Health of Houston Survey 2010

HEALTH OF HOUSTON SURVEY 2010

Methodology Report

Institute for Health Policy
The University of Texas School of Public Health

This report was prepared for the Health of Houston Survey by David Dutwin and Susan Sherr of SSRS.

SUGGESTED CITATION:

Health of Houston Survey. HHS 20010. Methodology Report. Houston, TX: Institute for Health Policy, 2011.

FUNDING:

The Health of Houston Survey 2010 was funded by the Houston Endowment, Inc. with outreach support from The University of Texas Health Science Center at Houston's (UTHealth) Center for Clinical and Translational Science.

TABLE OF CONTENTS

1. HHS 2010 DESIGN AND METHODOLOGY SUMMARY

1.1 Overview

The Health of Houston Survey 2010 (HHS 2010) is an address-based (AB) survey of Houston's population. HHS 2010 is based at the University of Texas Health Science Center at Houston (UTHealth) Institute for Health Policy (IHP). HHS 2010 collects extensive information for multiple segments of the population on health status, conditions, behaviors, insurance coverage, and access.

The study was designed to capture reliable data for a number of populations:

- Each of ACS 7 Super Public Use Microdata Areas (SuperPUMAs) in Harris County
- Whites, African Americans, Hispanics, Vietnamese, and Other Asians
- A standard range of age and income cohorts
- The total population of Harris County and the City of Houston

The HHS 2010 sample is representative of Harris County and the City of Houston's non-institutionalized population living in households.

1.2 Sample Design Objectives

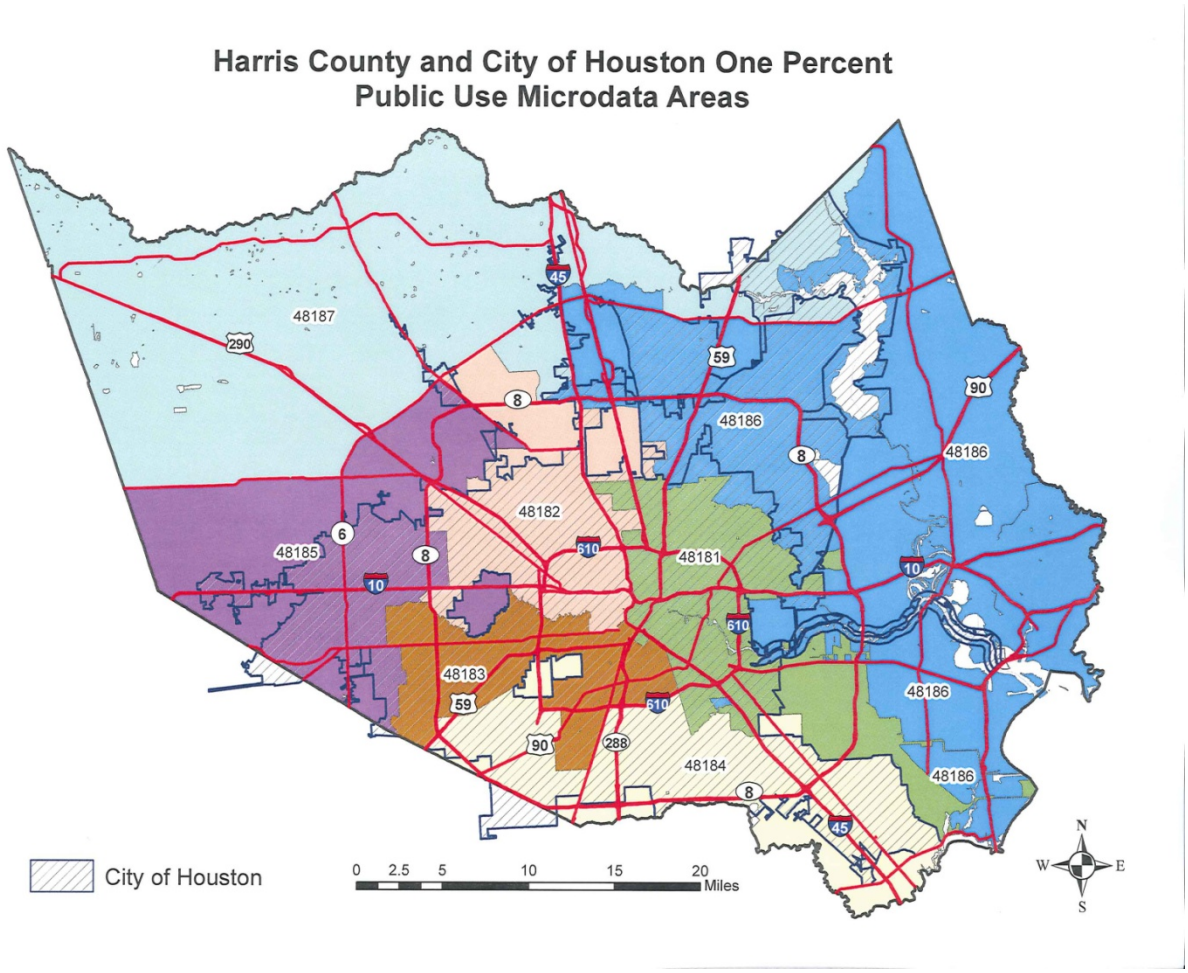
To achieve the sample design parameters stated above, HHS employed a multi-dimensional sample design. Specifically, the design stratified by both SuperPUMA and by concentration of ethnic populations by both household density and ethnic status of residents' surname. This resulted in 45 strata in a 7 x 7 design:

TABLE 1: SAMPLE STRATIFICATION

SuperPUMA	Strata	SuperPUMA	Strata
48181	Residual	48184	Residual
48181	Black High	48184	Black High
48181	Hispanic High	48184	Hispanic High
48181	Vietnamese High	48184	Asian High
48181	Asian Surname	48184	Vietnamese High
48181	Vietnamese Surname	48184	Asian Surname
48182	Residual	48184	Vietnamese Surname
48182	Black High	48185	Residual
48182	Hispanic High	48185	Black High
48182	Asian High	48185	Hispanic High
48182	Vietnamese High	48185	Asian High
48182	Asian Surname	48185	Vietnamese High
48182	Vietnamese Surname	48185	Asian Surname
48183	Residual	48185	Vietnamese Surname
48183	Black High	48186	Residual
48183	Hispanic High	48186	Black High
48183	Asian High	48186	Hispanic High
48183	Vietnamese High	48186	Asian Surname
48183	Asian Surname	48186	Vietnamese Surname
48183	Vietnamese Surname	48187	Residual
		48187	Black High
		48187	Hispanic High
		48187	Asian High
		48187	Asian Surname
		48187	Vietnamese Surname

The original design allowed for the attainment of approximately 575 interviews per SuperPUMA, while in aggregate, attaining a minimum of 200 interviews of Vietnamese; 250 other Asians; 700 African Americans, and 1,000 Hispanics, with 4,000 interviews overall. Based on a mid-project assessment, the target was increased to 4,200 overall interviews, of which 3,600 would come from web or telephone. The targets for SuperPUMA and race/ethnicity were proportionately increased for the new design as well.

The study relied on an address-based design. Because of the increase in cell phone only (CPO) households, researchers are faced with increasing challenges in terms of being able to cover an entire population. Over 25 percent of households are now, nationwide, without landline telephone service. Another 8 percent, it is believed, are part of "zero-bank" households, and most importantly, there are likely significant numbers of CPO households in Houston that have area codes outside of the Houston area. An address-based design circumvents these difficulties, given than the sample source is the U.S. Postal Service's Delivery Sequence File (DSF), a database that is considered to cover at least 98 percent of all households in the U.S., a number that is likely higher for an urban area like the city of Houston.



Note in the SuperPUMA map there are three areas of the city of Houston that fall outside of the 7 SuperPUMA targeted for the study. These areas were included in the sample and subsumed under the SuperPUMA most proximate geographically.

1.3 Data Collection

Because the sample is address-based, data collection methods differ from traditional telephone samples. The HHS 2010 study executed a data collection strategy designed to attain the highest response rate possible. This design combines telephone (CATI), web, and mail survey options, all offered in three languages.

Surveys were conducted in English, Spanish, and Vietnamese. These languages were chosen given our population of interest. Additional Asian languages were excluded due to generally low linguistic isolation rates and due to the complexity of administering an address-based design in a wide range of languages.

Further details on data collection are provided in the data collection section later in this report.

1.4 Response Rates

The overall response rate for HHS 2010 is a composite of the screener completion rate (i.e., success in introducing the survey to a household and randomly selecting an adult to be interviewed) and the extended interview completion rate (i.e., success in getting the selected person to complete the extended interview).

To maximize the response rate, especially at the screener stage, an invitation letter in three languages was mailed to all sampled addresses. A $2 bill was included with the invitation letter to promote cooperation. As well, the unmatched sample (sample for which a telephone number could not be identified) was offered a $20 incentive upon completion of the survey. Respondents were offered a chance to participate in a random drawing for a $200 VISA gift card.

