Developing model of toll road traffic forecasting during ramp-up period

Weka Indra Dharmawan1,*, Ade Sjafruddin2, Russ Bona Frazila2, and Febri Zukhruf2

1Doctoral Study Program of Civil Engineering, Faculty of Civil and Environmental Engineering Institut Teknologi Bandung, Bandung, Indonesia
2Transportation Engineering Research Group, Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Bandung, Indonesia

Abstract. The Feasibility of Toll Road project investment that uses Public Private Partnership (PPP) scheme is largely determined by the accuracy of traffic forecasting as a reflection of revenue streams. The accuracy level of traffic forecasting is needed to get a description of risks and uncertainty of Toll Road projects to be invested. Unfortunately, the international studies of forecasting show trend of overestimate, particularly occurred in the early years of the new Toll Road operation. It is the acute problem in the short term Toll Road investment or known as ‘ramp-up period’. The conventional model of aggregation based on socio-economic and demographic growth has not been able to anticipate the problem, since the ramp-up period is a process of learning and adaptation for regional travellers due to changes in travel behaviour after the new Toll Road infrastructure began to operate. Accordingly, the disaggregation model is considered the most realistic used to predict the potential traffic that occur during the ramp-up period. This paper provides a review of several studies dealing with traffic forecasting model for Toll Road projects during the ramp-up period.

1 Introduction

The Feasibility of Toll Road project investment that uses Public Private Partnership (PPP) scheme is largely determined by the accuracy of traffic forecasting as a reflection of revenue streams. This is due to some or all of the projects generally use loan funds with warranty in the form of confidence to the number of traffic forecasting that will occur, or even without using other guarantees. Therefore, accuracy and reliability in predicting the traffic volume are needed for risk and uncertainty of the Toll Road project to be invested.

Unfortunately, the results of international studies on the performance of Toll Road traffic estimates in several countries have shown been overestimate of 20% to 30% in the first few years of operation [1]. According to Ortuzar and Willumsen [2], this condition is the result of the ramp-up period, which is considered a short-term acute problem that occurs in the first few years when a new transportation facility or service starts operating [3].

The conventional model that is commonly used to predict the volume of Toll Road traffic flow is still not able to accommodate the influence of ramp-up. The conventional model approach is still aggregated based on the socio-economic and demographic growth of a study area, while during the ramp-up period it is more dominated by the traveler’s behavior of the regional movements. Travelers still don’t know what benefits the new Toll Road has to offer. It takes a learning and adaptation process for the travelers by maximizing the utility or minimizing the costs of the perceived benefits while using Toll Road during the journey.

The ramp-up period is important for Toll Road projects, especially for credit rating agencies or lending institutions that are very concerned about the performance of an investment asset in the early years of operation. This is because of the cumulative curve of a project's cash flow will be at its lowest point. All project debts have been withdrawn, but the revenue streams have begun to be generated. Potential difficulties that will be faced (even a failure in payment of debt) from a project can be predicted [4].

This paper is the result of review from several studies on ramp-up phenomena and traffic forecasting model related to Toll Road investments. The first section of this paper describes the tendency and magnitude of inaccuracy in the Toll Road traffic forecasting, then in the second section explains the understanding of the ramp-up period that occurred in the new transportation service infrastructure. The third section shows several methods that have been done to anticipate the effect of the ramp-up period in Toll Road traffic forecasting.
2 Frequency and magnitude of misprediction

Hartgen [5] conducted a study of existing travel demand prediction models over a span of 50 years related to accuracy and utilization of public decision making. It was concluded that the results of the model tended to be overestimate at around 30%, especially in the Toll Road. According to Flyvbjerg [6] the condition like this can take place throughout the cycle of the concession period. The weakness of the model is more due to not involving behavior content, inaccurate input and assumptions used, policies that are less sensitive and excessive complexity [5]. In addition, political institutions and environments that tend to encourage optimism biased, and too neglect the impact of risk that will occur so that it affects the quality of the model. There are fundamental problems not only from the technical side, but from the institutional sector (non-technical) also the cause of the inaccuracy of the model made.

