ESTIMATION OF DISCRETE CHOICE MODELS FOR ROUTE CHOICE FROM DATA COLLECTED USING THE RAPP-UP SMARTPHONE APPLICATION

G HAYES* and C VENTER**

Department of Civil Engineering, Centre for Transport Development, University of Pretoria;
*Tel: 063 618 1855; Email: u16401868@tuks.co.za; **Email: Christo.Venter@up.ac.za

ABSTRACT

A smartphone-based route choice stated and revealed preference application called RAPP-UP was used to collect route choice preference data from a small sample of Gauteng car commuters for their trip to work to demonstrate the application's proof of concept. The application presented choice sets with two route alternatives to survey participants on their smartphone using real-time values of travel time and cost. The route choice preference data was used to estimate discrete choice models to simulate route choice in congested urban conditions as well as willingness to pay (WTP) measures such as the non-work related value of travel time (VTT). The utility attributes included travel time (disaggregated into free-flow, slowed-down and stop-start travel time); trip petrol cost; trip toll cost (if any); and the probability of arriving at the trip destination on time. This paper presents the results of the discrete choice models that were estimated based on this data set. Confirming international evidence, the models showed that the sample of commuters perceived statistically significant differences between the attribute coefficients of the three trip time categories and their associated values of travel time. The trip petrol cost, toll cost and on-time arrival attribute coefficients were also significant. When a Toll Road Quality Bonus (TRQB) was included in the utility equation as a dummy attribute the coefficient was statistically significant but had a negative sign. The sample of survey participants thus perceived the use of Gauteng tolled freeways negatively for the unobserved factors of utility. The route choice models demonstrated the proof of concept of the RAPP-UP application as a convenient, practical, low-cost tool for collecting route choice preference data on congested urban road networks and provides the basis for its application with a larger sample of car commuters.

1. INTRODUCTION

The collection of traveller information is rapidly changing from traditional interview-based approaches such as computer-aided personal interviews (CAPI) to those based on smartphone applications. These applications (apps) are now widely used to collect transport related data. Most are passive data collection tools, i.e., they do not require any intervention by the participant other than to download the app onto their smartphone and activate it during their trip. While many of these apps are developed on a bespoke basis by researchers, platforms such as Itinerum provide survey templates that allow customisation for specific survey types that collect data passively and actively (Patterson, Fitzsimmons, Jackson & Mukai, 2019). Flocktracker is another smartphone-based trip survey instrument developed by the Massachusetts Institute of Technology (MIT, 2021). These apps can be used in place of traditional household travel surveys, travel diary surveys, trip satisfaction surveys and origin-destination surveys. The introduction of smartphones with reliable, embedded GPS devices has dramatically improved the ease with which to collect revealed preference (RP) survey data, especially because survey-specific applications can be developed and used on these devices (Shen & Stopher, 2014; Wang, He & Leung, 2018). The use of smartphone applications for the collection of stated preference (SP) data is, however, limited. RAPP-UP (Route Choice Application - University of Pretoria) fills this gap for conducting route choice based stated preference surveys.

RAPP-UP was developed to undertake vehicular route choice preference surveys using both SP and RP methods. The details of the RAPP-UP application and the methodology it employs are described by Hayes and Venter (2022). RAPP-UP generates route alternatives for a survey participant between their specified origin and destination based on real-time travel data derived from commercially available TomTom® traffic data sets. These traffic data sets are derived from cellphone tracking data are reported according to ISO intelligent transport systems standards (International Organization for Standardization (ISO), 2020). Interrogation of the TomTom® traffic data sets enabled the derivation of more disaggregated route characteristics, for example the classification and quantification of travel time into free-flow, slowed-down and stop-start travel time. The ISO intelligent transport systems specification for traffic flow level (tec001:EffectCode) is shown in Table 1 together with the RAPP-UP travel time classifications.

Table 1: Traffic Effects Code (TEC001) for Traffic Flow Descriptions and Travel Time Classifications

TEC Code	TISA English "Word"	Comment	Travel Time Classification
1	Traffic flow unknown	Shall be used if traffic flow is unknown. Note: This is often the case for local hazard warnings	Slowed-down time (sdt)
2	Free flow traffic	Traffic flow is not restricted	Free-flow time (fft)
3	Heavy traffic	Traffic flow is restricted due to a large number of vehicles	Stop-start time (sst)
4	Slow traffic	Traffic is slower than normal	Slowed-down time (sdt)
5	Queueing traffic	Traffic is in queues, but is still moving slowly	Slowed-down time (sdt)
6	Stationary traffic	Traffic is stationary or barely moving	Stop-start time (sst)
7	No flow	Traffic is completely stopped or there is no flow due to the road being closed / blocked; the cause-component may give more information about the reason for "no traffic flow". For roads with at-grade junctions, how the closure/blockage affects cross-road traffic maybe further specified with the attribute at *GradeJunctionClosure*	Stop-start time (sst)

Measures of other route characteristics, such as the average fuel consumption for a route and the probability of arriving at the selected destination on time, can also be derived.

