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

What's missing from Australian household travel surveys? – Off-peak travel!

Tim Veitch[1], Michael Paech[2], Jeffery Eaton[3]

[1,2,3]80 Jephson St, Toowong, QLD, 4066

Email for correspondence: tim.veitch@veitchlister.com.au

Abstract

When Veitch Lister Consulting (VLC) redeveloped its Zenith model of Victoria, we investigated any biases that might exist in the 2007 Victorian Integrated Survey of Travel and Activity (VISTA07). Simply put, a model estimated using a biased survey will itself be biased!

We assigned the weighted survey data to the model transport network and compared the outputs to other observed measurements of travel demand such as traffic and person count data for various modes.

VISTA07 accurately reflects peak period travel, especially in the AM Peak, but under-estimates travel during the remainder of the day, during both the inter-peak and off-peak periods. Unadjusted for, this could cause transport models to under-predict off-peak travel.

To properly understand and adjust for this effect, we needed to understand its source, which could be:

- under-reporting of travel by survey respondents
- sample bias not accounted for in the usual weighting method
- travel which is by definition not part of the VISTA survey, such as visitor travel, travel made by professional drivers and freight vehicles.

To explore these causes, similar analyses were done in the Sydney Greater Metropolitan Area using the Sydney Household Travel Survey. Comparisons were made with trip purpose information from the Melbourne Rail OD survey, and we performed regressions using traffic counts.

A significant amount of the "missing travel" during off peak periods may be related to work and tertiary education commuting, and not discretionary travel as we had previously thought. This indicates undetected sample bias which is not accounted for in the usual weighting method.

1. Introduction

Household travel surveys are a primary source of data for the development of strategic transport models. It is important then for transport modellers to fully understand the strengths and weaknesses of these surveys in order to apply their data in a more robust way during the development and validation of transport models.

Large scale metropolitan household travel surveys in Australia are funded primarily by state and local governments. When planning these surveys there is an inevitable trade-off between cost, time and quality. This has led to different methodologies and survey frequencies from state to state depending on the available funds and the level of understanding by funding bodies of the importance of these surveys to transport planning and modelling.

It is outside the scope of this paper to critique in detail the various methodologies employed in household travel surveys. Much work has been done in this area by household survey practitioners both nationally and internationally. For example, Ashley, Richardson and Young

(2009) review various recent Australian and New Zealand household travel surveys, comparing their methodologies and assessing the levels of under-reporting detected in each.

Veitch Lister Consulting (VLC) has developed strategic transport models for all the major Australian metropolitan areas. In the development of these models, we have had to acquire a detailed knowledge of each of the household travel surveys conducted in these metropolitan areas as we try to understand the differences between surveys and their individual strengths and weaknesses.

In this paper we outline the results of our efforts to investigate the cause of under-reporting generally. We focus on the off-peak travel periods, as these periods appear to suffer most from under-reporting. The catalyst for this investigation was the recent redevelopment of the Zenith model of Victoria, and as such, the paper is largely concerned with the household travel survey data available for that area. However our comparisons with similar analysis done in the Sydney Greater Metropolitan Area are presented to demonstrate that off-peak under-reporting is not necessarily methodologically or geographically specific. This indicates a common, but as yet unaccounted source of under-reporting.

We do not attempt to offer suggestions for methodological improvement, but rather suggest areas for further investigation into alternative weighting schemes to help ameliorate the effect of under-reporting of off-peak travel. We present work we have done thus far in this area. We hope to report outcomes of further work being undertaken at the time of writing, on presentation of this paper at the conference in September 2013.

2. Victorian Integrated Survey of Travel and Activity (VISTA07)

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

The survey was conducted in the Melbourne Statistical Division, as well as the regional centres of Shepparton, Bendigo, Ballarat, Geelong and La Trobe Valley.

The scope and methodology for VISTA07 included the following:

- The survey was restricted to residents of the above study area, living in occupied private dwellings;
- The primary survey instrument was a set of forms completed by survey respondents, which included forms collecting household, person, vehicle, and trip making information. A travel diary was provided for each household member aged five years and above;
- Each household was assigned a specific ‘travel day’ on which each in-scope household member was asked to record their travel for the day;
- Telephone motivational calls, telephone and postal reminders were also used 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 Collection 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 variety of demographic data, together with details of all trips made by residents of the household aged 5 and over, including:

- Demographics: age, gender, employment status, educational status, other activity status, occupation, industry, personal income, car ownership, licence holding
- 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. This represents about 1% of the 2006 population.

The response rate for the survey was 47%.

3. Survey weighting

Sample surveys, particularly those with complex sampling schemes that involve stratification and clustering such as VISTA07, will inevitably require sample and expansion weights to be calculated to allow more accurate estimates of population characteristics. VISTA07 employed a two stage weighting process at the household and person level.

3.1 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 local government authority (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 2006 in terms of the number of households of each dwelling type and ownership status for each LGA.

3.2 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 population as contained ABS Census 2006.

Three control variables were used to calculate person weights: Home LGA; Gender (2 types: Male and Female); and Age Group (18 groups: 5-year groupings from 0-4 up to 85+)

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

3.3 Trip weights

Trip weights are designed to correct for certain types of under reporting of travel.

In their report: "VISTA07 Survey Procedures and Documentation" (The Urban Transport Institute, 2009), TUTI discusses three types of non-reporting of travel:

1. Deliberate non-reporting of travel considered unimportant by the respondent, or to reduce the effort in filling the 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 asking 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.

4. Survey validation using Zenith network assignment

A key step in the recalibration of the Zenith model has been the "validation" of VISTA07 – i.e. establishing how closely estimates from the survey data reflect 'reality'[1].

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 (State Government of Victoria, 2006)
- The 2009/10 Rail OD Survey (which included 7am-7pm station counts) (State Government of Victoria, 2010)
- The 2008 Tram OD survey (State Government of Victoria, 2010)
- 2008 estimates of bus route boardings (State Government of Victoria, 2008)

Together, these data sources provide a comprehensive view of motorised travel in Melbourne.

