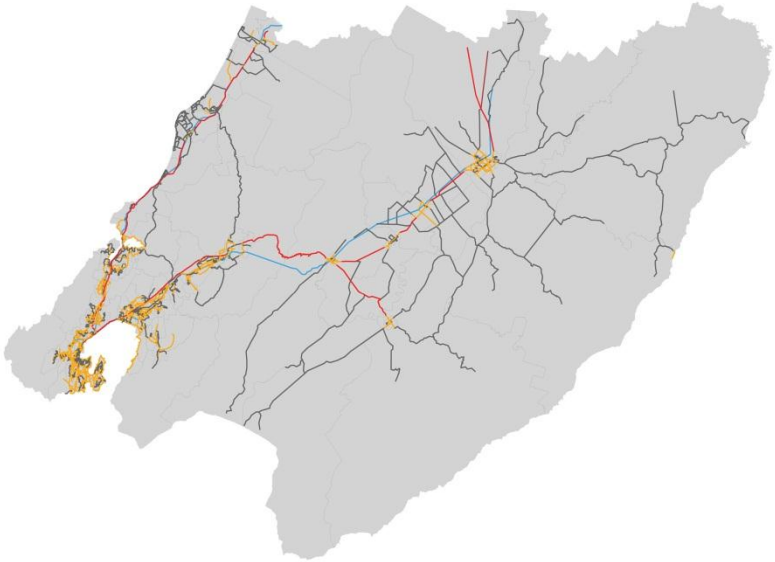
OPUS INTERNATIONAL CONSULTANTS AND ARUP

WELLINGTON TRANSPORT MODELS

Contract No C3079



TN2: Survey Sampling Methodology

Date: December 2012

ARUP OPUS

Wellington Transport Models

TN2 : Survey Sampling Methodology

prepared for

Greater Wellington Regional Council

Prepared By

Marius Roman (Arup)

Reviewed By

Bruce Johnson (Arup)

Opus International Consultants Limited
Wellington Office
Level 9, Majestic Centre, 100 Willis Street
New Zealand
Ph: +64 4 471 7000

ARUP

Arup
Level 17, 1 Nicholson Street
Melbourne VIC 3000
Australia
Ph: +61 3 9668 5500

Date:	December 2012
Reference:	g:\localauthorities\wrc\proj\5-c2050.00 - c3079 wtsm wptm\600 deliverables\630 final tech notes\tn2 survey sampling methodology final.docx
Status:	Final
Revision:	1

Document History and Status

Issue	Rev	Issued To	Qty	Date	Reviewed	Approved
Memo	-	Peter Dunn	-	28/07/2011	Marius Roman	Bruce Johnson
First draft	1	Nick Sargent - GW	-	8/03/2012	Bruce Johnson	David Dunlop
Final	1	Nick Sargent - GW	1 Hard & 1 CD	06/12/2012	Bruce Johnson	David Dunlop

This report takes into account the particular instructions and requirements of our client. It is not intended for and should not be relied upon by any third party and no responsibility is undertaken to any third party.

John Bolland: _____
(Peer Reviewer)

Nick Sargent: _____
(GWRC)

Contents

1 Introduction

Opus International Consultants Limited (Opus) and Arup Australia (Arup) were commissioned by Greater Wellington Regional Council (GWRC) to rebase the existing 2006 Wellington Transport Strategy Model (WTSM) to a new base year of 2011. Opus updated the WTSM while Arup developed a Wellington Public Transport Model (WPTM) based on figures from WTSM and detailed public transport surveys. The whole process of model updates and development is complex and involves several steps which have each been individually reported in a series of technical notes.

This document describes the methodology used in the selection of the rail and bus routes to be surveyed as part of the data collection exercise for the project. The resultant services that we consider important to capture, based on this methodology, have also been listed and a simple statistical analysis has also been undertaken to check how many services will need to be surveyed to obtain a reasonable sample size. Finally, estimated costs are provided as a means of checking that the data collection task is not beyond the allocated budget.

