An Efficient Solution to Toll Choices involving Multiple Toll Booths and/or Multiple Tolling Strategies

Keith. Long[1]
[1] Principal, Keith Long, Peregian Beach, QLD, Australia

1 Introduction

The recent opening of Sydney's Westlink M7 Motorway has crystallized deficiencies in traditional toll road demand forecasting approaches. The Westlink M7 distance-based, capped toll cannot be efficiently modelled using an assignment with the toll incorporated into the generalised cost, and traditional behavioural route choice models encounter convergence, run-time and potential market issues given the sheer number of toll segments to be evaluated (where a toll segment is the combination of toll booths available for a particular trip).

These deficiencies are exacerbated by the construction of more urban toll roads together with the possibility of more complex tolling strategies permitted by electronic tolling.

A *new* behavioural route choice model is proposed which provides an efficient solution to toll choices involving multiple toll booths and/or multiple tolling strategies (eg toll caps, toll discounts).

The solution is implemented in the commercial ***FASTOLL*** software package.

2 Current approaches to toll road demand forecasting

2.1 Introduction

There are two broad approaches to forecasting toll road demand:

- assignment with toll incorporated into generalised cost (toll delay penalty or TDP models); and
- behavioural route choice (BRC) models.

The explanatory variables and associated parameters (or weights) used as input to either approach are generally derived from stated preference / revealed preference (or SPRP) surveys.

2.2 Assignment with toll incorporated into generalised cost (TDP models)

The perceived out-of-pocket toll cost is incorporated into the generalised cost function which is used as the measure of impedance in the trip assignment process.

In generalised cost traffic modelling the most common method of representing tolls in a road network is to add a penalty to any link that includes a toll booth. If time alone is the basis for the generalised costs used in choosing a route, the *equivalent time penalty* can be calculated from the toll price using an average value of time adjusted for any toll road route bias. This penalty is than included in the generalised cost of the tolled route and the assignment algorithm is used to allocate trips between the tolled route and the untolled route. If there is more than one tolled route, penalties for each toll link may be included, reflecting the toll to be paid at each.

The equivalent time penalty which is added to any link that includes a toll booth is generally called the *toll delay penalty* (TDP), and is expressed in minutes for a particular toll level. A logical interpretation of the TDP is that it represents the point on a toll diversion curve (see Section 2.3.3) where 50% of users choose the tolled route and 50% of users choose the untolled route, where there is no difference in general utility terms.

TDP models are easy to implement, safe and reliable. However, there are two major weaknesses:

- there can be excessive sensitivity of demand responses as a tolled route changes from being the best (generalised time) route to the second best route, as toll prices are increased; and

- demand cannot be predicted under the more complex tolling strategies now permitted by electronic tolling such as:
    - distance-based tolls with toll caps;
    - unique entry-exit tolls; and
    - toll discounts.

This paper does not investigate TDP models further.

2.3 Behavioural route choice (BRC) models

Behavioural route choice (BRC) models deal with toll payments as a consumer choice and do not confound this choice with the overall route choice models used in assigning traffic to a network. That is, the payment of a toll is treated as a purchase of a range of road travel benefits.

2.3.1 Variables, parameters and utility functions

The choice of route for car drivers is influenced largely by variables associated with the particular trip (eg travel time, toll price, toll payment method, variability of travel time and perceptions of driving comfort offered by a high standard road), but can also be influenced by other variables (eg perceptions of total toll for a two-way trip and toll budgets over time).

Consider a simple road network, with a single toll class and a single toll booth. It is commonly, and reasonably, assumed that the utility for the tolled and untolled routes can be expressed as shown in Equation 1.

Equation 1

$$\begin{aligned} \text{Utility}_{\text{untolled}} \quad U_{\text{untolled}} &= \phantom{a0 +{}} a1 * \text{Time}_U + \phantom{a2 * \text{Toll}_T} + a3 * \text{Var3}_U + \ldots \\ \text{Utility}_{\text{tolled}} \quad U_{\text{tolled}} &= a0 + a1 * \text{Time}_T + a2 * \text{Toll}_T + a3 * \text{Var3}_T + \ldots \end{aligned}$$

Where:

$Time_U$ is the travel time for a trip via the untolled route;
$Time_T$ is the travel time for a trip via the tolled route;
$Toll_T$ is the total toll paid via the tolled route;
a1 and *a2* are parameters which reflect the influence of each of the respective variables above on route choice;
*'a3 * Var3 + …'* reflects the influence of other variables on a driver's route choice. These additional variables could apply to the untolled route only, the tolled route only or both routes. For example, a variable such as *toll payment type* would apply to the

tolled route only, whilst a variable such as *StopStart travel time* would apply to both the untolled and tolled routes; and

a0 accounts for the influence of those variables which are not included in the utility functions, and it can be considered to be a route constant for the particular tolled route being considered (and is also known as a toll bonus or ASC).

The parameters *a0..a3* in the above functions can be estimated through well established statistical procedures, which generally involve Stated Preference and Revealed Preference (SPRP) survey techniques. These techniques invite interviewees to respond to hypothetical choices between travelling on an untolled route versus travelling on tolled routes, with a particular aim of identifying the level of toll at which users would divert from the tolled route. By comparing the *switch* tolls with likely travel time savings from using the tolled route compared with the non-tolled route, an individual's *value of time* can be inferred for use in a toll choice model. The ratio of *a2/a1* gives an estimate of the *Value of Time* (VoT in min/$).

Thus, a key assumption of behavioural route choice modelling is that the willingness of potential toll road users to pay a toll is driven by a range of relative utilities, with the value of time being merely one component of route choice. These models assume that individuals will choose a route by considering the inherent utility they would derive from travelling via each route, and that the probability of selecting any one route depends on an assessment of the relative utilities of each route.

2.3.2 The toll choice model

For a simple urban road network with a single toll class (or market segment) and a single toll booth, and where the toll choice is limited to time and toll variables, a binary logit model would generally be adopted to predict choice between the tolled route and the untolled route. As the number of toll booths in an urban road network increases, leading to significant increases in the number of tolled routes (or toll segments), the relationship applied to predict choice between the untolled route and the tolled routes also becomes more complicated. Various model structures, including binary logit models, multi-nominal logit models and nested logit models are used in practice. This paper does not investigate the structure of the toll choice model further.

2.3.3 The toll diversion curve

The probability of choosing the tolled route can be plotted against the explanatory variables used in the utility function. For the above example, which specifies time (min) and toll ($) in the utility function, the probability of using the tolled route, at a particular toll price, can be plotted against time savings in a typical *toll diversion curve* shown in Figure 1.

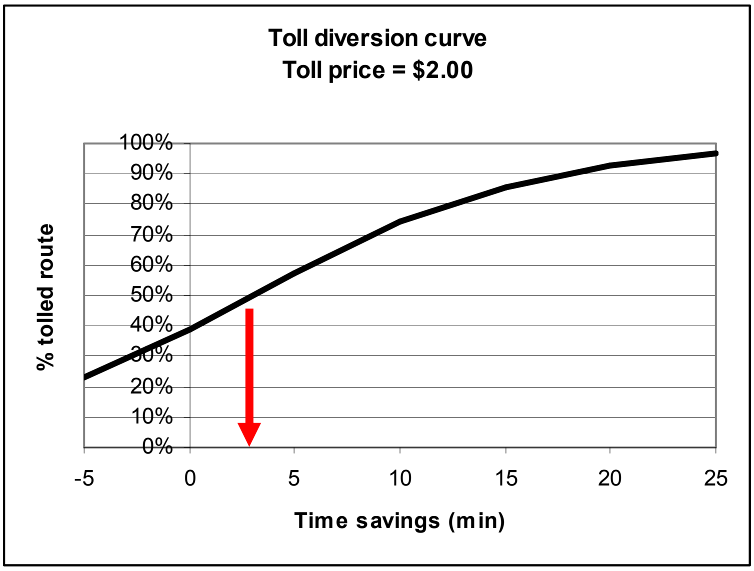


Figure 1

The probability of using the tolled route logically increases as the travel time savings increase. The slope (or λ) of the curve at the point of inflection is the point at which the choice is most sensitive to the changes in utility difference. The point on this toll diversion curve where 50% of users choose the tolled route and 50% choose the untolled route, where there is no difference between route options in general utility terms, is (by inspection) around 3 minutes of time savings. This value is typically adopted as the *toll delay penalty* (TDP) for use where the toll price is incorporated into a generalised cost assignment.

2.3.4 Deficiencies of traditional BRC models

McEvoy, Prince and Ferreira (1995) summarised the two major deficiencies of traditional BRC models as:

- “ the inability to account for the effect of congestion on travel speeds *practically* (my emphasis). Changes in link speed will alter the cost associated with toll and non-toll paths and hence route choice. To take account of this, several iterations of the toll road demand forecasting and speed updating are usually required. This process is very time consuming / cumbersome (or unrealistic where multiple toll plazas are in use) and comments by another user suggests that the fluctuations are not readily ‘damped’; and

- the task of defining trip matrix / paths of potential toll users is a fairly complex ‘art’ and open to error (particularly with multiple toll plazas).”

While there have been significant increases in computing power in recent years, these have largely been offset by increases in the complexity of the toll road demand forecasting task, with contemporary models having more zones, time periods, toll classes (or market segments) and explanatory variables to evaluate. In short, the above two comments are as relevant today as they were a decade ago.

Practitioners continue to use TDP models, even when the application is clearly inappropriate, favouring a safe and reliable (though flawed) approach. Their alternative is a BRC model with likely convergence issues, excessive run-times and arbitrary definitions of potential toll users.

Furthermore, there are two trends which will only exacerbate these major deficiencies:

- a trend towards more urban toll road projects, with more toll booths, leading in turn to more potential toll segments; and

- a trend to more electronic toll transactions, leading to the possibility of more complex tolling strategies (eg toll caps, unique entry-exit tolls, toll discounts, differential tolls).

