Zenith Model of Victoria

Technical Note 2

Review of VISTA07

Zenith Version 2.0.0

VEITCH LISTER CONSULTING PTY LTD

Date: 26-03-2013

Zenith Model of Victoria

Technical Note 2: Review of VISTA07

Date	Revision	Prepared By	Checked By	Approved By	Description
24-01-2011	A	TV	MV	MV	Draft – Work in Progress - VISTA07 (1.3)
04-04-2011	B	TV	AMA	MV	Draft for comment – revised with the release of VISTA07 (1.3.1)
24-01-2012	C	TV	AMA	MV	Final Report

Contents

1 Introduction

The Zenith travel model of Victoria is one of a family of models developed by Veitch Lister Consulting (VLC) for transport planning in Australian cities and regions.

This document is one in a series of technical notes that collectively describe the Zenith Model of Victoria.

1.1 Related Documents

This technical note is the second of eleven. The other working papers are:

- Working Paper 1: Model Validation Framework and Data Sources
- Working Paper 2: Review of VISTA07
- Working Paper 3: Home Based Trip Production Model
- Working Paper 4: Non-Home Based Trip Production Model
- Working Paper 5: Household Segmentation & Travel Market Segmentation Models
- Working Paper 6: Period Allocation and Vehicle Occupancy Models
- Working Paper 7: Mode Choice Model
- Working Paper 8: Destination Choice and Trip Attraction Model
- Working Paper 9: Overall Model Validation
- Working Paper 10: Backcasting and Sensitivity Testing
- Working Paper 11: Reference Case Model Assumptions

1.2 Scope and Aim of This Document

The aim of the document is to investigate the degree to which the VISTA07 survey reflects, and is consistent with, other observed measurements of travel in Melbourne, including:

- VicRoads 2006 screenline traffic counts
- Estimated rail network demands (2009/10 Rail OD survey),
- Estimated tram network demands (2008 Tram OD survey)
- Estimated bus boardings (based on validations)

We will also compare VISTA07 with other household travel surveys, including:

- The South East Queensland Travel Survey (SEQTS)
- The Sydney Household Travel Survey (Sydney HTS)

During the course of this document, we will also examine the methodology used to weight the VISTA07 data (a methodology developed by The Urban Travel Institute), and propose and describe a set of correction factors which we believe corrects for certain biases in the VISTA07 survey. These weights will form the basis of the Zenith model recalibration.

The remainder of the document is structured as follows:

Section 2 provides a brief overview of the VISTA07 survey methodology.

Section 3 validates the VISTA07 survey against a wide array of other data sources.

Section 4 examines the issue of under-reporting in VISTA07, and compares VISTA07 with the Sydney and South East Queensland household travel surveys.

Section 5 describes the correction factors developed by VLC, and updates the validation of VISTA07 against other data sources.

Section 6 concludes the document.

2 VISTA07 Survey Methodology

2.1 Overview

The 2007/08 Victorian Integrated Survey of Travel and Activity (VISTA07) was undertaken by The Urban Transport Institute (TUTI) between June 2007 and June 2008.

The survey was conducted in Melbourne, as well as a number of regional centres, as shown in Figure 1.

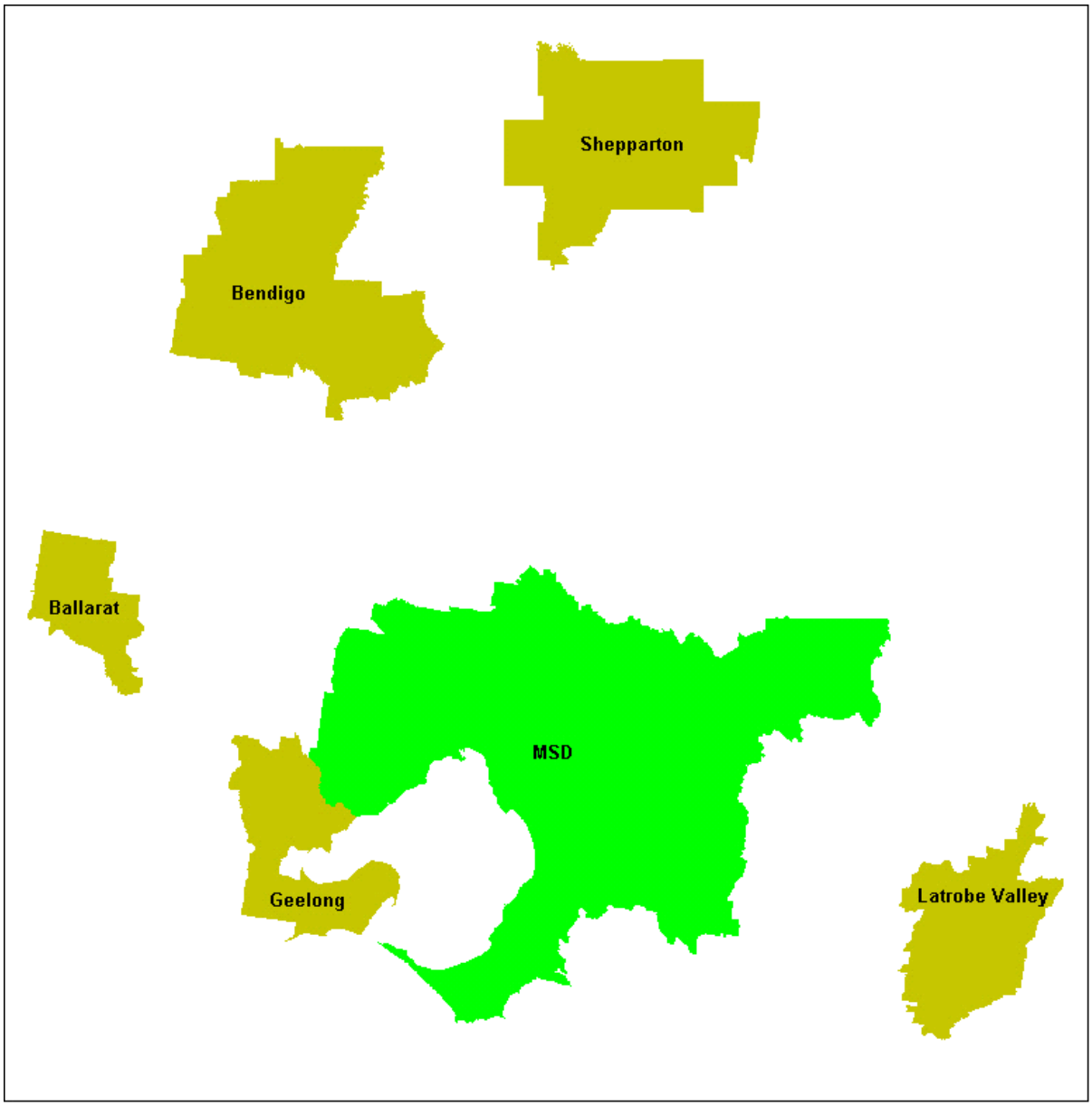

Figure 1 - VISTA07 Study Areas (Source: TUTI - VISTA07 Survey Procedures and Documentation)

Based upon the VISTA07 documentation provided by DOT and prepared by TUTI, VLC understands that:

- The survey was restricted to residents of the above study area, aged 5 and above, and living in occupied private dwellings.
- The survey was "self-completed" by survey respondents, with survey forms hand-delivered and hand-collected by survey staff. Telephone motivational calls, and telephone and postal

reminders were also undertaken to maximise the response rate. Where necessary, follow up telephone calls were also made to clarify survey responses, or to fill in missing information.

- Because of the cost of hand-delivering and collecting survey forms, a multi-stage sampling process was employed. This involved first randomly selecting Census Collector Districts (CCDs), and then randomly selecting a sample of households from each CCD. This enabled survey staff to visit a number of selected households within the same CCD in quick succession. The CCDs were chosen to provide a reasonable sample within each LGA.
- All residents of the household were asked to complete the survey.

The questionnaire itself was designed to capture a wide array of demographic variables, together with details of all trips made by residents of the household, including:

Demographics

- Age
- Gender
- Employment status
- Educational status
- Other activity status
- Occupation
- Industry
- Personal income

Trip Attributes

- Origin location
- Destination location
- Type of place – origin
- Type of place - destination
- Activity undertaken - origin
- Activity undertaken - destination
- Mode of travel
- Start and end time of trip

The overall sample size was 17,300 households, including 43,800 people.

The average response rate of the survey was 47%.

For more detailed information see the TUTI report: *"VISTA07 Survey Procedures and Documentation"*.

2.2 Survey Weighting

2.2.1 Person Weights

In total, the VISTA07 survey was completed by 43,800 people, which is approximately 1% of the 2006 population for the surveyed region (4,095,513 people).

For many applications, such as the comparison of VISTA07 with other measures of travel, or the derivation of population wide statistics, it is useful to "factor" or "weight" up the survey responses to the total population. A simple approach in this case would be to uniformly factor up all respondents by 100 (as it is a 1% sample).

However, the VISTA07 survey is not (and was not designed to be) entirely representative of the population, either spatially or socio-demographically. As such, a more refined set of weights can be derived by applying greater weights to person types that are under-represented in the sample.

Individual person weights were derived using a two-step expansion process which is now described.

2.2.1.1 Calculation of Household Weights

Each household was assigned a household weight, such that the weighted households would be spatially and demographically representative of households in the survey study area in 2006.

Three control variables were used to expand households:

- Home LGA
- Dwelling Type (2 types: Separate House and Other Type)
- Dwelling Ownership (2 types: Owned & Buying, and Renting & Other)

For each combination of Home LGA, Dwelling Type and Dwelling Ownership, the number of surveyed households was compared with the number of households of this type in the ABS Census 2006, with the ratio of these two quantities determining the household weight.

As a result, the weighted households should be consistent with the ABS Census in terms of the number of households of each dwelling type and ownership status for each LGA.

2.2.1.2 Calculation of Person Weights

Each person in each household was initially assigned a weight equal to their corresponding *household weight*. Individual person weights were then adjusted such that the weighted persons were spatially and demographically representative of the person types found in the ABS Census 2006.

Three control variables were used to adjust person weights:

- Home LGA
- Gender (2 types: Male and Female)
- Age Group (18 groups: 5-year groupings from 0-4 up to 85+)

As a result, the weighted persons should be consistent with the ABS Census 2006 in terms of the number of people of each gender and age group for each LGA.

2.2.2 Adjusted Trip Weights

The household and person weights described in the previous sections are designed to ensure that the weighted survey is spatially and demographically representative. TUTI have also calculated trip weights which are designed to correct for certain types of under reporting of travel.

In their report: *"VISTA07 Survey Procedures and Documentation"*, TUTI discuss three types of non-reporting of travel:

1. Deliberate non-reporting of travel, "either because they thought they were unimportant or because they wanted to minimise the effort in completing the travel diary",
2. Under reporting due to "proxy reporting", where the travel diary of an individual is filled out by another household member. The other household member may not have full knowledge of the individual's travel, and
3. Under reporting due to forgetfulness or oversight, which is more likely to occur when there is a significant delay between the travel day and when the travel diary is completed.

In light of these issues, VISTA07 included questions about who filled out each travel diary (targeting type 2 under-reporting), and the date when the travel diary was actually completed (targeting type 3).

TUTI's analysis of the survey responses found that travel diaries filled out by proxy exhibited lower trip rates on average. They also found that travel diaries not completed on the travel day generally exhibited lower average trip rates.

There is the potential, of course, for this to be due to demographic factors, rather than under-reporting. For example, it could be that people who do not fill out the form themselves (e.g. young children), do not make as many trips. To account for this, the effect was explored for different groupings of age and gender to see if the effect was present within each demographic group.

The effect was found, and adjustment factors were calculated for each combination of age, gender, travel mode and trip type (home based, and non-home based). In each demographic category, those who filled the form out themselves and on the travel day were considered to be the point of truth, with trip rate adjustment factors calculated relative to this group.

3 Survey Validation

A key step in the recalibration of the Zenith model has been the "validation" of the VISTA07 survey – i.e. establishing the degree to which the survey accurately reflects reality.

Our general approach has been to compare VISTA07 with other data sources which can provide a more accurate estimate of particular aspects of travel in Melbourne. These data sources have included:

- VicRoads 2006 screenline traffic counts
- The 2009/10 Rail OD Survey (which included 7am-7pm station counts)
- The 2008 Tram OD survey
- 2008 estimates of bus route boardings (based on validations)

Together, these data sources provide a comprehensive view of motorised travel in Melbourne; pedestrian and cycling counts would complete the picture, but are not readily available at this time.

Given that VISTA07 is only a sample survey, weights have been used to scale VISTA07 responses to "population level" demands. Naturally, the weighted survey is "lumpy", but it is nevertheless useful at higher levels of aggregation.

As described in Section 2, TUTI have calculated weights for VISTA07, including person weights which can be used to scale VISTA07 trips to population level demands, and adjusted trip weights, which are an attempt by TUTI to correct for certain types of under-reporting.

We have validated both weighting methodologies, to explore the role played by the adjusted trip weights.

3.1 VicRoads 2006 Screenlines

The weighted VISTA07 survey has been assigned to the Zenith 2008 road network, with the resulting loads compared with VicRoads 2006 Screenline Traffic Counts. VISTA07 has been weighted to 2006, which makes for a fair comparison.

Three key assumptions have been made during our comparisons:

- VISTA07 data has been restricted to weekday, during school term times, and excluding public holidays. The weights have been recalculated accordingly. This is consistent with our understanding of when the VicRoads screenlines are conducted.
- The Zenith traffic assignment process has been used to determine routes for the VISTA07 trips; separate assignments were performed for AM Peak (7-9am), PM Peak (4-6pm) and the rest of the day.
- Zenith model estimates of commercial vehicle traffic and visitor flows have been added to the VISTA07 numbers. This is because the screenline counts include commercial vehicles and visitors, but VISTA07 does not.

3.1.1 By Hour of the Day

An initial comparison between VISTA07 (including Zenith CVs and visitors) and the screenline counts is provided in Figure 2. The comparison is performed by hour of the day (where each VISTA07 trip is allocated to a single hour using its start time). Both the person weighted, and trip weighted VISTA07 data are presented.

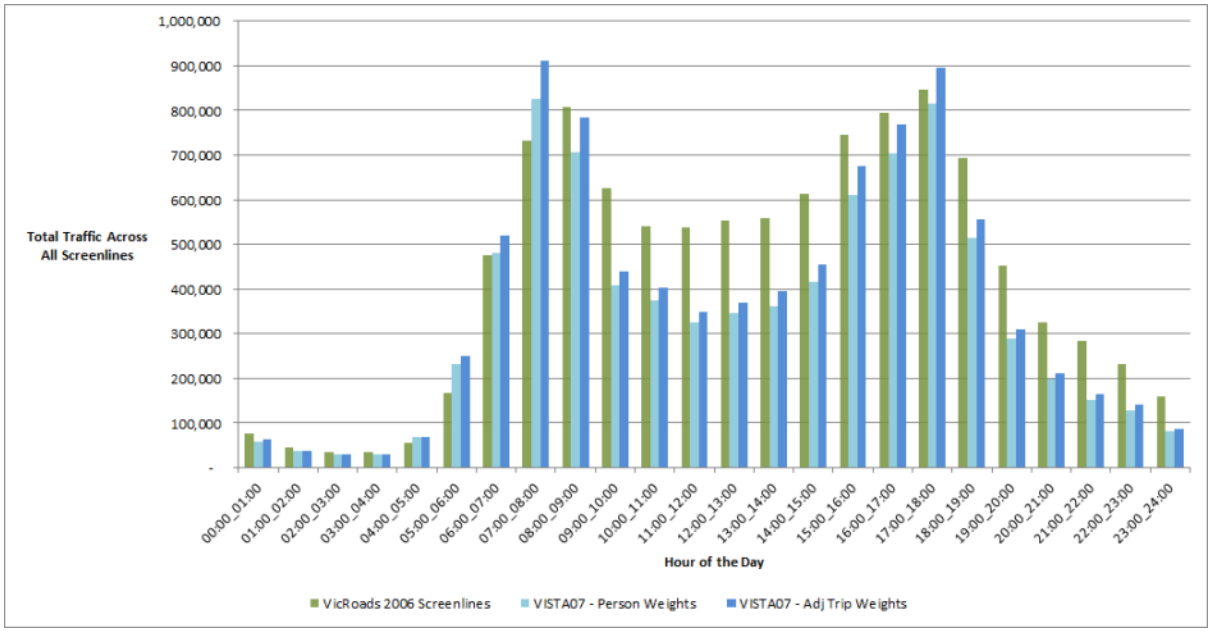

Figure 2 - Comparison of Total Screenline Counts and VISTA07 plus Zenith Visitors and CVs

Until 9am, VISTA07 and screenline counts match quite well, with the screenline counts marginally higher before 4am, and the VISTA07 survey higher during the period 4am – 8am. The person weighted VISTA07 matches screenline counts best during this period – it seems that the adjusted trip weights cause an "over-reporting" of travel in the AM peak.

During the middle of the day (9am – 3pm), VISTA07 is consistently lower than screenline counts – by 35% in the case of person weighted VISTA07, and by 30% with the inclusion of the adjusted trip weights.

Between 3pm and 6pm (the school and work peak), VISTA07 again matches screenline counts well. The person weighted VISTA07 are generally 10% lower than screenline counts, while the adjusted trip weights match screenlines well.

However, from 6pm onwards, VISTA07 is again lower than screenline counts by 30-40%.

Based on these comparisons, it appears that the VISTA07 survey accurately reflects travel during the peaks, but heavily under-reports during the rest of the day.

To understand just how low VISTA07 is during the off peak, recall that the Zenith model's estimates of commercial vehicle and visitor flows have been added to the VISTA07 survey in the above analysis. These CVs and visitors make up 30% of the total VISTA07 flow shown in Figure 2 above. If we remove the CVs and visitors, then the pure VISTA07 accounts for only half of all traffic crossing screenlines during the interpeak, as seen in Figure 3 below. While this is not a fair comparison (as the screenline counts include CVs and visitors), it illustrates that VISTA07 provides only a small base of trips to "factor up". In fact, to match screenline counts during the interpeak, we have to factor VISTA07 up by an average of 1.7.

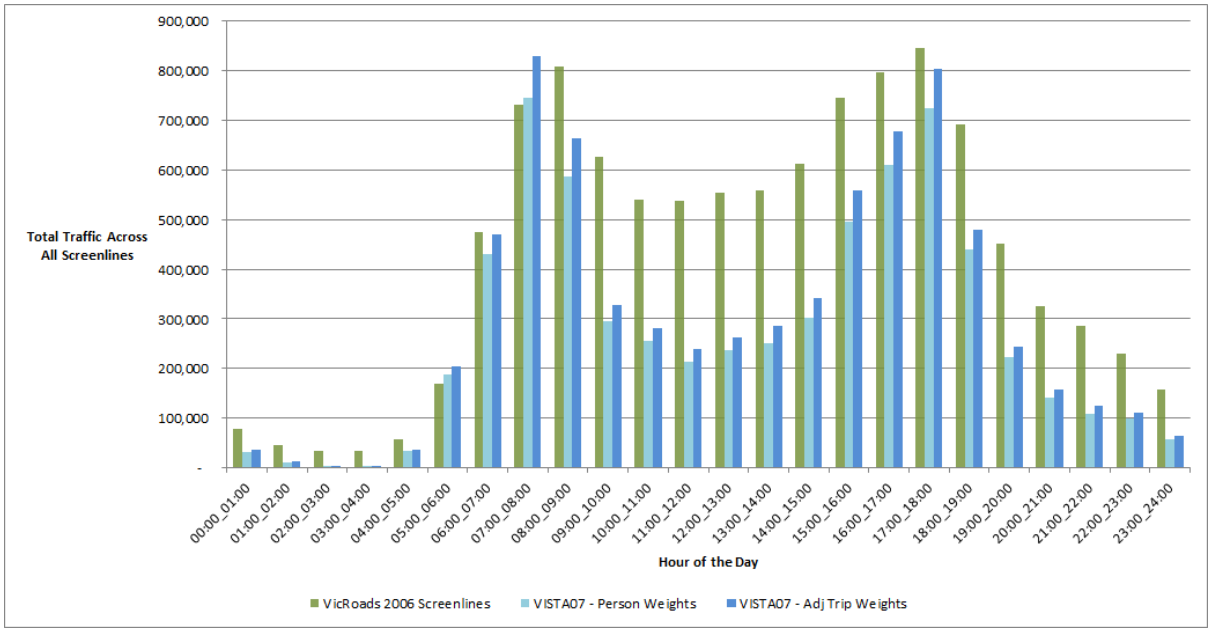

Figure 3 - Comparison of Total Screenline Counts and VISTA07

Of course, we are making an implicit assumption here that the Zenith model's estimates of CV and visitor travel are robust. To support this assumption, we have compared the Zenith model's prediction of CV flows with counted CVs, where classified counts exist. The comparison, which is presented in Figure 4 below, shows that the Zenith model tends to over-predict commercial vehicle traffic on low volume roads, and is approximately correct on larger roads. Certainly there is no evidence of the Zenith model under-predicting commercial vehicle traffic.

