1. PURPOSE

The purpose of this note is to document the procedure followed to calculate the model calibration regression equations developed for generating the attraction trip ends. This has been completed for both the vehicle driver (three-step) trip ends and the person (four-step) models.

2. INTRODUCTION

In this technical note the three-step and four-step model regression equations developed for generating the attraction trip ends are presented. These have been calibrated using trip end data across the 198 Waikato Regional Transportation Model (WRTM) zones inside the Household Interview Survey (HIS) Study Area from the 2008 Waikato Household Interview Survey. The location and extent of this Study Area is presented in Figure One.

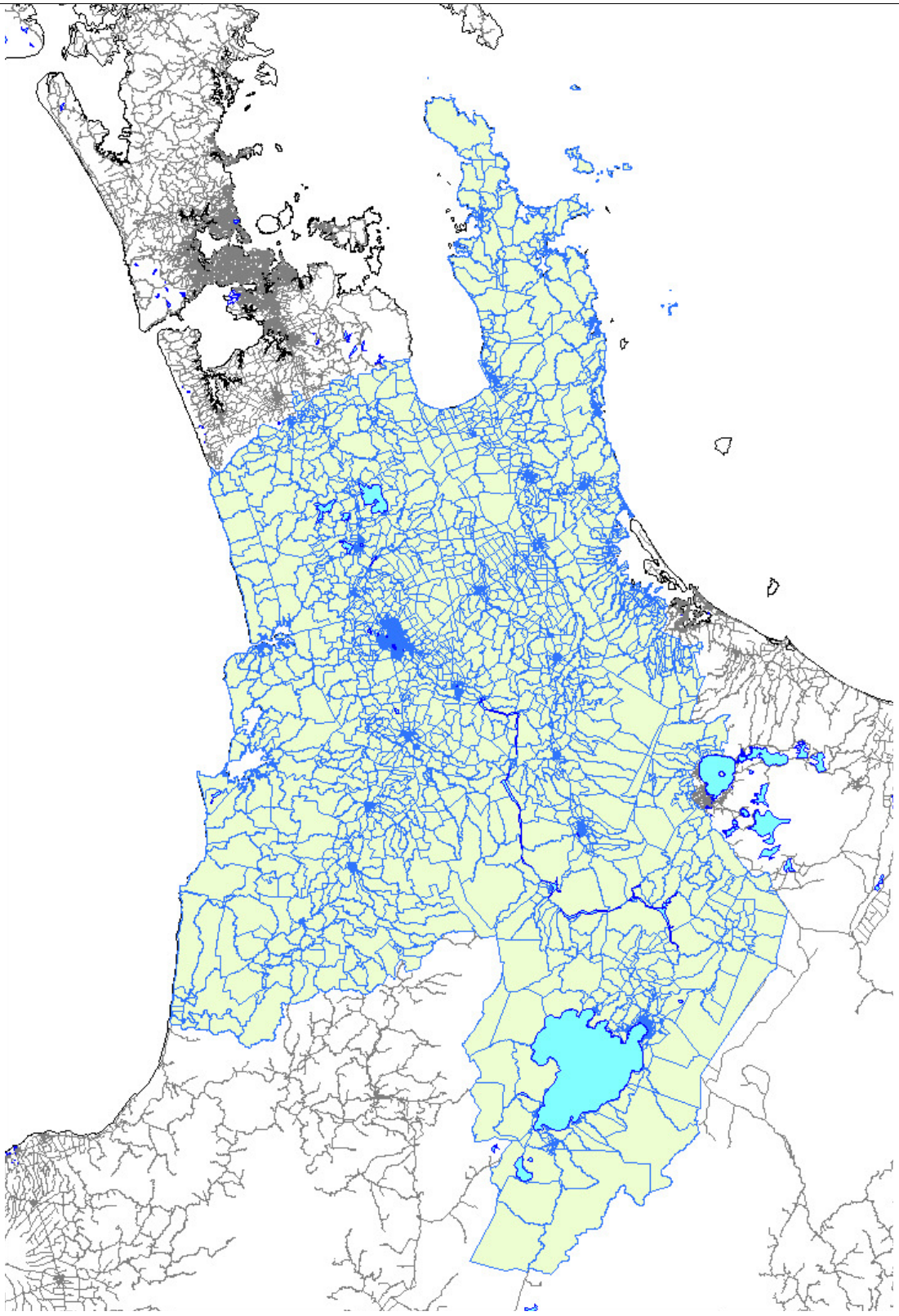
Landuse data has been taken from two sources. Firstly, landuse data relating to both residential and commercial activity has been sourced from Statistics New Zealand's 2006 Census data and education sector data has been sourced from the Ministry of Education, namely School Roll and Tertiary Equivalent Full-Time Students (EFTS) data.

3. DATA USED TO CALIBRATE ATTRACTION EQUATIONS

The trip end data for which regression equations were fitted against was sourced from the Waikato Region HIS (2008) and is defined as follows:

For trips whose origin and destination both lie within the Waikato Regional Transportation Model (WRTM) zones the total trip ends (not including the household trip end for home based trips) for the following trip purposes:

- HBW = Home based Work
- HBB = Home based Business
- HBE = Home based Education
- HBSh = Home based Shopping
- HBSR = Home based Social/recreation
- HBO = Home based Other
- NHB = Non Home Based



Waikato Regional Transportation Model
Survey Area

Figure 1

Some consideration was given as to disaggregating Non Home Based trips into trips that had a workplace as one trip end (i.e. Non Home Based – Work), and trips that did not (i.e. Non Home Based – Non Work). It was evident in the validation process that there was an excellent fit for Non Home Based trips (see Figure 7), and therefore it was not necessary to further split this purpose. Rather, the disaggregation of this variable was considered as a 'fall-back' position should the results have not been as encouraging.

The total number of trip ends by WRTM zone were isolated for each purpose to cover the three periods of the day being morning peak (7-9am), interpeak (9am-4pm) and evening peak (4-6pm) for car driver trip mode only, which resulted in a total of 21 origin-destination pairs of trip ends (i.e. 7 purposes by 3 time periods) by WRTM zone. Home to and to home trips have been aggregated to form Home based trips.

This analysis has then been repeated and included in this report for all modes as 'person trips', which includes:

- Car driver
- Car passenger
- Public transport
- Walking
- Cycling

The trip end data was regressed against the a total of 23 landuse variables as follows:

- Number of households in WRTM zone (total in HIS study area is 147243),
- School Roll – primary and secondary combined, totalling 76459 in HIS Study Area (source 2008 Ministry of Education July roll data)
- Tertiary Equivalent Full Time Students (EFTS), totalling 19998 in HIS Study Area (source 2008 Ministry of Education EFTS by campus by provider)
- Number of jobs by workplace zone for all 19 ANZSIC06 job industry classifications plus total jobs (i.e. this is 20 variables),
    - Job industry classifications are defined as (with HIS study area totals):

Classification	Number of Jobs
Agriculture, Forestry and Fishing	25960
Mining	842
Manufacturing	19879
Electricity, Gas, Water and Waste Services	1090
Construction	15790
Wholesale Trade	6372
Retail Trade	18072
Accommodation and Food Services	10243
Transport, Postal and Warehousing	5678
Information Media and Telecommunications	1477
Financial and Insurance Services	3198
Rental, Hiring and Real Estate Services	4246
Professional, Scientific and Technical Services	10522
Administrative and Support Services	4944
Public Administration and Safety	5685
Education and Training	13200
Health Care and Social Assistance	14959
Arts and Recreation Services	2901
Other Services	6532
Total Jobs	**186321**

4. CORRELATION ANALYSIS OF VARIABLES

The landuse variables specified in section 3 above were checked for cross-correlations. This was necessary to gain an understanding of which variables were highly correlated and therefore could either be aggregated or used as a proxy for other variables. The full correlation results are presented in Appendix One to this technical note.

Retail jobs have a high correlation with office jobs and other jobs of an administrative nature. This occurs as a result of theses type of jobs being centred in the commercial urban centres in the study area. There is also a high correlation between retail jobs, office jobs and total jobs. These may prove to be interchangeable to some extent and may result in total jobs not occurring as a variable in all administrative, office and retail roles.

Whilst not included in the table above, education has a relatively high correlation with Regional schools (0.688) and Regional Tertiary (0.767). The significance of these two role based education variables leads to reduction in the correlation of the Education variable, and hence being dropped from the analysis.

A high correlation (i.e. of over 0.8) is exhibited between the following variables:

Variable	High Correlation with	R^2		Variable	High Correlation with	R^2
Manufacturing	Construction	0.880		Information	Finance	0.969
	Wholesale	0.897			Rental	0.887
Construction	Wholesale	0.917			Administration	0.932
Wholesale	Other services	0.818			Public Admin	0.915
Retail	Accommodation	0.910			Arts	0.887
	Information	0.830			Total Jobs	0.810
	Finance	0.896		Finance	Rental	0.926
	Rental	0.919			Professional	0.963
	Professional	0.880			Administration	0.970
	Administration	0.873			Public Admin	0.950
	Public Admin	0.895			Arts	0.917
	Arts	0.832			Total Jobs	8.847
	Other services	0.825		Rental	Professional	0.942
	Total Jobs	0.925			Administration	0.919
Accommodation	Information	0.866			Public Admin	0.906
	Finance	0.892			Arts	0.894
	Rental	0.947			Total Jobs	0.894
	Professional	0.897		Professional	Administration	0.955
	Administration	0.877			Public Admin	0.939
	Public Admin	0.873			Arts	0.903
	Arts	0.861			Total Jobs	0.887
	Total Jobs	0.847				
Administration	Public Admin	0.945				
	Arts	0.923				
	Total Jobs	0.887				
Public Admin	Arts	0.914				
	Total Jobs	0.878				
Arts	Total Jobs	0.838				
Other services	Total Jobs	0.867				

5. METHODOLOGY

The regression analysis was completed using StatistiXL version 1.8, a specialist software add-on to Microsoft Excel that includes a wide-range of statistical tools.

Each of the sets of trip ends (i.e. by purpose by period for both vehicles driver and person models) was regressed against the nominated landuse variables in section 3 of this technical note.

The regression was a forward stepwise regression with a constant of zero forced in each instance. The criterion for the inclusion of variables is a probability of a type one error being 0.1 (i.e. there is a 10% probability that a variable is selected within the analysis but it is not a significantly significant variable). Variables were only permitted to have non-negative coefficients in the analysis.

The first pass of the regression model included all 23 variables to see which were significant. Some variables were then discarded if they should not logically be included (e.g. Mining included in Home Based Education) and the analysis was revisited without such anomalies.

The resultant R-Squared values were then considered and where an R-Squared of less than 0.5 was noted the data was checked to see if there was sufficient sample data to establish a regression equation. The measure for this assessment was the mean cell size, which equates to the average number of expanded trip ends per Regional Model zone from the HIS data.

In order to improve the quality of fit of the attractions equations, it was necessary to consolidate trip end data across all three periods of the day for some trip purposes. This was done for HBE trips, and then again for HBSR trips in the morning and interpeak periods. As such, the attraction equation developed for all three periods is identical for the HBE trip purpose and the attraction equation for HBSR is identical in the Morning Peak and Interpeak periods.