A *new* behavioural route choice (BRC) model is proposed to overcome these deficiencies and provide an efficient solution to the problem of multiple toll booths and/or multiple tolling strategies.

3 A new behavioural route choice model

A *new* BRC model is proposed with the following features:

- a toll connectivity matrix and associated logic;
- each toll booth represented by a dummy zone;
- an acceptance condition and cutoff parameter;
- matrix indexing of the toll choice model data;
- a tolled trip threshold; and
- a demand matrix disaggregated into component trip legs.

Each feature is now discussed.

3.1 Toll terminology

The key terms used in this paper are defined below:

Toll booth	is the location on the road network where details of a particular trip are processed so as to enable a toll to be applied to that particular trip. In closed tolling systems, each entry (or onramp) and each exit (or offramp) is a separate toll booth
Toll segment	the combination of toll booths available for a particular trip
Valid toll segment	a toll segment satisfying various logic tests
Valid OD pairs	for a valid toll segment, those OD pairs satisfying an acceptance condition
Toll class	market segment (eg car and truck, work and non-work)
Times	the travel times from an origin to a destination

3.2 Toll connectivity matrix and associated logic

In an urban road network with *n* toll booths, with a maximum of three toll booths per trip, the total number of *potential* toll segments is given by:

Trips using a single toll booth	n toll segments; plus
Trips using two toll booths	n^2 toll segments; plus
Trips using three toll booths	n^3 toll segments.

Relying on brute force to calculate the potential toll segments is possible, but hardly desirable, especially as the number of toll booths in a network increases. The solution is to determine the number of *valid* toll segments through the use of a *toll connectivity matrix* together with associated logic. The approach may be best understood using an example.

Consider an urban road network with three toll booths A, B and C. There are 39 potential toll segments:

• Single toll booth trips	A B C	3
• Two toll booth trips	AA AB AC BA BB BC CA CB CC	9
• Three toll booth trips	AAA AAB AAC ABA ABB ABC ACA ACB ACC BAA BAB BAC BBA BBB BBC BCA BCB BCC CAA CAB CAC CBA CBB CBC CCA CCB CCC	27

Clearly many of these potential toll segments can be dismissed (eg AA and AAA). Related toll segments can also be dismissed (eg AAB AAC BAA CAA etc). And depending where the toll booths are located, further potential toll segments can be dismissed. Assuming the configuration below, the 39 *potential* toll segments can be reduced to seven *valid* toll segments:

A -------→ B ------→ C

Single toll booth trips:

1	A	trips from an origin zone, through toll booth A (only), to a destination zone
2	B	trips from an origin zone, through toll booth B (only), to a destination zone
3	C	trips from an origin zone, through toll booth C (only), to a destination zone

Two toll booth trips:

4	AB	trips from an ozone, through toll booths A then B (but not C), to a dzone
5	BC	trips from an ozone, through toll booths B then C (but not A), to a dzone
6	AC	trips from an ozone, through toll booths A then C (but not B), to a dzone

Three toll booth trips:

7	ABC	trips from an ozone, through toll booths A then B then C, to a dzone

The above logic can be summarised using a toll connectivity matrix, which denotes whether a trip from one toll booth to another is valid. The toll connectivity matrix for our example is shown in Table 1.

Table 1: Example toll connectivity matrix

		To	
	A	B	C
From A	0	1	1
From B	0	0	1
From C	0	0	0

The key points are:

- the valid single toll booth segments are A B C (by inspection);
- the valid two toll booth segments are extracted directly from the toll connectivity matrix as AB AC BC; and
- the valid three toll booth segments are *derived* from the toll connectivity matrix. If AB is valid, and BC is valid, then ABC is also valid. Conversely, if AC is valid, but CB is not valid, then ACB is not valid.

If we considered it unlikely that a tolled trip would use A then C, but bypass B, then we would redefine the toll connectivity matrix as shown in Table 2.

Table 2: Example toll connectivity matrix

	To		
	A	B	C
From A	0	1	0
From B	0	0	1
From C	0	0	0

From Table 2, the six valid toll segments are now A B C AB BC ABC

In this example, there would be a minimal saving in model run-times by moving from seven to six valid toll segments, so we would probably leave the toll connectivity matrix as per Table 1. But as the number of toll booths in a network increases, the number of valid toll segments can increase dramatically. For example, the Sydney road network will have around 80 toll booths by late 2006, with thousands of potential toll segments. The toll connectivity matrix permits these potential toll segments to be reduced to a manageable number of valid toll segments.

All cells in the toll connectivity matrix can, of course, be set to ‘1’. In this case, the solution will simply grind through all the toll booth combinations, with the only penalty being additional model run-times.

3.2.1 Associated logic tests

In addition to the toll connectivity matrix, the solution applies several associated logic tests:

- for a toll segment consisting of a single toll booth, the toll booth cannot be the exit ramp (or offramp) of a distance-based toll system; and
- for toll segments consisting of more than two tollbooths:
    - if toll booth i = toll booth j then toll segment ij is removed (these combinations would typically be excluded in the toll connectivity matrix);
    - if toll segment ij is not valid (from the toll connectivity matrix), then toll segment ijk is also not valid; and
    - if toll segment jk is not valid (from the toll connectivity matrix), then toll segment ijk is also not valid.

3.3 Toll booth represented by dummy zone

The previous section outlined the approach for determining the number of valid toll segments. The next step is to calculate the utilities for each valid toll segment.

Consider the previous example with three valid toll segments (A B AB), and a simple utility function which includes the explanatory variables *time* and *toll.* For each valid toll segment, it is necessary to extract the OD times prior to calculating the utilities (given that the extraction of the tolls is a trivial exercise). We need to extract:

- Toll segment A $Time_A$, the times for valid OD pairs which pass through toll booth A, without passing through any other toll booths;
- Toll segment B $Time_B$, the times for valid OD pairs which pass through toll booth B, without passing through any other toll booths;
- Toll segment AB $Time_{AB}$, the times for valid OD pairs which pass firstly through toll

booth A, then through toll booth B, without passing through any other toll booths.

Traditionally, these times have been extracted from road networks using select links, and various other cumbersome network constructions. Whilst these methods have been possible for small numbers of toll segments, the process becomes increasingly cumbersome as the number of toll segments increases. Extracting the origin-destination explanatory variables specified in a utility function (such as times) for each valid toll segment is computationally the most difficult part of a toll choice model. The traditional BRC approach is discussed further in Appendix B.

The solution introduces a dummy zone which straddles the link containing each toll booth (the *toll link*), as shown in Figure 2. The model is seeded with a traditional network assignment with the toll price incorporated into the generalised cost. The equilibrium times on every link are saved, the toll links are overwritten with 999 minutes and the network is re-skimmed to produce the times without passing through a toll booth (or the untolled times).

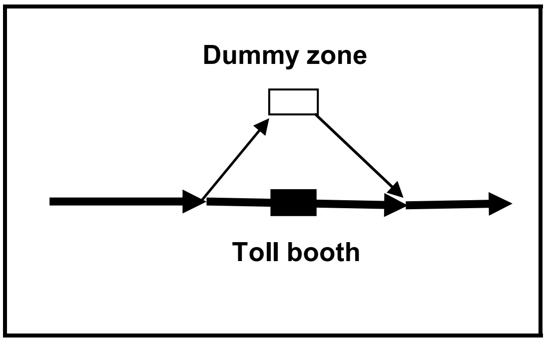


Figure 2 Toll booth construction

However, because a dummy zone now represents each toll booth, this matrix also contains the times:

- from each origin to each toll booth (without travelling through a toll booth);
- from each toll booth to each toll booth (without travelling through a toll booth); and
- from each toll booth to each destination zone (without travelling through a toll booth).

The size of the time matrix is equal to the number of zones in the road network *plus* the number of valid toll booths. The matrix of *times* for the example is shown in Figure 3.

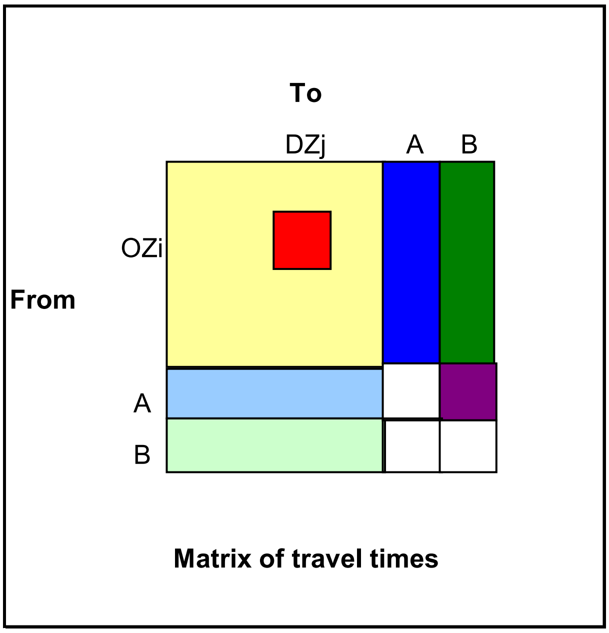


Figure 3

This matrix contains the *times* from:

- any origin to any destination (yellow square) without passing through a toll booth. The example (shown in red) is from origin i to destination j;
- any origin to toll booth A (dark blue vector);
- toll booth A to any destination (light blue vector);
- any origin to toll booth B (dark green vector);
- toll booth B to any destination (light green vector); and
- toll booth A to toll booth B (purple scalar).

All the origin-destination times for the *untolled routes* (shown in yellow in Figure 3) are retained. The OD times for each valid toll segment are slightly more complicated.