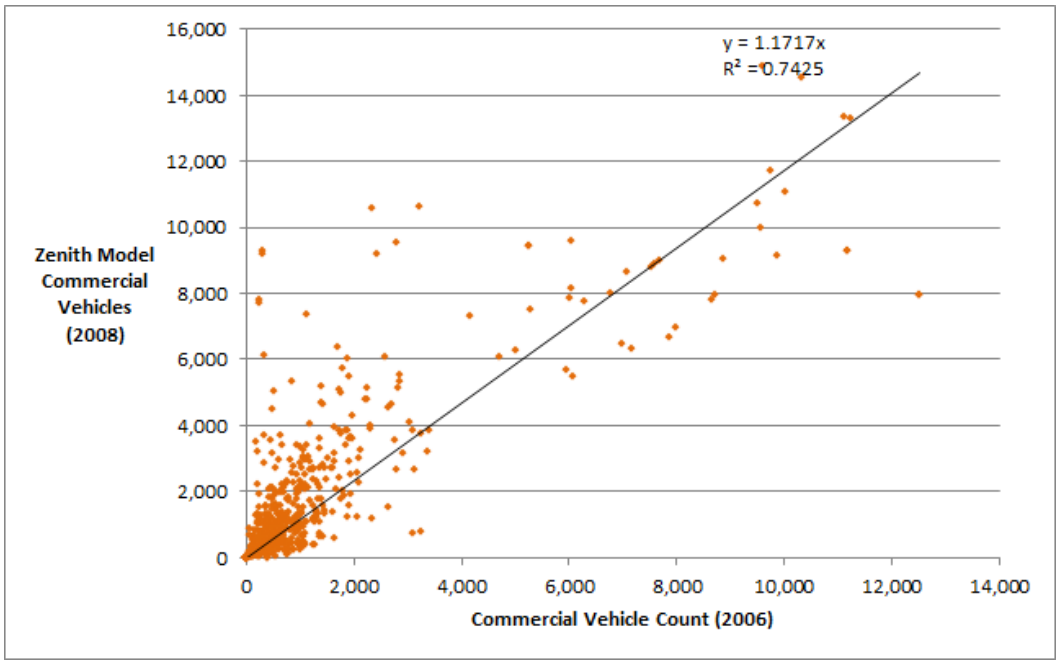

Figure 4 - Modelled vs. Counted Commercial Vehicles (24hr).

Less information exists for visitors, but we think it unlikely that they would be substantial enough to dramatically influence our analysis.

There are also other potential sources of traffic which might be counted at the screenlines, but excluded from VISTA07, including:

- "Dead running" taxis – i.e. taxis with no passenger
- Public transport vehicles (i.e. busses, trams)
- Intrazonal trips in VISTA07 (these trips have their origin and destination in the same travel zone. There are intrazonal trips in VISTA07, but they have not been assigned as part of this process)

While these markets may add a few percent to the traffic stream, we don't believe that they can explain the trips which appear to be missing from VISTA07 in the off peak. This assertion is supported by the fact that these markets also exist in the peaks, and yet VISTA07 accurately matches peak screenline counts.

Given all of the above, we conclude that VISTA07 is substantially under-reporting off peak traffic made by *residents of the study area*, perhaps by up to 30 - 40%.

3.1.2 By Screenline

In the previous section we concluded that VISTA07 consistently under-reports travel during the off-peak. In this section we explore whether there is a pattern among different screenlines.

In 2006, VicRoads collected traffic counts on 21 screenlines. The total counted traffic across each screenline has been compared with the corresponding VISTA07 traffic flow (including Zenith CVs and visitors, and using the VISTA07 adjusted trip weights).

Our analysis is presented in Figure 5, Figure 6 and Figure 7 below.

Two key observations can be made:

- There is remarkable *correlation* between the counted traffic and VISTA07 traffic across all screenlines, with an R-Squared in excess of 0.97 in all periods. This is a very positive result.
- Consistent with the previous section, VISTA07 over-reports the AM peak by about 13%, is about right in the PM peak, and is approximately 30% low in the inter peak.

In the inter peak especially, it is remarkable that such a high correlation is achieved, despite VISTA07 missing 30% of traffic. It suggests that the "missing traffic" – whatever the source – is at least systematically missing across all screenlines, and during all the inter peak hours. This rules out a lack of sample in VISTA07 as a plausible explanation, and hints at a more systematic effect.

In the AM peak, it is a little surprising that VISTA07 over-reports by 13%. We believe that this may be caused by the "correction factors" which are built into the adjusted trip weights. These correction factors attempt to correct for exactly the type of under-reporting that we are seeing in the off peak. However, the correction factors applied by TUTI do not differentiate between time periods, and so we believe that the correction factors are "over correcting" in the AM peak, and "under correcting" in the off peak. If we instead use the person weights (which have no built in correction factor), then VISTA07 is 2% higher than screenline traffic counts – a good result.

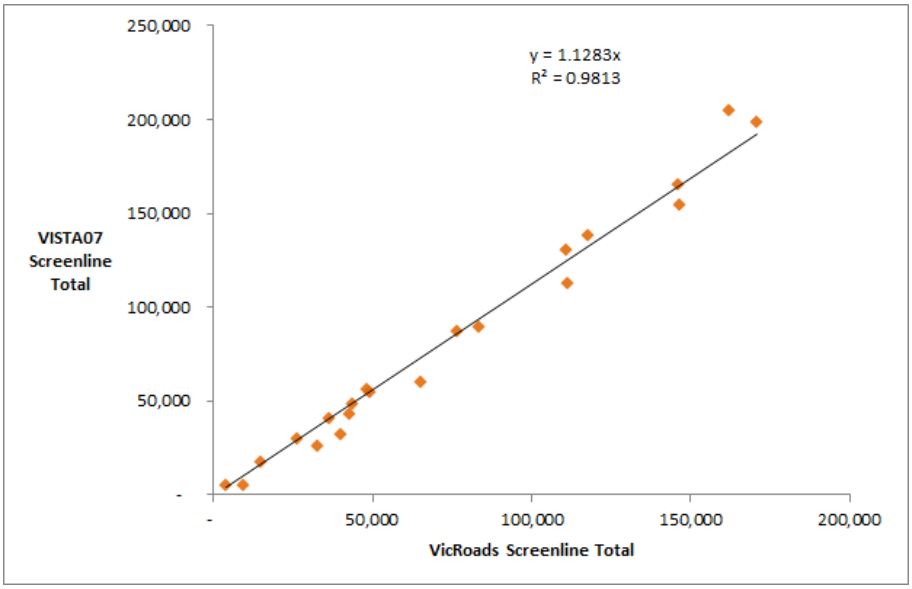

Figure 5 - Comparison of VicRoads VISTA07 Screenline Totals (7-9am)

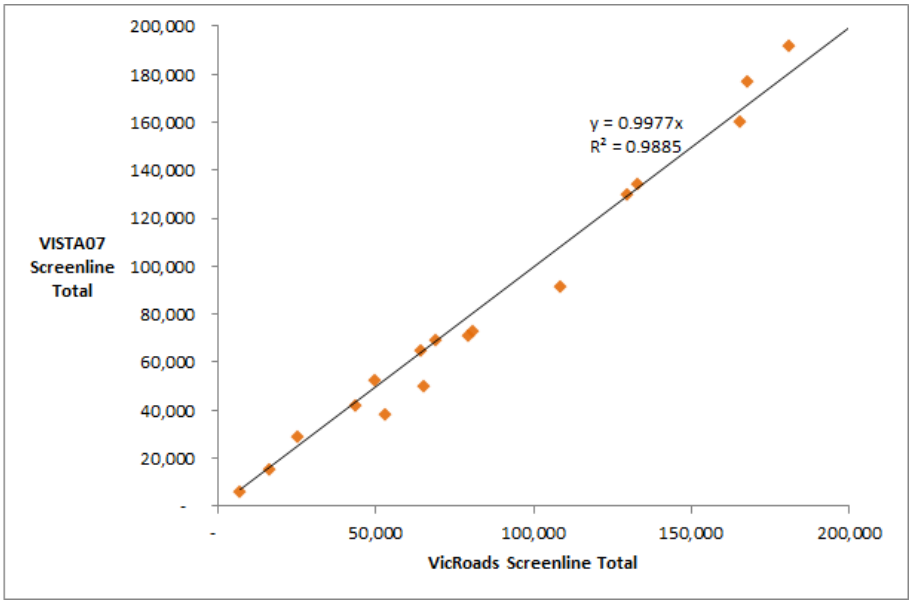

Figure 6 - Comparison of VicRoads and VISTA07 Screenline Totals (3-6pm)

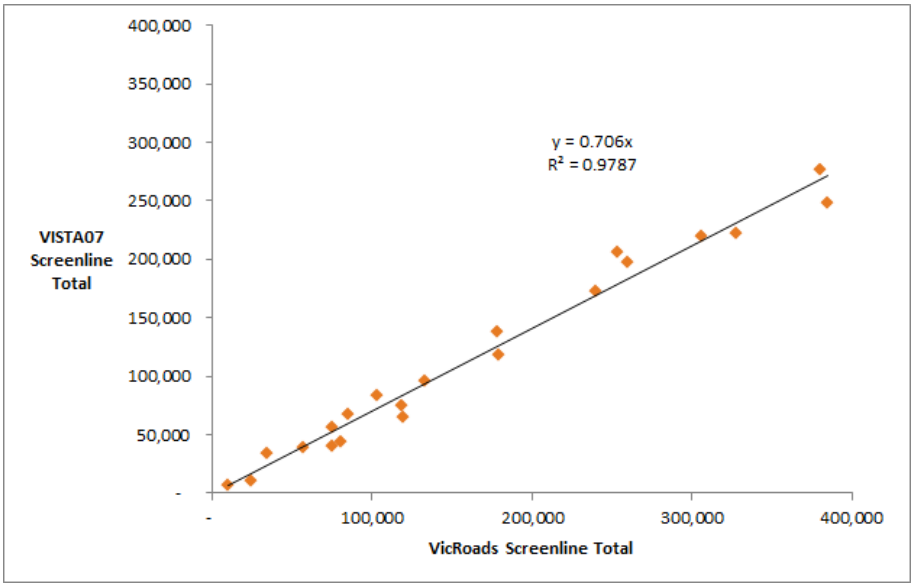

Figure 7 - Comparison of VicRoads and VISTA07 Screenline Totals (9am-3pm)

3.1.3 By Count Location

We have further explored the VISTA07 survey by comparing it with individual screenline traffic counts.

Given that VISTA07 is only a 1% sample (0.5% once school holidays, weekends and public holidays are removed), it would not be surprising in the least if there was considerable variation at the level of individual counts. However, this turns out not to be the case.

Figure 8, Figure 9 and Figure 10, below, present comparisons between VISTA07 traffic flows and individual screenline traffic counts, for the AM peak, PM peak and inter peak respectively. Consistent with earlier analysis, VISTA07 tends to be high in the AM peak, about right in the PM peak, and is quite low in the inter peak.

There are, however, some interesting patterns:

- The correlation between VISTA07 and screenline traffic counts in the AM peak is excellent (R-Squared of 0.88). This is very positive.
- In the PM peak, VISTA07 noticeably over-predicts on roads with large flows (i.e. major roads with flows excess of 9000 in the period 3-6pm). This is compensated by under-prediction on roads with small flows.
- A similar effect is found in the inter peak. While VISTA07 is on average 30% lower than screenlines in the inter peak, it is consistently *less low* on roads with large flows (i.e. major roads), and consistently *more low* on small roads. A dashed 45 degree line has been marked on Figure 10 to highlight this. This finding provides us with some clues as to what kind of traffic is missing during the inter peak.

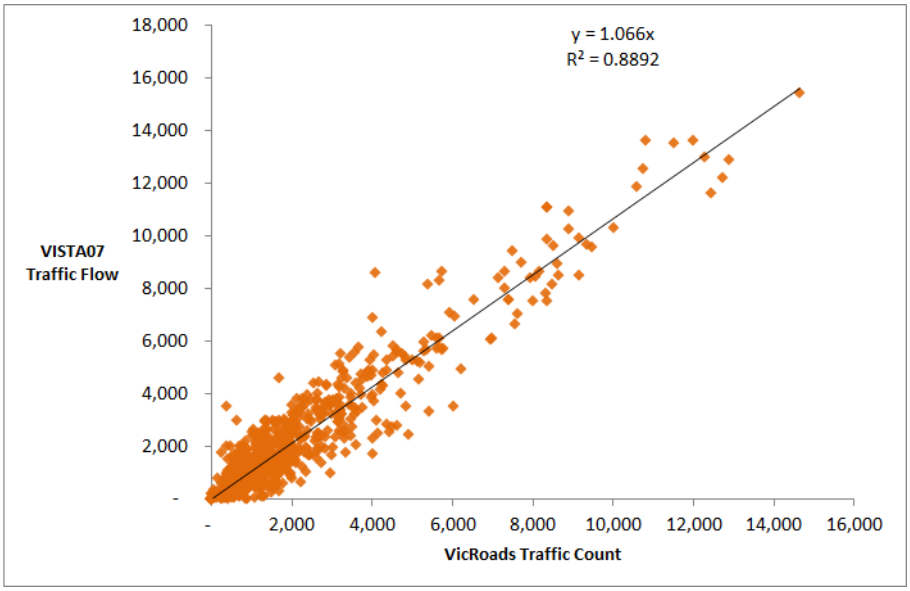

Figure 8 - VISTA07 vs. VicRoads Screenline Traffic Counts (7am-9am)

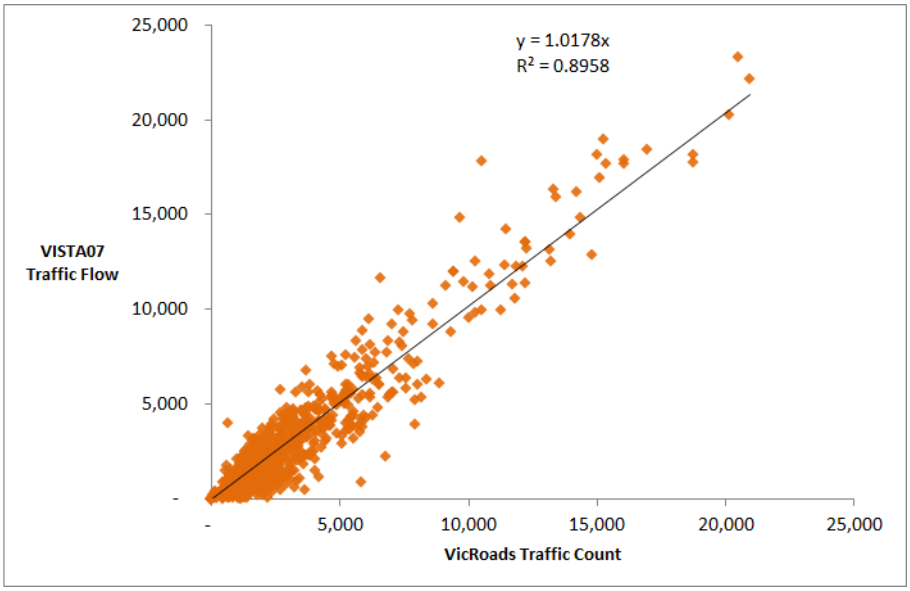

Figure 9 - VISTA07 vs. VicRoads Screenline Traffic Counts (3pm-6pm)

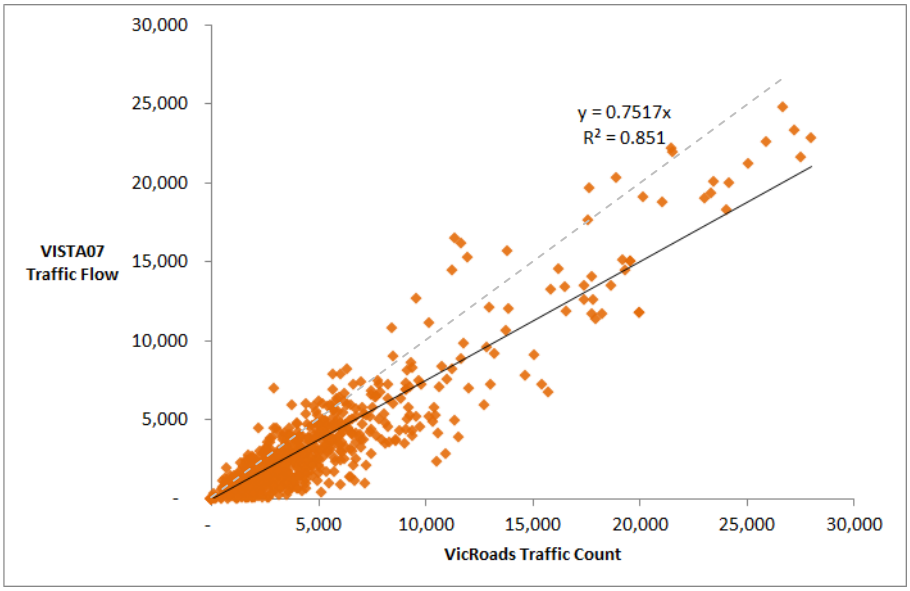

Figure 10 - VISTA07 vs. VicRoads Screenline Traffic Counts (9am-3pm)

We have also mapped the VISTA07 and screenline traffic flows, to explore any spatial patterns. Figure 11, Figure 12 and Figure 13 present the results.

We make the following observations:

- Referring to Figure 11, we can see again that VISTA07 tends to marginally over-report in the morning peak (again, we believe this to be an issue with the correction factors). Overall, however, the match between VISTA07 and screenline counts in the AM peak is superb, given the VISTA07 sample.
- Referring to Figure 12 we can again see that VISTA07 tends to slightly over-predict on major roads (i.e. major freeways) in the PM peak. Overall, however, the match is excellent.
- Referring to Figure 13 we can again see that VISTA07 under-reports travel in the inter peak. The pattern is consistent across the entire network. Figure 18 presents the same comparison, but without freeways, to highlight the pattern on non-freeway roads.

