Australasian Transport Research Forum 2013 Proceedings
2 - 4 October 2013, Brisbane, Australia
Publication website: http://www.patrec.org/atrf.aspx

A Case Study of Risk Quantification in Mesoscopic Modelling of Public Transport

Ting Yu[1]

[1]Planning and Programs, Transport for NSW, 18 Lee Street, Chippendale NSW 2008

ting.yu@transport.nsw.gov.au

Abstract

The failure of recent public transport projects raises concerns regarding the reliability of public transport demand models. One shortcoming of the practical demand models is that their results are presented using point estimation. By contrast, transport modellers could present their prediction using interval estimation. This would help decision-makers understand the whole distribution and quantify the downfall risk associated with public transport projects by controlling "Optimism Bias". This paper presents a case study to quantify the risk of using a mesoscopic model for public transport - Public Transport Project Models (PTPM).

In a nutshell, PTPM employs a Nested Logit Model (NLT) that takes two sources of inputs: level of services (e.g. travel distance) and parameters. The parameters are estimated using maximum likelihood methods based on the observed mode choices derived from the Household Travel Survey. The estimation of some parameters relies on only a handful of noisy observations which are barely adequate to faithfully represent the whole population. Modellers need to quantify the relationship between their choice of parameters and the patronage prediction. One approach is to rank the parameters according to their influence on the final result. However, one of challenges of parameter ranking is the difficulty of analytically deriving the first deviation of the given NLT. Numerical simulation, e.g. Monte Carlo, is a rather easy approach to calculate the first deviation. This paper demonstrates how the high-performance PTPM carries out thousands of simulation runs to construct the patronage distribution by responding to different distributions of parameters.

1. Introduction

Public transportation systems are increasingly complex, incorporating diverse travel modes and services. Of thirty-five public transit projects that Martin Wachs of the RAND Corporation, has studied in the US, thirty-three overestimated patronage and twenty-eight underestimated costs (Martin 2009). Transport modellers need to quantify the relationship between their choice of parameters and the resulting patronage prediction.

The Strategic Travel Model (STM) is a world class tool, used for projecting travel patterns in Sydney, Newcastle and Wollongong under different land use, transport and pricing scenarios (James, Andrew et al. 2011). In 2011, BTS developed a mesoscopic model, Public Transport Project Model (PTPM), to address a few limitation of the STM. The STM is inefficient when applied to local public transport projects because of its lack of detail and flexibility. The PTPM is built on the output of STM, and has a focus on public transport and public transport demands. The PTPM gives extra flexibility to model specific public transports projects by improved accuracy and a better level of detail (David 2011). First, it is based on good observed local surveys of public transport travel patterns and transaction records, and thus able to capture local travel behaviours and demonstrate an accurate representation of travel patterns in the corridor of interest. Secondly, it is specifically designed to have shorter run times making it easier to run many tests of alternative schemes and scenarios. Thirdly, it is capable of addressing behavioural factors not included in the strategic model, such as crowding and improved modelling of rail access by car (park-and-ride and kiss-and-ride). The

PTPM has been applied to major public transport projects, including the North West Rail Link. The study area of the North West Rail Link project covers some Greenfield areas, e.g. Rouse Hill and Kellyville. These areas are witnessing large population and employment growth. Due to the lack of existing public transport facilities, the local residents are using car access to the rail stations. All these features distinguish this project from other public transport projects in NSW.

The PTPM employs a Nested Logit Model (NLT) to forecast the public transport demand for the study area. The NLT combines the level of services (e.g. travel distance) and parameters to estimate the mode shares. The physical travel skim is directly observable, but the parameters are estimated by Maximum Likelihood Estimation methods (MLE) according to the observed mode choices collected by the Household Travel Survey (HTS).

The Household Travel Survey (HTS) is the largest and most comprehensive source of personal travel data for the Sydney Greater Metropolitan Area (GMA). This survey is a benchmark for best practice in travel surveys in Australia and around the world, as well as being the longest running continuous household travel survey in the country. However, the PTPM requires more detailed and comprehensive representations of the transport users, and some of requirements are beyond the current HTS:

1) Incompleteness of survey data

The HTS has low spatial coverage and low temporal coverage. The HTS interviewed approximately 8,500 people in 3,500 households annually over the last ten years (BTS 2012). Compared with a population of over 3 million people in Sydney, this survey covers only a small portion of travellers. It is also well-known that travellers have different preferences between work days and weekends, or between inter-peak and peak time within a day. This requires even more data to represent the variance of an individual traveller.