For each valid toll segment, the solution constructs a matrix of times by extracting and combining various arrays from the time matrix, as shown in Figure 4. The arrays are combined using the normal matrix rules for adding origin vectors, destination vectors and scalars. In our example, there are three valid toll segments (A B AB).

Toll segment A:

- extract an origin vector (from all origin zones to toll booth A); plus
- extract a destination vector (from tollbooth A to all destination zones).

Toll segment B:

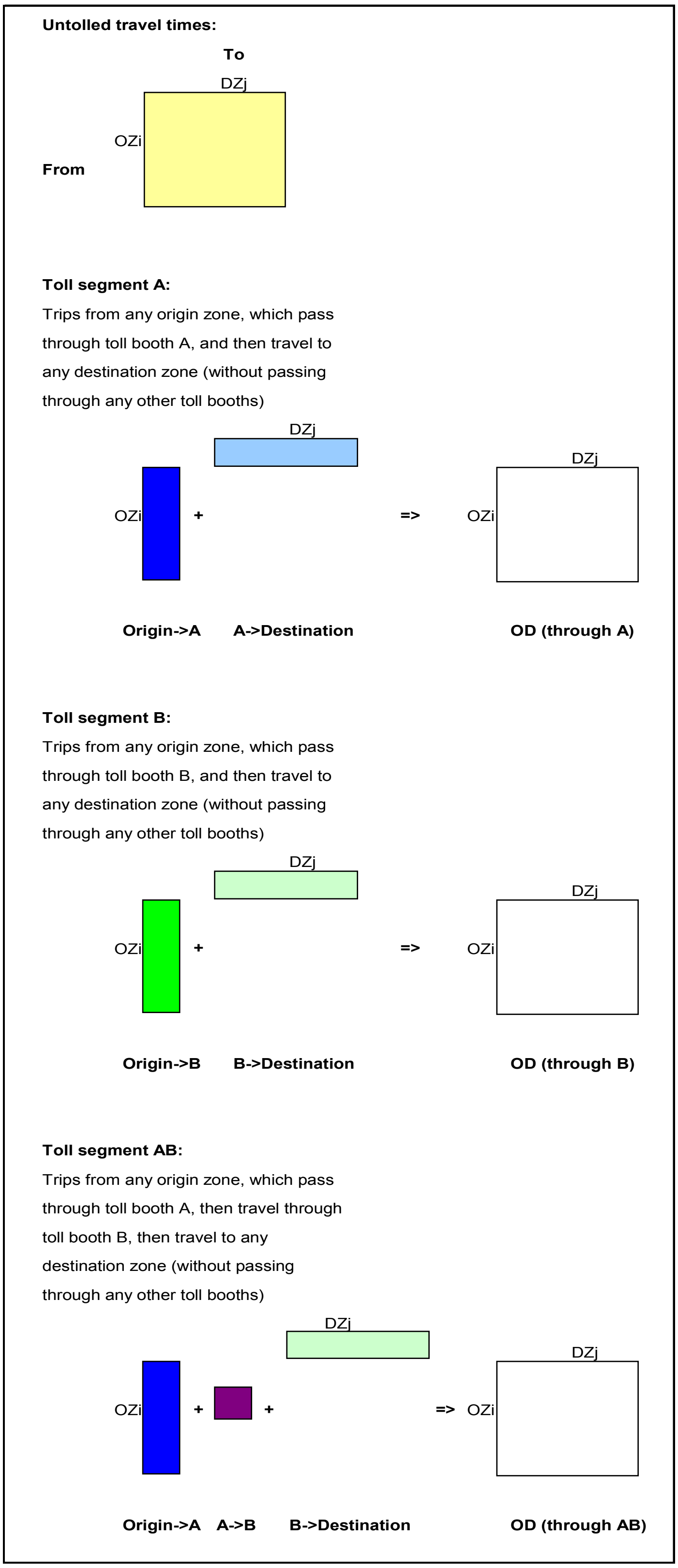
- extract an origin vector (from all origin zones to toll booth B); plus
- extract a destination vector (from tollbooth B to all destination zones).

Toll segment AB:

- extract an origin vector (from all origin zones to toll booth A); plus
- extract a scalar (from toll booth A to toll booth B); plus
- extract a destination vector (from toll booth B to all destination zones).

There are only two toll booths in our example. The general case is for *n* toll booths, where *n* is also the number of dummy zones shown in Figure 3. The solution constructs the origin-destination time matrix for each toll segment *on the fly* (ie in memory).

Our example is for a *time* variable, but other explanatory variables used in utility functions (such as StopStart time, FreeFlow time and Reliability of time) use the same approach.



Untolled travel times:

Toll segment A:
Trips from any origin zone, which pass through toll booth A, and then travel to any destination zone (without passing through any other toll booths)

Toll segment B:
Trips from any origin zone, which pass through toll booth B, and then travel to any destination zone (without passing through any other toll booths)

Toll segment AB:
Trips from any origin zone, which pass through toll booth A, then travel through toll booth B, then travel to any destination zone (without passing through any other toll booths)

Figure 4

The introduction of a dummy zone to represent each toll booth permits the explanatory OD variables used in a utility function (such as time) to be efficiently processed *on the fly* for each valid toll segment using simple matrix algebra. The solution allows the potential market for each valid toll segment to be the total trip matrix, thus overcoming one of the major deficiencies identified earlier by McEvoy, Prince and Ferreira (1995).

3.4 The acceptance condition and cutoff parameter

For each valid toll segment, the solution only retains those origin-destination (OD) pairs which satisfy an *acceptance condition*. The default acceptance condition is that the time via the tolled route *less* a cutoff parameter *less* the time via the untolled route must be less than '0' minutes. The toll segment is skipped (in the processing) if there are no valid ODs satisfying the acceptance condition. The skipped toll segment can be permanently removed by inclusion in the toll connectivity matrix, or temporarily skipped until the next iteration, where, after updating the times, there may now be OD pairs which satisfy the acceptance condition.

In addition to being part of the acceptance condition, the *cutoff parameter* (expressed in minutes) defines the level of negative time savings to be considered in the toll choice model. Consider a simple urban road network with one toll booth only. A binary logit curve, which shows the probability of choosing a tolled route compared with an untolled route, is shown in Figure 5. This particular curve shows that some motorists will continue to choose the tolled route even when there are negative time savings. These motorists are not only prepared to pay a toll, but also travel longer than the untolled route. Whilst there are plausible reasons for this (eg the tolled route may be the sign-posted route, and therefore the easiest to follow for driver's unfamiliar with the network), the extent of such behaviour is of obvious interest for sensitivity testing.

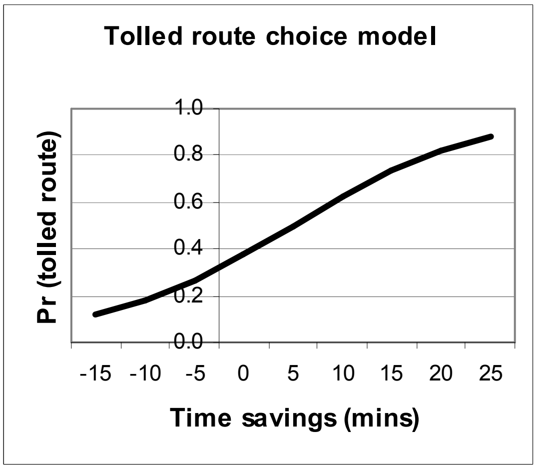


Figure 5

Cutoff parameters of 10 minutes and -5 minutes are illustrated in Figure 6.

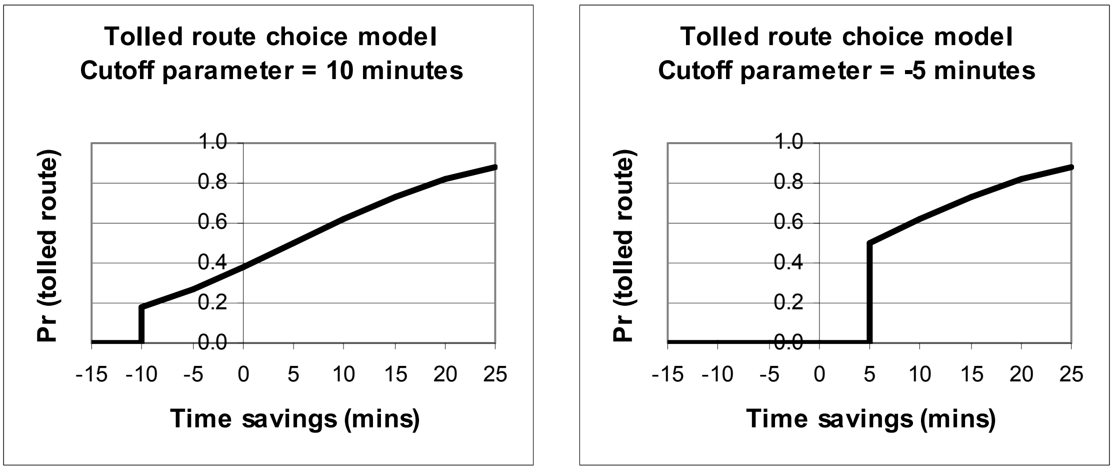


Figure 6

As the cutoff parameter increases, the number of OD pairs to be processed in the toll choice model will also generally increase, as will the model run-times.

3.4.1 Sparse toll segment matrix

Consider a city, bisected by a river and connected with one bridge, with 100 zones north of the river and 100 zones south of the river. The catchment for the southbound bridge crossing has 10,000 valid OD pairs, representing only 25 percent of the possible 40,000 OD pairs.

Similarly, although the number of valid OD pairs will vary across toll segments, the numbers will always be substantially less than the total number of cells in the (trip) matrix, even with large cutoff parameters.

