DEVELOPMENT OF A COMPREHENSIVE URBAN PUBLIC TRANSPORT MODEL FOR LONG-RANGE PATRONAGE FORECASTING

Paolo Delle Site
Roger L. Mackett
Centre for Transport Studies, UCL, London

1. INTRODUCTION

In this paper the introduction on a given urban route of a new public transport (PT) mode, which upgrades, on at least part of the route, an existing bus or rail service, is considered. PT in urban areas can be improved with measures including not only adjustment of short-range operating factors, such as frequency and fare, but also enhancement of some of the long-range factors, relating to the right-of-way (ROW) and technology. Interest in the latter is prompted by the potential of so-called intermediate modes, i.e. modes lying between conventional heavy rail and bus services (e.g. bus priority measures, guided busways, street running trams, light rail—LR). Evidence of this relatively new trend is provided by the number of bus priority and LR schemes that have been introduced in many urban areas. New ideas for PT in outer London are covered in a report by London Transport (London Transport Planning, 1995). It is likely that decision makers will have to cope increasingly in the future with the problem of choosing between alternative PT mode options for either heavily-used bus corridors or lightly-used rail services as well as new routes, as in the cases of the study for London.

Patronage forecasting is a key issue in strategic planning as it plays a role in both financial and economic assessment and, hence, finally, in the choice between alternative PT modes. The current debate on the accuracy of patronage forecasting (Mackett, 1996), although suggesting that cases of massive overestimates—like those shown by some LR schemes in the US—can be explained by the need to ensure funding, has led the authors of this paper to reconsider the problem from basic principles.

The paper reports on research work carried out with a grant by the European Commission within the Human Capital Mobility Programme, aimed at:

- first, developing a comprehensive theoretical framework for PT patronage forecasting,
- secondly, bringing it to the modelling stage.

Section 2 of the paper presents the theoretical framework, highlighting the relevant elements. The variables and the interactions are identified to illustrate the potential for the analysis of the policy and the responses to it. The PT modes are described in terms of their basic characteristics. Also the context of the policy, including the relevant exogenous elements that are independent of the policy but with an effect on patronage, is identified. The time scale of interactions is taken into account in order to cover the common length of time horizon adopted for this kind of forecasts. Section 3 addresses some modelling issues. The main components of the modelling framework are first
identified. To achieve higher policy sensitivity, the problem is formulated as a restricted area model. The general hallmarks of the approach are described. The modelling of the effects of changes in transport supply brought about by the policy, all other elements affecting patronage remaining equal, is then investigated. The potential of the contribution which can be represented by current modelling practice, together with major gaps in current practice and research are discussed. The features of a simplified version that has been implemented conclude the Section. Section 4 concludes the paper, summing-up contributions provided, suggesting interesting directions of research and highlighting work needed for exploitation of the proposed approach.

2. THEORETICAL FRAMEWORK

The theoretical framework in terms of the identification of the relevant components and their interactions is outlined in the next subsections. Figure 1 provides a synthesis of the main elements, and is useful as diagrammatic guideline for this Section.

2.1 Policy definition and scope

The policy addressed consists of changes in the supply of transport services and infrastructures represented by the provision of a new PT mode on a given route. This will be referred to as the PT linehaul mode.

The set of relevant modes which represents the supply side of the problem includes first:
1. the PT linehaul mode
and extends to:
2. access/egress modes, PT and non-PT, for which some of the characteristics are changed; the policy can include the redesign of part of the bus network to feed the PT linehaul mode—new feeder bus network (FBN)—and the introduction of park-and-ride (P&R) facilities;
3. PT and other modes in competition with the PT linehaul mode for which some of the characteristics are changed; the policy can include the introduction of reserved lanes which subtract road space and of priority at junction which reduces junction capacity;
4. PT and other modes of access/egress and in competition with the PT linehaul mode for which the characteristics are assumed unchanged.

The characteristics of the first three sets which are part of the policy are regarded as endogenous elements of the supply side, those of the last set as exogenous.