The PTPM estimates 4 demand matrices (for different purposes) each of which contains 2690*2690=7,236,100 OD pairs. By considering the relatively small sample size (85,000 out of 7,236,100), the estimation of some parameters indeed relies on only handful of observations which are barely enough to represent the whole population.

2) Noisy data

Even though there is a systematic method to minimize the data error, unavoidably, the HTS data contains noise due to human errors. These errors can be from data entry mistakes or from false claims from the interviewee.

3) Relevance of data to the future

All the HTS data represents only the past travel behaviours. It is well-known that travel patterns evolve over time due to various social and technology changes.

All these limitation of current HTS could lead to biased estimation of the underlying true parameters. Modellers need to see the impact of any bias introduced in the estimation process on the final forecasting. One of the key questions if parameters are improper, is what is its influence on the final result? Can modellers rank the parameters based on their influence on the final result? Then, based on the rank, modellers can focus on the most influential parameters.

More importantly, this paper demonstrates that the final prediction is presented as interval estimation (e.g. confidence intervals) instead of point estimation that is the expected value. In statistics, the expected value (the first moment) is inadequate to represent an arbitrary distribution. The higher order statistics gives adequate representation of the distribution of the patronage prediction. By understanding the whole distribution, the risks associated with the project are better understood (*Bliemer*, *Rose* et al. 2008). This paper demonstrates a method to discover potential over-estimation or under-estimation of patronage, and control the so-called "Optimism Bias". One of the key challenges of quantifying the risk and uncertainty is that it is difficult to analytically derive the closed-form first deviation of a given

NLT. Numerical simulation, e.g. Monte Carlo, is a rather easy approach to calculate the first deviation. This paper demonstrates the high-performing PTPM carries out thousands of simulation runs to construct the demand distribution by responding to the entire joint distribution of parameters.

A black swan is positive or negative event that is deemed improbable yet causes significant consequence (Taleb 2010). The choice of parameters becomes a kind of art, based on a sixth sense or superior expertise. Very often the choice is conservative and risk-averse. It tends to be average values with the assumption that prediction follows a normal distribution. In this way, the risk is often underestimated by the thin tail of the normal distribution. Additionally the lack of data availability, the tight project schedule, the complexity of the models, and the lack of computing power, all lead to difficulties in estimating the full distribution and measuring the full scale of risk.

2. Nested Logit Tree

Random utility models (RUMs) state that a decision maker, labeled n, faces a choice among J alternatives. The decision maker would obtain a certain level of utility (or profit) from each alternative. The utility that decision maker n obtains from alternative j is $u_{n,j}$, j = 1,..., J. This utility is known to the decision maker but not by the researcher. The decision maker chooses the alternative that provides the greatest utility (Train 2009). This utility function can be decomposed as two parts.

$$u_{n,j} = v_{n,j} + \varepsilon_{\mathrm{n,j}}$$

The first part is the *representative utility* $v_{n,j}$, and it depends on parameters that are unknown to the researcher and therefore estimated statistically. The second part of the utility is unobserved by the researchers. It captures the factors that affect utility but are not included in representative utility.

2.1 Unobservable Components of Utility

The error terms ε are unobserved random variables that are described by a probability distribution. In general, this may be a joint distribution of all the error terms, so we use the vector $\varepsilon_{\mathrm{n,j}} = [\varepsilon_{\mathrm{n,j,1}}, \varepsilon_{\mathrm{n,j,2}}, \varepsilon_{\mathrm{n,j,3}}]^T$, which aggregates the error terms for all products.

Apart from the error terms, other components which are not directly observable include the parameters. For example the bus utility function in the PTPM is:

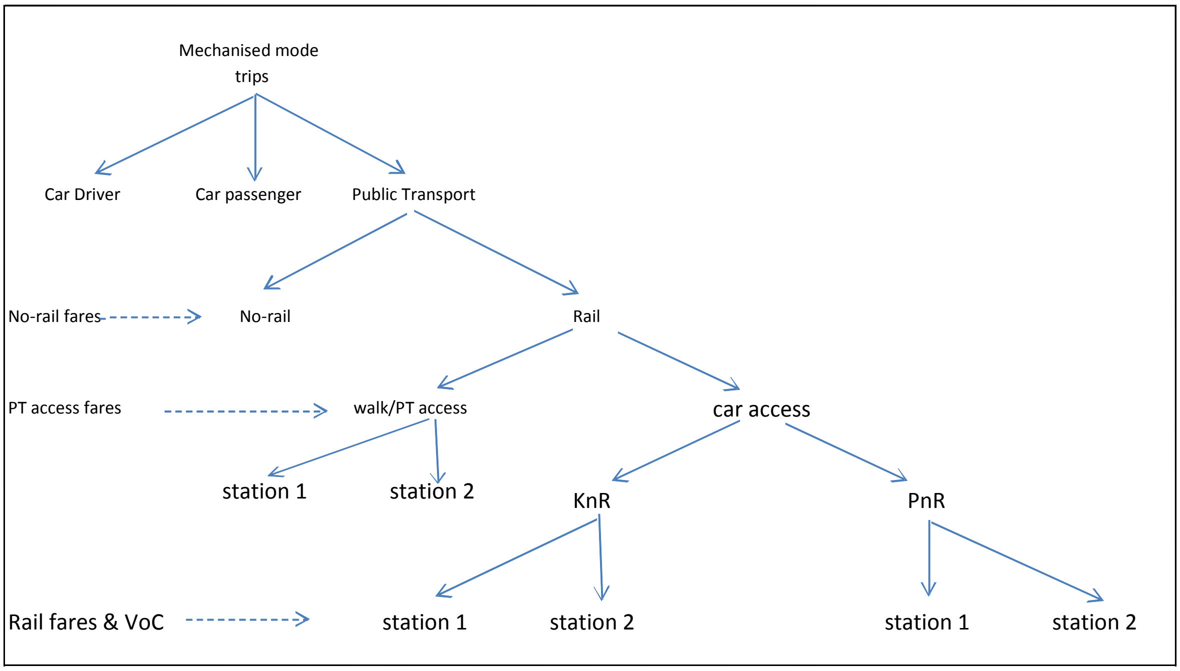
$$v_b = \lambda_{2p} * [\text{ivt} + \alpha_{p,b}*\text{wait} + \beta_{p,b}*(\text{access+egress time}) + \gamma_{p,b}*\text{Int} + \text{fare}_p/\text{VoT}_{p,b}]$$

where $\alpha_{p,b}$ is the wait time weight, $\beta_{p,b}$ is the acess-egress time factor, and is $\gamma_{p,b}$ the interchange penalty. All these parameters are estimated by observing the travel behaviours.

2.2 Public Transport Project Model

The Public Transport Project Model (PTPM) includes a Nested Logit Structure (NLT) for mode choice purposes. The NLT is illustrated in the diagram below, showing the options that will be covered. Level-of-service elements are included at all levels. Also shown is where the matrix-based public transport fares and vehicle operating costs are introduced. For car access and bus/walk access to rail, the best 2 station options are incorporated in the model.

Figure 1 The nested logit structure as implemented in the PTPM



The structure and parameters of the PTPM mode choice model are developed from existing local and international knowledge. STM 3 is the most recent multimodal modelling estimation exercise in Sydney and provides the most sensible starting point to determine the details of the PTPM mode choice model specification. The STM 3 coefficient values are given in Table 1. The STM 3 uses the Maximum Likelihood Estimation (MLE) to estimate the coefficients based on the Home-based Travel Survey (HTS). The STM 3 has 2690 internal travel zones which is the smallest measurement unit. The STM 3 is split into 7 purposes, so the STM estimates at least 7 demand matrices each of which contains 2690*2690 OD pairs. The HTS interviewed approximately 8,500 people in 3,500 households annually over the last ten years (BTS 2012). Clearly the data fed into the STM is not sufficient to cover the whole Sydney area. An under-sample can cause the final estimation to be biased towards very few observations. More importantly, the HTS collected data over the past 10 years. Due to dramatic demographic and economic changes (e.g. an immigration influx and financial crisis); the current travel behaviour can be significantly different from 2000s.

Table 1 STM3 Coefficients

Level-of-Service Parameter	Mode	Coefficient						
		HBW	HBEB	HBEd Pri	HBEd Sec	HBEd Tert	HBS	HBO
First (or overall) wait time	train	4.42	4.03	0.34	0.34	1.72	1.79	0.86
	bus	2.46	2.42	0.42	0.32	1.51	2.63	2.04
Other wait time	train	3.30					0.92	
	bus	1.83					1.35	
Access/egress time (walk, bus)	train	2.73	4.18	0.23	0.23	0.32	0.91	0.20
	bus	1.51	2.51	0.29	0.22	0.28	1.33	0.47
Access/egress time (car)	train	3.24	3.99		0.61	fixed	fixed	fixed

Boardings	train						16.1	12.4
	bus						23.7	29.7
Values of time	train	$4.5-9	$7-13		$16.4	$11.70	$6.18	$5.30
	bus	$7-14.5	$9-16		$16.2	$8.90	$3.14	$1.80
	car	$12-23	$22-43		$24	$32.70	$9.50	$7.20

Expected values of these coefficients are given in Table 2, based on national and international experience.