The solution employs sophisticated matrix indexing to ensure that only the valid OD pairs for each valid toll segment are processed, whilst retaining full matrix functionality. Many hundreds of toll segments can be processed efficiently and the process is limited only by the available computer memory.

3.5 Tolled trip threshold

The solution applies a tolled trip threshold after the completion of the first model iteration. For each valid toll segment, if the total number of tolled trips (summed over all toll classes) is less than the trip threshold, then the toll segment is skipped in subsequent iterations. For later model runs, the skipped toll segment can be permanently removed by inclusion in the toll connectivity matrix, or temporarily skipped until the next iteration, when, after updating the times, the toll segment may now satisfy the trip threshold condition. Judicious use of the trip threshold can significantly reduce model runtimes, especially in the early stages of a project.

3.6 Trip demand matrix by leg

The final step is to disaggregate, for each valid toll segment, the tolled trips by their component trip legs. For example, if the toll choice model estimated 5 trips for a particular OD pair of toll segment ABC, then these 5 trips are disaggregated into four component trip legs:

Origin zone to toll booth A (dummy zone)	5 trips
Toll booth A (dummy zone) to toll booth B (dummy zone)	5 trips
Toll booth B (dummy zone) to toll booth C (dummy zone)	5 trips
Toll booth C (dummy zone) to destination zone	5 trips

The 5 trips on each leg are inserted into the trip demand matrix at locations Oz_A, A_B, B_C and C_Dz respectively. The trip demand matrix, which includes the dummy zones, and being the untolled trips plus the tolled legs, is then assigned to a road network having all the toll links banned, using a single class equilibrium assignment. The procedure continues iteratively (using a successive averaging technique) with updated travel times input to the toll choice model, until the predicted toll road traffic converges.

Whilst it may be desirable to separate the tolled and untolled trips for reporting purposes, or separate particular toll segment(s) and/or toll class(es), assigning the combined demands using a single class equilibrium assignment is computationally the most efficient approach.

Disaggregating the trip demand matrix by the component legs overcomes the remaining deficiency identified by McEvoy, Prince and Ferreira (1995). The *new* BRC model converges readily because *all* the tolled trips must travel through their designated toll booths (or at least through the adjacent dummy zones).

4 Conclusion

The solution to toll choices involving multiple toll booths and/or multiple tolling strategies is a *new* BRC model, the key features of which are summarised below and illustrated in Figure 7:

- valid toll segments are determined using a toll connectivity matrix and associated logic;
- each toll booth is represented by a dummy zone;
- an acceptance condition is imposed which identifies, for each valid toll segment, the number of valid origin-destination pairs to be evaluated in the toll choice model. The acceptance condition includes a cutoff parameter, which also sets the amount of negative time savings for toll users;
- the solution employs sophisticated matrix indexing to ensure that only the valid OD pairs for each toll segment are processed, whilst retaining full matrix functionality. Many hundreds of toll segments can be processed efficiently and the process is limited only by the available computer memory;
- the use of a trip threshold after the first iteration of the model can significantly reduce the number of toll segments to be considered in later iterations of the model, by skipping those toll segments where the total tolled trips (summed across all toll classes) are less than the trip threshold; and
- disaggregating the tolled trips into their component legs, summing across all toll segments and adding the untolled trips, ensures that a computationally efficient single class equilibrium assignment can be performed. The model converges readily because *all* the tolled trips must travel through their designated toll booths (or at least through the adjacent dummy zones).

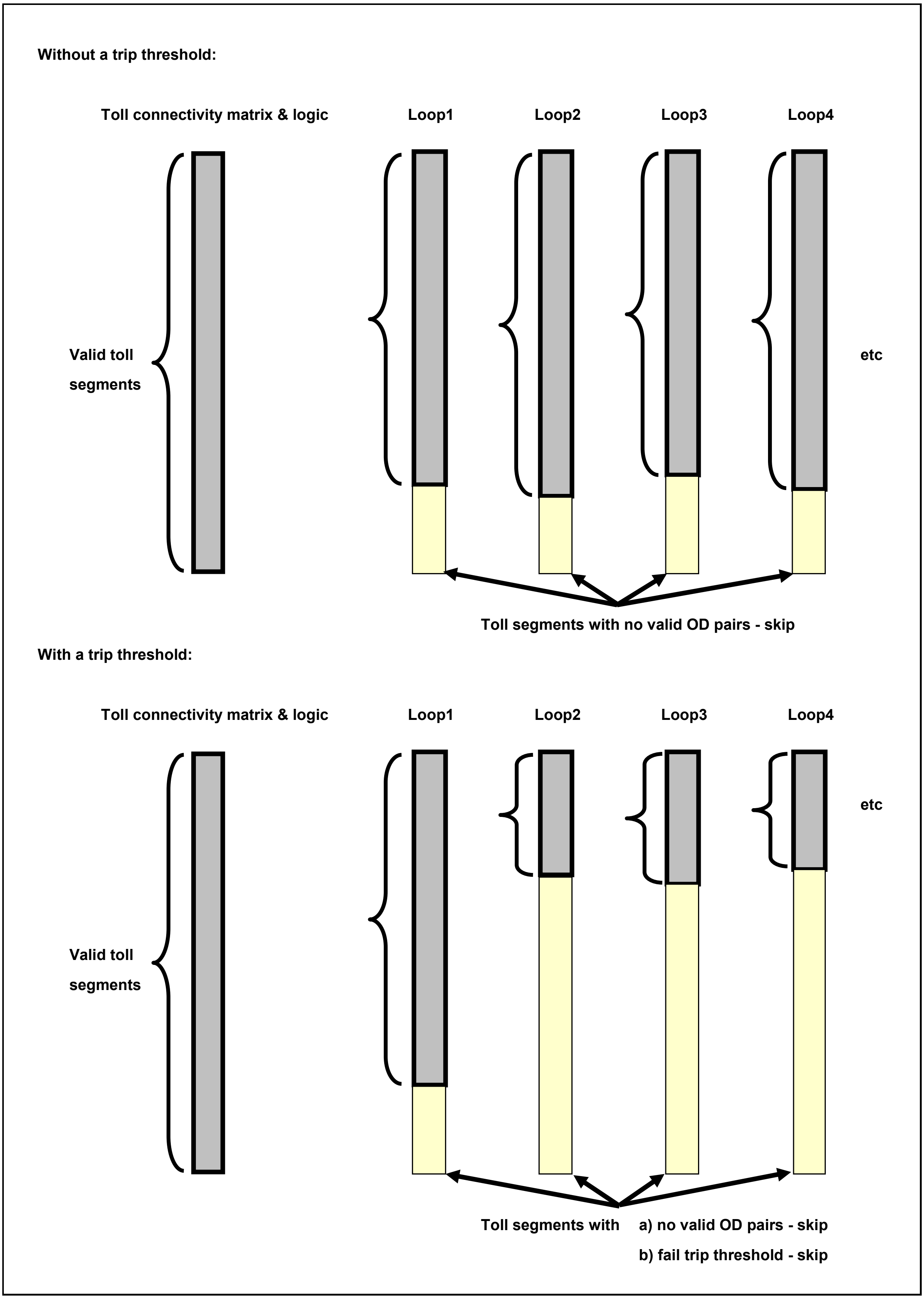

Without a trip threshold:
Toll connectivity matrix & logic
Loop1
Loop2
Loop3
Loop4
Valid toll segments
etc
Toll segments with no valid OD pairs - skip
With a trip threshold:
Toll connectivity matrix & logic
Loop1
Loop2
Loop3
Loop4
etc
Valid toll segments
Toll segments with a) no valid OD pairs - skip
b) fail trip threshold - skip

Figure 7

5 Appendix A

This example illustrates the following feature of the *new* BRC model:

- defining each toll booth as a dummy zone;
- for each valid toll segment, extracting the explanatory origin-destination variables (such as times);
- for each valid toll segment, retaining only the OD pairs satisfying the acceptance condition;
- calculating utilities for both the untolled route and the (valid OD pairs only of the) valid toll segments;
- calculating the probability of using a particular toll segment;
- calculating the number of tolled trips using a particular toll segment; and
- calculating the trip demand matrix by disaggregating the tolled trips into component legs, summing across valid toll segments and adding the remaining untolled trips.

Consider the road network shown in Figure 8 consisting of 4 zones, and two toll booths. To simplify the presentation, the toll booths are assumed to be one-way and times are assumed to be the same in either direction.

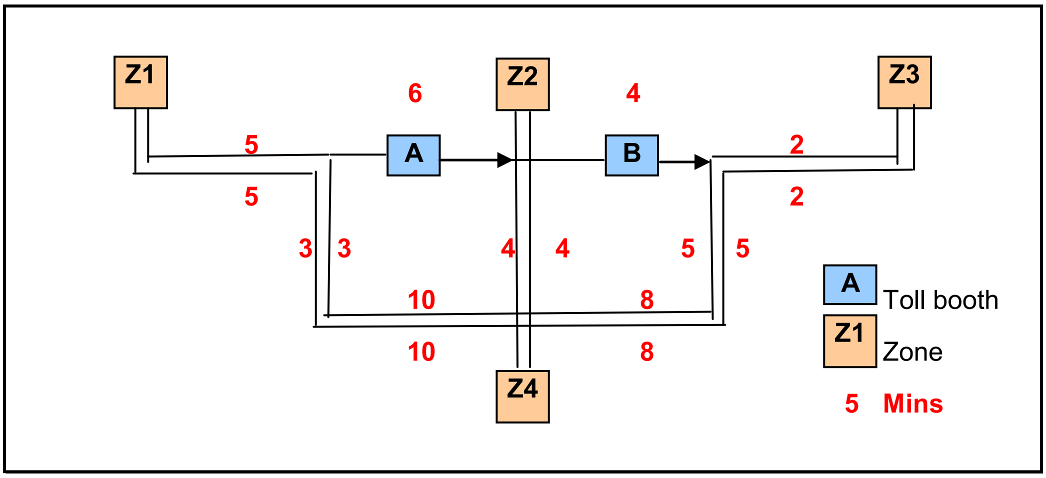


Figure 8

For the two toll booth system shown in Figure 8, there are six possible toll segments:

- A B AA AB BA BB

By inspection, trips using a single toll booth (either A or B) should be included; AA and BB can logically be removed, leaving four toll segments.

