# International Comparison of Urban Light Rail Systems:

# the roles of integrated ticketing, pedestrianisation, and population density

Graham Crampton,
Centre for Spatial and Real Estate Economics,
School of Business,
Reading University, Whiteknights, Reading, RG6 6AW, England

Paper (number 167) prepared for the ERSA2002 Conference, Dortmund, August 27th-31st, 2002.

## 1. Introduction

Light rail or modern tram has recently been enjoying something of a revival in a number of countries, including Britain. It seems to offer many of the advantages of a rail-based public transport mode, especially a greater degree of separation from road traffic resulting in greater reliability of service. It can also embody the `modern image' of a city for marketing purposes more effectively than the bus, while being much cheaper than full underground rail in terms of infrastructure and operating cost. Many European countries now have urban public transport strategies in which light rail plays a prominent part, and in North America too there have been notable examples of modern trams being at the heart of major cities attempting to come to terms with their transport congestion problems.

However, fundamental difficulties have also arisen. Adding rail-based public transport, which thrives best in a high-density urban structure, to car-dominated urban systems can often give disappointing results. Also, the interrelationships that exist with the rest of the public transport supply, and the way in which the whole system is priced and marketed can present a substantial group of difficult overlapping issues.

The research project reported in Hass-Klau and Crampton (2002) assembled a body of data on light rail and public transport in 24 case study cities from 7 countries. The data was used as background to a discussion of the role of light rail in modern urban transport systems, and the construction of a ranking of the 24 case study cities based on their performance in various dimensions detailed below.

In this paper we summarise the basic econometric analysis contained in the research report (Hass-Klau and Crampton, ibid), and develop further the quantitative description of performance, together with analysis of the inter-urban variability. In particular, we experiment with cluster analysis to get further insights into the groups of success, near-success and weakness that we find in such a relatively small group of cities, each of which has its own complex public transport history.

The topic of how best to quantify the level and quality of public transport service has not been a heavily researched topic. Much more well developed has been the literature on the links between transport investment and urban economic development, including the relationship with compact urban form (Dickens (1992)). British work in particular has often been more concerned with the inconsistencies of how light rail projects were financed in the 1980s and 1990s relative to financial decisions in the local bus sector. Experienced British public transport operators have expressed regret that a more consistent appraisal framework was not in place at that time in Britain (Tyson (1992)).

There is also a strand of published discussion of alternative methods of financing rail infrastructure, a debate has relevance both to inter-city and urban rail investment, and is extremely topical for Britain. Ferreira (1997), writing against an Australian inter-city background, suggests that the best way forward would be for rail infrastructure to be owned by a new joint venture company, with equity owned jointly by the operators and governments (whether Central or State-level). In Britain this has most urban relevance to the effective coordination of inter-city rail services with urban rail systems, which has been problematic in the major conurbations.

Recently, American transport academics have focused more closely on the best indicators with which to measure public transport (or `transit') performance (Li and Wachs (2000)). In particular, they examine the indicators used by a sample of the US urban rail systems; they suggest that a clearer distinction between the cost efficiency and measures of customer or service effectiveness is crucial both in monitoring existing systems and in rational decision-making over new investments. The multi-dimensional monitoring of public transport performance comes closest to the approach taken in this paper.

## 2. Econometric Analysis of the Average Ranking Performance Index

At first we were interested in obtaining an overall index of performance of each city's light rail system, in the context of its public transport system as a whole. We could then examine whether we could (despite the relatively small sample of cities) empirically explain a substantial proportion of the variation in performance by the `hard' and `soft' factors we would expect to play a role. For instance, what is the importance of average speed of light rail service compared to the fare levels charged by the light rail system in explaining performance? Or what is the relative role of performance features specific to the light rail service (such as headway at peak hours) compared to details of urban population density structure or city centre car parking? We will first discuss in more detail the dependent and explanatory variables we used, and then go on to present empirical results.