Table 2 Coefficient Standard Values

Level-of-Service Parameter	**Target Values**	**Issues**
Overall wait time	~2.0 1.4 (ATC)	Implications of non-generic values. There appear to be some unrealistically low and high values.
Access/egress time	~2.0 1.4 (ATC)	Implications of non-generic values. There appear to be some unrealistically low and high values.
Boardings	5-10 mins	Implications of non-generic values. It is not clear what has been assumed if there is no boardings penalty (is it 5mins?). The calibrated penalties seem very large.
Values of time	Business: ~$42 Non-business: ~$11-13	As discussed in the report, many of these implied values of time are different from expectations

3. Case Studies

It is difficult to derive a closed-form relationship between the parameters and mode shares. A Monte Carlo simulation approach is employed to demonstrate the joint distribution in this case study. One of major obstacles of carrying out the Monte Carlo simulation for a large and complex NLT is the computation time of the model. In this case study, 3,000 runs of simulation are carried out to generate sufficient samples to form the distribution. In order to get the results within a reasonable time period, the PTPM is written in Visual C++ to maximize its performance and reduce the runtime. The original PTPM is written in Python, and each run takes 30 mins. The C++ version of PTPM takes less than 1 min to complete one run.

Two sets of parameters are selected: wait time weights (i.e. Alpha) and boarding or interchange penalties (i.e. Gamma). Both wait time weights and boarding penalties are included at three levels of the NLT: Kiss and Ride, Bus access to rail, and Bus only.

- The Alpha for Kiss and Ride is for the wait time between the rail station and destination zone, including the wait time for rail and wait time for connecting bus at the egress leg etc.
- The Alpha for Bus access to rail is for the wait time between original zone and the rail station, including only the wait time for bus to access rail.
- The Alpha for Bus only is for the wait time between original zone and the destination zone, including all the wait time in the trip.
- The Gamma for Kiss and Ride is for the boarding between the rail station and destination zone, including the first boarding for rail and boarding for connecting bus at the egress leg etc.
- The Gamma for Bus access to rail is for the boarding between original zone and the rail station, including only the boarding for bus to access rail.

- The Gamma for Bus only is for the boarding between original zone and the destination zone, including all the boarding in the trip.

In this case study, a uniform distribution between 0 and 4 is assumed for the Alpha, and a uniform distribution between 0 and 20 is assumed for the Gamma. This assumption is based on the STM3 coefficients (see Table 1), which is estimated by MLE, and Coefficient Standard Values (see Table 2), which is specified by the experts using their experience.

Figure 1: Distribution of Alphas: the x-axis is the value of the parameters, and the y-axis is the frequency.