We could confidently say that BA is very unlikely ie a trip passing firstly through toll booth B and then looping around to pass secondly through toll booth A. This leaves *three valid toll segments A, B and AB*.

We could formalise this process by designing a *toll connectivity matrix*, as shown in Table 3.

Table 3: Toll Connectivity matrix

	TO toll booth A	TO toll booth B
FROM toll booth A	0	1
FROM toll booth B	0	0

The Toll Connectivity matrix has allowed us to move from six potential toll segments to three valid toll segments.

The untolled times, plus the tolled times for the three toll segments (A B AB), are shown in Figure 9. The untolled times choose the quickest path through the system *without* using either toll booth A or toll booth B. The tolled times for toll segment A choose the quickest path through the system but *must* pass through toll booth A and *cannot* pass through toll booth B (an example from Zone 3 to Zone 1 is shown in Figure 10). Similarly, the tolled times for toll segment B choose the quickest path through the system but *must* pass through toll booth B and *cannot* pass through toll booth A. Finally, the tolled times for toll segment AB choose the quickest path through the system but *must* pass firstly through toll booth A and secondly through toll booth B.

Untolled times

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		22	33	18
	Z2	22		19	4
	Z3	33	19		15
	Z4	18	4	15	

TOLL SEGMENT A

Travel times (mins)

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		11	30	15
	Z2	45		42	27
	Z3	56	34		38
	Z4	41	19	38	

Cutoff parameter (mins)

5

Acceptance condition

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		-16	-8	-8
	Z2	18		18	18
	Z3	18	10		18
	Z4	18	10	18	

Valid ODs

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		-16	-8	-8
	Z2				
	Z3				
	Z4				

TOLL SEGMENT B

Travel times (mins)

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		43	28	39
	Z2	35		6	17
	Z3	54	40		36
	Z4	39	25	10	

Cutoff parameter (mins)

5

Acceptance condition

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		16	-10	16
	Z2	8		-18	8
	Z3	16	16		16
	Z4	16	16	-10	

Valid ODs

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**			-10	
	Z2			-18	
	Z3				
	Z4			-10	

TOLL SEGMENT AB

Travel times (mins)

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		32	17	28
	Z2	58		29	40
	Z3	69	55		51
	Z4	54	40	25	

Cutoff parameter (mins)

5

Acceptance condition

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**		5	-21	5
	Z2	31		5	31
	Z3	31	31		31
	Z4	31	31	5	

Valid ODs

TO zone

		Z1	Z2	Z3	Z4
FRC	**Z1**			-21	
	Z2				
	Z3				
	Z4				

Figure 9: Acceptance condition and valid origin-destination pairs

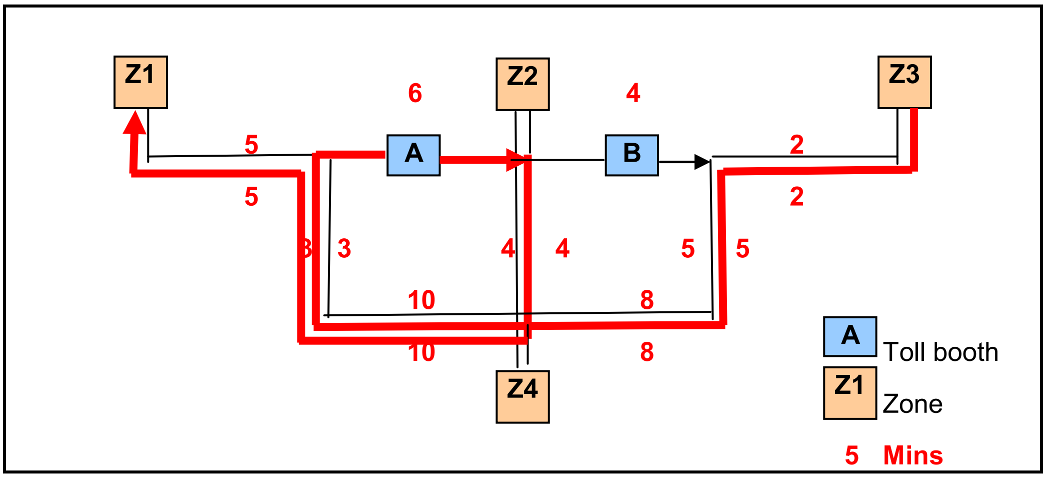


Figure 10: Toll segment A – Zone 3 to Zone 1

The key points are:

- some motorists are willing to pay a toll in exchange for travel time savings. Some motorists are still willing to pay a toll, even though they receive no (or even negative) time savings – presumably these motorists receive benefits other than travel time savings (eg safer, easier driving, sign-posted route etc);
- however, it would be reasonable to propose that there is a limit to the amount of negative time savings – this is specified using a *cutoff parameter*. For each valid toll segment, only those origin-destination pairs where the tolled time *less* the cutoff parameter *less* the untolled time is less than '0' minutes are retained;
- the number of valid origin-destination (OD) pairs is logically different for each toll segment, and also varies depending on the selected cutoff parameter. As the *cutoff parameter* is increased, the number of valid ODs also increases, as shown below:

	Cutoff parameter 5 mins	Cutoff parameter 10 mins
Toll segment A	3	12
Toll segment B	3	12
Toll segment AB	1	12
Total	7	36

Use of a cutoff parameter of 5 minutes has allowed us to reduce the number of origin-destination pairs from a possible 48 (ie 4 zones x 4 zones x 3 valid toll segments) to 7 valid OD pairs.

In summary, through the use of a *toll connectivity matrix* (Table 3) and a *cutoff parameter* (5 minutes), we have reduced the toll choice to:

- Toll segment A (three valid ODs);
- Toll segment B (three valid ODs); and
- Toll segment AB (one valid OD).

The reduced toll choice set also results in significant computational savings. These savings accrue exponentially as:

- the number of toll booths increases; and/or
- the number of toll booths allowed in a single trip increases.

In Figure 9, the untolled travel times are traditionally generated by skimming a loaded network, with both toll links banned. However, it has been difficult to generate the tolled travel times for each toll segment. The road network construction shown in Figure 11 shows how this can be achieved.

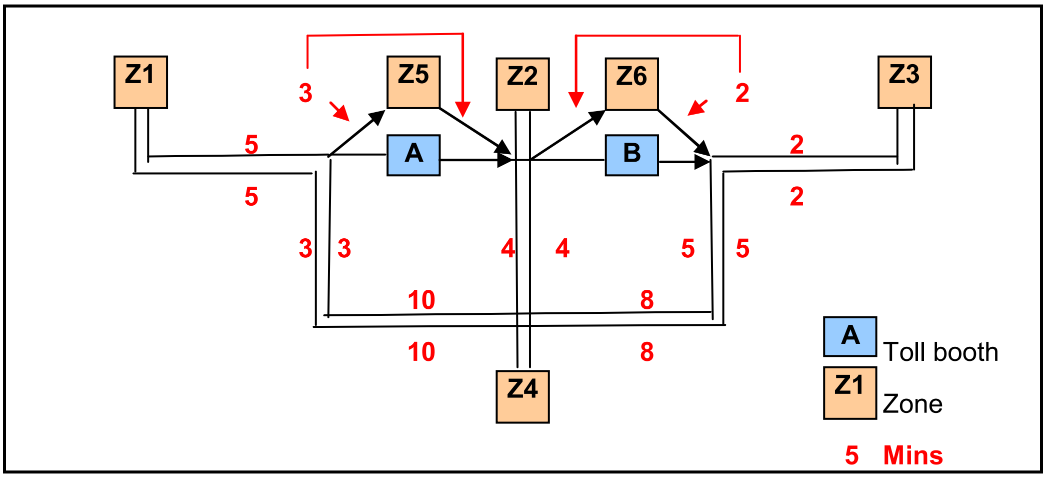


Figure 11

A *dummy* zone straddles each link containing a toll booth. In the example:

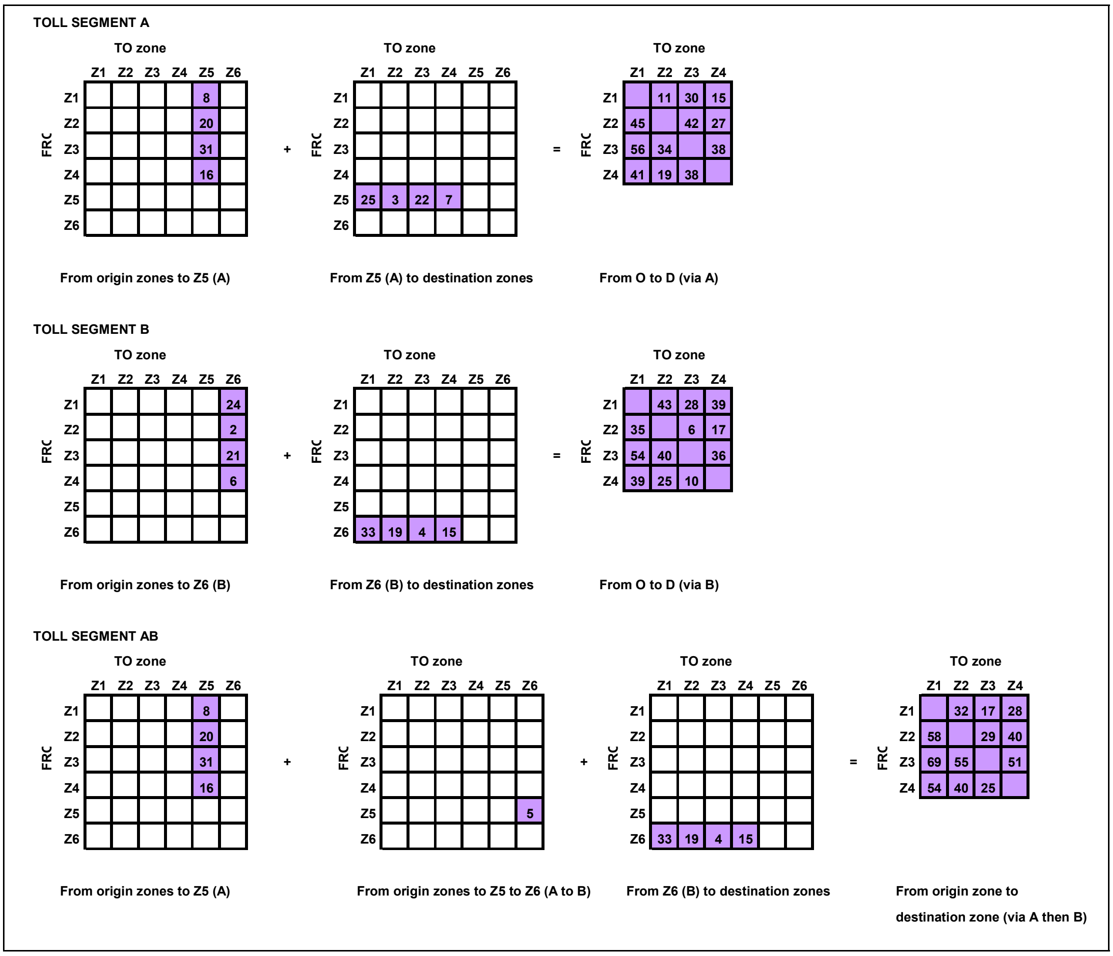
- Zone 5 represents toll booth A; and
- Zone 6 represents toll booth B.

The matrix of times shown in Table 4 can be generated by skimming a loaded network, again with both toll links banned. The block of times shaded yellow represents the untolled times, which are the same as those shown in Figure 9.

Table 4: Times

		TO zone					
		Z1	Z2	Z3	Z4	Z5	Z6
FRC	Z1		22	33	18	8	24
	Z2	22		19	4	20	2
	Z3	33	19		15	31	21
	Z4	18	4	15		16	6
	Z5	25	3	22	7		5
	Z6	33	19	4	15	31	

Interestingly, the tolled travel times for each toll segment can now be generated *on the fly* using the times in Table 4 and simple matrix algebra. The process is illustrated in Figure 12.



TOLL SEGMENT A

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1					8	
Z2					20	
Z3					31	
Z4					16	
Z5						
Z6						

From origin zones to Z5 (A)

+

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1						
Z2						
Z3						
Z4						
Z5	25	3	22	7		
Z6						

From Z5 (A) to destination zones

=

FRC / TO zone	Z1	Z2	Z3	Z4
Z1		11	30	15
Z2	45		42	27
Z3	56	34		38
Z4	41	19	38	

From O to D (via A)

TOLL SEGMENT B

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1						24
Z2						2
Z3						21
Z4						6
Z5						
Z6						

From origin zones to Z6 (B)

+

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1						
Z2						
Z3						
Z4						
Z5						
Z6	33	19	4	15		

From Z6 (B) to destination zones

=

FRC / TO zone	Z1	Z2	Z3	Z4
Z1		43	28	39
Z2	35		6	17
Z3	54	40		36
Z4	39	25	10	

From O to D (via B)

TOLL SEGMENT AB

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1					8	
Z2					20	
Z3					31	
Z4					16	
Z5						
Z6						

From origin zones to Z5 (A)

+

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1						
Z2						
Z3						
Z4						
Z5						5
Z6						

From origin zones to Z5 to Z6 (A to B)

+

FRC / TO zone	Z1	Z2	Z3	Z4	Z5	Z6
Z1						
Z2						
Z3						
Z4						
Z5						
Z6	33	19	4	15		

From Z6 (B) to destination zones

=

FRC / TO zone	Z1	Z2	Z3	Z4
Z1		32	17	28
Z2	58		29	40
Z3	69	55		51
Z4	54	40	25	

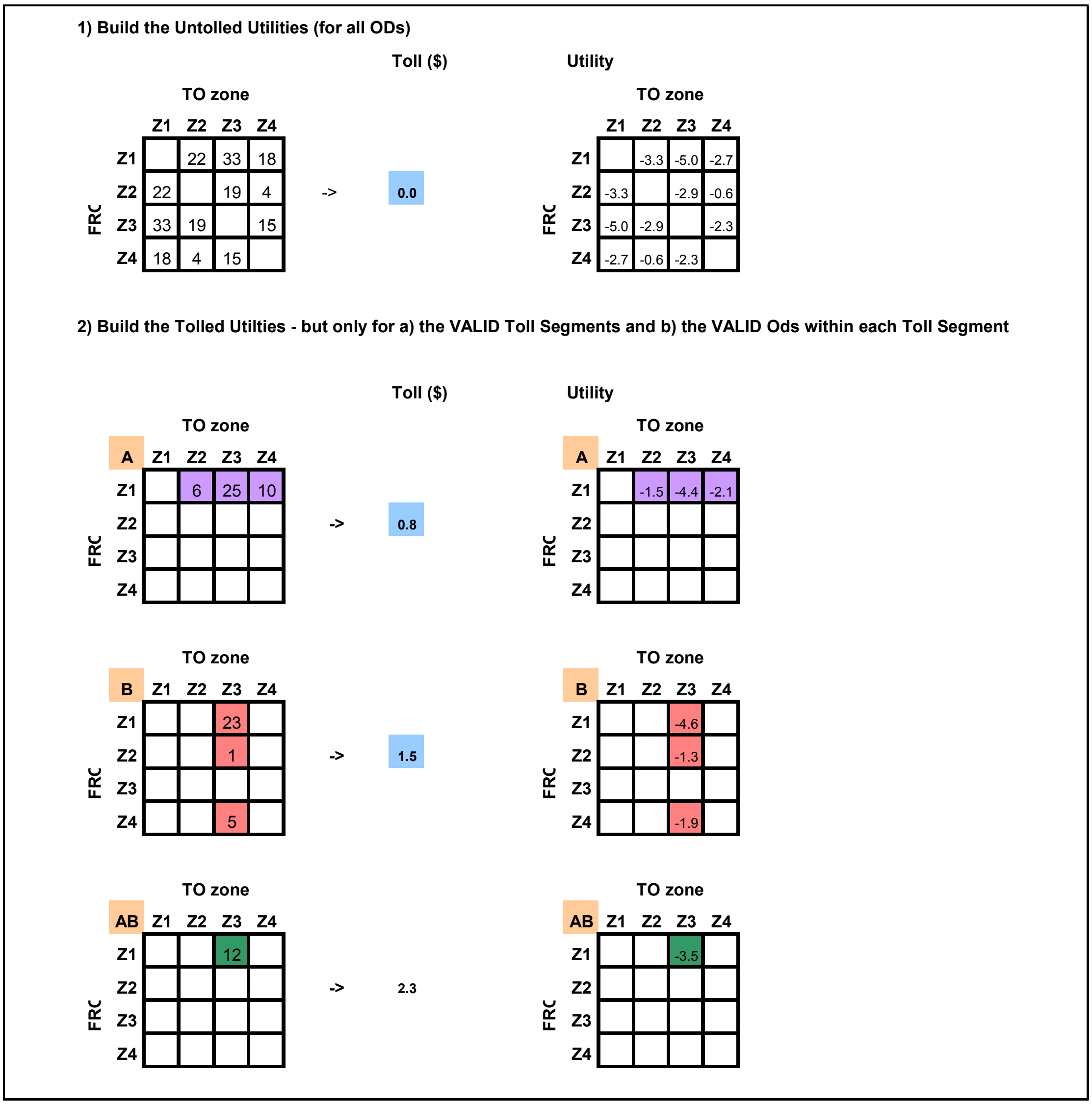
From origin zone to destination zone (via A then B)

Figure 12: Generated tolled times by valid toll segment

The next step builds the *utilities*, again *on the fly*. For our example, we will use a simple utility consisting of *time* and *toll* variables only, with illustrative parameters:

• Untolled	U_j	$= a1 \,.\, time_j$	time (mins)
• Tolled	U_i	$= a1 \,.\, time_i + a2 \,.\, toll_i$	time (mins), toll ($)
		$= a1 \,.\, (time_i + a2/a1 \,.\, toll_i)$	VoT (a2/a1) mins/$
		$= -0.15\ (time_i + 5 \,.\, toll_i)$	time (mins), toll ($)

In our example, the tolls are $0.80 (toll booth A) and $1.50 (toll booth B). The untolled utilities are calculated for *all* origin-destination pairs, but the tolled utilities are only calculated for the valid ODs of each toll segment. Generally, the toll for toll segment AB is $2.30 (ie the sum of toll booth A and toll booth B). However, it is possible for a user to overwrite with a more complex tolling strategy (such as a toll discount or a toll cap).



