Examining route section level tram-involved crash frequency using the random effects negative binomial model

Farhana Naznin¹, Graham Currie¹, David Logan² and Majid Sarvi¹

¹Department of Civil Engineering, Building 60, Monash University
²Monash University Accident Research Centre

Email for correspondence: farhana.naznin@monash.edu

Abstract

Safety is an overriding concern in design, operation and development of light rail systems including trams or streetcars as they impose crash risks on road users in terms of crash frequency and severity. The aim of this study is to identify the key traffic, transit and route factors that influence tram-involved crash frequencies along tram route sections in Melbourne. A random effects negative binomial (RENB) regression model was developed to analyse crash frequency data obtained from Yarra Trams, the tram operator in Melbourne tram network. The RENB modelling approach can account for spatial and temporal variations within observation groups in panel count data structures by assuming that the group specific effects are randomly distributed across locations. The results identify many significant factors affecting crash frequency. They are, in order of affect; tram stop spacing (-0.43), tram route section length (0.31), tram signal priority (-0.263), general traffic volume (0.17) and tram lane priority (-0.148). Platform stops (-0.09) and service frequency (0.004) also influence crash frequency. Findings provide useful insights on route section level tram-involved crashes in an urban tram or streetcar operating environment. The method described represents a useful planning tool for transit agencies hoping to improve safety performance.

Keywords: Safety, Tram-involved crash frequency, random effects negative binomial model

1. Introduction

Public transport is becoming important as mobility, accessibility and environmental problems are increasing in cities around the world (Gakenheimer, 1999, Fouracre et al., 2003). The growing needs for more efficient use of limited roadway space has led to the promotion of preferential treatments for high-occupancy transit vehicles (Slinn et al., 2005). The fundamental goal for improving transit is to improve transit travel time, provide efficient and reliable service to passengers and make transit a competitive travel option for commuters. However transit safety issues have received less priority, although transit-involved collisions represent a significant component of road crashes (Shahla et al., 2009, Cheung et al., 2008).

Trams/streetcars are light rail transit vehicles operating on tracks located on roads used by general road traffic. Tram systems have number of attractive features including their high passenger capacity, good comfort, and very low emission of pollutants compared to other transport systems (Anna and Bruce, 2001, Cliche and Reid, 2007). However, trams present a range of inherent safety issues regarding their design and operational characteristics, since they are large and heavy vehicles operating in confined, mixed and complex environments with pedestrians and cars, and even at low speed trams have been identified to have high crash risks compared to other vehicles (Grzebieta et al., 1999). Previous studies have also identified that trams impose more crash risks at intersections and along arterials than buses, and this is likely due to difference in operational methods between buses and trams (Shahla et al., 2009, Cheung et al., 2008).
The most common types of tram related incidents are collisions between trams, collisions of tram with road vehicles, trams hit person, trams hit infrastructure, tram collisions with obstructions, passengers fall on trams or platforms, and line derailments (Transport Safety Victoria, 2013). In Melbourne, 956 collisions were recorded during 2007–2008 and among these 736 collisions involved cars striking trams and 29 were trams hitting pedestrians (Dowling and Duzelovski, 2008). Mitra et al. (2010) explored tram related crash patterns and identified that ‘pedestrian hit by tram’ and ‘passenger fall down on tram’ are the common tram incident types. Grzebieta et al. (1999) pointed that the common type of tram and other vehicles related crashes occur when vehicles making U-turns or right turns in front of trams. They also identified that different tram classes has different front design and has different impact on cars and pedestrians. Candappa (2013) identified that the vehicles making U-turn through a median opening across tram tracks was largely struck by a tram approaching from behind.