*Dependent and explanatory variables*

We were first interested in trying out `direct' indicators of light rail performance, as well as those based on a ranking index that we had worked on in the course of the research. The direct measures, which are quite closely related to each other, were:

- Light rail passenger number per urban population;
- Light rail passenger number per light rail track-km;
- Light rail passenger-km per light rail track-km.
- Public transport passenger number per urban population;
- Annual growth of light rail passenger number;
- Growth over 1985-99 in public transport passenger number per urban population;

The simpler of these are probably those based just on an estimate of annual light rail trips, relative to either urban population or the physical size of the system measured by its track-km length. The third above is perhaps closer to the actual output of a light rail operator, given that `passenger-km' weights each trip by its length; however, the data would in many cases be based on surveys and estimates of travel distance which might be less accurate than simple passenger counts.

In fact our results below showed that it was easier to obtain a good statistical fit with performance indices based on the overall ranking of the cities' light rail systems. This in effect measures their performance not directly, but relative to each other. An overall ranking of the study cities had been computed based on the above 6 separate indicators, including light rail and total public transport trip numbers, and their growth rates. These indicators to some extent overlapped, and tended to `double count' light rail as part of the public transport total. In most cases, newer systems would do very well in terms of annual growth, as they built up passengers from a low initial level.

With this ranking indicator, we could use either the average ranking itself (starting at 4.33 with Freiburg, finishing with Dallas at 19.33), or even more simply place the cities in order and use the overall ranking itself (starting with Freiburg, Zürich, Köln at 1,2,3 …). This does not make a lot of difference in practice since the two are closely correlated with each other, but the average ranking itself works best empirically as the dependent variable. [Since a ranking of 1 would represent the best system, it was thought clearer to use a `performance index', so that variables affecting performance positively would get a positive sign. We therefore just `reversed the rank' by defining a Performance Index as (24 – Average Rank). This does not affect any of the statistical estimations other than reversing the signs of the explanatory coefficients so as to be more intuitive.]

The project had also collected information on a range of explanatory variables that could affect the performance of light rail. It is perhaps useful to group them under headings related to:

<u>Physical light rail operation, including:</u>

- average speed of light rail service;
- newness, i.e. percentage light rail vehicles less than 5 years old;
- peak service headway in minutes

<u>Price and marketing of light rail, including:</u>

- **% of light rail passengers using a Travel Card;**
- **monthly light rail fare relative to the country's GDP/capita;**

<u>Average accessibility of light rail routes</u>

- network density (total light rail track-km per urban population)**;**
- **estimated population in 300m light rail corridor either side of lines per km of track;**
- average stop distance on light rail system;

<u>Urban planning and restriction on car use:</u>

- **Pedestrian zone length per city population;**
- Number of park and ride spaces per km of light rail track

- Number of car parking spaces in the city centre relative to estimated city centre area.

Hours of Service in Light Rail

Number of hours run at peak frequency;

Total hours of service per weekday.

The four factors in bold are those which, on first analysis, seem to have a more significant effect on the overall indicator of success on their own, before considering their combined effect with other variables. The three strongest of these - travel card usage, the corridor density, and size of pedestrian area have positive effects, i.e. they improve the likelihood of the system scoring well in the combined measure of success.[1]

In contrast, the level of the monthly fare, which we calculated relative to the country's per capita GDP, tended to reduce the success of light rail, although we found this less significant than travel card use as a marketing variable. Experimenting with the monthly fare was intended to reflect the fact that many of the best systems deliberately set the single fare high, so that a travel card at the monthly fare is the most attractive option; not surprisingly the single fare was a much weaker variable. The use of GDP/capita tried to reflect what passengers in countries of different income levels would consider `high' or `low' fares.

With a small sample of 24 cities, the statistical estimations should be interpreted with caution. It is at first surprising that the variables relating to the physical operation of light rail, such as speed and headway, seem to be much less significant. Given that the Swiss and German light rail systems at the top of the performance ranking tended to be frequent, relatively slow and dense but with quite old vehicles, perhaps we should not be so astonished. We can now discuss the statistical results in more detail by firstly giving the simple correlation coefficients between the dependent variable we used above and the various explanatory variables.

It is a sobering thought for light rail operators that the simple correlation coefficients between average light rail performance and either average speed or vehicle newness are adverse (that is, the more successful systems have older, slower vehicles) and there seems to be little relationship with peak service headway.

The strongest simple correlations are between performance and travel card use, the pedestrian zone length, the population density in a 300m corridor either side of the light rail lines, a short average stop distance, and the adverse correlation with average vehicle speed.

