Building the Case: Improving the Credibility and Reliability of Travel Demand Models to Meet Changing Needs

David Kriger, P.Eng., MCIP
David Kriger Consultants Inc., Ottawa

Paper prepared for presentation at the
“Best Practices in Transportation Planning” Session
of the 2012 Annual Conference of the
Transportation Association of Canada

Fredericton, New Brunswick

April 2012

ABSTRACT

For over half a century, travel demand models have been used to project travellers' behaviour under different circumstances such as demographic and economic growth, network improvements and policy changes. The resulting forecasts have been used as the basis for road, transit and other infrastructure investment decisions, so much relies on getting the forecasts 'right.' Of course, forecasting is inherently a highly uncertain process: On the one hand, much effort has been spent over the past several years to replace the traditional four-step model with much more rigorous and representative algorithms, such as tour- or activity-based micro-simulation models.

On the other hand, even with new algorithmic improvements coming into practice, the application of travel demand models still has not fully addressed the need to improve the credibility and reliability of the outputs. Moreover, the types of issues for which decisions must be made, and the audience that constitutes the decision-makers themselves, are changing. Examples of the former concern changing traveller behaviour to use sustainable transportation modes. As for the latter, financial decision-makers now have a key role, along with engineers and planners, in determining whether or not a project proceeds to implementation: this is true both for projects that are financed entirely with public funds or externally through a public-private partnership.

The implications on model development and application are many. They include:

- The need to account for pricing (such as tolls) and how people make trade-offs (values of time, values of reliability, etc.)
- Changing needs for model validation, to ensure that the model is appropriate to the problem at hand and provides the necessary information at an acceptable level of reliability and credibility to satisfy information requirements.
- The types of data that are needed to support these applications.

In the end, models must be used to build a *credible* and *reliable* 'case' - whether it is a business case for internal or external funding approvals, or a case to demonstrate the GHG-reduction impacts of a proposed transit investment for public scrutiny.

The paper discusses the changing roles for models and their forecasts in the context of new planning requirements (e.g., sustainability and financing); who needs information and what kind of information they need from the models; the importance of improving model reliability and credibility; and the techniques and data for making these improvements. The paper draws from experience in Canada and the United States, where local and state authorities have had to start positioning themselves to accommodate alternative financing (such as P3) and address new initiatives (such as TDM). The paper will be of interest to Canadian government agencies at all levels that are charged with making investment decisions to meet growing transportation demand and emerging mobility needs.

1. INTRODUCTION

1.1 Objectives

This paper considers methods for improving the credibility and reliability of travel demand forecasting models, in order to better position Canadian transportation planning agencies to address emerging demands (such as sustainable transportation), allocate scarce resources, meet public scrutiny and help move projects towards implementation (whether by public or private financing). The discussion is distinct from, and complementary to, the many profound analytical and methodological improvements that are being made to the state-of-the-practice of modelling – notably, the movement from trip-based models to tour- and activity-based models. The distinction considers *how* models are being used; how they are validated; and *what* is being modelled.

1.2 Organization of paper

The remainder of the paper has four main sections. Section 2 provides a context for the discussion. Section 3 considers the issues of model *credibility* and *reliability*. Section 4 identifies seven key ways in which credibility and reliability can be improved. Section 5 concludes the paper.

2. CONTEXT

Over the last 10-15 years, much attention has been given to improving the algorithmic basis of travel demand forecasting models, and in turn, to the data and surveys upon which the models are based. This represents a significant transition from the traditional four-step model structure: trip generation, trip distribution, mode share and trip assignment. The four-step approach was developed in an era in which transportation plans focused on new road and, later, transit infrastructure requirements to meet forecasted shortfalls in capacity. The transition was driven by several factors, notably:

- Theoretical and methodological shortcomings and internal inconsistencies in the four-step approach. For example, the steps are treated as sequential decisions on the part of the traveller, rather than more correctly as combined or simultaneous decisions (e.g., whether or not to make [generate] the trip may depend on the availability of a particular mode). Another example is the theoretical concern that the Gravity distribution formulation does not realistically depict how people actually choose their destinations. These concerns are well documented in the literature, and are not discussed further.

