C. Household Interview Diary Survey Files

Household Interview Survey Data: The 1995 Household Interview Survey (HIS) collected complete travel information from a sample of 4,060 Oahu households. Households were recruited for participation by telephone using a random-digit-dial method that includes both listed and unlisted telephone numbers. The travel information was retrieved via telephone using a computer-assisted telephone interviewing (CATI) system.

Survey Procedures

The survey procedures included the following five steps.

1. Households were recruited to participate in the study via telephone. The initial telephone call asked basic questions about the household to maximize response rate to these crucial questions, to allow analysis of households that dropped out before completion, and to enable packaging of the proper number of travel diaries.
2. Households were sent via mail a set of survey materials, including a cover letter, household information sheet, and one travel diary for each household member.
3. The night before the assigned travel day, a reminder telephone call was placed to the recruited households to verify the materials' arrival and answer any questions the respondents might have had.
4. The day after their assigned travel day, telephone calls were made to the household to retrieve the travel information. The travel information was input by telephone interviewers directly into a Computer-Assisted Telephone Interview (CATI) system. The interviewers requested to speak with each member of the household to retrieve travel information. If a household member was unavailable or too young, another adult in the household read that person's travel information from his or her diary.
5. Respondent data was checked for both accuracy and consistency, and interviewers made clarification calls where appropriate.

Error-Checking Procedure

A custom-written computer program, called HISCHK, was designed to analyze the HIS data for accuracy and consistency. The two main groups of errors trapped include GETDAT errors, caused by problems in the structure or format of the data records; and CHECK errors, caused by incorrect or unexpected data (but in the expected format). Three levels of errors were defined: warnings (w), indicating either programming limitations or data problems which may be correctable with cursory checking; errors (e) which are more severe, usually indicating a problem requiring deeper investigation to solve; and fatals (f) which cause the data for that household to not be checked, and must be corrected for the data to be fully checked by the program. The specific errors HISCHK was programmed to catch are described below.

1. Messages referring to data formats (GETDAT)

- 9000 (f): record# rrrrr in household file has invalid hhnumber hhhhh
- 9001 (f): hh# hhhhh hhrecord# rrrrr hhid# changes within record string
- 9002 (f): recruit records not found for household# hhhhh
- 9003 (f): on recruit record# rrrrr hhid# changes within record string for household# hhhhh
- 9004 (e): hh# hhhhh cc cars rcrt, cc in hh
- 9005 (f): person records not found for household# hhhhh
- 9006 (f): hh# hhhhh per record# rrrrr hhid# changes within record string

- 9007 (f): on person rec rrrrr p# pp expected, qq found for hh# hhhhh
- 9008 (e): hhold# hhhhh has > 10 persons
- 9009 (e): hh# hhhhh pp pers rcrt, qq in per
- 9010 (f): activity records not found for household# hhhhh
- 9011 (f): on activity record# rrrrr hhid# changes within record string for household# hhhhh
- 9012 (f): on activity rec rrrrr a# aa expected, bb found for hh# hhhhh
- 9013 (e): hhold# hhhhh has > 50 activities

2. Messages referring to inconsistencies within one household's data (CHECK)

- 9000 (e): hpa# hhhhhppaa has travel time tttttt min. for activity _____
- 9001 (e): hpa# hhhhhppaa is a trip home with no travel info
- 9002 (e): hpa# hhhhhppaa has last endtime tttt purpose pp _____
- 9003 (e): hpa# hhhhhppaa activity _____ at location _____
- 9004 (e): hp # hhhhhpp stdnt claims schl on diary day, but no schl act
- 9005 (w): hp # hhhhhpp stdnt has no schl on diary day, but no reason
- 9006 (e): hp # hhhhhpp wrkr claims work on diary day, but no work act
- 9007 (w): hp # hhhhhpp wrkr has no work on diary day, but no reason
- 9008 (w): h # hhhhh diff #drop/pick @ sch: by drvrs mm, by stdnts nn

3. Messages referring to geocoding problems (GEOCOD)

- 9001 (e) : hpa# hhhhhppaa field type _____ not geocoded
  place : _____
  address : _____
  comm/zip : _____ zzzzz
  MapInfo code: ccccc

4. Messages summarizing status of a particular household

- 9901 (f): structural error in data -- file inspection needed nn
- 9901 (e): structural doubt in data -- file correction needed nn
- 9902 (e): logical inconsistencies -- call-back needed nn
- 9903 (e): geocoding problems -- call-back needed nn
- 9904 (e): invalid characters in data -- data correction needed nn
- 9905 (f): orphans in pers &/or acts -- file correction needed nn

Geocoding Procedures

All location data collected in the HIS was geocoded to latitude and longitude (X- and Y-coordinates). Geocoded data items include home, primary and secondary workplaces for workers, schools for students, activity locations, and bus boarding and alighting locations. Geocoding was performed using MapInfo™ and a commercial street database from GDT™. Household addresses were geocoded after recruitment, to ensure sufficient accuracy for delivery of the survey materials. Other locations were geocoded after retrieval of all data for a particular household, via the following steps.