Following this aggregation of periods, the only outstanding R-squared value of over 0.5 was in the morning peak period for the Home Based Shopping purpose where there was also a very low mean cell size. The variable that entered into the regression model was Retail Trade jobs, which was consistent with this purpose for the other periods of the day. For this reason this result was accepted.

A number of other hypotheses were tested in the regression analysis. The jobs from the Health Care and Social Assistance; Arts and Recreation Services; and Other Services ANZSIC06 industry categories were aggregated from Statistics New Zealand 2006 census data to form a “Community Jobs” variable. Subsequent regression analysis found a better fit of data using the Community Jobs variable rather than applying the three individual ANZSIC06 variables.

Similarly the Information Media and Telecommunications; Financial and Insurance Services; Rental, Hiring and Real Estate Services; Professional, Scientific and Technical Services; Administrative and Support Services; and Public Administration and Safety ANZSIC06 industry categories were aggregated from Statistics New Zealand 2006 census data to form an “Office Jobs” variable. As is consistent with the correlation analysis discussed in section 4 of this technical note this was highly correlated and

therefore largely interchangeable with the Retail Trade Jobs variable. Retail Trade Jobs provided a slightly better fit for most trip purposes, however by forcing the aggregated "Office Jobs" variable into the Home Based Business purpose there was negligible difference in the fit of the regression model so this was accepted.

Further tests were also undertaken to confirm that the Education Jobs ANZSIC06 category was redundant. In all cases no deterioration in the fit of the regression model was introduced by regressing the trip ends against the two Ministry of Education variables (school rolls and tertiary EFTS) as opposed to a combination of these and Education Jobs.

6. REGRESSION ANALYSIS RESULTS

The results of the regression analysis are summarised below. R-Squared values to measure the goodness of fit between the HIS data trip ends and the regression equation trip ends are included for each equation. T values for each accepted independent variable are included to indicate the extent to which each variable fits into the model. The values follow the Student's t-distribution curve and are indicative of the probability that the variable is significant in the equation. A value of 1.64 corresponds to a probability of 90% that the variable is significant and a value of 1.96 indicates a 95% probability. As these values become greater the probability that it is a significant variable increases.

The mean cell sizes are also reported which equate to the average number of expanded trip ends per Regional Model zone from the HIS data.

The trip purpose codes below are as follows:

- HBW – Home Based Work
- HBE – Home Based Education
- HBB – Home Based Business
- HBSh – Home Based Shopping
- HBSR – Home Based Social Recreation
- HBO – Home Based Other
- NHB – Non Home Based

7. 3 STEP MODEL REGRESSION ATTRACTIONS

Morning Peak Vehicle Trip Attractions			R^2	Mean Cell
HBW	=	0.302TOT I = 36.668	0.872	271
HBE	=	0.090SCH + 0.271TER I=6.178 I=18.317	0.699	65
HBB	=	0.104OFF I=17.019	0.595	25
HBSh	=	0.211RET I=17.922	0.535	27
HBSR	=	0.650RET + 0.215HH T=10.509 T=11.286	0.665	219
HBO	=	0.141COM + 0.071HH + 0.285SCH I=3.941 I=3.315 I=9.507	0.688	186
NHB	=	0.225WHOLE + 0.411 RET + 0.032HH + 0.077 SCH T=3.697 T =13.356 T=3.218 T=5.246	0.821	187
Interpeak Vehicle Trip Attractions			R^2	Mean Cell
HBW	=	0.241TOT + 0.088HH I=19.351 I=3.105	0.752	332
HBE	=	0.090SCH + 0.271TER I=6.178 I=18.317	0.699	65
HBB	=	0.276OFF + 0.404COM I=6.701 I=5.823	0.572	167
HBSh	=	3.199 RET I=29.022	0.810	356
HBSR	=	0.650RET + 0.215HH I=10.509 I=11.286	0.665	219
HBO	=	0.114COM + 0.143HH + 0.311SCH + 0.089TER I=1.964 I=4.569 I=7.048 I=2.683	0.649	246
NHB	=	3.806RET + 0.329HH T =29.039 T=4.899	0.854	616
Evening Peak Vehicle Trip Attractions			R^2	Mean Cell
HBW	=	0.294TOT I = 34.418	0.857	292
HBE	=	0.090SCH + 0.271TER I=6.178 I=18.317	0.699	65
HBB	=	0.054OFF + 0.120COM I=6.525 I=8.543	0.660	30
HBSh	=	1.018 RET T=23.057	0.730	116
HBSR	=	0.210COM+ 0.100HH I=9.231 I=10.226	0.620	97
HBO	=	0.171COM+ 0.193RET + 0.074SCH I=5.401 I=4.294 I=4.837	0.562	84
NHB	=	1.518RET + 0.207COM + 0.187HH T=21.241 T =3.984 T=4.273	0.871	241

8. 4 STEP MODEL REGRESSION ATTRACTIONS

Morning Peak Person Trip Attractions			R^2	Mean Cell
HBW	=	0.369TOT T = 44.640	0.910	323
HBE	=	0.609SCH + 0.117TER T=19.194 T=3.665	0.692	283
HBB	=	0.111OFF T=16.437	0.578	30
HBSh	=	0.265 RET T=14.633	0.521	37
HBSR	=	0.936RET + 0.436HH T=8.853 T=13.392	0.678	404
HBO	=	0.272COM+ 0.115HH + 0.481SCH T=4.233 T=3.086 T=8.924	0.665	313
NHB	=	0.335WHO + 1.216RET + 0.091HH + 0.230SCH T=1.806 T =13.013 T=2.962 T=5.133	0.795	281
Interpeak Person Trip Attractions			**R^2**	**Mean Cell**
HBW	=	0.135HH + 0.265TOT T = 4.119 T = 18.294	0.745	384
HBE	=	0.560SCH + 0.378TER T=13.106 T=8.777	0.622	312
HBB	=	0.775COM + 0.240OFF T=8.026 T=4.181	0.564	226
HBSh	=	4.182 RET T=30.765	0.828	463
HBSR	=	0.936RET+ 0.436HH T=8.853 T=13.392	0.678	404
HBO	=	0.563COM+ 0.364HH + 0.644SCH + 0.169TER T=4.420 T=3.844 T=6.637 T=2.305	0.652	536
NHB	=	11.066RET + 0.934HH T=24.984 T =6.843	0.829	1679

There are a few instances where landuse variables are included in some periods but not others. This often occurs because the nature of trip-making across the course of the day changes, depending on the hours of operation of various industry sectors and the business functions they serve at different times of day. Such results have been noted in other similar studies including the North Shore model calibration (from 1991 Auckland HIS) and model calibrations from Sydney's Transport Data Centre.

Where:

HH = Households (source Statistics New Zealand 2006 census)

SCH = School Roll (source 2008 Ministry of Education July roll data)

TER = Tertiary Equivalent Full Time Students (source 2008 Ministry of Education EFTS by campus by provider)

COM = Community Jobs (Health Care and Social Assistance; Arts and Recreation Services; and Other Services ANZSIC06 industry categories aggregated from Statistics New Zealand 2006 census)

OFF = Office Jobs (Information Media and Telecommunications; Financial and Insurance Services; Rental, Hiring and Real Estate Services; Professional, Scientific and Technical Services; Administrative and Support Services; and Public Administration and Safety ANZSIC06 industry categories aggregated from Statistics New Zealand 2006 census)

RET = Retail Trade Jobs (ANZSIC06 category from Statistics New Zealand 2006 census)

WHO = Wholesale Trade Jobs (ANZSIC06 category from Statistics New Zealand 2006 census)

TOT = Total Jobs (all 19 ANZSIC06 categories from (source Statistics New Zealand 2006 census)

9. A COMPARISON OF MODELLED VERSUS OBSERVED TRIP ENDS

The number of trip ends as generated in the model for each purpose and time period has been checked back against the WRTM Household Interview Survey expanded surveyed trip ends. The model specification report recommended that the modelled trip ends by purpose by period should be within 15% of surveyed trip ends in all cases.

A filter has been placed in the model to extract trip ends in the area covered by the HIS (see Technical Note 4 for HIS boundary) which in general terms removes the Tauranga, Rotorua and Western Bay of Plenty trip ends from the analysis. As such the survey and model can be compared on as equivalent data sets.

For each of the five models (AM and Interpeak models for the four-step model, and all three periods for the three-step model) the total HIS trip ends are included in **Table 1** and **Table 2** for HIS and modelled trips respectively. **Table 3** contains the percentage differences between the modelled and surveyed trip ends.

The only trip purpose which lies outside of the +/-15% target, is Business to Home in the interpeak period for the four-step model, whereby the modelled trip ends are 17% high. This is a trip purpose with a relatively small number of observations (6400 trips which is synonymous with 64 sampled trips in the HIS based on an average expansion factor of 100). To put this outlier into perspective, if only 10 more such trips were sampled in the interpeak period the residual would be removed.

The impact of the over-generation of these trips amounts to only 1000 trips over 205000 trips in total being 0.4% of total trip generation for the period. On this basis the outlier being only just outside of the set target is considered to be negligible.

Those trip purposes and periods where there is considered to be a low number of sampled trips (i.e. less then 100) are presented in bold in **Table 1**. These are the trip purposes within the corresponding time periods where it is likely to be difficult to get a good 'fit' against the relatively sparse HIS data.

Total Expanded HIS Trips by Trip Purpose and Time Period — **Table 1**

Trip Purpose	Four Step Model (Person Trips)		Three Step Model (Veh Driver Trips)		
	AM Peak	INT Peak	AM Peak	INT Peak	PM Peak
Home to Work	61770	**8014**	52096	**6430**	**3029**
Home to Education	55836	**2322**	**3720**	**609**	**565**
Home to Business	**4822**	**5453**	**4180**	**4230**	**2242**
Home to Shop	**5639**	10156	**4188**	**8196**	**5799**
Home to Social/Rec	**8932**	**8110**	**3780**	**4726**	**8598**
Home to Other	50302	**9046**	28260	**3181**	**5463**
Non Home Based	56343	97194	37546	63731	41333
Work to Home	**2877**	15301	**2282**	13522	55376
Education to Home	**291**	**3119**	**291**	**721**	**1702**
Business to Home	**1139**	**6403**	**842**	**4807**	**3731**
Shop to Home	**1812**	18394	**1321**	14316	17220
Soc/Rec to Home	**2593**	10420	**1493**	**5692**	10681
Other to Home	11719	10959	**8679**	**3765**	11402
Total All Purposes	264075	204891	148678	133926	167141

Total Modelled Trips by Trip Purpose and Time Period					Table 2
	Four Step Model (Person Trips)		Three Step Model (Veh Driver Trips)		
Trip Purpose	AM Peak	INT Peak	AM Peak	INT Peak	PM Peak
Home to Work	62297	7843	52680	6309	3035
Home to Education	56863	2637	3964	642	602
Home to Business	4786	5348	4133	4162	2177
Home to Shop	5618	10017	4083	8048	5989
Home to Social/Rec	9063	8252	3793	4652	8711
Home to Other	50523	8880	27973	3257	5327
Non Home Based	56617	98856	37682	61963	41615
Work to Home	3028	15494	2459	13697	55670
Education to Home	251	3155	251	771	1679
Business to Home	1224	7505	903	5154	3660
Shop to Home	1758	18322	1207	14049	17756
Soc/Rec to Home	2595	10669	1501	5611	10656
Other to Home	11958	11018	8815	3735	11216
Total All Purposes	266582	207997	149444	132050	168094