Table 1: Simple Correlation Coefficients between the Dependent Variable `Performance Index' based on average rank and each of the explanatory variables as on list above.

| Explanatory variable | Notation | Simple Correlation Coefficient r with Performance Index Dependent Variable |
|---|---|---|
| Average speed of light rail service | AVSPLR | - 0.49* |
| Vehicle newness (% light rail vehicles less than 5 years old) | PCNEWLRV | - 0.07 |
| Headway in minutes, average a.m. peak service | PHEADWAY | -0.09 |
| % of light rail passengers using a Travel Card; | TRAVCARD | +0.64** |
| Monthly fare relative to the country's GDP/capita; | MFAREGDP | - 0.19 |
| Light rail network density | NETDENS | +0.38 |
| Estimated population in 300m. light rail corridor either side of lines per km. of track; | CORDEN3 | +0.52** |
| Average light rail stop distance | AVSTDIST | -0.47* |
| Pedestrian zone length per city population. | PZLPP | +0.51* |
| Park and ride spaces per km. of track | PRPERTK | -0.01 |
| Number of car parking spaces in the city centre relative to estimated city centre area. | PSPKM2 | - 0.36* |
| Number of hours run at peak frequency | PHRSERV | 0.20 |
| Total hours of service per weekday | RUNTIME | 0.06 |
| Number of Suburban and S-Bahn rail stations relative to urban population; | SUBRLIND | - 0.02 |

** Statistically significant (2-tailed) at 99% confidence; * significant at 95% confidence.

The densest networks (in terms of total track length per capita) also tend to perform better, and those in cities whose centres are amply supplied with car parking perform worse, though at a lower level of significance.

We can also summarise here without going into detail the main correlations between the explanatory variables themselves:

- Pedestrian zone length is itself correlated with high light rail corridor density but low relative provision of city centre car parking;
- High light rail corridor population density is found with short average stop distances and low average speed;
- Systems with a high % of new vehicles tend on average to be faster;
- The systems with highest network density tend to be slower, with high travel card use, but have fewer park and ride spaces provided per km of track.

Out of the large number of regressions we carried out, we give a brief selection of results using the same Performance Index dependent variable (recalling that this was defined as [24 – Average Rank]). Regression results we have not given had either low levels of significance or the wrong sign.

**Table 2: Selected Multiple Regression Estimations using Performance Index as Dependent Variable.**

| Const. | TRAVCARD | PZLPP | CORDEN3 | MFAREGDP | PRPERTK | AVSPLR | PSPKM2 | NETDENS | % AdjRsq |
|---|---|---|---|---|---|---|---|---|---|
| 7.3 | 0.078<br>(4.3) | 0.67<br>(1.9) | 0.00094<br>(1.4) | -0.015<br>(1.1) | | | | | 60.2 |
| 6.5 | 0.076<br>(4.5) | 0.59<br>(1.6) | 0.0012<br>(1.6) | -0.016<br>(1.2) | 0.0076<br>(1.2) | | | | 61.1 |
| 8.7 | 0.074<br>(3.8) | 0.65<br>(1.8) | 0.00084<br>(1.2) | -0.013<br>(0.9) | | -0.053<br>(0.6) | | | 58.8 |
| 7.5 | 0.076<br>(3.5) | 0.63<br>(1.3) | 0.0011<br>(1.3) | -0.016<br>(0.9) | | | -0.00002<br>(0.2) | | 58.8 |
| 6.8 | 0.073<br>(3.4) | 0.71<br>(1.9) | 0.00090<br>(134) | -0.012<br>(0.8) | | | | 0.37<br>(0.5) | 58.6 |

(t-statistics are given in parentheses).

We see from Table 11 that the regression with a constant and travel card use, pedestrian zone length per capita, 300m. corridor density, and monthly relative fare had an adjusted R squared of about 60%. The t-statistics show that the statistical significance for the first two explanatory variables (travel card use and pedestrian zone length per capita) is 90% or better, and lower for the other two which have t-statistics of around 1.1 - 1.4. With such a relatively small sample, we should perhaps not be too demanding in terms of either the number of explanatory variables or their significance levels. It is still interesting and revealing that the strongest statistical explanation of light rail success comes through the role of successful marketing through travel card use, traffic-restraining pedestrianisation, and ease of physical access as represented by the population density structure.