1. All location data was passed through the automatic geocoding facility of MapInfo™.

2. For locations not successfully geocoded on pass 1, the interactive geocoding facility of MapInfo™ was employed. This generates a list of "best guesses", from which the user can select one.

3. Locations for which no "best guess" was deemed reasonable were given to a local firm (Mattson-Sunderland Research) for manual geocoding. This effort involved as little as correcting typographical information in the location record, or as much as examining local maps, telephone books and other resources to pinpoint the specific location reported.

4. Critical addresses (home, work, school) not geocoded by any of the above steps were referred back to interviewers for follow-up calls. Other addresses were held for clarification via call-back if some other more severe data problem warranted it.

5. Addresses not of sufficient quality to be resolved by the above steps, but with at least the community name where the activity took place, were geocoded to the Traffic Analysis Zone (TAZ) level by random allocation among all TAZs in that community. This process was not used for locations with the generic community name of "Honolulu".

The above steps succeeded in geocoding 100% of the household addresses, over 95% of workplaces and schools, and over 90% of activities, very few of which were geocoded directly to TAZ via step 5 above. The small number of activities with ungeocodable locations were not considered for trip-level analyses, as described under "Trip Expansion Factors" below.

Survey Expansion Procedure

To use the survey data for demographic and travel analyses, the sample's characteristics must reasonably match those of Oahu's population as a whole. The 1990 U.S. Census found 265,625 households on Oahu. Thus, to expand the sample to the most general characteristic of the population, its size, a factor of 265,625/4060=65.43 would be applied to each household. The same principle applies for matching on other characteristics, such as household size or area of residence, as long as data exist on the distribution of these characteristics within the population. In general, a household's expansion factor simply equals the inverse of the probability of its selection, but this depends on both sampling methods and response rates.

This effort followed a strategy of first developing factors based on available detailed data on the population. Sources included the 1990 Census and 1991 Transit On-Board Survey. An additional update factor then ensured that sampled households reflect known 1995 totals by geographic area, provided by the City and County of Honolulu Planning and Permitting Department. The various expansion factor components discussed below multiply to yield a final expansion factor for each household:

1990 Expansion Factor = Line Factor x Geographic Factor x Demographic Factor x Transit Factor

1995 Expansion Factor = 1990 Expansion Factor x Update Factor

<u>Line Factor</u>

This survey sampled telephone numbers to reach households. Thus, a household's probability of selection in the sample depends on the number of voice phone lines available to it. If a household has two voice lines, for example, its probability of selection is doubled and thus a

'line factor' of 0.5 is applied. Similarly, a household that shares a single voice line with another gets a line factor of 2.0, since the household that responded has a probability of selection half of that for a single-line household. In general, the line factor applied equals 1.0 divided by the number of voice phone lines to each household.

- 3680 households (90.6%) have either one or an unidentifiable number of voice phone lines per household, and thus receive a line factor of 1.0. Among these,
  - 3470 (85.5%) report one phone line not shared with another household;
  - 142 (3.5%) report one voice line and one or more data lines, not shared with another household;
  - 5 (0.1%) report two voice lines, shared with one household;
  - 63 (1.6%) have reporting errors, don't know or refuse to answer the questions.
- 343 more households (8.5%) have more than one voice line, not shared with another household. These receive line factors ranging from 0.17 to 0.5, depending on the number of voice line reported.
- 37 more households (0.9%) share their voice lines with other households, yielding line factors ranging from 1.5 to 5.0, depending on the numbers of voice lines reported and households they share with.

Geographic Factor

The Fall 1995 part of the HIS strategically oversampled households in areas with transit use by large, car-owning households, as revealed by tabulations of the households from Spring 1995. In addition, geographic response biases yielded under-representation of some areas. A set of geographic expansion factors corrected for these conditions. These factors expanded the number of sampled households in each of 69 Expansion Districts, specially developed for this purpose, to match the 1990 Census total housing units. The Expansion Districts are documented more fully below. Geographic factors range from 12.20 for Ewa nonCDP to 178.24 for Waikane CDP.