Modelled versus Surveyed Trips by Purpose and Time Period					Table 3
	Four Step Model (Person Trips)		Three Step Model (Veh Driver Trips)		
Trip Purpose	AM Peak	INT Peak	AM Peak	INT Peak	PM Peak
Home to Work	0.9%	-2.1%	1.1%	-1.9%	0.2%
Home to Education	1.8%	13.6%	6.5%	5.4%	6.6%
Home to Business	-0.7%	-1.9%	-1.1%	-1.6%	-2.9%
Home to Shop	-0.4%	-1.4%	-2.5%	-1.8%	3.3%
Home to Social/Rec	1.5%	1.8%	0.3%	-1.6%	1.3%
Home to Other	0.4%	-1.8%	-1.0%	2.4%	-2.5%
Non Home Based	0.5%	1.7%	0.4%	-2.8%	0.7%
Work to Home	5.2%	1.3%	7.8%	1.3%	0.5%
Education to Home	-13.7%	1.2%	-13.7%	6.9%	-1.4%
Business to Home	7.5%	17.2%	7.3%	7.2%	-1.9%
Shop to Home	-3.0%	-0.4%	-8.6%	-1.9%	3.1%
Soc/Rec to Home	0.1%	2.4%	0.5%	-1.4%	-0.2%
Other to Home	2.0%	0.5%	1.6%	-0.8%	-1.6%
Total All Purposes	0.9%	1.5%	0.5%	-1.4%	0.6%

10. 3 STEP MODEL VALIDATION OF ATTRACTION TRIP ENDS

The calibrated three step attraction equations have been included in the model and the resultant number of modelled trip ends by Regional model zone from the validated model have been compared against the HIS trips ends by zone. This comparison is intended

Note that given that the comparisons are included at regional zone level, there are a large number of zones with few survey households, meaning there will be a few outliers as a result especially for trip purposes with low levels of activity. The scatterplots for each modelled period are included in Figure 2 through Figure 8 for each trip purpose.

Each of the plots included an R-squared statistic, which measures goodness of fit. In general terms an R-squared of over 0.5 indicates there is a significant level of correlation between the two variables. It is important, however, to reiterate that with results calculated at zone level and with many of these zones having only a few surveyed households, it is difficult to address the issue of outliers.

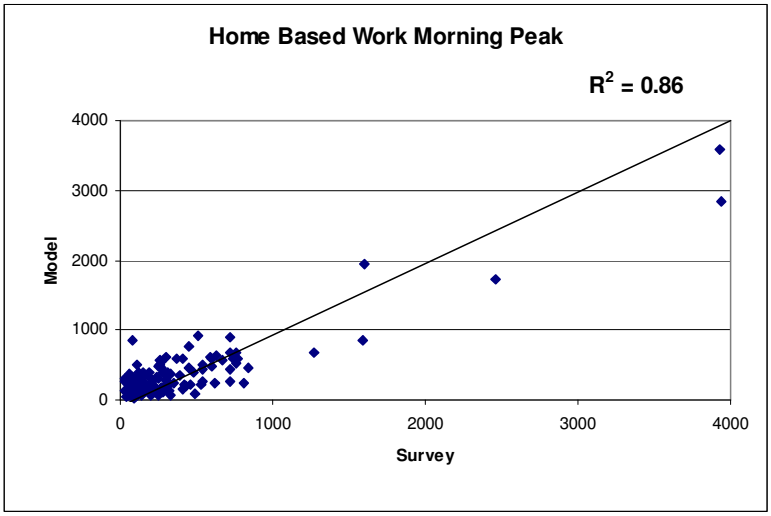
In **Table 1** the total number of expanded HIS trips are reported with those purposes with fewer than 100 sampled trips surveyed (based on an average expansion factor of 100) in the corresponding time period presented in bold. These are generally those with the worst fit in the following tables which is not altogether unexpected given that the scatterplot analysis has been undertaken over nearly 200 zones, meaning the average number of sampled trips per zone is no greater than 0.5 in these instances.

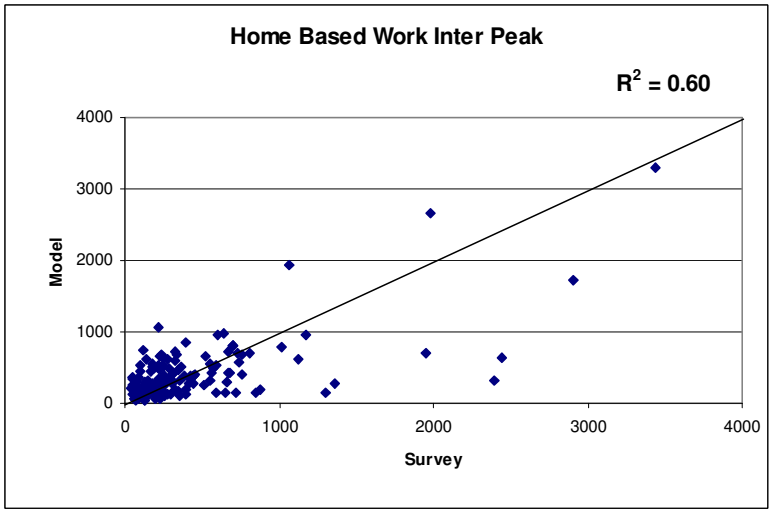
For those trip purposes with very few trips surveyed, the R-squared values are low. This was addressed in the calibration process by aggregating time periods for the Home Based Education and Home Based Social Recreational purposes. Putting these to one side the majority of R-squared values are over 0.5 with the critical Home Based Work peak periods and Non Home Based purposes over having R-Squared values of over 0.8.

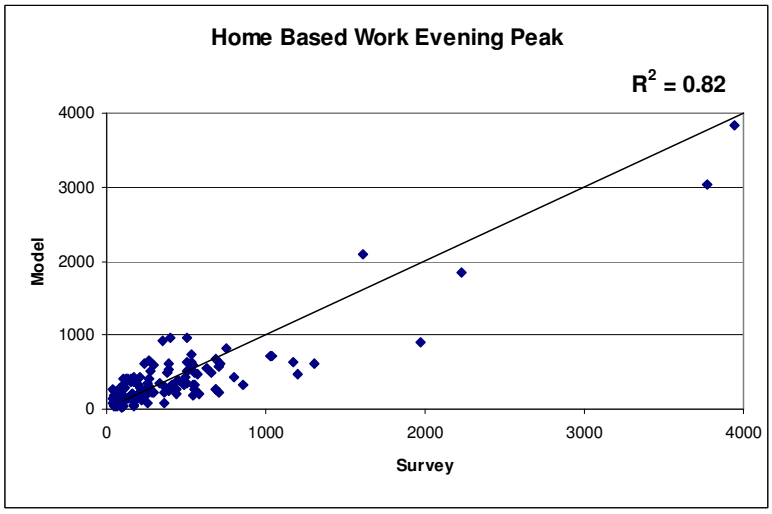
By means of comparison the North Shore calibration and validation of trip ends (from Auckland 1991 Household Interview Survey and reported in "North Shore 3 Stage and 4 Stage Model Build – Trip Generation and Trip Distribution Calibration Report" prepared by Gabites Porter in April 2006), yielded R-squared values for Home Based Work in peak periods in the range of 0.85-0.9 for the same three modelled periods which is similar to the Waikato results. North Shore Non Home Based results were in the order of 0.6-0.8 in the North Shore calibration with Waikato results in the 0.8-0.85 range. The other trip purposes in the North Shore Calibration process differed from those reported herein so a direct comparison is not possible, however R-squared values for other purposes were in the range of 0.2-0.8.

At time of writing the Christchurch Transport Model (CTM) equivalent trip end calibration and validation results have not been made available for comparison.

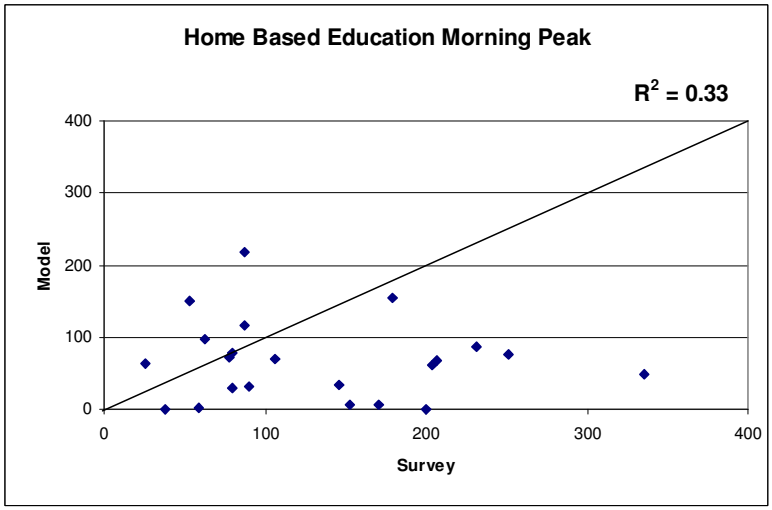
Also included is a scatterplot for all trip purposes aggregated for each of the three-modelled periods in Figure 9 with R-squared values of 0.85 to 0.89. These compare to values in the order of 0.8-0.9 from the North Shore model calibration. Whilst the Model Specification Report initially suggested R-squared values at least 0.9 to be the goal, in hindsight this is not practicable due to the 'lumpy' nature of HIS data. By comparison of R-squared values against the North Shore HIS it is proposed that a good measure of fit has been achieved.