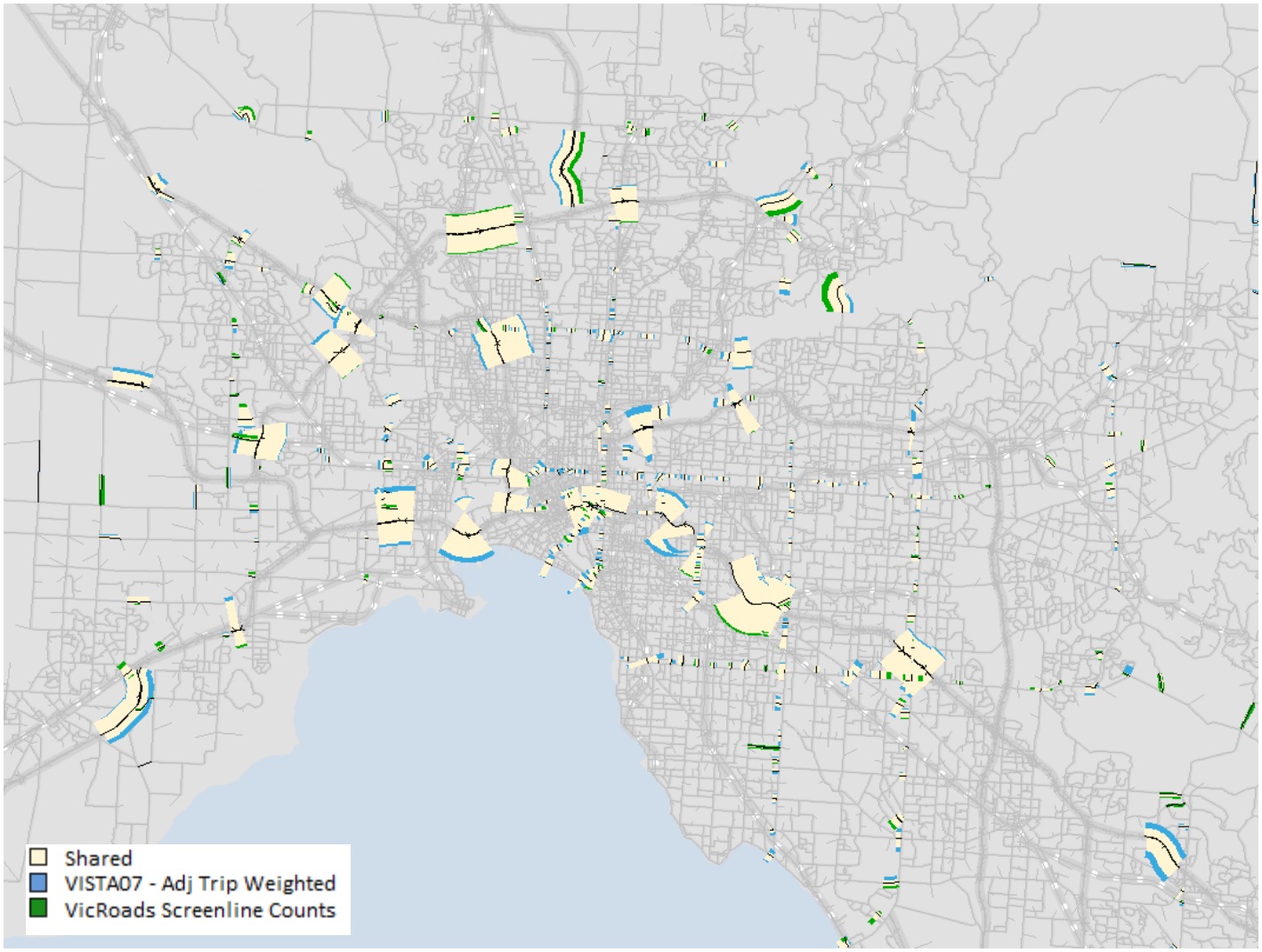

Figure 11 - Difference in Load by Screenline Count Location (AM Peak, 7am-9am)

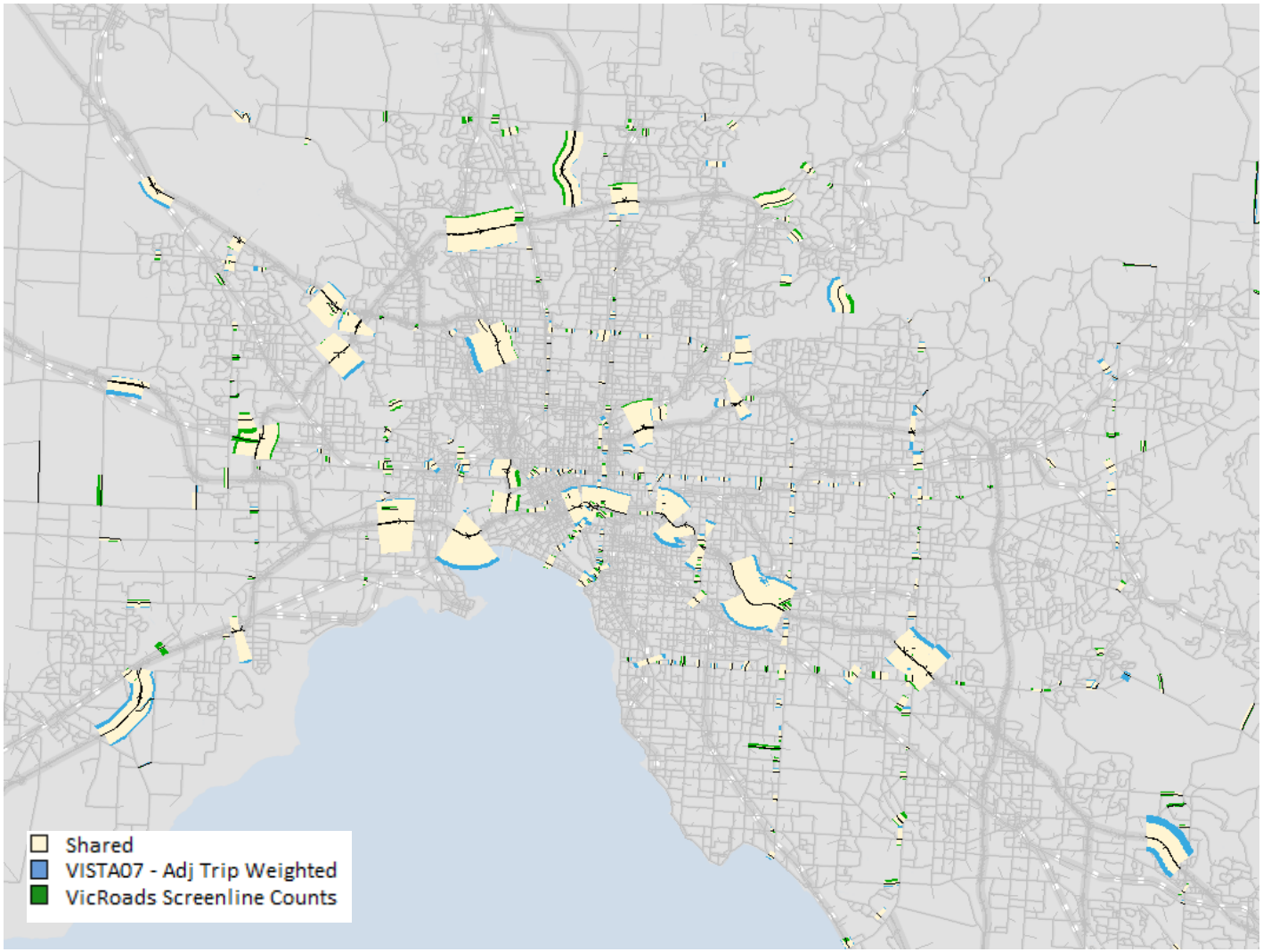

Figure 12 - Difference in Load by Screenline Count Location (PM Peak, 3pm-6pm)

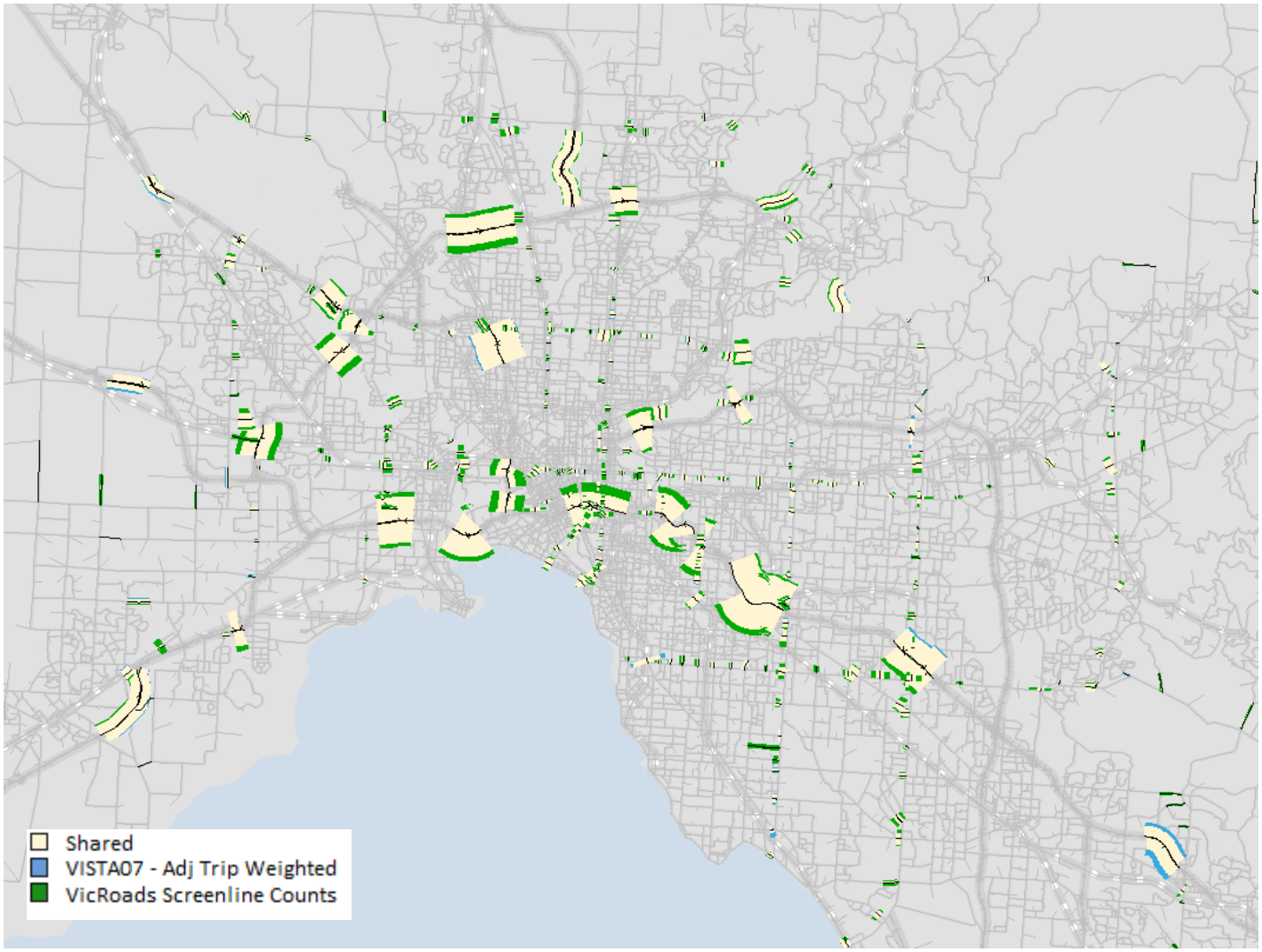

Figure 13 - Difference in Load by Screenline Count Location (Inter Peak, 9am-3pm)

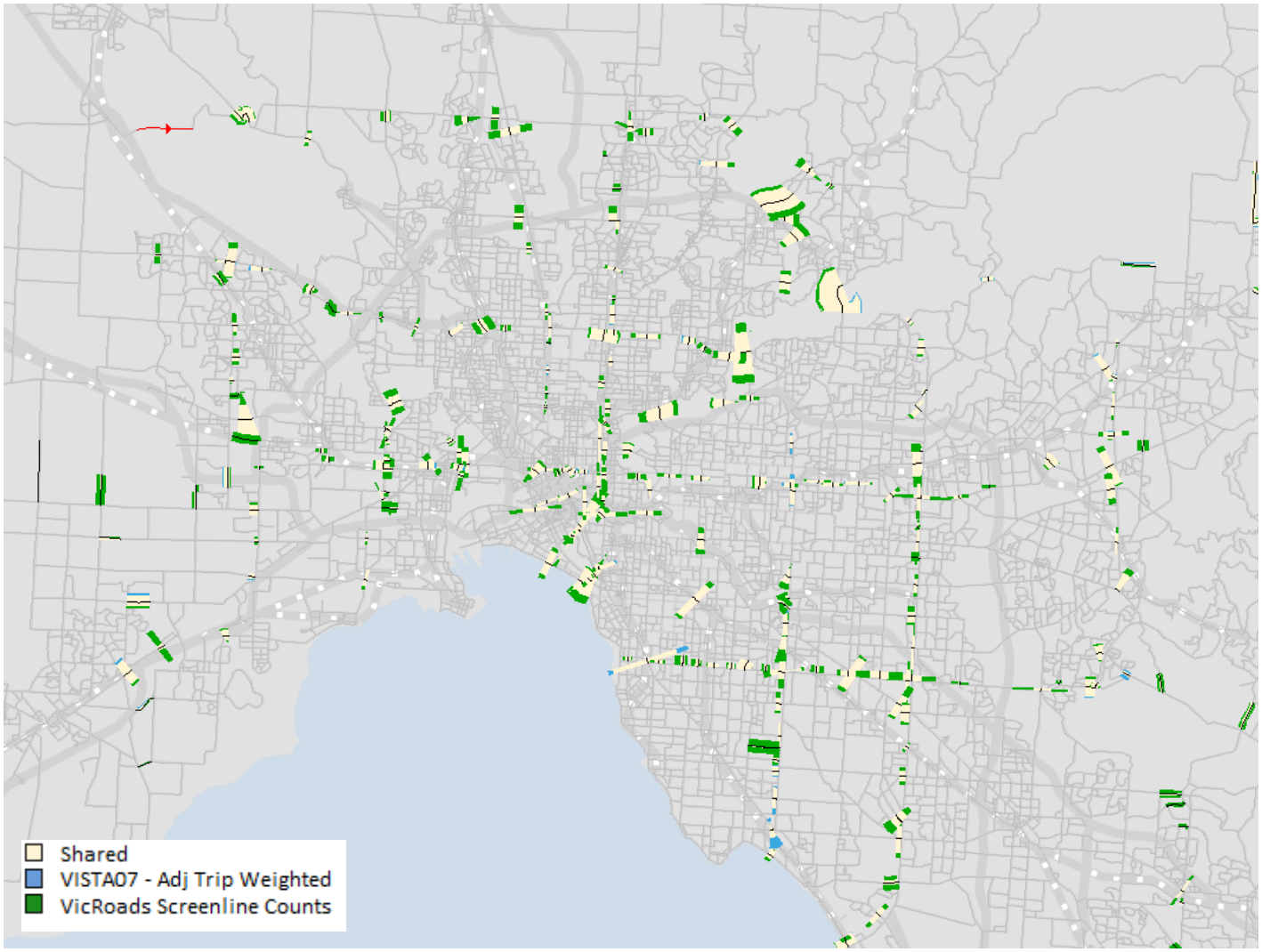

Figure 14 – Difference in Load by Screenline Count Location, Excluding Freeways (Inter peak, 9am-3pm)

3.1.4 By Level of Flow

In the previous sections we have noted that VISTA07 tends to be high (or at least less low) on major roads (e.g. freeways) during the PM peak and the inter peak. We believe that this provides some clues as to the types of trips which are under-reported. To further explore this effect, we have compared VISTA07 flows to screenline traffic counts, for different levels of road demand.

Figure 15 and Table 1 present the analysis for the AM peak. It can be observed that VISTA07 tends to over-report traffic on low volume roads (0 – 1000 vehicles in the 2 hour peak) by about 30%, and over-report by around 5 – 10% on all other roads. Once again, we believe VISTA07 to be high because of the correction factors – if we remove the correction factors, VISTA07 is spot on. This is an impressive result.

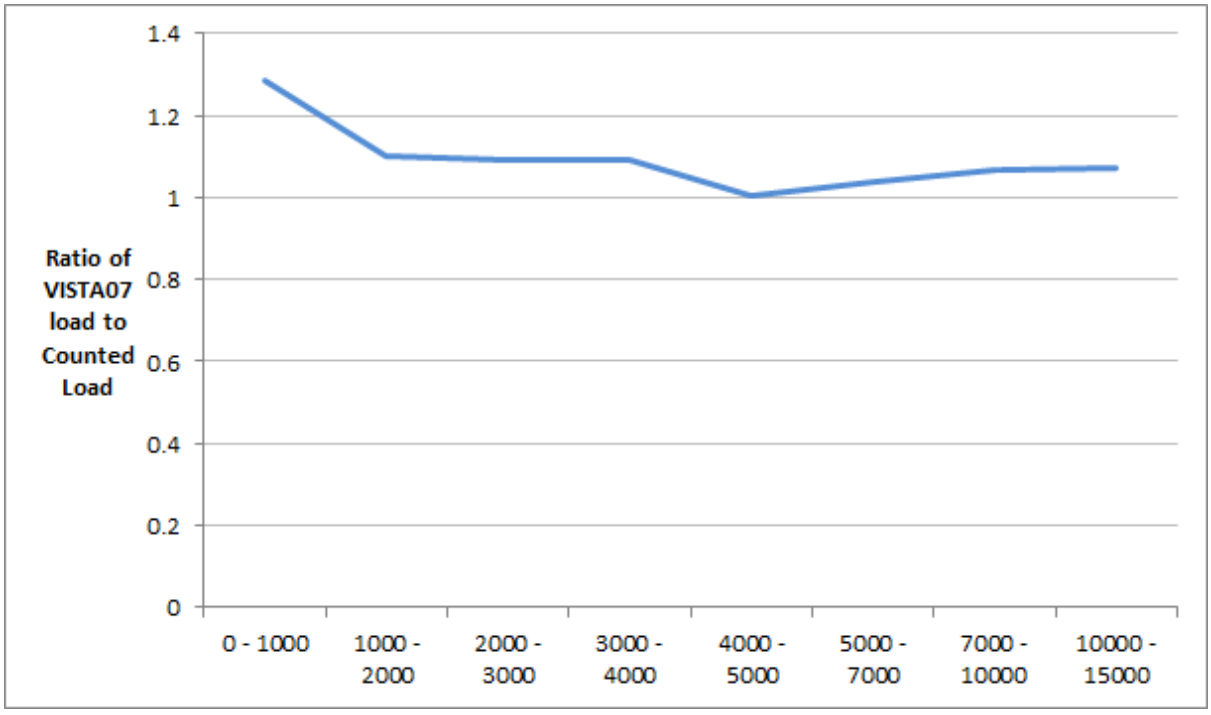

Figure 15 - Ratio of VISTA07 Traffic to Screenline Traffic, by Load Level (AM Peak)

Load Level	VicRoads 2006 Screenlines	VISTA07 - Adj Trip Weighted	Difference
0 - 1000	135,958	174,676	28%
1000 - 2000	314,899	346,489	10%
2000 - 3000	233,014	254,782	9%
3000 - 4000	204,436	223,255	9%
4000 - 5000	140,349	140,611	0%
5000 - 7000	148,431	153,840	4%
7000 - 10000	230,156	246,030	7%
10000 - 15000	130,703	140,324	7%

Table 1 – Traffic by Load Level (AM Peak)

In the PM Peak (Figure 16 and Table 2), a different pattern is evident. For low volume roads (0 – 2000 in the 3 hour peak), VISTA07 is approximately 10% low. However, as the load level increases, VISTA07 improves, and eventually over-reports on high volume roads (10000+ vehicles). This is consistent with our findings in Section 3.1.3.