We will now experiment with cluster and factor analysis techniques, which can sometimes add further insight to statistical models with many variables or highly collinear data.

## 3. Grouping of Case Study Cities using Cluster Analysis

When we developed the data base used in the project reported in Hass-Klau and Crampton (2002, op cit) we made no use of cluster analysis so as to keep the empirical work as accessible as possible to less technical readers. The overall ranking used in the OLS regressions above was based on rankings in 6 dependent variables, simply averaged and turned into a performance index. We noted that 4 of these performance indicators were `levels' variables, and the remaining 2 were `growth' variables.

It seems appropriate to follow this up by using the raw data in cluster analysis, to gauge whether intuitively appealing clusters appeared, and whether they differed much from the 3 broad groups that emerged from our overall rankings. Cluster analysis also allows us to see easily which case study cities were closest together in terms of the 6 indicators, that is which were grouped together earliest in the cluster `agglomeration schedule' produced as output. It also of course gives full detail on the order of clustering and the degree of closeness, as measured by the distance measure and clustering technique chosen. [2] It was also easy to separate the `levels' and `growth' indicators, so as to see the different clustering outcomes with these different groups of dependent variables.

The software allows a choice of 7 alternative clustering techniques and, although we experimented with all of them, we will report those using squared Euclidean distance, z-score standardisation and Ward's method.[3] If we impose a choice of 3 clusters over all 6 indicators, we find one remarkable outcome: Zürich is in a `class of its own', that is in a cluster by itself. The overall output is given in Table 3.

**Table 3. Cluster Analysis applied to all 6 dependent variables, 24 case study cities. (Ward's method, z-scores, Euclidean distance squared).**

Cluster 1: Zürich;

Cluster 2: Freiburg, Köln, Basel, Hannover, Düsseldorf, Dresden, Den Haag, Tyne and Wear, Leipzig;

Cluster 3: Calgary, Portland, Essen, Saarbrücken, Bremen, San Diego, Melbourne, Göteborg, Greater Manchester, Birmingham/West Midlands, Dallas, Sacramento, Strasbourg, Rouen.

We find some surprises in this use of the `raw numbers' and a formal clustering technique, rather than the use of simple ranking and a more ad hoc grouping. In particular, we might be surprised to find Dresden and Leipzig in the second cluster, and Tyne and Wear is the only British-American system to find its way there. Tyne and Wear is relatively successful only when monitored by Light Rail Passenger Km. per Track Km, reflecting the fact that it is a fairly spacious system with relatively long average trip length. Dresden and Leipzig, despite their substantial declines in passenger numbers, still have enough passengers remaining to put them in with the stronger systems.

Since the `agglomeration schedule' for this clustering technique allows an easy reading of what the outcome would have been with 4 or more clusters, it is worth noting that if there had been a $4^{th}$ separate cluster, it would have been Tyne and Wear, Dresden and Leipzig. (And if there had been a $5^{th}$ it would have been the `small light rail system but high growth' cluster of Portland, Saarbrücken, San Diego, Birmingham/West Midlands, Sacramento).

Zürich and its `cluster of one' is quite striking; it probably reflects that the Light Rail Passengers per capita is nearly twice as big as any other city, likewise the total public passenger numbers. If ever a simple factual confirmation were needed that public transport is not `poor people's transport', then Zürich provides it, although it is clearly the outcome of decades of investment, car restraint, and commitment to attract the professional and middle class user.

At the other end of the `agglomeration schedule', the first `pairings' made as part of the clustering process showing close similarity with respect to these 6 indicators are as follows: (one should recall that these use the 6 performance indicators standardised in z-score form, that is with each variable measured in terms of standard deviations away from its mean):

- Bremen with Göteborg;

- Köln with Hannover;
- Düsseldorf with Den Haag;
- Essen with Melbourne.

These pairings would not cause much surprise, except perhaps for Melbourne to discover that it is so similar to Essen, the heartland of the Ruhr.

Given that some questions might be raised about the mixing of the four level indicators (light rail passenger number per urban population, light rail passenger number per light rail track-km., light rail passenger-km per light rail track-km., public transport passenger number per urban population) with the two growth indicators in our overall clustering, we also experimented with running the same clustering analysis for the level indicators by themselves. Using the same cluster analysis specification, Table 4 shows the result of the 3 cluster outcome.