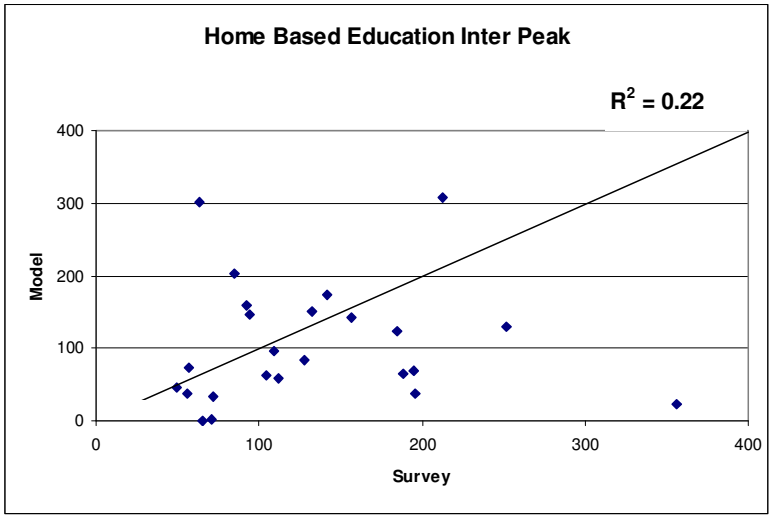


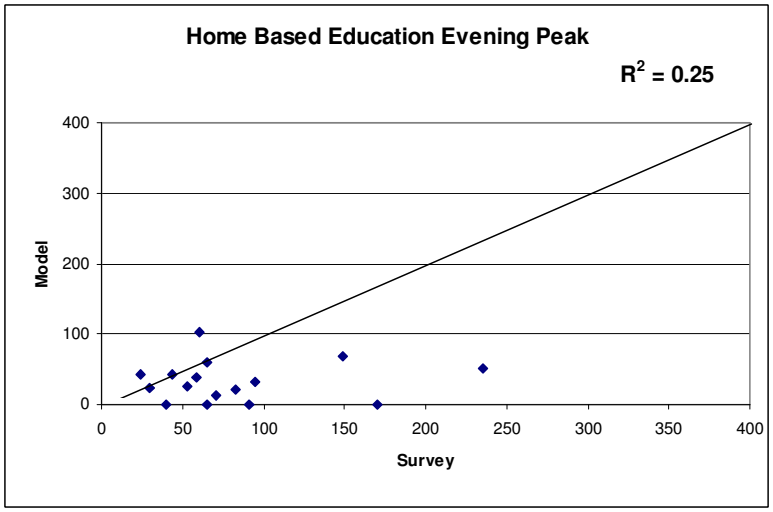




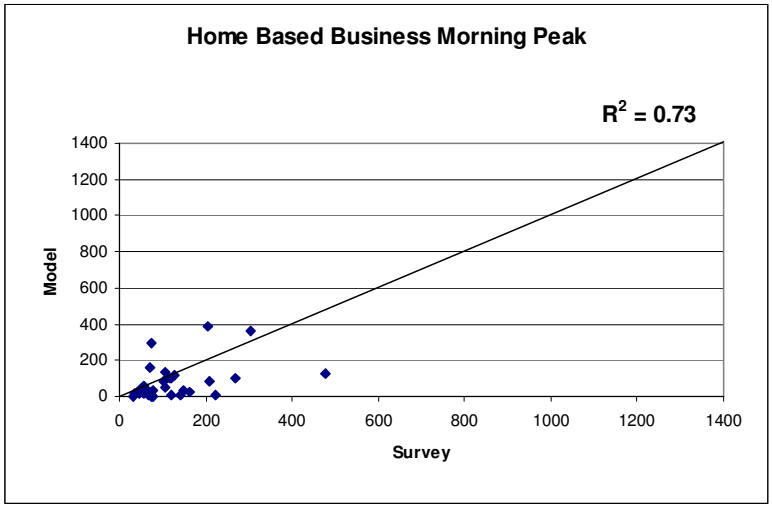
Home Based Work Attraction Trip End Validation | **Figure 2**

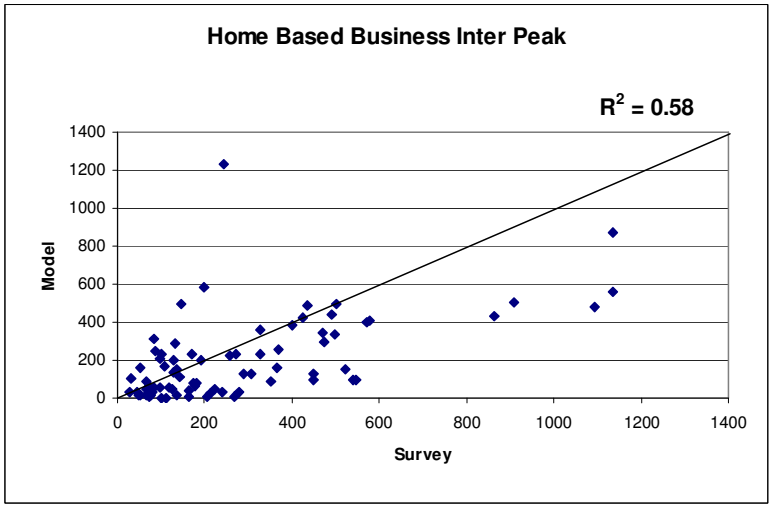


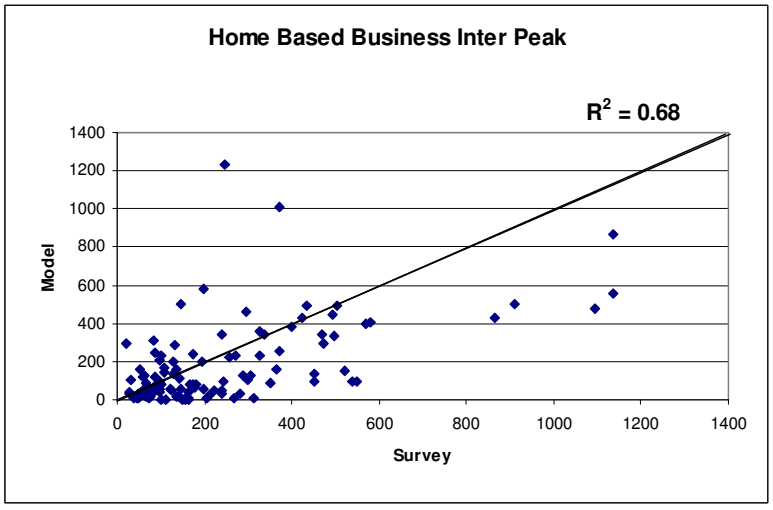




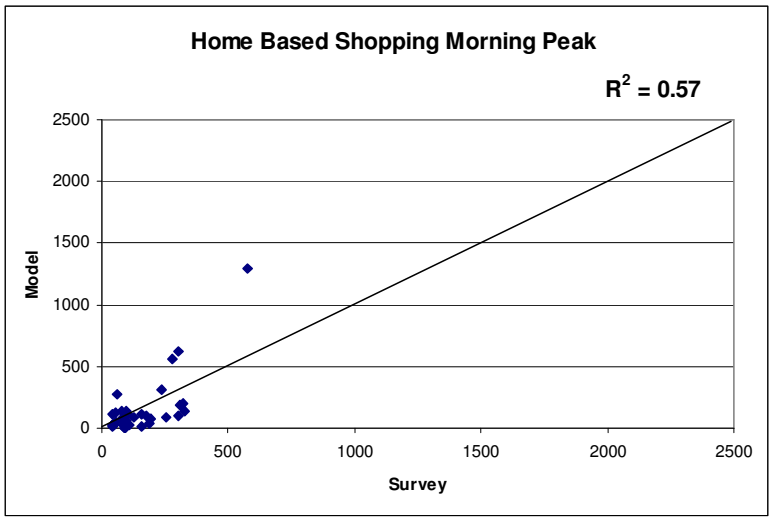
Home Based Education Attraction Trip End Validation	Figure 3

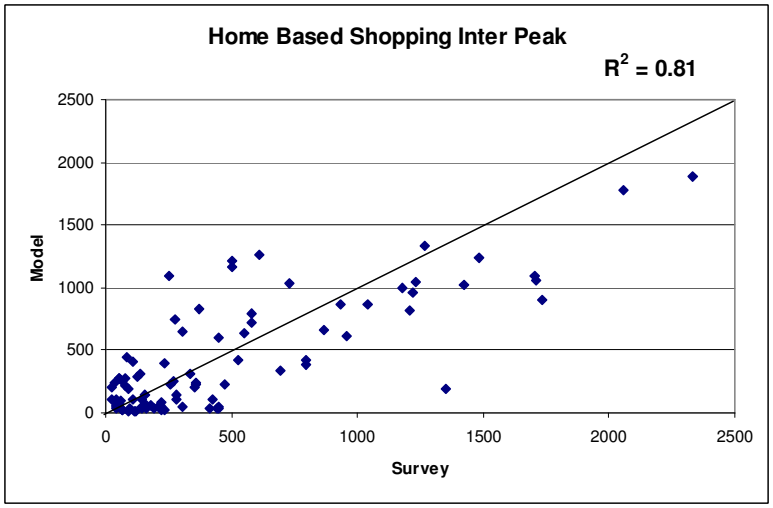


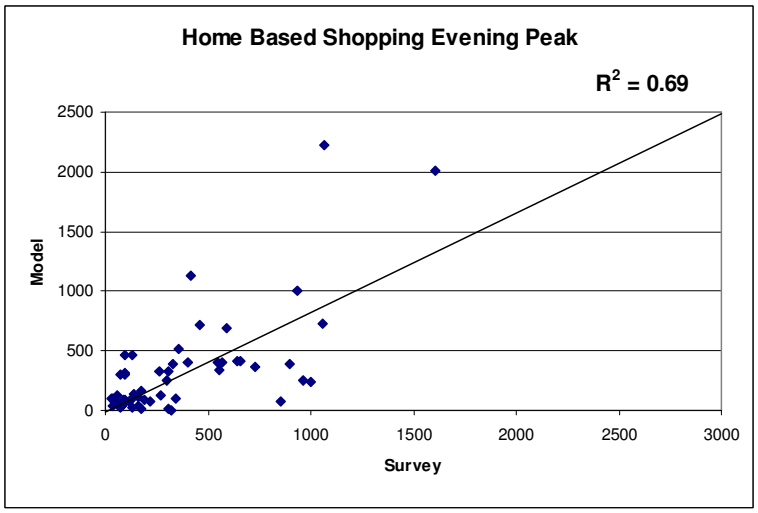




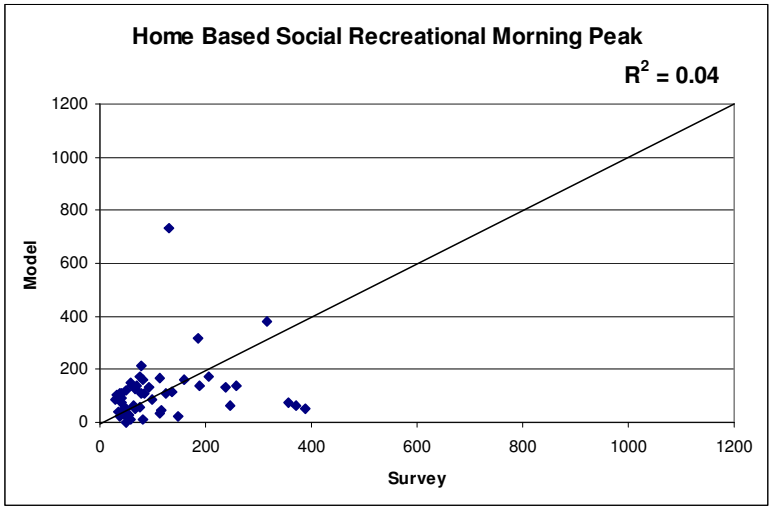
Home Based Business Attraction Trip End Validation | **Figure 4**

