C53

Paper written for presentation at the UTSG Annual Conference, held in Dublin, January 1998

# WHY ARE TRAVEL DEMAND FORECASTS SO OFTEN WRONG (AND DOES IT MATTER)?

Roger L Mackett
Centre for Transport Studies
University College London
Gower Street
London WC1E 6BT

**Abstract**

The theme of this paper is the examination of the quality of the outputs from the travel demand modelling process. Three topics are considered: the British national road traffic forecasts, local traffic forecasts, and forecasts used in the development of new urban public transport systems. Examples are taken from Great Britain and the United States. The forecasts are generally poor, and those that are good may be more due to coincidence that modelling skill. A number of reasons for the poor quality of the forecasts are put forward and some suggestions for improvement made.

## 1 INTRODUCTION

The objective of this paper is to examine the accuracy of travel demand forecasts, and if they are not accurate, to consider why they are not. Three topics are considered: the British national road traffic forecasts, local traffic forecasts, and forecasts in the development of new urban public transport systems. The examples are from Great Britain and the United States. The comparisons for the first case have all been made by the author, in the second case they are all based on secondary sources and in the third case there is a mixture. In the cases examined by the author no selectivity was applied, that is, no examples were obtained and ignored because they did not support the argument being put forward. As far as is known this is true of the work by other authors.

## 2 FORECASTING THE NATIONAL DEMAND FOR ROADS

The development of a national road system has been an integral part of the evolution of the transport system that exists in Britain today. Until the early 1960s future traffic levels on highways were estimated by adding a fixed percentage to existing flows, which often led to underestimates because they did not adequately take into account the factors underlying the growth in traffic. For example in 1945 two forecasts were made using 1933 vehicle registrations as a base. It was forecast that by 1965 vehicle numbers would have grown by 75% in urban areas and 45% in rural areas. In fact vehicle numbers had nearly doubled by 1950. In 1954 a new set of forecasts were

This paper is produced and circulated privately and does not constitute publication. It may be subject to revision before publication.

issued, anticipating a 75% increase by 1974. This figure was reached by 1962. In 1957 the 6.7 million vehicles was forecast to reach 8 million by 1960, whereas, in fact, there were 9 million vehicles (Starkie, 1982).

In 1962 national traffic forecasts were made by John Tanner of the Transport and Road Research Laboratory and these formed the basis of the official forecasts for a number of years (Tanner, 1965, 1974). For cars, Tanner's method involved the making of forecasts of car ownership per person, then population forecasts were used to give an equivalent number of cars. Assumptions were made about the number of kilometres per year to give estimates of total car kilometres. Similar methods were used for motorcycles and buses. The lorry figure was based on forecasts of ton-kilometres which were then split between road and other modes, with assumptions then made about lorry sizes and utilization rates to forecast the number of lorries and lorry kilometres. The number of light vans was assumed to be related to the number of lorries. The total number of vehicles, vehicle kilometres and total pcu (passenger car unit) kilometres was found by addition across the range of vehicles, allowing for miscellaneous vehicles. Tanner's method involved the use of a logistic curve (an S-shaped curve), which was uniquely defined by a suitable assumption about the saturation level which is the maximum level of car ownership which it is assumed will be reached. This was based on observations of data for parts of the United States. Gross Domestic Product (GDP) and motoring costs were incorporated into the procedures. The number of kilometres per car was later modified by the length of motorway (Tanner, 1974).

In the early 1970s the construction of new roads became the subject of public debate. Protestors started to disrupt the public inquiry system. They started to shift the focus of their protests from the line of a particular highway to broader questions about the need for new roads. Disruption of inquiries became common because the Department of the Environment (which included transport in its brief at this time), did not accept that it was the function of public inquiries to address such fundamental questions and so was not prepared to allow its staff to be questioned on these issues. The protestors continued to disrupt proceedings, culminating in the indefinite abandonment of the Airedale Trunk Road Inquiry in February 1976. The Department responded by announcing an independent review of its methods of scheme appraisal, including the traffic forecasts. To implement the proposal, the Advisory Committee on Trunk Road Assessment (ACTRA) under the chairmanship of Sir George Leitch, was set up.

The Leitch Committee made a number of criticisms of Tanner's approach. The report (Department of Transport, 1977) demonstrated the sensitivity of the forecasts to the saturation level chosen. Amongst the criticisms made was the lack of recognition of uncertainty in the forecasts, that the forecasts were self-fulfilling, and the lack of sensitivity to policy changes and fuel availability.

The Leitch Committee recommended that there should be a move away from Tanner's extrapolatory technique to a causal or disaggregate model, that is, one that represented the causal factors that underlie car ownership explicitly. It also recommended that the likely range of uncertainties involved in the forecasts should be indicated. Following these recommendations a new car ownership model was developed (Bates, Gunn and

Roberts, 1978). This model was based on the relationship between the probability of a household at a given level of income having a car. A similar relationship was used for the probability of a car-owning household having two or more cars. The model was calibrated using data from the Family Expenditure Survey for the period 1965 to 1975.