[1] It is of course understood that 'reality' in this context is based on other data sources which themselves have varying levels of inaccuracy.

Given that VISTA07 is 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 3, 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 VISTA using both the person weights and the adjusted trip weights, to explore the role played by the trip weight adjustments.

5. VISTA07 and screenline traffic counts

Weighted data from VISTA07 were assigned to the Zenith 2008 road network, with the resulting loads compared with VicRoads 2006 screenline traffic counts. VISTA07 has been weighted to 2006, allowing for a fair comparison.

The following assumptions are made in these comparisons:

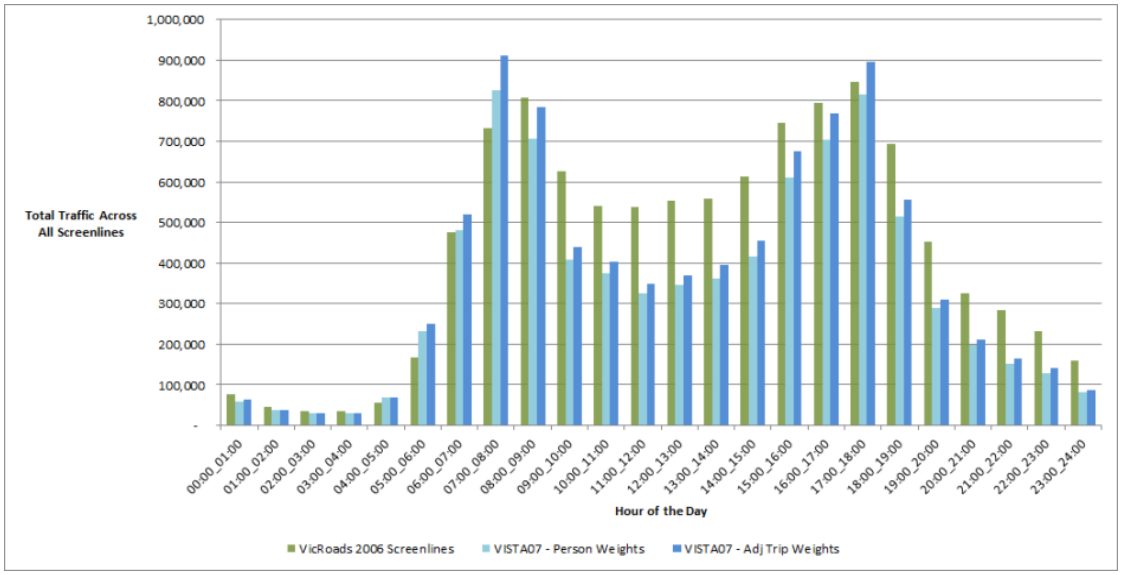
- VISTA07 data has been restricted to weekday, during school term times, and excluding public holidays and weights have been recalculated accordingly. This is consistent with our understanding of when the VicRoads screenlines are conducted.
- The Zenith traffic assignment process has been used to determine 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.
- Zenith model estimates of CV and visitor travel are robust.[2]

5.1 By hour of day

An initial comparison between the VISTA07 assignment (including Zenith CVs and visitors) and the screenline counts is provided in Figure 1. 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.

[2] To support this assumption, we have compared the Zenith model's prediction of CV flows with counted CVs, where classified counts exist. The Zenith model tends to over-predict commercial vehicle traffic on low volume roads, and is approximately correct on larger roads. The Zenith model does not appear to under-predicting commercial vehicle traffic.

Figure 1: Total Screenline Counts and Zenith Assignment of VISTA07 trips, Zenith Visitors and CVs



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

Between 9am and 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.

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

Based on these comparisons, it appears that the scale of traffic demands present in the weighted VISTA07 survey is consistent with VicRoads screenline traffic counts during the peaks, and in the pre-AM peak, but is 30-40% lower during the rest of the day.

5.2 By screenline

In 2006, VicRoads collected traffic counts on 21 screenlines. The total counted traffic across each screenline has been compared with the corresponding traffic flows generated using VISTA07 weighted data (using the adjusted trip weights), combined with Zenith CVs and visitors. The results (by time of day) are presented in Figures 2, with each data point representing a VicRoads screenline.

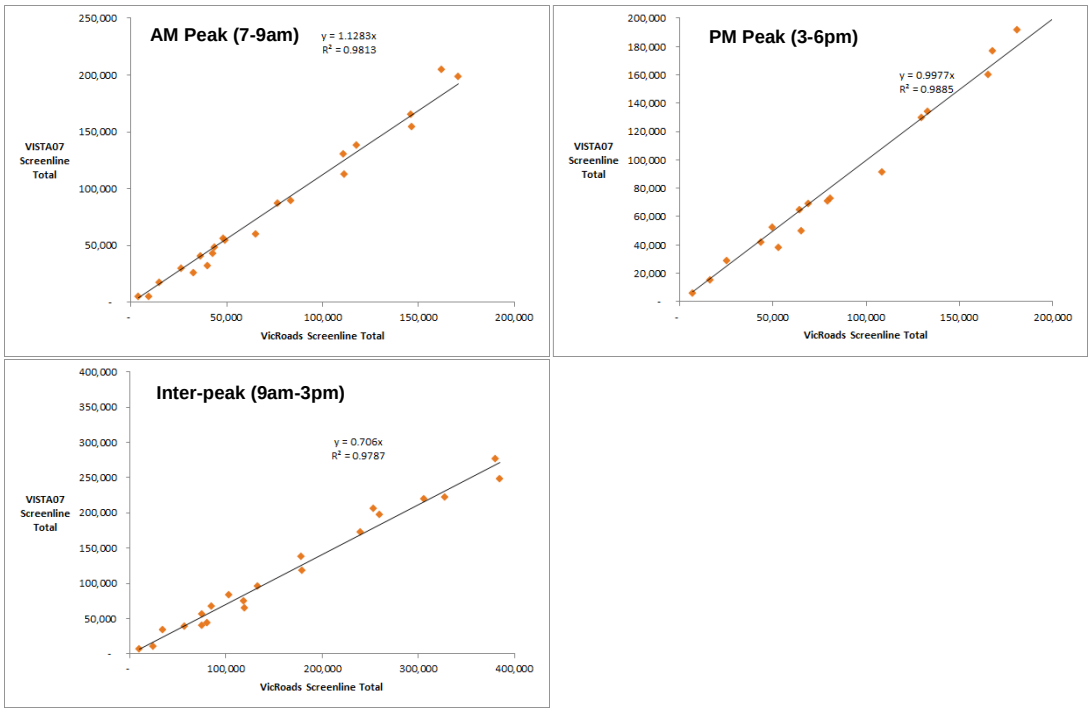
Two key observations can be made:

