—Red Line	Manchester—Bury, Altrincham and Eccles Lines[a]
		Dallas, TX—DART[a]	
		Minneapolis, MN—Hiawatha Line	Nottingham—Hucknall Lyon—Line 1, Line 2
		Tacoma, WA—Tacoma Link	Montpellier—Line 1
		Buffalo, NY—Niagara Frontier	Rouen—Line 1
		Tampa, FL—Hillsborough	
		Portland, OR—MAX[a]	
		Sacramento, CA—Regional Transit	
		Sand Diego, CA—trolley[a]	
		Saint Louis, MO—Metrolink	

[a] These lines analysed as a group due to poor data availability

Australian light rail systems are characterised by low service levels, slow speeds, short stop spacing, a very low share of accessible stops and a low share of segregated right of way compared to both US and European systems. However there are more integrated fare systems in Australia. Interestingly the US systems have the highest service levels measured as vehicle trips p.a. They also have the longest stop spacing (841 m which is just over 3 times that in Australia). European service levels are almost as high as the US but con-siderably above those in Australia (there is 48% more service in European systems than in Australia and 80% more in the US than Australia). European systems have the highest speeds, a complete system of accessible stops and more than double the track segregation in Australia. However it has the lowest share of integrated fares.

Figure 1 illustrates the relationship between average speed, stop spacing and % seg-regated right of way. Neither relationship for the aggregate data is strong; the correlation

Table 4 Average of light rail route variable statistics by continent

	Australia	North America	Europe	
Dependent variables (ridership)				
Boardings/route Km	Mean	433,820	582,320	879,754
	SD	219,522	458,220	470,945
Boardings/vehicle Km	Mean	6.4	5.2	9.5
	SD	2.0	2.3	3.8
Explanatory variables				
Residential density	Mean	3,713	3,222	2,484
	SD	942	3,948	1,439
Employment density	Mean	7,611	2,500	1,506
	SD	2,455	3,296	1,098
Car ownership	Mean	434	531	396
	SD	53	156	78
Service level (vehicle trips/annum)	Mean	64,260	114,877	94,679
	SD	15,341	58,811	18,208
Average speed (kph)	Mean	17	18	25
	SD	2	7	6
Stop spacing	Mean	279	841	722
	SD	98	642	251
% Accessible stops	Mean	21	54	100
	SD	26	50	0
% Segregated right of way	Mean	24	39	54
	SD	23	47	40
Integrated fares	Percent	96	76	50

between stop spacing and speed is R = 0.61 and between stop spacing and share of ROW is R = 0.39.

Results

The following independent variables were submitted in a step-wise regression (as described in the section ‘‘the statistical model’’): average stop spacing, average speed, percent of each route with segregated track, employment density, residential density, car ownership, vehicle trips per annum, continent (dummy variables) and integrated ticketing (dummy variable). The dependent variable was boardings/route km.[1]

Table 5 shows the result set from the multiple regression analysis. The model proved statistically significant at the 99% level with five significant explanatory variables and an adjusted R^2 of 0.76 suggesting that 76% of the variance in Boardings/Route Km are explained by the model.

The most important factor influencing ridership is Service Level measured as vehicle trips p.a. (b = 0.74). This is closely followed by routes being in Europe (b = 0.72) a

[1] A log transformation of Boardings/Route Km (and, later, boardings/vehicle km) resulted in a lower R^2 so the untransformed variable was used.