A distinction arises within the endogenous elements of the supply side. Some of the characteristics of the PT modes (linehaul and access/egress), typically those relating to operations (e.g. frequency and fare of the new FBN and of the PT linehaul mode), can be adjusted in the short-range and are determined by the operator according to a given objective and maybe subject to constraints. The supply characteristics of the modes included in the first three sets above are grouped into two classes: the first including the PT optimisable characteristics, the second all the others. The former class can be extended to include some long-range characteristics (e.g. stop spacing and vehicle capacity). The forecast of patronage is obtained within an optimal design framework for the former class, within a ‘what if’ approach (more conventional in PT strategic
planning) for the supply characteristics of the latter class. Implicitly, for the former class the ‘what if’ approach is shifted from the characteristic to the criterion (objective function) that the operator adopts to set the characteristic.

The settings of the supply characteristics determine the level of service (LOS) attributes of the relevant modes. These represent the whole set of variables the users respond to in making travel decisions. Conventionally-used proxies for LOS are the (generalised) cost and the (dis)utility. The LOS is also affected by demand through congestion effects. Whenever interactions occur between modes (e.g. between bus and cars on mixed-traffic ROW) the LOS of one mode depends on demand for other modes also (i.e. there are non-separable cost or utility functions).

A taxonomy of the relevant set of supply characteristics that are endogenous to the policy (divided into PT and other road modes) is shown in Table 1. The result of a similar effort for the set of LOS attributes that can be affected by changes in the endogenous supply characteristics is shown in Table 2. The table shows also which attributes are affected by demand through congestion. A list of the possible objectives adopted by the operator to optimise some of the supply characteristics, differentiated according to the actors (restricted to users and the operator) for which allowance is made, and that of the financial constraints that the optimisation is subject to, are shown in Table 3. The financial constraint applies only when objectives including both users’ and operator’s perspectives are adopted. The underlying hypothesis is that benefits and costs to non-users are not relevant in the decision regarding the setting of the optimisable supply factors, although they are included in the subsequent assessment stages.

2.2 Responses to the policy

The change in LOS brought about by the policy yields responses which fall into two different domains: travel decisions and land-use.

The former domain includes the dimensions of demand that are affected by the change in the LOS. Different social groups are to be considered because of the different perceptions of LOS and different mode alternatives available. For given settings of socio-economic and land-use factors, the demand for PT of all social groups at equilibrium, i.e. the demand that is consistent with the produced LOS, is the patronage level sought. Demand for each mode is generally affected by the LOS of other modes also. Hence, due to the non-separability on both the supply and demand sides, PT patronage is the result of a multi-modal equilibrium which yields demand and LOS at equilibrium of all relevant modes for given settings of socio-economic and land-use factors. The values taken at equilibrium by these basic quantities, together with others relevant to benefits and costs for the users and the operator, determine the values of the quantities in the objectives and in the constraints which the operator uses to set the PT optimisable supply characteristics.

The latter domain of responses to the policy includes land-use changes that are produced by the change in LOS and, hence, in accessibility. These responses occur after a time lag and, in turn, affect travel decisions. The land-use factors affected by the policy represent endogenous determinants of demand, while all other socio-economic
and land-use factors that affect demand but are set independently of the policy, are regarded as exogenous.

The classes of responses differentiated by domain, type (direct/indirect), strength and time scale are shown in Table 4. It is implicitly assumed that trip frequency and car ownership are inelastic to the policy.

2.3 Scenarios for the forecast procedure

The forecast of patronage over a predefined time horizon is obtained by subdividing the time horizon into time periods each giving rise to a separate forecast. Each time period is identified by a scenario, where a given setting of the exogenous elements (referred to in Section 1 as the context outside the policy) and of the land-use factors affected by the policy is maintained for the whole time length (e.g. one year).

A further disaggregation with respect to time arises within each time period. This is subdivided into subperiods to allow for demand and supply variations and subsequent changes in PT service arrangements (e.g. over a week—weekday and weekend service, and over a day—peak and off-peak service).

The patronage and the other relevant quantities (demand for other modes, LOS on all modes, users’ benefits, fare revenues and operating costs) at equilibrium provided for each time period or scenario represent the output sought (the former quantity as the main product, the latter as by-products of the framework) which is then fed into the assessment stages.