**Table 4. Cluster Analysis applied to 4 dependent variables on levels, 24 case study cities. (Ward's method, z-scores, Euclidean distance squared).**

Cluster 1: Zürich;

Cluster 2: Freiburg, Köln, Basel, Hannover, Düsseldorf, Dresden, Den Haag;

Cluster 3: Tyne and Wear, Leipzig, Calgary, Portland, Essen, Saarbrücken, Bremen, San Diego, Melbourne, Göteborg, Greater Manchester, Birmingham/West Midlands, Dallas, Sacramento, Strasbourg, Rouen.

The only difference is that Tyne and Wear (Newcastle) and Leipzig move into cluster 3. Zürich is still in a `class of its own'. Cluster 2 in terms of the four levels variables becomes smaller and exclusively German-Swiss-Dutch.

Finally, a similar cluster analysis using the other two `growth' variables (Annual growth of light rail passenger number, and growth over 1985-99 in public transport passenger number per urban population) gives a 3-cluster breakdown that is somewhat different.

**Table 5. Cluster Analysis applied to 2 dependent variables on light rail and public transport growth rates, 24 case study cities. (Ward's method, z-scores, Euclidean distance squared).**

Cluster 1: Portland, Saarbrücken, San Diego, Birmingham/West Midlands, Strasbourg;

Cluster 2: Zürich, Freiburg, Köln, Basel, Hannover, Düsseldorf, Calgary, Essen, Den Haag, Bremen, Melbourne, Göteborg, Greater Manchester, Dallas, Sacramento, Rouen;

Cluster 3: Dresden, Tyne and Wear, Leipzig.

We now find the newer, smaller, but very rapidly growing systems as a small light rail high growth group (Birmingham's rapid light rail growth gets it into this group despite its overall public transport decline). Tyne and Wear joins the two former East German cities in the severity of its overall light rail and public transport decline since 1985.

We also experimented with the other 6 clustering techniques available under SPSS, with simple unsquared Euclidean distance and with more than 3 clusters, but do not have space to discuss this further here.

## 4. Use of Factor Analysis in Grouping and Weighting of Dependent and Explanatory Variables

Factor analysis can prove useful when a limited sample size places restrictions on the use of explanatory or a range of dependent variables. In particular, it can provide an appropriate weighting within a group of explanatory variables (such as `user accessibility', or `physical performance'), using data extraction, eigenvalues, and rotation techniques that have by now become standard statistical techniques.

We have the possibility first to carry out factor analysis for the six variables which make up our overall performance discussed previously (Table 6); then we considered it appropriate to group the explanatory variables together in order to carry out further factor analysis on them.

**Table 6. Factor Analysis of Six Dependent Variables entering into Light Rail and Public Transport Performance Index, defined above.**

Communalities

| Variable | Initial | Extraction |
|---|---|---|
| LRPASSPC | 1 | .891 |
| PTPASSPC | 1 | .896 |
| LRPPERTK | 1 | .882 |
| ANNGRLRP | 1 | .744 |
| PTGRPP14 | 1 | .778 |
| LRPKPRTK | 1 | .709 |

Total Variance Explained (2 components with eigenvalues greater than 1)

| | Initial Eigenvalues | | |
|---|---|---|---|
| Component | Total | % of Variance | Cumulative % |
| 1 | 3.40 | 56.66 | 56.66 |
| 2 | 1.50 | 25.01 | 81.67 |
| | Extraction Sums of Squared Loadings | | |
| 1 | 3.40 | 56.66 | 56.66 |
| 2 | 1.50 | 25.01 | 81.67 |
| | Rotation Sums of Squared Loadings | | |
| 1 | 3.307 | 55.11 | 55.11 |
| 2 | 1.593 | 26.56 | 81.67 |

Component Matrix (2 components extracted by principal component analysis)

| | Component | |
|---|---|---|
| | 1 | 2 |
| LRPASSPC | .943 | -.052 |
| PTPASSPC | .946 | -.024 |
| LRPPERTK | .885 | .314 |
| ANNGRLRP | -.412 | .758 |
| PTGRPP14 | -.096 | .877 |
| LRPKPRTK | .808 | .236 |