The 1980 National Road Traffic Forecasts (NRTF) (Department of Transport, 1980) were produced by combining the results from the two car ownership forecasting models discussed above, that is the model developed by Tanner (1977), as used previously, but with revised assumptions about economic growth and motoring costs, and the causal model. Initial tests with the causal model showed it to be inaccurate, because it predicted a decline in car ownership, which did not happen (this happened because car prices had risen much faster than incomes in real terms over the preceding period). Hence the model was revised with the introduction of a term reflecting the acquisition of a driving licence which gave a more acceptable answer. High and low forecasts of car ownership were produced by each model. The results from the two models were combined by means of a pragmatic mixture of averaging and taking the more extreme values. The car ownership figures were fed into a car use model which used population projections to estimate the total number of cars. This was then multiplied by the average number of kilometres per car which had been calculated by applying income and fuel price elasticities to the existing average car use. The elasticity value of distance travelled against fuel prices was increased in the light of the recommendations of the Leitch Committee. This produced upper and lower estimates of car use. Simpler methods were used for commercial vehicles, and buses and coaches. The results were added together to give the road traffic forecasts for each year up to 2010.

In 1984 new National Road Traffic Forecasts were issued (Department of Transport, 1984). These used a similar methodology, but recalibration against more up-to-date data led to considerably higher forecasts than those calculated in 1980.

In 1987 the document *Policy for Roads in England 1987* was issued. It contained some modifications and minor expansions of the road programme. This was a period of relative economic expansion, which led to increased traffic on the road and hence more traffic congestion. In 1989 the Department of Transport (1989) produced new road traffic forecasts. The methodology was modified to some extent. The causal model was retained, but Tanner's model was dropped on the grounds that the high and low car ownership forecasts were very similar for the high and low GDP forecasts. It was replaced by a model based on the elasticity of car ownership with respect to income.

The 1989 NRTF suggested that the previous forecasts were an underestimate and that traffic demand would increase by between 83 and 142 per cent in 2025 over the 1988 levels. Part of this concern had been stimulated by the completion of the M25 motorway around London in 1986. Several sections were overloaded soon after opening, raising questions about the traffic forecasting methods used. In order to provide sufficient road space to cope with all the forecast traffic, the Department issued the *Roads for Prosperity* White Paper in which it announced a doubling of the road programme. The detailed programme was spelt out in *Trunk Roads, England: into the 1990s* issued in 1990. This included widening the M25.

In 1994 growing opposition to the various road schemes and the need to reduce public spending led to a severe reduction in the roads programme, with many schemes being abandoned, as discussed in *Trunk Roads in England, 1994 Review.* In 1997, shortly before the general election, the Sir George Young announced the abandonment of many road proposals to reduce the risk of planning blight.

In November 1997 new National Road Traffic Forecasts were issued (Department of the Environment, Transport and the Regions, 1997). These were carried out in two stages: an 'unconstrained' or 'background' forecast which was comparable with the 1989 NRTF, and then a 'fitting on' process which involved applying the background traffic forecast to the existing road network to take into account of drivers' reactions to changes in journey time and the physical capacity of the network. The car traffic was modelled as the product of car ownership and use. The car ownership model incorporated the concept of saturation. Instead of using a single value as Tanner did, separate values were used for eight different household types. Year-specific constants related to the growth in the holding of driver licences were used to modify the time trend. The estimates of car use allowed for the predicted changes in fuel price, but with increasing efficiency of fuel use by cars being taken into account (by 2026 these effects are assumed to cancel one another out, so the real price of fuel in pence per kilometre is the same as in 1996, having increased in the intervening period, with a peak in 2001). The 'fitting on' process involved the application of speed/flow relationships to the distribution of traffic categorised by vehicle type, road and area types, locations and times. Travel time elasticities for various trip purposes were applied to estimate the effects of increasing amounts of traffic on a fixed road network.

The 1997 NRTF are lower than those for 1989. For example from 1996 to 2025 the 1997 NRTF forecast growth was 52% compared with 68% in 1989 NRTF (for the central forecasts). The 10% difference (in the total traffic presumably) is attributed to 7.5% from lower GDP growth and 2.5% from the fitting-on process, that is allowing for the effects of congestion.

It is clear that the National Road Traffic Forecasts have varied. It is legitimate to ask how good they are. Table 1 shows the forecasts to 1995 from Tanner (1965, 1974) and each of the NRTF, plus the actual figures. (It should be noted that the base year figures for NRTF have been revised retrospectively twice since 1978 because of perceived weaknesses in the survey methodology. The values in Table 1 are based on the latest methodology and are consistent across all the base years plus 1995. It means that they are not consistent with the values in the NRTF documents. However, the NRTF forecasts are presented as scaling factors. These have been applied to the revised base values.)

Comparing the actual flows in 1995 of 430.9 billion vehicle-kilometres it can be seen that the NRTF forecasts from 1980 and 1984 were significant under-estimates, while the 1989 forecasts were an over-estimate. The percentage errors become less with time, but, of course, the forecast period is shorter, so it should be easier to be more accurate. The annual percentage rates of growth forecast increased from 1.7 in 1980 to 2.9 in 1989. The actual figures increased slightly from 1980 to 1984, but then decreased significantly. In fact the actual flows in 1995 do not lie between any pair of high and low values. The switch from underforecast to overforecast may seem curious, but the period preceding 1989 was one of relative economic boom in which the total distance travelled increased annually by about 25 billion vehicle-km. It was followed by a period of during which it grew by about 1 billion vehicle-km each year. The actual annual average growth was 8 billion vehicle-km compared with the forecast average growth each year of 12 billion vehicle-km. However, it can be seen that the forecasts by Tanner (1965, 1974) were remarkably accurate. This seems ironic given the criticisms made in the ACTRA report and the recommendations that led to revised methods which turn out to be less accurate. However, it should be recognised that Tanner's method with its simple assumption about a single saturation level was very limited in scope, and that the accuracy of the forecast must be regarded as largely coincidental. It is worth noting that Tanner (1974) recognised that future traffic growth could be influenced by policy decisions which is why he made three alternative forecasts. It is also interesting to note that Tanner's inclusion of the future length of motorway in his 1974 forecasts as a determinant of the amount of traffic seems to be reflected in some of the thinking in the 1997 NRTF about the sensitivity of the forecasts to policy variables.

