1. PURPOSE

The purpose of this note is to document the procedure followed to calculate the first level expansion factors, and to check the sample against Stats data for the household categories outlined in Technical note 4/1. It deals with weekday travel only.

2. INTRODUCTION

The data base was supplied by Opus on 24th November 2008, and updated on 9th January 2009. There are five levels of expansion that can be applied to the data.

- A factor that relates the number of household sampled to the total number of households in the study area.
- A factor that is used to adjust for any bias in the sample representation of the household categories when checked against Stats data.
- A factor which accounts for under sampling of people in high occupancy households
- A factor to adjust the survey day of each household back to a March average weekday
- A factor to account for any under reporting when checked using the Roadside Interview surveys.

The derivation of each of these factors is discussed below

3. FIRST LEVEL EXPANSION

When the sample frame was established in June 2008, Opus were initially asked to sample 2000 households with a sample to be taken from every area unit in the study area, with an average sample rate of 1.4% (2000/147,000). On inspection, there were a number of area units that would have been sampling only a few households, and some were aggregated such that at least 10 households would be sampled in each. These aggregated area units were termed survey areas.

There are 249 area units in the area to be covered by the HIS. These reduced to 148 after aggregation[1].

In the event, Opus did not survey at least 10 households in each of the survey areas, and further aggregation was necessary to areas that were termed expansion areas..

[1] The area units make up the zones in the regional model and have been numbered 650 to 849. In the spreadsheets that calculate the expansion factors, aggregated zones are numbered 1 to 34, with the other zones retaining there original numbers.

Once that was done, expansion factors were calculated. Several were very high, and a third round of aggregation was undertaken so that all expansion factors would be less than twice the mean – approx 150. One expansion factor fell outside this criterion. The result is a set of factors that range from 23.0 to 174.9, with the lowest sample size ranging from 10 to 53.

This final round of aggregation reduced the number of sample areas to 96. The geographic location of the expansion areas is shown on Figure 1 to 4, with the meshblocks that were sampled in each also highlighted. Figures 5 to 8 show the first level expansion factors for each area. The expansion factor was applied to the interviews from all households in the aggregated area.

The next step was to allocate the expansion factors to each of the sample households according to the meshblock in which they occurred, and to transfer that information to each household record in the database.

During the model validation process it was noted that there were inconsistencies between urban and rural trip generation, and upon further investigation it was discovered that by expanding urban and rural households by the aggregated sample areas, they were not being fairly represented. Some areas were being over-represented and others under-represented.

As such the first level expansion was revisited with the application of factors for urban and rural areas. Urban areas were defined as those with at least 10 dwellings per hectare. The urban and rural regions were further split into Hamilton and surrounds and the remainder of the study area. As such there were four areas as indicated on Figures 9 and 10, which are:

- Hamilton and surrounds urban
- Hamilton and surrounds rural
- Regional urban
- Regional rural

An expansion factor for each of these four areas was calculated and applied to all of the households within each. The expansion factors are shown in **Table 1**.

Household Categories — **Table 1**

Region	Total HH in Area	Sampled HH in Area	Expansion Factor
1 Hamilton Urban	57852	877	65.966
2 Hamilton Rural	10230	197	51.929
3 Regional Urban	24096	402	59.940
4 Regional Rural	55056	536	102.716

4. SECOND LEVEL EXPANSION

The second level of expansion was to check that the distribution of the sample households by household category was representative of that in the 2006 Census data. The number of households in each of the 18 categories noted in Table 1 below was requested from the Stats department, stratified by cars available, for each area unit.

This request produced a number of cell where the observations in the stats data breached their confidentiality policy and the 18 categories were reduced to 12, also shown on Table 1.

Household Categories		Table 2
Description	**Category**	**Grouped category**
1 Employed	1	1
1 Student	2	2
1 Not working	3	
2 Both working	4	3
2 Both Education	5	4
2 Both not working	6	
2 - 1 work, 1ed	7	5
2 - 1 work, 1 n/w	8	
2 - 1 ed, 1 n/w	9	4
3+ at least one working	10	6
3+ at least one student	11	
3+ all not working	12	
1+teens	13	7
1+kids	14	8
1+both	15	9
2+teens	16	10
2+kids	17	11
2+both	18	12

As discussed in the technical note on the household distribution model, it proved impossible to calibrate probability functions for the 12 categories, and a revised household category definition was formed, with the first cut being whether or not a household had children, and then categorisation by household size.

The sample distribution was then compared against regional totals in each of the categories. Table 2 below shows the expanded household totals against regional numbers for the 10 categories.

Category expansion comparison with regional totals						Table 3
Category	Description	Expanded Sample		Stats Population		Factor 2
1	One Adult working	14,858	10.2%	15,820	10.8%	1.065
2	One adult not working	17,513	12.0%	16,226	11.0%	0.927
3	Two Adults working	22,496	15.4%	26,487	18.0%	1.177
4	Two adults not working	17,960	12.3%	12,308	8.4%	0.685
5	Two adults one working	11,918	8.2%	9,078	6.2%	0.762
6	Three + adults	8,274	5.7%	9,788	6.7%	1.183
7	Two Person Family	5,340	3.7%	4,992	3.4%	0.935
8	Three person Family	14,842	10.2%	16,198	11.0%	1.091
9	Four Person Family	18,989	13.0%	19,590	13.3%	1.032
10	Five + Person Family	13,581	9.3%	16,614	11.3%	1.223
Total		***145,770***		***147,101***		

The ratio of stats population against the expanded sample was then calculated as the second level expansion factor and applied to the interviews of each household in the category

5. THE THIRD LEVEL EXPANSION FACTOR

The third level expansion factor was required to adjust for a small under sampling of people in high occupancy households – in this case categories 6, 8,9 and 10. The total number of people in the HIS area was 392,871 from the 2006 census. After the second level expansions were applied, the number of people in the sample was 389,352 – less then 1% light. Accordingly, the 251,576 people (and trips) in the 3+ households were increased by 1.4% so that the number of people in the HIS data matched the Stats total.

6. THE FOURTH LEVEL EXPANSION FACTOR

All of the traffic count data was normalised to an average March weekday, being the design day for the model and the time when the census data was collected.. The HIS surveys were undertaken during July to October 2008, and the trips made on the survey day were also normalised to a March weekday using the appropriate factor from the table include as Appendix One. The factors were calculated from all of the telemetry count sites in Waikato and were applied to each trip that started in each period. The Interpeak factors show a significant degree of variability, and almost all weeks are less than an average week in March except for the holiday periods of January and February.