This note is intended to give guidance as to where resources should be directed and the number of responses required to provide meaningful results. It is not intended to provide specific advice on planning or conduct of the surveys and it is anticipated that the service provider conducting the surveys will evaluate staffing requirements to produce the best outcome for the given resources available.

2 Methodology

The aim of the methodology adopted was to identify the routes to be surveyed which will allow reasonable conclusions regarding the travel characteristics and origins and destinations of rail and bus users to be made whilst minimising the resource / cost implications of the data collection exercise.

The methodology is primarily based on a qualitative assessment of the attributes of individual rail and bus services. The attributes of individual services that were considered important to capture as part of the overall routes to be surveyed are as follows:

- Geographic coverage;
- Routes servicing special generators such as hospitals, educational institutions and the airport;
- Stopping pattern of services (e.g. express versus all-stoppers);
- Service frequencies; and
- Bus routes servicing modal interchanges.

Building on this approach it was also identified that it was reasonable to preferentially omit some services where usage characteristics could be reasonably inferred from nearby services with similar characteristics. For example, the catchment and user characteristics for local station feeder bus services in outlying areas can reasonably be estimated based on the limited geographic spread of catchment zones and applying user patterns relating to nearby services of a similar nature.

A supporting quantitative assessment of required sample sizes has also been undertaken to understand the number of service hours required to be surveyed. The quantitative analysis uses rail boardings from surveys undertaken in June 2011 and Electronic Ticketing Machine (ETM) bus patronage data along with specified confidence intervals and margins of error to estimate the required size of the population to be sampled. This is then compared to the number of completed forms that are anticipated to be returned based on the results of the bus and rail pilot surveys.

2.1 Statistical Considerations

A statistical approach to guide selection of appropriate sampling size has been applied and is described below. It is emphasised throughout that this methodology is for high level planning purposes only and does not represent the statistical level of confidence that will be achievable in estimating some of the more disaggregate elements of the model.

At the lowest level of detail the modelling process will be representing trips by purpose between individual origin destination pairs across the whole network. The sparseness of this data, especially in relation to public transport trips, where many cells will be of low or zero value, means within practical budget constraints it is impractical to define a statistically significant data collection approach at this level of detail. This issue is common to strategic models in general and already applies to specific elements of WTSM.

Given the above we have adopted the general approach that we wish to be reasonably confident of key parameters of the population using individual stops or routes on the

network. In the case of trains this is directly related to survey methodology which collects data at each stop with a known population size in each peak period – using this approach we can estimate, for example, the sample size required to be confident about the distribution of access modes for users of individual stations. For bus routes, where population estimates relate to the overall route usage, the approach is less directly related to individual stops, but provides some measure of confidence in determining how characteristics such as trip purpose may vary across individual routes. Such information will be useful in determining appropriate strategies for ultimate weighting of the sample data against ETM data of stop to stop travel.

The formula[1] used for the calculation of the required sample size is:

$$n = \frac{X^2\ N\ P\ (1-P)}{(ME^2\ (N-1)) + (X^2\ P\ (1-P))}$$

Where,

n = sample size

X^2 = Chi-square for the specified confidence level at 1 degree of freedom. For the analysis presented below a 95% confidence level has been specified.

N = population size

P = population proportion for the parameter of interest. The worst case for sample size is for a 50% proportion and this value has been used in the analysis presented below.

ME = desired margin of error. A value of 10% has been adopted in the analysis.

2.2 Survey Period

The public transport (PT) intercept surveys are envisaged to commence in early August 2011 and are intended to capture weekday rail and bus users travelling during the AM and Inter peak travel periods. These are defined as 7am-9am for the AM peak and 11am-1pm for the Inter peak (IP). Within this document references to AM and IP periods refer to these times. The analysis of ETM data is based on reporting trips with a boarding time falling within these specified times.