Several macro level transit crash frequency models have been developed in previous studies and the most of them focused on bus transit, and evaluated the key factors associated with bus-involved collisions (Goh et al., 2014, Quintero et al., 2013). Only two previous studies have attempted to identify the factors associated with tram-involved collisions; but rather than considering tram-involved collisions solely, they combined both tram and bus-involved collisions to identify the crash causation factors (Shahla et al., 2009, Cheung et al., 2008). Cheung et al. (2008) developed zonal level transit-involved collision prediction models for urban transit in Toronto, Canada using a negative binomial regression structure. The results showed that vehicle kilometres travelled, transit kilometres travelled, arterial road length, transit stop density, percent of near side stops and average posted speed are significant indicators for transit-involved collisions. Shahla et al. (2009) developed a negative binomial crash prediction model for intersections in Toronto, Canada using five years of transit collision data and pointed out that annual average daily traffic (AADT), public transit and pedestrian traffic volumes, turn movement treatments, public transit stop location, mode technology and availability of transit signal priority technology have significant associations with public transit related collisions at signalized intersections.

It is clear from the above discussion that the factors associated with tram-involved collisions at macro level are still unclear. In addition, the previous studies only focused on North America transit network, so very little is known about the validity of this prediction models in other countries, where traffic and transit environment vary considerably. Previous studies adopted the traditional negative binomial structure to model crash counts and ignored the spatial and temporal variations within the observation groups, even though the available crash data followed a cross sectional panel data structure. Clearly, these are the areas worthy of further research. The aim of this study is to identify the key factors that influence tram-involved crash frequency at the route-section level. For analytical preciseness, the random effects negative binomial (RENB) modelling approach was employed that can account for spatial and temporal variations within the observation groups in cross sectional panel count data structures.

This paper is structured as follows. The next section provides the details of data used in this study. Development of the statistical regression model is then provided. The paper closes with a discussion of the research findings and conclusions.

2. Data

Crash data used in this study was obtained from “Tram Incident Database” which is the ‘Yarra Trams’ crash reporting system (Yarra Trams, 2014). The database contains all tram-involved incidents which happened along all tram routes in Melbourne. The incidents are reported in two categories as ‘A’ and ‘B’ type incidents. The category ‘A’ type incident represents serious injury or death. Police and Yarra Trams authority always investigate this crash type after the crash has been reported by the respective tram driver. Whereas the category ‘B’ type incident represents minor incident and tram drivers are responsible for
Examining route section level tram-involved crash frequency using the random effects negative binomial model

reporting this type of incident, and Police and Yarra Trams authority usually do not investigate this type of incident. For the present study a total of 1,177 tram-involved all injury crashes for 5 years (between years 2009 to 2013) were analysed which occurred along 101 tram route sections of a selected 7 tram routes (tram routes number 1, 6, 19, 67, 70, 96 and 112). Each of the selected tram routes covers both inner and middle Melbourne, and comprises of tram route sections which have some tram lane and signal priority measures as well as some platform tram stops. Tram signal priority includes hook turns and turn bans for general traffic and ‘T’ lights for tram exclusive movement. Whereas, tram lane priority includes tramways, full-time and part-time tram lanes where traffic is excluded at selected times of the day (usually the peak).

Among 1,177 tram-involved crashes, 93% were reported as tram to other road vehicles collisions including cars, cyclists, motor cyclists; whereas 5% were tram to pedestrian collisions and 2% were tram to tram collisions. The tram incident database includes the time and date of incident, sign posts location of incident, tram route number, direction of travel and several tram driver and tram related features.

Depending upon the signpost locations tram route sections have been defined for this study i.e. the distance between two consecutive signposts has been defined as one section of tram route. Tram drivers are responsible to report an incident at a sign post location that occurred in the last route section they travelled. Figure 1 shows the selected 7 tram routes in Melbourne with sign post locations.

FIGURE 1: Location of Selected Tram Routes in Melbourne for present study

Traffic volume data along selected tram route sections have been extracted from the VicRoads’ information system (VicRoads, 2014). In addition to that tram service frequency, stop density and average speed of tram routes were collected form the Yarra trams network development information system (Yarra Trams, 2011). Locations of platform stops and location of tram lane and signal priority were obtained from Yarra Trams authority (Yarra Trams Database, 2014) and VicRoads.
Table 2 provides a brief description and summary statistics of covariates used in the crash frequency model for the present study. Key features of the inputs are that:

- On average there were 2.33 crashes per annum (p.a.) per section ranging from a minimum of zero to a maximum of 12 p.a.
- Traffic volume (AADT) was an average of 9,585/link but ranged from 1,100 to a maximum of 36,000.
- The average route section length was 890m, tram frequency was 672 trams a week and speed was 15.85kph. The average stop spacing was 250m (maximum 610m/link).
- The average platform stop ratio was 33%. The platform stop ratio was calculated as the number of platform tram stops among the total number of tram stops along a selected route section.
- Lane priority was available on 62% of links tested, while the ratio of intersections with tram signal priority was 44%.