This level of accuracy according to Bain [1] is an uncertainty and risk from a Toll Road investment. Traffic forecasting is generally used to measure bankability of an investment in Toll Road infrastructure projects whose construction is carried out by the private sector. Bain [1] used the 2005 research data [7] showing the ratio between the actual to the traffic prediction results in the first year presented in graphical form as in Figure 1.

![Fig. 1. Distribution of actual to forecast traffic volumes (Source: Bain and Polakovic, 2005)](image)

Data is normally distributed with an average ratio of 0.77 and a deviation standard of 0.25. Values ranging from 0.14 to 1.51 mean 86% of actual traffic volume is below the predicted results and 51% are above predictions. This fairly large range of errors illustrates the value of possible of uncertainty as part of the traffic risk that must be guaranteed by the investor. Welde and Odeck [8] denied what Bain [1] feared that when the actual traffic volume in the first few years operating below the predicted results would be difficult to catch up in the following years, it takes 5 years to catch up.

Baeza and Vassallo [9] used ramp-up phenomena to evaluate the actual traffic to forecasting proposed in the Toll Road investment bid documents in Spain by investors. The Toll Road traffic forecasts are generally consistently overestimate and occur during the ramp-up period. The overestimate tendency according to Baeza, and Vassallo [9], is a deliberate strategy carried out by investors in the auction process of the Toll Road project and not based on systematic errors in transportation modeling, with the hope that the government will renegotiate when the actual traffic conditions is far below what is expected. The assessment of the risk of Toll Road infrastructure investment projects is associated to the ramp-up factor according to Bain and Polakovic [1] as in Table 1.

<table>
<thead>
<tr>
<th>Years</th>
<th>Toll Road Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low Risk (%)</td>
</tr>
<tr>
<td>1^t</td>
<td>80</td>
</tr>
<tr>
<td>2^nd</td>
<td>90</td>
</tr>
<tr>
<td>3^rd</td>
<td>100</td>
</tr>
<tr>
<td>4^th</td>
<td>85</td>
</tr>
<tr>
<td>5^th</td>
<td>88</td>
</tr>
<tr>
<td>6^th</td>
<td>90</td>
</tr>
<tr>
<td>7^th</td>
<td></td>
</tr>
<tr>
<td>8^th</td>
<td></td>
</tr>
<tr>
<td>9^th and later.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 showed that the lowest risk of Toll Road investment occurs when the length of the ramp-up period is under 5 years, medium risk is between 5 and 10 years, while it is considered high risk in investing in Toll Roads when the ramp-up period is unfulfilled more than 10 years [1].

2.1 Risk typology of toll road

Research on the reliability of predictions of traffic flow volume and Toll Road revenues is carried out by Mueller [10]. The sample used was 20 Toll Roads in eight US states. Muler [10] divides Toll Roads into 4 groups based on several categories, which according to Kriger et al. [3] there are some overlapping, are as follows:

- Group 1: High Congestion, Suburban Areas (President George Bush Expressway, Dallas, Texas);
- Group 2: Outlying Roads of Metropolitan Areas and Established Agencies (Foothill North, Orange County, California);
- Group 3: Developed Corridors, Parallels of Existing Roads and/or Faulty Economic Forecasts (Sam Houston Toll Road, Houston, Texas); and
- Group 4: Least Developed Areas (E470, Denver, Colorado).

The average ratio of actual income to prediction for each Toll Road from the four groups during the first five years of operation (ramp-up period) can be seen in Figure 2. The ratio value of one means between actual and prediction is the same. The ratio value of less than one is meant the actual income less than the prediction and vice versa.
Figure 2 showed that the prediction of Toll Road revenues is consistently overestimate, occurring in Groups 2, 3 and 4 in the first five years of the ramp-up period, while Group 1 has a prediction of underestimate Toll Road revenues to actual. Muler [10] added, the interesting thing is the value of the Toll Road revenue ratio did not increase substantially during the five-year operation in the ramp-up period, except in Group 4 which was still possible to be studied.