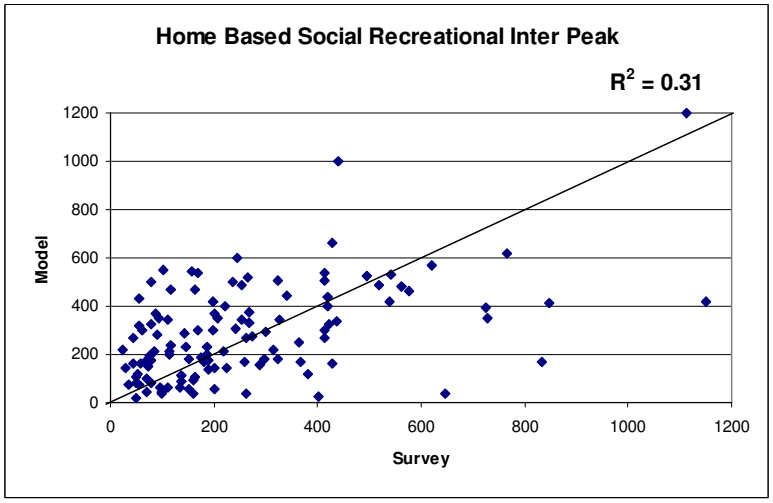


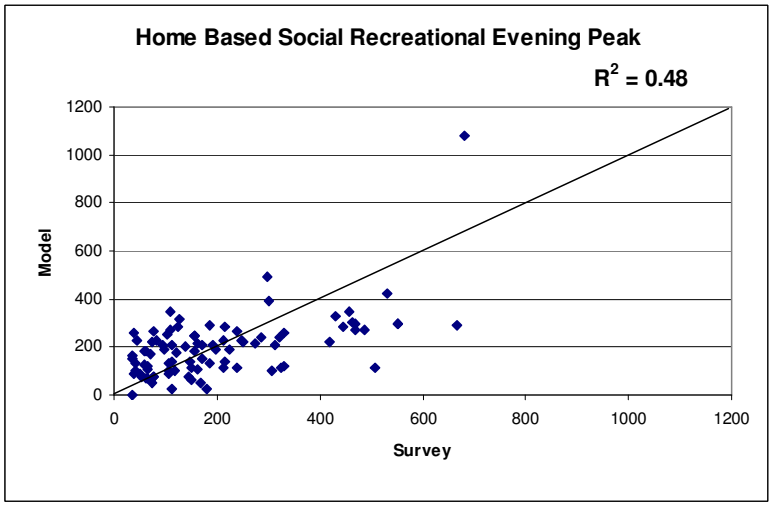




Home Based Shopping Attraction Trip End Validation | **Figure 5**

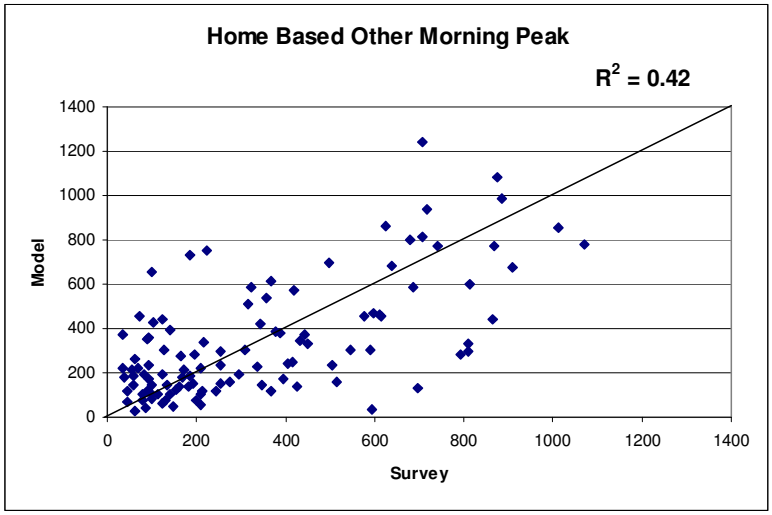


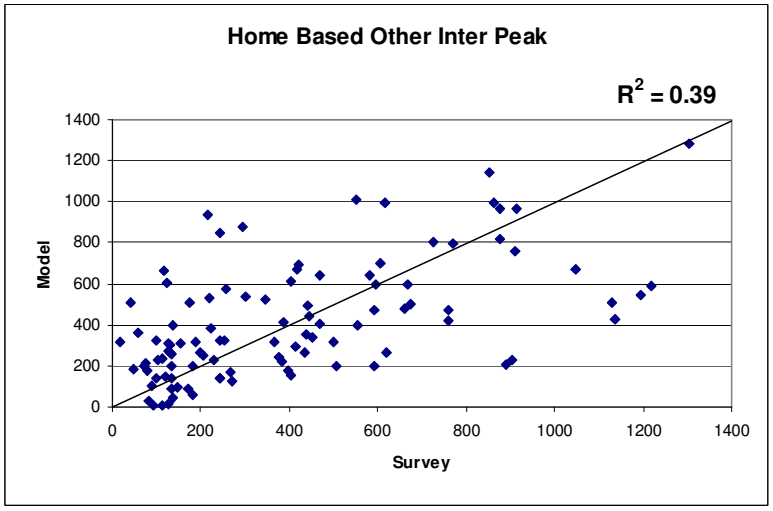


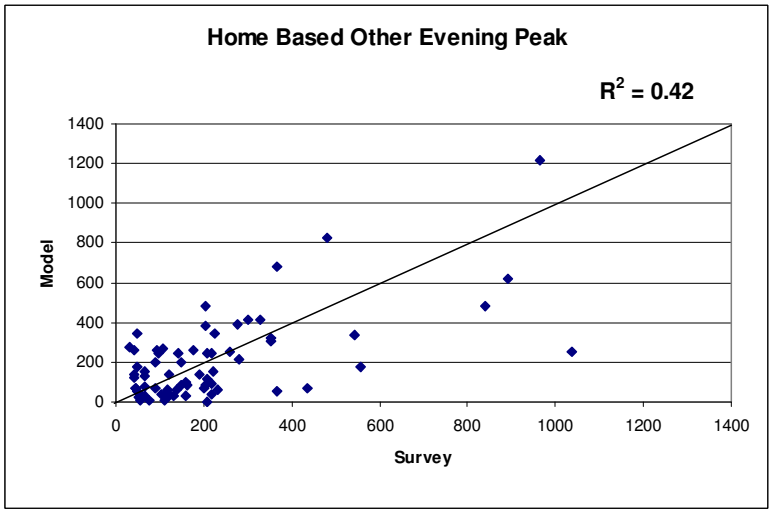


Home Based Social Recreational Attraction Trip End Validation

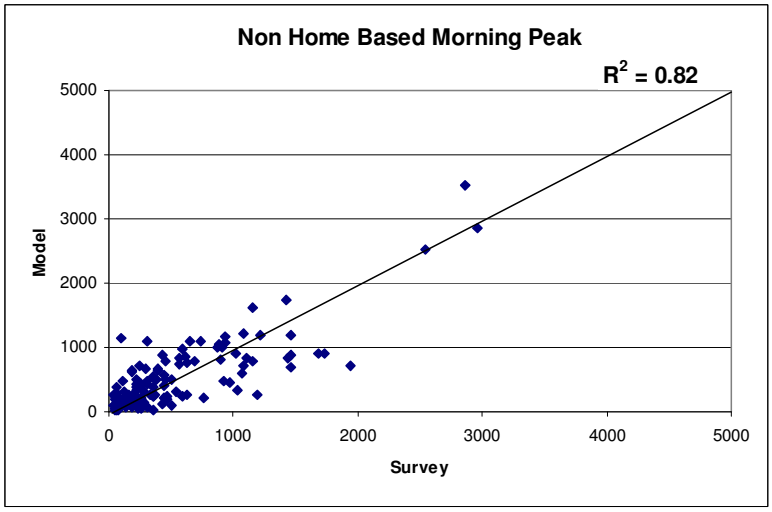
Figure 6

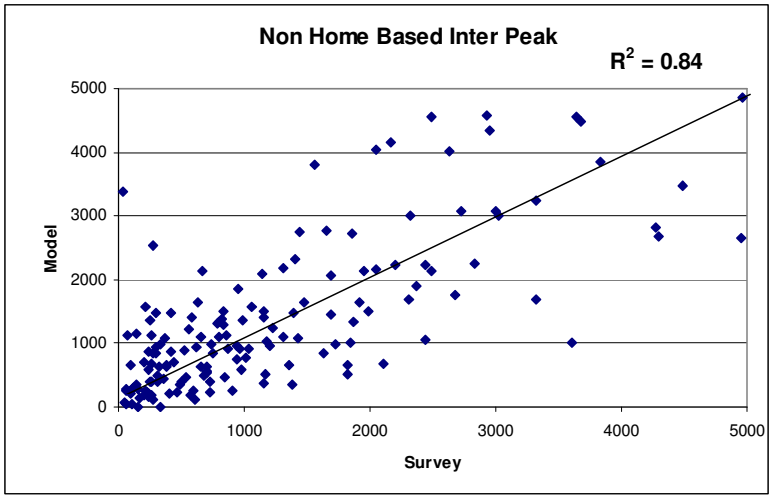


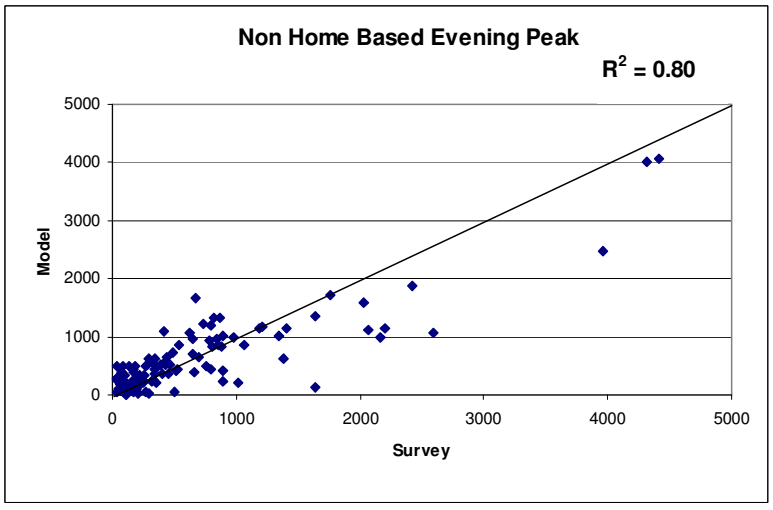




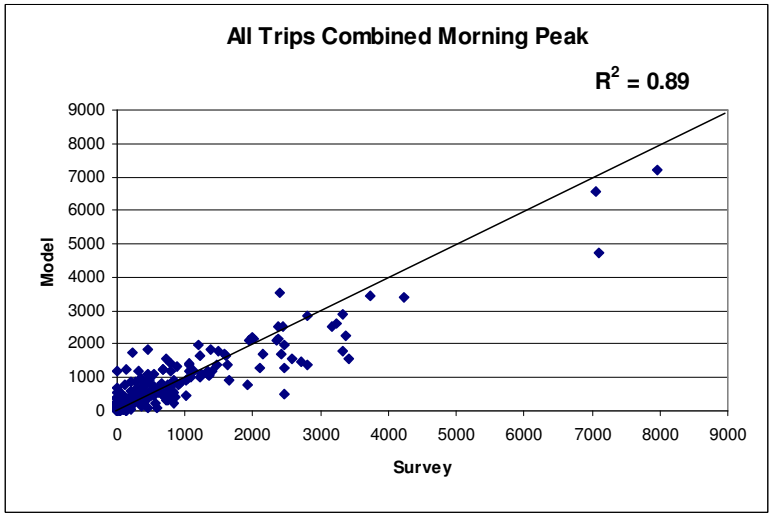
Home Based Other Attraction Trip End Validation | **Figure 7**