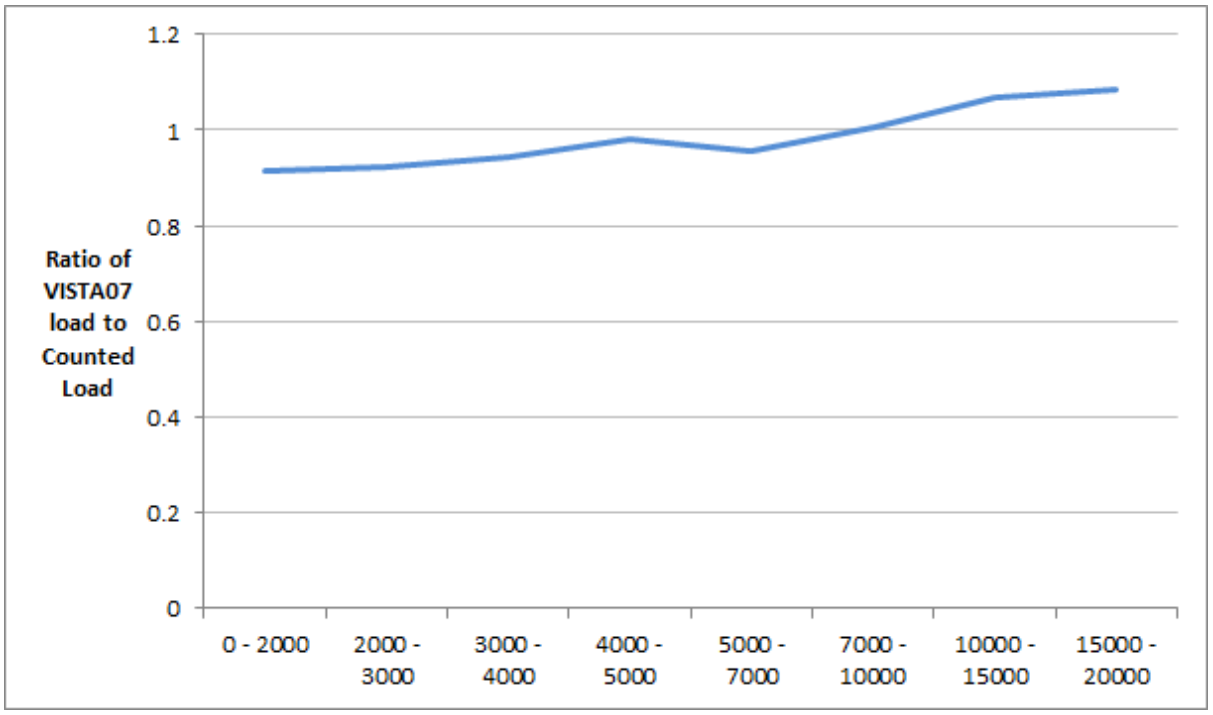

Figure 16 - Ratio of VISTA07 Traffic to Screenline Traffic, by Load Level (PM Peak)

Load Level	VicRoads 2006 Screenlines	VISTA07 - Adj Trip Weighted	Difference
0 - 2000	358,221	327,703	-9%
2000 - 3000	348,413	321,580	-8%
3000 - 4000	299,831	282,978	-6%
4000 - 5000	208,626	204,395	-2%
5000 - 7000	394,689	377,762	-4%
7000 - 10000	242,696	244,115	1%
10000 - 15000	339,427	362,114	7%
15000 - 20000	132,335	143,382	8%
20000 - 25000	61,633	65,665	7%

Table 2 - Traffic by Load Level (PM Peak)

In the Inter Peak (Figure 17 and Table 3), this pattern is exaggerated further. For low volume roads (< 2000 in the 6 hour inter peak), VISTA07 is 40% lower than screenline counts. As the load level increases, VISTA07 gradually improves, and is only 25% low on roads carrying 10,000 to 20,000 vehicles, and 13% low on roads carrying in excess of 20,000 vehicles. This is a fascinating pattern, particularly when contrasted with the AM peak, where the opposite effect was present. It seems to suggest that VISTA07 predominantly missing the trips which use smaller volume roads.

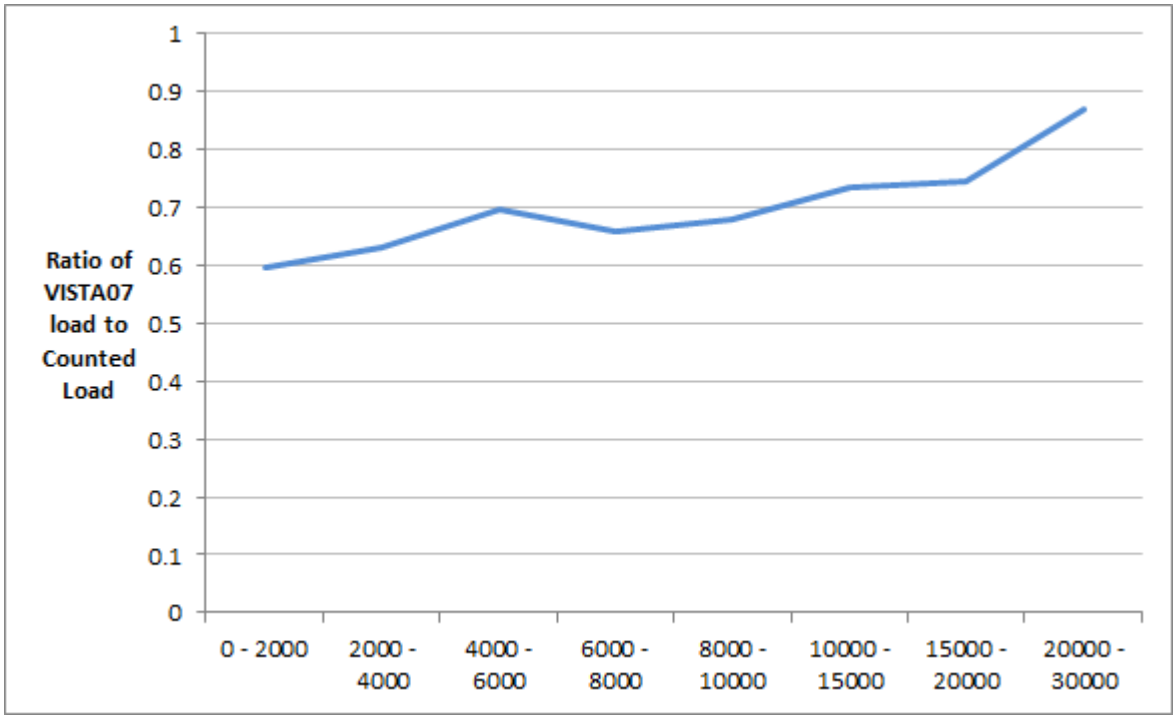

Figure 17 – Ratio of VISTA07 Traffic to Screenline Traffic, by Load Level (Interpeak)

Load Level	VicRoads 2006 Screenlines	VISTA07 - Adj Trip Weighted	Difference
0 - 2000	218,663	130,696	-40%
2000 - 4000	625,096	395,262	-37%
4000 - 6000	687,446	478,797	-30%
6000 - 8000	453,557	299,395	-34%
8000 - 10000	295,095	200,920	-32%
10000 - 15000	353,091	259,453	-27%
15000 - 20000	426,186	317,656	-25%
20000 - 30000	362,713	314,827	-13%

Table 3 – Traffic by Load Level (Inter Peak)

3.2 Public Transport

Public transport demands in VISTA07 have also been validated against available public transport counts and surveys.

Public transport trips in VISTA07 have been "assigned" to the Zenith base year public transport network using a specially developed path finding algorithm which finds the most likely path for each public transport trip, given the information provided in the VISTA07 survey. For example, survey respondents were asked to define the location of their boarding and alighting stops, as well as their travel mode, for each leg of their journey.

As with car trips, public transport trips have been weighted to 2006 population levels using two distinct approaches – a person weighted approach, which weights the survey sample to the population by LGA, gender and age, and an adjusted trip weighted approach developed by TUTI, which includes extra "correction factors" which aim to correct for certain types of under-reporting in VISTA07.

3.2.1 Rail Patronage

3.2.1.1 DOT 2008 Boarding Estimates

From the assignment of the VISTA07 survey to the public transport network, we can calculate a number of boardings per train station, for each hour of the day, and trip purpose. Of course, the VISTA07 estimates are unreliable at this level of disaggregation, as the survey is only a small sample. As such, we have aggregated the VISTA07 boarding estimates to line groups, and compared them with DOTs estimates of rail boardings for 2008.

Table 4 and Figure 18 present a comparison of DOTs 2008 estimates of rail boardings with estimates from VISTA07, across the entire day.

Overall, VISTA07 is approximately 10% lower than DOTs estimates, though it is noticeably low on the Burnley and Northern Corridors.

Line Group	DOT Estimates (2008)	VISTA07 (2006)	Difference
Burnley	127,007	98,585	-22%
Caulfield	148,160	145,039	-2%
Clifton Hill	63,902	60,112	-6%
Northern	112,773	87,129	-23%
Inner City Interchange Stations	231,274	225,670	-2%
Total	**683,116**	**616,536**	**-10%**

Table 4 - Rail Boardings by Line Group (Daily)

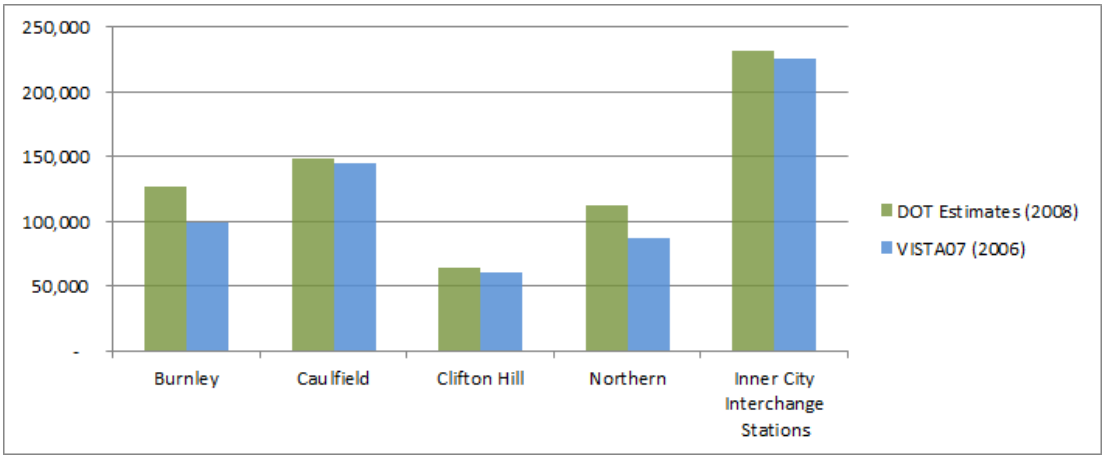

Figure 18 - Rail Boardings by Line Group (Daily)

Part of this difference could be brought about by the difference in year. Part could also be due to public transport trips made by visitors, which do not form part of VISTA07.

However, despite these factors, the story is quite different in the AM Peak (7-9am), as presented in Table 5 and Figure 19 below. The Burnley and Northern lines, which were 22% and 23% low daily, are 2% low, and 11% high in the AM peak, respectively. Overall boardings are 7% high in the AM peak.

In other words, where the daily boardings were low, the AM peak boardings are high.

Line Group	DOT Estimates (2008)	VISTA07 (2006)	Difference
Burnley	44,947	44,058	-2%
Caulfield	51,382	61,433	20%
Clifton Hill	23,267	24,604	6%
Northern	42,180	46,682	11%
Inner City Interchange Stations	10,551	7,928	-25%
Total	**172,327**	**184,705**	**7%**

Table 5 - Rail Boardings by Line Group (AM Peak)

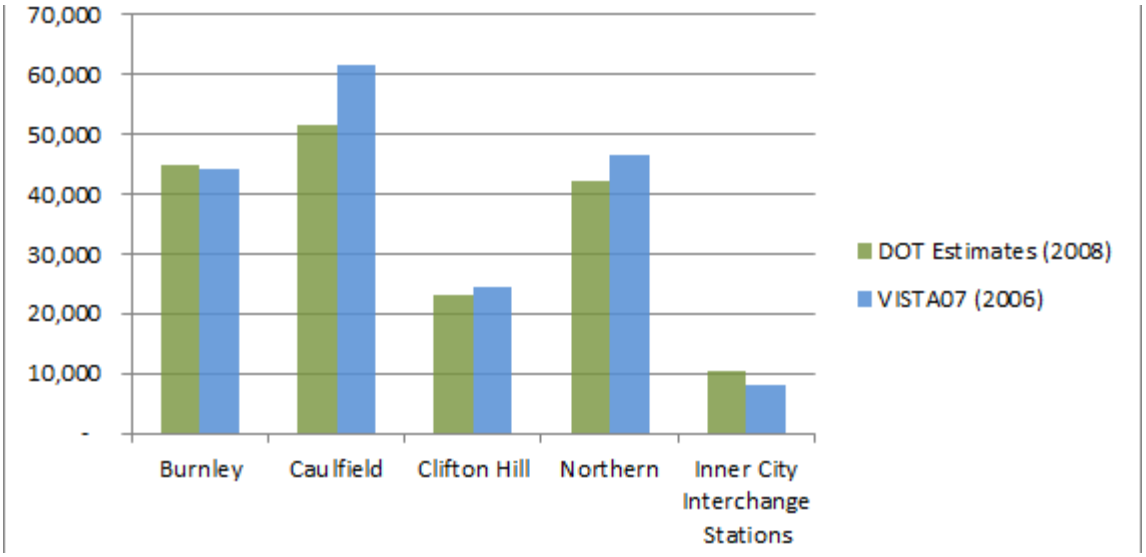

Figure 19 - Rail Boardings by Line Group (AM Peak)

The comparison for the PM peak (4-6pm) is presented in Table 6 and Figure 20 below. Overall, boardings in the PM peak are approximately equal in VISTA07 and the DOT estimates. However, there are significant differences among individual line groups. The Inner City Interchange Stations (the loop plus North Melbourne, Richmond and South Yarra) are approximately 22% high, while boardings on the Burnley, Caulfield and Northern lines are approximately 40% low.

Line Group	DOT Estimates (2008)	VISTA07 (2006)	Difference
Burnley	18,170	10,672	-41%
Caulfield	19,505	11,913	-39%
Clifton Hill	7,600	6,754	-11%
Northern	11,487	6,364	-45%
Inner City Interchange Stations	96,531	117,598	22%
Total	153,292	153,300	0%

Table 6 - Comparison of DOT estimated and VISTA07 Rail Boardings by Line Group (PM Peak)

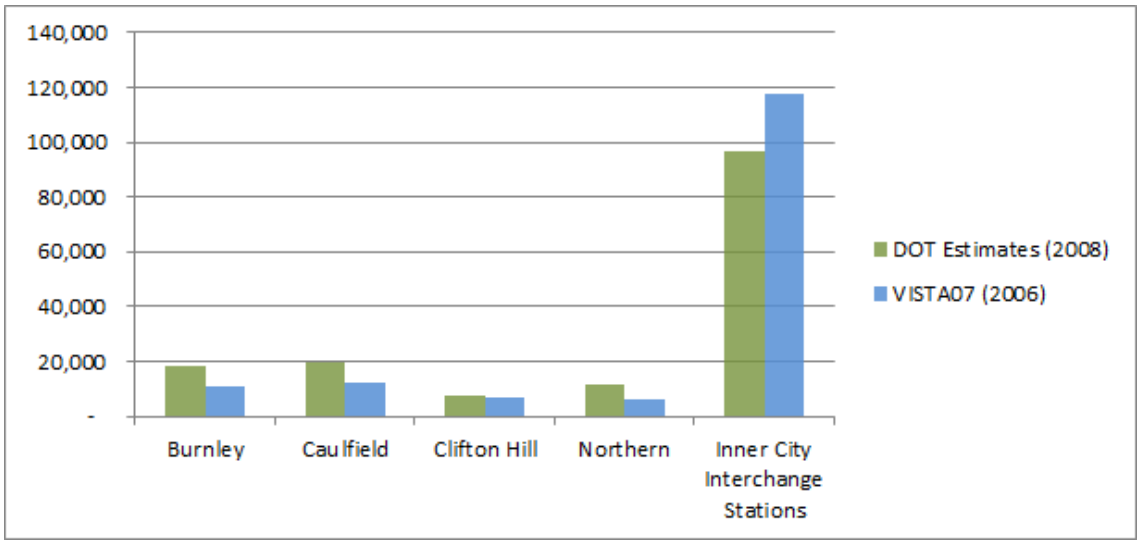

Figure 20 - Comparison of DOT estimated and VISTA07 Rail Boardings by Line Group (PM Peak)

In the Off Peak (Table 7 and Figure 21), VISTA07 boardings are on average 22% lower than DOTs estimates across all line groups. VISTA07 is especially lower on the Burnley and Northern lines (31% and 42% respectively).

Line Group	DOT Estimates (2008)	VISTA07 (2006)	Difference
Burnley	63,890	43,854	-31%
Caulfield	77,273	71,694	-7%
Clifton Hill	33,035	28,755	-13%
Northern	59,106	34,083	-42%
Inner City Interchange Stations	124,193	100,145	-19%
Total	357,497	278,530	-22%

Table 7 - Comparison of DOT estimated and VISTA07 Rail Boardings by Line Group (Off Peak)

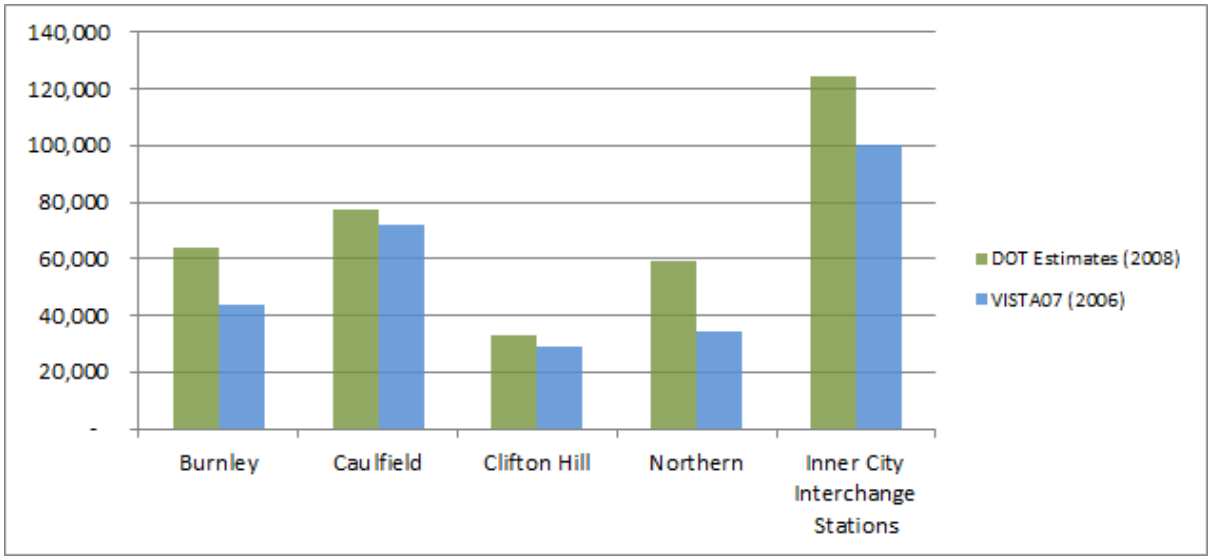

Figure 21 - Comparison of DOT estimated and VISTA07 Rail Boardings by Line Group (Off Peak)

Pulling this together, we can observe that:

- At a high level, VISTA07 is over-reporting rail demand in the AM peak by 7%, is about right in the PM Peak, and is 22% low in the off peak.
- This pattern is remarkably similar to that observed for road traffic.
- There is a fair degree of variation at the level of individual line groups. In particular, the Northern and Burnley lines tend to be low in VISTA07. We will attempt to explain that shortly.
- In the PM peak it appears that VISTA07 over-reports boardings in the CBD and under-reports on the suburban lines.

We will now explore these effects further.

AM Peak Over-Reporting

As with road traffic, we believe that VISTA07 is over-reporting AM peak rail boardings because of correction factors built into the adjusted trip weights. We believe that these factors, which are designed to correct for certain types of under-reporting (and are not time period specific), are over-correcting in the AM peak. If we remove the correction factors (i.e. use person weights rather than adjusted trip weights), then VISTA07 is 4% lower than DOTs 2008 estimates. Given that VISTA07 has been weighted to 2006 populations (but is representative of behaviour in 2007/08), this is about where we'd expect it to be.