3. MODELLING ISSUES

3.1 Modelling framework

The modelling framework must be able to represent both the responses to the policy and the effects of relevant exogenous elements which are independent of the policy. The former include the travel decisions and the land use factors that are affected by the endogenous supply characteristics identified in the previous Section. The latter include the travel decisions that are affected by changes in socio-economic and land-use factors (e.g. car ownership) and by changes in supply characteristics other than those included in the policy (e.g. increase in car operating cost due to car pricing schemes).

According to the different time scale involved for responses (travel decisions in the short-range, land use effects in the long-range), the modelling framework includes:

- a supply-demand equilibrium model (SDEM) to represent responses that occur in the short range,
- a location model (LM) to represent responses occurring in the long-range.

The SDEM represents travel decisions including destination, mode, route and time of departure and provides the equilibrium between demand and supply for a given time period or scenario. The LM represents the location effects of the demand-supply equilibrium that is established with the policy and provides the destination choices brought about by these effects. The two models are linked over time with the LM...
feeding the SDEM of the subsequent time period. The effects of changes in exogenous elements are taken into account in the SDEM, which is extended, if needed, to represent travel decisions relating to trip frequency (e.g. if changes in income occur).

3.2 Modelling approaches

There are two approaches determined by the population which is being represented: conventional full-scale planning models, and sub-area or corridor marginal models. Examples of the latter, less used, approach are in Talvitie (1978), De Cea et al. (1986), and in Dehghani and Harvey (1994). Marginal models have the advantage of a higher sensitivity to the relevant decision elements of the problem because of their potential in the representation of some issues, relating to the supply-demand equilibrium (represented in the framework here in the SDEM), for which consolidated modelling solutions at the full-scale level do not yet exist. The relevant issues include the following.

(i) Setting of the service offered by the operator. Both short-range operating parameters (frequency and fare) and long-range supply characteristics are assumed given, independent of demand. The consequences are that the line capacity is also assumed and, due to the absence of capacity constraints which are not usually included in PT assignment models, check for consistency with resulting flows is necessary. Practically, capacity is adjusted to flows iteratively but there is no proof that convergence is achieved. Models have been formulated where PT assignment is combined with optimisation of frequency by Schöele (1980) and LeBlanc (1988), with optimisation of vehicle capacity for a bus network by Shih and Mahmassani (1994).

(ii) Car-based feeder modes (P&R). Combined car-PT modes are not usually represented at the assignment level. Multi-modal network equilibrium modelling including combined modes was first addressed only recently by Fernandez et al. (1994).

(iii) Congestion effects within the PT system itself. PT cost functions independent of passenger flows are used in assignment models. A few congested PT assignment models have been formulated, where in-vehicle time and, most recently, waiting time increasing with flows are included (there is a review in Wu et al., 1994).

3.3 Basic features of a restricted-area model

The drawbacks of the conventional full-scale approach highlighted above can be overcome, to some extent, by a restricted-area approach at the cost of a simplified representation of other elements.

The relevant population is represented by that belonging to the PT linehaul catchment area. This can be identified, by using the hypothesis that a maximum of two transfers can occur for each trip to reach the destination, with the area where access/egress to the PT linehaul mode is made possible by at least one of the following: walk, existing PT interconnected, new FBN and P&R.
The relevant PT network includes both access/egress and linehaul modes. In the hypotheses that the existing PT linehaul mode is not retained and that overlapping of other routes is negligible, the relevant PT network is restricted to the linehaul mode, the existing PT routes interconnected with the PT linehaul mode and the new FBN. An example is shown in Figure 2 where the traffic zones of an existing system that contain the PT linehaul catchment area are also represented.

The relevant non-PT modes are those for which demand can be subtracted from or added to and include car as access/egress mode (P&R) and walk and car as competing linehaul modes (taxi and motorcycles, which usually represent a minor fraction of demand, are omitted). A simplified treatment of car mode can be adopted where the effects brought about by changes in endogenous and exogenous supply characteristics are represented sequentially in a two-step procedure. The effects on route choice of reduction of arc and junction capacity and of other changes that are exogenous to the policy are modeled within a background conventional full-scale model by assignment of the car OD matrix to the road network. This provides the new flow pattern together with average road speed and trip length for cars. The effects of changes in other endogenous supply characteristics (relating to both PT and other road modes, as shown in Table 1) are supposed to have negligible effects on route choice and average trip length for the car mode and are modeled within the SDEM.