7. COMPARISON WITH THE ROADSIDE INTERVIEW SURVEY

The Roadside Interview survey covered 18 sites in the model area, plus sites from The Auckland and Tauranga Roadside Interviews. Broadly these can be arranged into a cordon that represents the external roads entering and leaving the study area, an internal cordon covering the main roads on the edge of the Hamilton urban area – the area defined as the urban model area, and a screen line using the Waikato River.

In order to check the level of under-reporting, (if any) the 24 hour trip matrix extracted from the HIS was compared against the manual classified counts taken at each of RIS stations. Only light vehicles were compared. The RIS did not cover the bridge at Ngaruawahia and the automatic traffic count was used at that location adjusted for the same proportion of heavy vehicles as were counted on the Wairere Bridge.

The HIS only contains those trips that are made by residents of households within the model area – it doers not account for trips made into and out of the model area by people who live outside the area.

Those trips were abstracted from the RIS interview data for the stations on the external cordon, and were added into the HIS trip matrix. Because the RIS was in only one direction these external trips were transposed and added in to get the reverse movement.

For the record, there are 1,003,409 light vehicle trips in the matrix (6.8 trips per household) with an average trip length of 13.6 km, and 13.3 minutes – an average speed of 58.7 km/hr. This is a little faster than in a normal urban network, but about half of the travel is in the faster rural part of the model.

Table 4 below shows the results of the comparison between assigned HIS trips and the Roadside Interview count stations. At this stage it is only useful to report cordon and screen line totals as the 'lumpy' nature of the surveyed matrix does not necessarily permit the correct assignment to each particular surveyed location. The assignment issue is particularly noticeable at the river screenline.

Nevertheless, it is clear from the table that there is minimal, under reporting, particularly at the internal cordon around Hamilton. There is no basis to apply any further factors.

Comparison of Assigned HIS trips with classified counts (vehicles per average weekday)			Table 4
	External Cordon	**Internal Cordon**	**River crossings**
Inbound			
Counted	26,052	28,215	84,921
Assigned	24,910	28,139	71,538
Diff	-1142	-76	-13,383
%	95.6%	100.0%	84.2%
GEH	0.5	0.1	9.8
Outbound			
Counted	27,821	28,818	84,599
Assigned	24,217	28,239	72.089
Diff	-3604	-579	-12,510
%	87.0%	98.0%	85.2%
GEH	4.6	0.7	9.1
Total			
Counted	53,873	57,033	169,520
Assigned	49,127	56,378	143,627
Diff	-4746	-655	-25,893
%	91.2%	98.9%	84.7%
GEH	4.3	0.6	13.4

8. THE DATABASE

The Home Interview data is stored in two formats. The first is in a Microsoft Access data base (.mdb) with the expansion factors described above included. Any address information, apart from geocodes has been removed.

The second is an ASCII (text) dataset organised in the form of Household 1, person 1, trips for person 1, person 2 trips per person 2 etc. The data in the ASCII file is a subset of the Access database in that it only includes the information needed for model calibration. A purpose written FORTRAN program is used to prepare the household trip rates reported in Technical Note 7. The file includes all of the expansion factors except the conversion from survey day to design day factors which are read separately by the software.

Table 5 includes the key data that needs to be matched by any analysis using either dataset. Note that 5 trips were removed from the ASCII dataset because of invalid formatting, so it has 26,015 records in it, rather than 26,020 in the Access database.

Key HIS data totals		Table 5
	Unexpanded	***Expanded***
Households	2,012	147,101
Persons	5,189	392,997
Valid Trips	18,801	1,650,299
Invalid trips[2]	18	1,141
Total records	26,020	N/A

9. CONCLUSIONS

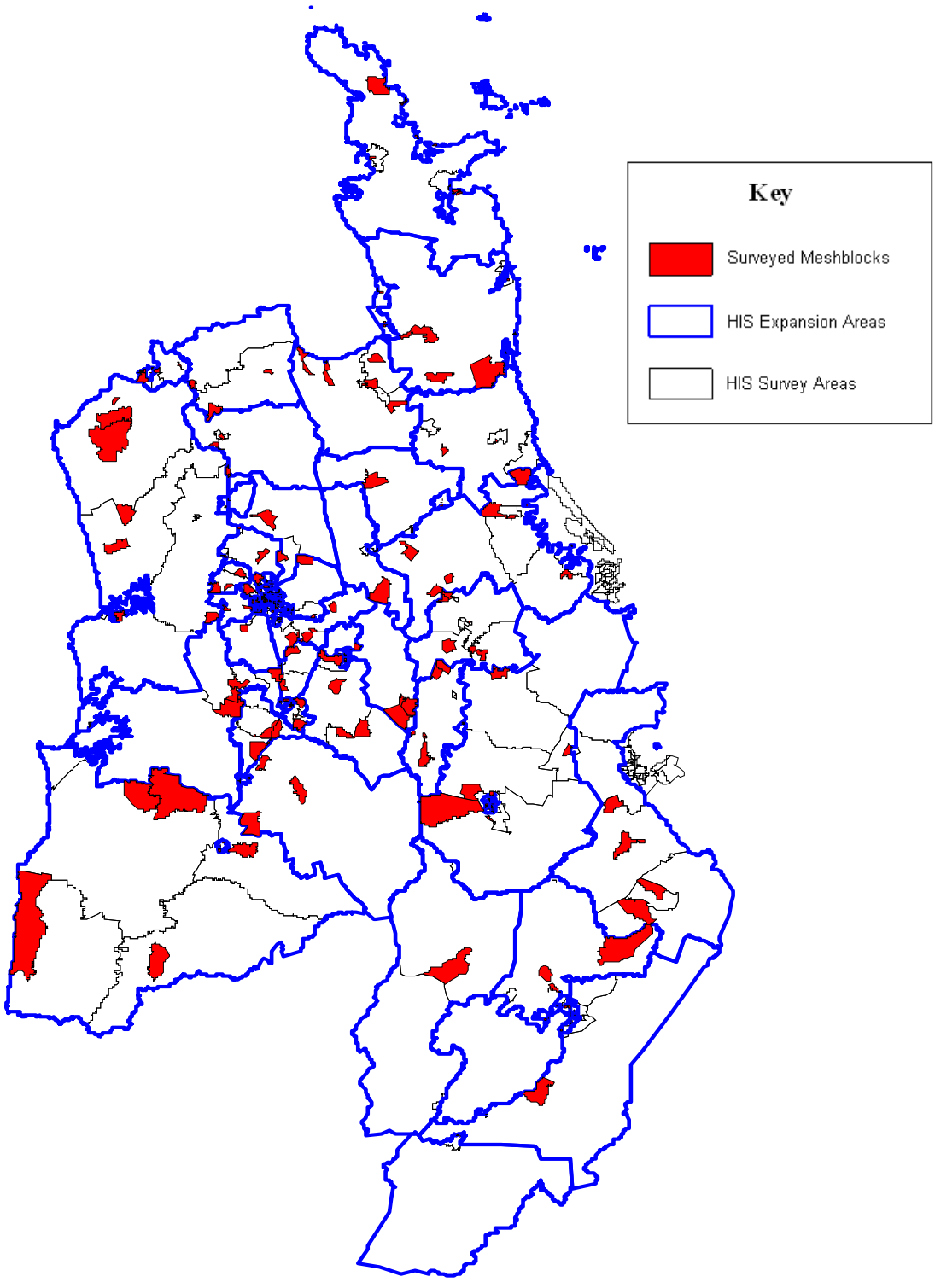
With the application of the various expansion factors as described above, the Home Interview Survey data is suitable for use for model calibration.