Of course, removal of the correction factors also reduces VISTA07 in the other periods, causing VISTA07 to be lower than DOTs 2008 estimates by 9% in the PM peak, and 30% in the off peak.

Variation Between Line Groups

There is a surprising amount of variation between VISTA07 and the estimated boardings at the line group level. More specifically, the Burnley and Northern groups are significantly lower than the Caulfield and Clifton Hill groups, as shown again in Table 10 below.

These variations are worth exploring because they could have implications for the development of the Mode Choice models based on VISTA07.

Line Group	DOT Estimates (2008)	VISTA07 (2006)	Difference
Burnley	127,007	98,585	-22%
Caulfield	148,160	145,039	-2%
Clifton Hill	63,902	60,112	-6%
Northern	112,773	87,129	-23%
Inner City Interchange Stations	231,274	225,670	-2%
Total	**683,116**	**616,536**	**-10%**

Table 8 - Comparison of DOT estimated and VISTA07 Rail Boardings by Line Group (Daily)

In the case of the Northern line, we believe a great deal can be explained in terms of population growth. According to DOT estimates, AM peak rail demand on the Northern Group (at the CBD cordon) grew by 31% between 2006 and 2008.

Part of this growth will have been due to population growth in the corridor, and part will be due to other factors – fuel prices, road congestion, CBD employment growth, environmental and health awareness, etc. Having been collected during 2007 and 2008, VISTA07 will already be reflective of these other factors (at least to the extent that they were present during the survey period). However, VISTA07 has been weighted to 2006 population, and therefore, any growth in rail demand caused by population growth will be missing from VISTA07.

We believe that this helps to explain why the Northern Line is low in VISTA07 – considerable population growth occurred in this corridor between 2006 and 2008.

However, while this explanation fits in terms of daily boardings, it fails to explain why the Northern line is in 11% higher than the DOTs 2008 estimates in the AM peak, suggesting that the under-reporting is confined to the off peak and PM peak. We have no explanation for this at present, but it is possible that we are stretching the VISTA07 sample too far.

The Burnley line is more puzzling, given its more stable population. We have explored whether this may be due to sample bias. In Figure 22 below, we look at the number of people living within certain distain distances of their nearest train station on the Burnley line. This analysis is restricted to those people who have the Burnley group as their closest group. It can be seen that VISTA07 includes less people living within 1km of a train station on the Burnley line than are present in the Census. Conversely, VISTA07 has a larger number of people living between 1 and 4km from a station. Those living close to train stations are obviously more likely to travel by train, and this may cause a bias. However, we have also examined average rail mode shares by distance from the nearest train station, and concluded that this is unlikely to be the cause.

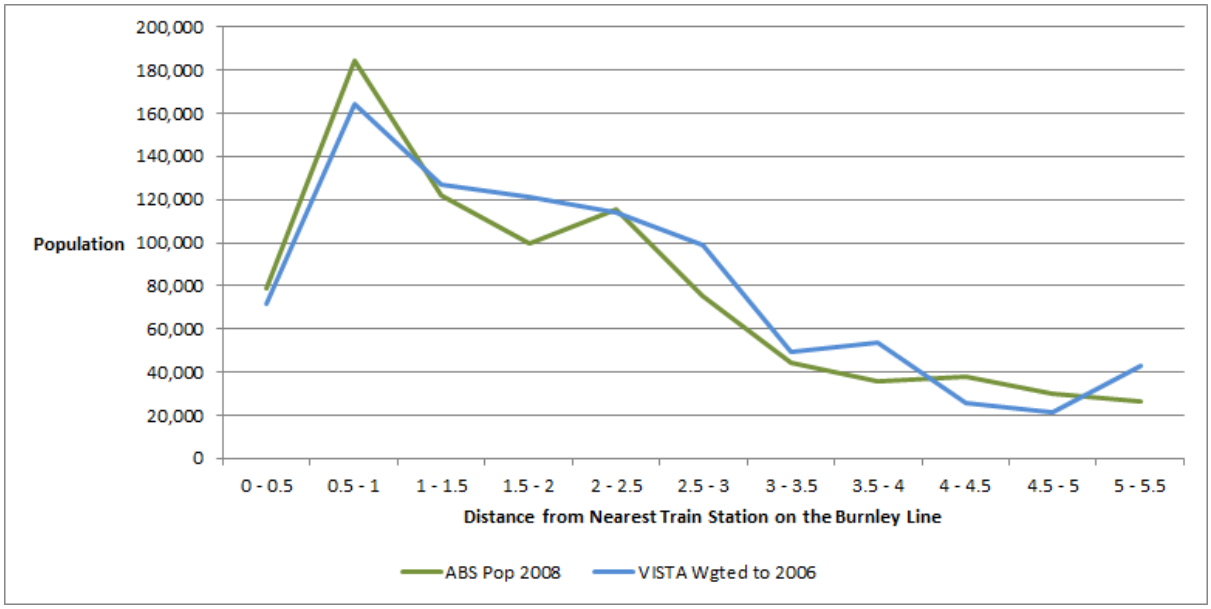

Figure 22 - Population by Distance from Nearest Train Station on the Burnley Line

We have further explored the Burnley group by examining individual line corridors and line segments. The results, presented in Table 9 below, show that VISTA07 is considerably lower than DOTs 2008 estimates on most line segments, the major exception being the East Camberwell to Ringwood segment (which is of course, the trunk of this line group).

Line Corridor	Line Segment	DOT Estimates (2008)	VISTA07 (2006)	Difference
Camberwell Corridor	Riversdale-Alamein	5,883	1,503	-74%
	Ringwood East-Lilydale	9,990	8,945	-10%
	Heathmont-Belgrave	13,464	6,939	-48%
	East Richmond-Camberwell	29,167	19,141	-34%
	East Camberwell-Ringwood	41,874	40,649	-3%
Glen Waverley Line	Heyington-Glen Waverley	26,630	21,408	-20%

Table 9 - Rail Demands for the Burnley Group (DAILY)

Given this analysis, we are unsure as to why the Burnley group seems to be so under-reported in VISTA07. However, as with the Northern group, we can again observe that the under-reporting is entirely confined to the off peak and PM peak; the AM peak seems to be reported accurately.

This is clearly going to take more work to explain, but at a high level, these findings are almost entirely consistent with what we found when examining road traffic.

VISTA07 seems to reflect AM peak travel very accurately, but is missing a very large chunk of off peak travel, and in the case of rail, PM peak boardings on the suburban lines.

Viewed in this light, perhaps we should not be asking why the Burnley and Northern groups are low in the off peak, but instead asking why the Caulfield and Clifton Hill groups are not (or are at least significantly less low).

3.2.1.2 Rail OD Survey 2009/10

The Rail OD survey, conducted in 2009 and 2010, collected information that turns out to be very helpful in examining the under-reporting of travel in the off peak.

Figure 23 below compares VISTA07 boardings with counts performed as part of the Rail OD survey, by hour of the day. Counts were only performed between 6:30am and 7pm – for simplicity, we have left out the 6:30 – 7:00am period.

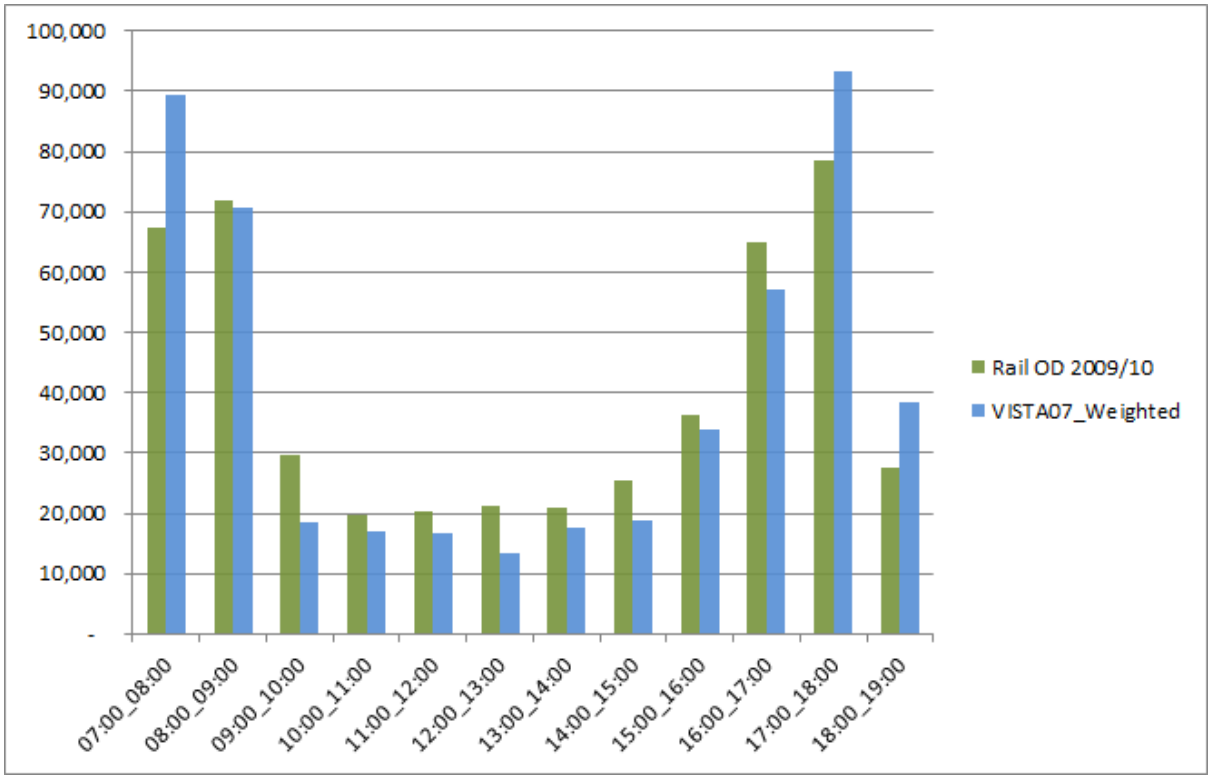

Figure 23 - Total Rail Boardings by Hour of the Day

The pattern should look very familiar – it is almost identical to that observed for traffic, though the level of under-reporting in the inter peak is not quite as severe.

Besides counts, the Rail OD survey also involved on platform interviews designed to capture information relating to:

- Destination station (including any en-route interchanges)
- Trip purpose
- Mode of access / egress
- Origin and destination location

This information gives us a pathway to exploring what types of trips are missing from the off peak in VISTA07. The trip purpose information is particularly useful.

Figure 24, Figure 25 and Figure 26 present boardings by trip purpose for the AM peak, PM Peak and Inter Peak respectively, for both VISTA07 and the Rail OD survey.

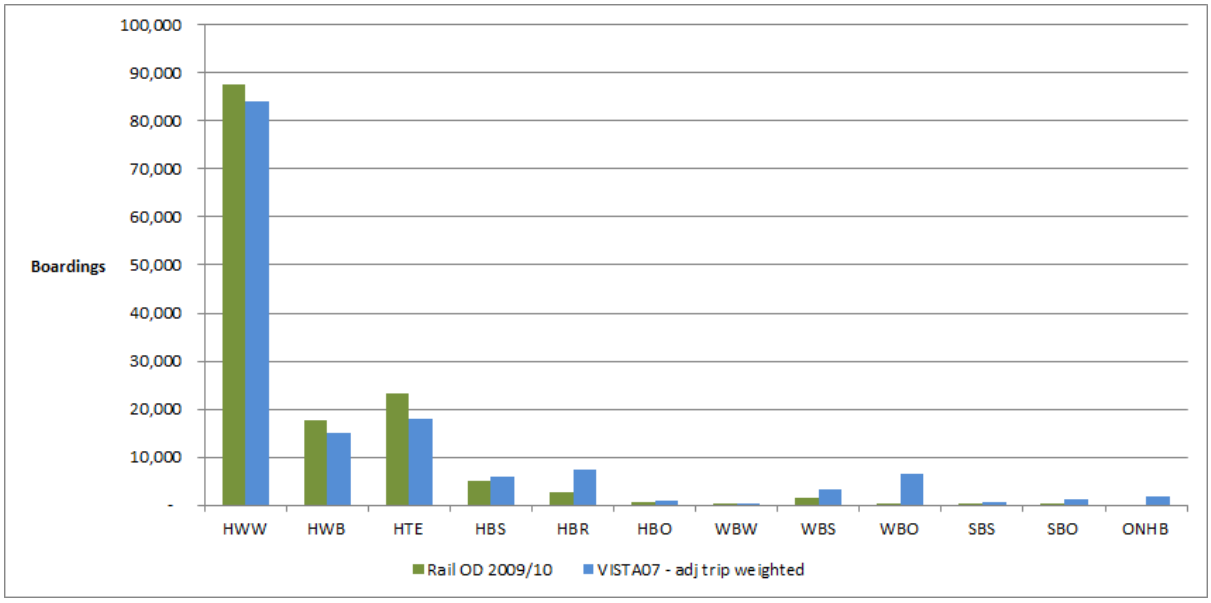

Figure 24 - Boardings by Trip Purpose (AM Peak, 7am – 9am)

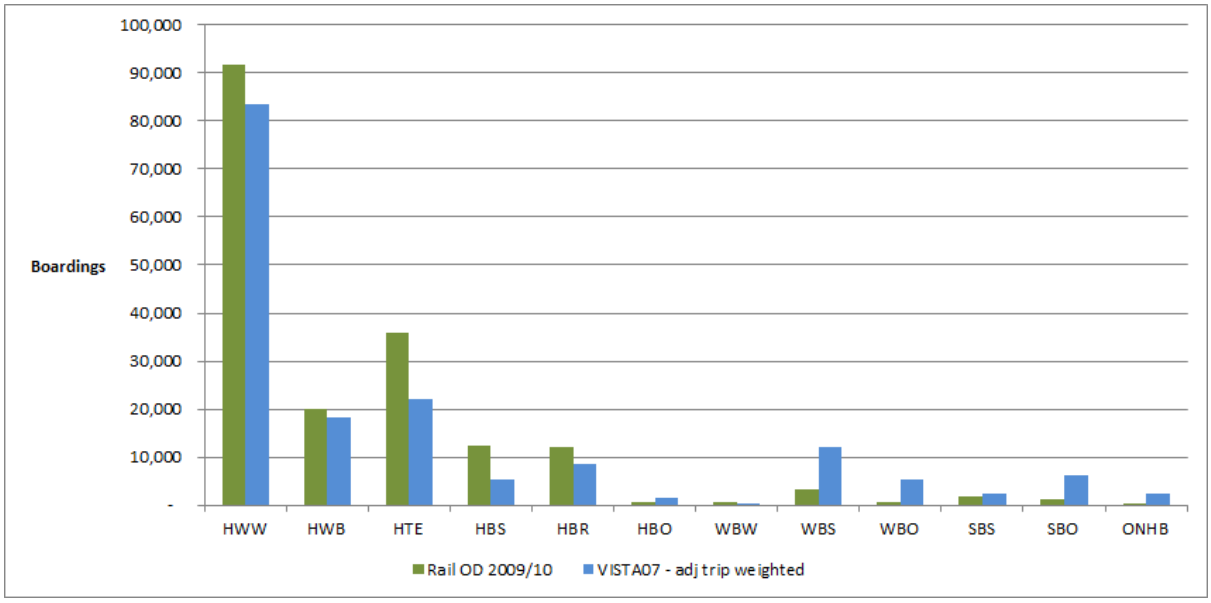

Figure 25 - Boarding by Trip Purpose (PM Peak, 3pm - 6pm)

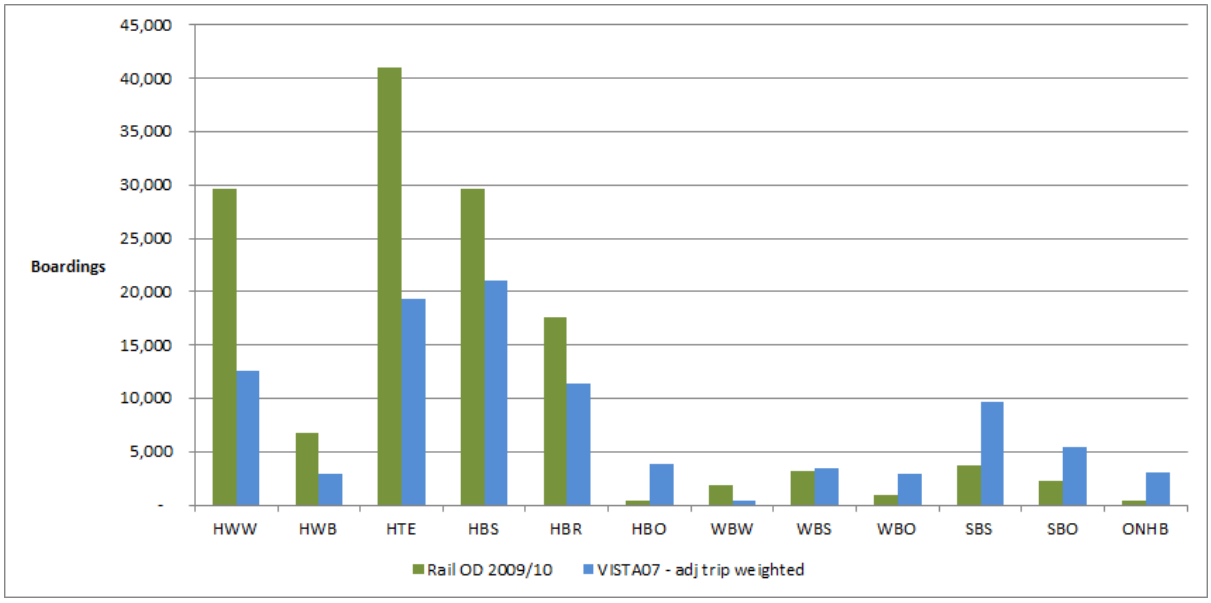

Figure 26 - Boardings by Trip Purpose (Inter Peak, 9am - 3pm)

In the AM peak (Figure 24), VISTA07 and the Rail OD survey are in very close agreement – yet again, VISTA07 does a fantastic job of reflecting the AM peak.

In the PM peak (Figure 25), VISTA07 appears to be missing a fair number of tertiary education, shopping and recreation trips. However, it compensates by having a higher number of non-home based trips. We believe this is caused by differences in survey method and definitions.

In the Inter Peak (Figure 26), we have our first hard evidence of the types of trips which might be missing from VISTA07 in the off peak. Perhaps surprisingly, the Rail OD survey suggests that VISTA07 is missing a large number of home based work trips, and a large number of tertiary trips in the off peak. It also appears to be missing some home based shopping and recreation trips, but these are compensated by their non-home based counter-parts: shopping based shopping, shopping based other and other non-home based.

Because of definitional and methodological differences between VISTA07 and the Rail OD survey, we have further aggregated these trip purposes into three groups: "work related", "tertiary education" and "other". The result (for the Inter Peak only) is presented in Figure 27 below.

This result came as quite a surprise to us; until this point our working hypothesis had been that "commuting" trips – Home Based Work and Home Based Education – were reported accurately in VISTA07 (explaining why the peaks were reflected accurately), and that the off peak under-reporting was mostly caused by discretionary travel such as Home Based Shopping, Home Based Recreation, Shopping based Shopping, etc.

If correct, the comparison with the Rail OD survey suggests instead that off peak commuting trips are the missing piece of off peak travel, at least for rail demand. An obvious question is whether the same principle could apply to road traffic. In Section 4.3 we will present evidence that suggests that to a large extent it does.

Figure 27 - Boardings by Grouped Purpose (Inter Peak, 9am - 3pm)

3.2.2 Tram Patronage

VISTA07 tram demands have been compared with tram demands from the 2008 Tram OD survey.