- Travel demand forecasting become the subject of important lawsuits in the United States in the 1980s, in which model forecasts that showed the air quality impacts of proposed new highways to be minimal were successfully challenged. Subsequently, the 1990 *Clean Air Act* and the 1991 *Intermodal Surface Transportation Efficiency Act* (and its successors) have contributed to significant changes in modelling and planning; and – although the legislation applies to the United States – these impacts have propagated around the world.[1]

- Public pressure to promote sustainable transportation as an alternative to constantly increasing road capacity. That is, sustainable transportation – in the form of transit infrastructure and services, as before, but also in walking, cycling, Transportation Demand Management techniques, trip avoidance, telecommuting, etc. – was established as an increasingly viable solution set for transportation plans. Hence the need arose to account for these with greater precision and sophistication in forecasts.

- Acknowledgement of the ‘maturity’ in the travel demand ‘markets’ in many Canadian urban areas (and in those of other developed countries). For example,

[1] For a detailed discussion, see M. Garrett and M. Wachs, *Transportation Planning on Trial: The Clean Air Act and Travel Forecasting*, Sage Publications, Thousand Oaks, California, 1996.

growth in the transit share to the central business district - a focus of many transportation planning objectives – has achieved the desired targets but has since stagnated (more precisely, it increases only with growth in population). In the meantime, population and jobs have both dispersed to the suburbs, such that the traditional downtown orientation of the transit system has been replaced by the need to accommodate multiple urban and suburban foci.

- Changing finances mean that the availability of funds for new transportation improvements is increasing constrained. Funds must compete with other public needs, such as health and education; and the effects of this scarcity are amplified by changing economic and demographic conditions in Canada and elsewhere.

From a methodological perspective, the tendency towards tour- and activity-based models represents a significant and, likely, permanent change in direction. However, notwithstanding the many methodological improvements, the four-step model, or its derivations, remains the dominant approach in Canada, the United States and around the world. The 2005 Ottawa-Gatineau model is a prominent Canadian example of a tour-based model (and this model is now being updated); however, virtually all other applications in Canadian practice are trip-based. The availability of budgets to fund new models – in the United States, the development of activity-based models at many Metropolitan Planning Organizations (regional planning agencies) is measured in years – and the intensive data and computational requirements means that the transition will happen but only over time.

3. CREDIBILITY AND RELIABILITY

The aforementioned issues and approaches are well documented in the literature. However, another factor is permeating into practice with much less attention and at a much slower pace: this concerns the need to improve model *credibility* and *reliability*. The focus here is on model outputs – that is, model forecasts – and their accuracy in projecting travel over time. The concern is much more prominent in forecasts for public-private partnerships (P3), in which significant amounts of money, and the associated risk, are at stake. The concerns have been raised mainly by the financial community: commercial investors (such as banks and pension funds), bond rating agencies (which determine the costs of borrowing money for financing new construction) and monoline insurers (which provide 'guarantees' to issuers of credit [financing] against default). As is discussed in the next section, the financial community has effectively imposed new requirements on modelling in order to improve credibility and reliability.

Public agencies in Canada have used P3s as a means of implementing social and other infrastructure when traditional public funds are not available, or in order to advance the introduction of the new facility. In transportation, P3s have been used to implement new highways and crossings or rehabilitate older structures in Nova Scotia, Prince Edward Island, New Brunswick, Québec, Ontario, Manitoba, Alberta and British Columbia. The private partner develops the facility, and has the responsibility for operating, maintaining and rehabilitating the facility over a fixed period of time, in exchange for a monetary payment. In some cases – such as the Highway 407 ETR in the Greater Toronto-Hamilton Area – the concessionaire charges a toll directly to users. In other cases, such as Calgary's Stoney Trail and Edmonton's Anthony Henday ring roads, the owner (in this case, the Province of Alberta) makes an 'availability' payment to the concessionaire (that is, the concessionaire makes the facility 'available' to serve a traffic volume that is guaranteed by the owner, in return for a fixed rate of compensation). Some transit infrastructure has been implemented as a P3: notably, Vancouver's Canada Line, which opened in 2009, and the upcoming Ottawa LRT project.

Although P3s can be controversial, they are an established 'fact' in Canada and elsewhere. Moreover, other tolled initiatives are being considered in Canada's urban areas, at least conceptually: these include cordon pricing and preferential parking pricing. As a result, they cannot be ignored in transportation planning or, it follows, in travel demand forecasting. Finally, in recent years HOT (high-occupancy tolled) and managed lanes are being implemented instead of new greenfield tolled roads in several urban areas in the United States, in part to make better use of existing underutilized HOV lanes in locations where road widening is not feasible: although HOT lanes do not exist in Canada, their eventuality should not be ruled out.