It is clear that the recent forecasts of road traffic have not been very accurate. The road building programme was expanded considerably to meet a demand that did not materialise. The fact that the economic recession and a desire by the government to cut public expenditure led to a subsequent reduction in the roads programme, thereby avoiding the over-provision of road-space, does not excuse the poor quality of the forecasts.

There is another argument against the National Road Traffic Forecasts. This is that they are self-fulfilling prophecies. The argument is that the forecasts are made and a programme of road construction is carried out. The availability of more fast roads encourages people to use them, so that the total amount of travel increases. There is anecdotal evidence of longer journeys, for example around the M25 motorway to the Lakeside shopping centre at Thurrock, presumably replacing shorter trips to other centres. It seems very unlikely that the greater availability of road space was the only factor. Increasing incomes and car ownership, plus associated shifts in lifestyle are causing many people to become more car dependent and so willing and able to travel further. This is complemented by other related trends of a decline in public transport usage and decentralisation of cities. It seems likely that one factor contributing to these trends is the increase in road-space, which is based on the NRTF, which as shown above are based on some rather inaccurate modelling work.

## 3 LOCAL TRAVEL DEMAND

As well as errors at the national scale, there can be poor forecasts more locally. In 1988 the National Audit Office (1988) examined the road-building process. It was concerned with trunk roads. According to the report, the Department of Transport is reasonably satisfied if a forecast is within ± 20% of the actual flow one year after opening. Of the 41 schemes analysed by the Department 22 met the criteria. This means that nearly half did not. The National Audit Office examined a total of 161 Department of Transport and Welsh Office Schemes in terms of the original forecasts for the design year (the 15th year after opening) with the predicted flows for that year calculated from the most recent traffic counts. On 39 of the schemes the predicted flows were more than 20% over the forecast, and in 40 cases predicted flows were more than 20% below the original forecast. The following possible reasons for the differences were put forward:

a) overestimation of traffic changing from the old road to the new;

b) no allowance for redistributed and generated traffic;

c) underestimation of national traffic growth;

d) wide regional and national variations in traffic growth;

e) unexpected developments.

The following year, the Public Accounts Committee (1989) was even more critical. It made a number of recommendations, stating that the traffic forecasting procedures must be improved.

Harris (1993) has analysed of the monitoring of the Department of Transport's traffic forecasts for 151 schemes. The errors in the forecasts vary from underprediction of 53% to overprediction of 112%. 24 schemes had forecasts within 5% of the actual flows, and 61 within 25%, so that 90 (about 60%) had errors of over 25%. Harris (1993) argues that the main reason for the errors is that they are based on NRTF for 1980 and 1984 which did not anticipate the high economic growth rates of the mid-1980s. In the report of the Standing Advisory Committee on Trunk Road Assessment (Department of Transport, 1994) on the traffic generation effects of new trunk roads, these differences between the forecast and actual flows were examined in the search for evidence of induced traffic resulting from the construction of new highways. The mean result for all schemes was that the flows were underestimated by 12%, which could be interpreted as evidence of traffic induction (that is, the extra 12% trips arose from the construction of the highways and were not included in the forecasts because the method used only considered changes of route and ignored other effects such as change of mode). In order to allow for the higher than anticipated economic growth all the forecasts were increased by an average of 16%. This had the effect of making the range of errors range from 47% underprediction to 129% overprediction. It reduced the difference between the averages of the forecast and actual traffic flows, but did nothing to reduce the dispersion in the results. In other words this scaling did nothing to improve the forecasts.

Harris (1993) examined one scheme in detail (the by-pass round the village of Prees in Shropshire). In this case the forecast was that the by-pass would carry an annual average of 4800 vehicles per day (vpd) in 1986, while a count of traffic on a weekday in April 1990 showed 6400 vehicles. The forecast was scaled to 1990 values by using the NRTF 1984 scenario which more closely match what actually happened over the period, giving a value of 5390 vpd. The actual factors were converted to an annual average flow of 6600 vpd by means of suitable factors. The 18.3% difference is explained as being made up of an underlying underprediction of 9.5%, a national forecast underprediction of 11.6% (because of the underestimate of economic growth) and 2.1% due to incorrect assignment (due to a smaller share of the corridor flow going along the by-pass than forecast).

It is not just interurban transport schemes that are subject to errors. Some time ago the quality of some of the forecasts for urban areas began to be questioned. Mackinder and Evans (1981) examined the results from 44 British transport studies by comparing the study forecasts with the actual changes that had taken place. It was found that nearly all the forecast items considered were overestimated. Mackinder and Evans showed that keeping constant various items that were forecast outside the transport model, such as population, employment and household income, would not have produced significantly greater forecast errors in any of the outputs and in many cases the average errors would have been considerably less. They also found evidence of specification errors particularly in the forecasting of public transport trips. These errors included variables not being used that should have been, and the assumption that the average number of trips made each day by each type of household is constant over time.

These types of errors in forecasting do not happen only in Great Britain. Mierzejewski (1995) has looked at some of the travel demand forecasts made in the United States. For example he examined the forecasts for 1985 from the Tampa Urban Area Transportation Study carried out in 1970. He examined the forecasts for the 87 links where a comparison could be made with actual flows for 1985 and found errors ranging from -87% to +281%. He also considered a 1980 study by the US Institute of Transportation Engineers (1980) in which eight studies were reviewed. He suggests that many of the results were poor, despite the fact that the report says that they were quite good. For example, in the South-Eastern Wisconsin, the report states that the population growth in Milwaukee was projected to grow by 10%, whereas it declined by 2%. This was claimed to be an error in the population forecast of -11.4%, but in terms of the error in the population growth which is what was being forecast, the error was 120%.

