An Explorative Econometric Model of Sydney Metropolitan Rail Patronage

By Neil Douglas[1] and George Karpouzis[2]

[1] Neil Douglas is Managing Director, Douglas Economics.
[2] George Karpouzis is Chief Economist, Rail Corporation (RailCorp) NSW.

Abstract

An explorative time-series econometric model of the demand for Sydney metropolitan rail is presented that uses annual data from 1969 to 2008. The model uses regression to explain the variation in the year on year change in rail demand. The dependent variable is specified in terms of trips per head using factored population. A set of explanatory variables including gross state product, office employment, fare, train kilometres and 'event' variables are introduced to explain the variation in rail trip rate. Of the variables modelled, office employment and gross product per head are shown to be important determinants of rail patronage.

The model is explorative in nature and a more detailed model is under-development. The ultimate aim is a short to medium term forecasting model of rail patronage to guide timetable, operational and business planning.

1.   Introduction

This paper presents an econometric model of rail demand in metropolitan Sydney using 38 years of patronage data. The model is explorative in nature and a more detailed model is under-development. The ultimate aim is a short to medium term forecasting model of rail patronage to guide timetable, operational and business planning.

Section 2 briefly reviews some of the types of model that have been used to explain and forecast rail demand. Section 3 presents the specification of the model that was fitted to Sydney data. Section 4 describes the dependent variable which was rail trip rate (trips divided by factored population). Sections 5-9 describe the explanatory variables that were introduced to explain the variation in rail trip rate. Some comments are made in section 10 regarding variables that would be expected to affect demand but were not included such as petrol prices and car ownership.

Section 11 presents the fitted model and reviews the parameter estimates. Of the variables considered, office employment is discussed in section 12 to be the most important influence on the level of rail demand. Section 13 makes some concluding remarks.

2.   Literature Review

Many econometric models of rail demand have been undertaken over the last forty years. Studies have included time series models such as the intercity models in the UK by Owen and Phillips (1986) and by McLeod et al (1985) of commuter rail services; cross-sectional or 'gravity type' models such as those fitted by Tyler and Hassard (1973) for inter-urban rail services; hybrid time-series - cross-sectional approach for example Douglas (1987). In addition, many studies have estimated price, service level and service quality elasticities using mode share models such as the Sydney Travel model estimated by Hague Consulting Group (2001) and maintained by the Transport Data Centre. As an alternative to using 'revealed preference' accounts data there have been models based on 'stated preference' market research. Booz Allen Hamilton (2000) summarized some of the results from this type of research.

This study of Sydney metropolitan rail patronage has used revenue accounts data. The time-series model fitted to the data is similar to an econometric model fitted by TfL (2002) to underground and bus travel in London. Where the models differ is data periodicity. The TfL model used four weekly data whereas this study has used annual data. It is hoped to replace the annual data with four weekly data in future to increase the number of observations and increase the precision of parameter estimation.

The model is also similar to a model fitted by Charles River Associates (CRA) in 2008 as part of an assessment of the net benefit of CityRail services. CRA used a shorter time series and attempted to explain the absolute level of rail patronage whereas this study has modelled the year on year change in rail demand.

3.   Model Equation

The model attempts to explain the variation in the year on year change in Sydney metropolitan rail patronage in terms of a set of socio-economic, fare, service level and 'event' variables. 38 years of accounting data dating from 1969-70 is used.

Equation 1 presents the model. The dependent variable is rail patronage divided by factored Sydney metropolitan population. In other words, the year on year change in rail trip rate was modelled rather than rail trips.

The model was specified in ratios to reduce auto correlation; the tendency for estimation errors in successive years to be related.

$$Q = C + \beta_i RF + \beta_i KMS + \beta_i MOE + \beta_i GSPP_t + \beta_i MI_t + \beta_i OLY_t + \beta_i AFC_t \qquad \text{....1}$$

where: $Q = \ln\left[\dfrac{Q_t/(Pop_t)^{0.52}}{Q_{t-1}/(Pop_{t-1})^{0.52}}\right]$ and Q_t is the number of rail trips (millions) and Pop_t is population of Metropolitan Sydney in millions factored by 0.52 to reflect the journey profile of CityRail users.

$RF = \left[RF_t - RF_{t-1} \right]$ where RF_t = average real fare per trip.