3 Ramp-up

When a new transportation infrastructure or service begins to operate, it takes some time for travel demand to grow and stabilize, as travelers adjust to changes in their new pattern of movement. This initial adjustment process was defined by D'Este [11] as a ramp-up. During this period, actual travel demand is usually lower than what is expected, but over time, the difference continues to decrease and eventually disappears. D'Este [11] and Kriger et al. [3] consider that the ramp-up effect is important for Toll Road infrastructure as an early indicator to assess the performance of traffic forecasting on actual conditions.

According to Kriger et al. [3] the ramp-up period is the time needed for new transportation infrastructure to reach their full potential demand, without considering the influence of socio-demographic growth and regional economy. Nunez [12] argument that the ramp-up period is a reflection of not knowing the traveler to the new transportation infrastructure and the benefits of what will be felt.

Wibowo [13] defined the character of the ramp-up period with the high growth rate of abnormal traffic volume but then gradually slows with over time. After the ramp-up period ends, the volume of traffic will usually grow stably as other alternative routes in the same corridor. Jenks et al. [14] linked the ramp-up period with the learning and adaptation processes of regional travelers to the condition of new trips due to the Toll Road infrastructure that began operating.

Figure 3 showed a diagram to understand the concept of a ramp-up phenomenon. During the ramp-up period, the traffic volume experienced an increase and tends to fluctuate (oscillation). These characteristics will be different when entering a steady state, with graphical oscillations of the volume of traffic flow that is almost or not even happening [15]. The steady state indicates that the Toll Road infrastructure has reached its equilibrium and then enters the maturity phase. In this phase, the growth characteristics of travel demand are more influenced by socio-demographic conditions, regional economics, as well as changes in monetary policies and conditions (such as changes in toll rates, fuel prices, or other operational costs).

Another thing that happened during the ramp-up period in Toll Road infrastructure was the occurrence of diverted traffic and induced traffic [14]. Diverted traffic is the transfer of traffic from other routes to the new route, while induced traffic is the increase in the volume of new vehicle traffic due to an increase in road capacity (or new road segments). According to Mokhtarian et al. [16], this event shows the existence of a travel request that was not seen before or known as latent demand.

Bain and Polakovic [1] argued that the ramp-up period is not as aggressive as what is predicted so that the actual volume of traffic is difficult to pursue predictions that tend to overestimate. These influences are generally considered as a source of risk in traffic forecasting in transportation projects [3].

It could be said that there is no strong consensus on the theory of what is the exact cause of the ramp-up phenomenon, but there are three reasons that have been identified [15,17]. First, the effect of the learning curve is most often used to explain the ramp-up phenomenon. The process of maximizing utility or minimizing the cost of transportation service choices is understood by travelers to build their final decisions by learning from mistakes they have experienced and felt before. Second, someone sometimes changes their travel behavior to respond to opportunities offered by a new travel option. Although the cause is not completely separate from the effect of the learning curve, it is seen more often to be perceived as switching demand and induction demand. Finally, operational teething troubles is one of the reasons for several reasons for the ramp-up phenomenon. Two previous reasons, namely the learning curve and behavioral change are viewed from the point of view of the traveler, while the last issue is associated to the operator. These three reasons can be understood as the learning and adaptation processes of the travelers.
Cheung and Polak [19] to analyse the elements of risk and uncertainty in predicting the traffic volume. Case studies were done on Toll Roads in Spain. Stochastic simulations are used the bootstraping techniques by calculating capacity constraints in predicting long-term traffic flow volume (during the concession period). The dynamic model is determined by the speed adjustment associated with the ratio between the actual traffic and the maximum capacity of the Toll Road.

Response Surface Methods (RSM) are used by Adler, et al. [21], to analyse the quantitative risk of traffic forecasting. Model parameters and conditions in the future are considered as an element of uncertainty of input data by including the effect of ramp-up factors. The probability distribution of the uncertainty factor was developed based on previous experience. The traffic forecasting model can be synthesized by using RSM into a mathematical analysis equation model applied to Monte Carlo simulations to obtain the probability distribution of the desired output.

5 Agent based model approach

The ramp-up period is more related to the learning and adapting process of regional travelers due to changes in travel patterns caused by the existence of toll roads that have started operating [14]. There are changes in travel patterns assumed to lead to dynamic traffic conditions from day to day, as a result of the behavior of individual travelers in choosing to learn and adjust to new routes [22]. Dynamic model of route choice from day to day, according to Wei, F., et al., (2014) can be approached to an agent-based simulation model (ABSM) involving reinforcement learning.