Equally important, even if private sector financing is not involved, increasingly public agencies are subjecting new road and transit infrastructure to benefits case analysis: at its heart, this process is based upon a quantitative economic benefit-cost analysis but it also includes other criteria, such as social and environmental measures, which may be qualitative, quantitative or both. For example, as applied by Metrolinx for transit infrastructure, the benefits case analysis comprises five "accounts" (in turn based upon the BC Ministry of Transportation and Infrastructure's Multiple Accounts Evaluation [MAE] framework): The first two, transportation user benefits (e.g., travel time savings) and financial impacts (e.g., ridership and revenues), are used for the benefit-cost analysis. The other three accounts, environmental impacts (e.g., greenhouse gas emissions), economic impacts (e.g., new jobs created) and socio-community (e.g., shaping of land use). The Metrolinx process compares a proposed initiative with other alternatives.[2] BC's MAE framework, established in the 1990s, has been applied to both roads and transit. Two key points to note:

- The benefits case analysis is required by Metrolinx and BC MoTI as a prerequisite for funding. It provides a systematic and consistent framework within which different projects can be assessed.

- Many of the inputs required for the process are derived from travel demand models: using the examples cited above, these include travel time savings, ridership and greenhouse gas emissions.

In sum, the need for improved credibility and reliability in travel demand forecasts is not strictly a 'private matter:' it has important implications for public sector planning as well.

4. SEVEN WAYS TO IMPROVE CREDIBILITY AND RELIABILITY

The P3 experience has several points to offer public sector modelling. These are:

1. Recognize that the financial community is now in the decision-making group. Financial approvals commonly were separated from planning and engineering approvals – for example, a functional plan or environmental assessment is conducted prior to submission for funding. Now, however, the financial feasibility

---

[2] Several examples of benefits case analyses may be found on Metrolinx's website: http://www.metrolinx.com/en/regionalplanning/projectevaluation/benefitscases/benefits_case_analyses.aspx

of a project is determined as part of the planning process. It is not sufficient to say that the model process responds to the requirements, often mandated by planning legislation, of an environmental assessment or a transportation master plan. The implication is that the model must be able to respond to different questions, covering different time periods, geographies and tighter model validations. These are described below.

2. Short-term is as important as the long-term. Traditional transportation plans commonly focus on long-term conditions, looking 10 and 20 years into the future (and often beyond). Aside from these fixed intervals – current base year and these future years – there tends to be little consideration of the pace of growth between horizons or, especially, of the impacts of short-term business cycles. Also important is the short-term ramp-up period: this reflects the build up of usage from the time a facility is opened until stable flows are realized (and it is steady-state stable conditions that models forecast). Ramp-up is characterized by very rapid growth, and commonly lasts from 3 to 6 years. P3 initiatives pay particular attention to ramp-up volumes and the resultant revenue stream, because significant 'front-end' costs are incurred during construction. However, ramp-up occurs for both tolled and non-tolled facilities.

   The short-term is important for three reasons:

   a. The base year calibration may be a misleading indicator of future travel. Horizon year demographic and socio-economic forecasts, upon which travel demand forecasts are based, are concerned only with the end points and not with the intervening years and, accordingly, tend to 'wash out' short-term fluctuations. However, a calibration year that occurs in an economic peak or trough will result in relationships that may not make sense in the future: for example, trip generation rates that are unusually low or high because of a business cycle. These discrepancies are not simply obviated by forecasted demographic and socio-economic inputs or by adjustment factors.
   b. There is some evidence that if a short-term forecast is lower than expected, then the long-term forecasts similarly may never catch up (i.e., if the forecast is 'off' now, it cannot always be assumed to be addressed in the future); or that the forecast at a particular horizon may be delayed.
   c. The factors that influence traveller behaviour in the short-term may differ from those of the long-term. In particular, changes in fuel prices, increases in fares or parking cost, GDP growth and other factors that are related to (or define) business cycle fluctuations have been shown to be better indicators of short-term traveller behaviour than demographic and socio-economic which are better indicators of long-term behaviour. One way of addressing this need is through the use of econometric models, which predict short-term travel behaviour as a function of the aforementioned price and cost indicators.