The whole set of possible mode choices, for both access/egress and linehaul, is represented in Figure 3. For car, PT interconnected and FBN, access and egress are always assumed to be by walking. The choice of a mode alternative including PT for at least one of the trip components (access/egress or linehaul) is coincident with that of a route on the relevant PT network defined above. Thus, route choice responses, for both PT and car modes, are not modeled in the approach here and the set of travel decisions within the SDEM is restricted to time of departure, destination and mode choice.

For the difficulties of modeling time-of-departure choices within an equilibrium framework where steady-state conditions are considered, the set of choices represented in the SDEM is further restricted to destination and mode. The SDEM is formulated within the restricted-area approach outlined above as a mathematical programming problem where:

- the optimisable variables are represented by the short-range PT supply characteristics (frequency and fare) and the vehicle capacity of both the FBN and the PT linehaul mode; the former are differentiated for each sub-period to account for demand and supply variations;
- the objective function is represented by any of the objectives (Table 3) the operator can use to set the variables above;
- the constraints include:
  - demand and LOS simultaneous equations, representing, in each subperiod, the multi-modal equilibrium and providing destination and mode choices consistent with the LOS produced,
  - operational constraints: financial (Table 3), if consistent with the objective, and capacity.
The drawbacks (i), (ii) and (iii) of the full-scale approach mentioned in the previous subsection can find a solution in the restricted-area approach here where (i) the problem is formulated within an optimal design framework for the FBN and the PT linehaul mode, (ii) combined car-PT modes are represented at the mode choice level and (iii) congestion effects are included in the cost or utility functions of the relevant PT modes.

The general mathematical formulation of the SDEM is:

\[
\begin{align*}
\max \ W(D, L, S, \bar{S}, \bar{A}) \\
\text{subject to:} \\
D = d(L, \bar{A}) \\
L = l(D, S, \bar{S}) \\
y(D, L, S, \bar{S}, \bar{A}) \leq 0
\end{align*}
\] (1)

where (quantities given are overlined):

- \( W \) = objective function
- \( d, l, y \) = vector-valued functions, respectively, demand, LOS and constraints
- \( D \) = demand vector of components \( D_{jm}^t \) where the superscripts refer: \( t \) to time subperiod and \( g \) to social group; the subscripts: \( j \) to origin-destination (OD) pair and \( m \) to mode
- \( L \) = LOS vector of components \( L_{jm} \) with superscript and subscripts as above
- \( S, \bar{S} \) = vectors of, respectively, optimisable PT supply characteristics and all other supply characteristics
- \( \bar{A} \) = vector of socio-economic and land-use factors.

### 3.4 Modelling the effects of PT supply changes

The specification of the functions \( d, l, y \) and \( W \) in (1) is obtained as follows. The case where the SDEM applies to a context with unchanged exogenous factors, compared to the existing system or to the conditions represented in the scenario of the previous time period, is considered.

**Demand functions** \( d \) modelling simultaneously destination and mode choices are obtained within the random utility theory. Usually, logit forms are adopted, either multinomial or nested.

The preliminary identification of catchment area for each stop of both access/egress and linehaul PT modes and for P&R facilities is needed. The choice of the access/egress mode for areas where at least two modes are available (e.g. areas having both direct access to the FBN by walk and to the PT linehaul mode by P&R) can either be modelled jointly with the linehaul mode choice or separately according to a separate model or arbitrary criteria.
Incremental forms are used to represent changes in supply characteristics brought about by the policy. The base condition is either the existing system or the scenario of the previous time period. New modes (e.g. P&R) can be tackled within the incremental formulation as it can be shown easily that the share of the new modes can be expressed as function of changes in LOS referred to any existing mode (a proof for the multinomial logit is given in Delle Site, 1996).