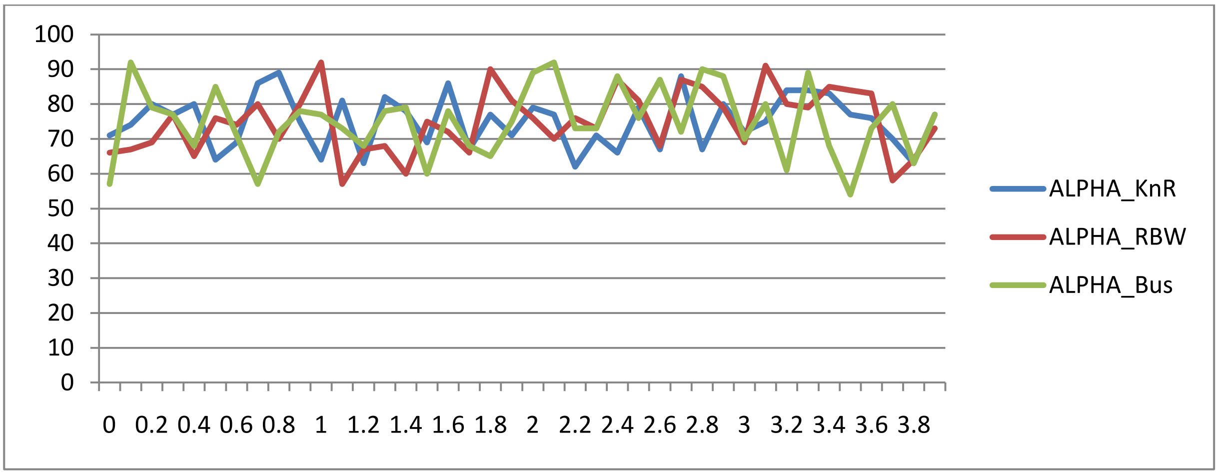
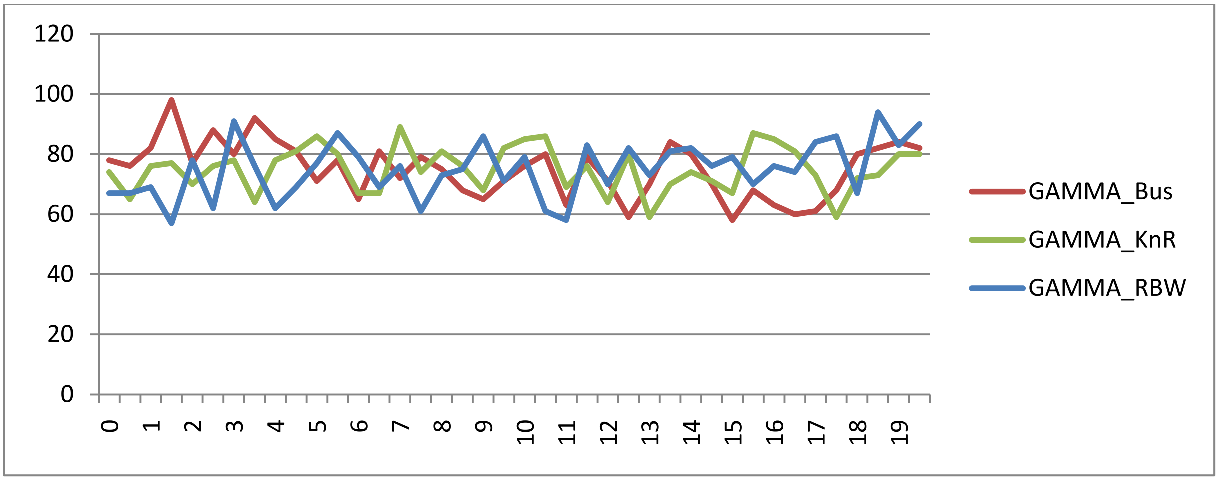


Figure 2: Distribution of Gammas: the x-axis is the value of the parameters, and the y-axis is the frequency.



These parameters must be mutually independent each other. The correlation between parameters are measured and displayed at the table below (Table 3). All the correlation values are well less than 0.05. It indicates the independence between the parameters.

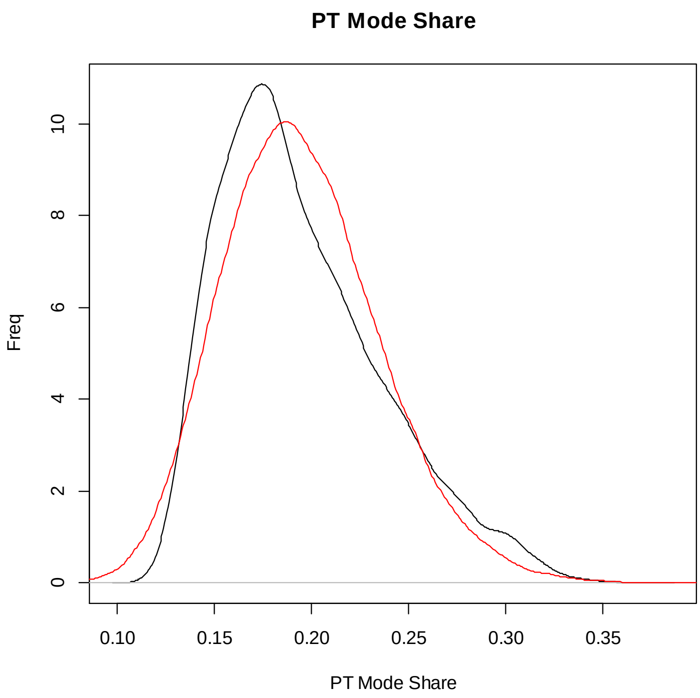
Table 3 Correlation between Choices of Parameters:

	Alpha Knr	Alpha Rbw	Alpha Bus	Gamma Bus	Gamma Knr	Gamma Rbw
Alpha Knr	1	0.015266	0.010492	0.007133	0.011663	-0.02407
Alpha Rbw	0.015266	1	0.000761	-0.00347	0.002467	-0.0098
Alpha Bus	0.010492	0.000761	1	0.01372	0.017285	0.03607
Gamma Bus	0.007133	-0.00347	0.01372	1	-0.01385	-0.02945
Gamma Knr	0.011663	0.002467	0.017285	-0.01385	1	-0.02366
Gamma Rbw	-0.02407	-0.0098	0.03607	-0.02945	-0.02366	1

The PT mode share is a demand-weighted average of all PT mode shares across the 2690*2690 OD pairs. After 3,000 runs, the expected value is 19.52% and standard deviation is 4.18%. The figure below displays the distribution of average PT mode shares of these 3,000 simulation runs. Clearly, the distribution is not a normal distribution. The best fitted distribution is the Beta distribution (alpha = 18.44 and beta = 75.98). When the Alphas are set as 2 and the Gammas are set as 10, the estimated PT mode share is 17.6%. The majority of this distribution is between 12% and 32%, while its peak skews towards the lower end. More importantly, there is still around a 16% chances to have PT share less than 15%, which is a standard deviation away from the expected value at the lower end.

One explanation is that while the wait time weight and boarding penalty increase from zero, the PT share will decrease rapidly from around 36%. But at the high end of the parameter range (4 for wait time weight, and 20 for boarding penalty), the impact of the parameter change diminishes. Therefore, a large portion of PT mode shares concentrates on the lower half. For the decision-makers, this distribution indicates that the PT model share is around 19.25%, and is within the range 12% - 32%.

Figure 3: Distribution of Mode Shares: the x-axis is the value of the mode shares, and the y-axis is the frequency.



The second part of the case study is the parameter ranking, and demonstrates an experiment to measure the impact of an individual parameter choice on the mode shares in the PTPM. This experiment addresses an open question: which parameter has the highest influence on the mode share and demand?
The same two sets of parameters as shown in the previous section are chosen to demonstrate their influence on the average mode share cross the whole Sydney: The first set consists of the wait time weights (i.e. Alpha) for Bus only, Bus/Rail, and Kiss and Ride, and follows a uniform distribution between 0 and 4. The current setting in the PTPM is 2.5 for all three parameters. The second set of parameters consist of boarding penalties (i.e. GAMMA) for Bus only, Bus/Rail, and Kiss and Ride, and follow a uniform distribution between 0 and 20. The current setting in the PTPM is 5 for bus only, and Kiss and Rail, and 10 for Bus/Rail.

The PTPM calculates the difference between the mode shares of base year and future year to estimate the PT demand of future year. In the following experiment, the formula is:

Demand = Base year STM demand * (Future Year Mode Share – Base Year Mode Share)

Figures 4 and 5 display the relationship between PT share and individual parameters. Apart from that, a linear regression between PT share and parameters is displayed below (Table 4) to rank the parameters. According to Table 4, the wait time weight of the Kiss-and-Ride has the largest influence on the PT share. That is around -0.025. That means if we increase the wait time weight of the kiss-and-ride by 1, the PT share will drop 2.5%.

Moreover, the relationship between the Alpha of Kiss and ride and mode share is clearly non-linear. The mode share drops more quickly when alpha is set less than 2.5 than after alpha is set larger than 2.5. In contrast, the Gamma of Bus only almost has no influence on the mode share.

Figure 4 Impact of individual parameters on PT Share: the x-axis is the value of the parameters, and the y-axis is the PT mode share.