Fig. 1 Average speed, stop spacing and segregated right of way share

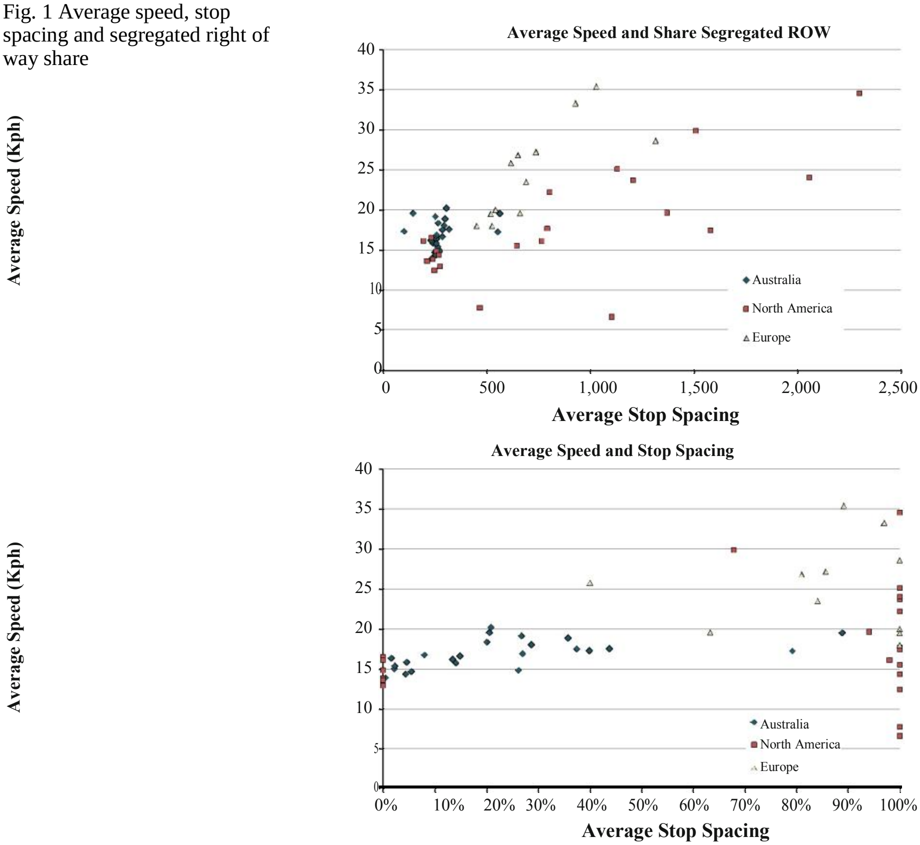

Table 5 Multiple regression results—boardings/route km

R^2 (adjusted) = 0.76; $F(5,51) = 37.0$, p \ 0.001				
Variable	Unstandardised weight	Standard error B	Beta (b), standardised	t-Value
Constant	34,336.831	148,520.297		
Explanatory variables				
Vehicle trips/annum	6.912	0.632	0.741	10.938
Europe	714,593.149	80,030.637	0.720	8.929
Speed	-27,728.595	5,525.895	-0.395	-5.018
Employment density	25.643	8.200	0.236	3.127
Integrated ticketing	241,967.756	72,057.231	0.244	3.358

dummy variable added to the analysis to test if regional circumstances could explain variations in performance. In this case the results firmly suggest that there is something special about the European context which makes ridership higher than in North America and Australia. Speed has the third strongest link with Boardings/Route Km (b = -0.40). Here slower speed is associated with higher ridership; although this coefficient appears to fail the logic test of expectations, previous work has found similarly paradoxical findings regarding running speed (Crampton 2002; Hass-Klau and Crampton 2002; Currie and Delbosc 2010). The reason is that dwell times increase as loadings rise. As a result busier routes tend to be slower.

Integrated ticketing (b = 0.24) and Employment Density (b = 0.24) were the other two significant explanatory variables in the model.

None of the other explanatory variables explored proved to be statistically significant in the modelling. Variables such as car ownership and residential density were excluded from the final model for this reason.

Statistical assessment of this model showed that collinearity was not a concern as all VIF values were below 2. Three data points were outside of ±2 residuals but this is within 5% of the sample size. No cases had a Cook's distance near 1 nor did any values have a leverage over 0.32 (three times the average leverage). Six cases had Mahalanobis Distances over the conservative cut-off of 11 but considering they were not suspect under the other criteria they are unlikely to be cause for concern. The Durbin–Watson value was 1.5, within accepted values. Plots of residuals versus predicted values show some degree of heteroscedasticity, which is unsurprising with the small sample size. This suggests that caution should be used when generalising these results to all light rail systems.

Secondary analyses

A series of secondary analyses were undertaken to further explore the results.

The first tested a different explanatory variable: Boardings/Vehicle Km, a measure of system 'service effectiveness'; i.e. the relative ridership performance for a given level of service (after Fielding 1987). The logic of this test was to try and control for the dominating influence of service level (measured as vehicle kilometres per annum) from the initial regression and to see if a different set of variables influences service effectiveness.