The duration of the overall survey period and the scheduling of the surveys to be undertaken on the suggested services will be at the discretion of the data collection services provider.

[1] Krejcie & Morgan, *Determining Sample Size for Research Activities*, Educational and Psychological Measurement, #30, pp. 607-610

3 Selected Routes

3.1 Rail Services to be Surveyed

Given the proposed survey method (i.e. surveyors positioned on platforms as opposed to in-vehicles) all stations within the Metlink service area will need to be included in the rail intercept surveys. This excludes the stations on the Johnsonville Line surveyed during the AM peak period as part of the rail pilot survey undertaken on Wednesday 29th June 2011. This implies a total of 45 stations during the AM peak and 53 stations during the IP period.

The results of the pilot survey on the Johnsonville Line during the AM peak period are given in Table 3-1. This information was used as part of the quantitative analysis undertaken for the scoping of the rail surveys reported in Table 3-2. The values in this table are derived by determining how many valid questionnaires would be returned at each station if the proportion completed is the same as in the pilot.

Table 3-1: Results of Pilot Survey on Johnsonville Rail Line, AM Peak Period

		Number of Forms			
Boarding Station*	**Total in Pilot**	**Distributed**	**Refused**	**Completed[1]**	**Total in Pilot**
Johnsonville	245	153	92	91	37%
Raroa	71	55	16	43	61%
Khandallah	161	143	18	73	45%
Box Hill	54	32	22	19	35%
Simla Crescent	199	186	13	57	29%
Awarua Street	178	174	4	117	66%
Ngaio[2]	166	139	27	80	48%
Crofton Downs	175	172	3	92	53%
Total	**1,249**	**1,054**	**195**	**572**	**46%**

* Excludes Wellington Station

1 'Completed' defined as responses with valid trip purpose and valid origin and destination responses at better than suburb level

2 Buses replaced trains on the 7:31am and 8:25am services to Wellington

Table 3-2: Rail Stations to be Surveyed and Required Sample Sizes

	Total Boardings*		Required Sample Size[1] (Percentage of Total Boardings)		Anticipated Completed Returns[2]		Anticipated less Required	
Station/Line	**AM**	**IP**	**AM**	**IP**	**AM**	**IP**	**AM**	**IP**
Johnsonville								
Johnsonville	n/a	66	n/a	41 (62)	n/a	30	n/a	-10
Raroa	n/a	8	n/a	9 (114)	n/a	4	n/a	-5
Khandallah	n/a	5	n/a	9 (183)	n/a	2	n/a	-7
Box Hill	n/a	9	n/a	9 (102)	n/a	4	n/a	-5
Simla Crescent	n/a	25	n/a	23 (92)	n/a	12	n/a	-12