Rotated Component Matrix (Method: Varimax with Kaiser Normalisation, converged in 3 iterations)

| | Component | |
|---|---|---|
| | 1 (PLEVELSFAC) | 2 (PGROWTHFAC) |
| LRPASSPC | .908 | -.260 |
| PTPASSPC | .917 | -.233 |
| LRPPERTK | .933 | .110 |
| ANNGRLRP | -.234 | .830 |
| PTGRPP14 | .100 | .876 |
| LRPKPRTK | .840 | .051 |

We see that the factor analysis on the dependent variables gives two components after extraction (having retained just those with eigenvalues exceeding one). The first component clearly weights heavily on the `levels' variables, and the second on the `growth' variables. This gives us the possibility of storing the component 1 and 2 factor values, and doing regression estimates on the level or growth elements of performance.

For the explanatory variables, we grouped them first and then did factor analysis on each group. In most cases, only one component was extracted and so no rotation took place. In other cases, only two variables were present in a group of explanatory variables, so that the formation of factor was statistically trivial (the factor is an equal weighting of the two variables).

**Table 7. Factor Analysis of Explanatory Variables**

(a) Light Rail Physical Performance Factor

Communalities

| Variable | Initial | Extraction |
|---|---|---|
| AVSPLR | 1 | .572 |
| PCNEWLRV | 1 | .637 |
| PHEADWAY | 1 | .397 |

Total Variance Explained (one component with eigenvalue greater than 1)

| | Initial Eigenvalues | | |
|---|---|---|---|
| Component | Total | % of Variance | Cumulative % |
| 1 | 1.606 | 53.55 | 53.55 |

Component Matrix

| | Component (denoted LRPERFAC) |
|---|---|
| | 1 |
| AVSPLR | .756 |
| PCNEWLRV | .798 |
| PHEADWAY | -.630 |

Here in Table 7 we find that the single light rail performance factor loads strongly and positively on high average speed, high percentage new vehicles, and short operating headways.

Continuing through the explanatory variables, the `marketing variables' (TRAVCARD and MFAREGDP, i.e. the % of passengers using travel cards and the monthly fare relative to GDP per capita) form a trivial factor with equal weighting (+0.711 for TRAVCARD and – 0.711 for MFAREGDP). Although using this factor (denoted MKGFAC) has the effect of imposing a restriction that the coefficients are equal and opposite, we experimented with it for consistency's sake. We do of course have the option of including variables individually, without the restriction.

Table 7(b) shows the formation of a factor for variables related to car use restriction, namely PRPERTK, PSPKM2, PZLPP (i.e. park and ride spaces per track km., city centre parking spaces per km2 city centre area, pedestrian zone length per person); again only one factor had an eigenvalue over unity and hence one was extracted.

(b) Car Use Restriction Factor

Communalities

| Variable | Initial | Extraction |
|---|---|---|
| PRPERTK | 1 | .154 |
| PSPKM2 | 1 | .756 |
| PZLPP | 1 | .639 |

Total Variance Explained (one component with eigenvalue greater than 1)

| | Initial Eigenvalues | | |
|---|---|---|---|
| Component | Total | % of Variance | Cumulative % |
| 1 | 1.549 | 51.64 | 51.64 |

Component Matrix

| | Component (denoted CRESTFAC) |
|---|---|
| | 1 |
| PRPERTK | .393 |
| PSPKM2 | -.869 |
| PZLPP | .799 |

The car use restriction factor (CRESTFAC) is loaded strongly on pedestrian zone length and negatively on city centre parking, and less strongly on park and ride provision.

We experimented as an alternative to CRESTFAC with the presence of suburban rail stations (both as a 0-3 index and in per capita form), but the results were not interesting enough to give in detail here.

We also experimented with the two `hours of service' variables, one measuring the length of peak hour service (PHRSERV) and the other measuring the total running time on a weekday (RUNTIME). The factor combining the two would simply be an average of the two (details not given). We thought it would be more interesting and appropriate to include these two variables in the more general accessibility factors, reported next, which would then contain elements of physical accessibility to the lines, density of network and stop density, together with hours-of-service accessibility.