TABLE 2: SURVEY TOPICS

Topics	Randomly Selected Adult in Household	Randomly Selected Child in Household
Demographics I (Age, Gender, Race/Ethnicity)	Adults	Child
General Health Status	Adults	Child
Health conditions (Obesity, Diabetes, Asthma, Cancer, Cardiovascular Disease, Hypertension)	Adults	
Health Conditions (Obesity, Physical, behavioral or mental conditions)		Child
Health and Dental Insurance Status	Adults	Child
Health and Dental Care Access	Adults	Child
Mental Health Assessment	Adults	
Mental Health Access and Utilization	Adults	
Mammography	Females Age 40-74	
Pap Test	Females	
Colorectal Cancer	Adults Age 50-76	
Behavioral Risk Factors I (Smoking, Second Hand Smoke, Alcohol Abuse)	Adults	
Prenatal Care/Breastfeeding	Females Age 18-50	
Employment	Adults	
Income	Adults	
Economic Hardship	Income <150,000K	
Public Programs (Food Stamps, Supplemental Security Income, Social Security/Pensions, WIC, Child Support)	Adults	
Behavioral Risk Factors II (Diet, Physical Activity)	Adults	Child
Sedentary Behavior		Child
Neighborhood, Environment & Housing	Adults	
Transportation	Adults	
Social Cohesion	Adults	
Environmental Risks	Adults	
Interpersonal Violence	Adults	
Demographics II (Country of Origin, Languages Spoken at Home, Citizenship)	Adults	
Household Phone Status	Adults	
Sexual Identity/Orientation	Adults	
Social Support	Adults	

1.5 Weighting the Sample

Survey data are weighted to adjust for differential sampling probabilities, to reduce any biases that may arise because of differences between respondents and non-respondents (i.e., nonresponse bias), and to address gaps in coverage in the survey frame (i.e., coverage bias). Survey weights, when properly applied in surveys can reduce the effect of nonresponse and coverage gaps on the reliability of the survey results (Keeter et al. 2000, Groves 2006). Details are provided in the section regarding weighting.

2. SAMPLING METHODS

2.1 Overview

Historically, Random Digit Dialing (RDD) telephone interviewing has been the method of choice for many survey data collection efforts given the strength of its randomization method, ease of administering complex questionnaires using computerized interviewing systems, excellent coverage of the overall population (given that less than 2% of Americans live in a household without telephone service), and relatively low cost. Survey coverage refers to the extent to which the sample frame for a survey includes all members of the target population. A survey design with a gap in coverage raises the possibility of bias if the individuals missing from the sample frame (e.g., households without landline telephones) differ from those in the sample frame. Unfortunately, the coverage of the overall population in RDD surveys is changing as more and more households are relying on cell phones and giving up their landline telephones. Cell phone numbers are typically not called in RDD surveys.

Cell phone-only households are increasing rapidly in the United States, with 24.9% of households estimated to be cell phone-only in the first half of 2010, as compared to 20.2% in 2008 (Blumberg & Luke, 2011). While there is limited data available on the share of cell phone-only households within each state, a recent model-based approach (combining survey data and synthetic estimates) was used to generate state-level estimates of cell phone-only households using the National Health Interview Survey (NHIS). Based on that work, an estimated 20.1% of households in Harris County were cell phone-only in 2007, a figure that was revised to be 32.4% in 2010.

In order to capture cell phone-only households in the sample frame for the HHS 2010, the decision was made to utilize an address-based sample (AB sample) for the survey. The AB sample captures households with landline phones, cell phone-only households and non-telephone households. One limitation of both AB sample and RDD sample is that they both miss homeless persons which are estimated to be between 10,000 and 15,000 persons based on HUD estimates.

The AB sample was developed in the following steps:

1. A file was generated of all Harris County and City of Houston residential addresses currently in use based on the United States Postal Service Delivery Sequence File (DSF). The DSF is a computerized file that contains information on all delivery addresses serviced by the USPS, with the exception of general delivery.[1] The DSF is updated weekly and contains home and apartment addresses as well as Post Office boxes and other types of residential addresses for mail delivery.

2. That address file was run against databases from InfoUSA, Experian, TargusInfo, and Acxiom that include all listed landline telephone numbers in the state to identify addresses with a listed telephone number.

In order to facilitate the fielding of the survey, the final AB sample was divided into two segments: addresses with a listed landline telephone number and addresses without a listed landline telephone number.

[1] See http://pe.usps.gov/text/dmm300/509.htm.

The overall sampling design contained a number of features across several dimensions that can be described in terms of sample stratification, household selection criteria, and within household selection criteria. These are summarized below and then furnished in more detail later in this section.

1) Sample stratification

- Set interview targets per Super Public Use Microdata Areas (SuperPUMAs).
- Within SuperPUMA creation of strata of addresses by listed Vietnamese and Asian-non-Vietnamese surnames.
- Stratification of residual (households without an Asian surname) households by Census block group aggregate incidence of Hispanic, percent African American, and percent Asian.

2) Household-level selection

- Screening households with respondents under 18 years of age.
  - If the person on the phone is younger than 18, interviewer asks for another household member who is 18 or older.
  - If there is no household member 18 or older, the household is not eligible, and the interview is terminated.
- Screening households where every adult was age 65 and older.
  - If the household contained only adults ages 65 and older, the interview was terminated in 33 percent of such instances. That was designed to balance for the fact that such households more readily respond to surveys compared to other households.