Station/Line	Total Boardings* AM	Total Boardings* IP	Required Sample Size[1] (Percentage of Total Boardings) AM	Required Sample Size[1] (Percentage of Total Boardings) IP	Anticipated Completed Returns[2] AM	Anticipated Completed Returns[2] IP	Anticipated less Required AM	Anticipated less Required IP
Awarua Street	n/a	74	n/a	41 (55)	n/a	34	n/a	-7
Ngaio	n/a	22	n/a	17 (76)	n/a	10	n/a	-7
Crofton Downs	n/a	15	n/a	17 (111)	n/a	7	n/a	-10
Wellington	n/a	76	n/a	44 (58)	n/a	35	n/a	-9
Subtotal	n/a	300	n/a	209 (70)	n/a	138	n/a	-71
Wairarapa								
Note: Stations north of Upper Hutt not surveyed								
Already on train	835	-	86 (10)	-	461	-	375	-
Upper Hutt	63	-	37 (59)	-	35	-	-2	-
Waterloo	237	-	69 (29)	-	131	-	62	-
Petone	23	-	17 (73)	-	13	-	-4	-
Wellington	11	-	9 (83)	-	6	-	-3	-
Subtotal	1,169	-	218 (19)	-	645	-	427	-
Hutt Valley								
Upper Hutt	175	39	63 (36)	28 (73)	97	22	34	-7
Wallaceville	154	13	59 (38)	9 (70)	85	7	26	-2
Trentham	181	28	63 (35)	23 (82)	100	15	37	-8
Heretaunga	78	4	44 (56)	0 (0)	43	2	-1	2
Silverstream	305	17	74 (24)	17 (98)	168	9	95	-7
Manor Park	41	6	28 (69)	9 (152)	23	3	-6	-6
Pomare	73	10	41 (56)	9 (91)	40	6	0	-4
Taita	271	16	71 (26)	17 (104)	150	9	79	-8
Wingate	51	6	33 (65)	9 (152)	28	3	-5	-6
Naenae	196	36	65 (33)	28 (79)	108	20	43	-9
Epuni	94	12	47 (50)	9 (76)	52	7	5	-3
Waterloo	1,245	103	89 (7)	49 (48)	687	57	598	8
Woburn	277	36	72 (26)	28 (79)	153	20	81	-9
Ava	249	12	70 (28)	9 (76)	137	7	68	-3
Petone	241	46	69 (29)	33 (72)	133	25	64	-8
Ngauranga	2	1	0 (0)	0 (0)	1	1	1	1
Kaiwharawhara	1	2	0 (0)	0 (0)	1	1	1	1
Wellington	122	117	54 (44)	54 (46)	67	65	14	11
Subtotal	3,756	504	940 (25)	333 (66)	2,073	278	1,133	-54
Kapiti								
Waikanae	243	62	69 (28)	37 (60)	134	34	65	-3
Paraparaumu	473	95	80 (17)	49 (52)	261	52	181	3
Paekakariki	147	15	59 (40)	17 (111)	81	8	22	-8

Station/Line	Total Boardings* AM	Total Boardings* IP	Required Sample Size[1] (Percentage of Total Boardings) AM	Required Sample Size[1] (Percentage of Total Boardings) IP	Anticipated Completed Returns[2] AM	Anticipated Completed Returns[2] IP	Anticipated less Required AM	Anticipated less Required IP
Muri	-	-	-	-	-	-	-	-
Pukerua Bay	126	16	55 (44)	17 (104)	70	9	14	-8
Plimmerton	235	19	69 (29)	17 (88)	130	10	61	-6
Mana	151	12	59 (39)	9 (76)	83	7	25	-3
Paremata	345	12	76 (22)	9 (76)	190	7	115	-3
Porirua	910	124	87 (10)	54 (43)	502	68	415	15
Kenepuru	31	10	23 (74)	9 (91)	17	6	-6	-4
Linden	221	28	67 (30)	23 (82)	122	15	55	-8
Tawa	207	35	66 (32)	28 (81)	114	19	48	-9
Redwood	254	20	70 (27)	17 (83)	140	11	71	-6
Takapu Road	167	10	62 (37)	9 (91)	92	6	31	-4
Kaiwharawhara	1	4	0 (0)	0 (0)	1	2	1	2
Wellington	70	139	41 (58)	57 (41)	39	77	-2	20
Subtotal	3,581	601	881 (25)	352 (59)	1,977	332	1,096	-20
Melling								
Melling	245	17	70 (28)	17 (98)	135	9	66	-7
Western Hutt	72	2	41 (57)	0 (0)	40	1	-1	1
Petone	160	3	60 (38)	0 (0)	88	2	28	2
Ngauranga	0	1	0 (0)	0 (0)	0	1	0	1
Kaiwharawhara	1	1	0 (0)	0 (0)	1	1	1	1
Wellington	39	12	28 (73)	9 (76)	22	7	-7	-3
Subtotal	517	36	199 (38)	26 (72)	285	20	86	-6
Total	**9,023**	**1,441**	**2238 (25)**	**920 (64)**	**4,981**	**768**	**2,743**	**-152**