[2] Invalid trips are home to home – generally recreational.

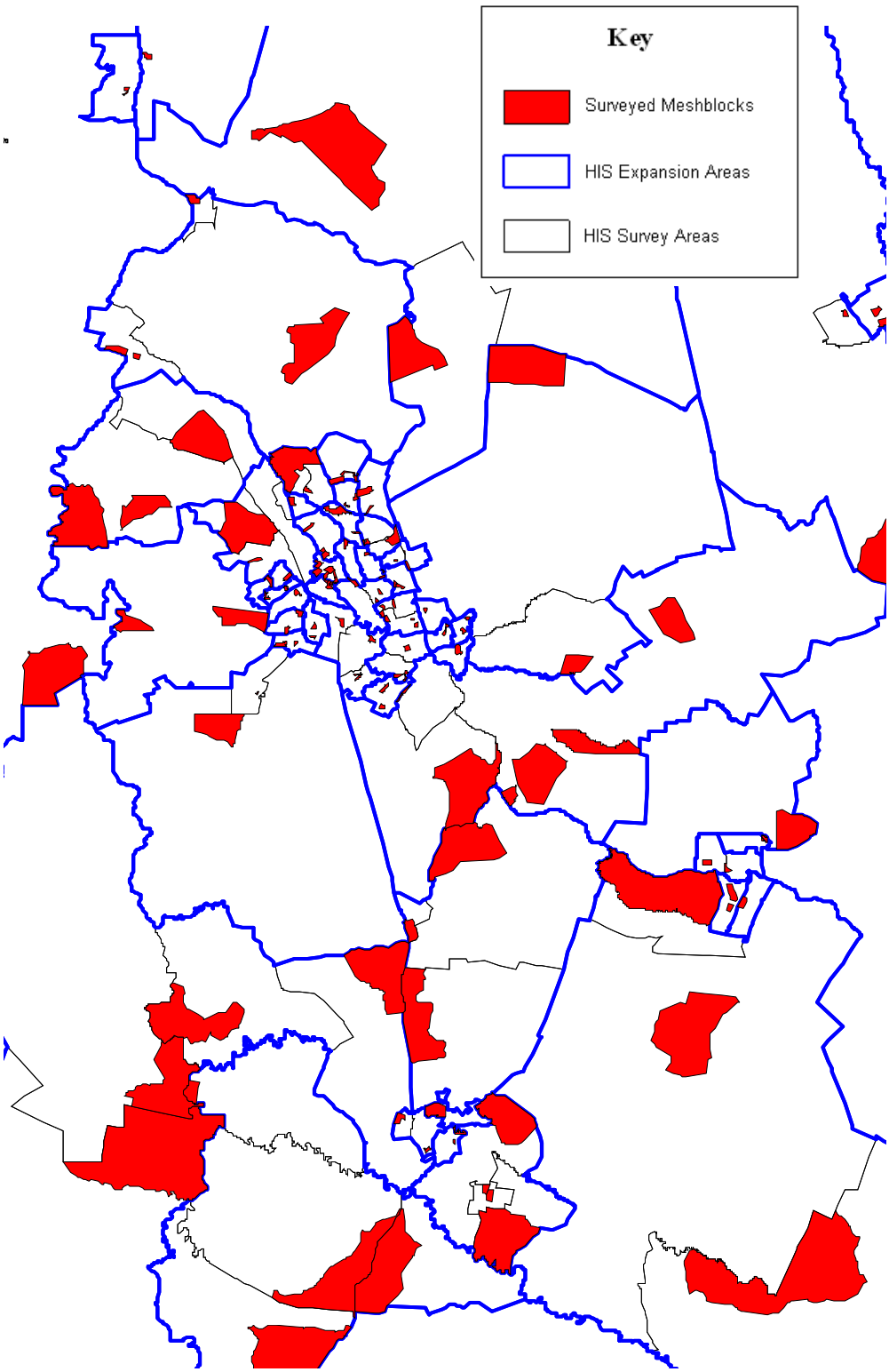
Conversion of HIS Survey week to March.

Date	Factors		
	AM Peak	Inter peak	PM Peak
8-Jan	1.3399	0.7854	0.8948
15-Jan	1.1267	0.8396	0.9224
22-Jan	0.9927	0.9097	0.948
29-Jan	1.0289	0.9076	0.936
5-Feb	1.1295	0.8961	0.9145
12-Feb	1.113	1.046	1.0264
19-Feb	1.0056	1.0465	1.0351
26-Feb	0.9674	1.0237	1.01
5-Mar	0.9754	1.0408	1.0332
12-Mar	0.9877	1.0644	1.0273
19-Mar	0.9733	1.0842	1.0326
26-Mar	0.9644	0.9283	0.9686
2-Apr	1.0851	0.9121	0.9356
9-Apr	1.0419	1.1712	1.0838
16-Apr	0.9913	1.1544	1.0767
23-Apr	1.0211	1.1368	1.1174
30-Apr	1.0487	0.8915	0.9741
7-May	1.066	0.9985	1.003
14-May	1.0345	1.2037	1.1377
21-May	0.9974	1.1989	1.1034
28-May	1.0286	1.2482	1.1444
4-Jun	1.0349	1.1691	1.064
11-Jun	1.2059	1.1118	1.0472
18-Jun	0.8818	1.1839	1.0309
25-Jun	1.0933	1.2888	1.1999
2-Jul	1.1018	1.2236	1.1764
9-Jul	1.0936	1.2301	1.1657
16-Jul	1.1638	1.0616	1.049
23-Jul	1.1279	1.0515	1.0476
30-Jul	1.1224	1.3298	1.2408
6-Aug	1.1783	1.3564	1.3022
13-Aug	1.0618	1.2784	1.179
20-Aug	1.1078	1.2938	1.2338
27-Aug	1.0444	1.2542	1.1752
3-Sep	1.0158	1.2416	1.1249
10-Sep	1.0214	1.2181	1.1188
17-Sep	1.0527	1.2514	1.1749
24-Sep	0.9935	1.211	1.1152
1-Oct	0.9822	1.1938	1.0975
8-Oct	1.0753	1.0108	1.0143
15-Oct	1.1121	1.026	1.0445
22-Oct	1.0249	1.232	1.1435