LOS functions \( I \) are expressed incrementally. Thus, the LOS attributes represented need to be coincident with those for which values taken in the base condition are known. A background conventional full-scale model provides the values needed in the existing system and those resulting from the changes, if included in the policy, in supply characteristics that are treated separately as indicated in section 3.3.

The LOS attributes included explicitly in cost or utility functions fall in the performance and out-of-pocket cost categories (Table 2) and refer to walk, car and PT, as both access/egress and linehaul modes. The attributes relating to performance are usually restricted to average travel time components, thus excluding PT reliability (except for the effects represented in the average waiting time) and safety.

Travel time by walk is obtained for a given zoning system as a function of the density distribution of the ODs within the stop catchment area.

Travel time by car is obtained from area-wide speed flow relationships (a review is provided in Olszewski et al., 1995). These provide the average road speed as function of total vehicle-kms, given by those affected by car and those by PT in mixed-traffic ROW. They apply to trips using car as both linehaul mode and access/egress mode to and from P&R facilities. The distance travelled is given by the preliminary assignment in the former case (linehaul), is estimated on the basis of the minimum length route in the latter (P&R), where an additional time component is added to represent time spent in the parking lot.

Travel time by PT results from the different arcs represented in the network: transfer (between different PT modes and from/to P&R facility), wait and in-vehicle. Constant values independent of demand are assumed for travel time on transfer arcs to facilitate application of the model. Waiting time functions for random passenger arrivals, which depend on average frequency and standard deviation of headway, are used in the conventional manner. The effects of vehicle capacity are not taken into account and a solution for this shortcoming has not yet been developed. Waiting time functions increasing with flows to account for vehicle capacity are included in recent transit assignment models but neither empirical nor theoretical evidence of the form of the proposed functions exist. The issue of schedule following can be considered a minor concern in a frequency-setting framework also for long average headways given that, even if schedule followers may not incur an actual physical wait time at the stop, they incur a schedule delay relative to the actual time at which they would have wanted to depart. In-vehicle time is subdivided into a component exclusive of stops and dependent on car flows on mixed-traffic ROW (obtained equalising car and PT speed), and a component representing delay at stop and dependent on boardings and alightings (different forms of stop service time functions are reviewed in Vandebona and Richardson, 1985).
Out-of-pocket costs relating to fare (PT unimodal and integrated with parking fee) are modelled using various structures including flat, distance-related and zone. The operating cost for car is assumed proportional to the distance travelled, with the fraction of the unit cost that relates to fuel consumption dependent on average speed.

The attributes excluded relating to service quality of PT are represented by the mode constant which can be interpreted as the contribution of the unmeasured random effects and, in particular, of the unmeasured qualitative attributes affecting demand.

Financial constraints \( y \) and objective functions \( W \) include operating costs which are assessed by the conventional multi-variable route allocation models. Estimates that are sensitive to the vehicle size and to subperiods (peak and off-peak) for the relevant unit costs should be included (as shown respectively by Oldfield and Bly, 1988 and by Cherwony and Subhash, 1978).

Users benefits are assessed by rigorous measures applied to multi-modal context depending on the demand function used (there is a review in Jara-Díaz, 1988) rather than by the commonly used approximate rule-of-a-half which, moreover, breaks down when new modes are included.

The set of constraints can be extended to include capacity constraints for PT and P&R (the extension to all modes would raise the issue of consistency between OD matrices and network capacity). The capacity constraint for P&R can be dropped if functions which increase when flows approach capacity are considered to represent time spent in the parking lot.

3.5 Implementation

A simplified version of the SDEM has been implemented (Delle Site, 1996) to model mode choice responses to changes in supply characteristics relating to the PT linehaul mode only. A solution approach based on penalty function and global optimisation methods is used to provide the optimal frequency, fare per unit length and vehicle capacity. A hypothetical case study referable to a bus route in London is used to demonstrate the effects on patronage and the policy implications derived from different objectives and different sensitivity of demand to qualitative attributes in the two cases of improved bus service and LR.

4. CONCLUSION

The work reported on here highlights the theoretical underpinnings of the problem of strategic patronage forecasting for PT projects and provides a critical review of current state of modelling and proposes a new approach.