* Based on data from Rail Passenger Boarding and Alighting Surveys, June 2011

\- No data available for these routes

[1] Using 95% confidence and 10% margin of error

[2] Based on information from pilot survey. Uses 46% completed returns for stations on the Johnsonville line and assumes a 55% completed returns rate for stations along all other lines. This 20% increase is based on the assumption that issues identified during the Johnsonville pilot survey will be reduced, if not eliminated altogether, improving the response rate

The information in Table 3-2 shows that sampling the rail stations over a 2 hour period is not anticipated to yield the required sample sizes for the given confidence and margin of error specified for all stations. The majority of stations in the AM peak will be satisfied by a 2 hour sample period (due to the larger population size during this period), however a 2 hour survey period during the IP period would appear to be insufficient at most locations. The information also highlights that, from a statistical perspective, attention needs to be paid to maximising response rates at low demand stations in order to achieve confidence in estimating particular characteristics of the station users.

Should budget allow, it would be desirable to collect surveys during the IP period over a 3 to 4 hour duration to increase response numbers.

It is noted that in the situation where sample sizes are limited by practical budget considerations then alternative strategies may be employed in the model estimation process to counter this. For example, while losing some level of confidence in model accuracy, it may be reasonable to estimate some elements of IP behaviour (such as distribution of access zones) based on AM behaviour.

3.1.1 Other Issues to Address

It is assumed that administration of the survey will be designed to provide an unbiased sample of users and that all users provide similar levels of response rate. In this respect:

- It will be important to consider whether travellers undertaking short trips have sufficient time to complete and return survey forms on their trip; and
- There were some comments in relation to the pilot surveys that data returned from school children was less reliable in terms of full completion of origin or destination details.

Consideration needs to be given to practical issues of survey collection and resourcing and analysis implications:

- For train station surveys the methodology is based on collection of forms at the end of a trip. Some in scope surveys with boardings within the defined peak period will end outside the time period so staff resources need to be allocated beyond the survey duration; and
- On buses the sampling statistical approach is based on travelling across the whole length of route. It may be possible to consider on board surveys covering just a proportion of the route, however this adds issues of missing collection of distributed forms for uncompleted trips, and complicating assessment of response rates and weighting processes.

3.1.2 Special Services

Wairarapa Trains

Due to travel time issues, the stations north of Upper Hutt could involve significant additional resources to gather data in comparison to other locations, with limited additional value for insight into key travel issues. Rather than deploying staff to each station on the Wairarapa Line it may be advantageous to place drop boxes at the Wairarapa Stations and for surveyors to board Wairarapa services at Upper Hutt to handout surveys. Surveyors could disembark at Waterloo Station before returning to Upper Hutt to repeat the process. Passengers disembarking from Wellington bound trains at Upper Hutt Station would need to be intercepted to complete the survey. Completion rates for this method would be lower but should provide sufficient data to assign passengers to model zones.

Capital Connection

The Capital Connection service running between Wellington City and Palmerston North does not represent a large portion of rail trips but if budget allows it would be advantageous to survey this service. Due to the nature of this service it may be preferable to place surveyors on-board the service at the Waikanae Station and position a surveyor on the Waikanae Platform to intercept any passengers disembarking. No analysis has been undertaken to determine how many surveyors would be appropriate for this service however a minimum of 2 on board and one on platform should be utilised. Currently, this service has not been included in our analysis.

NB: It is noted that Tranz Metro has already been approached regarding conducting surveys on-board trains and further negotiations may be required to allow this to occur for the Wairarapa services. If the movement of surveyors between carriages is the major concern then it may be feasible to use separate staff in each carriage. Tranz Scenic (operator of the Capital Connection) may have similar safety concerns that will need to be addressed.

Cable Car

For completeness the cable car should be surveyed. Given the low capacity of the cable car it is expected that a single surveyor be deployed onboard to undertake these surveys during a 2 hour AM and IP period.