### TABLE 1: Summary statistics of variable in RENB model

<table>
<thead>
<tr>
<th>Panel Structure Variables</th>
<th>Parameters</th>
<th>Min</th>
<th>Max</th>
<th>Average</th>
<th>St. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year (2009=1, 2010=2, 2011=3, 2012=4, 2013=5)</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Section (section 1=1 to section 101=101)</td>
<td>1</td>
<td>101</td>
<td>51</td>
<td>29.18</td>
</tr>
<tr>
<td>Crash Frequency Model Variables</td>
<td>crash frequency (collisions/year)</td>
<td>0</td>
<td>12</td>
<td>2.33</td>
<td>2.36</td>
</tr>
<tr>
<td>Traffic volume (AADT) of section *</td>
<td>Traffic volume (AADT) of section *</td>
<td>1,100</td>
<td>36,000</td>
<td>9,585</td>
<td>6,001</td>
</tr>
<tr>
<td>Section Length (km)</td>
<td>Section Length (km)</td>
<td>0.1</td>
<td>2.45</td>
<td>0.89</td>
<td>0.61</td>
</tr>
<tr>
<td>Service frequency (Number of trams/week)</td>
<td>Service frequency (Number of trams/week)</td>
<td>517</td>
<td>911</td>
<td>671.62</td>
<td>125</td>
</tr>
<tr>
<td>Average Speed (Km/hr)</td>
<td>Average Speed (Km/hr)</td>
<td>15</td>
<td>17</td>
<td>15.85</td>
<td>0.64</td>
</tr>
<tr>
<td>Stop Spacing (km/stop)</td>
<td>Stop Spacing (km/stop)</td>
<td>0(^b)</td>
<td>0.61</td>
<td>0.25</td>
<td>0.09</td>
</tr>
<tr>
<td>Platform stop ratio</td>
<td>Platform stop ratio</td>
<td>0</td>
<td>1</td>
<td>0.33</td>
<td>0.42</td>
</tr>
<tr>
<td>Presence of tram lane priority (yes= 1, no =0)</td>
<td>Presence of tram lane priority (yes= 1, no =0)</td>
<td>0</td>
<td>1</td>
<td>0.62</td>
<td>0.49</td>
</tr>
<tr>
<td>Ratio of intersections with tram signal priority</td>
<td>Ratio of intersections with tram signal priority</td>
<td>0</td>
<td>1</td>
<td>0.44</td>
<td>0.41</td>
</tr>
</tbody>
</table>

\(^a\) The weighted average method was applied to compute the AADT value for sections that comprise more than one road segments.

\(^b\) One tram route section with length of 100m does not include any tram stop at the beginning or end.

### 3. Methodology

The aim of this research is to identify the key factors that influence tram-involved crash frequencies along tram route sections. A wide range of crash frequency models have been developed over past decades to gain a better understanding of the factors that affect crash causality and frequency (Washington et al., 2010). There are several strengths, weaknesses and flexibility associated with all crash frequency models. Crash count is considered to be discrete, non-negative, infrequent and likely to be over-dispersed. The most commonly used crash data modelling approaches are poisson and negative binomial (NB) regression models.
Examining route section level tram-involved crash frequency using the random effects negative binomial model

(Washington et al., 2010). However the poisson model cannot deal with over-dispersed crash data, while the negative binomial model can successfully model count data with over-dispersion. Over-dispersion in crash counts occurs when the variance exceeds the mean. If a significant amount of zero crash observations are found for any road entity for a specific period of time, the most popular model used by transport safety analyst is the zero inflated poisson or negative binomial models (Washington et al., 2010, Lee and Mannering, 2002, Shankar et al., 2003, Lord et al., 2005). Several studies have been identified the problems associated with the zero inflated models due to long term mean equal to zero (Lord et al., 2007, Malyshkina et al., 2009), as this model cannot properly reflect the crash-data generating process (Lord and Mannering, 2010).