- 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.
- Consistent with the previous section, the VISTA07 assignment over-estimates 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. 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 is an argument against a lack of sample in VISTA07 as a plausible explanation, and hints at a more systematic effect.

Figure 2: Regression analysis of modelled loads using VISTA07 assignment and VicRoads screenline counts



5.3 By level of flow

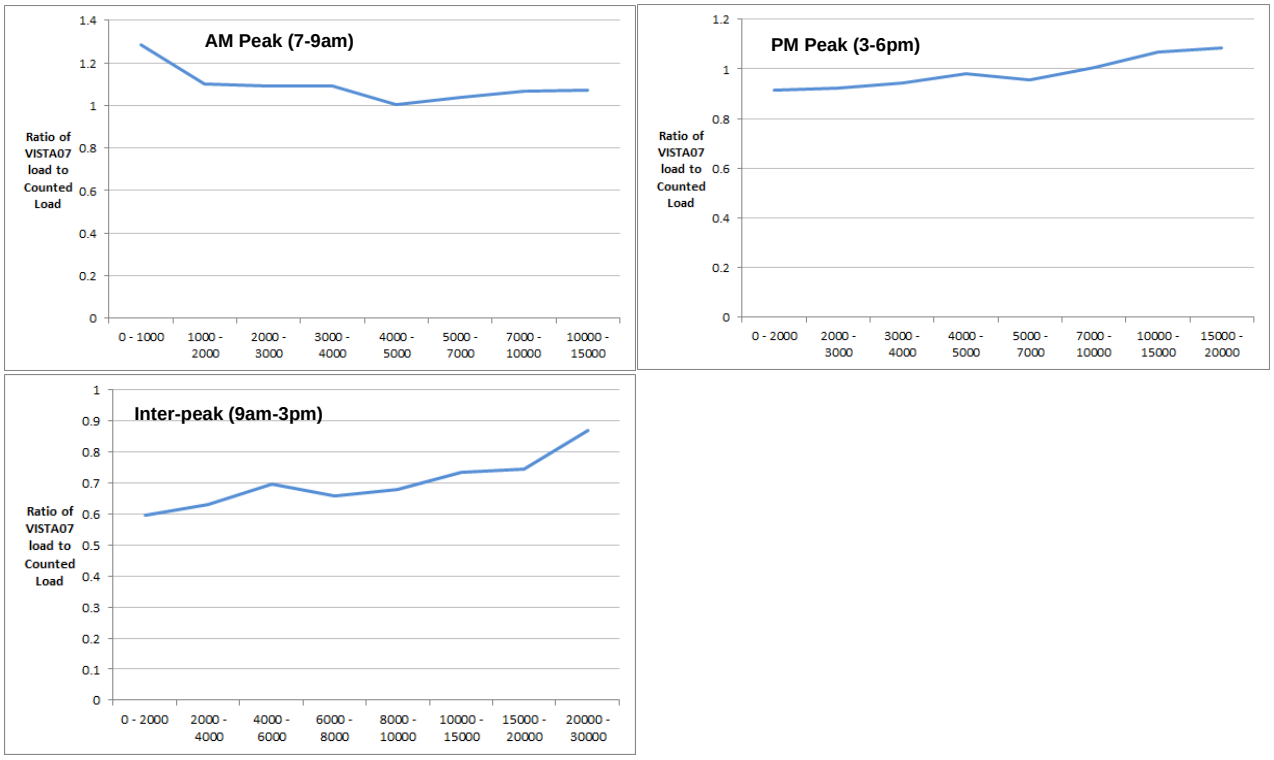
An exploration of differences between VISTA07 modelled traffic flows and screenline traffic counts for different levels of demand reveals some interesting patterns. This was done by calculating the ratio of aggregate volumes of VISTA07 traffic flows to screenline traffic counts, grouping roads by their load levels. The results are shown in Figure 3.

In the AM peak it can be observed that VISTA07 tends to over-estimate 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.

In the PM peak 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).

In the inter-peak this pattern is even more evident. 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. Combined with the previous analysis, it seems to suggest that the current weighting mechanism may be over-weighting the AM peak, and under weighting the inter peak, particularly trips which make use of lower order roads. We suspect that the lower order roads cater mostly for travel to-and-from the home, though we do not support that assertion here.

Figure 3: Ratio of VISTA07 modelled traffic to screenline counts, by time of day and load level



6. VISTA07 and recorded public transport demand

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.