$TKMS = \ln \left[\dfrac{TKMS_t}{TKMS_{t-1}} \right]$ where $TKMS_t$ =Train kilometres (millions) run.

$MOE = \ln \left[\dfrac{MOE_t}{MOE_{t-1}} \right]$ where MOE_t Metropolitan office employment (millions).

$RGSPP = \ln \left[\dfrac{RGSP_t}{RGSP_{t-1}} \right]$ where $RGSP_t$ = Real Gross State Product of NSW per capita.

MI_t = Major Incident on CityRail taking a value of 1 if an incident occurred else zero.

OLY_t = Sydney Olympics taking a value of 1 if financial year = 2000-01 else zero.

AFC_t = Automatic Fare Collection taking a value of 1 if financial year is 1989-90 or later.

C = constant indicating the average year on year ratio change in patronage level.

β_i = parameters to be estimated with t denoting year.

4.   Patronage Trip Rate

Rail patronage was collated from a variety of sources. From the mid 1980s, patronage data was obtained from CityRail *"Compendiums"*. Earlier data was obtained from annual reports.

Figure 1 plots the trend in CityRail patronage over the modelled period (1969-70 to 2007-08). Patronage was defined as the number of passenger trips on CityRail services. All passenger trips were included including *'unremunerative'* trips i.e. passengers who travel free.
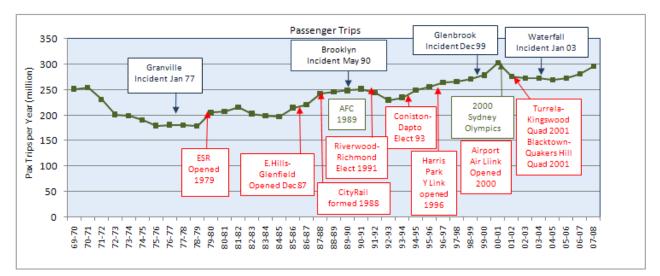
Figure 1: Trend in CityRail Patronage 1969-2008

It should be noted that there is no exact figure for passenger trips. Instead, RailCorp estimates passenger trips from ticket issues. Ticket type multipliers are applied to calculate the number of passenger trips. For example, a single ticket represents one passenger trip whilst a return represents two trips. There is some uncertainty regarding the number of trips made on periodical tickets. A weekly ticket has been estimated to represent 11 trips.
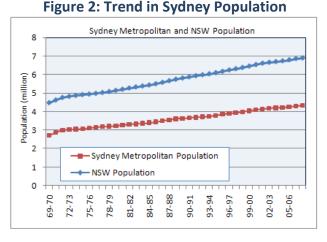
Patronage trended downwards during the 1970s largely due to increasing car ownership. From 1979 onwards however patronage increased in line with growth in population, Sydney office employment and train kilometres.

Significant events likely to have affected rail patronage are also shown. Two positive impacts were: Automatic Fare Collection and the 2000 Sydney Olympics. Four major train incidents that caused loss of life are also shown: Granville (January 1977); Brooklyn (May 1990); Glenbrook (December 1999) and Waterfall (January 2003). The remaining boxes highlight other factors that may have affected patronage but were not assessed in the model.

As equation 1 showed the dependent variable specified in the model was not rail trips but rail trips divided by population (factored). This was because there was little variation in the change in population by year. As a result, population tended to be correlated with the underlying time trend in rail patronage which was negative.

Figure 2 presents the growth in population for metropolitan Sydney and NSW which includes Sydney and Illawarra statistical divisions and the Newcastle Statistical subdivision.

Figure 2: Trend in Sydney Population

Sydney Metropolitan population grew from 2.7 in 1969 to 4.5 million in 2008 at a reasonably constant annual rate (1.2% compound). By comparison, NSW population increased from 4.3 million to 6.9 million (1.1%). Of the two possible population variables, Sydney Metropolitan population was used as the trip denominator since it more closely related to the catchment area of CityRail.

Rather than assume all trips were population driven, a factor of 0.52 was applied to population to reflect the journey purpose profile of CityRail patronage whereby 39% of CityRail trips are commuting to/from work and therefore relating to employment rather than population; 5% are work related; and, 4% of trips are made by non-Sydney residents.