This paper presents the outcome of the deployment of RAPP-UP with a small sample of Gauteng car users (48 participants) for their trip to work during the weekday morning peak hour between June and August 2022. The objective of the survey was to illustrate the proof of concept for RAPP-UP, i.e., that the application could be successfully deployed on a sample of Gauteng commuters and based on their route preference data, various forms of discrete choice models could be estimated to simulate route choice on congested urban road networks in Gauteng.

2. FORM OF UTILITY EXPRESSION IN RAPP-UP

A fundamental requirement of the development of RAPP-UP was the definition of trip utility. Trip utility for route choice experiments has most commonly been defined as a weighted linear combination of trip time, trip cost and trip time reliability attributes (Hensher, Ho, & Liu, 2016; Ortuzar & Willumsen, 2011). The form of expression that was adopted in RAPP-UP is as follows for individual *i* using route alternative *j* out of *J* alternative routes with attributes *k*:

$$U_{ij} = V_{ijk} + \varepsilon_{ij} = \sum_{j=1}^{J} \beta_k X_{ijk} + \varepsilon_{jj} \qquad (1)$$

where U_{ij} is the utility of individual *i* using route *j*; β_k are the attribute coefficients to be estimated; X_{ijk} are the observed utility attributes for route *j* and attribute *k*; and ε_{ij} is the utility error term for individual *i* using route *j*. The observed utility V_{ijk} is also known as the deterministic or representative component of utility.

In urban areas, trip time has been disaggregated into free-flow, slowed-down and stop-start time, as demonstrated by Wardman (1986; 2016; 2012) and Hensher (2005; 2004; 2016). They showed that motorists' perceptions of the value of travel time (VTT) varies between these categories of travel time. International evidence (Hensher & Rose, 2005; Wardman & Ibanez, 2012) showed that the value of travel time associated with stop-start travel conditions is higher than that for slowed-down and free-flow conditions. Wardman and Ibanez (2012) refer to the ratios of attribute coefficients for free-flow, slowed-down and stop-start-time as congestion multipliers. This form of the utility expression used in RAPP-UP filled two gaps in the understanding of motorists' route choice behaviour. Firstly, the congestion multiplier ratios have never been quantified in Gauteng (or South African) urban conditions. Secondly, the VTTs derived from the ratios of the travel time categories and the petrol and toll cost attribute coefficients provide insights into the sample of motorists' willingness to pay for travel time savings in the urbanised areas of Gauteng Province. The last SP experiment to collect route preference data in the urban road context in South Africa was by van Zyl and Raza in 2004 (van Zyl & Raza, 2006).

In mode and route choice models the perceived (and observable) monetary out-of-pocket trip cost incurred by an individual using their private vehicle are the petrol cost and the toll and parking costs (if any). International evidence has shown that trip time reliability can also be a contributor to trip utility (Brownstone & Small, 2005; Carrion & Levinson, 2012). The attributes included in the RAPP-UP utility expression in this study included the three travel time categories related to the levels of service, i.e., the time spent in free-flow (fft), the slowed-down time (sdt) and the time spent in stop-start conditions (sst). Two trip cost attributes were included, i.e., tolls costs on the Gauteng Freeway Improvement Project (GFIP) freeways and other freeway tolls, and the trip petrol cost. Route travel time reliability was included and expressed as the probability of arriving at the destination on time. The estimated time of arrival was provided by the TomTom® route generation algorithm. The travel time reliability attribute level for a route was derived from the proportions of the stop-start, slowed down and free-flow times on that route (Hayes & Venter, 2022) and expressed as a percentage. The utility expression was thus defined as the linear-in-parameters sum of six attributes, viz., the trip time, trip cost and route reliability attributes, for individual *i* using route *j* as follows:

$$U_{ij} = b_1{*}fft_{ij} + b_2{*}sdt_{ij} + b_3{*}sst_{ij} + b_4{*}petcost_{ij} + b_5{*}tollcost_{ij} + b_6{*}pota_{ij} + \varepsilon_{ij} \qquad (2)$$

where U_{ij} is the trip utility for individual *i* using route *j*; b_1, b_2, … are attribute coefficients to be estimated; fft_{ij} is the free flow time on route *j* in minutes; sdt_{ij} is the slowed-down time on route *j* in minutes; sst_{ij} is the stop-start time on route *j* in minutes; $petcost_{ijk}$ is the petrol cost for route *j* in Rands; $tollcost_{ij}$ is the toll cost of route *j* in Rands (if applicable); $pota_{ij}$ is the probability of on-time arrival using route *j* in percent, and ε_{ij} is the error term for individual *i* using route *j*. The petrol cost of the route was calculated based on the route length, an average vehicle fuel consumption of 8 litres per 100 kms (International Energy Agency, 2021) and the price of petrol when the surveys were undertaken (R23.50 per litre in Gauteng for 95 Octane petrol). Note that the experiment was unlabelled, and an Alternative Specific Constant (ASC) was therefore not included the utility expression.