Errors in land use forecasts, that is, the distribution of population and employment, seem to be a major cause of errors in travel demand modelling. A good example of poor employment forecasts are the Dallas CBD (central business district) forecasts discussed by Kain (1990). Kain presents estimates of actual employment for the years 1956 to 1988 from a variety of sources, and forecasts to 1990 and 2000 from several sources under various assumptions. Kain regards some of the earlier estimates of actual employment with suspicion, but the more recent ones as reasonably reliable. Two things could be deduced: CBD employment was decreasing and by 1988 was about 90,000, and so by 1990 was somewhere below this figure. It is clear that all the forecasts assumed a large growth even when it was clear that the actual trend was downwards. By way of example, Kain regards the forecasts by NCTCOG (North Central Texas Council of Governments) as reasonably reliable. In 1982 when it estimated the 1980 employment to have been 127,000, it forecast the employment in 1990 would be 209,370. In 1987 it revised its estimate of the 1980 employment to 102,386, which implied CBD employment was decreasing. It also revised its 1990 forecast downwards to 120,206. Thus it recognised that the level of employment 1990 would be lower than originally forecast, but still forecast that there would be an increase.

Kain (1990) also examined the employment forecasts for Dallas CBD used to make the case for the Dallas Area Rapid Transit (DART), and argues that the very high values were used to produce high estimates of patronage on DART. In fact he claims that the figures used for DART were deliberately falsified in order to ensure that the system was built. The use of such forecasts in the development of urban public transport systems is the subject of the next section.

## 4 DEVELOPING NEW URBAN PUBLIC TRANSPORT SYSTEMS

There is clear evidence from the United States of proposals for rail-based schemes in which the forecast levels of patronage were well in excess of what actually occurred. A major study by Pickrell (1990, 1992), funded by the US Department of Transportation, found that the forecasts used to justify the building of many urban rail systems in the United States were significantly wrong. He examined the forecasts of patronage, capital costs and operating costs for ten systems which opened during the period 1976 to 1987. Table 2 shows the patronage on these systems (except the Atlanta Metro for which no patronage forecasts were available), in the first nine columns. It also shows the St Louis system using data from Warren (1995) plus three of the British light rail systems that have opened since 1980, namely those in Manchester, Sheffield and Newcastle. The figures for actual patronage for the US systems as quoted by Pickrell (1990) are also shown. These have been updated using figures from Dunphy (1995) and Bushell (1994).

Pickrell (1990) found that in most cases the patronage was less than half the predicted values and the capital costs over 50% higher than predicted. In two cases (the Washington and Atlanta Metro systems) the annual operating expenses were over three times the forecast figure. The total costs per rail passenger for the seven systems for which such forecasts were made exceed the forecast values by between 188% for the Washington Metro to about 800% or more for the Detroit Downtown People Mover and the Miami Metro.

It can be seen from Table 2 that several years after the actual patronage figures quoted by Pickrell (1990) none of the US systems that he examined have reached the forecast values. The Washington Metro comes closest, but even by 1993 was still 8% below the forecast for 1977, despite the expansion of the system. Of the other systems examined only the Portland and Sacramento light rail systems, and to a lesser extent the Miami Metro, have had significant increases in patronage. The other systems examined by Pickrell (1990) are still well below the patronage forecasts, which were often made for several years earlier. In fact the two downtown people-movers (Miami and Detroit) have both lost significant patronage since the 1988 figures cited by Pickrell (1990). This may be due to decentralisation of activities or a downturn in the economy. In fact the Buffalo and Portland light rail systems have both lost patronage since 1993. (According to Bushell (1994) weekday patronage was 29,700 in Buffalo and 27,200 in Portland.)

Pickrell examined in detail the forecasting methods used and concluded that the errors must have arisen from the structure of the forecasting models, the ways in which they were used, and the misinterpretation of the numerical outputs. A detailed study of the Sacramento light rail transit scheme (Johnston et al, 1988) shows that 'overstated assumptions and irregular manipulations of data' led to the choice of light rail. However, they argue that the choice of light rail reflected the local concerns of the politicians, and in that sense it was valid.

However, there are exceptions to the cases considered by Pickrell (1990). For example, on the St Louis light rail scheme which opened in July 1993, the forecast daily patronage of 17,000 at the end of the first year has been greatly exceeded with a weekday daily patronage of 44,414 with higher figures at the weekend (Warren, 1995). The forecast values are much lower than those predicted for the light rail systems examined by Pickrell (1990, 1992) on which the forecast daily weekday patronage ranged from 42,500 to 92,000. The actual values for these systems quoted by Pickrell (1990) ranged from 14,400 to 30,600. The more recent figures for these systems range from 23,000 to 31,100, which are well below the forecasts. This suggests that the forecasts have become much more conservative over time, and that the St Louis system is much more successful in practice than those in Buffalo, Pittsburgh, Portland, and Sacramento examined by Pickrell (1990, 1992). It is not known whether the financial case for the St Louis light rail system was made on the basis of a forecast of only 17,000 passengers per day. If so, that was a very low figure compared with other systems.

It is interesting to consider whether the same overforecasting issue arises outside the United States, for example by examining some quantitative information on the British systems that have opened in recent years, by comparing forecasts with the actual patronage.