3.2 Bus/Trolley Services to be Surveyed

Table 3-3 below lists the bus / trolley routes that are considered to be adequate in determining the travel patterns and travel purposes of a representative sample of the bus users in Greater Wellington. It is therefore recommended that all of the routes listed below are surveyed.

A summary of the proportion of number of routes and total boardings selected to be surveyed compared to the total number of routes and boardings for the Greater Wellington area is given in Table 3-4. Boarding data for each route in the Greater Wellington area along with route length and travel times is also given in Appendix A.

It is important to note that as Confidentiality Agreements between Arup, Greater Wellington and Mana Buses (operators of Kapiti and Porirua bus services) have not been signed, ETM data for these services is not available.

Paragraph removed for confidentiality reasons.

Table 3-3: Bus Routes to be Surveyed and Required Sample Sizes

Table 3-3 *removed for confidentiality reasons.*

Table 3-4: Proportion of Routes to be Surveyed Compared to Totals for Greater Wellington

Table 3-4 removed for confidentiality reasons.

It is important to note that routes listed in Table 3-3 only include services where the ETM data showed boarding during the peak periods. Regional services with limited operating times or frequencies (e.g. Te Maura – Wellington (92) and Timberlea – Wellington (93)) are consequently not included. ETM data for these services is available outside of the peak periods and will be used to inform the modelling process, however, surveys should not be undertaken on these routes as resources deployed to other services will provide more value.

Using the NZ Bus ETM data, Figure 3-1 and Figure 3-2 below represent the overall boarding profiles through time for all bus services (excluding Kapiti and Porirua) during the AM and IP, respectively. This data may be useful in informing overall strategies for staff deployment across the survey periods. It is noted that on individual routes profiles vary and in the AM, routes starting remote from the CBD have a higher proportion of passengers boarding earlier in the period than shown by the overall profile.