(c) General Accessibility Factors

Communalities

| Variable | Initial | Extraction |
|---|---|---|
| PHRSERV | 1 | .535 |
| RUNTIME | 1 | .748 |
| CORDEN3 | 1 | .663 |
| NETDENS | 1 | .762 |
| AVSTDIST | 1 | .889 |

Total Variance Explained (2 components with eigenvalues greater than 1)

| | Initial Eigenvalues | | |
|---|---|---|---|
| Component | Total | % of Variance | Cumulative % |
| 1 | 2.31 | 46.14 | 46.14 |
| 2 | 1.29 | 25.79 | 71.93 |
| | Extraction Sums of Squared Loadings | | |
| 1 | 2.31 | 46.14 | 46.14 |
| 2 | 1.29 | 25.79 | 71.93 |
| | Rotation Sums of Squared Loadings | | |
| 1 | 2.104 | 42.07 | 42.07 |
| 2 | 1.493 | 29.86 | 71.93 |

Component Matrix (2 components extracted by principal component analysis)

| | Component | |
|---|---|---|
| | 1 | 2 |
| PHRSERV | .717 | -.145 |
| RUNTIME | .289 | .815 |
| CORDEN3 | .560 | -.591 |
| NETDENS | .730 | .478 |
| AVSTDIST | -.929 | |

Rotated Component Matrix (Method: Varimax with Kaiser Normalisation, converged in 3 iterations)

| | Component | |
|---|---|---|
| | 1 | 2 |
| PHRSERV | .706 | .191 |
| RUNTIME | -.106 | .859 |
| CORDEN3 | .765 | -.278 |
| NETDENS | .439 | .754 |
| AVSTDIST | -.903 | -.271 |

We find from the factor analysis of the five variables grouped under general accessibility that two components could be retained, having eigenvalues over 1. The first weights heavily on light rail corridor population density, negatively on average stop distance, and positively on length of peak hour service. The second component weights after rotation most heavily on weekday running time and network density. We can call these two components `general accessibility' factors 1 and 2, denoted ga1 and ga2, and can now use the stored values in further regressions.

## 5. Use of Factor Scores in Regression Analysis on Light Rail Performance

We will now finally report results from the use of the factor analysis on both performance indicators and explanatory variables. We will limit our attention to three basic models, without going into further detail over excluding variables, or experimenting with functional forms. First, we repeat the regression with the overall performance index dependent PERFIND used in the research report. In addition, we use the two alternative `levels' and `growth' factors of performance obtained above. The growth measure of performance is expected to have more statistical `noise', reflecting the impact of shorter periods of change and urban structures that have had minimal time to adapt to new routes and light rail capacities.

**Table 8: Selected Multiple Regression Estimations using Performance Index and Performance Factors as Dependent Variables, with Explanatory Factor Variables.**

| Dependent Variable | Const. | LRPERFAC | MKGFAC | CRESTFAC | GAFAC1 | GAFAC2 | % AdjRsq |
|---|---|---|---|---|---|---|---|
| PERFIND | 11.6 | -0.008<br>(0.01) | 1.90<br>(2.9) | 1.17<br>(1.8) | 1.86<br>(2.8) | -0.04<br>(0.1) | 56.6 |
| PLEVELSFAC | 0.02 | -0.045<br>(0.2) | 0.52<br>(2.7) | 0.14<br>(0.8) | 0.39<br>(2.0) | -0.04<br>(0.2) | 40.9 |
| PGROWTHFAC | -0.07 | 0.21<br>(0.9) | 0.08<br>(0.4) | 0.225<br>(1.1) | 0.09<br>(0.4) | -0.366<br>(1.7) | 36.3 |

(t-statistics are given in parentheses).

The results summarised in Table 8 confirm that with the same performance index dependent variable based on overall ranking used above in Table 2, the most significant explanatory variables are the marketing factor, the car restriction factor, and the first general accessibility factor (which mainly represented light rail corridor population density, average stop distance, and length of peak hour service). The physical performance factor and the second general accessibility factor were not significant.

We then examine as alternatives the levels and growth measures of performance, taking advantage of the stored factor scores from the factor analysis reported above. This has the advantage of using the `raw data' on performance, rather than rankings of these data. It also explicitly separates out `levels' performance, which would normally represent long term transport policy commitment, from the normally shorter term growth performance of light rail or overall public transport.