When crash data is collected from N (1, . . , N) locations for T (1, . . , T) time periods, the traditional poisson and negative binomial model assume that there are N*T independent observations. However, under this circumstance the crash data may contain random location specific effects and temporal correlations, and the best way to model these data is by treating them in a time series cross sectional panel data structure with N location groups and T time periods.

Several methods have been developed in previous studies to model crash frequency for panel crash data. For instance the random and fixed effects models (Chin and Quddus, 2003, Hausman et al., 1984, Shankar et al., 1998), random parameter models (Ukkusuri et al., 2011, Anastasopoulos and Mannering, 2009), mixed effects models (Goh et al., 2014, Baayen et al., 2008), multinomial models (Guo, 1996, Ulfarsson and Shankar, 2003), bivariate/multivariate models (Miaou et al., 2005, Song et al., 2006) and multilevel models (Jones and Jørgensen, 2003, Kim et al., 2007). However each of these models have strengths and weaknesses depending upon sample size, estimation process and transferability to other datasets (Lord and Mannering, 2010).

In addition to the statistical regression models, the neural network and Bayesian network models are also used by previous studies which do not require any functional form to link the dependent and independent variables (Xie et al., 2007, Chang, 2005), but criticized for being black boxes because they cannot generate explicit functional relationships and statistically interpretable results (Xie and Zhang, 2008).

For the present study the random and fixed effects negative binomial models have been considered due to their easy estimation process and transferability to other data set and can handle panel and over dispersed crash data. Hausman et al. (1984) examined both random effects (RE) and fixed effects (FE) models for panel data analysis to research and development of patents. The random effects (RE) model assumes that the location specific unobserved effects (αi) are distributed over the spatial/temporal units and they are independent of regressors (explanatory variables). However, the fixed effects (RE) model assumes that the unobserved effects are accounted by indicator variables and correlated with regressors (explanatory variables) (Lord and Mannering, 2010). The Hausman specification test (1978) is also available to test whether the fixed or random effects model is appropriate and this test has been adopted for this study.

The random effects negative binomial (RENB) model has been deployed for this research which incorporates a random location and time specific effects on parent NB model by assuming that over dispersion parameter is randomly distributed across groups without considering constant across locations (Shankar et al., 1998).

The structure of the RENB model used in this study is as follows:

\[ E(A_{it}) = \exp(\beta X_{it} + u_i + \varepsilon_{it}) \]  

(1)

where \(E(A_{it})\) represents the predicted number of crashes along tram route section i in year t, \(X_{it}\) is the vector of explanatory variables, \(\beta\) is the vector of estimable parameter, \(\varepsilon_{it}\) is the
vector of residual errors, and \( u_i \) represents the random effects for \( i^{th} \) location group and \( \exp(u_i) \) is gamma distributed with mean 1 and variance \( \alpha_i \), where \( \alpha_i \) is the over dispersion parameter in the NB model. The RENB model allows over dispersion parameter to vary randomly across groups and assume that \( 1/(1+ \alpha_i) \) follows beta distribution \( \text{Beta}(r, s) \) (Hausman et al., 1984). Estimation of two distribution parameters \( r \) and \( s \), and model parameter \( \beta \) are done by maximum likelihood technique.

Model's goodness-of-fit has been assessed by \( R^2_\alpha \) as proposed by Miaou et al. (1996). \( R^2_\alpha \) determines how well the variance of data is captured by the model compared to fundamental model with no variables.

\[
R^2_\alpha = 1 - \frac{\alpha}{1 + \alpha_{max}}
\]

(2)

where \( \alpha \) is the estimated over dispersion parameter for the selected model and \( \alpha_{max} \) is the estimated over dispersion parameter for the fundamental model containing only constant term.