1) Build the Untolled Utilities (for all ODs)

Toll ($) -> 0.0 -> Utility

TO zone (FRO)

	Z1	Z2	Z3	Z4
Z1		22	33	18
Z2	22		19	4
Z3	33	19		15
Z4	18	4	15	

TO zone (FRO)

	Z1	Z2	Z3	Z4
Z1		-3.3	-5.0	-2.7
Z2	-3.3		-2.9	-0.6
Z3	-5.0	-2.9		-2.3
Z4	-2.7	-0.6	-2.3	

2) Build the Tolled Utilties - but only for a) the VALID Toll Segments and b) the VALID Ods within each Toll Segment

Toll ($) -> 0.8 -> Utility

TO zone (FRO)

A	Z1	Z2	Z3	Z4
Z1		6	25	10
Z2				
Z3				
Z4				

TO zone (FRO)

A	Z1	Z2	Z3	Z4
Z1		-1.5	-4.4	-2.1
Z2				
Z3				
Z4				

-> 1.5 ->

TO zone (FRO)

B	Z1	Z2	Z3	Z4
Z1			23	
Z2			1	
Z3				
Z4			5	

TO zone (FRO)

B	Z1	Z2	Z3	Z4
Z1			-4.6	
Z2			-1.3	
Z3				
Z4			-1.9	

-> 2.3 ->

TO zone (FRO)

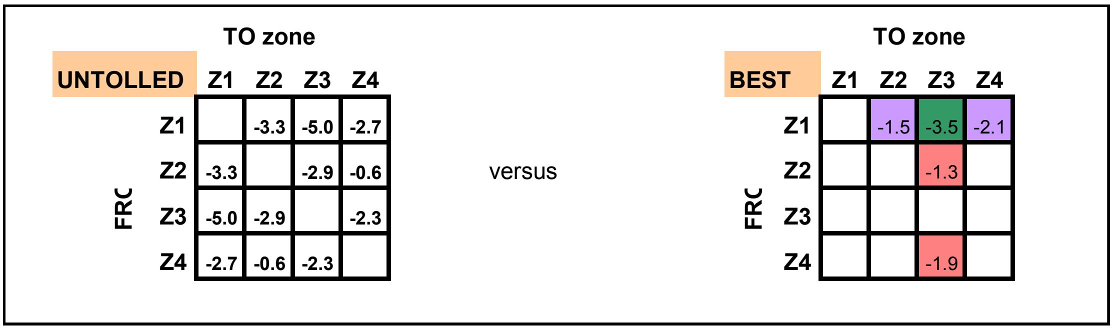
AB	Z1	Z2	Z3	Z4
Z1			12	
Z2				
Z3				
Z4				

TO zone (FRO)

AB	Z1	Z2	Z3	Z4
Z1			-3.5	
Z2				
Z3				
Z4				

Figure 13

In our example, we will use a toll choice model where the upper level choice is between the untolled utilities and the *best* tolled utilities, as shown in Figure 14.



TO zone (FRO)

UNTOLLED	Z1	Z2	Z3	Z4
Z1		-3.3	-5.0	-2.7
Z2	-3.3		-2.9	-0.6
Z3	-5.0	-2.9		-2.3
Z4	-2.7	-0.6	-2.3	

versus

TO zone (FRO)

BEST	Z1	Z2	Z3	Z4
Z1		-1.5	-3.5	-2.1
Z2			-1.3	
Z3				
Z4			-1.9	

Figure 14

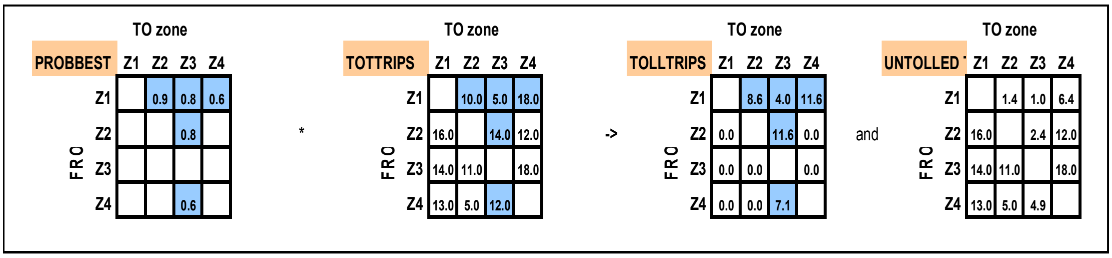
For the valid ODs of each valid toll segment, and a cutoff parameter of 5 minutes, the *best* tolled utilities are (from Figure 9):

Zone 1 to Zone 2	Toll segment A
Zone 1 to Zone 3	Toll segment C (followed by A then B)

Zone 1 to Zone 4	Toll segment A
Zone 2 to Zone 3	Toll segment B
Zone 4 to Zone 3	Toll segment B

In our example, the probability of making a tolled trip is made by comparing the untolled and *best* Tolled utilities using a binary logit, as shown in Equation 2 and detailed in Figure 15. The tolled trip matrix is derived by multiplying the tolled probabilities with an example *total* trip matrix. The untolled trip matrix is simply derived by subtraction.

Equation 2 $\text{Pr (tolled)} = \exp (U_{\text{tolled}}) / (\exp(U_{\text{tolled}}) + \exp(U_{\text{untolled}})$



PROBBEST	Z1	Z2	Z3	Z4
Z1		0.9	0.8	0.6
Z2			0.8	
Z3				
Z4			0.6	

TOTTRIPS	Z1	Z2	Z3	Z4
Z1		10.0	5.0	18.0
Z2	16.0		14.0	12.0
Z3	14.0	11.0		18.0
Z4	13.0	5.0	12.0	

TOLLTRIPS	Z1	Z2	Z3	Z4
Z1		8.6	4.0	11.6
Z2	0.0		11.6	0.0
Z3	0.0	0.0		0.0
Z4	0.0	0.0	7.1	

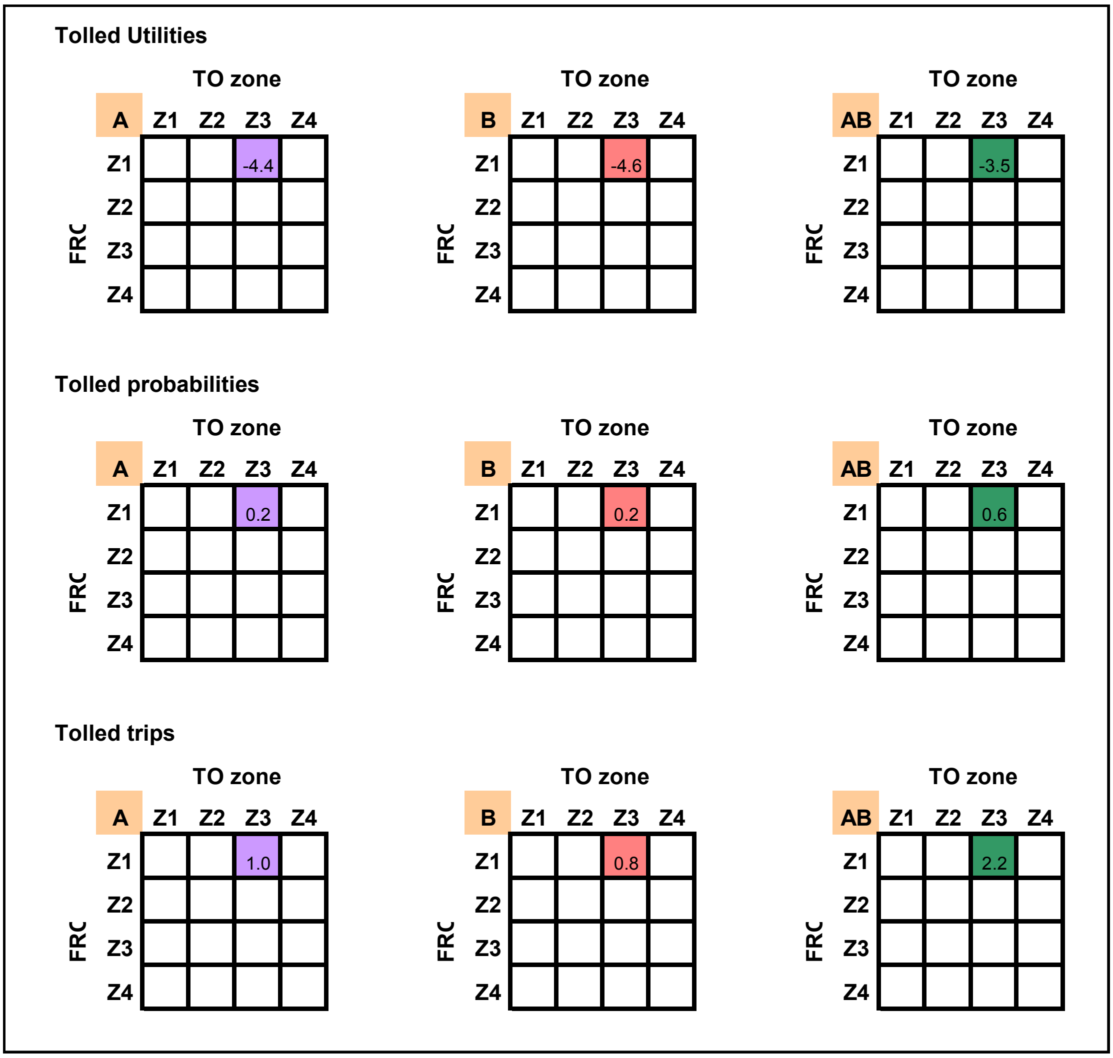
UNTOLLED	Z1	Z2	Z3	Z4
Z1		1.4	1.0	6.4
Z2	16.0		2.4	12.0
Z3	14.0	11.0		18.0
Z4	13.0	5.0	4.9	

Figure 15

In our example, there are 4.0 tolled trips from Zone 1 to Zone 3. Although toll segment C has the *best* tolled utility for this OD pair, there are also valid utilities for toll segments A and B. For this OD pair, it would seem reasonable to allocate the 4.0 tolled trips across the three valid toll segments. In our example, we will allocate these trips using a multinomial logit model, as shown in Equation 3 and detailed in Figure 16.