While real-time data were used for the various travel time and petrol cost attribute levels, a degree of analyst control over the attribute levels was implemented by introducing attribute factors for the toll cost and trip time reliability factors. This was implemented by factoring the actual levels of toll cost and trip time reliability by either one of two values, i.e., 0.9 or 1.1. This range was considered practical as it provided a realistic range of toll cost attribute levels that was confirmed in the pilot survey stage of testing RAPP-UP. The factors were applied to the attribute levels in the choice sets according to a fractional factorial experimental design that required the survey participants to undertake eight surveys, i.e., eight separate trips over a period of time between origin and destination (O-D). The survey days did not have to be consecutive, nor did the O-Ds need to be the same for each trip.

An economic experiment was also introduced to reduce the no-consequence drawback of SP experiments by means of a survey toll account. At the start of the experiment each participant was given a survey toll account with a positive balance of R300.00. Choosing a tolled route reduced the survey toll account balance by the toll fee amount. The balance of the survey toll account was paid out to participants at the end of the survey.

3. RAPP-UP SURVEY PARTICIPANT AND TRIP DATA DESCRIPTION

To demonstrate the proof of concept for RAPP-UP, a small sample of 48 participants were recruited from the consumer panel of a market research company. The participants were required to meet four survey participation criteria, i.e., ownership of a recent model Android smartphone; ownership of their own vehicle; making regular commute trips to work using their vehicle in the weekday morning peak period; and commute to work in the Johannesburg / Pretoria and Ekurhuleni urban regions of the Gauteng Province. Each participant was required to complete a short socio-demographic and trip-related questionnaire. The key characteristics of the survey participants are shown in Figure 1. While there was heterogeneity in the characteristics, most participants were in the 31 to 40-year age band; had gross monthly incomes of between R30,000 and R40,000; departed for work between 07:00 and 07:30; and perceived their commuting trip time to be between 30 and 40 minutes.

It was recognised that the survey sample was biased toward high income earners and did not reflect the broader car commuting public in Gauteng or any other urban area in South Africa. The purpose of the survey was primarily to demonstrate the proof of concept for RAPP-UP. The route choice model outputs should only be interpreted in this context.

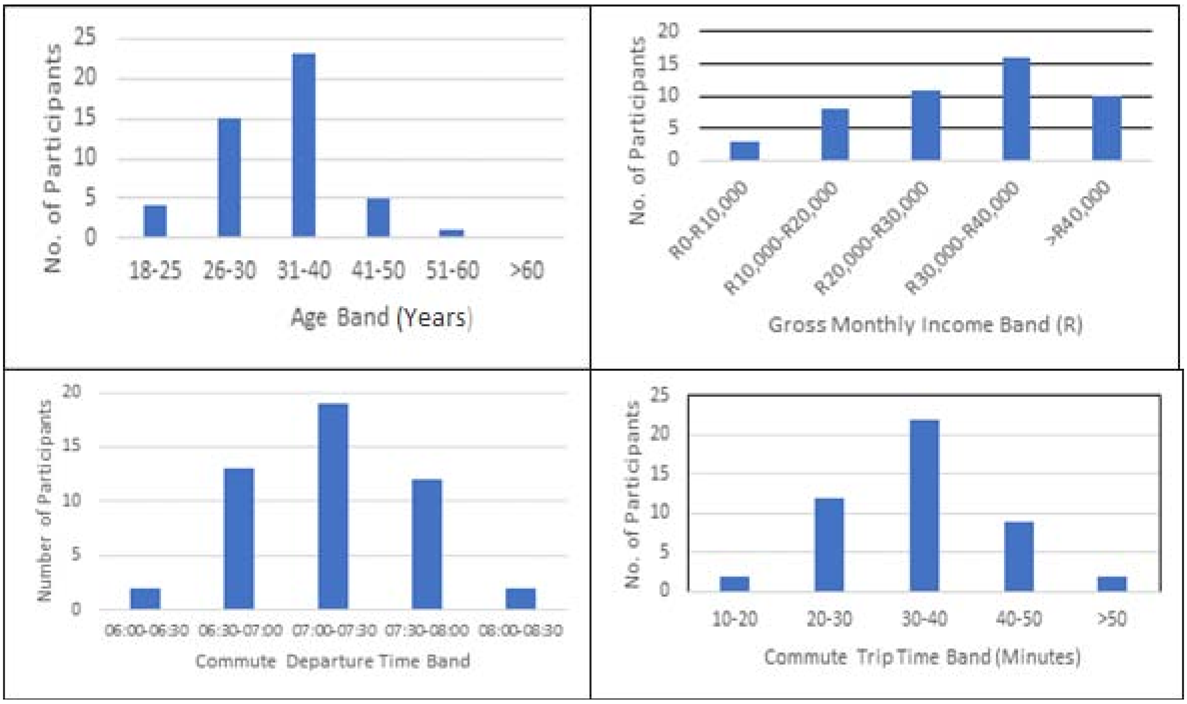

Figure 1: Key RAPP-UP survey participant characteristics

The survey participants were required to drive their preferred route after selection in the SP component of the survey. It was found that this self-validating RP mechanism was well adhered to, with little variation between chosen route and driven route. Over 90% of the chosen routes were adhered to.

The participants were required to complete eight trips, and while most completed this number, five participants completed only four trips. Thus, a total of 364 observations (choice sets) were used to estimate the multinomial logit (MNL) and C-Logit discrete route choice models, for a total of 728 routes. For the panel based random parameters models (RPL) 344 observations were used i.e., the participants who completed all eight trips.