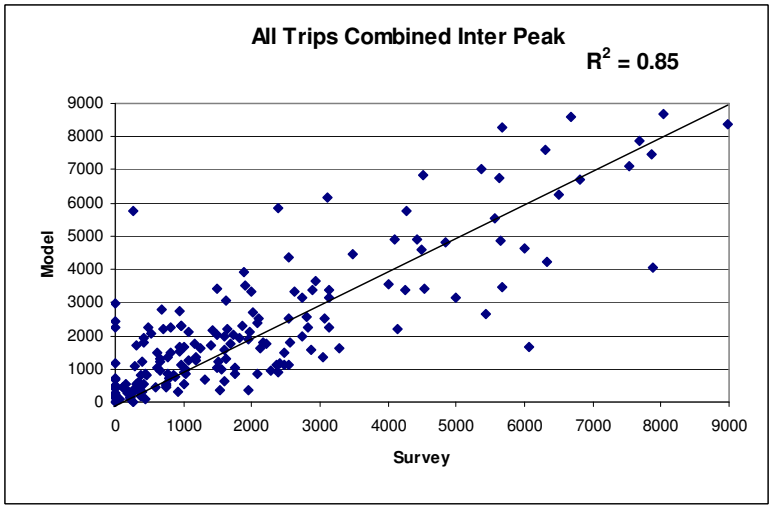


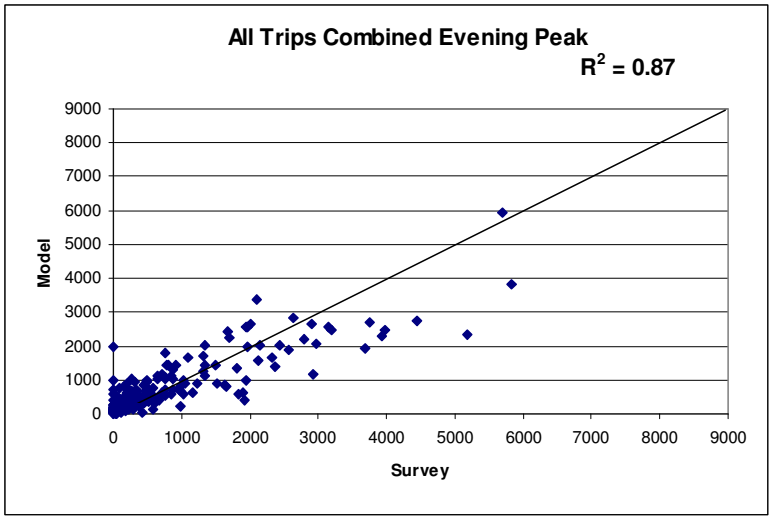




Non Home Based Attraction Trip End Validation	Figure 8







All Trips Combined Attraction Trip End Validation | **Figure 9**

11. 4 STEP MODEL VALIDATION OF ATTRACTION TRIP ENDS

The calibrated four step attraction equations have been included in the model and the resultant number of modelled trip ends by Regional model zone from the validated model have been compared against the HIS trips ends by zone. This comparison is intended to show how well the modelled trip ends validate against survey.

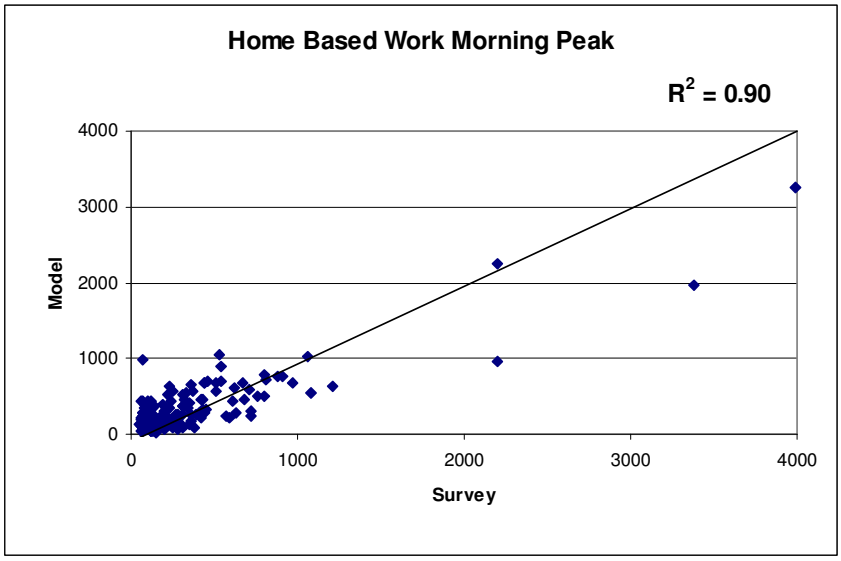
Note that given that the comparisons are included at regional zone level, there are a large number of zones with few survey households, meaning there will be a few outliers as a result especially for trip purposes with low levels of activity. The scatterplots for each modelled period (morning peak and interpeak only) are included in Figure 10 through Figure 16 for each trip purpose.

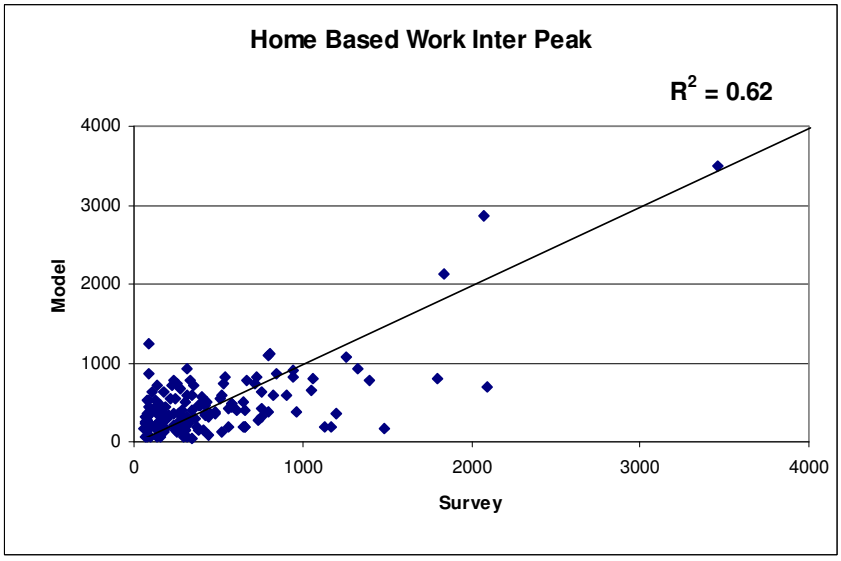
Each of the plots included an R-squared statistic, which measures goodness of fit. In general terms an R-squared of over 0.5 indicates there is a significant level of correlation between the two variables. It is important, however, to reiterate that with results calculated at zone level and with many of these zones having only a few surveyed households, it is difficult to address the issue of outliers.

In **Table 1** the total number of expanded HIS trips are reported with those purposes with fewer than 100 sampled trips surveyed (based on an average expansion factor of 100) in the corresponding time period presented in bold. These are generally those with the worst fit in the following tables which is not altogether unexpected given that the scatterplot analysis has been undertaken over nearly 200 zones, meaning the average number of sampled trips per zone is no greater than 0.5 in these instances.

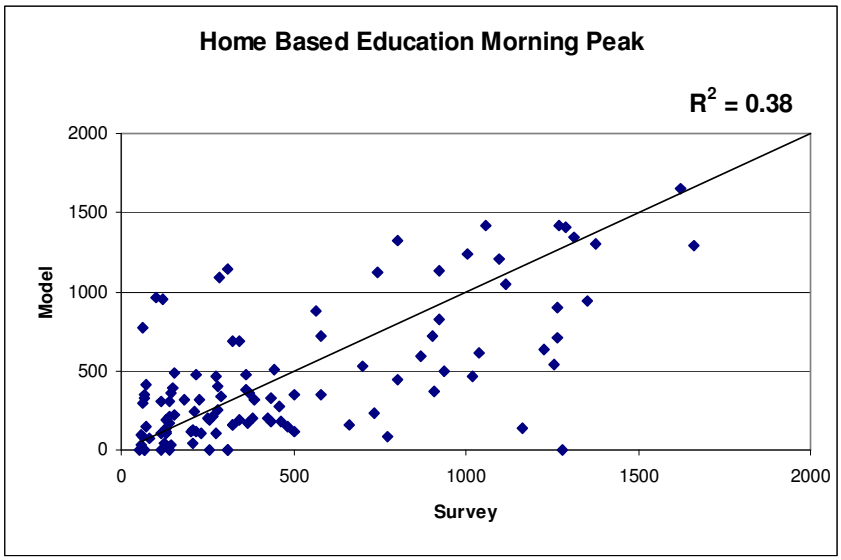
For those trip purposes with very few trips surveyed, the R-squared values are low. This was addressed in the calibration process by aggregating time periods for the Home Based Education and Home Based Social Recreational purposes. Putting these to one side the majority of R-squared values are over 0.5 with the critical Home Based Work peak periods and Non Home Based purposes over having R-Squared values of over 0.8.

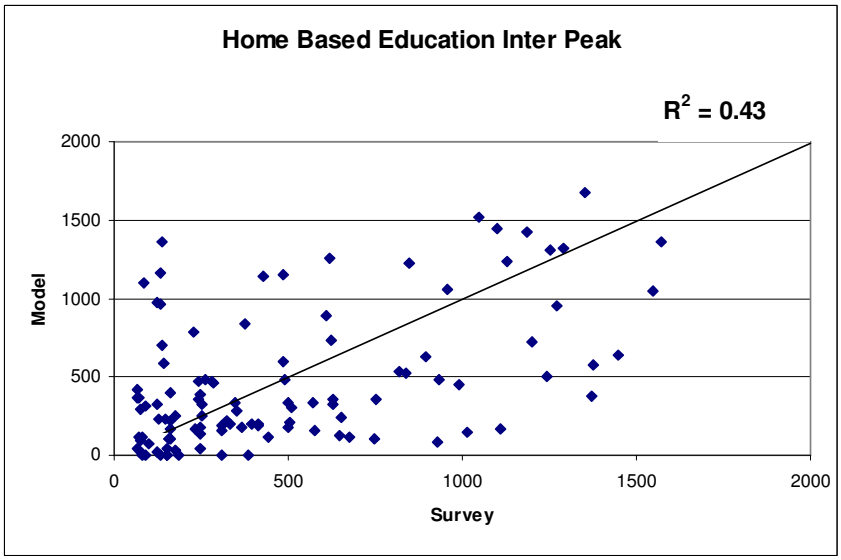
Also included is a scatterplot for all trip purposes aggregated for each of the three-modelled periods in Figure 17 with R-squared values of 0.85 to 0.89. These compare to values in the order of 0.8-0.9 from the North Shore model calibration. As mentioned in the previous section, the Model Specification Report initially suggested R-squared values at least 0.9 to be the goal, however in hindsight this is not practicable due to the 'lumpy' nature of HIS data. By comparison of R-squared values against the North Shore HIS it is proposed that a good measure of fit has been achieved.