Demographic Factor

The next set of expansion factors reflects the island-wide joint distribution of households by their number of persons, vehicles owned, and 1989 household income, all as reported in the 1990 Census. A three-way cross-tabulation was developed from the 1990 Public Use Microdata Sample (PUMS), a Census product that provides detailed demographics on individual households at the expense of detailed geography. When divided by a matching cross-tab of sampled households (already expanded by the line and geographic factors), this cross-tabulation yielded demographic expansion factors. All households reported their size and number of vehicles, but 17% failed to report their household income. These were assigned a demographic factor equal to the average for their size/vehicles stratum. Demographic factors range from 0.09, for one-person households with no vehicles and 1989 income in the range $75-100,000, to 10.32, for seven-or-more-person households with three-or-more vehicles and 1989 income in the range $60-75,000. These imply that the former households had the highest response rates, after adjusting for geography, while the latter had the lowest response rates.

Transit Factor

To ensure that the sample reflected known levels of transit use on Oahu, the final set of 1990/1991 factors reflect total trips by trip purpose made by residents. The 1991 On-Board Survey reported 73,900 home-based work trips, 77,600 home-based non-work trips, and 21,500 non-home-based trips made by residents. Each HIS household reported a number of transit trips by each of these purposes, ranging from zero to twelve. An iterative method was applied to estimate a transit factor for each household, since a single household might have any number of transit trips for multiple purposes. This process yielded household-level factors ranging from 0.79 to 1.87, applied on top of the factors described above.

Second Geographic Factor and Closure Check

The process outlined was sequential, in that each factoring step depends on the results of previous steps. A second set of geographic factors ensured that demographic and transit factoring did not skew the geographic distribution of households. These factors range from 0.13 for Iroquois Point CDP to 12.28 for Ewa nonCDP. Tabulation of households by demographics and transit use revealed that application of the second geographic factors still allowed the sample to reflect known characteristics.

Update Factor

The process thus far ensured that the HIS households reflect characteristics of the population in general in 1990/1991. The update factors ensured that the sample matches the known 1995 geographic distribution of households. These factors depend on the number of housing units reported in 1990 and 1995 by the City & County of Honolulu, as well as an assumed constant and uniform occupancy rate for housing units. Update factors by Expansion District ranged from 0.93 for Iwilei D23 to 5.57 for Waipahu nonCDP. These imply that Iwilei D23 had the largest proportional loss in housing units (3943 to 3669) and Waipahu nonCDP the largest gain (495 to 2757) from 1990 to 1995.

Trip Expansion Factors

Expansion factors for activities or trips depended on the household factors, since detailed information on overall trip-making by residents was not available outside the HIS. However, 351 of the 4060 households collected cannot be used for trip-level analyses because of suspected under-reporting or mis-reporting of travel. The entire expansion process described above was redone for this remaining 3709 households to ensure this subsample matches the known population. In addition, some activity records did not yield usable trip information because of missing data on key questions, or ungeocodable location information. The trip factors ensured that all remaining trips match the geographic and demographic distributions reflected by the whole sample. The 32,749 trips reported by HIS households yielded 30,703 usable trip records, having trip factors ranging from 2.43 to 556.93, with 80% between 31.67 and 138.89, and only 1% greater than 323.16.

Expansion Districts

Development of geographic expansion factors required a district system at the level of detail for which geographic matching was desired. The island of Oahu contains roughly 200 Census Tracts, averaging 1400 total households and 20 sampled households, too few for reliable

factoring. The 23 districts used for summarizing analytic results are too large to appropriately target underrepresented areas. A desirable system would yield somewhere between 50 and 150 districts, possibly differing in size depending on the density of population in different areas. A system of 69 Expansion Districts was developed to expand the 1995 Home Interview Survey (HIS), using the following steps:

1. The 1990 U.S. Census used 48 'Census Designated Places' (CDPs) on the island of Oahu. Each of these CDPs except Honolulu comprised a single Expansion District. The size of Honolulu CDP warranted further subdivision, as described below. Consequently, 47 of the total 69 Expansion Districts, listed in the following table as '*name* CDP', were identical to the CDPs found outside the Honolulu CDP.

2. Work supporting the Honolulu Rapid Transit Program developed a system of 23 summary districts for reporting results. Twelve of these districts were entirely within the Honolulu CDP. Ten of these were used as individual Expansion Districts. The other two, Kakaako and Ala Moana, were combined to form Expansion District 'Kakaako D23'. These eleven Expansion Districts are listed in Table 1 as '*name* D23'.