Table 2 summarises the characteristics of the route alternatives generated by RAPP-UP for each utility attribute. The range of attribute levels that contains 90% of the observed values are shown together with the average value. The characteristics were generated using all 728 observations. The average route petrol cost is five times the toll cost when only tolled routes are considered. This is due to the toll cost being relatively low for the tolled routes.

Table 2: Description of route attributes generated by RAPP-UP

Route Attribute	90% Observed Range	Average Value
Total travel time*	11 min – 53 min	31.5 min
Free-flow travel time	0 min – 46 min	27.6 min
Slowed-down travel time	0 min – 12 min	2.6 min
Stop-start travel time	0 min – 25 min	1.1 min
Petrol cost	R5.80 – R45.60	R24.03
Toll cost	R0.00 – R12.71	R3.43
Route length*	5.0 km – 41.0 km	21.4 km
Probability of On Time Arrival	0.64 – 0.99	0.86
*Attributes not included in utility expression		

From Table 2, it is noticeable that the slowed-down and stop-start travel times make up relatively low proportions of the overall trip time. This effect is illustrated in Figure 2 that shows the trip time distributions for the total trip time and the free-flow, slowed-down and stop-start times. The slowed-down and stop-stary trip time distributions are skewed toward lower travel times.

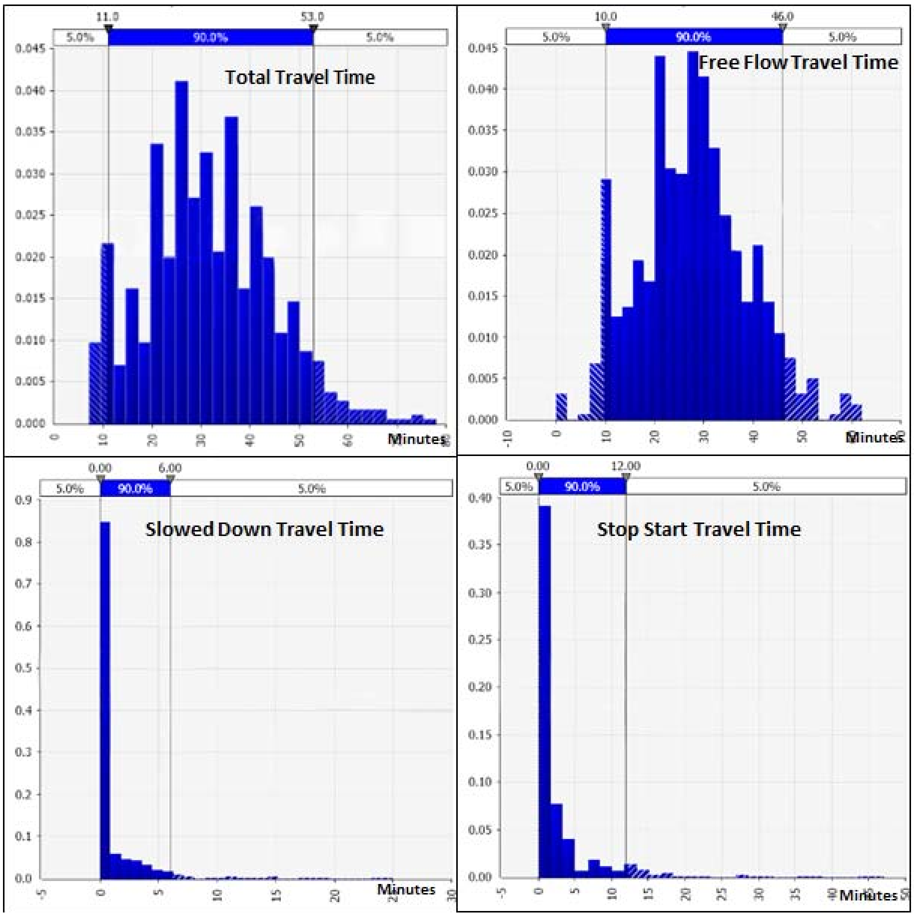

Figure 2: Distributions of total, free-flow, slowed-down and stop-start travel time

4. ROUTE CHOICE MODEL ESTIMATION

Three forms of discrete choice models were estimated i.e., the Multinomial Logit (MNL) model; the Random Parameters Logit (RPL) model (also known as the mixed logit model); and the Commonality Logit (or C-Logit) model. The C-Logit model specifically accounts for route correlation due to the overlap of two route alternatives. The model accounts for the correlation by adjusting the utility expression with a route commonality factor. Various forms of this factor have been suggested, but all are quantified by the proportion of common distance or time overlap for each alternative. The route commonality factor is included in the utility expression as an additional attribute with its own coefficient and is used to proportionally adjust the representative utility of a route and hence account for correlation between routes (Prato, Rasmussen & Nielsen, 2014). When the route alternatives do not overlap, the commonality factor is zero and the C-Logit model collapses to the MNL form. Most routes generated by RAPP-UP had no overlap and the C-Logit model commonality attribute coefficient was not significant, and they are hence not reported in this paper.