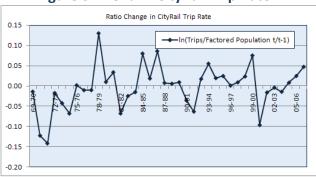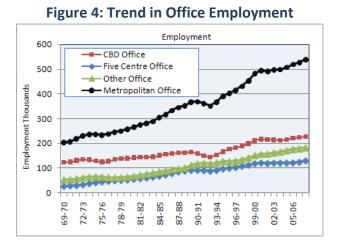**Figure 3: Trend in CityRail Trip Rate**

Figure 3 presents the transformed rail trip rate variable (after taking natural logarithms). The graph shows four periods of decline and four periods of increase with the periods of decline steeper and shorter than the periods of increase.

5. Office Employment

Office employment is an important determinant of CityRail patronage especially of peak travel which accounts for over one half of total rail demand. A 2004 survey of CityRail patronage estimated that one half of employed respondents (51%) had a managerial, professional or associated professional occupation with clerical occupations accounting for a further 33%, Douglas Economics (2004). Thus 84% of respondents had 'white collar' occupations. Blue collar occupations (trades + labour and related workers) accounted for 12% of employed trips.

Figure 4: Trend in Office Employment

The model used Sydney metropolitan wide office employment as an explanatory variable. Over the 39 years, office employment growth averaged 2.5% per year. Office employment was provided by the Property Council of Australia, NSW branch and is calculated by reference to office floor space (net absorption) rather than direct enumerations of office employment.

As Figure 4 shows, the trend in office employment has varied by centre. At 1.6% per year, Sydney CBD office employment grew at the slowest rate and it is Sydney CBD where CityRail achieves the highest Journey to Work (JTW) share of 51%, TDC (2006). By contrast, office employment at the five major non CBD centres (North Sydney, Chatswood, Crows Nest, Parramatta and North Ryde) grew at a much higher rate of 4.1% per year with growth at 'other' centres increasing at 3.2% per annum. JTW shares to these centres is lower however than to Sydney CBD for example 43% North Sydney, 31% Chatswood and 25% Parramatta.

It is intended that further work will explore a weighted employment variable to take account of variations in office employment across the metropolitan region. The data series will also be continued to include the impact of the 2008-09 global economic downturn.

6. Gross State Product per Head

Real Gross State Product per head (RGSPPH) is a proxy for changes in real personal income. RGSPP was calculated by dividing NSW Gross State Product by population and deflating by the NSW GDP deflator (2007-08 = 1). The NSW GDP deflator allowed for a general decline in the value of money over the thirty year period. Figure 5 shows the change in RGSPPH year on year. The economic recessions of 1981-82 and 1990-91 are shown as sharp dips.

RGSPPH was considered likely to have two counteracting effects on rail patronage. Firstly, shopping and leisure activity will tend to rise as RGSPP increases. Increases in shopping and leisure, other things being equal, should stimulate an increase in rail use. However, higher RGSPP might also stimulate greater car use thereby suppressing rail demand.

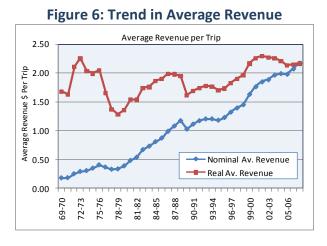**Figure 5: Ratio Change in Real Gross State Product per Head**

7.   Rail Fare

Nominal average fare was calculated by dividing annual ticket revenue by the estimated number of annual trips. The nominal average fare was then converted into a 'real' fare by deflating by the NSW GDP deflator.

Figure 6 shows the trend in nominal and real fare over the thirty year period. Nominal fares rose throughout the period whereas real fare oscillated between a low of $1.29 per trip in 1978-79 and a high of $2.30 in 2001-02. Goods and

Figure 6: Trend in Average Revenue

Service Tax (GST), introduced on July 1st 2000, was included in the revenue and average fare estimates since most passengers cannot claim back GST.

It is worth stressing that there are limitations in using average revenue to predict patronage response to fare changes. Clearly, the actual impact will depend on how individual ticket prices change. A different response is likely if short distance off-peak fares rise by more in percentage terms than long distance peak fares. Also the use of annual data effectively averages 'fares' over a year and therefore precluded estimation of immediate (monthly) to shorter term response (quarterly-six month effects). In this regard it is intended to replace the annual data series with four weekly accounts data.