3) Individual-level selection

- Respondent is randomly selected from all household members using the "Rizzo" method[2] of selection.
  - First, the number of people in household is determined.
  - If it is a single-person household, that person is the respondent.
  - If it is a two-person household, one of those two people is randomly selected to be the respondent.
  - If it is a three or more-person household, a random selection of household members is performed by the Web/CATI program. If the current respondent is selected, he or she is the respondent. If another household member is selected, we asked for the household member, other than the current respondent, with the most recent birthday.

2.2 Sample Stratification

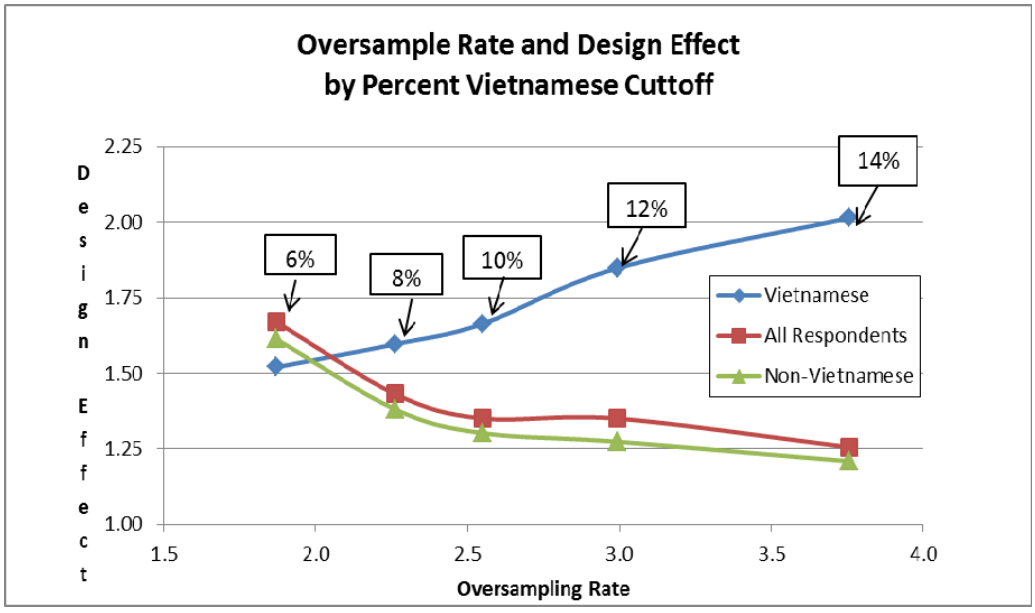
The number of interviews by SuperPUMA was set to ensure adequate statistical power within each. Stratification by racial/ethnic surname and aggregate Census block group incidence of minority households was done to maximize the number of interviews of African Americans, Asians, and specifically Vietnamese, while maintaining an acceptable number of interviews of both Caucasians and Hispanics.

[2] Rizzo, L.J., Brick, J.M., and Park, I. (2004). A minimally intrusive method for sampling persons in radon digit dial surveys. Public Opinion Quarterly, 68, 267-274.

Census block groups were defined as being high Hispanic if 50 percent or more households were Hispanic; high African American if 50 percent or more African American; high Vietnamese if 10 percent or more Vietnamese, and high Asian if Asian-non-Vietnamese incidence was 15 percent or higher.

Since Vietnamese was the most critical group, as well as the group that required the most aggressive oversampling strategies to meet interviewing targets, we analyzed whether the definition of Vietnamese high would be most effective if defined at 6, 8, 10, 12, or 14 percent Vietnamese.

As shown in the graph below, we decided upon 10 percent as the optimal cut point in terms of keeping the design effect within Vietnamese households to a minimum while also keeping the overall design effect low (1.68):



In addition, the 10 percent cut point provides a respectable number of Census block groups to work with (34 of 1,947) whereas the 12 percent cut point only contained 25 and the 14 percent cut point only 14 Census block groups. While the 6 percent and 8 percent cut points held 49 and 87 block groups respectively, these cut points would have pushed the overall design effect for the stratification over 1.3 which was not deemed to be optimal.

Below are the final strata used for the survey:

TABLE 3: STRATIFICATION PLAN – HOUSEHOLD RACE*

Super PUMA	Strata	Total Households	White Households	Black Households	Asian Households	Other Households	Hispanic Households	Vietnamese Households
48181	Residual	28,891	16,928	2,245	256	9,021	8,491	38
48181	Black High	30,542	850	24,453	10	5,068	4,778	10
48181	Hispanic High	111,711	16,220	8,125	10	86,456	85,576	31
48181	Vietnamese High	1,174	158	24	10	199	188	646
48181	Asian Surname	972	97	24	583	24	49	194
48181	Vietnamese Surname	875	88	22	175	22	44	525
	TOTAL	**174,165**	**34,340**	**34,894**	**1,044**	**100,791**	**99,125**	**1,444**
48182	Residual	95,965	54,225	10,557	1,675	27,025	25,294	429
48182	Black High	25,957	2,159	18,394	8	4,970	4,731	75
48182	Hispanic High	70,294	13,463	6,216	224	49,032	48,317	221
48182	Asian High	2,305	1,031	379	242	457	377	5
48182	Vietnamese High	4,286	748	1,437	61	1,316	1,250	359
48182	Asian Surname	1,696	263	66	1,576	66	131	525
48182	Vietnamese Surname	2,627	170	42	339	42	85	1,018
	TOTAL	**203,130**	**72,058**	**37,091**	**4,125**	**82,908**	**80,184**	**2,632**
48183	Residual	144,198	87,011	18,489	6,653	28,368	24,263	850
48183	Black High	32,608	1,784	25,061	335	4,916	4,497	10
48183	Hispanic High	29,164	2,965	3,215	718	21,619	21,130	137
48183	Asian High	34,727	8,889	8,167	6,259	9,381	8,055	692
48183	Vietnamese High	14,908	2,604	2,860	1,597	4,996	4,625	1,180
48183	Asian Surname	3,652	359	90	2,156	90	180	719
48183	Vietnamese Surname	3,593	365	91	730	91	183	2,191
	TOTAL	**262,850**	**103,978**	**57,972**	**18,449**	**69,460**	**62,932**	**5,779**
48184	Residual	94,538	56,631	11,198	2,100	20,945	19,019	484
48184	Black High	63,458	3,805	49,524	255	9,053	9,268	10
48184	Hispanic High	36,228	5,775	5,822	130	23,948	23,543	150
48184	Asian High	11,493	6,777	998	1,652	1,346	979	167
48184	Vietnamese High	3,318	811	380	68	1,485	1,401	277
48184	Asian Surname	2,421	312	78	1,874	78	156	625
48184	Vietnamese Surname	3,123	242	61	484	61	121	1,453
	TOTAL	**214,579**	**74,353**	**68,062**	**6,564**	**56,916**	**54,487**	**3,165**
48185	Residual	155,153	86,269	19,566	5,601	37,509	34,462	21
48185	Black High	1,353	183	768	32	289	270	7
48185	Hispanic High	13,754	3,402	1,442	167	8,089	7,870	10
48185	Asian High	14,741	6,340	2,595	1,459	2,730	2,825	163
48185	Vietnamese High	6,365	1,844	1,369	480	1,233	1,063	191
48185	Asian Surname	3,512	665	166	3,990	166	333	1,330
48185	Vietnamese Surname	6,651	351	88	702	88	176	2,107
	TOTAL	**201,528**	**99,054**	**25,995**	**12,432**	**50,104**	**46,998**	**3,829**
48186	Residual	123,976	75,179	15,713	955	30,636	28,136	57
48186	Black High	22,220	1,884	14,560	71	5,484	5,258	5
48186	Hispanic High	32,797	6,512	3,314	28	22,732	22,347	10
48186	Asian Surname	1,105	88	22	527	22	44	176

Super PUMA	Strata	Total Households	White Households	Black Households	Asian Households	Other Households	Hispanic Households	Vietnamese Households
48186	Vietnamese Surname	878	111	28	221	28	55	663
	TOTAL	**180,976**	**83,774**	**33,637**	**1,801**	**58,902**	**55,840**	**911**
48187	Residual	199,169	142,872	16,612	4,050	31,315	27,962	501
48187	Black High	6,177	813	3,960	95	1,158	1,048	8
48187	Hispanic High	866	178	133	42	484	460	10
48187	Asian High	2,441	1,045	511	187	412	334	70
48187	Asian Surname	2,193	303	76	1,819	76	152	606
48187	Vietnamese Surname	3,032	219	55	439	55	110	1,316
	TOTAL	**213,878**	**145,430**	**21,347**	**6,631**	**33,501**	**30,065**	**2,510**
	GRAND TOTAL	**1,451,106**	**612,988**	**278,997**	**51,047**	**452,582**	**429,632**	**20,271**

** Household counts based on Claritas 2010.*

The sampling plan is as follows:

TABLE 4: STRATIFICATION PLAN – EXPECTED INTERVIEWS BY RACE AND POVERTY STATUS

Super PUMA	Strata	Percent of Households	Allocation of Interviews	Expected Total Interviews	Expected Black Interviews	Expected Asian Interviews	Expected Hispanic Interviews	Expected Vietnamese Interviews	Expected Other Interviews	Expected Below Poverty Interviews
48181	Residual	16.6%	13%	78	6	1	23	0.10	48	8
48181	Black High	17.5%	15%	90	72	0	14	0.03	4	35
48181	Hispanic High	64.1%	52%	312	23	0	239	0.09	50	102
48181	Vietnamese High	0.7%	8%	48	1	0	8	26	13	12
48181	Asian Surname	0.6%	5%	30	1	18	2	6	4	9
48181	Vietnamese Surname	0.5%	7%	42	1	8	2	25	5	12
	TOTAL	**100%**	**100%**	600	104	28	287	58	124	178
48182	Residual	47.2%	32%	192	21	3	51	1	116	17
48182	Black High	12.8%	13%	78	55	0	14	0	8	22
48182	Hispanic High	34.6%	30%	180	16	1	124	1	39	51
48182	Asian High	1.1%	5%	30	5	3	5	0	17	4
48182	Vietnamese High	2.1%	9%	54	18	1	16	5	15	8
48182	Asian Surname	0.8%	5%	30	1	28	2	9	-11	5
48182	Vietnamese Surname	1.3%	6%	36	1	5	1	14	16	6
	TOTAL	**100%**	**100%**	600	117	40	213	29	200	113
48183	Residual	54.9%	28%	168	22	8	28	1	109	11
48183	Black High	12.4%	10%	60	46	1	8	0	5	16
48183	Hispanic High	11.1%	10%	60	7	1	43	0	8	22
48183	Asian High	13.2%	20%	120	28	22	28	2	40	15
48183	Vietnamese High	5.7%	20%	120	23	13	37	10	37	23
48183	Asian Surname	1.4%	5%	30	1	18	1	6	4	6
48183	Vietnamese Surname	1.4%	7%	42	1	9	2	26	5	9
	TOTAL	**100%**	**100%**	600	127	71	149	45	209	103
48184	Residual	44.1%	24%	144	17	3	29	1	94	9
48184	Black High	29.6%	29%	174	136	1	25	0	12	41
48184	Hispanic High	16.9%	17%	102	16	0	66	0	19	26
48184	Asian High	5.4%	12%	72	6	10	6	1	48	4
48184	Vietnamese High	1.5%	6%	36	4	1	15	3	13	6
48184	Asian Surname	1.1%	5%	30	1	23	2	8	-4	6
48184	Vietnamese Surname	1.5%	7%	42	1	7	2	20	14	8
	TOTAL	**100%**	**100%**	600	181	45	146	33	195	101
48185	Residual	77.0%	32%	192	24	7	43	0	118	14
48185	Black High	0.7%	3%	18	10	0	4	0	4	2
48185	Hispanic High	6.8%	7%	42	4	1	24	0	13	8
48185	Asian High	7.3%	15%	90	16	9	17	1	47	7

Super PUMA	Strata	Percent of Households	Allocation of Interviews	Expected Total Interviews	Expected Black Interviews	Expected Asian Interviews	Expected Hispanic Interviews	Expected Vietnamese Interviews	Expected Other Interviews	Expected Below Poverty Interviews
48185	Vietnamese High	3.2%	18%	108	23	8	18	3	55	6
48185	Asian Surname	1.7%	7%	42	2	48	4	16	-28	3
48185	Vietnamese Surname	3.3%	18%	108	1	11	3	34	58	7
	TOTAL	**100%**	**100%**	600	81	84	112	55	268	47
48186	Residual	68.5%	57%	342	43	3	78	0	218	36
48186	Black High	12.3%	14%	84	55	0	20	0	9	26
48186	Hispanic High	18.1%	18%	108	11	0	74	0	23	33
48186	Asian Surname	0.6%	5%	30	1	14	1	5	9	9
48186	Vietnamese Surname	0.5%	6%	36	1	9	2	27	-4	11
	TOTAL	**100%**	**100%**	600	111	26	175	32	256	116
48187	Residual	93.1%	77%	462	39	9	65	1	348	22
48187	Black High	2.9%	4%	24	15	0	4	0	4	2
48187	Hispanic High	0.4%	2%	12	2	1	6	0	3	3
48187	Asian High	1.1%	5%	30	6	2	4	1	16	2
48187	Asian Surname	1.0%	5%	30	1	25	2	8	-6	3
48187	Vietnamese Surname	1.4%	7%	42	1	6	2	18	15	4
	TOTAL	**100%**	**100%**	600	64	44	83	29	381	36
	GRAND TOTAL			4,200	786	338	1,164	280	1,633	695
	Adjusted for Non-Response			**4200**	**762**	**265**	**1086**	**198**	**1889**	**666**

As the above table illustrates (given differences between the percent of households to the allocation on interviews), Asian, African American, and Vietnamese strata are oversampled significantly, while Hispanic areas on average are proportionate and Residual strata are under-sampled. The overall goal, as expressed by IHP, was to attain as close to 200 Vietnamese interviews as possible, as well as an additional 250 Asian interviews.

While the design "on paper" should attain 280 Vietnamese and 338 Asian interviews, in fact, we know from prior survey research that certain ethnic and racial populations tend to attain higher nonresponse than others. The adjusted-for-nonresponse figures are what we expected to attain "on the ground."