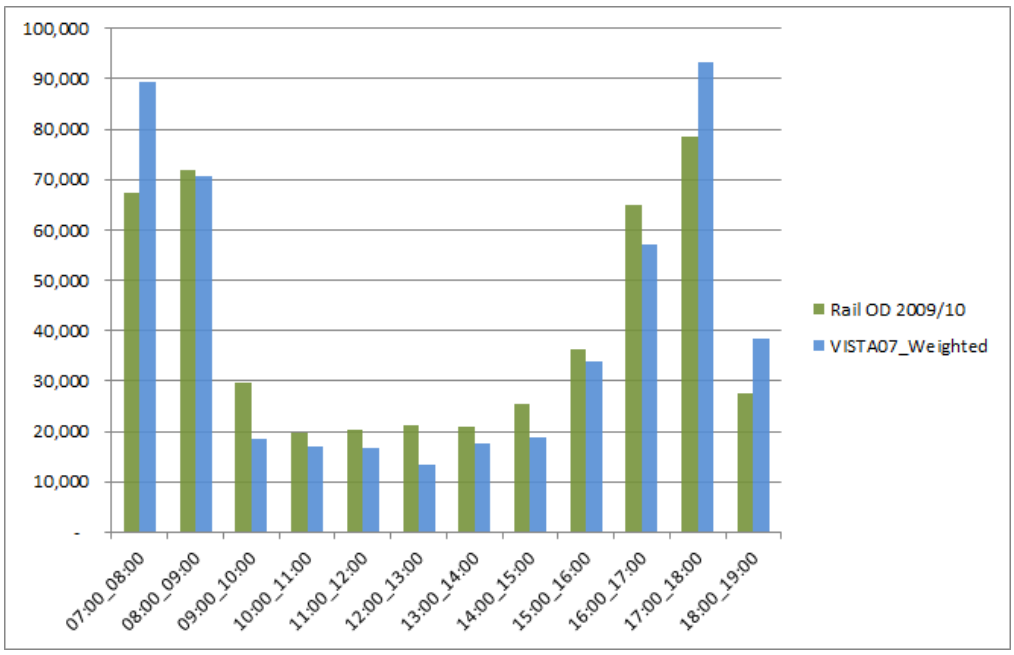
As with car trips, public transport trips have been weighted to 2006 population levels using two distinct approaches – a person weighted approach, which weights the survey sample to the population by LGA, gender and age, and an adjusted trip weighted approach, which includes extra correction factors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

6.1 Rail demand compared to the 2009/10 Rail OD Survey

The Rail OD survey, conducted in 2009 and 2010, collected information that is helpful in examining the under-reporting of travel in the off peak. The survey included a full count of station entries / exits between 7am and 7pm, as well as on-platform interviews with a sample of riders.

Figure 4 compares VISTA07 estimated boardings with counts performed as part of the Rail OD survey, by hour of the day, aggregated over all train stations.

Figure 4: Total rail boardings by time of day



An almost identical pattern emerged to that observed for traffic, though the level of under-reporting in the inter peak was not quite as pronounced.

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 may be missing in the off peak in VISTA07. The trip purpose information was particularly useful.

Figure 5, Figure 6 and Figure 7 show boardings by trip purpose for the AM peak, PM Peak and inter-peak respectively, for both VISTA07 and the Rail OD survey.

Figure 5: Total rail boardings by trip purpose (AM Peak)

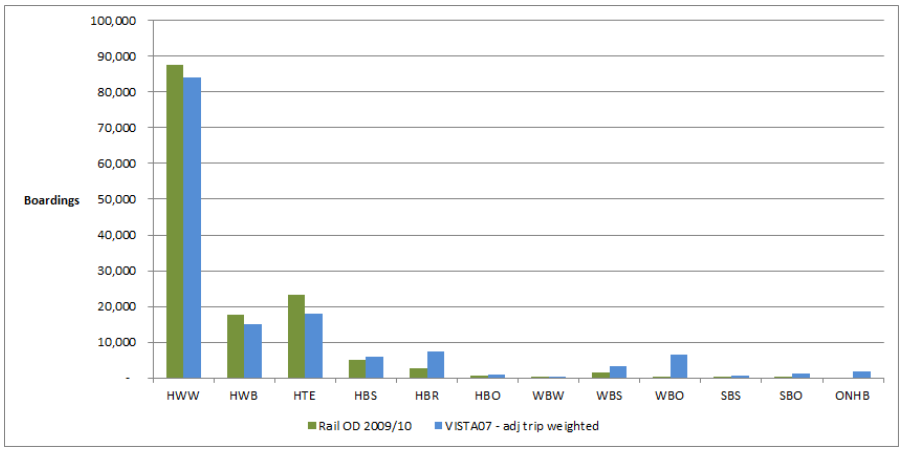


Figure 6: Total rail boardings by trip purpose (PM Peak)

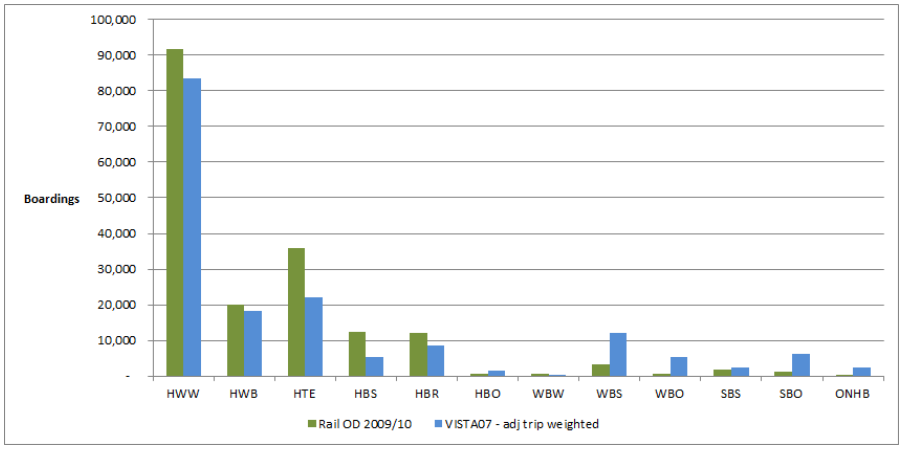
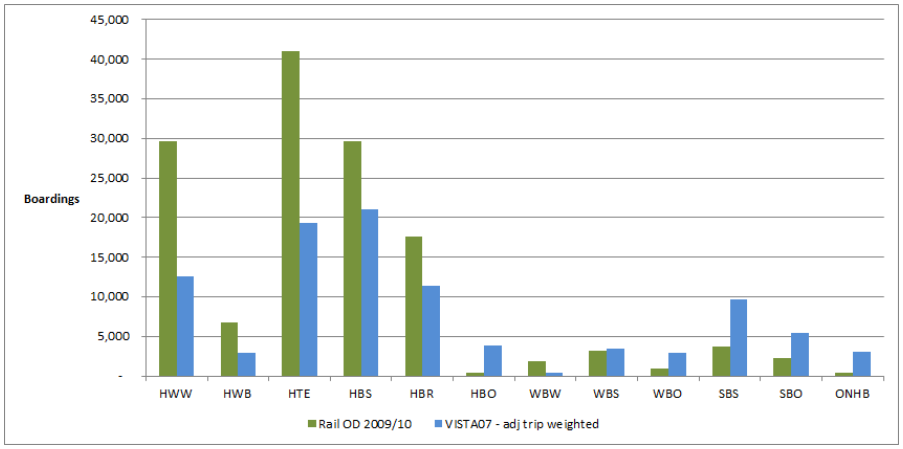