The main areas of research in PT modelling that will provide significant contributions to both the conventional full-scale approach and the restricted-area approach proposed here, relate to time of departure responses, PT reliability and qualitative attributes, and waiting time functions depending on vehicle capacity.
Further work needed to make the proposed approach operational relates to both the modelling and implementation stages. The modelling issues to be addressed are the representation of changes in exogenous factors (both socio-economic and land-use and supply sides) within the SDEM, and the development of the LM.

REFERENCES


Table 1. Endogenous supply characteristics

- infrastructure-related
  - right-of-way (A fully separated, B longitudinally separated, C mixed traffic)
  - access facility
  - short-distance transport
  - parking
  - platform design
  - platform capacity
  - stop/station spacing

- vehicle-related
  - technology
  - design
  - capacity
    - seated
    - total

- operations-related
  - scheduling
    - headway
    - headway control
    - stopping
    - longitudinal spacing control
  - other modes interference management
  - real time user information
  - fare
    - collection
    - structure
    - level

OTHER ROAD MODES

- infrastructure-related
  - arc capacity
  - P&R facility capacity

- traffic management system-related
  - junction capacity
  - integrated multimodal fare

Table 2. Level of service attributes affected by endogenous supply characteristics, affecting demand shown in brackets for relevant attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>PT</th>
<th>Performance</th>
<th>Service Quality</th>
<th>Convenience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access/exit time</td>
<td>(PT, ORM)</td>
<td>* travel time (ORM)</td>
<td>* vehicle occupancy (PT)</td>
<td>* platform occupancy (PT)</td>
</tr>
<tr>
<td>Waiting time</td>
<td>(PT)</td>
<td>* in-vehicle time variation</td>
<td>* quality of waiting</td>
<td>* quality of waiting</td>
</tr>
<tr>
<td>In-vehicle time</td>
<td>(PT, ORM)</td>
<td>* vehicle occupancy (PT)</td>
<td>* awareness/understanding</td>
<td>* aesthetics</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>* waiting time variation</td>
<td>* personnel appearance and behaviour</td>
<td>* passenger behaviour and image</td>
</tr>
<tr>
<td>Safety</td>
<td></td>
<td>* in-vehicle time variation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Out-of-pocket cost</th>
<th>Fare</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTHER ROAD MODES</td>
<td></td>
</tr>
</tbody>
</table>

Key: ORM = other road modes

Table 3. Objectives and financial constraints

<table>
<thead>
<tr>
<th>OBJECTIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>USERS &amp; OPERATOR PERSPECTIVE</td>
</tr>
<tr>
<td>* max net benefit (users benefits + fare revenues - operating costs)</td>
</tr>
<tr>
<td>* max patronage</td>
</tr>
<tr>
<td>* max usage (passenger-kms)</td>
</tr>
<tr>
<td>OPERATOR PERSPECTIVE ONLY</td>
</tr>
<tr>
<td>* max profit or min loss</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FINANCIAL CONSTRAINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>* subsidy as given amount</td>
</tr>
<tr>
<td>* subsidy as given fraction of operating costs</td>
</tr>
</tbody>
</table>

Table 4. Responses to the policy

<table>
<thead>
<tr>
<th>DIRECT EFFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>* travel decisions in the short range (first order)</td>
</tr>
<tr>
<td>* mode</td>
</tr>
<tr>
<td>* destination (activities with no transitional cost only)</td>
</tr>
<tr>
<td>* time of departure</td>
</tr>
<tr>
<td>* land use in the long range</td>
</tr>
<tr>
<td>* location of dwellings and workplaces within the existing stock (second order)</td>
</tr>
<tr>
<td>* new development of dwellings and workplaces (third order)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INDIRECT EFFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>* travel decisions due to land-use changes</td>
</tr>
<tr>
<td>* destination (all activities)</td>
</tr>
</tbody>
</table>
Figure 1. Theoretical framework: main components and interactions

Legend
Access/egress modes:
1 walk
2 existing PT interconnected
3 new FBN
4 P&R

- P&R facility
- PT linehaul
- existing PT interconnected
- new FBN
- traffic zone boundary

Figure 2. PT network and traffic zones containing catchment area

Figure 3. Mode choice alternatives