3. Of the other eleven summary districts, nine were entirely outside the Honolulu CDP. The area in each of these NOT covered by any CDP comprised a separate Expansion District. These nine are listed in Table 1 as '*name* D23: nonCDP'.

4. Two of the 23 summary districts straddled the border of Honolulu CDP. Within each, the parts in the Honolulu CDP and not in any CDP have been combined to form Expansion Districts. These two are listed in Table 1 as '*name* D23: HonCDP+nonCDP'.

5. Housing unit totals for the 69 Expansion Districts were developed by aggregating from 1990 U.S. Census Blocks. Total housing units by district ranged from 47 in Aiea nonCDP to 20,566 in Kalihi D23, with a total of 281,374 housing units reported on Oahu in the 1990 Census. The attached map ("HIS Expansion Districts with Sequential Numbers") shows the 69 Expansion Districts.

Table 1: Expansion Districts

Expansion District Name	Seq Num	Source Num	HU 1990	Geog Fac 1	Dem Fac	Tran Fac	Geog Fac 2	Upd Fac
Downtown D23	1	1	6848	86.89	.80	.98	1.28	1.14
Kakaako D23	2	2	2082	83.48	.86	.97	1.20	1.95
Beretania D23	3	4	7067	108.97	.83	1.04	1.16	1.01
Makiki D23	4	5	17525	78.07	.80	1.01	1.24	1.01
Waikiki D23	5	6	18901	125.46	.74	.98	1.39	1.09
McCully D23	6	7	15779	81.53	.91	1.04	1.07	1.01
UH/Manoa D23	7	8	4852	52.86	1.04	.99	.97	1.06
Kaimuki D23	8	9	18478	67.11	.92	.99	1.10	1.00
Iwilei D23	9	10	3943	87.82	.95	1.17	.90	.93
Kalihi D23	10	11	20566	49.08	1.07	1.05	.89	1.01
Airport D23: Hon/nonCDP	11	12	3985	46.99	1.03	.96	1.18	1.02
Salt Lake D23:Hon/nonCDP	12	13	11012	86.91	.89	.96	1.20	1.03
Aiea D23: nonCDP	13	14	47	53.00	.85	.95	1.14	1.00
Waipahu D23: nonCDP	14	15	519	18.94	1.09	1.00	.81	5.57
Mililani D23: nonCDP	15	16	414	13.03	.91	.96	4.29	2.86
Ewa D23: nonCDP	16	17	1237	29.17	1.04	.95	2.14	2.73
Waianae D23: nonCDP	17	18	1025	74.71	.84	.92	1.77	1.09
North Shore D23: non	18	19	930	32.97	1.13	1.00	.95	1.68
Koolauloa D23: nonCD	19	20	351	52.57	.80	.93	1.67	.98
Kaneohe D23: nonCDP	20	21	768	134.67	.77	.95	1.26	1.04
Kailua D23: nonCDP	21	22	418	110.00	.61	1.20	1.82	1.09
East Honolulu D23	22	23	15644	63.48	.86	.97	1.18	1.04
Ahuimanu CDP	23	400	2537	58.00	1.00	.96	.88	1.05
Aiea CDP	24	550	2678	79.76	.94	.93	1.06	1.03
Aliamanu CDP	25	1925	2353	78.60	1.13	.95	.94	1.00
Barbers Point Housing CDP	26	2500	866	98.00	.88	.95	.15	1.09
Ewa Beach CDP	27	7450	3426	96.93	1.09	1.04	.82	1.00
Ewa Gentry CDP	28	7470	752	70.92	.78	1.01	1.18	2.21
Ewa Villages CDP	29	7485	939	106.30	1.99	1.29	.36	1.14
Fort Shafter CDP	30	8350	917	41.77	1.04	.91	1.06	1.00
Halawa CDP	31	10150	4094	102.96	1.08	1.01	.84	1.02
Haleiwa CDP	32	10750	866	98.89	1.04	.90	1.05	1.02
Hauula CDP	33	12400	1021	107.10	1.49	.95	.89	1.01
Heeia CDP	34	13900	1557	74.41	.85	.94	1.02	1.01
Hickam Housing CDP	35	14200	1841	63.15	1.13	.92	.50	1.05
Iroquois Point CDP	36	19100	1180	39.26	1.09	.96	.13	1.00
Kaaawa CDP	37	19550	452	79.00	.93	.91	1.48	1.05
Kahaluu CDP	38	21200	960	56.06	1.16	.91	.81	1.02
Kahuku CDP	39	22250	586	87.86	1.56	1.00	.79	.96
Kailua CDP	40	23150	12225	52.90	.91	1.01	.93	1.03
Kaneohe CDP	41	28250	10849	77.08	.96	.97	.91	1.07
Kaneohe Station CDP	42	28425	2030	47.44	1.03	.93	.82	1.16