Figure 7: Total rail boardings by trip purpose (inter-peak)



In the AM peak, VISTA07 and the Rail OD survey are in very close agreement for home based travel, with some discrepancies in the amount of non-home based travel. These discrepancies are assumed to be related to differences in the way trip purpose information was collected in the two surveys. Overall in the AM Peak, the VISTA07 and Rail OD surveys are highly consistent.

In the PM peak, 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. Again, differences in survey methodology may be partly responsible for these differences.

In the inter-peak, the Rail OD survey suggests that VISTA07 is missing a large number of home based work (commuting) trips, and a large number of tertiary education trips. 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, the trip purposes were then aggregated into three groups: “work related”, “tertiary education” and “other”. The result (for the inter-peak only) is presented in Figure 8 below.

A common assertion amongst users of household travel survey data has been that “commuting” trips – Home Based Work and Home Based Education – are reported accurately in HTS (explaining why the peaks were reflected accurately), and that the off peak

under-reporting is mostly caused by under-reporting of discretionary travel such as Home Based Shopping, Home Based Recreation, Shopping based Shopping, etc.

This comparison with the Rail OD survey suggests that VISTA07 may instead be under-estimating the number of work and tertiary education related trips during the inter-peak, at least on the rail network. An obvious question is whether the same principle could apply to road traffic. In Section 8, evidence is presented that suggests it does.

6.2 Tram and bus demand

VISTA07 tram and bus demands have been compared with surveyed tram demands from the 2008 Tram OD survey and estimates of bus boardings. Due to space constraints we do not include a lengthy discussion of these findings here.

However we can note that despite significant differences in scope that make comparisons between VISTA07 and Tram OD data particularly difficult, a similar pattern can be observed. Table 1 shows the differences between modelled demand using VISTA07 data and tram and bus boarding estimates derived from other data sources. In general they support the comparison of VISTA07 modelled demand to rail patronage.

Table 1: Differences in tram and bus demand by period (VISTA07, TRAM OD, Bus boarding estimates)

Period	Tram	Bus
AM Peak	-11%	3%
PM Peak	-27%	-35%
Off-peak	-48%	-13%

7. Examination of under-reporting in VISTA07 and Other HTS

Weighted VISTA07 estimates of inter-peak and evening off peak travel appear to be considerably lower than estimates made using other observed measures such as traffic counts, rail and tram passenger counts and bus validations. This pattern has been found consistently across all motorised modes of travel (road, rail, tram and bus), and has been found to be 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 difference, and develop a working hypothesis about the types of trips that may be under-estimated in the weighted VISTA07 estimates.

7.1 Why VISTA07 may report lower off-peak travel demand

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

A. Under-reporting by survey respondents, such as:

i. The respondent deliberately leaving 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. Sample bias not accounted for in the usual weighting method (the survey sample being biased toward people who travel less during the middle of the day)

C. Travel which is by definition not part of the VISTA survey, such as visitor travel, travel made by professional drivers and freight vehicles.

TUTI correct for effects A) (ii and iii) through the weights which are built into the VISTA07 adjusted trip weights as described in Section 3.

VLC has also made some effort to compensate for C) by including Zenith estimates of visitor and commercial vehicle traffic.

While it is impossible to rule out these effects as possible causes, the degree of difference during the off peak suggests that A)(i) and B) must be at least partly responsible.

Explanation A(i) is an obvious explanation, given that VISTA07 uses a self-completion survey method. It is technically quite easy for respondents to leave out trips.

VISTA07 included extra data quality steps to attempt to deal with A) (i and ii) through the use of targeted clarification interviews. As data is entered after collection of response forms, the data entry system is designed to identify anomalies such as very long trip times, or excessive time spent in a single location amongst many others. These anomalies triggered phone calls in which the interviewer inquires as to whether the respondent visited other places not reported in the diary.

Even with the extra effort in clarifying travel diaries with responding households, it is highly plausible that trips that have been left out of the original diaries are not successfully captured.

On the matter of explanation B) (sample bias), VISTA07 has a response rate of 47%, and as such, sampling bias could have an impact. TUTI's weighting methodology attempts to correct for demographic and spatial biases related to age, gender, household ownership type, and dwelling type. However, it is possible that sampling bias remains.

The following sub-sections attempt to further explore the above issues.

7.2 GPS validation in VISTA07

During VISTA07, GPS devices were given to 85 survey respondents. 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 trip legs were missing from the travel diary but present in the GPS data, while 20% of trip legs recorded in the diary were missing from the GPS device. It is clear that the use of GPS devices had, at this stage, some technological and methodological issues to iron out.

The TUTI analysis (TUTI, 2009) 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.

This suggests that under-reporting of travel by survey respondents does occur to some degree, with non-work and non-home based trips making up the bulk of un-reported trips. However the final report for VISTA07 concludes that "the weighting and analysis procedures built into the VISTA07 survey accounted for any Non-Reported-Stops that might exist in the raw data collection procedures" (Ashley, Richardson, Young, 2009, p.6).