Various forms of utility expressions were tested using MNL and RPL models, starting with the base utility expression previously presented as Equation 2. All the models were estimated using the NLOGIT software platform.

4.1 Multinomial Logit Models (MNL)

The MNL base model output is shown in Table 3. All the utility attributes' coefficients are significant at the 99% confidence level, and all have the expected sign (i.e., positive for a utility; negative for a disutility). The Pota coefficient sign is positive, as this attribute reduces disutility as the probability of on-time arrival increases. The McFadden R^2 value of 0.51 is high for a non-linear model, corresponding to a linear regression R^2 value of over 0.80.

Table 3: Base MNL model outputs

Attribute	Coefficient	Std. Error	t-ratio	P-value	Lower CIL	Upper CIL
Free-flow time	-0.177***	0.036	-4.90	0.000	-0.247	-0.106
Slow-down time	-0.218***	0.055	-3.94	0.000	-0.326	-0.109
Stop-start time	-0.397***	0.090	-4.42	0.000	-0.574	-0.221
Petrol cost	-0.394***	0.068	-5.79	0.000	-0.527	-0.260
Toll cost	-0.288***	0.079	-3.66	0.000	-0.443	-0.134
Pota	+2.779***	1.072	2.59	0.010	0.678	4.880
Log-Likelihood	-122.07					
Sample size	364					
McFadden R^2	0.51					
***, **, * ==> Significance at 1%, 5%, 10% level. CIL = 95% confidence interval limit						

The key observations from the model are:

- The free-flow coefficient is lower in magnitude than the slowed-down and stop-start travel time values. This means that the survey participants associate less disutility with travel in free-flow time. The ratios of the slowed-down and stop-start coefficients with the free-flow value are termed congestion multipliers and are discussed below.
- The petrol cost coefficient is larger in magnitude than the toll cost value by a factor of 1.37. This means that for every R1.00 spent on either of these attributes for a particular route, the petrol cost contributes more to disutility by a factor of 1.37.
- The absolute value of the probability of on-time arrival (Pota) coefficient is substantially larger than the time and cost coefficients. However, there is a scale effect with this attribute. The Pota value is measured as a decimal value less than 1. To illustrate this effect, if a route has a Pota value of 0.8 and a petrol cost of R30, then the Pota contribution to utility is +2.22 and the petrol cost disutility contribution is -11.82.
- The base MNL can correctly replicate the route choices made by the participants 89% of the time and can hence be considered to be a reliable predictor of route choice for the sample of survey participants.

Table 4 presents a summary of the congestion multipliers determined by Wardman and Ibanez (2012) for various cities and countries. In the table congested conditions are defined as those where *"vehicle speed is noticeably restricted with frequent gear changes* *required"*. The congestion multipliers are the ratios of the congested: uncongested attribute coefficients and the stop-start: free flow and slowed-down: free-flow coefficients. The values of the RAPP-UP MNL model for Gauteng are included in the table. The sample of Gauteng commuters have a slowed-down ratio value lower than the international values, while the stop-start value is significantly higher than the international values. This result implies that the study's sample of survey participants associated higher levels of disutility with stop-start travel conditions than those found internationally.

Table 4: Summary of congestion multipliers for commuting trips (ratio of congested and free-flow coefficients)

City / Country	Year	Travel Time Type	Congestion Multiplier
UK	1986 - 2008	Congested	1.59
Canada	1993	Congested	1.37
New Zealand	2001	Slowed down	1.33
Copenhagen	2002	Congested	1.31
Brisbane	2005	Stop-start	1.34
Copenhagen	2007	Congested	1.15
Sydney	2008	Slowed down	1.35
Montreal	2005	Stop-start	1.65
Riga	2005	Stop-start	1.00
Serbia	2007	Stop-start	1.73
Brisbane	2007	Stop-start	1.19
Sydney	2010	Slowed down	1.68
Singapore	2009	Light & heavy congestion	1.16
South Africa (Gauteng Province)	2022	Slowed down	1.23
South Africa (Gauteng Province)	2022	Stop-start	2.24

The associated values of travel time derived from the ratio of the time and cost coefficients for the MNL are shown in Table 5.

Table 5: Commuting Values of Travel Time (VTT) from base MNL Model (rands/hour)

VTT (Rands/Hour)	VTT Free-Flow Time (Rand/hour)	VTT Slowed Down Time (Rand/Hour)	VTT Stop-Start time (Rand/hour)
VTT (Petrol)	R26.93	R33.17	R60.53
VTT (Toll)	R36.78	R45.30	R82.67

The VTT derived from the free-flow time coefficient is relatively low in the context of the high participant income levels, but is intuitively correct, i.e., commuters would not be willing to pay much for travel time savings when the trip is made in free-flow conditions and time savings are small. However, the VTT for the slowed-down and stop-start portions of a trip increase to R33.17 and R60.53/hour respectively. No recently derived South African private vehicle commuter VTT values are available for benchmarking purposes. However, for high income public transport users in Gauteng, i.e., Gautrain users, the VTT values are approximately R40/hour for the station access (i.e., first mile) of their Gautrain trip and R80/hour for the station egress part (i.e., last mile) (Watts, 2020).