8.   Train Kilometres Operated

Train kilometres operated provides a crude measure of service frequency and network size. Service frequency in turn measures the convenience of the timetable to passengers. Increasing the frequency of services reduces the 'displacement' between when passengers want to travel and when the timetable allows them to travel. For high frequency services, when passengers do not

Figure 7: Trend in Train Kilometres

need to consult the timetable, service frequency reflects the wait time.

Figure 7 shows that over the 39 year period, train kilometres doubled from just over 15 million kilometres to 35 million in 2007-08. Although the trend was upwards, there was variability in the year on year change.

It should be recognised that as well as influencing demand, train kilometres may be influenced by demand as a result of CityRail putting on more peak services to reduce crowding. The implication is that the estimated parameter may be inflated.

9.   Other Explanatory Variables

The model included dummy variables to model major train incidents, the introduction of automatic fare collection and the Sydney Olympics. These events were 'sign posted' in Figure 1.

Four major train incidents resulting in loss of life and a public inquiry occurred over the thirty year period: Granville January 1977; Brooklyn May 1990; Glenbrook December 1999 and Waterfall January 2003. The model included a variable taking a value of 1 if there was a major incident occurring in the financial year else zero.

The model also took into account the introduction of Automatic Fare Collection (AFC) in financial year 1989-90. AFC resulted in a magnetic stripe tickets being introduced with ticket barriers at major stations. As a result, fare evasion reduced and reported patronage increased. AFC was modelled using a dummy variable that took a value of 1 if the financial year was 1989-90 or later.

The Sydney Olympics was modelled using a dummy variable which took a value of 1 for financial year 2000-01 else zero.

10.   Omitted Variables

Several variables were omitted that could be expected to influence rail patronage. Currie (2006) has shown petrol prices to be positively related to public transport usage. In this study it was found that petrol prices and the CPI index had moved in tandem due to the use of annual data. Other omitted variables influencing the level of 'competition' include car ownership and bus fares with TfL included in their model of underground patronage.

With respect to rail service level, onboard travel time or speed, rolling stock quality, seasonality, tourism, retail activity advertising spend and access to the rail network have been shown to influence demand. The UK Passenger Demand Forecasting Handbook produced by the Passenger Demand Forecasting Council (2002) provides a comprehensive summary of factors influencing rail demand in the UK.

It is intended to include more variables (dependent on the availability of suitable data) as the model is developed further and the data is disaggregated into quarterly or four weekly accounting periods.

11.   Correlation between Variables

Table 2 presents a correlation matrix showing the strength of linear correlation between the variables. The strongest correlation between the dependent variable (trip rate) and the explanatory variables was Real Gross State Product per Head (RGSPPH) at 0.38. The weakest was Major Incidents (MI) at 0.04.

There was no particularly strong correlation between the explanatory variables. The strongest was a positive correlation of 0.39 between RGSPPH and metropolitan office employment (MOE). Second strongest was a negative correlation between major incidents (MI) and real fare difference (RF) at -0.34.

Table 2: Correlation Matrix

		Q	RF	TKMS	MOE	RGSPPH	MI	OLY	AFC
Passenger Trip Rate	Q	1.00	-0.22	0.24	0.25	0.38	0.04	0.24	0.13
Real Fare Difference	RF	-0.22	1.00	0.06	0.18	-0.16	-0.34	0.09	-0.01
Train Kilometres	TKMS	0.24	0.06	1.00	0.11	0.02	0.02	0.17	-0.27
Metropolitan Office Employment	MOE	0.25	0.18	0.11	1.00	0.39	0.11	0.00	-0.15
Real GSP per Head	RGSPPH	0.38	-0.16	0.02	0.39	1.00	0.15	-0.04	0.12
Major Incidents	MI	0.04	-0.34	0.02	0.11	0.15	1.00	-0.06	0.17
Olympics	OLY	0.24	0.09	0.17	0.00	-0.04	-0.06	1.00	0.16
Automatic Fare Collection	AFC	0.13	-0.01	-0.27	-0.15	0.12	0.17	0.16	1.00

Note: Q,TKMS,MOE,RGSPPH specified as ratios (t/t-1) with logarithms taken. Rf specified as difference (t-(t-1))

11.   Estimated Model

Table 3 presents the estimated model. The table provides the estimated parameters (β), standard errors and |t| values (independent of sign).[1]

All the parameters were of the expected sign. None of the parameters were significant at the 95% confidence level however (|t| ≥ 2.021) except the constant. The level of significance is considered to reflect the use of annual data. It is expected that statistical significance will improve with the use of four weekly or quarterly data.