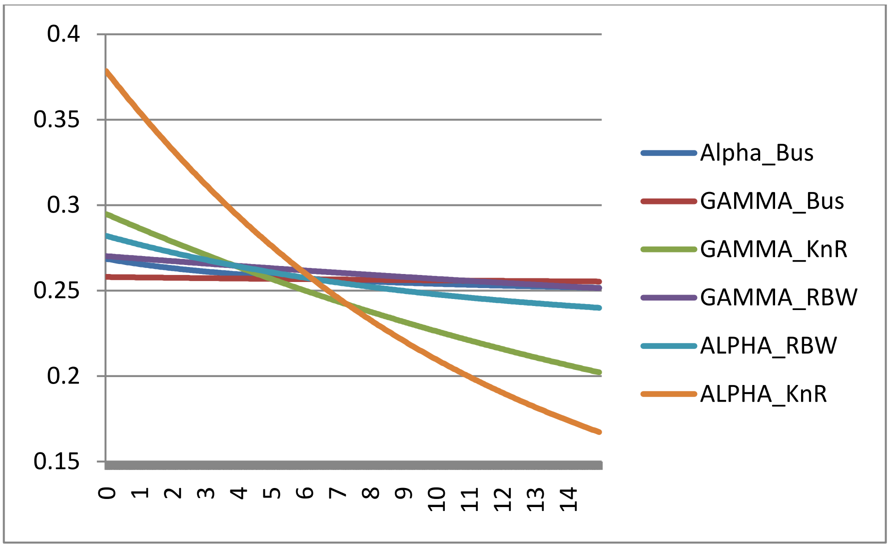


Figure 5 Impact of individual parameters on PT Demand: the x-axis is the value of the parameters, and the y-axis is the number of PT trips.

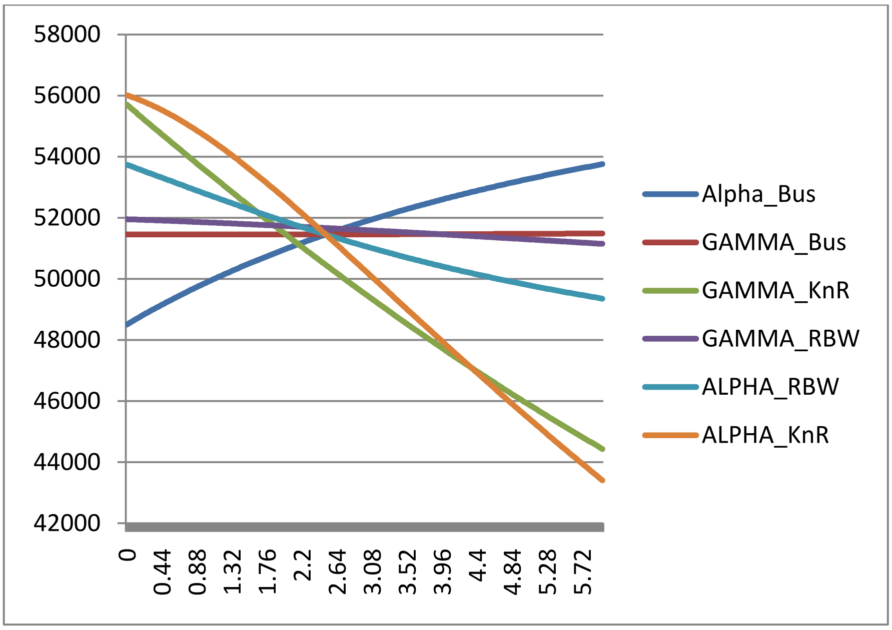


Table 4 Linear Regression between PT shares and Parameters

Regression Statistics	
Multiple R	0.975439983
R Square	0.95148316
Adjusted R Square	0.951369651
Standard Error	0.009212553
Observations	3000

	Coefficients
Intercept	0.320115166
ALPHA_KnR	-0.025739749
ALPHA_RBW	-0.005321281
ALPHA_Bus	-0.003690901
GAMMA_Bus	-0.000270005
GAMMA_KnR	-0.004514769
GAMMA_RBW	-0.000867965

The relationship between rail share and individual parameters are displayed in Fig 6 and 7. The rail mode is one level below the PT mode in the NLT structure. The bus mode competes with the demand for rail mode. The lower values of Alpha and Gamma of Bus mode reduce the rail mode share, as more people shift from rail to bus. But the shift is very small.

Figure 6 Impact of individual parameters on Rail Share: the x-axis is the value of the parameters, and the y-axis is the rail mode share.