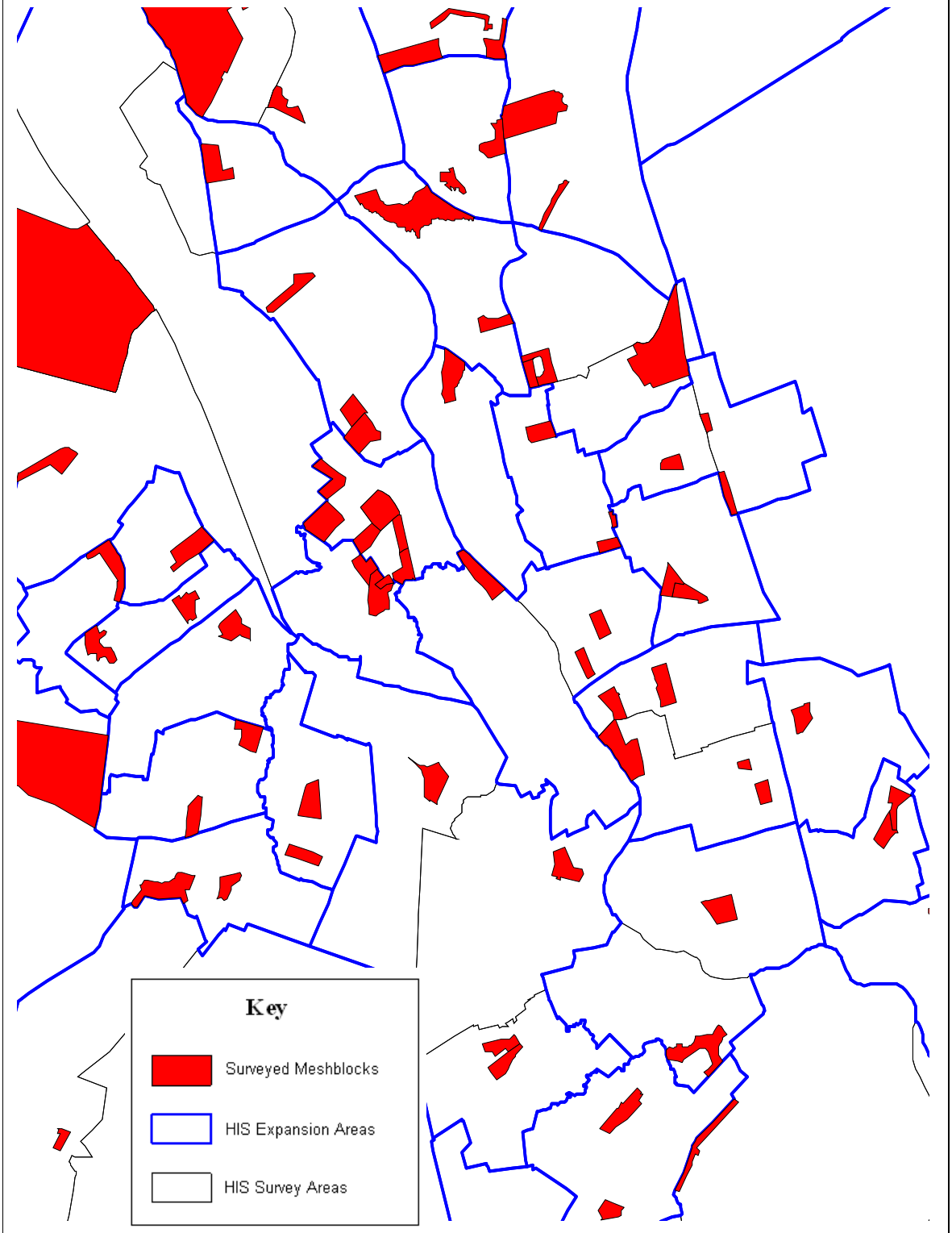
29-Oct	0.9905	1.0973	1.0277
5-Nov	1.1169	1.0059	0.9611
12-Nov	1.0161	1.2239	1.1693
19-Nov	0.9633	1.1606	1.0901
26-Nov	0.9873	1.1497	1.0917
3-Dec	1.0001	1.1246	1.0546
10-Dec	0.9943	1.1535	1.0887
17-Dec	0.9843	1.1159	1.0782
24-Dec	0.9895	1.0766	1.088
31-Dec	1.2114	0.7625	0.9022