The willingness to pay a toll for travel time savings is higher than that for petrol costs, but still relatively low for the high-income survey participants. This result is somewhat unexpected, as there has been high public resistance to the Gauteng GFIP freeway e-tolls. However, the Gauteng public's unwillingness to pay tolls has been affected by other factors including a successful civil disobedience campaign. This result should also be viewed in the context of the relative magnitude of average toll costs compared to petrol costs for a trip using a tolled section of freeway. The average petrol: toll cost ratio was five when considering tolled routes only, meaning that toll fees make up a significantly lower proportion of total trip disutility than petrol costs. Nevertheless, the result indicates that the sample of survey participants would be willing to pay a toll if they were offered route travel time savings on freeways.

A Toll Road Quality Bonus (TQRB) was introduced into the utility expression as a dummy attribute. The TQRB assumed a value of zero when there was no GFIP freeway component in a route and a value of one when a GFIP freeway component was included. The purpose of the TRQB was to determine whether there were any unobserved factors of utility associated with GFIP freeways. The related MNL model results are presented in Table 6. As with the base MNL, the coefficient signs are as expected and are all significant. The toll cost, Pota and TQRB coefficients are significant at the 95% level. The loglikelihood is -119.5, which is an improvement over the base model value of -122.1; and the McFadden R^2 is also slightly improved. The same congestion multiplier effect is reflected in the travel time coefficients. The introduction of the TQRB, therefore, improves the overall significance of the model. The TQRB has a negative sign, which means that the survey participants had a negative perception of the GFIP freeways when the unobserved factors of utility were considered. This negative perception likely reflects the adverse view of the GFIP scheme amongst the Gauteng public that has recently led to the scrapping of the e-tolls.

Table 6: Base MNL Model with Toll Road Quality Bonus Attribute (TQRB)

Variable	Coefficient	Std. Error	t-ratio	P-value	Lower CIL	Upper CIL
Free-flow time	-0.193***	0.037	-5.19	0.000	-0.265	-0.120
Slow-down time	-0.226***	0.055	-4.12	0.000	-0.334	-0.119
Stop-start time	-0.374***	0.090	-4.16	0.000	-0.551	-0.198
Petrol cost	-0.360***	0.068	-5.30	0.000	-0.493	-0.227
Toll cost	-0.196**	0.084	-2.32	0.020	-0.361	-0.030
Pota	+2.800**	1.091	2.57	0.010	0.661	4.939
TRQB	-0.786**	0.351	-2.24	0.025	-1.474	-0.098
Log-Likelihood	-119.5					
Sample size	364					
McFadden R^2	0.52					
***, **, * ==> Significance at 1%, 5%, 10% level. CIL = 95% confidence interval limit						

4.2 Random Parameters Logit (RPL) Models

The RPL model (also known as the mixed logit model) was considered the “model for the new millennium” by Ortuzar and Willumsen (2011). This is because the RPL is a highly flexible model that can approximate any random utility model (McFadden & Train, 2000). It overcomes the three limitations of the closed form of MNLs by firstly allowing for random taste variation, thereby relaxing the independent and identical distribution (IID) constraint; secondly it allows unrestricted substitution patterns and relaxes the independence from irrelevant alternatives (IIA) (i.e., correlation between alternatives) constraint; and thirdly it considers the correlation of unobserved factors over time, i.e., panel effects that take into account the variation in choice by individual decision makers making repeated choices. Simulation is required for unique RPL model identification. To enable this, the analyst is required to define distribution functions (i.e., mixing distributions) for each of the non-random attribute coefficients in the utility expression. RPL models yield estimates of the first and second moments (i.e., the mean and standard deviation) of the distribution of preferences cross the population of interest, and these are reported in the RPL model outputs. The identification of RPL models also requires careful consideration of the ratios of the time and cost coefficients to estimate VTT, i.e., the ratio of two distributions. Fixing the cost attribute coefficient is an approach to overcome this limitation (Train, 2009). Alternatively, the distribution defined for the cost attribute can be defined such that only positive VTT values are derived, for example using one-sided triangular or log-normal distributions.

The outputs of this study’s base model with the Toll Road Quality Bonus (TQRB) included in the utility expression are shown in Table 7. The first model has all coefficients randomised with normal distributions and the second model has the cost coefficients fixed. 1,000 Halton draws were used in the simulation for both models. The utility expression for these models is as shown in Equation 2, but with the TRQB included.