The most important element in ABSM is the signal of the process of decision making from the traveller. The volume of traffic that occurs every day is not the result of one decision making in a choice offered, but is a decision process carried out repeatedly. The repetition process, according to Grether and Nagel [23] is considered as part of the learning mechanism and also as adjustment behaviour when there is a new route offered [11], as new Toll Roads operate during the ramp-up period found on the road network.

Travelers will be able to feel the level of benefits offered from their experience of making repeated trips using the new Toll Road. Psychologically, there is a process of competition or competition in cognition (in the form of awareness or feeling) of the traveler to take the best decision on the route to be chosen during the journey, this is due to the human bounded rationality (BR) and uncertainty of environmental. Thus, stochastic rules in the route selection process are very influential [11].

If the value of utility from a segment or road route used before is felt to be a maximum (satisfactory outcome) for the traveler, it is very likely that the same decision will be chosen the next time. Otherwise, it will be avoided when what is perceptive is unsatisfactory as expected. This characteristic is identical to the theory of reinforcement learning (RL), illustrated when an activity provides maximum results, the probability of doing the same activity will continue to increase in the current time round. In other words, the behavior associated with an activity will get reinforcement [24]. RL asserts that decision-making behavior is directed to a random pattern caused by the human bounded rationality and the uncertainty of environment [4].

In agent-based modeling and simulation, travelers are considered as an entity in the form of agents who have the ability to learn and make decisions autonomously, with their basic behavior patterns as users of Toll Road. At the macro level, a review is carried out on a set of agents (multi-agents) that interact with a group of other agents or their environment (see Figure 4). A typical agent-based model has three elements:

1. **Agents**, their attributes and behaviors.
2. **Agent’ relationships** and methods of interaction. An underlying topology of connectedness defines how and with whom agents interact.
3. **Agents’ environment**. Agents live in and interact with their environment in addition to other agents.
5.1 Utility function

As with other consumer products [2], travelers will choose a path that is considered to provide maximum satisfaction. This assumption as perception and preference based on the characteristics of the route, then bound to the personal attributes of travel agents such as the amount of income, age, gender, level of education, and other socio-economic characteristics. The concept is generally related to discrete choice models, where the probability of each individual in choosing an option offered is a function of their socio-economic characteristics and the relative attractiveness of the choice [25]. To express the attractiveness of an alternative, the concept of utility (something maximized by each individual tripper) is used. The Random Utility Maximization formula must take into account the time and constraints [25]. The general formula of the traveler choice model is like in Equation (1) as follows:

$$ U_{ir} = V_{ir} + \varepsilon_{ir} $$

where:
- $U_{ir}$ is the utility observed from a route $r$ to traveler $i$ on day $t$.
- $V_{ir}$ is the function of the traveler characteristics ($Z_i$) and the alternative attribute of the route ($A_{ir}$) and can be written as $f(Z_i, A_{ir})$.
- $Z_i$ is traveler characteristics.
- $A_{ir}$ is the attributes of an alternative route $r$ for the individual tripper $i$ on day $t$.
- $\varepsilon_{ir}$ is an error value or utility that is unknown or involved from an alternative route $r$ to traveler $i$ on the day $t$.

$V_{ir}$ is a representation of traveler $i$ with respect to the alternative route $r$ and is a function of attribute $x$ on the day $t$, which may vary from individual to each other. However, certain homogeneous assumptions are needed for the population of the traveler. All travelers are considered to face the same set of alternatives and constraints [25]. The general formula of $V_{ir}$ as in Equation (2) as follows:

$$ V_{ir} = \sum_k \theta_{ir} x_{ir}^k $$

where:
- $\theta_{ir}$ is the parameter constant value for a group of all travelers who are homogeneous trip, but may differ from the attribute $k$ from the alternative segment $r$ that exists on day $t$.
- $x_{ir}^k$ is the vector attribute $k$ against the alternative route $r$ of the traveler $i$ on day $t$.