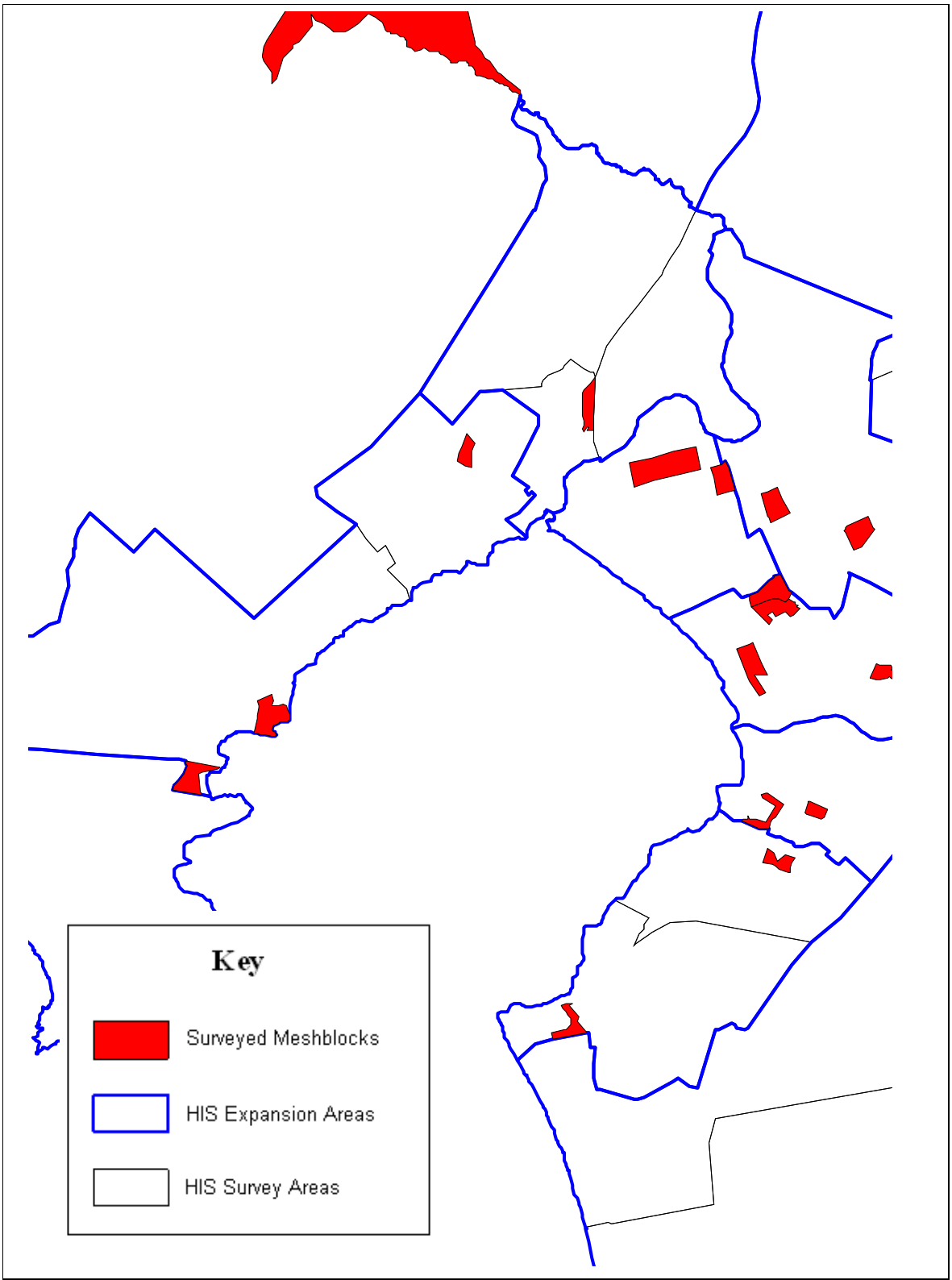
Waikato Regional Transport Model Household Interview Expansion Factors	**Overview of Waikato**	**Figure 1**



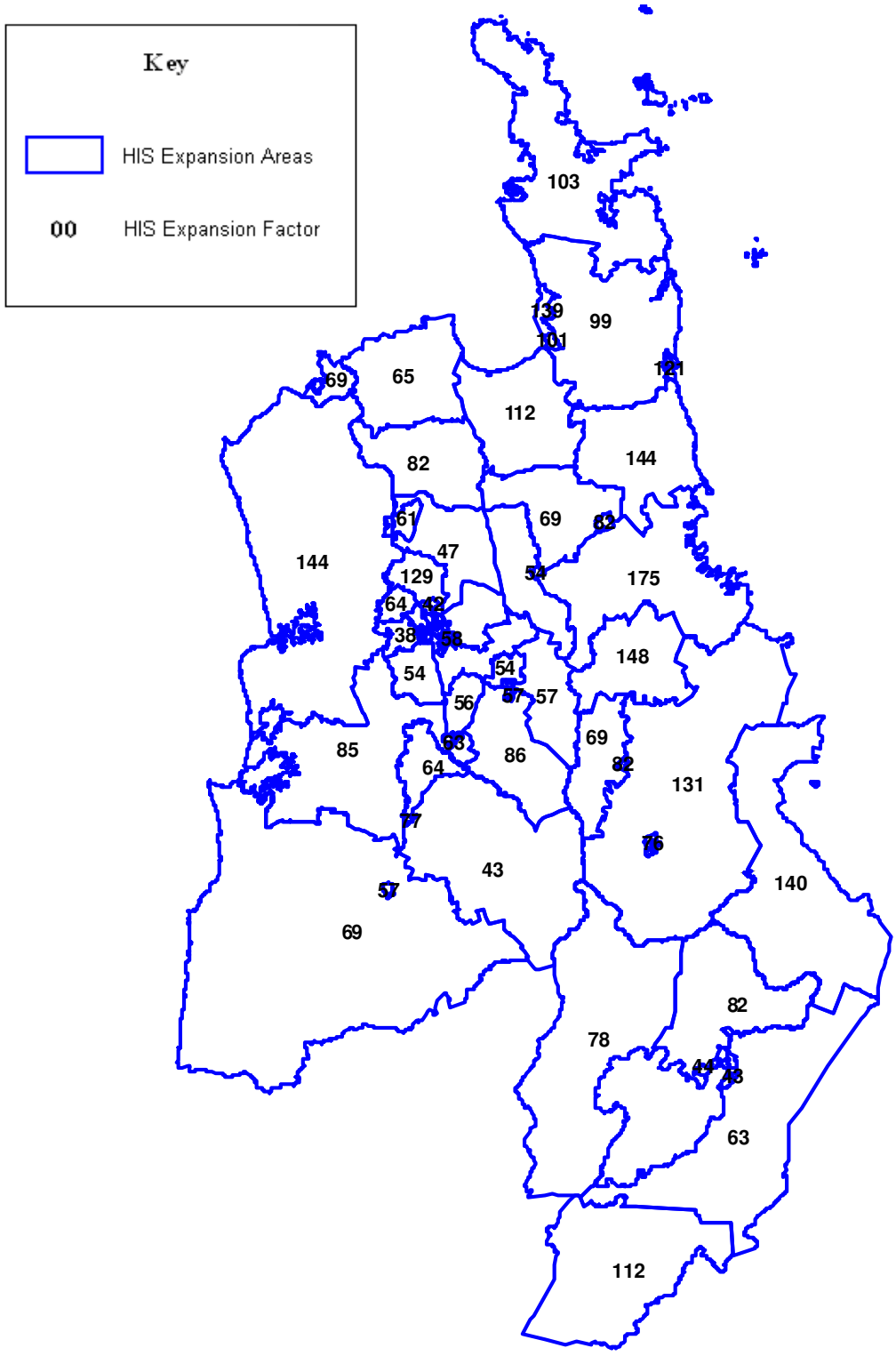
Waikato Regional Transport Model	**Overview of Hamilton Urban Study Area**	**Figure 2**
Household Interview Expansion Factors		