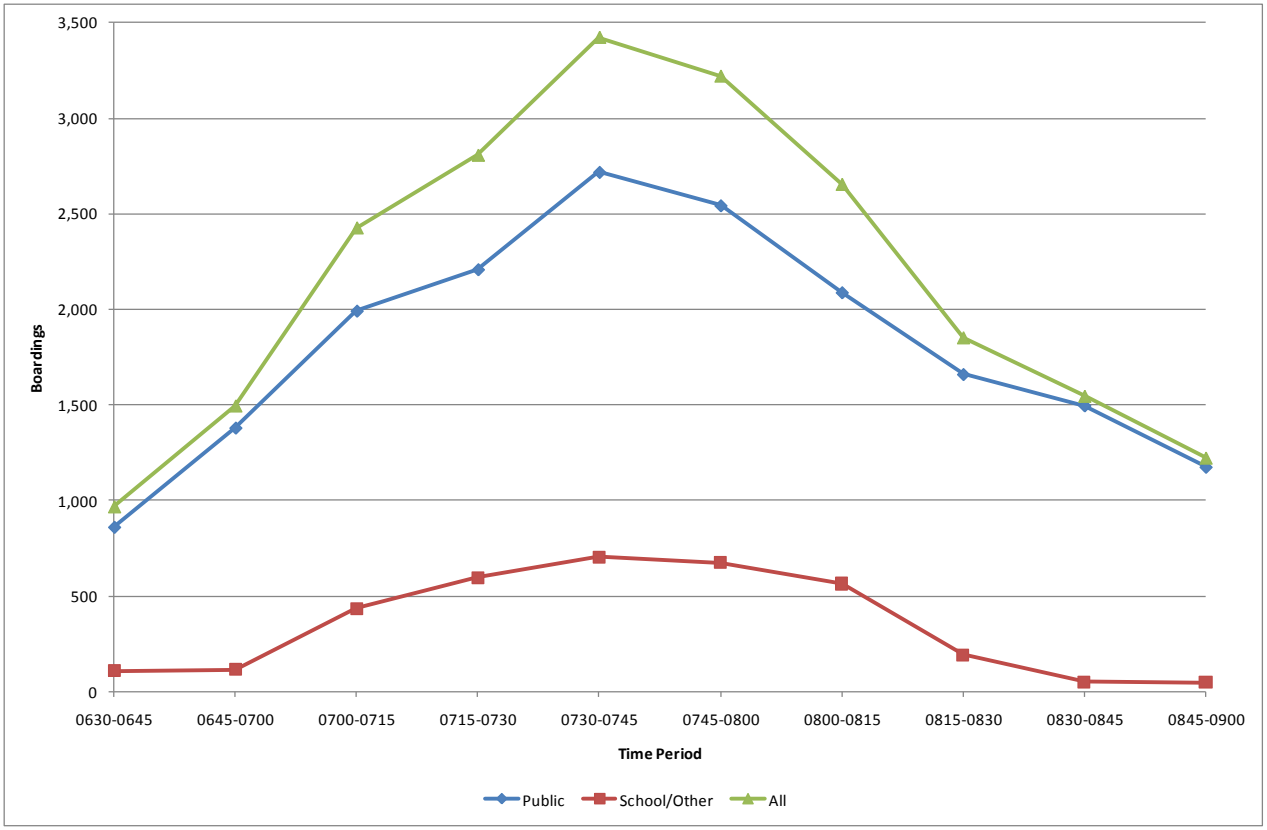


Figure 3-1: AM Peak Period Boarding Profile (Extended to include 6.30am – 7.00am)