Referring to Table 10, it can be seen that VISTA07 is 11% low in the AM peak, 27% low in the PM peak, and 48% low in the Off Peak. Overall it is 37% low across the day.

Period	Tram OD 2008	VISTA07 - Adj Trip Weighted	Difference
AM Peak	101,259	90,499	-11%
PM Peak	103,490	75,034	-27%
Off Peak	361,994	189,723	-48%
Total	566,742	355,256	-37%

Table 10 – Tram Demands by Period

A part of this can be explained in terms of the count year – VISTA07 has been weighted to 2006 population levels, but the Tram OD survey was conducted in 2008. This cannot account for a 37% difference daily, but it may help explain the 11% difference observed in the AM peak.

A part may also be explained by noting that the CBD did not form part of the VISTA07 sampling frame, even though people live there.

Visitors are also not included in VISTA07, and may be heavy tram users.

However, these explanations are unlikely to explain all of the apparent under-reporting.

To explore this further, we have disaggregated by hour of the day, as presented in Figure 28 below.

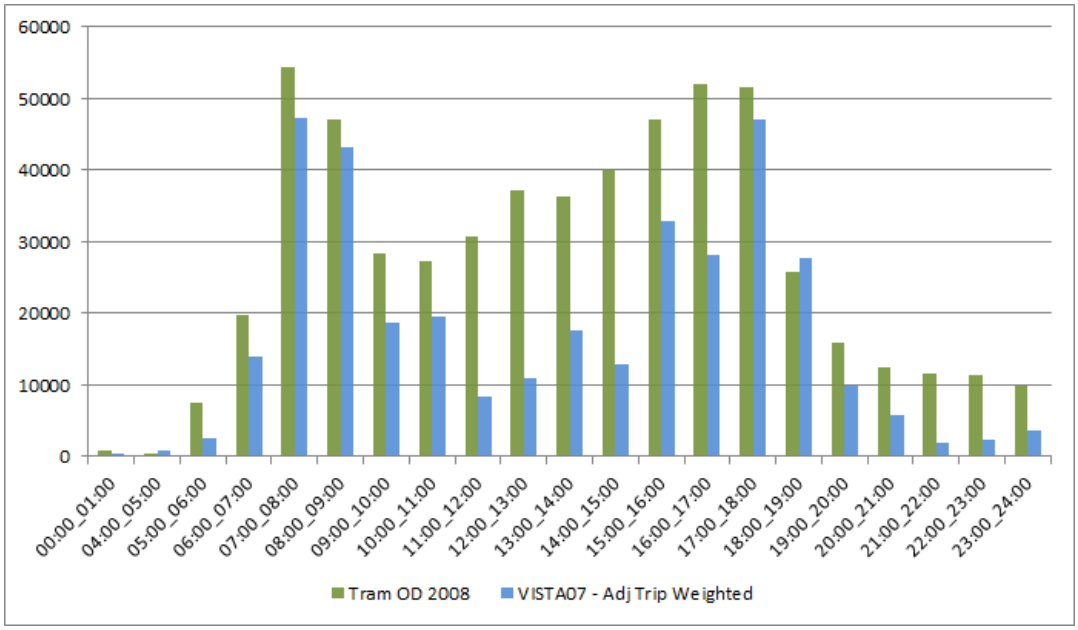

Figure 28 - Tram Boardings by Hour of the Day

This pattern is remarkably similar to the pattern observed for both rail and road traffic demands. Peak demands (especially the AM peak) are reflected accurately, while the off-peak is heavily under-reported in VISTA07. The hours between 11am and 3pm, and after 7pm are especially under-reported.

Using the on's and off's from the Tram OD survey, and our assignment of the VISTA07 survey, we have also compared loads throughout the network. Figure 29, Figure 30 and Figure 31 present the results for the AM peak, PM peak and Inter Peak respectively.

Unsurprisingly, there is a fair degree of variation between VISTA07 and the Tram OD survey. We think that some of the differences appear to be caused by changes in the network which may have occurred in between VISTA07 and the Tram OD survey, though we are yet to confirm this.

The key thing to note is that VISTA07 is significantly under-reporting travel in the Inter Peak, and it appears to be under-reporting across the whole network. In other words, the problem is not confined to travel within the CBD. Figure 32 provides a closer look at tram under-reporting in the inner city.

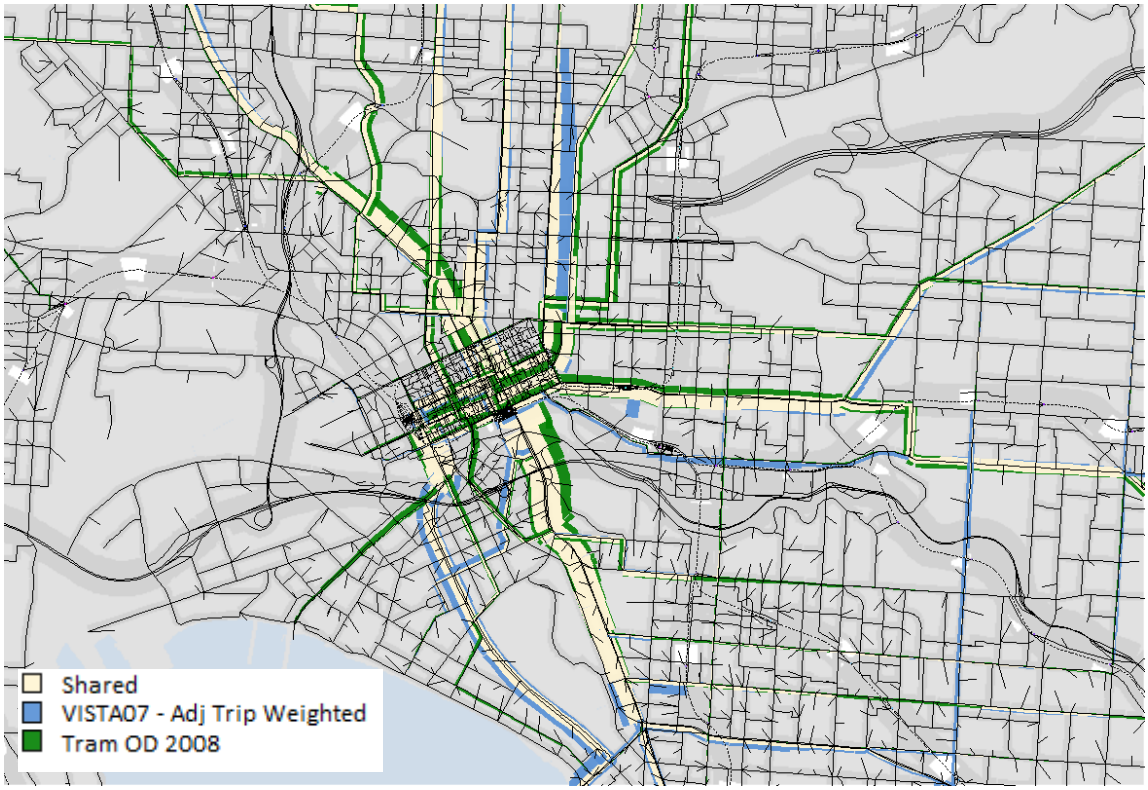

Figure 29 - Tram Loads (AM Peak, 7am – 9am)

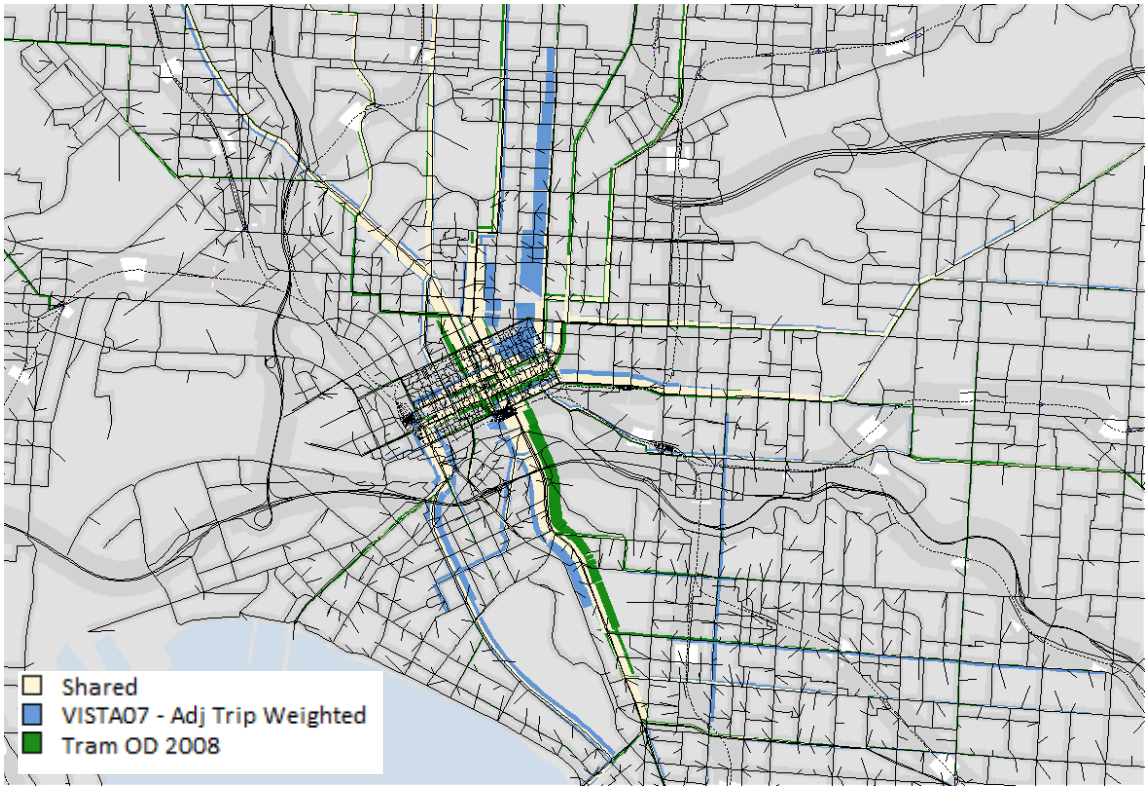

Figure 30 - Tram Loads (PM Peak 5pm - 7pm)

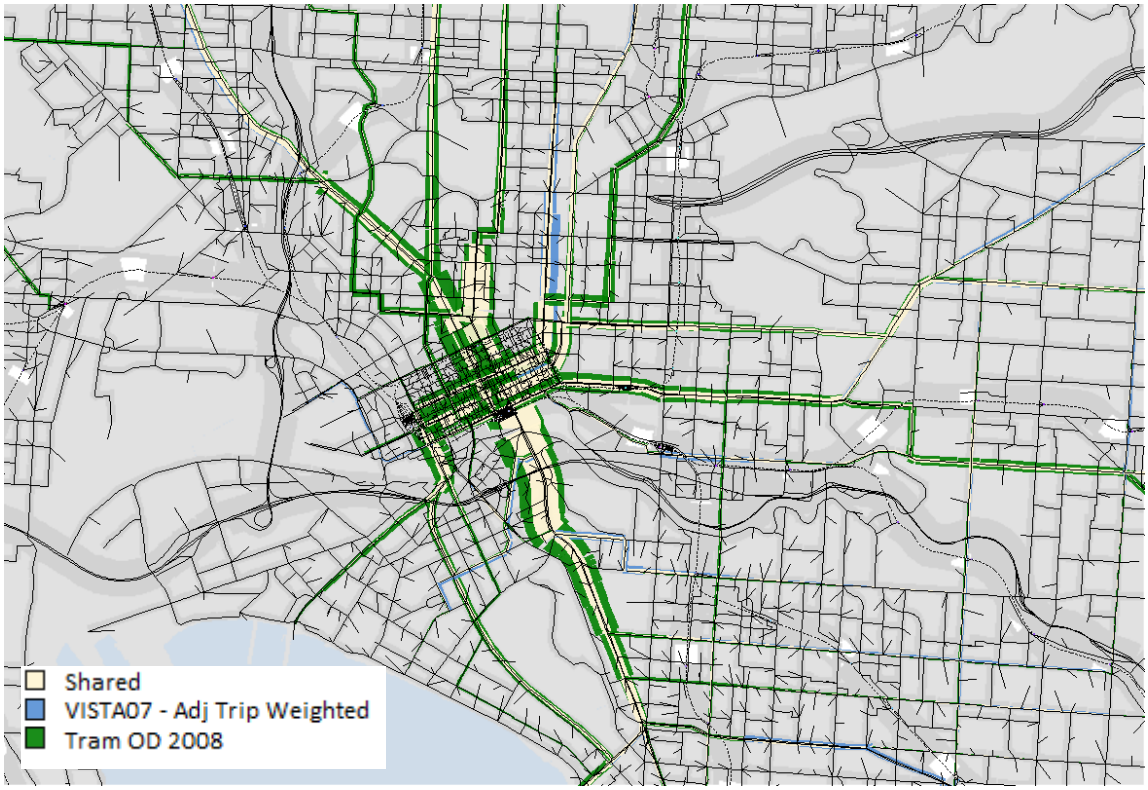

Figure 31 - Tram Loads (Inter Peak, 9am - 4pm)

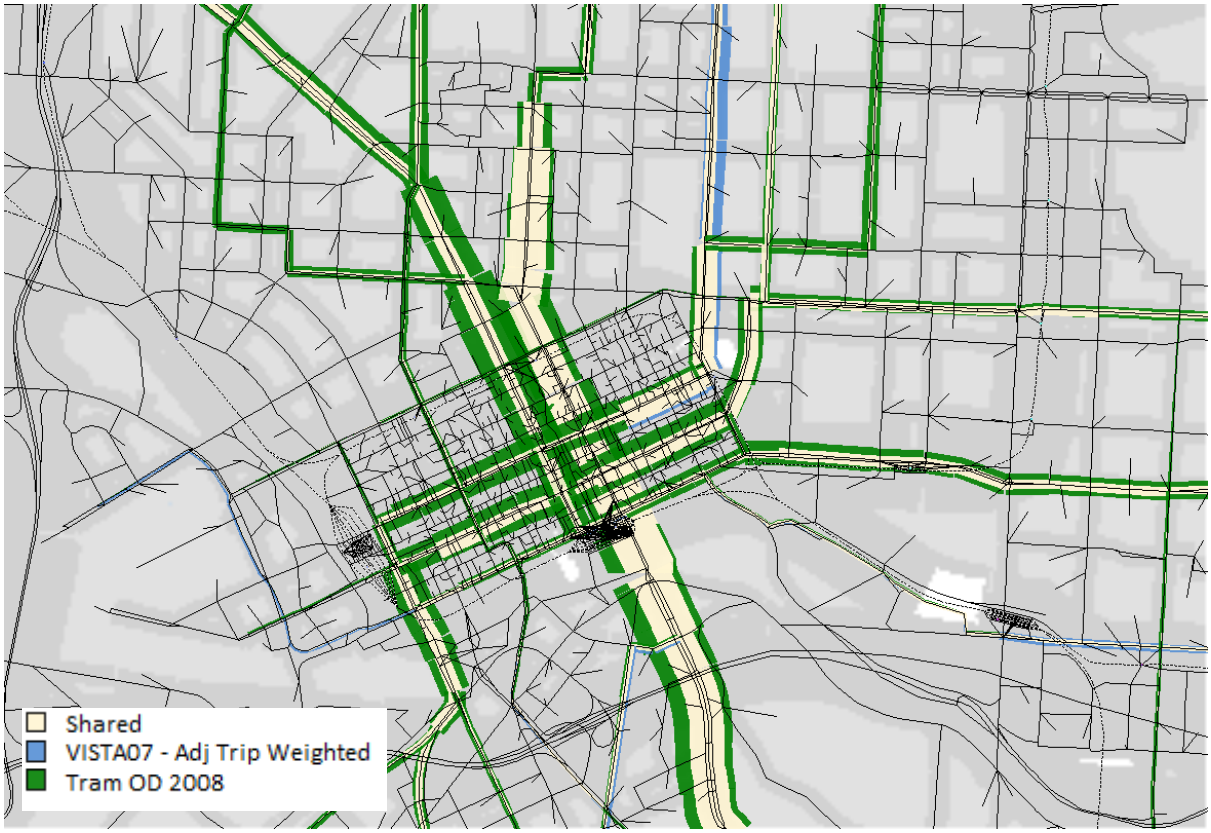

Figure 32 - Tram Loads in the Inner City (Inter Peak, 9am - 4pm)

3.2.3 Bus Patronage

VISTA07 demands on the bus network have been compared with 2008 estimates of bus boardings, based on validations during March, May, August and October, 2008, and counted bus boardings from 2005.

Table 11 below provides a comparison of bus boardings by route series (200s, 300s, etc) and by time of day.

There is considerable variation at the route series level. For example, across the entire day, the 600s series has 31% more boardings in VISTA07, while the 800s series has 54% less. This may be partly due to changes in the public transport network (such as the implementation of SmartBuses); it may also be to do with a lack of sample in VISTA07. For example, assuming an average weight of approximately 200, the 22,228 boardings of 800 series busses in VISTA07 is probably the result of around 110 recorded boardings in the VISTA07 survey. This is not a large sample.

The results are potentially more meaningful if aggregated over all routes during a period. Across the entire day, and all bus routes, VISTA07 has 13% less bus boardings than estimated for 2008. In the AM Peak, VISTA07 has 3% *more* bus boardings, while in the PM Peak and Off Peak, VISTA07 has 35% and 13% *less* boardings, respectively.

While the daily difference could be partly explained in terms of growth from 2006 to 2008, and visitors, who are missing from VISTA07, it is likely that VISTA07 is under-reporting bus travel, particularly in the PM peak, and to a lesser degree the off peak.

The PM peak under-reporting of bus demands is curious, and might be linked in some way to the under-reporting of boardings of suburban rail line boardings in the PM peak.

AM Peak

Route Series	Estimated Boardings (2008)	VISTA07 Estimated Boardings (2006)	Difference
200's	10,734	8,254	-23%
300's	5,759	5,102	-11%
400's	12,902	12,936	0%
500's	11,105	13,964	26%
600's	8,321	13,696	65%
700's	13,205	16,079	22%
800's	9,898	3,903	-61%
900's	2,183	2,685	23%
Total	**74,106**	**76,620**	**3%**

PM Peak

Route Series	Estimated Boardings (2008)	VISTA07 Estimated Boardings (2006)	Difference
200's	9,486	5,459	-42%
300's	3,994	3,443	-14%
400's	9,897	7,145	-28%
500's	9,163	5,793	-37%
600's	6,471	6,387	-1%
700's	11,550	8,402	-27%
800's	8,501	1,573	-82%
900's	2,281	1,820	-20%
Total	**61,344**	**40,023**	**-35%**

Off Peak

Route Series	Estimated Boardings (2008)	VISTA07 Estimated Boardings (2006)	Difference
200's	31,700	27,439	-13%
300's	9,026	5,626	-38%
400's	35,549	31,209	-12%
500's	31,914	33,641	5%
600's	20,099	25,747	28%
700's	37,321	30,685	-18%
800's	30,405	16,752	-45%
900's	7,962	6,708	-16%
Total	**203,975**	**177,807**	**-13%**

Daily

Route Series	Estimated Boardings (2008)	VISTA07 Estimated Boardings (2006)	Difference
200's	51,920	41,152	-21%
300's	18,779	14,171	-25%
400's	58,347	51,291	-12%
500's	52,772	53,399	1%
600's	34,891	45,830	31%
700's	62,076	55,165	-11%
800's	48,804	22,228	-54%
900's	12,426	11,214	-10%
Total	**340,016**	**294,449**	**-13%**

Table 11 - Bus Boardings by Period

4 Examination of Under-Reporting in VISTA07

In Section 3 we found that there is considerable under-reporting of travel in VISTA07, particularly in the Off Peak, and to a lesser extent, the PM peak. By contrast, the AM peak seems to be reported accurately. This pattern is very consistent across all motorised modes of travel (road, rail, tram and bus), and is relatively consistent across the entire network. It is also consistent from hour to hour during the inter peak and evening peak.