It therefore appears that item non-response (trip under-reporting) may not be the main contributor to the observed shortfall in off-peak travel.

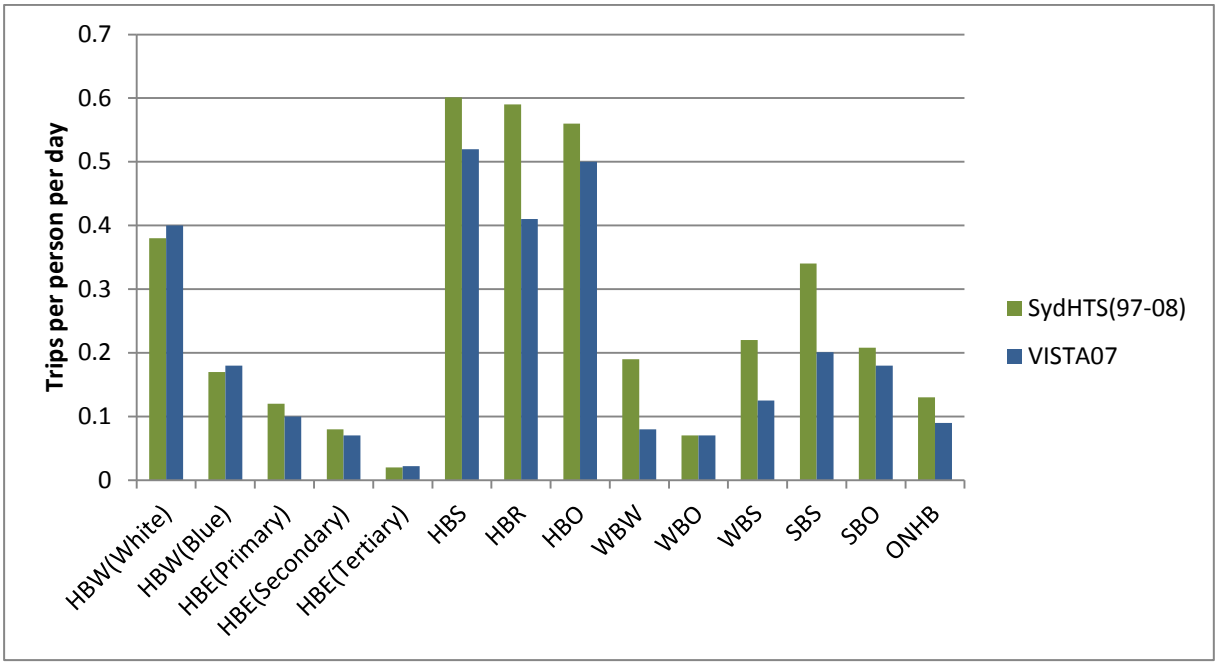
7.3 Comparison with Sydney household travel survey 2008

We compared VISTA07 with the Sydney Household Travel Survey 2008, incorporating data collected from 2006-08 (Sydney HTS). The Sydney HTS uses a hybrid methodology, with the primary data collection method being an in-house interview on the day of travel, using travel diaries as memory joggers during the interviews. It is assumed that this methodology should significantly reduce the amount of under-reporting.

As expected, the Sydney HTS reports a higher trip rate (3.7 trips per person per day) than VISTA07 (just over 3 trips per person per day), suggesting the more intensive interview methodology captures more trips.

In Figure 8, 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 between the surveys.

Figure 8: Person trip rate by purpose (Sydney HTS 2008 and VISTA07)



At first glance, this seems a likely explanation for the under-reporting of Off Peak travel in VISTA07, given that trips for these purposes are predominantly undertaken during the Off Peak.

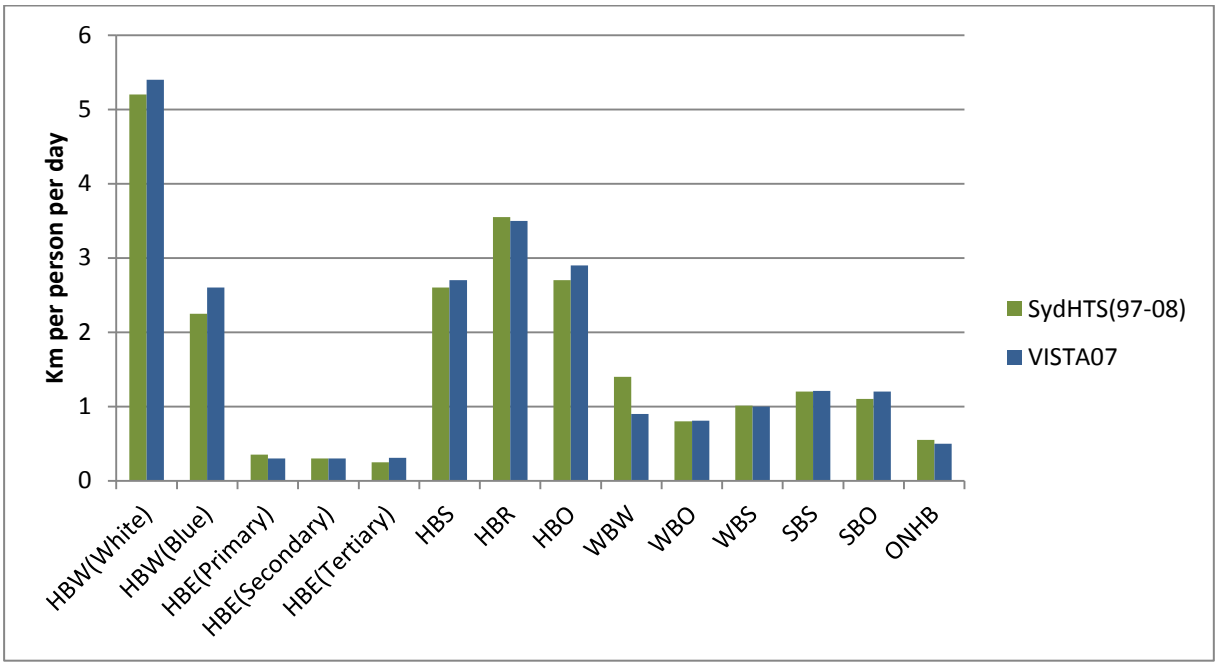
However, when the same analysis was performed using distance travelled, rather that trip rate, a completely different picture emerges.