3. Validation is as important as calibration. Several methods are described modelling texts for validating models. The point to add here is that the *appropriateness* of the model for the problem at hand also must be considered. This is more than the validation criteria, or how well the model has been calibrated. Models commonly must be adapted for the specific planning initiative that is under consideration: it is not sufficient to take the model as-is, regardless

of how well it has been calibrated, and simply add a new link to it. For example, if boarding and alighting volumes by station along a proposed rapid transit are being analyzed, then a validation along the screenlines that the corridor crosses is at too aggregate a level to be sufficient. Similarly, the distribution within a screenline becomes important. This can mean ensuring that:

a. The model zone system and network have been coded to incorporate the planned initiative *and* to the level of detail that is appropriate to the initiative. This is typically done for a sub-area model, in which attention must be focused upon a specific corridor or geographical area.

b. Demographic and socio-economic inputs are at a sufficient level of spatial detail and are up to date (both current year and forecast values [the latter meaning that the latest trends have been taken into account in the forecasts]).

c. The demand model for the sub-area must be reviewed, so that it reflects conditions within the area. This may require the refinement of demand model formulations within the study area; for example, recognizing that the proportion of through traffic may be on par with that generated by the sub-area zones, or that mode choice for this sub-area may differ from that of the model as a whole.

d. The potential markets for the initiative have been identified, distinguished from other demand, and their characteristics and motivations understood. In other words, the total potential travel demand from which users of a planned facility must be understood to be meaningful; it is not simply a matter of assuming that demand in excess of supply would be met by a new facility. Choices – mode, route and time of day – must be clearly understood and distinguished from each other: in other words, the impacts of an initiative on mode choice must be analyzed, with route and time choice considerations held constant, and so on.[3]

4. Modelling travel behaviour under pricing is fundamental. The crux of a P3 forecast is the relationship between the fee (toll) paid and the perceived benefit; that is, savings in travel time. The resultant value of time (VoT) can vary by trip purpose, mode, time of day and road conditions (congested v. free-flow). The VoT relationship, measured in $/hour, can be discerned from observed conditions where tolled facilities exist; however, VoT can change with time and for different facilities / corridors in the same area. The tolling scheme also has an

[3] Some models consider toll diversion as a mode choice, while others incorporate it only in trip assignment (as a route choice). Tolling as a route choice assumes that drivers will stay in their vehicles when a toll is imposed, and so the only choice is between tolled and free routes. Tolling as a mode choice recognizes that drivers could have the option of switching modes (taking transit, for example), rather than paying a toll. The mode choice approach also maintains an internal consistency within the model; on the other hand, if transit or other alternatives to driving are not realistic options, then a more practical approach is to focus only on the route choice. Time of day choice allows for drivers to shift their travel times rather than take a tolled route: in other words, even though a tolled route may provide a faster and more reliable alternative, some travellers may simply elect to stay on their current route and advance or delay their trip outside a peak period to a time when congestion is lower.

impact – for example, electronic tolling (in which transactions are recorded electronically and the driver is billed each month) versus a cash toll at a toll booth. Less easily quantified are drivers' perceptions regarding comfort and safety (e.g., the tolled road has up-to-date geometrics because it is newer than the existing alternatives).

Moreover, current public transit riders may value their current trip, which involves a slow but transfer-free bus ride, over a planned alternative (e.g., a new rapid transit line fed by bus), which may provide faster door-to-door trips but requires a transfer. Mode choice models commonly account for the 'penalty' associated with transfer (and with other attributes of transit use, such as waiting time); however, for new systems with which there is no prior experience in a region, current experience may not be applicable. For example, the observed transfer penalty from one bus to another bus may not correctly the penalty associated with transferring from a bus to a new LRT.

P3 initiatives often consider infrastructure that does not yet exist in a region; that is, infrastructure for which its travellers have no prior local experience. Without this experience, using existing VoT or transfer experience (such as may be available in the first place) is not considered by lenders to be sufficiently reliable indications of how travellers are likely to behave with the new infrastructure in place.

As a result, "bankable" or "investment-grade" traffic/ridership and revenue studies (that is, studies that are used as the basis for securing project financing) generally require the conduct of "stated preference" (SP) surveys. These are studies that are used to quantify consumers' responses to goods with which they have no prior experience. Originally developed for consumer products market research, SP surveys provide quantifications of VoT and travellers' perceptions of new modes. The results can be combined with origin-destination (or "revealed preference") surveys to improve mode choice models, or integrated into volume-delay functions to improve trip assignment.