IHP was also concerned with projecting the number of interviews by poverty status and by age. As shown in the above table, we expect about 666 interviews of persons under poverty, which is slightly lower (16.5 percent) than the actual rate of poverty in Harris County (17.8% based on the 2009 American Community Survey). Percents by age are provided below. While there was concern over the number of interviews of persons ages 65 and older, the counts broke at 60 and older as provided in the Table 5.

TABLE 5: STRATIFICATION PLAN – EXPECTED INTERVIEWS BY AGE AND HOUSEHOLDS WITH CHILDREN

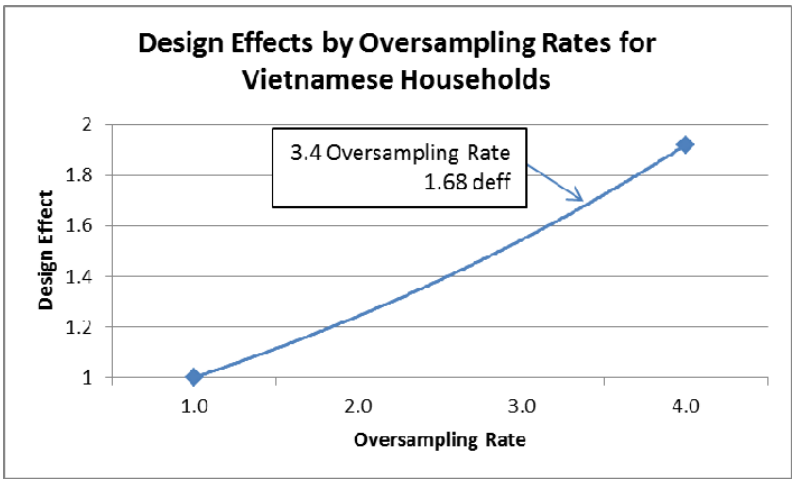
Super PUMA	Strata	Percent				Interviews			
		18-34	35-59	60+	With Kids	18-34	35-59	60+	With Kids
48181	Residual	35%	46%	18%	39%	28	36	14	30
48181	Black High	32%	42%	26%	40%	29	38	23	36
48181	Hispanic High	39%	45%	16%	52%	122	140	50	162
48181	Vietnamese High	31%	47%	21%	34%	15	23	10	16
48181	Asian Surname	37%	45%	18%	40%	11	13	5	12
48181	Vietnamese Surname	36%	45%	19%	40%	15	19	8	17
	TOTAL					220	269	111	274
48182	Residual	29%	49%	22%	28%	56	95	42	54
48182	Black High	35%	46%	19%	46%	27	36	15	36
48182	Hispanic High	39%	46%	15%	49%	70	83	27	89
48182	Asian High	30%	52%	18%	39%	9	16	6	12
48182	Vietnamese High	33%	53%	14%	47%	18	28	8	25
48182	Asian Surname	34%	48%	18%	43%	10	14	5	13
48182	Vietnamese Surname	34%	48%	18%	51%	12	17	6	18
	TOTAL					147	195	66	246
48183	Residual	34%	47%	19%	19%	57	79	33	32
48183	Black High	38%	42%	20%	37%	23	25	12	22
48183	Hispanic High	45%	45%	10%	47%	27	27	6	28
48183	Asian High	42%	46%	13%	27%	50	55	15	33
48183	Vietnamese High	33%	47%	20%	38%	40	56	24	46
48183	Asian Surname	37%	46%	17%	29%	11	14	5	9
48183	Vietnamese Surname	36%	46%	18%	43%	15	19	7	18
	TOTAL					166	196	69	188
48184	Residual	29%	50%	21%	36%	42	72	30	52
48184	Black High	34%	46%	20%	49%	60	80	34	86
48184	Hispanic High	39%	46%	14%	51%	40	47	15	52
48184	Asian High	28%	56%	16%	48%	20	40	12	34
48184	Vietnamese High	36%	50%	14%	48%	13	18	5	17
48184	Asian Surname	31%	49%	19%	50%	9	15	6	15
48184	Vietnamese Surname	32%	50%	19%	53%	13	21	8	22
	TOTAL					156	221	79	278
48185	Residual	31%	53%	16%	50%	59	102	31	95