In this Section we will examine possible causes of this under-reporting, and develop a working theory about the types of trips that are being under-reported. Correction factors are also developed, and implemented in the Zenith Trip Generation and Period Allocation models.

4.1 Potential Sources of Under-Reporting

There are two broad reasons why travel might be under-reported in VISTA07:

a) Survey respondents under-report their travel, either because:
   I. The respondent deliberately left out certain trips, so as to reduce the time taken to complete the form, or because they didn't think the trips were important,
   II. The respondent forgot certain trips. This effect would presumably increase if there was a lengthy gap between the respondent's travel day, and the completion of the travel diary,
   III. The travel diary was completed by another household member (i.e. by proxy), and the proxy doesn't have full knowledge of the travel made by the person in question.
b) Those who respond to the VISTA07 survey travel (on average) *less* than those who do not respond.

TUTI have attempted to correct for effects a) (ii and iii) through the correction factors which are built into the VISTA07 adjusted trip weights. It is our opinion, however, that these correction factors may have over-corrected in the AM peak, and under-corrected in the Off Peak. Adjusting for this may help reduce the level of under-reporting in the Off Peak slightly!

Therefore, we are left with a(i), and b) as possible explanations.

Explanation a(i) is highly plausible given that VISTA07 uses a *self completion* survey method. It is quite easy for respondents to leave out trips.

In light of this issue, GPS devices were given to survey respondents during the pilot phase of the VISTA07 survey. Results from the GPS devices were later compared with the corresponding travel diaries, to measure the degree of under-reporting. This work found that approximately 12% of trips were missing from the travel diary (but present in the GPS data), while 20% of trips recorded in the diary but were missing from the GPS device! It is clear that GPS devices have (or at least, had) some technological issues to iron out.

The TUTI analysis showed that 23% of the "GPS Only" trips were "Home Based Other" trips (where other means non-work), 11% were "Non-Home-Based Work" and 66% were "Non-Home-Based Other". No "Home Based Work" trips were missing.

Furthermore, the analysis suggested that shorter trips were more likely to be left out of the travel diary.

One can imagine, however, that respondents may have filled out their travel diaries more accurately, knowing that the GPS devices were being used to check their diaries. Given this, it is hard to say to what extent the survey respondents might deliberately under-report their travel in the main survey, where no GPS device was provided.

As such, it is believable that this work under-estimates the degree by which the self-completed travel diary approach under-reports travel.

Given this, we have conducted a separate analysis which compares VISTA07 with the Sydney Household Travel Survey (Sydney HTS). The Sydney HTS uses an "interview" approach, rather than a "self-completion" approach. This analysis is presented in the next section.

On the matter of explanation b), that those who successfully respond to VISTA07 travel (on average) less than those who do not respond, we have no data on which to base an analysis. However, given that VISTA07 has a response rate of only 47%, it is believable that sampling bias could have a significant impact.

It is also believable that people who lead busy lives (involving lots of travel) are less likely to respond to a survey such as VISTA07. It is also plausible that those who lead less regular lives (i.e. not 9-5) might be less likely to fill in a survey form, or at least not be home when other household members are completing the form (remember that a full household must respond). This may help to explain the under-reporting of off peak travel.

4.2 Comparison with Other Household Travel Surveys

The results of VISTA07 have been benchmarked against two other recent Australian household travel surveys:

- The Sydney Household Travel Survey – (97-08) (Sydney HTS), and
- The South East Queensland Travel Survey (06-08) (SEQTS).

The Sydney Household Travel Survey employs an "interview" approach, designed to elicit an accurate picture of the travel made by the respondent.

By contrast, the South East Queensland Travel Survey uses a "self-completion" methodology much like that used in VISTA07.

Figure 33, below, shows a comparison of total person trip rates between the three surveys, while Figure 34 shows the same comparison split by trip purpose.

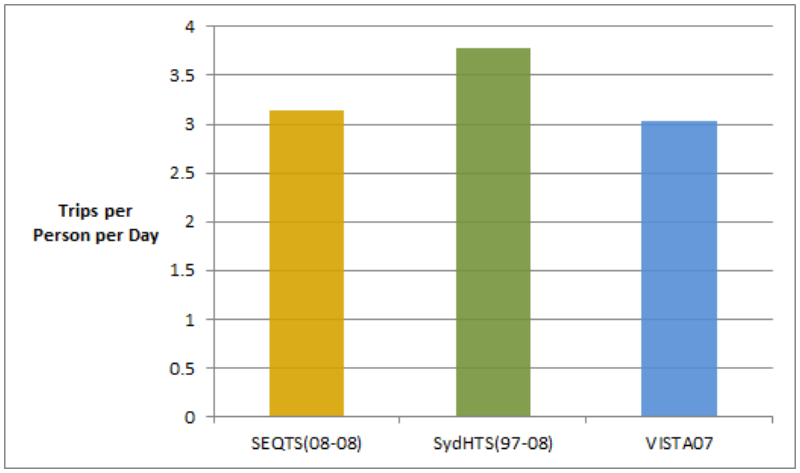

Figure 33 – Person Trip Rates

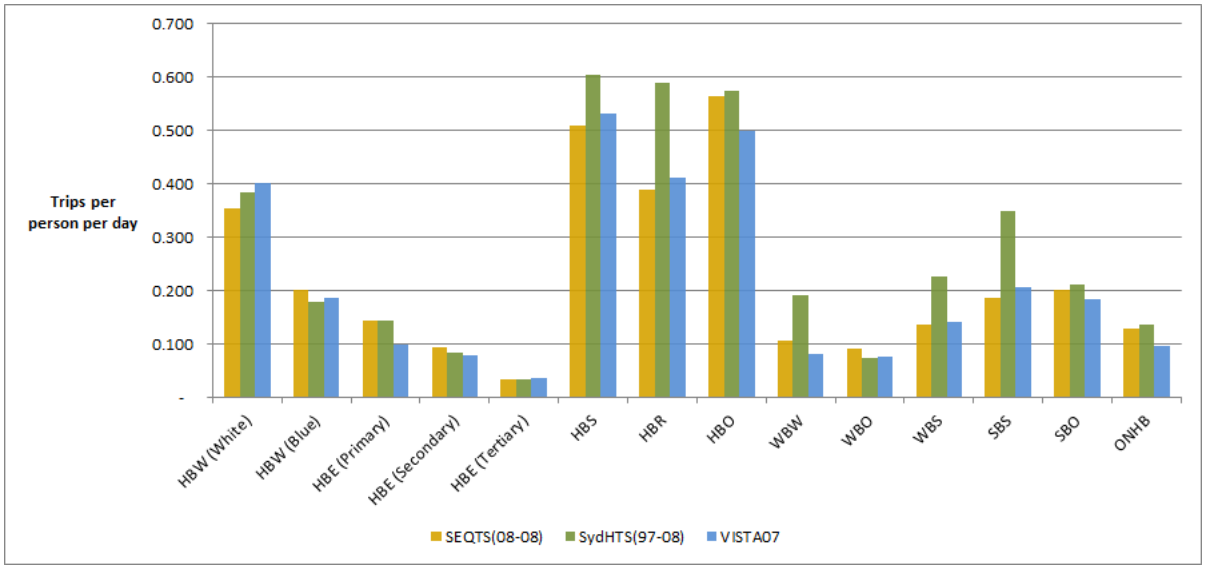

Figure 34 – Person Trip Rates by Trip Purpose

Referring to Figure 33, it can be seen that VISTA07 and SEQTS exhibit similar reported trip rates, while the Sydney HTS reports approximately 20% more trips.

This suggests that the "interview" approach employed in the Sydney HTS does indeed capture more trips. Moving to Figure 34, we can see that the Sydney HTS has higher trip rates for "discretionary" trips – Home Based Shopping, Home Based Recreation, Work Based Shopping, Shopping based Shopping – and Work Based Work. The trip rates for the "commuting" trip purposes – Home Based Work and Home Based Education – are generally similar across all three surveys.

It is notable that the trip rates in the VISTA07 and SEQTS surveys are very similar across purposes, consistent with their common methodology.

At first glance, this seems a likely explanation for the under-reporting of Off Peak travel in VISTA07, given that these purposes are predominantly undertaken in the Off Peak. However, the reality turns out to be rather more complicated.

If we perform the same analysis by distance travelled, rather that trip rate, then a completely different picture emerges. The distance travelled per person is in fact lowest in Sydney, with those in SEQ travelling significantly further than those in Melbourne and Sydney.

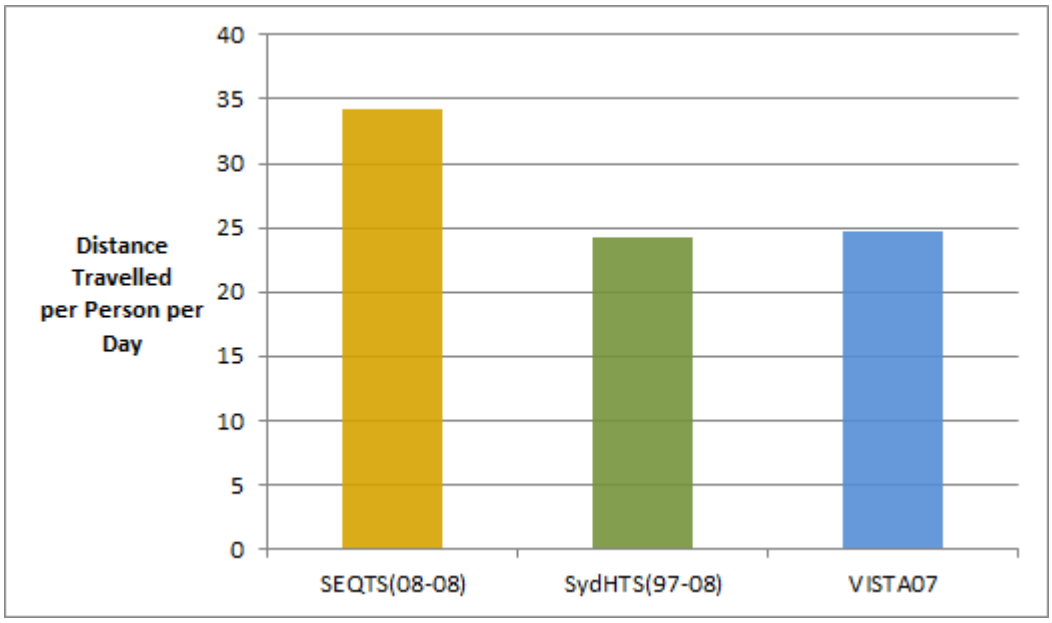

Figure 35 - Distance Travelled per Person

If we further disaggregate these trips by trip purpose (Figure 36), we find that the Sydney HTS and VISTA07 are remarkably similar across all purposes. This is quite astounding given the large differences in their trip rates.

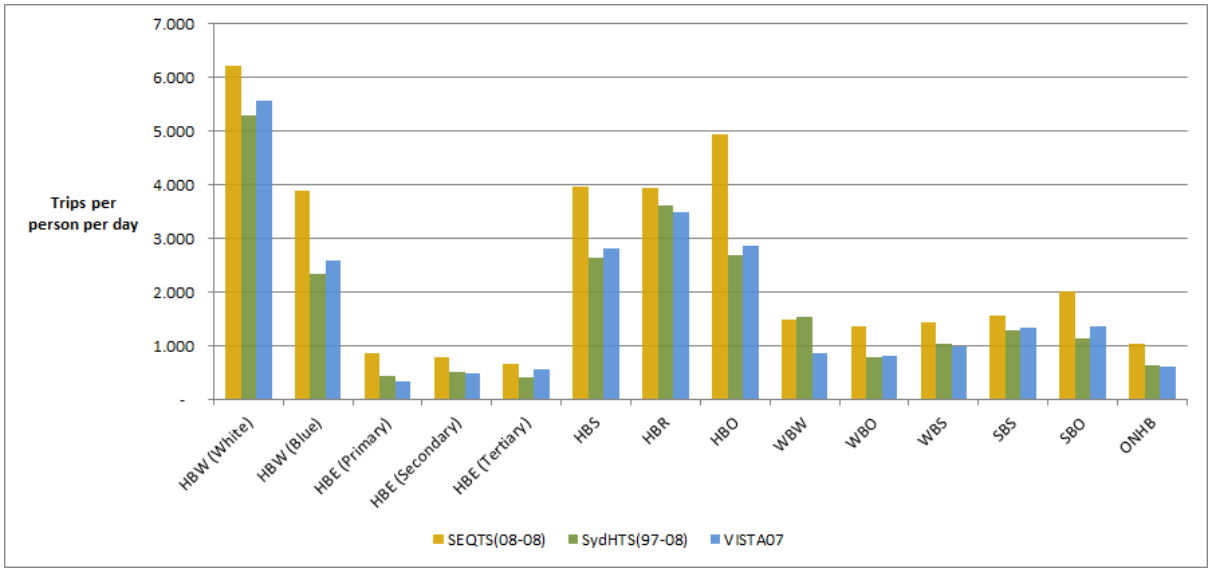

Figure 36 - Distance Travelled per Person per Trip Purpose

Clearly, the geometry of the survey region will have an impact on distances travelled. The SEQ region is particularly sprawling, which we expect explains the larger distances travelled by SEQTS respondents.

Nevertheless, the similarity between VISTA07 and the Sydney HTS is uncanny.

For example, the trip rate for Shopping Based Shopping trips in the Sydney HTS is 0.35, compared with 0.20 in VISTA07. However, the average trip length of these trips is 3.7km in the Sydney HTS, compared with 6.5km in VISTA07, so that overall, the distance travelled for SBS trips is almost identical in the two surveys.

While it is possible that Melbournians travel less frequently, but for longer distances, we suspect it more likely to be an artefact of the survey method.

We suspect that respondents to VISTA07 are leaving out small linked trips, but without corrupting the overall geometry of their journey. For example, take a trip to work with a stop at McDonalds on the way. In VISTA07, this might be reported simply as a trip to work (one trip), but in the Sydney HTS it might be reported as two linked trips (a home based shopping, and a work based shopping), each with a lower trip length.

This hypothesis implies certain substitution patterns among trip purposes. For example:

- A Home Based Work trip in VISTA07 might be represented as a Home Based Shopping + a Work Based Shopping trip in the Sydney HTS. The same goes for recreation or other
- A single Shopping Based Shopping trip in VISTA07 might be represented as a number of smaller Shopping Based Shopping trips in the Sydney HTS, as the respondent moves from one shop, or group of shops, to another

If we accept this hypothesis, then we would logically infer that the Sydney HTS is capturing roughly the same amount of travel as VISTA07 (at least in terms of distance), and should thus exhibit the same types of under-reporting as we have found in VISTA07.

Following this logic, we performed an assignment of the Sydney HTS using the Zenith model of Sydney, and compared the results to screenline counts (the same approach as that taken in Section 3.1, but excluding visitors). The result is shown in Table 12 below.

Period	Screenline Counts	Sydney HTS	Difference
AM peak	1,107,434	1,053,393	-5%
PM Peak	1,131,554	957,767	-15%
Off Peak	5,183,306	4,045,295	-22%
Total	7,422,294	6,056,455	-18%

Table 12 - Screenline Traffic Volumes by Period – Sydney HTS

The pattern is very similar to that observed in VISTA07. If we perform a consistent comparison (use VISTA07s person weights, consistent with the weights used for the Sydney HTS, and excluding visitors), we obtain the results in Table 13. The results are strikingly similar to those observed in Sydney, though it is notable that the degree of *off peak* under-reporting in the Sydney HTS is *less (22% instead of 33%).*

Period	VicRoads Screenlines	VISTA07 (person weighted excl. visitors)	Difference
AM Peak	1,539,946	1,493,762	-3%
PM Peak	1,641,379	1,474,821	-10%
Off Peak	7,205,675	4,838,706	-33%
Total	10,387,000	7,807,289	-25%

Table 13 - Screenline Traffic Volumes by Period - VISTA07 (person weighted, excluding visitors)

This suggests that the Sydney HTS is more accurate than VISTA07 when it comes to capturing off peak traffic. However, there is still a high degree of under-reporting in the Sydney HTS.

In light of this result, and also in light of the results of the GPS study described in the previous section, we suspect that those who *don't* respond to household travel surveys might be the cause of this seemingly universal under-reporting of off peak travel.

4.3 Regression Analysis

4.3.1 Inter Peak (9am – 3pm)

An alternative approach to asking what is "missing" in the off peak is to ask "what would we need more of" to best match the observed data we have. For example, if VISTA07 had twice the number of shopping trips, would it fix the issue of survey under-reporting?

Using statistical techniques, we can determine how much more of each trip purpose we would need in order to minimise the differences between VISTA07 and the VicRoads screenline traffic counts.

Given that under-reporting seems to be confined to the off peak, we have also limited our analysis to the off peak. The results for the inter peak period (9am – 3pm) are presented in Figure 37 and Table 14 below.

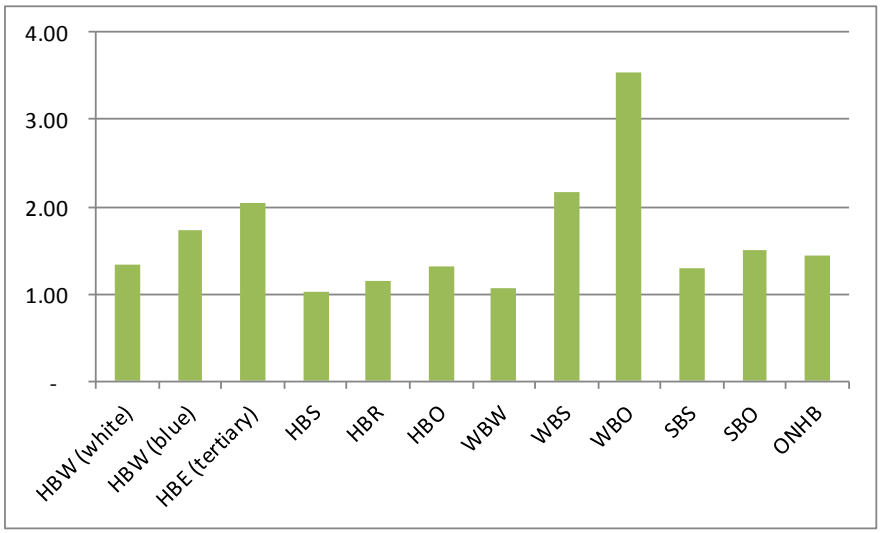

Figure 37 - Regression Based Correction Factors, by Trip Purpose (9am - 3pm)

Trip Purpose	Factor
HBW (white)	1.34
HBW (blue)	1.74
HBE (tertiary)	2.05
HBS	1.02
HBR	1.15
HBO	1.32
WBW	1.07
WBS	2.16
WBO	3.53
SBS	1.29
SBO	1.51
ONHB	1.44

Table 14 - Regression Based Correction Factors, by Trip Purpose (Inter Peak, 9am - 3pm)

The regression analysis suggests that to optimally match screenline traffic counts during the inter peak, we would need approximately 35% more home based work (white collar) trips, 75% more home based work (blue collar) trips, 15% more home based recreation trips, and 30% more home based other trips. We would also need 5% more work based work trips, 115% more work based shopping trips, 250% more work based other trips, 30% more shopping based shopping, 50% more shopping based other and finally, 45% more other non-home based trips.

Interestingly, the factor for home based shopping trips is only 1.02.

Summarised by grouped purposes, we obtain the following factors:

Grouped Purpose	Factor
Work related	1.62
Tertiary education	2.05
Shopping / rec / other	1.21

Table 15 - Regression Based Correction Factors, by Grouped Purpose (Inter Peak, 9am - 3pm)

The first thing to observe is that these factors align remarkably with the factors implied by our analysis of the Rail OD survey. As described in 3.2.1.2, the Rail OD survey collects trip purpose information during on-platform interviews. The results of these interviews implied that VISTA07 was only capturing half of work related and tertiary trips during the inter peak, but was approximately correct for shopping / recreation / other trips.