The fare parameter was negative indicating that as real fare increases rail patronage reduces (population held constant). The functional form produces a demand elasticity proportional to real fare. In 2007-08, the average fare was $2.17 (including GST). Measured from this level, the elasticity is -0.245 (-0.113 x 1 x 2.17) implying that a 10% fare increase reduces patronage by 2.45%.

Conventionally, fare elasticities are reported for fare paying passengers. Unremunerative trips are excluded. About 15% of CityRail trips are unremunerative. Excluding unremunerative fares increases the fare elasticity to -0.283. Figure 8 presents the fare

---

[1] The |t| values provide a measure of the accuracy of the parameter estimates. A value of 2.021 with 40 observations indicates that the parameter is significant at the 95% confidence level (two tailed test) i.e. there is less than a 1 in 20 chance that the parameter is not significantly different from zero.

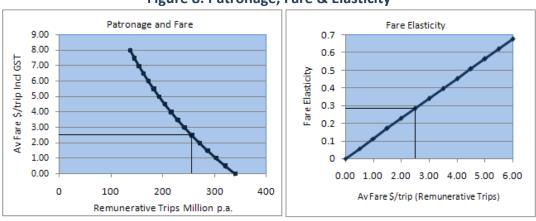demand curve for remunerative trips on the left hand side and the fare elasticity on the right hand side.

Table 3: Estimated Model

Var	Variable Description	β	Standard Error	|t|	≈ Prob β > 0 (1)
AFC	Automatic Fare Collection (1 if 1989-90 or greater, else 0)	0.021	0.018	1.170	74%
OLY	Olympics (1 if 2000-01, else 0)	0.064	0.052	1.226	76%
MI	Major Incident (1)	-0.028	0.029	0.971	64%
RGSPPH	Real Gross State Product per Head	0.740	0.528	1.401	83%
MOE	Metropolitan Office Employment	0.613	0.452	1.357	81%
TKMS	Train Kilometres Operated	0.246	0.156	1.574	87%
RF	Real Fare (real fare per trip)	-0.113	0.061	1.860	93%
C	Constant	-0.042	0.017	2.559	98%
R Squared		0.35			
Number of Observations		38			
Degrees of Freedom		30			

(1) Prob > 0 gives the percentage probability for a 2 tailed test that the parameter estimate is greater than zero (with 40 observations)

(2) MI took a value of 1 if year of Granville, Brooklyn, Glenbrook or Waterfall else zero

The fare elasticity of -0.28 is three-quarters the size of the elasticity of -0.38 reported in the CityRail Compendium 8th Edition which was based on stated preference market research surveys and market share information. The elasticity also compares with a value of -0.24 estimated by CRA (op cit) using 1977-78 to 2006-07 accounts data.

Figure 8: Patronage, Fare & Elasticity

The parameter for operated train kilometres was positive but relatively small in size. In contrast to fare, the model specification meant that the estimated parameter was also the demand elasticity (i.e. constant with respect to the level of train kilometres). At 0.25, a 10% increase in operated kilometres is forecast to increase rail patronage by 2.5%. Compared to TfL, the elasticity is higher than the estimate of 0.08 for Underground services.

Metropolitan office employment (MOE) had a strong positive relationship with rail patronage. A 10% increase in metropolitan office employment was estimated to result in an increase in rail patronage of 6.1%. The elasticity compares with an elasticity of 0.71 for London Underground patronage with respect to Central London employment.

The relationship between RGSPPH and rail trip rate was relatively strong with an estimated elasticity of 0.74. By comparison, TfL estimated lower elasticities of 0.54 for bus and 0.4 for Underground. The lower TfL elasticities may reflect the inclusion of retail sales and car ownership variables in the London model which were omitted in the Sydney model.

The 'event' variables estimated that Automatic Fare Collection (AFC) was associated with a 2.1% increase in patronage (reflecting less fare evasion) and the Sydney Olympics (OLY) in 2000 to have produced a one-off 6.4% increase in annual rail patronage.

The effect of a major train incident (MI) causing loss of life produced an average drop in annual patronage of 2.8% decline in annual patronage for the four incidents or $18 million in lost 'goodwill' and service disruption. Parameter significance was relatively weak however with a |t| value of 0.97.