In Figure 9, we find that the Sydney HTS and VISTA07 are remarkably similar across all purposes. This is surprising given the large differences in trip rates.

While it is possible that this is a real effect, caused by differences between Sydney and Melbourne, we suspect that it is more likely an artefact of the survey method. We suspect that VISTA07 respondents might be leaving out short trips that make up an overall "trip-chain", without corrupting the overall geometry of their journey. This is consistent with TUTI's comparison of GPS captured trips and travel-diary trips.

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

Figure 9: Person distance rate by purpose (Sydney HTS 2008 and VISTA07)



To test this, 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 5, but excluding visitors). The result is shown in Table 2 below, with equivalent VISTA07 results presented in Table 3.

Table 2: Screenline traffic volumes by period – Sydney HTS

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 3: Screenline traffic volumes by period – VISTA07

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%

The pattern is remarkably similar between the two surveys, though the Sydney HTS does appear to capture more off peak (and less peak) travel. Nonetheless, the Sydney HTS may still under-estimate off peak travel.

In light of this result, and considering the results of the GPS study described in the previous section, it appears that sample bias could be a common source of difference between household travel surveys and other observed sources of data.

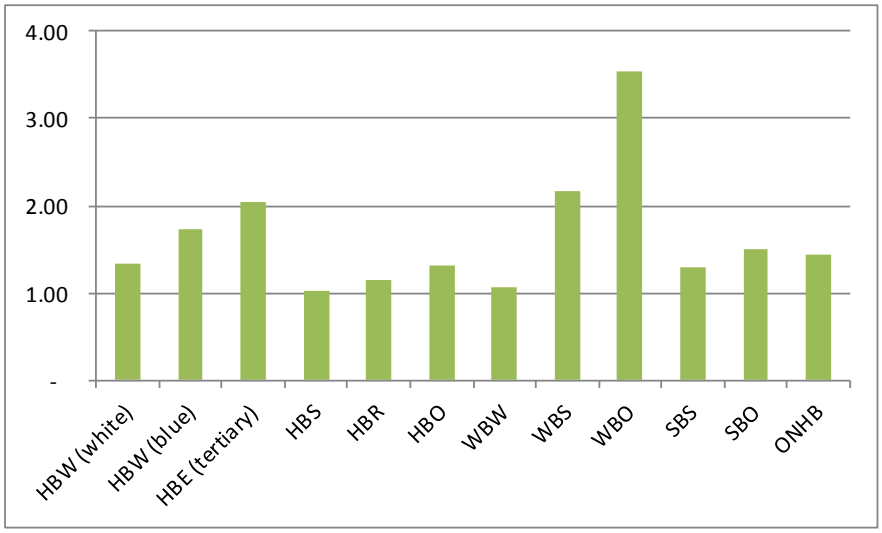
8. Using regression analysis to derive under-reporting factors

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 other observed measures of demand. 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. Figure 10 shows factors the (by trip purpose) which if applied to the VISTA07 inter-peak trips, would minimise the differences between VISTA07 and the VicRoads screenline traffic counts, at the level of individual counts per hour of the day.

Figure 10: Regression based correction factors, by trip purpose (9am – 3pm)



Summarised by grouped purposes, we obtain the following factors:

Table 4: Regression based correction factors, by grouped purpose (9am – 3pm)

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

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 section 7, 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.

The regression analysis presented above also suggests that non-home based travel (especially work based travel) is under-estimated.

A similar regression has been run for the evening off peak, from 6pm to midnight. Table 5 shows the correction factors adopted.

Table 5: Adopted correction factors

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

9. Post factor adjustment validation

Figure 11 presents the total traffic across screenlines by hour of the day, with the regression based 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 12 presents the same analysis for total rail boardings compared with the 2009/10 Rail OD survey. 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 11: Traffic across all screenlines by hour of the day