A traditional public sector transportation planning application may not require its forecasts to be at a bankable quality. However, the use of SP-derived parameters to complement observed behaviour (from OD surveys) can provide improved insight into travellers' expected behaviour with new transit modes or facility types with which they have no prior experience. Moreover, SP surveys have been used to consider the viability of other hard-to-model policy initiatives, including transportation demand management measures.

As well, Canadian models are now considering household income as a determinant of travel behaviour, especially mode choice: greater household income is associated with higher vehicle ownership but also with trip generation rates, trip distribution and mode choice. Traditionally, vehicle availability or household type served as proxies for this behaviour: these variables tended to be easier to gather in origin-destination surveys or were available from other sources. However, income has now been asked on some Canadian origin-destination surveys, and so it can be expected that modelling techniques will start to incorporate income.

Finally, the importance of including prices and costs in short-term forecasts also has been noted.

5. Risk analysis is more than sensitivity testing. Sensitivity tests commonly examine the impacts of higher or lower values in certain model inputs – e.g., increasing or decreasing a horizon-year population by 10%. While such tests are important to understanding the sensitivity of the model to different parameters, one by one - they can which of several different parameters the model has the greatest impact - they maintain the same basic assumptions and relationships and are focused on the inputs. In contrast, risk analysis addresses the inherent uncertainties in the model, its inputs and the forecasts. Statistical techniques can be used to assess the likelihood of achieving certain outcomes – e.g., a particular volume on a tolled expressway. Risk analysis also allows different combinations of parameters to be combined, which reflects a more likely scenario than the one-by-one analysis of parameters in a sensitivity test. (The rationale is that changes in one parameter likely will impact changes in another; however, the changes may not be at the same rates.) Moreover, risk analysis also allows for externalities to be considered: for example, changes in fuel prices (although assumptions may be required for these, they are explicitly identified). In sum, sensitivity tests and risk analysis are not substitutes for each other: both have roles to play in assessing the robustness of a forecast. Finally, it is important to note that neither sensitivity testing nor risk analysis substitutes for model and input validation.

6. Look at more than the commuter peak hour. Time of day shifting is a choice for some travellers, as much as route choice or mode choice. As noted, advancing or postponing a trip outside a congested peak period may be as valid a response as switching routes or modes. Moreover, growth in inter-peak travel means that the traditional basis of model forecasts in the commuter peak period may be missing an important component of the travel market in terms of both absolute numbers and in different trip purposes, such as work-related trips and medical trips, which are relatively unimportant during the commuter peaks. Similarly, commercial activity can be significantly higher during the inter-peak and evening / nighttime periods than during the two commuter peaks. Finally, benefit-cost analyses and even other tabulations, such as fuel consumption and GHG emissions, are considered at daily and annual levels: since these levels must be expanded from model outputs, the availability of several time periods, each with its different composition of trip purposes and modal uses, provides a more robust basis for expansion.

7. More emphasis on calibrating travel time. Models are calibrated in terms of trips, at both the matrix and network levels. Travel time and (the related) speed are important determinants of mode and route choice; however, these factors tend not be considered in the validation as well if at all. A P3 forecast commonly requires a validation of the travel time along the corridor in question and along parallel alternatives. Also important is the consideration of stop-start times during the trip, meaning that delays at intersections become important. Finally, values of time can increase with congestion.

5. CONCLUSION: WHY THIS IS IMPORTANT

In conclusion, scarcer finances for projects coupled with an increased need for justifying large-scale infrastructure expenditures to a sometimes-skeptical public (and politicians) suggest that financial considerations will become more important in transportation planning. This is true for both public and P3 initiatives. The experiences of models that have been developed for P3 initiatives provide some insight into how public sector models could account for road pricing considerations; and the aforementioned discussion suggests that these considerations also could be applied to the analysis of other policy issues, such as TDM, and for new rapid transit initiatives. Parallel to this is a need to consider pricing and costs in more profound ways in models, whether through using household income as a determinant of travel behaviour or accounting for the impact of prices and short-term business cycles on demand.

It can be expected that tour- and activity-based models will provide the fundamental methodological basis for improving the depiction of travel behaviour. This, of course, is essential. However, it is important to note that there will still remain a need to improve *credibility* and *reliability* of the model outputs; and it is hoped that the experiences and approaches described in this paper will facilitate those improvements.