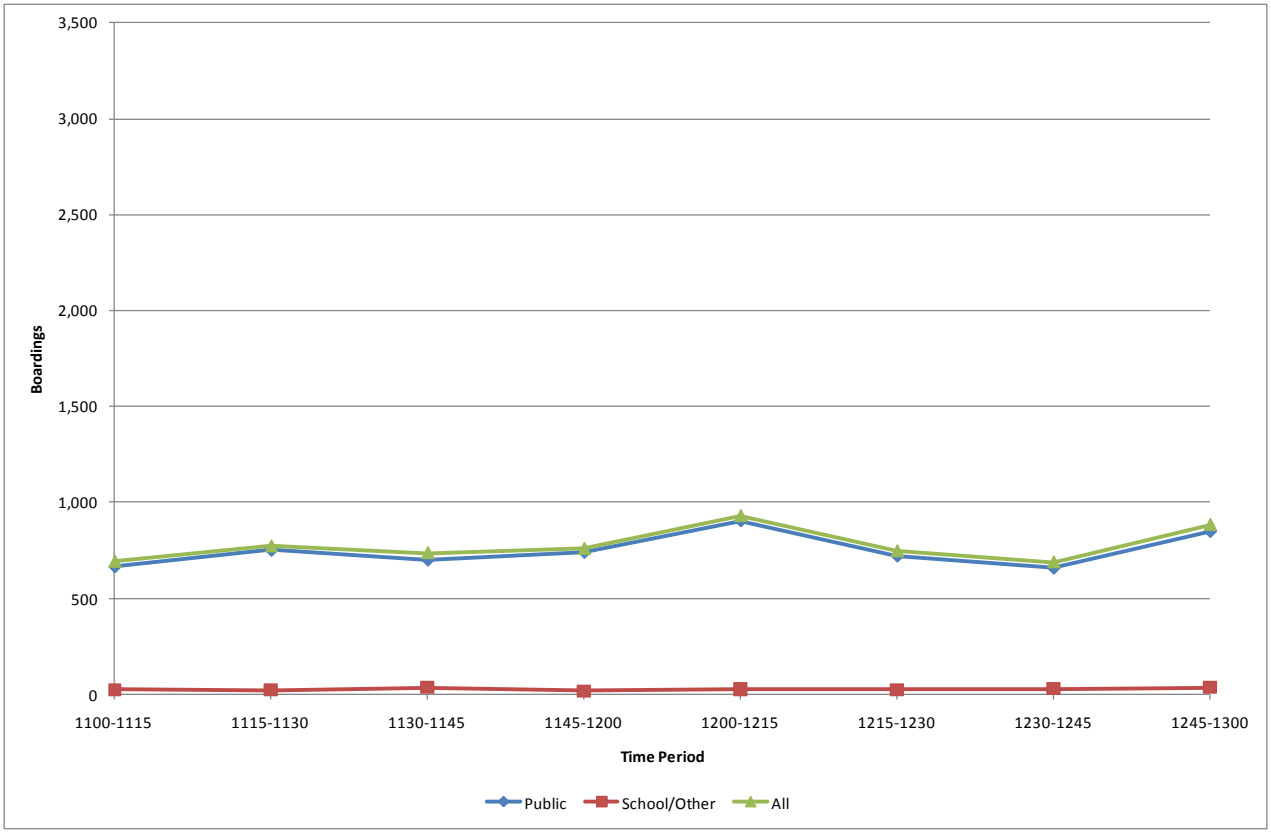


Figure 3-2: IP Period Boarding Profile

3.2.1 Special Services

School Buses

School buses have not previously been coded into WTSM; however, education trips that are made by school bus are included in the trip generation process. Modelling school buses specifically within WTSM would require a structural change to the model which would include separating primary / secondary education trips from tertiary trips and assigning the primary / secondary education trips as an additional class.

The changes required to include school bus routes within WTSM fall outside the scope of this study and the value in doing so would be limited. Maintaining consistency between WTSM and WPTM is integral to the modelling process which we have outlined and consequently prohibits the inclusion of school bus service within WPTM at this time.

Surveying of school buses raises privacy issues that would require addressing through communications with schools and parents. Given the proportion (17% of total boardings during the AM peak) of trips made on school buses it is felt that limited value would be added through this process and resources would be better deployed elsewhere. ETM data and the annual school travel surveys conducted by GWRC will provide information to help inform the update of WTSM and creation of WPTM in the absence of origin-destination survey data. Information is also being sought from the Ministry of Education which funds these services.

4 Estimated Costs

As an initial step in the survey planning process estimated costs to undertake the surveys have been calculated to gauge whether or not these costs can be accommodated by the budget allocated for the data collection task. The costs have been based on converting anticipated person-hours required to undertake the surveys using an assumed working period for each surveyor and a per hour dollar rate for labour. The travel survey consultant will need to review this memo and advise of actual costs. Some iteration in the planning process to refine the sample to match available budget is likely to be required.

The estimated cost for the data collection task is a ball-park figure only and assumes the following:

- Bus surveys require a team of 3 surveyors per bus (likely to be lower on low volume services, particularly during the IP period);
- Rail surveys require an average team of 3 surveyors per station (likely to be lower during the IP);
- Each surveyor to work for 3 hours per peak period (allows for half an hour travel each way), with the exception of IP surveys on rail platforms where surveyors are assumed to work for 4 hours; and
- Each surveyor is paid NZ$40.00 per hour.

Based on the above assumptions and the resource allocation defined in Table 3-2 and Table 3-3, the overall costs identified in Table 4-1 have been identified. It is important to note that the costs identified do not include consideration of additional costs associated with logistical constraints, data processing, analysis costs or survey preparation and supervision costs.

Table 4-1: Estimated Costs of the Recommended Rail and Bus Routes to be Surveyed

	Estimated Cost (NZ$)		
Mode	**AM Peak**	**Inter Peak**	**AM+IP**
Bus	22,680	15,840	38,520
Rail	16,320	25,600	41,920
Bus & Rail	39,000	41,440	80,440

Appendix A - Data for Greater Wellington Bus Services*

Appendix A *removed for confidentiality reasons*