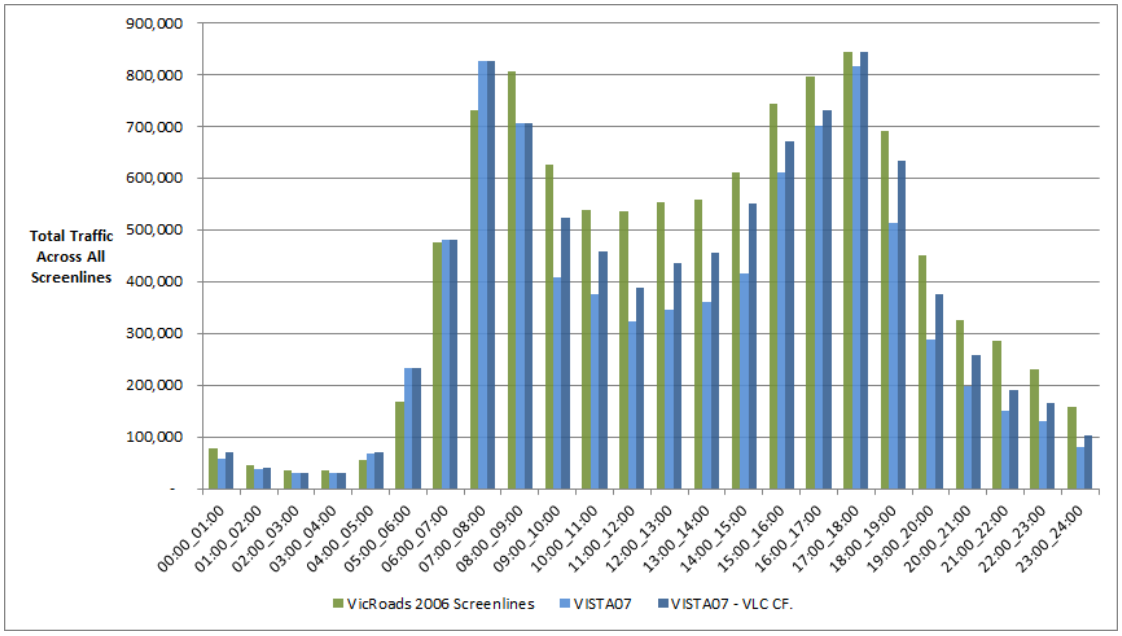
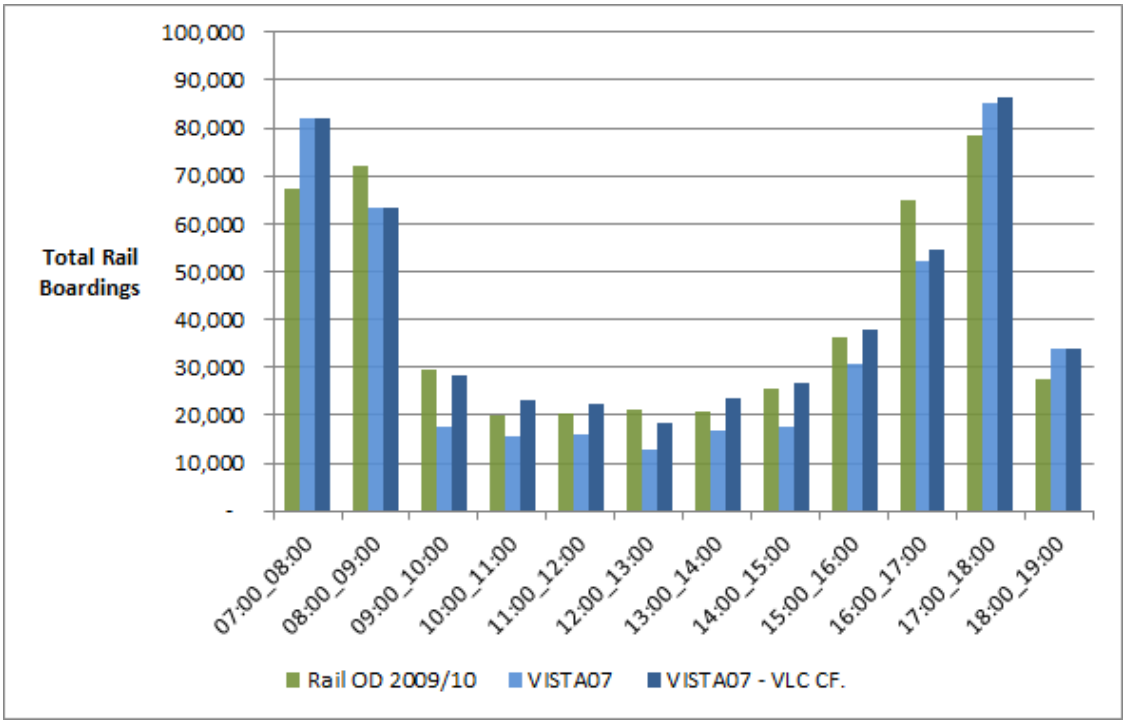


Figure 12: Total rail boardings by hour of the day



10. Conclusions

The aim of this paper has been to explore the extent to which the household travel surveys reflect travel in their respective cities. It has focused on one area, Melbourne. However a comparison with a similar in Sydney has been informative.

When link loads are derived by assigning weighted HTS data to model networks and these are compared to independent data sources such as traffic counts and public transport passenger surveys it has been found that:

- The surveys investigated accurately reflect travel in the AM peak, against almost all measures of AM peak travel demand, including traffic, rail boardings, tram boardings, and bus boardings.
- With some exceptions, they also reflect travel in the PM peak, though there does appear to be some degree of under-reporting, consistent across all modes.
- There appears to be significant, systematic under-reporting of 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. In Melbourne's case, 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).

The authors do not believe the surveys investigated are deficient either in methodology or in execution. However it appears that, despite the survey practitioner's best efforts, bias may be introduced by higher non-response by people who travel to and from work in the middle of the day and tertiary students, amongst other cohorts. The 'standard' means for dealing with non-response (household and demographic weighting) and under-reporting (trip-leg weight calculation) do not appear to account for this shortfall.

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

Further work is required investigating more recent surveys in the cities discussed in this paper as well as other areas in Australia to reach more concrete conclusions on whether this is a problem in across household travel surveys generally. The authors also intend to continue work on more robust weighting schemes to supplement the standard weighting schemes currently in use.

Acknowledgements

Policy and Communications Division, Department of Transport, State Government of Victoria

Information Services Team, Public Transport Division - Department of Transport, State Government of Victoria

Bureau of Transport Statistics, New South Wales Government

Tony Richardson, The Urban Transport Institute

References

Ashley, D, Richardson, A.J. and Young, D., 2009. Recent Information on the Under-reporting of Trips in Household Travel Surveys, *32nd Australasian Transport Research Forum*, Auckland, NZ, Sept. 2009.

Richardson, A.J., 2009. VISTA07 Procedures Report, Report to Victorian Department of Transport, *TUTI Report 66-2009*, The Urban Transport Institute, Victoria.

State Government of Victoria, 2006. Traffic count data at screenlines, *unpublished data*

State Government of Victoria, 2008. Victorian Integrated Survey of Travel and Activity 2007 - 2008 (VISTA07), unpublished database

State Government of Victoria, 2008. 2008 Tram Origin Destination Survey, unpublished database

State Government of Victoria, 2008. Bus boarding estimates (supplied), unpublished data

State Government of Victoria, 2010. 2009/10 Rail Origin Destination Survey, unpublished database

Bureau of Transport Statistics, New South Wales Government, 2008, Sydney Household Travel Survey 2008, unpublished data (supplied)