Boardings/vehicle km model

A model using Boardings/Vehicle Km as the dependent variable and with the same explanatory variables proved to be statistically significant. Again collinearity was not a concern in this model as all VIF values were below 2 and only 2 data points were outside of

±2 residuals. The Durbin–Watson value was 1.4, within accepted values. Plots of residuals versus predicted values showed acceptable levels of heteroscedasticity.

No cases had a Cook's distance near 1 but one case (Toronto 509/510) had a Mahalanobis Distance of over 18 and a leverage value of concern. These values indicate this data point is placing undue influence on the model so it was removed from this analysis. Table 6 shows the result set from the multiple regression analysis focussing on Boardings/Vehicle Km. This model was significant at 99% level with an overall adjusted R^2 of 0.67.

Almost all of the same variables were significant in this model however their relative influence was very different. Being in Europe was the most important influence on Boardings/Vehicle Km (b = 0.96) with speed the second most influential factor (b = -0.51). Employment density fourth (b = 0.47) and integrated ticketing came third (b = 0.37). Track segregation was significant in this model where it was not significant in the last (b = 0.28). Interestingly service level (vehicle trips p.a.) was again significant (b = 0.17) implying that a higher frequency results in higher service effectiveness.

Table 6 multiple regression results—boardings/vehicle km

R^2 (adjusted) = 0.67; F(6, 49) = 19.2, p \ 0.001				
Variable	Unstandardised weight	Standard error B	Beta (b), standardised	t-Value
Constant	4.104	1.509		
Explanatory variables				
Europe	6.959	0.706	0.961	9.863
Speed	-0.259	0.049	-0.505	-5.342
Employment density	0.00039	0.0001	0.471	3.910
Integrated ticketing	2.688	0.649	0.371	4.140
Track segregated	0.020	0.009	0.279	2.213
Vehicle trips/annum	0.000012	0.00001	0.174	2.037

Non-European models

A separate analysis was undertaken of non-European systems however the validity of this analysis is questionable given the much smaller sample size in the non-European data (45 compared to 57).

Conclusion and discussion

This paper has explored the relative influence of factors affecting light rail ridership on 57 light rail routes in Australia, Europe and North America through an empirical examination of route level data. A multiple-regression analysis predicting route level ridership (boardings per route km) using residential and employment density, car ownership, service level, speed, stop spacing, share of accessible stops, share of segregated right of away and integrated fares was undertaken. This established a statistically significant model (99% level, R^2 = 0.76) with five significant variables including, in order of influence: service level (b = 0.74), routes being in Europe (b = 0.72), speed (b = -0.40), integrated tick-eting (b = 0.24) and employment density (b = 0.24). A model predicting boardings per vehicle km (R^2 = 0.67) resulted in almost the same set of variables but in a different order of influence: Europe (b = 0.96), speed (b = -0.51), employment density (b = 0.47), integrated ticketing (b = 0.37), segregated right of way (b = 0.28) and vehicle trips/ annum (b = 0.17). In general statistical tests of the model have confirmed its reliability however the sample is considered small and caution should be adopted in using findings for other light rail systems.

The strength of service level in these results is consistent with a range of findings from previous route level research (including Stopher 1992; FitzRoy and Smith 1998; Kain and Liu 1999). In effect the quantum of service provided acts to drive the ridership that results, largely irrespective of other factors. It is interesting to note that service level was a significant (though small) influence on boardings per vehicle km, which suggests routes with a higher service level are more efficient.

The strength of the European dummy variable (b = 0.72) is intriguing. On first glance one would expect that factors such as residential and employment density or car ownership would explain why European routes have higher ridership. However this analysis has already allowed for these differences, suggesting the European ‘‘bonus’’ to light rail ridership is independent of these influences. Previous models (Hass-Klau and Crampton 2002) suggest that high light rail ridership could be associated with pedestrianisation which may explain the ‘European factor’ in this analysis. Another possible explanation is that public transport mode share is considerably higher in Europe (12%/15% in France/UK) compared to 5% in Australia, 3% in the US and 8% in Canada (Kenworthy and Laube 2001). Higher mode share may be a proxy for a greater network effect or a culture of transit use. Either way the results point to important non-measured influences in the European context which are worthy of further research.