4. Results

Table 3 presents the parameter estimates obtained from the random effects negative binomial (RENB) model using maximum likelihood technique in STATA version 13 (STATACorp. 2013, 2013) statistical software.

**TABLE 2: RENB model results for tram crash frequency**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimated value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.92</td>
<td>0.035</td>
</tr>
<tr>
<td>Ln(AADT)</td>
<td>0.17</td>
<td>0.026</td>
</tr>
<tr>
<td>Ln(Section Length)</td>
<td>0.31</td>
<td>0.000</td>
</tr>
<tr>
<td>Services per week</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Average Speed</td>
<td>0.09</td>
<td>0.036</td>
</tr>
<tr>
<td>Stop Spacing</td>
<td>-0.43</td>
<td>0.038</td>
</tr>
<tr>
<td>Platform stop ratio</td>
<td>-0.09</td>
<td>0.043</td>
</tr>
<tr>
<td>Tram lane priority (yes=1)</td>
<td>-0.148</td>
<td>0.032</td>
</tr>
<tr>
<td>Proportion of tram signal prioritised intersection</td>
<td>-0.263</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Likelihood-ratio test vs. pooled: chibar2= 67.08: Prob>=0.000
Hausman Test: chi2(8) = 10.76: Prob>chi2 = 0.2154>0.05

Log-likelihood: \(-893.42\)
Dispersion parameter, \( \alpha \): 0.272
95% CI for \( \alpha \): 0.187~0.394
Vuong test of zinb vs. standard negative binomial: \( z = 1.02 \) Pr>z = 0.1549
\( R^2_\alpha \): 0.84

The dispersion parameter estimate was found to be significantly different from zero (\( \alpha = 0.272 \)), which suggests that the negative binomial model structure is more suitable than the poisson structure. The ‘Vuong’ test was carried out to evaluate the appropriateness of the
Examining route section level tram-involved crash frequency using the random effects negative binomial model

zero-inflated negative binomial over the negative binomial regression model and an insignificant z-test value (p=0.154) indicated that the zero-inflated model is not appropriate for this data modelling (Shankar et al., 1997). The likelihood-ratio test indicated that the panel estimation is significant compared to the pooled estimation (STATA, 2014). In addition, the Hausman specification test is insignificant, which suggests using the random effects model instead of the fixed effects model. The value of $R^2$ was found as 0.84, which is more than 0.7 and indicates that the most of the necessary covariates are included in the model (Miaou et al., 1996).

The analysis results indicated that tram-involved crashes considerably decrease with the increase in tram stop spacing ($\beta=-0.43$). The presence of more tram stops along tram route sections i.e. less spacing between stops would mean trams have to brake and accelerate at stops more frequently and increase the chance of collisions between trams and other road users. This finding is in agreement with previous study, where transit stop density was found to be positively correlated with crash occurrence (Cheung et al., 2008). In mixed traffic tram operating environment, other vehicles have to stop at rear of trams at the older design tram stops i.e. at curb side stops to allow tram passengers' boarding and alighting (Currie et al., 2011). This may increase rear end collisions between tram and other vehicles at/near stops. In addition to that, safety zone stops, the another older design stop, have narrow waiting area for passengers adjacent to a metal barrier at the middle of the road, known to have potential tram-involved crash risks and the most common type of crash risk is passenger being stuck by tram (Currie and Reynolds, 2010). Less number of tram stops along tram route sections can help to reduce above mentioned potential tram-involved crashes.

The analysis results also showed that tram-involved crashes increase with the increase of tram route section length ($\beta=0.31$) and general traffic volume ($\beta=0.17$). These results are as expected, because longer tram route section and higher traffic volume are associated with the higher exposure between trams and other road users; and these have been shown to be reliable predictors of crash frequency by previous study (Cheung et al., 2008).