In Manchester the Metrolink light rail system opened in 1992. Annual patronage of 10.09 million per year (35,700 per day) had been forecast (Greater Manchester Passenger Transport Executive, 1985). The actual patronage was 12.3 million in 1994/95 (43,500 per day), 12.6 million in 1995/96 (44,500 per day), and 13.4 million in 1996/97 (47,400 per day) (Department of Transport, 1997), suggesting that the forecasts were actually an underestimate. It should be borne in mind that the light rail system was replacing two existing railway lines which were carrying 6.05 million passengers annually (21,400 per day), so the forecasting procedure was not as difficult as might otherwise have been the case. Most of the new patronage was expected to come from bus. Knowles (1996) has shown that much more of the patronage came from car or were new trips than was forecast. Much less than forecast came from the old rail services than was anticipated, partly because of the eight-month period between the closure of the rail line and the opening of the light rail system. Knowles (1996) shows that the travel patterns are very different to those forecast, with many fewer trips in the morning peak than forecast and many more off-peak trips, especially at weekends. This suggests that the closeness of the forecast to the actual patronage may be largely coincidence with an overforecast largely balancing an underforecast. Knowles (1996) comments that it is rather surprising that the system has attracted more passengers than forecast because several factors suggest that the opposite should have happened: it opened during a recession, many of the proposed car parks at stations were not built, the service frequency was lower than anticipated, and the population of the three districts served by Metrolink declined by 1.1% between 1991 and 1993.

In Sheffield where the Sheffield Supertram system was opened in 1994 the picture is not so bright. According to Harris (1994) the patronage in the first full year of operation, 1994/95, was predicted to be 7 million passengers (24,700 per day), with 20 million in 1995/96 (70,700 per day). This was on the assumption that 29 km of track would be open. In fact, the actual patronage in 1994/95 was 2.2 million (7,800 per day), but only 7 km of track was open for much of this time, and 22 km by the end of the financial year (Department of Transport, 1996). In 1995/96 actual annual patronage had reached 5.3 million (18,700 per day), but the full 29 km of the current system was not operational for all this time. However, even by 1996/97 with the full system open patronage only reached 7.8 million (26,600 per day). Fox (1996) has suggested a number a reasons why patronage has been lower than that forecast. The main factor seems to be that the system was planned to serve a number of high density residential developments, but by the time the system was opened, the three tower blocks at Herdings Park, to which a branch line was built, had been emptied, it had been decided to demolish the massive Kelvin development, and the Norfolk Park Estate was being allowed to depopulate because of its unpopularity. There was a clear lack of communication between the South Yorkshire Passenger Transport Executive building the Supertram system, and the City Council responsible for local authority housing. Other factors that Fox (1996) notes may have caused patronage to be lower than forecast include the excessive running times because of interaction with other traffic, not all the proposed park-and-ride sites being opened, deregulation of buses leading to competition from buses rather than providing complementary services, very high fares for some journeys and much of the anticipated urban regeneration not happening. Recently some of these issues have been addressed, for example cheaper fares have been introduced. It is clear that in Sheffield patronage is much lower than forecast. One reason appears to be that the exogenous inputs to the modelling process such as the pattern of population were wrong. This appears to be more a political issue than a technical one, but this does not excuse the poor forecasts of patronage, and does not mean that if the exogenous inputs had been more accurate the forecasts would have been correct.

Fullerton and Openshaw (1985) examined the forecasts of operating costs and revenues for the light rail system in Newcastle upon Tyne, known as the Tyne and Wear Metro, which opened in 1980. They found that the revenue forecasts were very optimistic, but that the cost forecasts were fairly accurate. They argue that the low revenue figures were due to the low patronage which they show to be much lower than forecast at, typically, one-third of the forecast value on the various Metro links. They compared the exogenous inputs of future population, employment and car ownership, and argue that these were fairly good, with the inner area population too high but the employment too low and the car ownership figures much too high, which should have made the forecast lower than the real figures. According to Hall and Hass-Klau (1985), the Tyne and Wear Metro system was forecast to carry 30 million passengers in 1983 (106,000 per day), increasing to 62 million (219,000 per day) when it opened fully. It actually carried 49.8 million passengers in 1983 (176,000 per day), which increased to 59.1 million in 1985 (208,900 per day), the first year of full operation (Department of Transport, 1996). This implies that the forecasting procedure was fairly accurate. However, patronage has declined steadily since then, reaching 35.9 million in 1995/96 (122,600 per day) and 35.4 million (120,900 per day) in 1996/97. This illustrates that even a system that is successful can lose passengers over time as car ownership increases and decentralisation occurs. There appears to be a difference between the forecasts being cited by these two sets of authors, but they are not presented in a way that is directly comparable. Often forecasts are revised which means that there is a variety to choose from.

The Docklands Light Railway was constructed to stimulate development in London Docklands and so no cost-benefit analysis was carried out. Much of the justification was based on the generation of jobs. It was predicted that 9300 extra jobs would result from the construction of the DLR (Docklands Public Transport and Access Steering Group, 1982). There has been considerable growth in the number of jobs in the area but it is impossible to identify which ones would not be there if the DLR had not been built. Patronage reached 14.5 million in 1995/96 (51,300 per day) following expansion of the system capacity (Department of Transport, 1996).

Thus, the evidence of the accuracy of the forecasts for recent British light rail systems is limited. There was significant overforecasting on only one out of the three operational systems for which forecasts were made. It may be that public expenditure in Britain is so limited that the forecasts are subject to stricter scrutiny than was the case in the United States when the systems examined by Pickrell (1990, 1992) were developed.