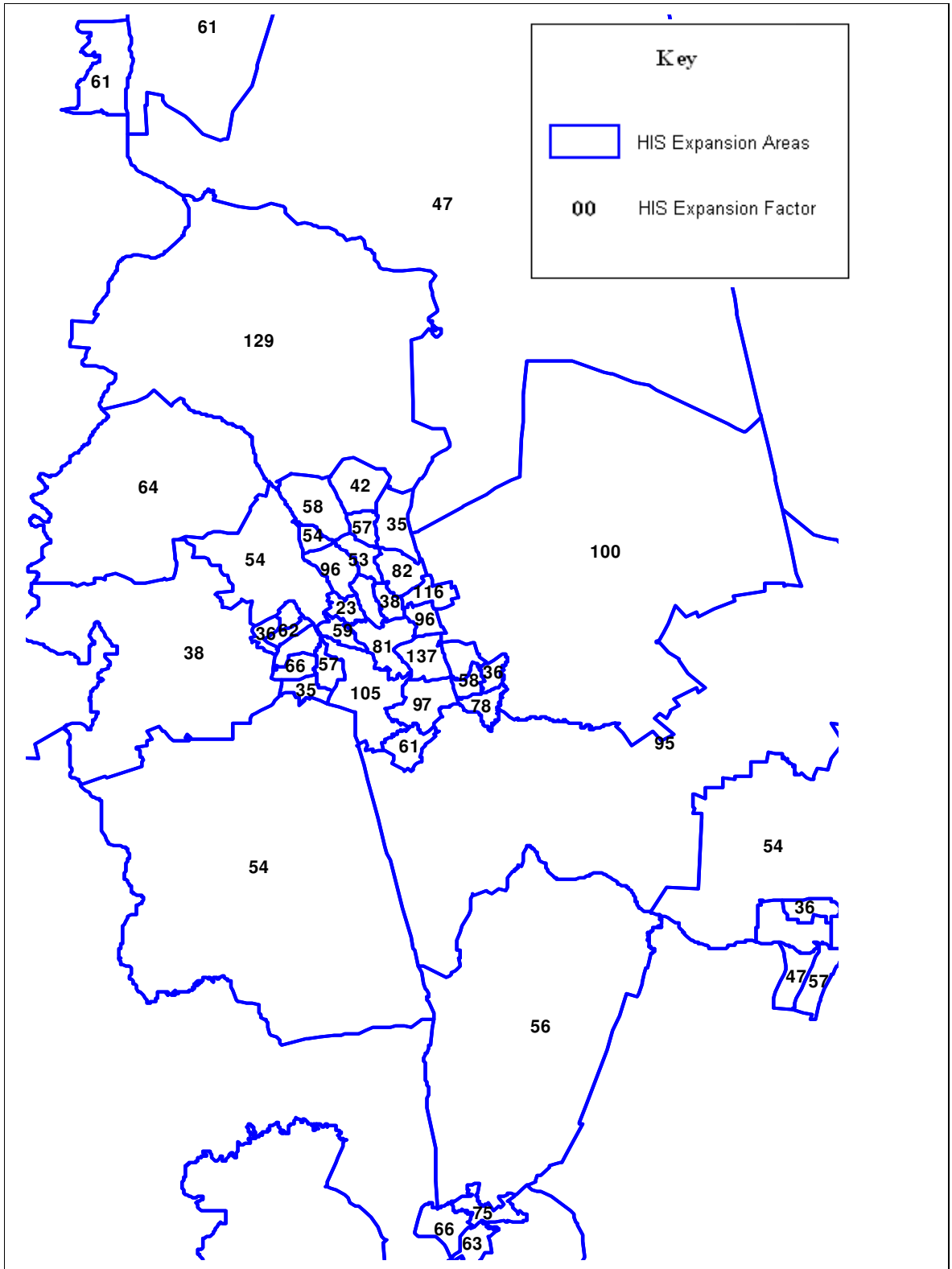
Waikato Regional Transport Model	**Close up Of Hamilton City**	**Figure 3**
Household Interview Expansion Factors		



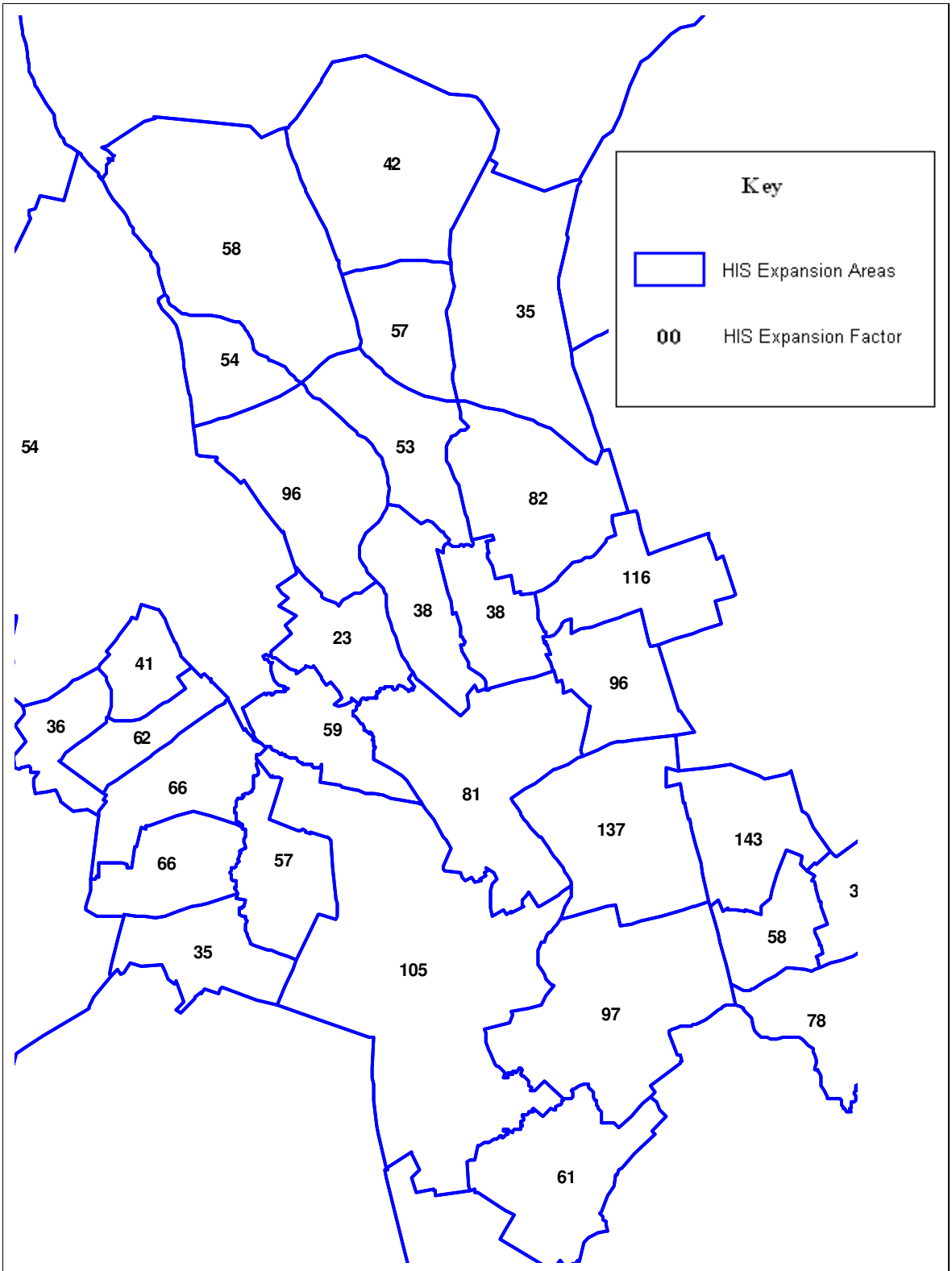
Waikato Regional Transport Model	**Close Up of Taupo**	**Figure 4**
Household Interview Expansion Factors		