The negative link between speed and ridership is also consistent with previous research (Crampton 2002; Hass-Klau and Crampton 2002; Currie and Delbosc 2010). In effect light rail systems with lower speed have higher ridership. This is likely to be a combined influence of longer boarding/dwell times due to higher ridership and the fact that inner urban areas (e.g. CBD’s) have high ridership, shorter stop spacing and hence slower speeds.

The strength of integrated ticketing ($b = 0.24$) as a ridership driver is also consistent with previous research (e.g. Hass-Klau and Crampton 2002) while employment density ($b = 0.24$) has also been found to be an important driver (Kain and Liu 1999). It is interesting to note that employment density, not residential density, was the significant predictor in this model. Residential density has been identified as a significant route-level ridership driver in previous research (Seskin and Cervero 1996; Johnson 2003). Emerging evidence is sug-gesting that residential density is not a primary driver of transit ridership and that employment density may be the more important influence (Chen et al. 2008; Mees 2009).

An additional secondary analysis of the non-European data was undertaken but had statistical concerns due to the lower number of data points.

Overall the research findings stress the importance of providing a high level of service as a major driver of light rail ridership. The ‘European Factor’ is also important but requires further research to clarify the specific aspects of European systems which result in higher ridership. Mode share and cultural/behavioural influences have been suggested as possible reasons for this influence. This would be a fruitful area for exploration in future research. Further research, including a wider range of routes and data points, would also improve concerns over the reliability of the modelling.

Acknowledgments Parts of this research was conducted while Dr. Ahern was on sabbatical at the Institute of Transport Studies, Monash University. The authors would like to thank the Urban Institute in University College Dublin for part-funding this sabbatical. The authors would also like to acknowledge the help of Susanna Schmidt of the Urban Institute Dublin for help with processing data. In addition Karen Woo of the University of Toronto assisted collecting some of the North American data whilst on a visit to Monash from the University of Toronto. We would also like to thank Prof Patrick Bonnel Professor of Transport Economics, ENTPE France for advice on sources for collecting data from France.

References

Australian Bureau of Statistics (2006). Census of population and housing

Babalik-Sutcliffe, E.: Urban rail systems: analysis of the factors behind success. Transp. Rev. 22, 415–417 (2002)

Balcombe, R., Mackett, R., et al. The Demand for Public Transport; A Practical Guide. Transport Research Laboratory, Crowthorne (2004)

Barnett, V., Lewis, T.: Outliers in statistical data. New York, Wiley (1978)

Capineri, C., Kamann, D.: Synergy in networks: concepts. In: Button, K., Nijkamp, P., Priemus, H. (eds.) Transport Networks in Europe—Concepts, Analysis and Policies. Edward Elgar Publishers, Cheltenham, UK (1998)

Chen, C., Gong, H., et al.: Role of the built environment on mode choice decisions: additional evidence on the impact of density. Transportation 35(3), 285–299 (2008)

Cook, S.A., Weisberg, S.: Residuals and Influence in Regression. Chapman & Hall, New York (1982)

Crampton, G.: International Comparison of Urban Light Rail Systems; the Role of Integrated Ticketing, Pedestrianisation and Population Density. European Regional Science Association (ERSA) 2002. Dortmund, Germany (2002)

Currie, G., Delbosc, A.: Understanding ridership drivers for bus rapid transit systems in Australia. Australasian Transport Research Forum, Canberra, Australia (2010)

Currie, G., Shalaby, A.: Success and challenges in modernising streetcar systems—experience in Melbourne and Toronto. Transp. Res. Record J. Transp. Res. Board Wash. DC 2006, 31–39 (2007)

Denant Boemont, L., Mills, G.: Urban light rail: intermodal competition or coordination? Transp. Rev. 19, 241–253 (1999)

Durbin, J., Watson, G.S.: Testing for serial correlation in least squares regression, II. Biometrika 30, 159–178 (1951)

Edwards, M., Mackett, R.: Developing new urban public transport systems: an irrational decision-making process. Transp. Policy 3(4), 225–239 (1996)