Equation 3

$$\Pr(A) = \exp(U_A) / (\exp((U_A)+\exp(U_B)+\exp(U_C))$$
$$\Pr(B) = \exp(U_B) / (\exp((U_A)+\exp(U_B)+\exp(U_C))$$
$$\Pr(C) = \exp(U_C) / (\exp((U_A)+\exp(U_B)+\exp(U_C))$$



Tolled Utilities

TO zone

	A	Z1	Z2	Z3	Z4
FRC	Z1			-4.4	
	Z2				
	Z3				
	Z4				

TO zone

	B	Z1	Z2	Z3	Z4
FRC	Z1			-4.6	
	Z2				
	Z3				
	Z4				

TO zone

	AB	Z1	Z2	Z3	Z4
FRC	Z1			-3.5	
	Z2				
	Z3				
	Z4				

Tolled probabilities

TO zone

	A	Z1	Z2	Z3	Z4
FRC	Z1			0.2	
	Z2				
	Z3				
	Z4				

TO zone

	B	Z1	Z2	Z3	Z4
FRC	Z1			0.2	
	Z2				
	Z3				
	Z4				

TO zone

	AB	Z1	Z2	Z3	Z4
FRC	Z1			0.6	
	Z2				
	Z3				
	Z4				

Tolled trips

TO zone

	A	Z1	Z2	Z3	Z4
FRC	Z1			1.0	
	Z2				
	Z3				
	Z4				

TO zone

	B	Z1	Z2	Z3	Z4
FRC	Z1			0.8	
	Z2				
	Z3				
	Z4				

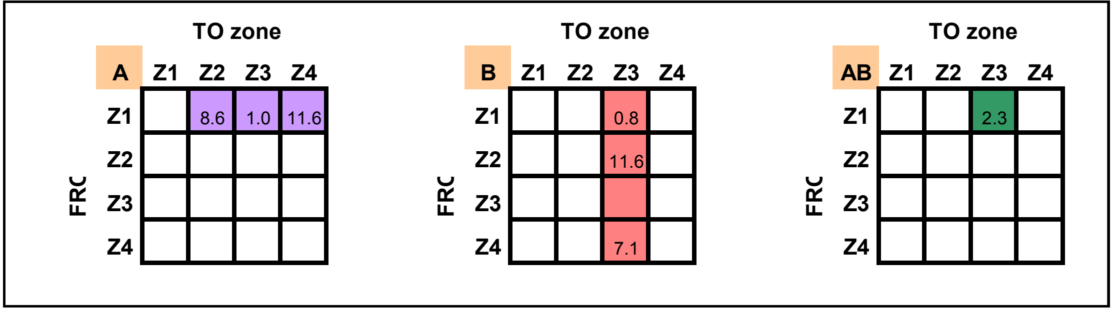
TO zone

	AB	Z1	Z2	Z3	Z4
FRC	Z1			2.2	
	Z2				
	Z3				
	Z4				

Figure 16 Allocate tolled trips (Z1->Z3) across valid toll segments
Note: numbers may not add due to rounding

The total tolled trips for each toll segment can now be summarised in Table 5.

Table 5 Total tolled trips



TO zone

	A	Z1	Z2	Z3	Z4
FRC	Z1		8.6	1.0	11.6
	Z2				
	Z3				
	Z4				

TO zone

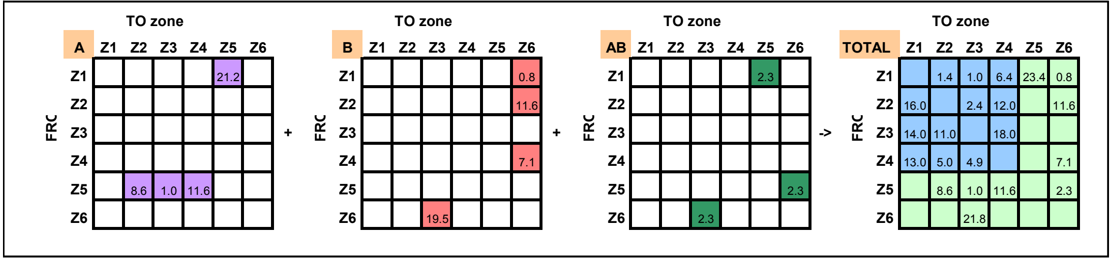
	B	Z1	Z2	Z3	Z4
FRC	Z1			0.8	
	Z2			11.6	
	Z3			🟥	
	Z4			7.1	

TO zone

	AB	Z1	Z2	Z3	Z4
FRC	Z1			2.3	
	Z2				
	Z3				
	Z4				

Finally, for each valid toll segment, the tolled trips can be converted to *toll legs*, and the untolled trips added, as shown in Table 6. Whilst it may be desirable to separate the tolled and untolled trips for reporting purposes (or separating a particular toll segment(s) and/or toll class(es)), assigning the combined demands using a single class equilibrium assignment is computationally the most efficient approach.

Table 6: Final trip demand matrix



TO zone

A	Z1	Z2	Z3	Z4	Z5	Z6
Z1					21.2	
Z2						
Z3						
Z4						
Z5		8.6	1.0	11.6		
Z6						

+

TO zone

B	Z1	Z2	Z3	Z4	Z5	Z6
Z1						0.8
Z2						11.6
Z3						
Z4						7.1
Z5						
Z6			19.5			

+

TO zone

AB	Z1	Z2	Z3	Z4	Z5	Z6
Z1					2.3	
Z2						
Z3						
Z4						
Z5						2.3
Z6			2.3			

->

TO zone

TOTAL	Z1	Z2	Z3	Z4	Z5	Z6
Z1		1.4	1.0	6.4	23.4	0.8
Z2	16.0		2.4	12.0		11.6
Z3	14.0	11.0		18.0		
Z4	13.0	5.0	4.9			7.1
Z5		8.6	1.0	11.6		2.3
Z6			21.8			

6 Appendix B

This appendix discusses how traditional BRC models are generally applied in practice, and then compares and contrasts this with the proposed *new* BRC model.

In a traditional BRC model, the potential market for each toll road is first estimated, by assigning the total demand to a network assuming an untolled facility, and extracting the users of each of the tolled movements by extracting select links. Because the facility is likely to be heavily over-saturated and delivering poor performance when operating as an untolled freeway, various techniques (such as seeding the facility with fixed, high speeds) are used to generate a potential market which is larger than for the untolled facility. However, this enlarged potential market includes those users who might use the toll road despite incurring some negative time savings (or time dis-benefits), to receive other benefits from the tolled route such as reliability, safety and higher quality driving environment.

The variables included in the toll choice model (eg travel time, toll etc) are calculated for each potential trip and a generalised cost calculated for the tolled and untolled routes. The time savings by using the toll road can be readily calculated.

The toll choice model is then applied to each potential market and total demand is split into toll road users and non-users for each market segment. The toll choice model calculates for each origin-destination pair, the generalised costs to travel via the untolled and tolled routes, and determines the proportion of trips which are prepared to pay the toll. These are applied to the set of potential trip matrices.

The resulting tolled and untolled users are then re-assigned to the network. The usual speed-flow curves are re-introduced to the tolled routes and travel speeds are updated depending on the traffic volume. The tolled trips are generally assigned first to the whole network (which has no toll penalties) and then the remaining untolled trips are assigned to the network where links containing toll booths are banned. There is no guarantee that the tolled trips will be assigned to the correct routes.

The procedure continues iteratively with updated travel times input to the toll choice model, until the predicted toll road traffic converges.

The calculation of generalised costs for the tolled route and the next best untolled route relies on using network skims. Iteration procedures can now be used where previously they were too time-consuming and cumbersome to implement.

The proposed *new* BRC model reverses the traditional method. Firstly, a traditional network equilibrium assignment is performed with the toll incorporated into the generalised cost. The equilibrium travel times on every link are saved, the toll links are overwritten with 999 minutes and the network is re-skimmed to produce, for each origin-destination pair, the travel times on the untolled route.

However, because dummy zones are incorporated to represent each toll booth, this matrix also contains the times:

- from each origin zone to each toll booth, *without travelling through a toll booth*;
- from each toll booth to each toll booth, *without travelling through a tollbooth*; and
- from each toll booth to each destination zone, *without travelling through a tollbooth*.

The potential toll road market for each market segment consists of *all* origin-destination pairs. It is thus possible to assess the tolled route options, whether using a single toll booth, two or more toll booths, against the untolled route. The approach loops around the valid toll segments (or valid tolled routes), only processing the utilities of the valid OD pairs. Therefore, for each valid OD pair, there may be a number of valid toll segments (or tolled routes), where each toll segment has its own utility.

A toll choice model then calculates the probability of making a tolled trip by a particular toll segment, depending on the relative untolled and tolled utilities and the structure of the toll choice model.

The final step is to disaggregate the tolled trips by component trip legs. For example, if the toll choice model estimated 5 tolled trips for toll segment ABC, then these 5 trips would be disaggregated into 4 trip legs:

• origin zone to toll booth A (dummy zone)	5 trips
• toll booth A (dummy zone) to toll booth B (dummy zone)	5 trips
• toll booth B (dummy zone) to toll booth C (dummy zone)	5 trips
• toll booth C (dummy zone) to destination zone	5 trips

The total demand matrix, which includes the dummy zones, and being the untolled trips plus the tolled legs, is then assigned to a network having all toll links banned. The procedure continues iteratively (using a successive averaging process) with updated travel times input to the toll choice model, until the predicted toll road traffic converges.

7 References

McEvoy, J Prince, A and Ferreira, L (1995) *Toll road modelling approaches for Queensland* Road and Transport Research Vol 4 No 3 Sept 1995