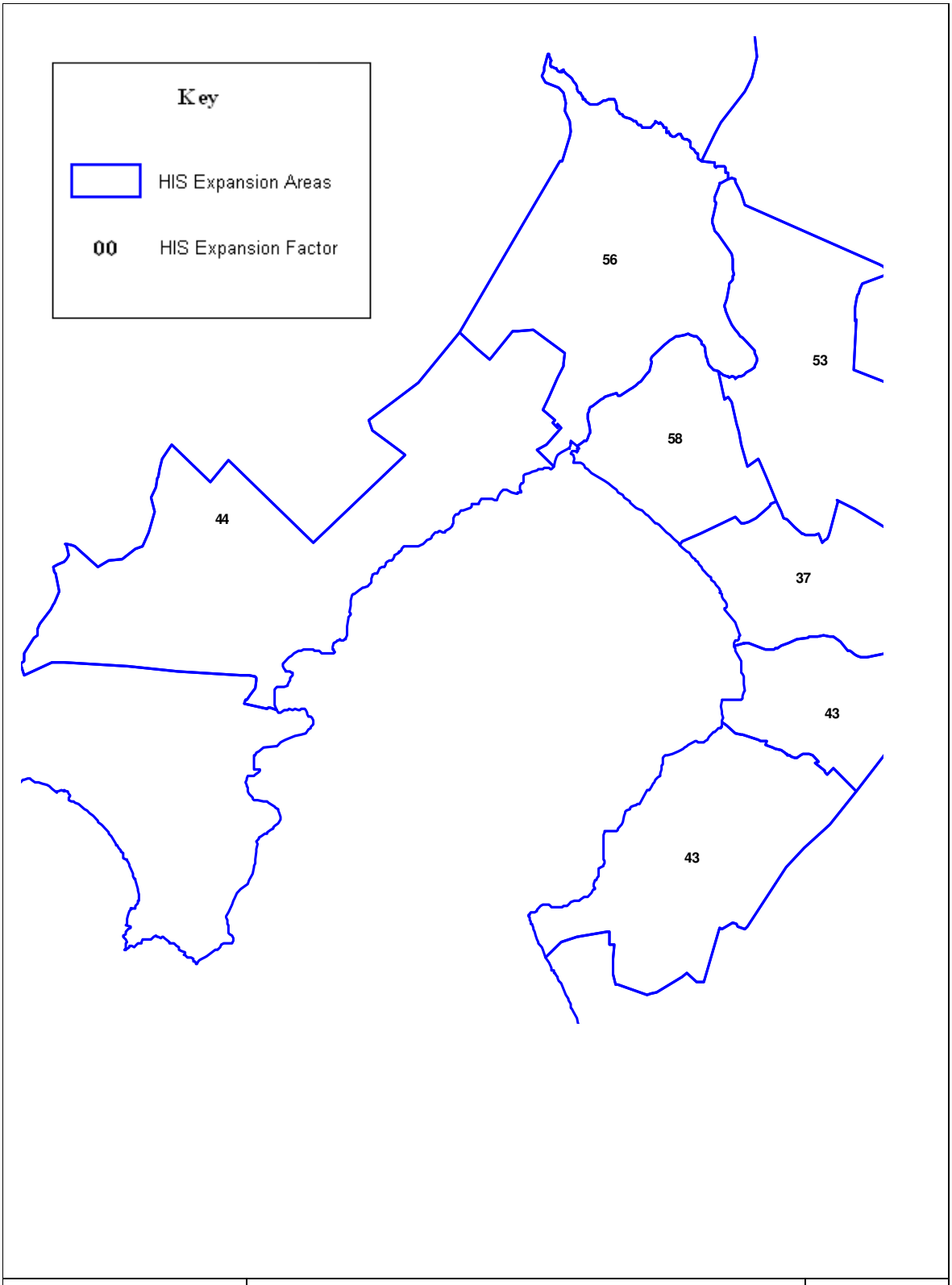
Waikato Regional Transport Model	**Overview of Waikato**	**Figure 5**
Household Interview Expansion Factors		