Egis Semaly Ltd and Faber Maunsell. Comparative Performance Data from French Tramsway Systems—Final Report, South Yorkshire Passenger Transport Executive (2003)

Evans, J.E.: Traveler Response to Transportation System Changes Chapter 9—Transit Scheduling and Frequency. Transportation Research Board, Washington, DC (2004)

Fielding, G.J.: Managing Public Transit Strategically. Jossey-Bass Inc Publishers, San Francisco (1987)

FitzRoy, F., Smith, I.: Public transport demand in Freiburg: why did demand double in a decade? Transp. Policy 5, 163–173 (1998)

Green, S.B.: How many subjects does it take to do a regression analysis? Multivar. Behav. Res. 26(3), 499–510 (1991)

Hair, J., Black, N., et al.: Multivariate Data Analysis, 6th edn. Prentice Hall, New Jersey (2006) Hass-Klau, C., Crampton, G.: Future of urban Transport; Learning from Success and Weakness; Light Rail. Environment and Transport Planning, Brighton, UK (2002)

Johnson, A.: Bus transit and land use: illuminating the interaction. J Public Transp. 6(4), 21–39 (2003)

Kain, J., Liu, Z.: Secrets of success: the large increases in transit ridership achieved by Houston and San Diego transit providers. Transp. Res. A 33, 601–624 (1999)

Kain, M., Barranda, A., et al.: Factors influencing light-rail station boardings in the United States. Transp. Res. A 38, 223–247 (2004)

Kelley, K., Maxwell, S.: Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. Psychol. Methods 8(3), 305–321 (2003)

Kenworthy, J., Laube, F.: The Millennium Cities Database for Sustainable Transport. International Union of Public Transport (UITP), Brussels (2001)

Knowles, R.: What future for light rail in the UK after ten year transport plan targets are scrapped? Transp. Policy 14, 81–93 (2007)

Mackett, R., Babalik-Sutcliffe, E.: New urban rail systems: a policy based technique to make them more successful. J. Transp. Geogr. 11, 151–164 (2003)

McCollom, B.E., Pratt, R.H.: Traveler Response to Transportation System Changes Chapter 12—Transit Pricing and Fares. Transit Cooperative Research Program. Washington DC, Transportation Research Board, TCRP Report 95 (2004)

Mees, P.: How dense are we? Another look at urban density and transport patterns in Australia, Canada and the USA. Road Transp. Res. 18(4), 58–67 (2009)

Myers, R.: Classical and modern regression with applications, 2nd edn. Duxbury, Boston, MA (1990) Pratt, R.H., Evans, J.E.: Traveler Response to Transportation System Changes Chapter 10—Bus Routing and Coverage. Transit Cooperative Research Program, Washington, DC (2004)

Seskin, S., Cervero, R.: Transit and urban form. Federal Transit Administration, Washington, DC (1996)

Statistics Canada: Ontario: 2006 Community Profiles, 2006 Census, Ottawa (2007)

Stevens, J.P.: Applied Multivariate Statistics for the Social Sciences, 4th edn. Erlbaum, Hillsdale, NJ (2002) Stopher, P.: Development of a route level patronage forecasting method. Transportation 19, 201–220 (1992) U. S. Census Bureau: Profile of selected social characteristics (2000)

University of Toronto: Transportation Tomorrow Survey. Data Management Group, University of Toronto, Toronto, Canada (2006)

Author Biographies

G. Currie is Chair of Public Transport and leads the Public Transport Research Group at the Institute of Transport Studies Monash University. He has over 30 years experience as a public transport planner, researcher and consultant and is developer and director of the ‘World Transit Research’ clearinghouse (www.worldtransitresearch.info).

A. Ahern graduated from Civil Engineering in Trinity College Dublin in 1998 before completing a PhD in the Centre for Transport Studies, University College London in 2001. She is now a senior lecturer in University College Dublin. Her research interests include the provision of transport infrastructure and its impacts on development, urban form, travel patterns and social exclusion.

A. Delbosc graduated with a research-based Master of Social Psychology from Harvard University. Her psychology work has been published in the Proceedings of the National Academy of Sciences. She is presently a research fellow in the Monash University Institute of Transport Studies (Melbourne, Australia) researching the social implications of public transport and transport disadvantage.