The traveler $i$ selects the maximum utility alternative, and the possibility of an alternative route $r$ will be chosen by the traveler $i$ on the day $t$ if and only if:

$$ U_{is} \geq U_{ir}, \forall s \in R $$

$$ P(r) = \frac{P(U_{ir} \geq U_{is}, \forall s \in R)}{P(U_{ir} = \text{max}_{s \in R} U_{is})} $$

where:
- $P(r)$ is a random route selection probability index of $r$.
- $R$ is a set of routes in the network.

Wei et al. [11] and Shang et al. [26] define utility as a travel time or travel cost.

5.2 Route choice model (RCM)

The behavior mechanism that underlies the traffic assignment model is a choice or process of decision making. When a traveler travels from the origin $o$ to the destination $d$, it will be faced with several alternative route options found on the road network. There is only one route that is considered to be significantly different from other routes by the traveler, and it is decided to be used to reach a common destination. However, not all travelers in the journey will use the same route, they will rationalize the availability and preferences they have.

According to Akiva et al. [27] RCM involves evaluating the perception of travel about the characteristics of the route, including travel time, travel costs, distance, security, reliability, and so on. The traveler $i$ will always maximize the utility or minimize the cost, so the alternative route will be chosen by the traveler $i$ on day $t$ if the probability of route selection is like in Equation (4).

The logit model is used to describe the random route selection of each traveler as in Equation (5) as follows:

$$ P_i^r(\text{route}_r) = \frac{\exp(V_{ir})}{\sum_{r'} \exp(V_{ir'})} $$

where:
- $r'$ is the logit for all preferred routes (in this case Toll Roads and not Toll Roads) that connect OD pairs.

5.3 Reinforcement learning

The learning process (reinforcement learning) is based on the interaction between the traveler and the Toll Road that is used and felt as a result of the chosen route decision making process. A new decision and a reward for the behavior of the Toll Road route choice indicates the suitability of what is expected (or a penalty if the reward is negative). Figure 5 can explain the mechanism of learning from interactions between agents and the environment.
It has become a consensus that the prediction of the toll road traffic volume with PPP scheme tends to overestimate 20% to 30% of the actual, and has become an optimism bias. One of the causes of inaccuracy in traffic forecasting on Toll Road infrastructure is the ramp-up effect or transition period which is rarely taken into account in the modeling. Ramp-up is the period of time needed for the volume of traffic to reach their full potential, without considering growth after the new Toll Road infrastructure facility operates. Then the probability is updated again when the path is selected repeatedly.

5.3 Bounded rationality

The basic assumption of the deterministic model is that the route selection behavior of the traveler is perfect rationality, who always tends to switch to the route with the lowest travel costs [26]. Under equilibrium conditions, travel costs on all routes available from the origin to the destination are equal and minimum. But in reality, travelers do not always choose the route that has the least cost of travel [28] due to various reasons such as lack of information and inertial behavior. The uncertainty is assumed that the traveler will follow the stochastic route selection rules, the greater the probability of the path chosen, the greater the likelihood that the path will be used [11].

6 Summary and recommendations

It has become a consensus that the prediction of the toll road traffic volume, a stochastic method approach with Monte Carlo simulation is used in order to accommodate the risk and uncertainty traffic in the form of probability distributions, while the learning curve that describes the characteristics of toll road traffic volume during the ramp-up period has not been seen. Disaggregation model approach by agent-based model and simulation involves reinforcement learning in traffic forecasting as a potential toll road user during the ramp-up period is considered quite realistic. The hope is that it can provide a picture of the learning curve on the traveler’s decision making attitude during the process of route choice.

References

22. F. Wei, S. Ma, N. Jia, “A Day-to-Day Route Choice Model Based on Reinforcement Learning” (2014)

27. B. Akiva, Moshe, M. Bierlaire, Discrete Choice Methods and Their Applications to Short Term Travel Decisions, (1999)


29. B. Flyvbjerg, **39** (6), 522-530, (2005)


34. K. Ozbay, A. Datta, P. Kachroo, Travel Patterns and Behavior, Effects of Communications Technology, **1752**:38-46 (200)


43. X. Di, and H. X. Liu, Transportation Research Part B, **85**, Elsevier Ltd. 142-179 (2016)