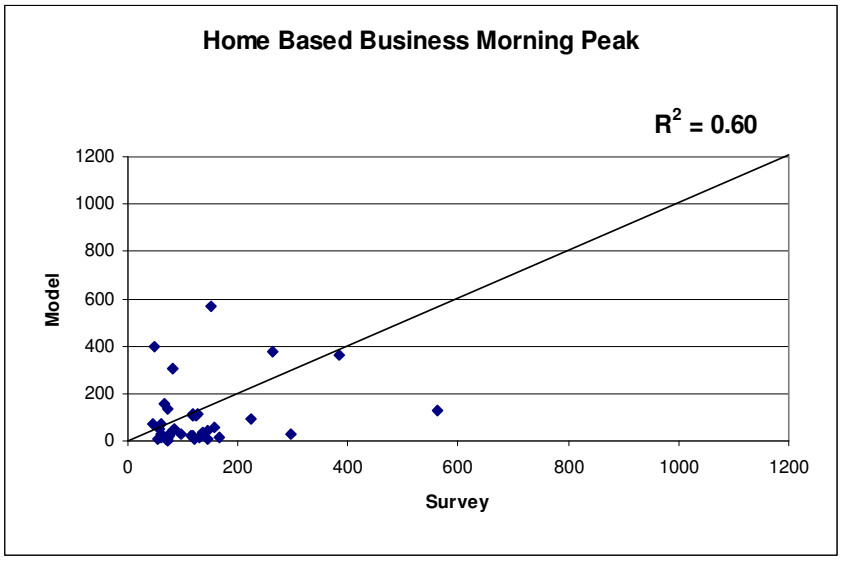


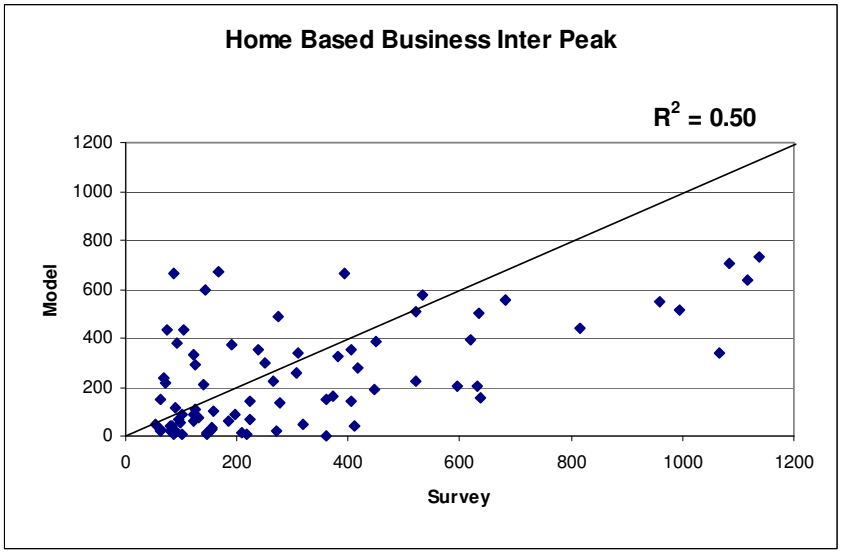
Home Based Work Attraction Trip End Validation	Figure 10



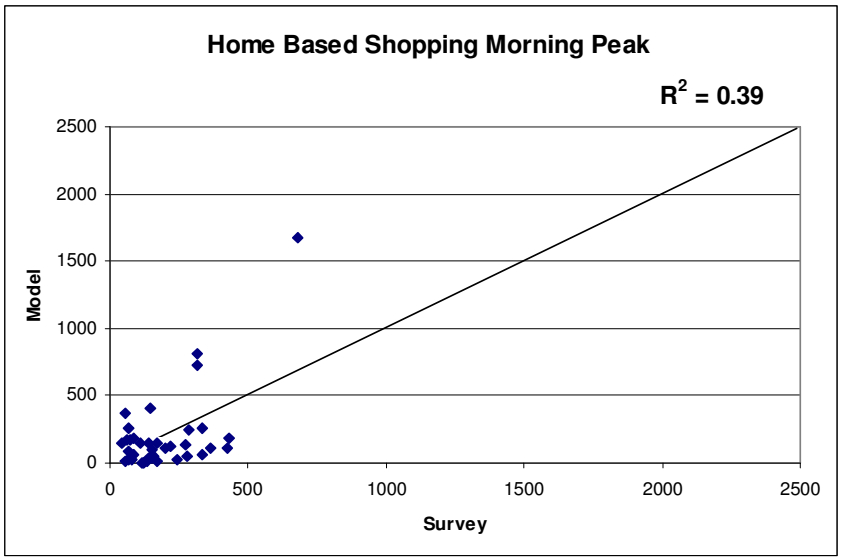


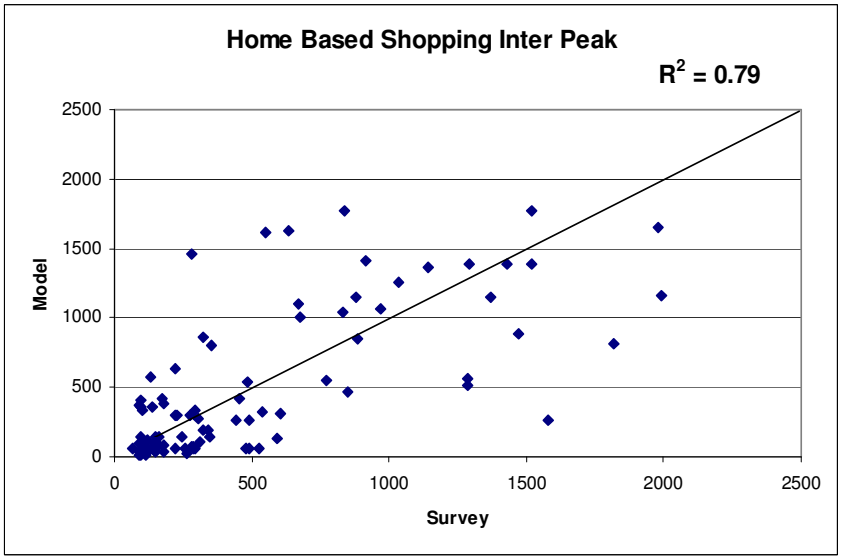
Home Based Education Attraction Trip End Validation	Figure 11



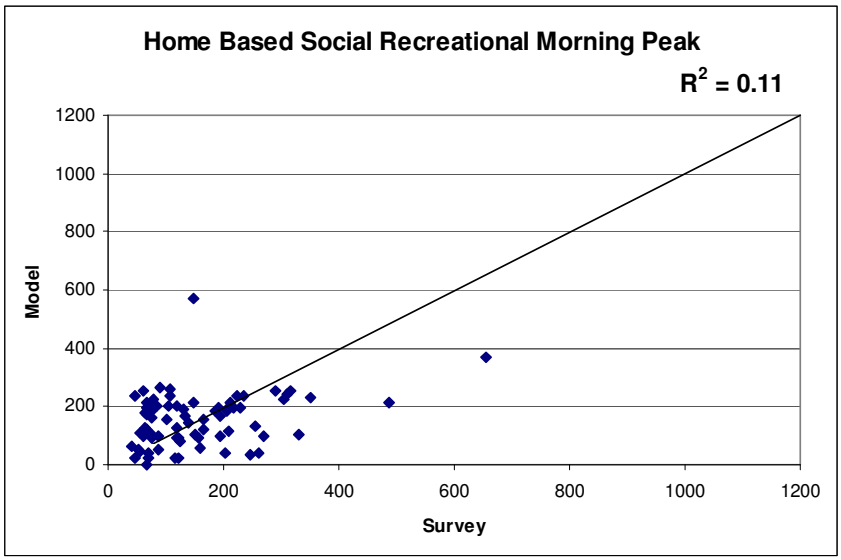


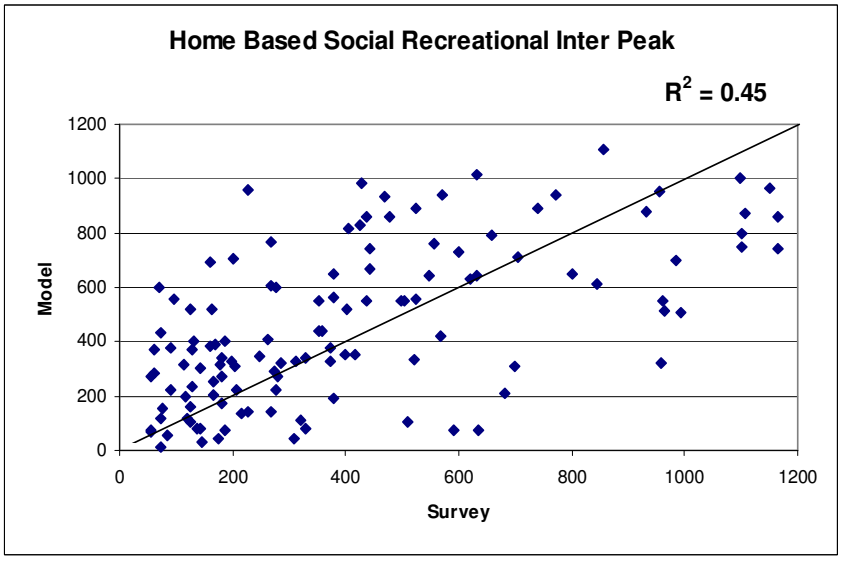
Home Based Business Attraction Trip End Validation	Figure 12



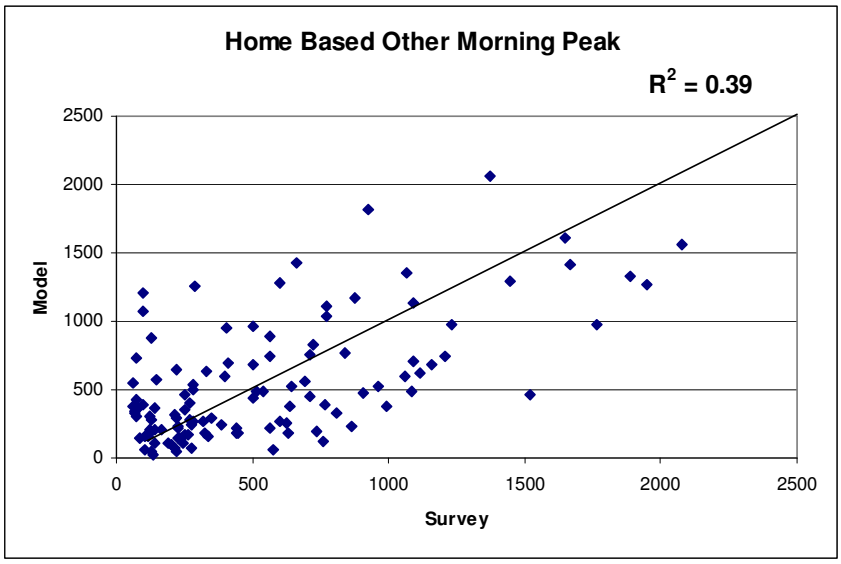


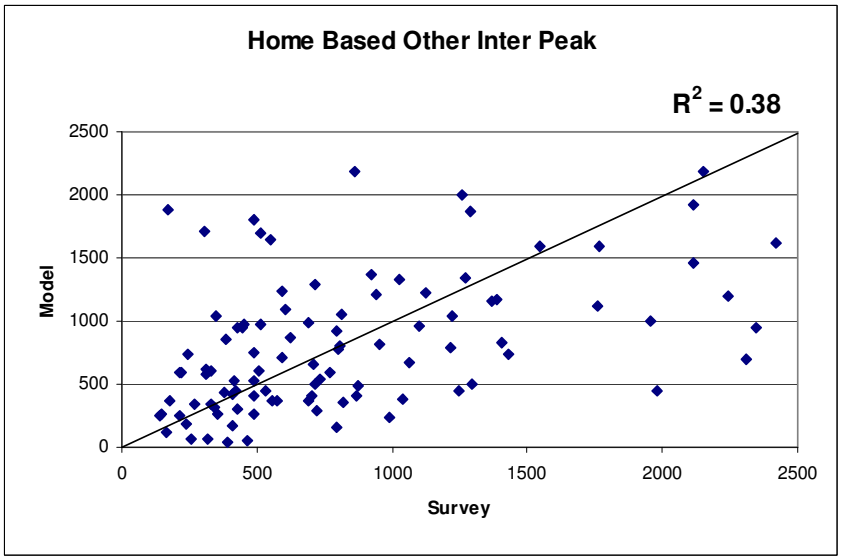
Home Based Shopping Attraction Trip End Validation	Figure 13