## 5 WHY ARE THE FORECASTS WRONG?

The evidence presented above suggests that forecasts of travel demand are often wrong. The National Road Traffic Forecasts have all been clearly wrong, despite a number of significant refinements over the years. The methods by Tanner which they replaced appear to give very accurate answers but it is difficult to accept models with such simplifications are correct. The local transport studies in Britain reviewed by Mackinder and Evans, the National Audit Office and the Department of Transport seem to be generally very poor, with similar observations by Mierzejewski for the United States. The forecasts used in the development of new urban public transport systems tend to be poor, particularly in the United States. Overall the picture is very disappointing. In none of the three cases is there evidence of good forecasts. It is worth stressing that in carrying out this work no examples of good forecasts were found and ignored. It is simply the case that there were very few examples of good forecasts and some of those that were good were probably right for the wrong reason., for example, Tanner's forecasts and those for the Manchester Metrolink where the overall patronage figure was about right but the distribution between peak and off-peak was quite wrong.

Identifying that there is a problem is not difficult. Deciding why there is a problem is more difficult. Perhaps the real question is 'Why has nothing been done about it?' because this problem has been known about for years: for example, Mackinder and Evans published their work in 1981 and the National Audit Office published its report in 1988. Several reasons can be put forward: firstly, it is difficult to admit that previous work is wrong, whether you are the person who carried it out or the person who commissioned the work and accepted it. Secondly, much effort has been put into developing the techniques, and those who can make money out of selling their services using them are unlikely to be interested in unremunerative activities like model development. Thirdly, forecasts are required in order to development transport infrastructure, and it is easier to continue with the existing techniques than to say that they are so flawed that they must be disposed of. Fourthly, there is no new set of techniques available to replace the existing ones, and so there are likely to be arguments along the lines 'It's better to use the existing techniques with all their weaknesses than nothing at all'. Fifthly, there is a desire for continuity so that, for example road schemes being appraised now can be compared with schemes appraised previously.

One difficulty with forecasting is that it is not a value-free process. Whilst most planners are trying to give an honest statement of what they believe will happen, this can be confused with a different issue of what they want to happen. It is a natural human characteristic to want to be associated with success rather than failure. This may manifest itself in terms of unwillingness to admit that a city is in decline or it may be shown in terms of a desire to obtain funding for a system that will improve a city because the planner believes that is what is required to solve the problems. This issue may become exacerbated by politicians who want to obtain central government money or maintain their own popularity and so will only accept optimistic forecasts. A different dimension to the process is the claim that forecasts can be self-fulfilling: an accusation that has been made about the National Road Traffic Forecasts: a forecast of the amount of traffic is made, the roads are built to accommodate it, and the traffic level grows in response to the greater journey opportunities offered.

The reasons why the forecasts are wrong can be divided into two groups: internal to the modelling process and external to it. There may be overlap between the two cases. Reasons internal to the modelling process include:

a) **Inputs are wrong**, for example population, employment and GDP. These other factors are usually forecast outside the transport modelling process by other people. This does not excuse poor transport forecasts. However, it must be recognised that forecasts of, for example, the national economy are extremely difficult to make. There is little doubt that travel demand is strongly influenced by the state of the national economy. There may be significant changes in the factors in reality over time, for example the spatial distribution of population in Sheffield before Supertram opened. This implies that the forecasting of exogenous inputs to the transport modelling process may require rather pragmatic procedures rather than pure forecasting techniques.

**b) Variables are missing.**

**c) Model mis-specification.** There may be major relationships missing from the model, for example traffic induced by road building in the procedures used for appraising road schemes in Britain until very recently. Another problem may be attempts to represent a dynamic process with a static model.

**d) Models being used for unsuitable purposes.** It is quite possible for a model to be taken beyond its capabilities simply because no other technique is available. One example is a model which fits well at a single point in time is assumed to be capable of forecasting without any testing of its forecasting ability.

**e) Insufficient disaggregation.** A model may treat everybody in the same way whereas variation in behaviour may be quite significant.

Reasons external to the modelling process include:

**a) Desire to achieve a specific result,** such as obtaining funding from elsewhere for a local transport scheme which might lead to ignoring less desirable outputs from the model.

**b) Political pressure** which makes certain outcomes unacceptable.

**c) Technical incompetence** because the operators of the models do not really understand what they are doing.

A number of reasons why models may be wrong have been put forward. A key question is, does the fact that travel demand models are so often wrong really matter. The answer to this question is 'yes' for a number of reasons:

a) Travel demand models have a strong relationship with policy (Mackett, 1998), and so affect the lives of many people both directly through affecting their mobility, and indirectly through public expenditure.

b) There are major expectations of the transport system and if these are not met because of poor forecasting, credibility may be lost (Mackett and Edwards, 1998).

c) It is very unprofessional. Travel demand models have existed for about 40 years. There is little evidence that recent models are significantly better at forecasting than those produced many years ago. This is not a good reflection on those working in the field.

There follows another question: 'What can be done about it?' The answer cannot be easy because if it is it would have been taken a long time ago. Similar problems have been recognised in the US and the 'Travel Model Improvement Program' has been set up, partly in response to the mandatory requirement that travel demand forecasts are correct. A similar programme would be useful here, particularly in the light of the new approach to transport policy being adopted with its emphasis on integration. It is likely that better policy will be formulated if the tools to assess the impacts are reliable. It would be particularly useful to carry out an assessment of which models are reliable and which are not so that effort for improvement can be put in the right place. There is probably need for a change of culture that makes assessing the validity of a model much more central to the development process than has been the case up to know. This also requires the making public the poor results as well as the good as part of the collective learning process. There may need to be a shift of emphasis within the transport modelling field from the detailed, such as traffic flow and route choice to broader issues such as the total level of travel demand because there is little point in having a perfect assignment model if the traffic being assigned contains huge errors. There is also a need for a dialogue with the modellers who provide the exogenous inputs, such as population and GDP, in order to improve those forecasting techniques.