Table 1: Expansion Districts

Expansion District Name	Seq Num	Source Num	HU 1990	Geog Fac 1	Dem Fac	Tran Fac	Geog Fac 2	Upd Fac
Kawela Bay CDP	43	32700	418	73.17	.80	.95	1.65	.99
Laie CDP	44	43250	1122	53.50	1.10	.96	1.20	1.01
Maili CDP	45	47300	1490	101.33	.92	.94	1.22	1.09
Makaha CDP	46	47450	3178	104.58	1.02	1.01	.91	.98
Makaha Valley CDP	47	47470	604	123.20	.60	1.15	1.49	1.09
Makakilo City CDP	48	47750	3050	80.28	.85	.96	1.12	1.29
Maunawili CDP	49	50750	1443	44.62	1.08	.98	.70	1.00
Mililani Town CDP	50	51050	8900	69.47	1.01	.99	.92	1.02
Mokuleia CDP	51	52550	798	58.57	.97	.96	1.12	1.04
Nanakuli CDP	52	53900	2128	144.73	1.04	.93	1.06	1.01
Pearl City CDP	53	62600	8999	78.95	1.02	.98	.92	1.00
Punaluu CDP	54	66500	466	163.00	.61	.95	2.25	1.10
Pupukea CDP	55	66800	1488	67.87	.88	.94	1.52	1.05
Schofield Barracks CDP	56	69050	3556	65.83	1.05	.96	.71	1.04
Village Park CDP	57	72255	2176	94.73	1.04	.94	.95	1.19
Wahiawa CDP	58	72650	5765	79.71	1.07	1.00	.84	1.07
Waialua CDP	59	74000	1205	77.44	1.07	1.00	.99	1.00
Waianae CDP	60	74450	2264	82.50	1.14	.94	.94	1.26
Waikane CDP	61	75800	189	199.00	.77	.95	1.16	1.02
Waimalu CDP	62	77750	10613	94.57	.87	1.01	1.05	1.03
Waimanalo CDP	63	78050	832	67.31	1.22	1.19	.59	1.08
Waimanalo Beach CDP	64	78200	965	72.50	.98	1.12	.72	1.08
Waipahu CDP	65	79700	7739	121.65	1.10	1.05	.80	1.04
Waipio CDP	66	79860	4087	92.52	.91	1.00	1.01	1.01
Waipio Acres CDP	67	80000	1836	103.90	.89	.97	1.11	1.07
Wheeler AFB CDP	68	80600	704	40.47	1.08	.91	.23	1.00
Whitmore Village CDP	69	80900	839	57.25	1.22	1.21	.61	1.19

Key: Seq Num = expansion district sequential number; Source Num = summary district of CDP from which the expansion district is derived; HU 1990 = total housing units from 1990 Census; Geog Fac 1 = first geographic factor; Dem Fac = average demographic factor; Tran Fac = average transit factor; Geog Fac 2 = second geographic factor; Upd Fac = 1990 to 1995 update factor.

HIS Expansion Districts

The Data Library CD contains the data from the 1995 HIS, as well as a copy of the survey instrument. The data files contain household (HLD), person (PER), activity (ACT), vehicle (VEH), and point geocoding (GEO) information. Within each name, the extension indicates the type of information in that file: SAV is the data in SPSS format, DAT is the data in ASCII format, and INF is a data dictionary listing the variables along with the values, definitions, and locations. Lastly, the SPSS syntax file GEOMRG.SPS shows how to use SPSS to merge the point geocoding information, should that ever be necessary. 761-TAZ geocoding is included with each data file.

File	Contents
HIS_FORM.PDF	Survey instrument
HLD.DAT	Household data in ASCII format
HLD.SAV	Household data in SPSS format
HLD.INF	Description of data set
PER.DAT	Person data in ASCII format
PER.SAV	Person data in SPSS format
PER.INF	Description of data set
ACT.DAT	Activity data set in ASCII format
ACT.SAV	Activity data set in SPSS format
ACT.INF	Description of data set
VEH.DAT	Vehicle data in ASCII format
VEH.SAV	Vehicle data in SPSS format
VEH.INF	Description of data set
GEO.DAT	Geocoding information in ASCII format
GEO.SAV	Geocoding information in SPSS format
GEO.INF	Description of data set
GEOMERG.SPS	Instructions on how to merge the point geocoding information