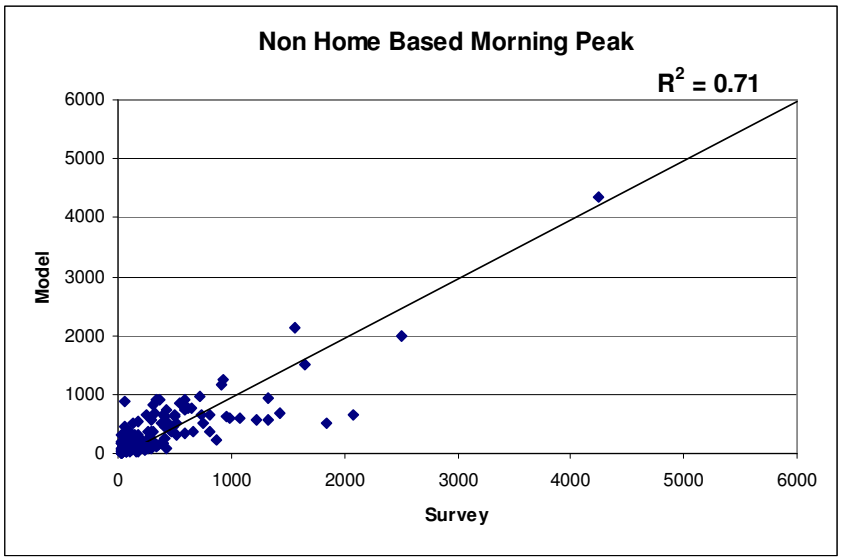


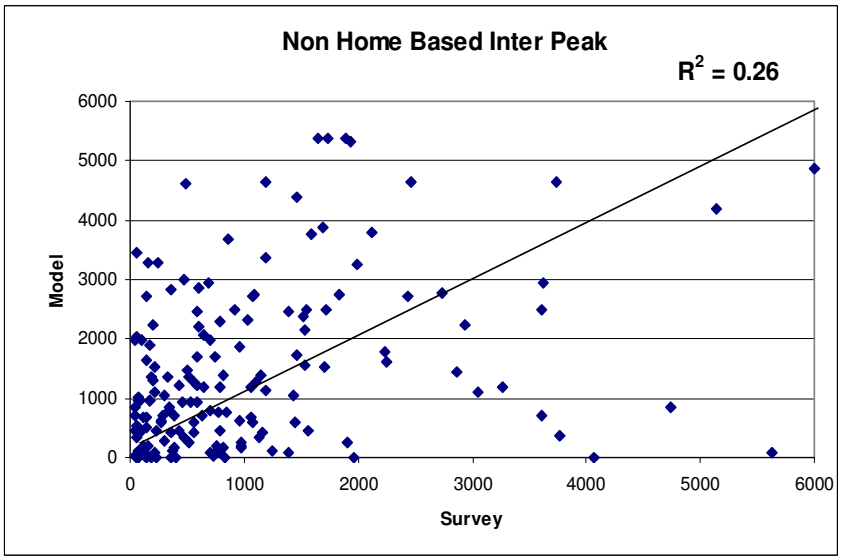
Home Based Social Recreational Attraction Trip End Validation	Figure 14



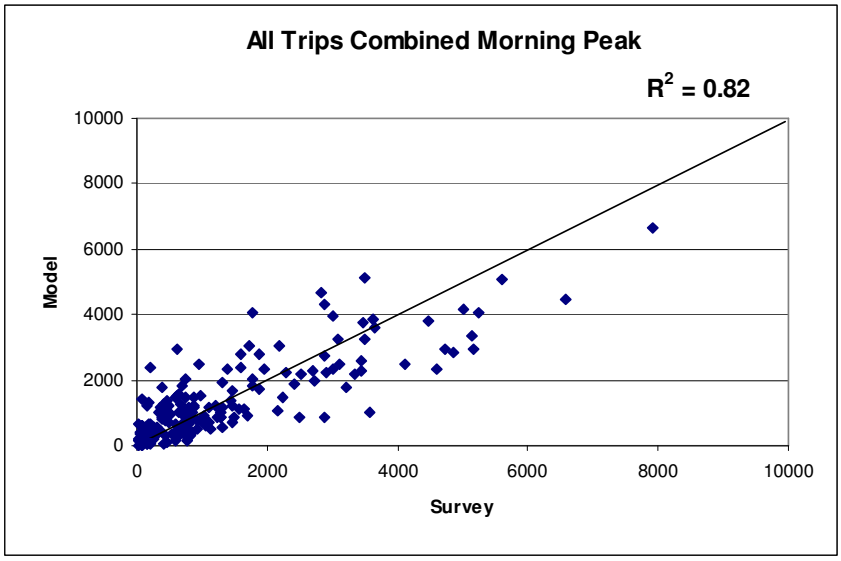


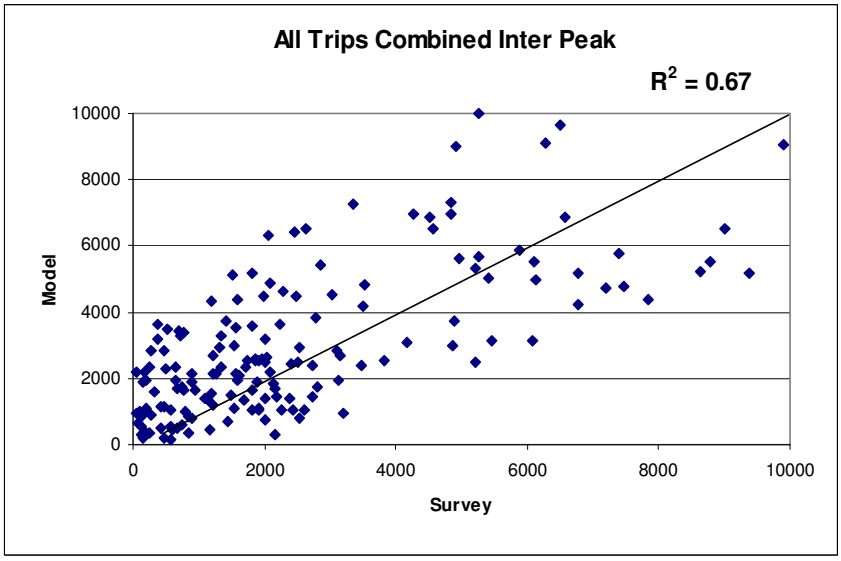
Home Based Other Attraction Trip End Validation | **Figure 15**





Non Home Based Attraction Trip End Validation	Figure 16





All Trips Combined Attraction Trip End Validation	Figure 17

APPENDIX ONE

Correlation Matrix (R)											
	Agric	Mining	Manu	Elect	Const	Whole	Retail	Accom	Transp	Info	Finance
Agriculture											
Mining	0.006										
Manufacturing	0.003	0.092									
Electricity	-0.081	0.470	0.263								
Construction	-0.058	0.091	0.880	0.278							
Wholesale	-0.086	0.046	0.897	0.352	0.917						
Retail	-0.141	0.025	0.451	0.490	0.530	0.612					
Accommodation	-0.119	0.047	0.224	0.486	0.332	0.397	0.910				
Transport	-0.060	0.062	0.661	0.407	0.638	0.732	0.649	0.538			
Information	-0.072	0.003	0.170	0.524	0.180	0.332	0.830	0.866	0.550		
Finance	-0.070	0.006	0.169	0.529	0.226	0.372	0.876	0.892	0.543	0.969	
Rental	-0.085	0.037	0.282	0.493	0.389	0.468	0.919	0.947	0.584	0.887	0.926
Professional	-0.066	-0.003	0.228	0.517	0.292	0.422	0.880	0.897	0.571	0.932	0.963
Administration	-0.039	0.013	0.280	0.541	0.334	0.469	0.873	0.877	0.618	0.951	0.970
Public Admin	-0.041	-0.007	0.278	0.509	0.319	0.454	0.895	0.873	0.578	0.915	0.950
Education	-0.183	0.018	0.155	0.378	0.248	0.294	0.656	0.677	0.385	0.633	0.664
Health Care	-0.110	-0.002	0.174	0.239	0.185	0.242	0.426	0.424	0.317	0.402	0.439
Arts	-0.009	-0.011	0.246	0.488	0.309	0.415	0.832	0.861	0.576	0.887	0.917
Other Services	-0.133	0.079	0.641	0.486	0.709	0.818	0.825	0.718	0.757	0.632	0.677
Total Jobs	0.029	0.057	0.594	0.530	0.630	0.722	0.925	0.847	0.746	0.810	0.847
Households	0.080	0.111	0.018	0.066	0.204	0.053	0.223	0.242	0.086	0.092	0.128
School	-0.116	0.040	0.001	0.138	0.097	0.056	0.307	0.316	0.148	0.219	0.247
Tertiary	-0.079	-0.022	0.121	0.367	0.169	0.259	0.621	0.694	0.356	0.708	0.731

Note:
Orange Highlight = Correlation coefficient > 0.8
Yellow Highlight = Correlation coefficient between 0.5 and 0.8

	Rental	Profes	Admin	Pub Admn	Edu	Health	Arts	Other	TotJobs	Hholds	School
Professional	0.942										
Administration	0.919	0.955									
Public Admin	0.906	0.939	0.945								
Education	0.735	0.704	0.671	0.640							
Health Care	0.439	0.517	0.440	0.429	0.376						
Arts	0.894	0.903	0.923	0.914	0.653	0.393					
Other Services	0.760	0.714	0.722	0.684	0.497	0.404	0.613				
Total Jobs	0.894	0.887	0.887	0.878	0.653	0.572	0.838	0.867			
Households	0.248	0.191	0.159	0.126	0.455	0.252	0.134	0.203	0.246		
School	0.342	0.302	0.270	0.214	0.688	0.208	0.223	0.225	0.281	0.548	
Tertiary	0.755	0.728	0.711	0.696	0.767	0.301	0.730	0.449	0.633	0.134	0.182

Note:
Orange Highlight = Correlation coefficient > 0.8
Yellow Highlight = Correlation coefficient between 0.5 and 0.8