Table 7: Base RPL Model outputs (all coefficients randomised with normal distributions)

Attribute	Base RPL Model with All Coefficients Randomised with Normal Distributions			Base RPL Model with Cost Coefficients Non-Randomised		
	Coefficient	Std Error	t-Ratio	Coefficient	Std Error	t-Ratio
Free-flow time	-0.249***	0.054	-4.62	-0.251***	0.052	-4.85
Slowed-down time	-0.361***	0.115	-3.13	-0.363***	0.112	-3.25
Stop-start time	-0.633***	0.162	-3.91	-0.590***	0.137	-4.30
Petrol cost	-0.481***	0.144	-3.33	-0.420***	0.091	-4.60
Toll cost	-0.224*	0.129	-1.74	-0.219*	0.121	-1.77
Pota	+2.463	1.600	1.57	+2.355	1.511	1.56
TRQB	-1.153**	0.505	-2.28	-1.189**	0.492	-2.42

Table 7: Cont'd

Standard Deviations						
Free-flow time	0.009	0.061	0.14	0.003	0.066	0.04
Slowed-down time	0.294**	0.105	2.80	0.290***	0.105	2.75
Stop-start time	0.007	0.195	0.04	0.012	0.267	0.05
Petrol cost	0.148	0.144	1.03	-	-	-
Toll cost	0.019	0.225	0.09	-	-	-
Pota	4.502*	2.496	1.80	4.213	2.513	1.68
TRQB	0.014	0.512	0.03	0.003	0.516	0.01
Log-Likelihood	-106.9			-101.8		
McFadden Pseudo R^2	0.56			0.56		

For the fully randomised model the three travel time coefficients and the petrol cost coefficient are significant at the 99% level, the toll cost at the 90% level and the TQRB at the 95% level. Pota is significant at the 85% level. All coefficients have the expected sign. For the non-randomised cost model, the same coefficient significance patterns were found.

Other important observations are:

- The McFadden R^2 value in both models of 0.56 is highly significant, showing a good fit between the observed and modelled choices. The log-likelihood of the non-random cost coefficients model (-101.8) is a significant improvement over the fully randomised model (-106.9), indicating a better fit between observed and modelled choices. (The R^2 and log-likelihood values should not be compared to the equivalent MNL model as the sample sizes are not consistent).
- While there is some variation around the mean for the petrol coefficient (standard deviation is 0.148) in the fully randomised model, it is not significant (t-ratio is 1.03). There is no variation in the toll coefficient (standard deviation is 0.019), but it is also not significant (t-ratio is 0.09). These low and insignificant variations around the mean coefficient values for these attributes suggest that the VTT values do not vary significantly between participants.
- The Pota coefficient has a significant standard deviation and hence there is heterogeneity associated with this coefficient. However, the mean Pota coefficient value is not significant.
- The low coefficient standard deviations indicate that it is not necessary to use mixing distributions such as the log-normal to ensure positive willingness-to-pay measures.

The values of travel time savings derived from the petrol and toll cost numeraires for each model are summarised in Table 8. The RPL model VTTs are higher than those of the base MNL model, especially the willingness to pay a toll for travel time savings that are nearly double those estimated with the MNL model. Several authors have shown that RPL model estimate higher willingness to pay measures than MNL models. Amador et al. (2005) indicate that it is common for RPL models to estimate higher VTTs than the more restrictive MNL models. Bhat (1998), and Train (2009) found that willingness to pay measures for all attributes are higher with RPL models than with MNL models and suggest that MNL models underestimate the willingness to pay.

Table 8: Values of Travel Time (VTT) from Base RPL Model with TRQB (Rands/Hour)

VTT (Rands/Hour)	**Free-Flow VTT (Rand/hour)**	**Slowed-Down VTT (Rand/Hour)**	**Stop-Start VTT (Rand/hour)**
RPL Model Fully Randomised			
VTT (Petrol)	R31.06	R45.03	R78.96
VTT (Toll)	R66.70	R96.70	R169.55
RPL with Cost Coefficients Fixed			
VTT (Petrol)	R35.86	R51.86	R84.29
VTT (Toll)	R68.76	R99.45	R161.64

5. CONCLUSIONS

The Route Choice Application - University of Pretoria (RAPP-UP) was successfully developed and is now a stable platform for the execution of large-scale route choice surveys. The application was designed to be flexible and can hence be easily modified to accommodate a desired number of utility attributes. The attribute traffic data can either be derived directly from platforms such as TomTom® in real time or can be externally determined and input into the application by means of look-up tables. Lastly it is possible to factor the attribute levels to put more experimental control in the hands of the analyst.

Feedback from the survey participants was that the application was reliable, robust, user-friendly and easy to understand. The collection and collation of choice preference data during and after the survey from the cloud-based database was efficient and effective. The data format in the database was designed to allow for easy formatting for input into the discrete choice modelling platform NLOGIT. Big data sets from large samples can easily be accommodated.

The MNL and RPL models that were estimated from the small sample of Gauteng commuters were statistically significant. However, the C-Logit models defaulted to the MNL model form due to the low levels of route overlap in the choice sets. The MNL and RPL models confirmed that for the sample of participants in the survey there was a significant difference in perception of utility between free-flow, slowed-down and stop-start travel time. The congestion multipliers determined from the models were in line with international findings, except that the sample of Gauteng commuters valued the disbenefit of stop-start driving conditions higher than that found in international studies. The petrol cost coefficient was consistently higher than the toll cost value, but this was influenced by the route petrol costs being significantly higher than the toll cost by a factor of five in the choice sets for tolled routes.

The inclusion of a Toll Road Quality Bonus (TRQB) improved the overall significance of the models. However, the TQRB coefficients were consistently negative, indicating a negative perception of the GFIP toll scheme when non-observed factors of utility are considered. This is likely related to the adverse perception of the GFIP system held by Gauteng motorists.