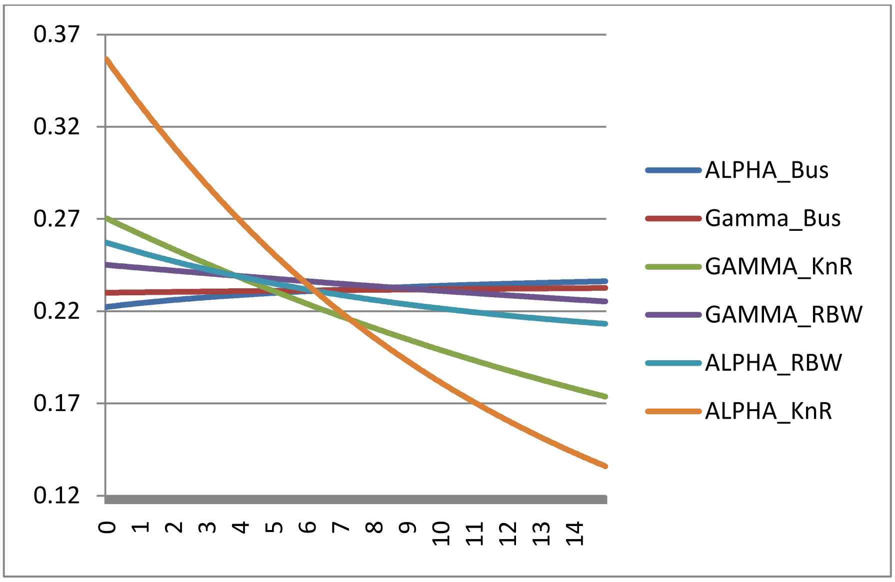
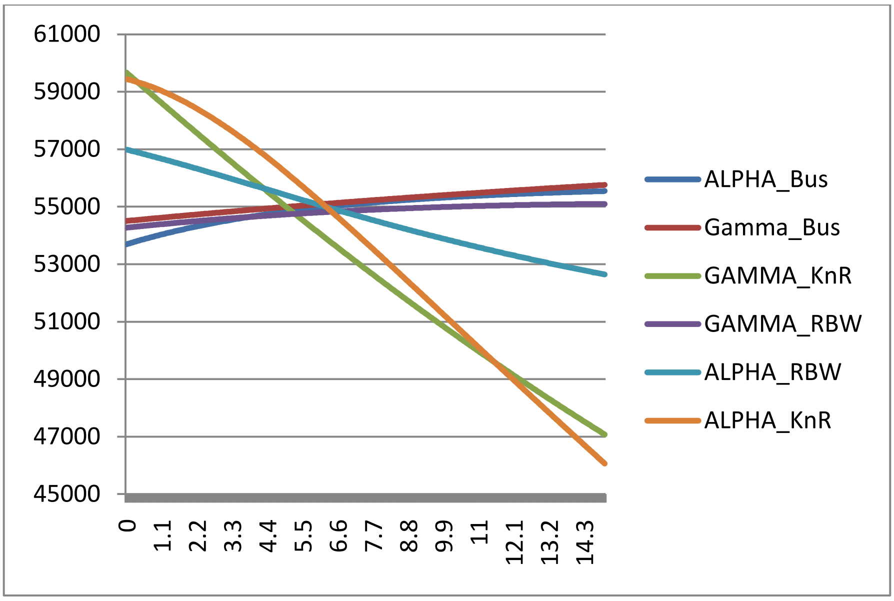


Figure 7 Impact of individual parameters on Rail Demand: the x-axis is the value of the parameters, and the y-axis is the number of rail trips.



5. Conclusion:

This paper demonstrates a case study of quantitatively measuring the risk (or uncertainty) associated with a mesoscope model of public transport project. The experiments clearly demonstrate: first, the model share does not follow a normal distribution, and the assumption of the normal distribution may help to understand the risk; secondly, the relationship between individual parameters and mode share is always non-linear or close to linear; thirdly, the influence of parameters on the mode share are not the same. Some parameters have far greater influence than others. For example, the influence of Alpha of Kiss and Ride is 95 times of the influence of the Gamma of Bus only. Instead of examining all the parameters, transport modellers could focus on the most important parameters.

The major obstacle of carrying out the Monte Carlo simulation is the runtime. It requires modellers to have advanced computing skills and high performance computers. This paper demonstrates the benefit of the simulation, and measures the downside risk for public transport projects to avoid large scale failure.

References:

Bliemer, M. C. J., *J. M. Rose* and S. Hess (2008). "Approximation of Bayesian Efficiency in Experimental Choice Designs." Journal of Choice Modelling, Vol 1, No 1 **1**(1).
BTS (2012). Household Travel Survey 2010/11 Technical Document T. f. N. Bureau of Transport Statistics, Australia.
David, A. (2011). NWRL Model, Technical Specification, Sinclair Knight Merz.
James, F., D. Andrew, P. Bhanu and M. Frank (2011). Extending the Sydney Strategic Model to represent toll road and park-and-ride choices Australasian Transport Research Forum 2011. Adelaide, Australia.
Martin, W. (2009). Political Aspects of Forecasting: Explaining and Controlling “Optimism Bias” in Transportation Forecasts. USDOT Workshop for Transportation Forecasters. U.S. Department of Transportation.
Taleb, N. N. (2010). The black swan : the impact of the highly improbable. New York, Random House Trade Paperbacks.
Train, K. (2009). Discrete choice methods with simulation. Cambridge ; New York, Cambridge University Press.