12.   Impact of the 2008-09 Global Recession

Australia has remained relatively resilient to the 2008-09 down turn in economic activity. However the most recent July 2009 estimates of CBD white collar office employment produced by the Property Council of Australia, NSW branch have shown a decline of 4% on July 2008 levels.
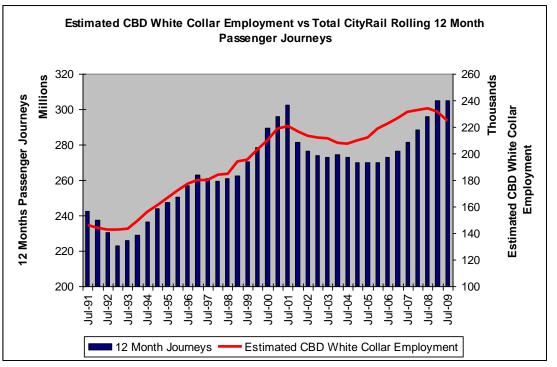
Figure 9: Office Employment and CityRail 12 Month Journeys

As can be seen from Figure 9, rail patronage and CBD office employment have tracked each other closely over the eight year period.

The elasticity of 0.63 (Table 3) would predict that the 4% reduction would reduce rail trips by 2.4% July 2009 compared to July 2008. However, this decline effect is not been picked up by the graph probably reflecting the use of a rolling 12 month average.

13.   Conclusions

Annual patronage data for Sydney metropolitan rail services over 38 years was modelled by reference to socio-economic, fare, train operating kilometres and other data. The estimated fare and train kilometres elasticities were of intuitive size and of reasonable accuracy.

Two variables describing the level of economic activity were introduced: Metropolitan office employment (MOE) and Real Gross State Product per Head (RGSPHH). Both variables were found to have a strong positive relationship with rail patronage. A 10% increase in MOE was estimated to increase rail patronage by 6.1% with a 10% in RGSP raising rail patronage by 7.4%.

The model is exploratory in nature. Several variables known to affect rail demand were not able to be included in the model. It is hoped to introduce some of the omitted variables in future work when the data series is extended and after the annual is replaced by four weekly or quarterly data.

References

Booz Allen Hamilton *"Value of Public Transport Attributes"* report to Transfund NZ June 2000.

CRA (2008), *"Value of CityRail Externalities and Optimal Government Subsidy"* report by CRA International for the Independent Pricing and Regulatory Tribunal (IPART), June 2008.

CityRail (2008), *"2008 Compendium of CityRail Statistics"* (Sixth Edition) produced by RailCorp.

Currie G. and Phung J. *"Aggregate and Disaggregate Analysis of Fuel Price and Interest Rate Impacts on Public Transport Demand"* ATRF 2007.

Douglas N.J. (1984) *"An Econometric Investigation into the demand function of Express Coach Travel"*, Paper presented at TRB Conference "Beyond Deregulation, at MIT Boston 1984.

Douglas Economics (2004), *"Value of Rail Travel Time"* report by Douglas Economics for RailCorp Train Services Rail Development, May 2004.

Hague Consulting Group (2001) *"Sydney Strategic Travel Model – Stage 1 Estimation"*, Hague Consulting Group & ITS Sydney, Report to NSW Department of Transport, August 2001.

McLeod G., Abbie E.G., Fotheringham S. and Everest J .T. (1985) *"Analysis and forecasting of London Commuter Rail Passenger Flows using Box-Jenkins Methods"*, T.R.R.L IN 0404/85 July 1985.

Owen and Phillips (1986), *"An Econometric Investigation into the Characteristics of Railway Passenger Demand"*, University of Leeds working paper February 1986.

Passenger Demand Forecasting Council (2002), *"Passenger Demand Forecasting Handbook"*. Report edited by C. Nash.

Transport Data Centre (2006) Analysis of the 2006 Census reported in the RailCorp Compendium 8th Edition.

TfL (2002), *"London Underground and Bus Demand Analysis 1970-2000"*, August 2002, Internal Note prepared by David Edwards (LUL). Geoff Hobbs (LUL), Ian Kincaid (SDG) and Alex Mitrani (SDG).

Tyler and Hassard (1973) *"Gravity / Elasticity Models for the Planning of the Inter-Urban Rail Passenger Business"* PTRC proceedings, Summer 1973.