The last definitive values of travel time determined for car commuters in South African urban areas were derived in 2004, making it difficult to compare and benchmark the values derived from the MNL and RPL models. However, when benchmarked against recent high-income Gautrain user VTTs for the first and last mile trip, the free-flow value was less, the slowed-down value similar, and the stop-start value higher.

While the models were estimated from a small sample of Gauteng commuters, they nevertheless point to important implications for urban road planning and economic appraisal in urban areas in South Africa. Firstly, car commuters are willing to pay tolls for travel time savings on urban freeways, but there is variation in their willingness to pay depending on the congestion levels they experience. The ability of urban toll schemes to cater for variable toll tariffs depending on congestion levels seems to be a fundamental requirement for a successful scheme. Secondly, there appears to be a utility benefit for trip time reliability. The willingness to pay for this reliability requires further investigation, keeping in mind that trip time reliability is affected by freeway and non-freeway roads.

The deployment of the RAPP-UP application with a significantly larger and representative sample of commuters is required to further explore the route preferences and choice behaviour in urban areas.

6. REFERENCES

Amador, F, Gonzalez, R & Ortuzar, J. 2005. Preference heterogeneity and willingness to pay for travel time savings. *Transportation*, 627-647.

Bhat, C. 1998. Accommodating variations in responsiveness to level-of-service variables in travel choice models. *Transportation Research A*, 455-507.

Brownstone, D & Small, K. 2005. Valuating time and reliability: Assessing the evidence from road pricing demomstration. *Transportation Research: Part A Policy and Practice*, 279-293.

Carrion, C & Levinson, D. 2012. Value of travel time reliability: A review of current evidence. *Transport Research Part A: Policy and Practice*, 720-741.

Hayes, G & Venter, C. 2022. An Innovative Method to Collect Route Choice Preference Data using a Smartphone Application. *40th Southern African Transport Conference.* Pretoria: Southern African Transport Conference.

Hensher, D & Rose, J. 2004. Determination of Value of Travel Time for Brisbane River Crossing. Sydney: Hensher & Rose.

Hensher, D & Rose, J. 2005. Behavioural Value of Travel Time Savings: Gateway Upgrade (GUP) Project for the Brisbane Gateway Consortium and the North-South Bypass Tunnel (NSBT) Project. Sydney: Hensher Group (Pty) Ltd.

Hensher, D, Ho, C & Liu, W. 2016. How much is too much for tolled road users: Toll saturation and the implications for car commuting value of travel time savings? *Transportation Research Part A*, 604-621.

International Energy Agency. 2021. Vehicle Fuel Economy in Major Markets 2005-2019 Working Paper 22. IEA.

International Organization for Standardization (ISO). (2020). *ISO 21219 -15:2020(E):* Intelligent transport systems - Traffic and travel information (TTI) via transport protocols experts group, generation 2: Part 15 Traffic event compact (TPEG2-TEC). Switzerland: International Organization for Standardization (ISO).

McFadden, D & Train, K. 2000. Mixed MNL models of discrete response. *Journal of Applied Econometrics*, 447-470.

MIT. (September 2021). Mobility Futures Collaboration. Available at: https://mfc.mit.edu/flocktrackerabgerufen

Ortuzar, JD & Willumsen, L. 2011. Modelling Transport 4th Edition. Chichester: John Wiley & Sons.

Patterson, Z, Fitzsimmons, K, Jackson, S & Mukai, T. 2019. Itinerum: The open smartphone travel survey platform. *SoftwareX*, 1-7.

Prato, C. 2009. Route choice modelling: past, present and future research directions. *Journal of Choice Modelling*, 65-100.

Prato, C, Rasmussen, T & Nielsen, O. 2014. Estimating Value of Congestion and of Reliability from Observation of Route Choice Behaviour of Car Drivers. *Transportation Research Record: Journal of the Transportation Research 9999.*

Shen, L & Stopher, P. 2014. Review of GPS Travel Survey and GPS-Data Processing Methods. *Transportation Reviews*, 316-334.

Train, K. 2009. Discrete Choice Models with Simulation. Cambridge: Cambridge University Press.

van Zyl, N & Raza, M. 2006. In Search of the Value of Time from South Africa to India. *Travel Survey Methods* (S. 457-484). Emerald Group Publications.

Wang, Z, He, S & Leung, Y. 2018. Applying mobile phone data to travel beaviour research: A literature review. *Travel Behaviour and Society*, 141-155.

Wardman, M. 1986. Route Choice and the Value of Motorists' Travel Time: Empirical Findings Working Paper 224. Leeds: Institute of Transport Studies, University of Leeds.

Wardman, M & Ibanez, N. 2012. The congestion multiplier: Variations in motorists' valuation of travel time with traffic conditions. *Transportation Research Part A*, 213-225.

Wardman, M, Chintakayala, V & de Jong, G. 2016. Values of travel time in Europe: Review and meta-analysis. *Transportation Research Part A*, 93-111.

Watts, D. 2020. Exploring the Access/Egress Behaviour of Rapid Rail Passengers through the Application of Nested Logit Models. Pretoria: University of Pretoria MSc Thesis.