## 6 CONCLUSIONS

This paper has demonstrated that many forecasts in the field of travel demand are wrong. This is clearly unsatisfactory. A number of reasons for the errors have been put forward, plus some suggestions for remedial action. The credibility of the transport profession is at stake here. It would be very risky to ignore the warning signs and just carry on using the same techniques despite their demonstrable weaknesses. To solve the problems will require resources, but first it needs recognition that this is an important issue and that action is required. Only then can the process of developing a better range of techniques start.

## 7 REFERENCES

BATES J, GUNN H AND ROBERTS M (1978) *A Disaggregate Model of Household Car Ownership*, Department of Transport, London.

BUSHELL C (ed) (1994) *Jane's Urban Transport Systems,* Jane's Information Group, Coulsdon.

DEPARTMENT OF THE ENVIRONMENT, TRANSPORT AND THE REGIONS (1997) *National Road Traffic Forecasts 1997.*

DEPARTMENT OF TRANSPORT (1977). *Report of the Advisory Committee on Trunk Road Assessment (The Leitch Report)* HMSO, London.

DEPARTMENT OF TRANSPORT (1980) *National Road Traffic Forecasts 1980.*

DEPARTMENT OF TRANSPORT (1984) *National Road Traffic Forecasts 1984.*

DEPARTMENT OF TRANSPORT (1989) *National Road Traffic Forecasts 1989* HMSO, London.

DEPARTMENT OF TRANSPORT (1994) *Trunk Roads and the Generation of Traffic: Report of the Standing Advisory Committee on Trunk Road Assessment (SACTRA)* (HMSO, London).

DEPARTMENT OF TRANSPORT (1996) *Transport Statistics Great Britain 1996*, HMSO, London.

DEPARTMENT OF TRANSPORT (1997) *Transport Statistics Great Britain 1997*, HMSO, London.

DOCKLANDS PUBLIC TRANSPORT AND ACCESS STEERING GROUP (1982) *Public Transport Provision for Docklands: Summary of the Assessment of Schemes.*

DUNPHY R T (1995) Review of recent American light rail experiences. *Proceedings of the Seventh National Conference on Light Rail Transit,* held in Baltimore, Maryland, USA, November 1995, **1**, pp 104-113.

FOX P (1996) Tram operation in Sheffield: progress, problems and a crisis, *Light Rail Review*, **7**, pp 63-71.

FULLERTON B and OPENSHAW S (1985) An evaluation of the Tyneside Metro, in K J Button and D E Pitfield (eds) *International Rail Economics*, Gower, Aldershot.

GREATER MANCHESTER PASSENGER TRANSPORT EXECUTIVE (1985) *Proposed Light Rapid Transit System: Application to the Department of Transport for Section 56 Grant Aid*, Manchester.

HALL P AND HASS-KLAU C (1985) *Can Rail Save the City? The Impacts of Rail Rapid Transit and Pedestrianisation on British and German Cities*, Gower, Aldershot.

HARRIS M (1994) Sheffield Supertram. *Modern Railways.* **51**, pp 279-283.

HARRIS R C E (1993) Monitoring Department of Transport traffic forecasts, Proceedings of Seminar D, 21st Summer Annual Meeting, PTRC, held at UMIST.

INSTITUTION OF TRANSPORTATION ENGINEERS (1980) Evaluation of the accuracy of past urban transportation forecasts, *ITE Journal*, **50**.

JOHNSTON R A, SPERLING, D, DELUCHI M A and TRACY, S (1988) Politics and technical uncertainty in transportation investment analysis, *Transportation Research.* **21A,** pp 459-475.

KAIN J F (1990) Deception in Dallas: strategic misrepresentation in rail transit promotion and evaluation, *Journal of the American Planning Association,* **56,** 184-196.

KNOWLES R D (1996) Transport impacts of Greater Manchester's Metrolink light rail system, *Journal of Transport Geography*, **4**, pp 1-14.

MACKETT R L (1998) The role of travel demand models in policy making, *Project Appraisal* (forthcoming).

MACKETT R L and EDWARDS M (1998) The impact of new urban public transport systems: will the expectations be met?, *Transportation Research A* (forthcoming).

MACKINDER I H and EVANS S E (1981) The predictive accuracy of British transport studies in urban areas, *TRRL Report SR 699,* Transport and Road Research Laboratory, Crowthorne.

MIERZEJEWSKI E A (1995) *A New Strategic Urban transportation Planning Process*, Center for Urban Transportation Research, University of South Florida, Tampa.

NATIONAL AUDIT OFFICE (1988) *Department of Transport, Scottish Development Department and Welsh Office: Road Planning*, Report by the Comptroller and Auditor General, HMSO, London.

PICKRELL D H (1990) *Urban Rail Transit Projects: Forecast Versus Actual Ridership and Cost*, DOT-T-91-04, Urban Mass Transportation Administration, U.S. Department of Transportation, Washington D.C.

PICKRELL D H (1992) A desire named streetcar: fantasy and fact in rail transit planning, *Journal of the American Planning Association.* **58**, pp 158-176.