The model suggested that tram route sections with higher number of signalised intersections with tram signal priority experience less crash rate ($\beta=-0.26$). The result is in agreement with the study conducted by authors (Naznin et al., 2015a, Naznin et al., 2015b), which specified that the presence of tram signal priority increases overall road safety at intersections in Melbourne. However, the result is opposite to the findings by Shahla et al. (2009), who examined transit safety at intersections in Toronto and found that transit-involved collisions increased at intersections due to the presence of tram signal priority. Several tram signal priority features are different in Melbourne compared to Toronto. In Melbourne, separate 'T' light is provided for tram movement, which is a possible reason for reduced crash occurrence, as it creates awareness among other road users that tram is expected to cross the intersection. Also separate 'T' light segregates tram movement from general traffic movement at intersection. In addition, hook turns and turn bans in Melbourne represent potential reasons for reduced crash occurrence between trams and other vehicles, and backed up by Currie and Reynolds (2011). Moreover, crash exposures at intersections along tram routes are higher for Toronto compared to Melbourne, as the tram headways are less than 5 minutes in Toronto, whereas in Melbourne typical tram headways are 7.5 minutes (Currie and Shalaby, 2008). Aforementioned discussions could be the reasons why the introduction of tram signal priority had led to the reduced tram-involved crash occurrence in Melbourne.

The model result also suggested that the crash rate along tram routes with tram lane priority have approximately $\exp(-0.141)$ or 0.86 times of the crash rate for routes without tram lane priority assuming all other variables are constant. That means, the presence of tram lane priority is associated with a 14% reduction in tram-involved crash occurrence at all severity levels. This finding is in agreement with authors previous study (Naznin et al., 2015a, Naznin et al., 2015b), which showed that tram lane priority is 19.4% effective in reducing total crashes along tram routes. Also a recent study by Richmond (2014) suggested
that the presence of dedicated streetcar right-of-way can reduce the rate of collisions. Removing general traffic from tram tracks can help to reduce collisions between trams and other vehicles.

The model also indicated that higher tram average speed increases the tram-involved crash frequencies (β=0.09). Tram which is classified as heavy vehicle requires more stopping distance to stop after braking compared to cars (Candappa et al., 2013). Tram travelling at slower speed reduces the stopping distance and increases the probability that the tram can come to a complete stop prior to collision with other road users.

The analysis result revealed that the presence of platform stops can reduce tram-involved crash frequency (β=-0.09), as the platform stops can alleviate the issues associated with the older design tram stops by providing safe and raised waiting platform for passengers and can reduce the possibility of passengers being hit by trams (Currie and Reynolds, 2010). The analysis result also showed that tram-involved crashes increase with the increase in tram service frequency (β=0.004), because the higher tram service frequency are the associated with higher exposure between trams and other road users.

5. Conclusion

An analysis of tram-involved crash data was carried out using the random effects negative binomial (RENB) model to explore the safety effects of the key traffic, transit and route factors in Melbourne, Australia. The results showed that tram stop spacing, tram route section length, tram signal priority, general traffic volume and tram lane priority have significant association with tram–involved crash occurrences. The average speed of the tram, the proportion of platform tram stops and tram frequency are also linked to tram-involved crash frequency but at a lower level.

The study findings provide useful insights on route section level tram-involved crash occurrence and present useful planning tools for transit agencies. An important finding of this research is that both tram lane and signal priority measures in Melbourne’s context act to improve road safety and should be a major consideration for the road management agencies when implementing tram priority and road schemes.

There are numerous opportunities to carry out further research in this field. In particular, further collection of tram incident data could be undertaken as a mean to improve model validity and possibly identify other factors that can be significant in explaining tram-involved crash frequency. Identifying the key factors associated with different crash severity levels will also be a potential area for further research.

Acknowledgement

This research is a part of wider Australian Research Council Industry Linkage Program project LP100100159, ‘Optimizing the Design and Implementation of Public Transport Priority Initiatives’ Institute of Transport Studies, Monash University in association with the Transport Research Group, University of Southampton, UK. The Principal Chief Investigator is Professor Graham Currie, the Chief Investigator is Associate Professor Majid Sarvi and the Partner Investigator is Dr. Nick Hounsell.

The authors would like to thank the Australian Research Council for financial support of the research and the Yarra Trams for providing access to their ‘tram incident database’.

References

Examsing route section level tram-involved crash frequency using the random effects negative binomial model


CLICHE, D. & REID, S. 2007. Growing Patronage-Think Tram?