Both approaches suggest that only half of inter peak tertiary travel is being captured, and that work related trips are also highly under-reported. Both approaches also suggest that other trips are not highly under-reported (relatively speaking). Interestingly, the regression analysis suggests that non-home based trips are highly under-reported in VISTA07, which seems sensible.

The second thing to note is that these factors do not fully correct for inter peak under-reporting. Without the correction factors, VISTA07 is 35% low; with the correction factors, VISTA07 is 14% low. This is an interesting result, and suggests that our simple "by purpose" factors can only correct for so much.

To see why, recall from Section 3.1.4 that the under-reporting of traffic in the off-peak is greatest on low volume roads, and least on high volume roads. The key finding is re-presented in Figure 38 below. The updated curve, with the inclusion of the regression based correction factors, is then presented in Figure 39. Traffic on low volume roads is now only 20% low (rather than 40% low). However, traffic on the highest volume roads is now high, by 11%. In other words, our correction factors are *over-correcting* on high volume roads. If the factors were pushed higher, this would only worsen.

Over-correction is a good indicator that we are stretching the "by purpose" factors as far (and perhaps a bit further) than they should go. To go further, we will need to adopt a more refined factoring approach.

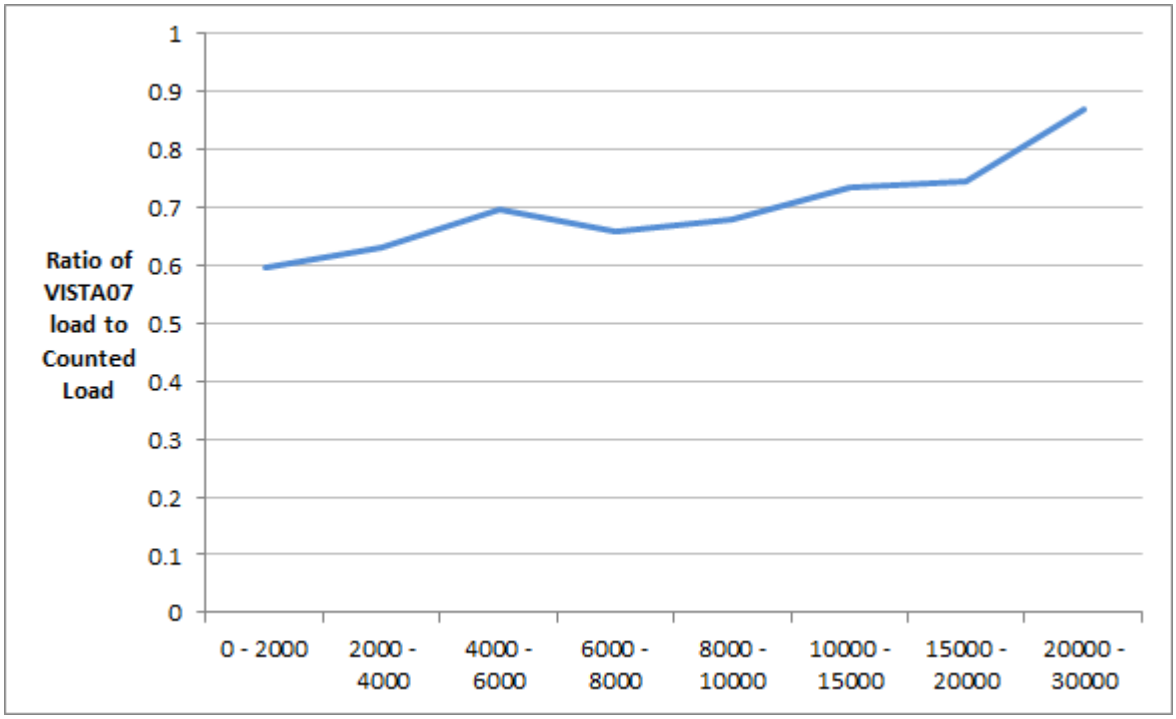

Figure 38 - Ratio of VISTA07 Traffic to Screenline Traffic, by Load Level (Inter Peak)

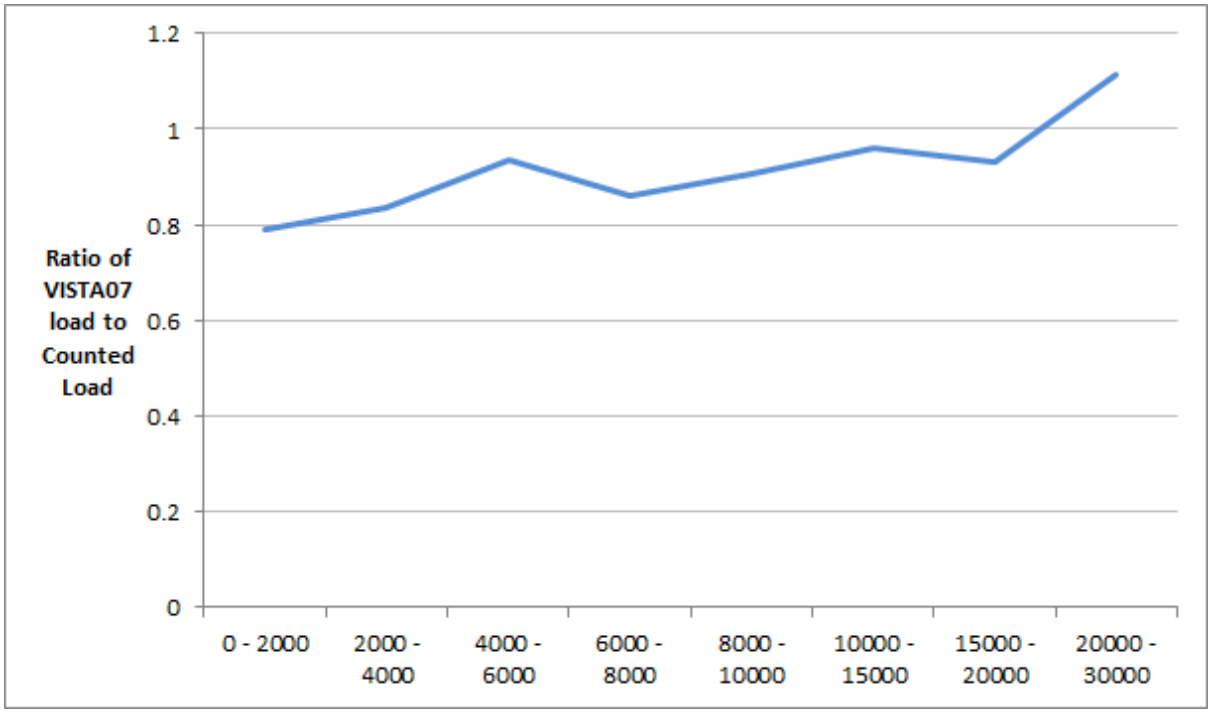

Figure 39 - Ratio of VISTA07 (with regression based correction factors) to Screenline Traffic, by Load Level (Inter Peak)

Take home based shopping as an example. There are several types of home based shopping trips:

- The grocery shop – *generally made close to home*
- The major shop – *generally at a high street or shopping centre*
- Shopping for pleasure – *generally made on high streets or in the CBD*
- Dropping in at the shops on the way to or from work – *generally made near the origin or destination, and without significant detour*

It could be that VISTA07 is missing a particular type of shopping trip – say grocery shopping. However, our simple "by purpose" factors cannot account for this.

It could also be that VISTA07 is missing shorter trips rather than longer trips, within certain trip purposes. Again, our current set of factors are limited.

The third thing to note is that these regressions are quite volatile – we are really stretching the limits of the VISTA07 sample. Because of this, the accuracy of these factors should not be overstated, and more work is required. In particular, we would like to improve the sample by including VISTA09 (when it is available), and by including additional count locations.

4.3.2 Post PM Peak (6 – midnight)

A similar regression has been run for the evening off peak, from 6pm to midnight. The following factors were derived:

Trip Purpose	Factor
HBS	1.77
HBR	1.50
HBO	1.55
WBS	1.81
SBS	2.00
SBO	1.47

Table 16 - Regression Based Correction Factors, by Trip Purpose (Evening Off Peak,6pm - midnight)

Purposes for which no factor is supplied are assumed to have a factor of 1.0.

The results indicate that shopping and recreation related purposes are under-reported in the evening off peak.

5 Adopted Correction Factors

5.1 Correction Factors

At this point in time we intend adopt the regression based approach as the basis for correcting VISTA07 in developing the Zenith model. However, we are fully aware that a lot work needs to be done in this area, if only because these factors only correct for about half of the off peak under-reporting.

Table 17 presents the "correction factors" which will be applied to VISTA07 in developing the Zenith model.

Trip Purpose	Inter Peak (9am - 3pm)	Evening Peak (6pm - midnight)
HBW (white)	1.35	1.00
HBW (blue)	1.75	1.00
HBE (tertiary)	2.00	1.00
HBS	1.05	1.75
HBR	1.15	1.50
HBO	1.30	1.50
WBW	1.10	1.00
WBS	2.10	1.80
WBO	3.00	1.00
SBS	1.30	2.00
SBO	1.50	1.50
ONHB	1.40	1.00

Table 17 - Correction Factors by Trip Purpose and Period

5.2 Validation

In this Section we briefly illustrate the effect of our correction factors, and re-validate against the various sources of count data.

Figure 40, below, presents the total traffic across screenlines by hour of the day, with our correction factors included in the "VISTA07 – VLC CF." data. It can be seen that the correction factors have significantly reduced the level of off peak under-reporting in VISTA07. Overall, they have corrected for about half of all under-reporting.

Figure 41 then presents the same analysis for total rail boardings compared with the 2009/10 Rail OD survey. Encouragingly, the factors appear to have corrected for off peak under-reporting on the rail network. Further analysis is needed to determine if we have "over-corrected" for certain types of rail travel.

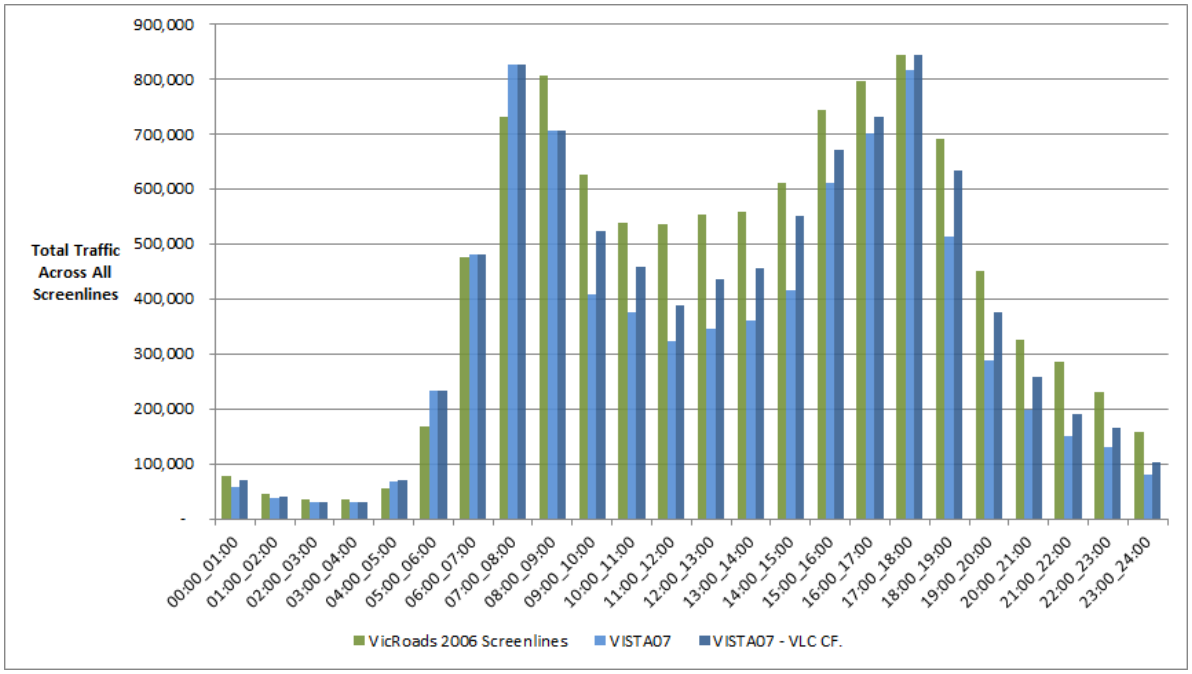

Figure 40 - Traffic Across All Screenlines by Hour of the Day

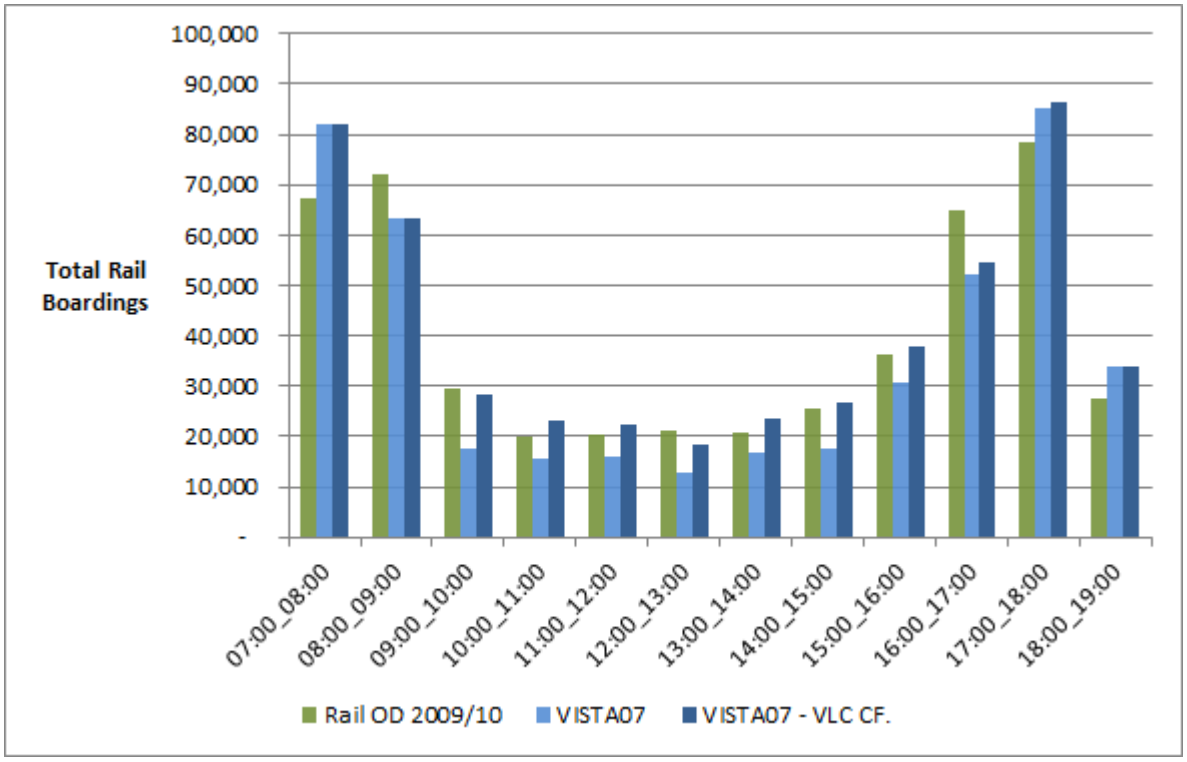

Figure 41 - Total Rail Boardings by Hour of the Day

Figure 42, below, presents an hour by hour analysis of total boardings on the tram network.

It can be seen that while the correction factors have helped somewhat, considerable under-reporting remains. The correction factors have particularly helped in the period between 9am and 11am. However, VISTA07 is so low during the period between 11am and 3pm, and post 7pm, that unreasonably high factors would be needed to match the Tram OD survey.

Further work is needed to explore whether there are problems with the Tram OD survey, or to understand why VISTA07 is so low.

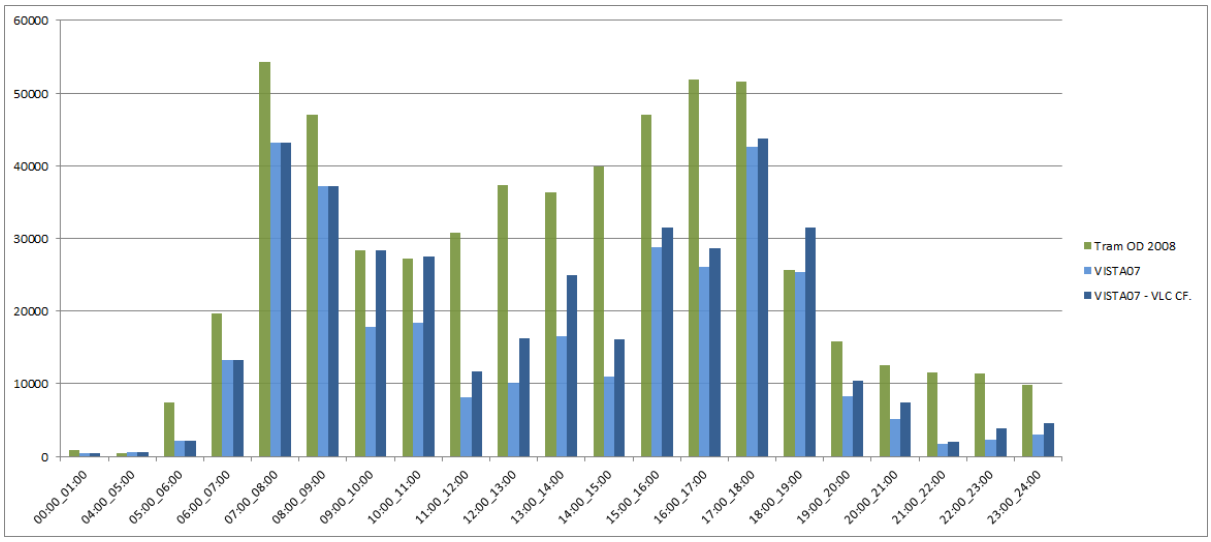

Figure 42 - Total Tram Demands by Hour of the Day

Finally, Table 18 presents total bus boardings by period. It can be seen that the correction factors have helped somewhat in the off peak. Further work will be required to determine why VISTA07 reports such low numbers of bus trips in the PM peak.

Period	Estimated Boardings (2008)	VISTA07	VISTA07 - VLC CF.
AM peak	74,106	62,463	62,463
PM peak	61,344	34,536	35,094
Off peak	203,975	147,138	175,593

Table 18 - Total Bus Demands by Period

6 Conclusion

The aim of this working paper has been to explore the extent to which the VISTA07 survey reflects travel in Melbourne.

We conclude that:

- VISTA07 accurately reflects travel in the AM peak, against almost all measures of AM peak travel demand, including traffic, rail boardings, tram boardings, and bus boardings.
- VISTA07 mostly reflects travel in the PM peak, though there does appear to be some degree of under-reporting, consistent across all modes. In particular, rail boardings on the suburban rail network, and bus boardings appear to be severely under-reported in VISTA07.
- VISTA07 significantly and systemically under-reports travel in the off peak. This pattern is consistent across all off peak hours, including the inter peak, and the evening off peak. The level of under-reporting is such that it is very difficult to infer exactly what types of travel are missing. Evidence from the Rail OD survey, and from regression analysis using screenline traffic counts, indicates that home based work travel might be under-estimated by between 30% and 50% in the inter peak, while travel for tertiary education appears to be under estimated by 50%. Non-home based travel also appears to be significantly under-reported, though it is unclear whether this applies to business travel (work based work).

In response to these findings, VLC has developed correction factors which make an attempt to correct for under-reporting. These factors are moderately successful, correcting for about half of all off peak under-reporting.

We do not, however, think we have gotten to the bottom of the under-reporting problem, and hope to further explore the issue with a larger survey sample that includes VISTA09.