PUBLIC ACCOUNTS COMMITTEE (1989) *Road Planning,* 15th Report of the Public Accounts Committee, HC 101, HMSO, London.

STARKIE D (1982) *The Motorway Age: Road and Traffic Policies in Post-war Britain*, Pergamon Press, Oxford.

TANNER J C (1965) Forecasts of vehicle ownership in Great Britain, *Roads and Road Construction*, **43**, pp 341-347, 371-376.

TANNER J C (1977) Car ownership trends and forecasts, *TRRL Report LR 799*, Transport and Road Research Laboratory, Crowthorne.

WARREN W D (1995) Why success in St Louis? *T R News,* **180,** pp 22-26.

*Table 1. Forecasts of 1995 motor vehicle traffic flows from NRTF*

| | 1965 Tanner | 1974 Tanner | 1980 NRTF | 1984 NRTF | 1989 NRTF | 1995 actual |
|---|---|---|---|---|---|---|
| Base year | 1965 | 1974 | 1978 | 1982 | 1988 | |
| Length of period to 1995 (years) | 31 | 21 | 17 | 13 | 7 | |
| Actual traffic flow (billion vehicle-km) | 152.3 | 229.8 | 256.5 | 283.9 | 375.7 | 430.9 |
| Factor to 1995 - low | | | 1.20 | 1.19 | 1.17 | |
| Factor to 1995 - high | | | 1.47 | 1.39 | 1.28 | |
| Forecast traffic flow in 1995 - low | | 389 | 307.8 | 337.8 | 439.6 | |
| Forecast traffic flow in 1995 - high | | 455 | 377.0 | 394.6 | 480.9 | |
| Forecast traffic flow in 1995 - mean or medium | 429.6 | 429 | 342.4 | 366.2 | 460.2 | |
| % difference between mean forecast and actual flows | -0.2 | -0.4 | -25.8 | -17.7 | +6.4 | |
| Forecast % growth to 1995 | 182.1 | 86.7 | 33.5 | 29.0 | 22.5 | |
| Actual % growth to 1995 | 182.9 | 87.5 | 68.0 | 51.2 | 14.7 | |
| Forecast % growth each year to 1995 | 3.4 | 3.0 | 1.7 | 2.0 | 2.9 | |
| Actual % growth each year to 1995 | 3.4 | 3.0 | 3.1 | 3.2 | 1.9 | |

*Table 2: Comparison of forecast and actual patronage on new urban public transport systems*

| | Washington | Baltimore | Miami | Buffalo | Pittsburgh | Portland | Sacramento | Miami | Detroit | St Louis | Manchester | Sheffield | Newcastle |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Metro | Metro | Metro | LR | LR | LR | LR | DPM | DPM | LR | LR | LR | LR |
| **Forecast** | | | | | | | | | | | | | |
| Year | 1977 | 1980 | 1985 | 1995 | 1985 | 1990 | 2000 | 1985 | 1985 | 1994 | 1996 | 1996 | 1985 |
| Patronage | 569.6 | 103.0 | 239.9 | 92.0 | 90.5 | 42.5 | 50.0 | 41.0 | 67.7 | 17.0 | 35.7 | 70.7 | 219.1 |
| **Actual** | | | | | | | | | | | | | |
| Year | 1986 | 1987 | 1988 | 1989 | 1989 | 1989 | 1989 | 1988 | 1988 | 1994 | 1995 | 1995 | 1985 |
| Patronage | 411.6 | 42.6 | 35.4 | 29.2 | 30.6 | 19.7 | 14.4 | 10.8 | 11.3 | 44.4 | 43.5 | 7.8 | 208.9 |
| % difference | -28% | -59% | -85% | -68% | -66% | -54% | -71% | -74% | -83% | +161% | +22% | -89% | -5% |
| Year | 1993 | 1991 | 1992 | 1995 | 1992 | 1995 | 1995 | 1992 | 1992 | | 1996 | 1996 | 1996 |
| Patronage | 526.6 | 45.2 | 48.4 | 29.0 | 31.1 | 24.0 | 23.0 | 9.5 | 8.8 | | 44.5 | 18.7 | 122.6 |
| % difference | -8% | -56% | -80% | -68% | -66% | -43% | -54% | -77% | -87% | | +25% | -74% | -44% |

Note: Patronage is in thousands of passengers on a weekday. The percentage differences are between the actual values and the forecast values. The dates shown are the year in which the financial year ended. LR stands for light rail. DPM stands for downtown people-mover. The classification for the US systems apart from St Louis is based on that by Pickrell (1990). The forecasts and the earlier actual figures for all the US systems except St Louis are taken from Pickrell (1990). The more recent actual figures for Buffalo, Portland and Sacramento are from Dunphy (1995), while those for all the other US systems except St Louis are from Bushell (1994). The figures for St Louis are from Warren (1995). The forecast for Manchester is from Greater Manchester Passenger Transport Executive (1985), the forecast for Sheffield from Harris (1994), and the forecast for Newcastle upon Tyne from Hall and Hass-Klau (1985). The actual figures for Manchester, Sheffield and Newcastle upon Tyne are from Department of Transport (1996). For Washington the actual annual patronage for 1993 was divided by 281.6, and for Buffalo the annual figure for 1993 was divided by 276 to convert to daily figures, using figures from Pickrell (1990). The daily figures for the later date for all US systems except Buffalo, Portland, Sacramento and St Louis, plus both the actual figures and the forecasts for Manchester, Sheffield and Newcastle upon Tyne have been obtained by dividing the annual figures by 282.9, which is the average of the values for Washington, Atlanta and Buffalo given in Pickrell (1990).