With the `levels' factor (PLEVELSFAC) as dependent variable, we find that the marketing factor and the first general accessibility factor are still powerful and significant explanatory variables, with the car restriction factor weaker. We do not report detailed results after omitting the weakest explanatory variables; the adjusted R squared rose to 47%, and the t-statistic on the car restriction factor rose to 0.9.

With the performance `growth' factor (PGROWTHFAC) as dependent variable, we find a lower goodness of fit, with a strong coefficient but wrong sign for the second general accessibility factor. Recalling that this second accessibility factor loaded heavily on network density and weekday run time, this would seem to be picking up the low growth found in the mature systems with high network density and long run times. Other explanatory variables are weaker with the performance growth factor dependent, though the t-statistic for the car restriction factor is 1.1, giving some suggestion that car restriction measures may play a useful part is generating light rail and public transport growth, especially in highly congested modern conditions.

## 6. Concluding Remarks

We have reached some contrasting findings on the development of light rail and public transport in a wide range of urban transport systems, using first conventional OLS regression, then grouping our dependent and explanatory variables using factor analysis and conducting further regression analysis. We also gave further detail on the grouping of the cities themselves using cluster analysis.

One of the most robust findings is that physical or technical performance by itself seems to play very little role in either the accumulated level of public transport and light rail performance, or its recent growth. The strongest variables explaining the accumulated level of performance are those capturing the public transport marketing strategy (monthly fare level and travelcard use), and general passenger accessibility (in particular light rail corridor population density, stop density, peak hour service length).

Car restriction measures, such as length of pedestrianised street and modest amounts of city centre car parking, were also important explanatory variables of the overall ranking of cities in public transport performance. They too can be taken as providing key underlying elements of a successful public transport and light rail strategy, but require local government with the political will to risk alienating motorists by reducing routine private car road and parking capacity. Evidence from the Swiss and South German cases covered here suggests that it is a long haul to create a strong public transport culture that can attract professional and middle class car owners into everyday use in commuting and leisure.

## References.

Dickens, I. (1992) Transport investment, economic development and strategic planning: The example of light rail transit, **Planning Practice and Research**, 7(2), pp. 9-12.

Ferreira, L. (1997) Rail track infrastructure ownership: Investment and operational issues, **Transportation**, 24, pp. 183-200.

Hass-Klau, C. and G. Crampton (2002) **Future of Urban Transport. Learning from Success and Weakness: Light Rail** (Brighton: Environmental and Transport Planning), ISBN:0951962078, available through the website <www.etphassklau.co.uk>.

Li, J. and M. Wachs (2000) A test of inter-modal performance measures for transit investment decisions, **Transportation**, 27, pp. 243-267.

Rogerson, P.A. (2001) **Statistical Methods for Geography** (London: Sage).

Tyson, W. (1992) Appraisal of bus and light rail projects, **Project Appraisal**, 7(4), pp. 205-210.

---

[1] We also experimented with an index of the total availability of conventional suburban rail, to try to capture the extent to which this acts as a substitute for light rail. We calculated both the number of suburban rail stations per urban population, and also converted this into a 0,1,2,3, index representing zero, low, medium, and high availability of suburban rail. However, this variable had no significant quantitative impact, and apart from giving its correlation coefficient with the Performance Index is not discussed further.

[2] Cluster analysis was computed using SPSS version 9, using data on the 6 indicator variables. It should also be noted that the data for the Light Rail Passenger Km. was not available to us for Strasbourg, and this meant that the indicator LRPKPRTK (Light Rail Passenger Km. per Track Km.) would have been a missing value. We elected to impute a value for this variable by using the known Strasbourg value of LRPPERTK (Light Rail Passengers per Track Km) and then applying the ratio of LRPKPRTK to LRPPERTK taken from Rouen, a French city of similar urban and system size. We are reasonably confident that this imputed value is correct to 10%.

[3] Ward's method in cluster analysis is a commonly used hierarchical cluster analysis technique. "Ward's method chooses that merger (of clusters) that results in the smallest increase in the within sum of squares … (The within sum of squares may be thought of as the amount of scatter about the group centroids.) (Rogerson (2001), p. 201.)