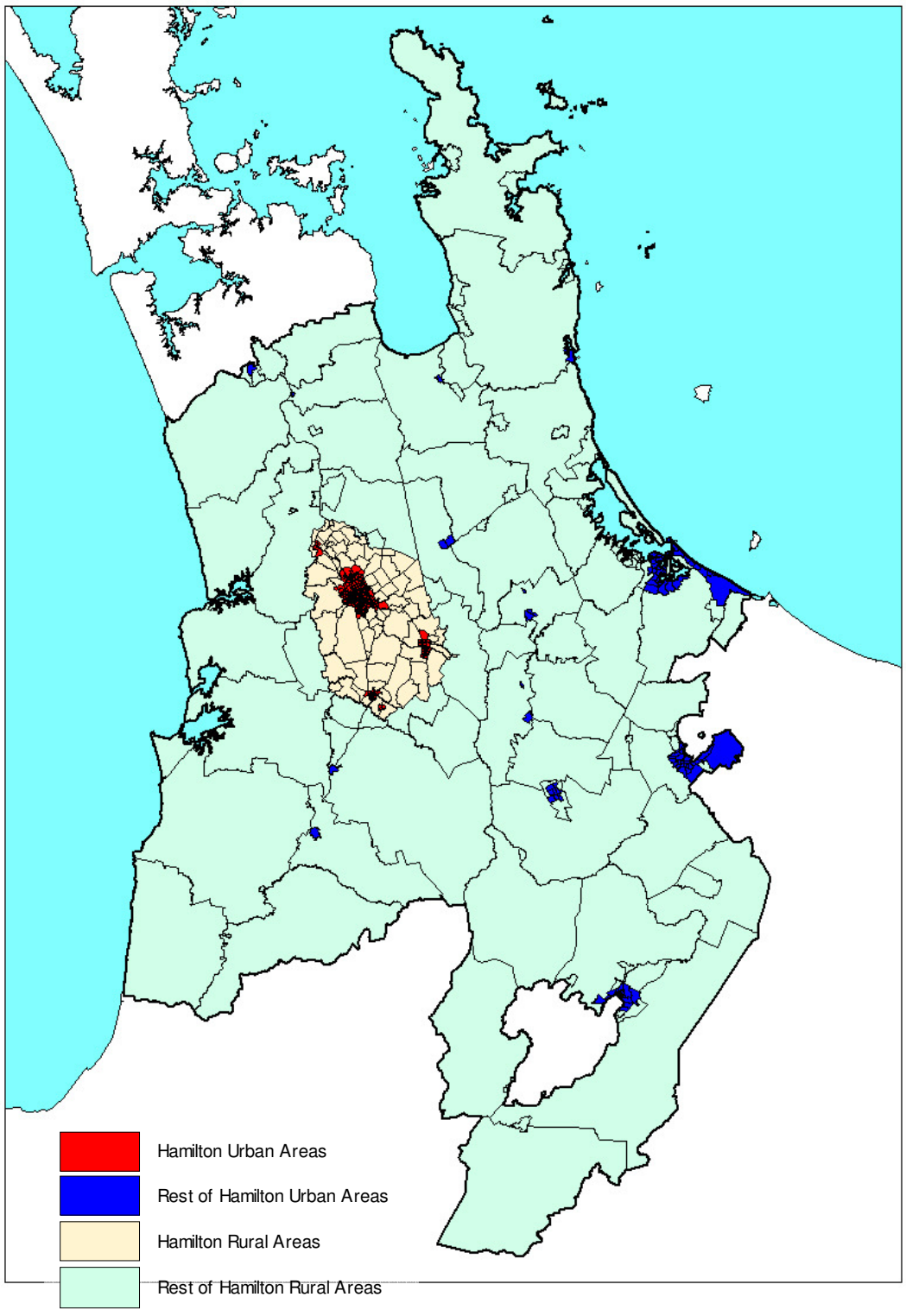
Waikato Regional Transport Model	**Overview of Hamilton Urban Study Area**	**Figure 6**
Household Interview Expansion Factors		



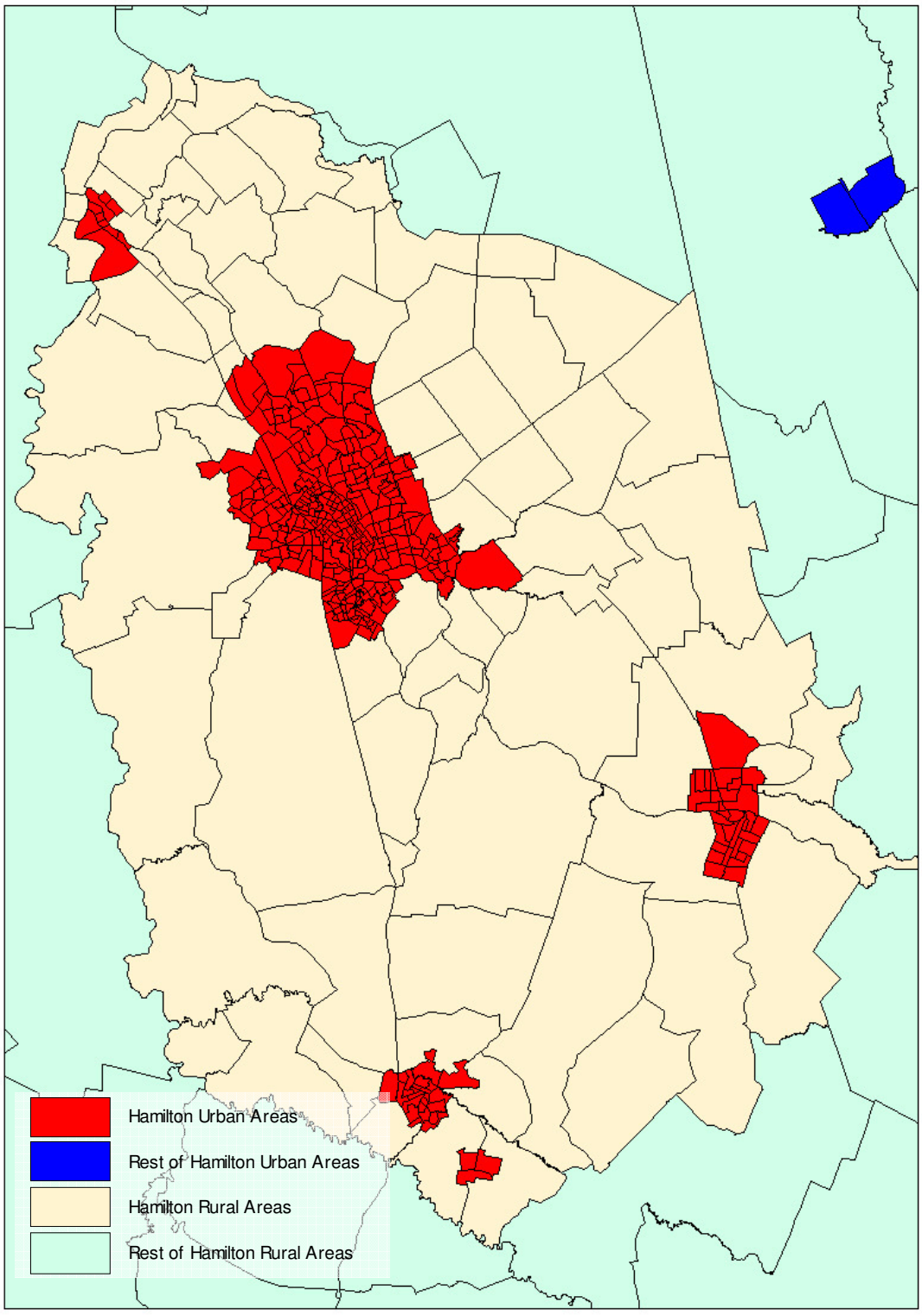
Waikato Regional Transport Model	**Close up Of Hamilton City**	**Figure 7**
Household Interview Expansion Factors		



Waikato Regional Transport Model	**Close Up of Taupo**	**Figure 8**
Household Interview Expansion Factors		



Waikato Regional Transport Model	**Definition of Urban and Rural Areas**	**Figure 9**
Household Interview Expansion Factors		



Waikato Regional Transport Model	**Definition of Urban and Rural Areas Inset**	**Figure 10**
Household Interview Expansion Factors		