Super PUMA	Strata	Percent				Interviews			
		18-34	35-59	60+	With Kids	18-34	35-59	60+	With Kids
48185	Black High	34%	51%	15%	43%	6	9	3	8
48185	Hispanic High	36%	51%	13%	52%	15	21	5	22
48185	Asian High	31%	56%	13%	80%	28	51	12	72
48185	Vietnamese High	30%	55%	15%	39%	32	60	16	42
48185	Asian Surname	31%	54%	15%	89%	13	23	6	38
48185	Vietnamese Surname	31%	54%	15%	49%	34	58	16	53
	TOTAL					128	222	58	329
48186	Residual	32%	49%	18%	45%	111	169	62	152
48186	Black High	38%	46%	15%	51%	32	39	13	43
48186	Hispanic High	40%	45%	15%	54%	43	48	16	58
48186	Asian Surname	35%	48%	17%	45%	10	14	5	14
48186	Vietnamese Surname	34%	49%	17%	45%	12	18	6	16
	TOTAL					337	510	161	283
48187	Residual	30%	53%	17%	48%	138	244	80	222
48187	Black High	43%	49%	7%	44%	10	12	2	11
48187	Hispanic High	41%	48%	11%	47%	5	6	1	6
48187	Asian High	29%	57%	14%	57%	9	17	4	17
48187	Asian Surname	30%	53%	17%	63%	9	16	5	19
48187	Vietnamese Surname	30%	53%	17%	63%	13	22	7	26
	TOTAL					184	317	99	300
	GRAND TOTAL					1,338	1,931	643	1,899

Overall, this was developed to attain the following design effects:

TABLE 6: PLANNED DESIGN EFFECTS OF STRATIFICATION

SuperPUMA	Overall	Vietnamese
48181	1.21	1.71
48182	1.23	1.95
48183	1.46	1.91
48184	1.31	1.90
48185	1.97	1.11
48186	1.11	1.85
48187	1.16	1.66
TOTAL	**1.35**	**1.68**



Estimates for the sampling plan were derived from Claritas estimates of households, since Claritas provides such data down to the Census block group level (post-stratification weighting percent frequencies, however, utilize U.S. 2009 Census American Community Survey data, with totals based on the 2010 U.S. Census).

2.3 Household Selection

Households were required to have at least one person over the age of 18. If the person answering the phone was not 18, we asked to speak to someone over the age of 18. If a household contained only adults ages 65 and older, the interview was terminated in 33 percent of such instances to balance for the fact that such households respond more readily to surveys compared to other households.

2.4 Individual Level Selection

One randomly selected adult age 18 and older was selected from each household to participate in the survey. Within-household selection was conducted using a modified Rizzo selection method. Respondents were first asked how many adults 18 or older lived in their households. If the respondent lived alone, the interview would begin immediately. If two people lived in the household, the computer would randomly select one of these two people, either the current respondent or the other person in the household. The interviewer would then ask to speak with the randomly selected person.

In households with more than two people, either the current respondent or any other adult in the household other than the initial respondent was selected by the computer program. If it was another adult, the interviewer would ask the respondent to name the person in the household, other than themselves, who had the most recent birthday. If the person on the phone did not know who had the most recent birthday, the respondent would be asked to roster all individuals in the household by initials and age so that the computer could randomly select one person. However, this process never became necessary because, in every relevant case, the original respondent was able to identify the household member with the most recent birthday, who then became the individual selected to be the final survey respondent.

3. DATA COLLECTION

3.1 Overview

Data collection relied on three interview modes: telephone (CATI), web, and mail. The survey options were explained to those sample members in advance letters and reminder letters. Advance letters and reminder letters in three languages were mailed to all in the sample, offering the options of telephone and web survey models. In addition, sample for which listed telephone numbers could be obtained, traditional telephone interviewing methods are used as well.

The specific steps for the data collection process were as follows.

1. Advance letters in three languages were sent to all households. The advance letter invited the household to participate in the study and offered the option of calling in to the survey center using a toll-free telephone number or completing a web-based survey. Unmatched sample also had the option of sending their phone number by filling out a postcard that was sent with the advance letter. Letters for AB sample with a listed telephone number also notified people that they would be receiving a call in the next few weeks to complete the survey. Advance letters included a $2 pre-incentive.

2. Telephone interviews were attempted with all households for which we had a telephone number. The initial calls commenced one week after the mailing of the advance letters.

3. Reminder notices were sent to all non-responding households.

4. A final reminder notice was sent to all non-responding households. A copy of the mail questionnaire was included in this final reminder notice

The advance letters and reminder postcards included The University of Texas School of Public Health logo and were signed by the Principal Investigator for the study, Dr. Stephen H. Linder, PhD from the Institute for Health Policy (IHP). All of the letters and reminder postcards included a 1-800 toll-free number that the respondent could call for additional information on the survey or to complete the survey by telephone.

3.2 Timeline

The study timeline was as follows:

TABLE 7: TIMELINE

Milestone	Date
Project Award	April 5, 2010
Sampling Plan Approved	July 20100
Draft Instrument Received by SSRS	April 29, 2010
Instrument CATI English Programming	July 1, 2010
Instrument WEB English Programming	August 18, 2010
Instrument Translation	August 2010
Instrument CATI Spanish/Vietnamese Programming	August 2010
Instrument WEB Spanish/Vietnamese Programming	August 2010
Advance Letter Development and Approval	May - June 2010
Advance Letter Translation	May - June, 2010
CATI Pilot Test	September 22, 2010
Web Pilot Test	September 16-22, 2010
Instrument Mail Development	January - March 2011
Mail Pilot Test	December 30, 2010
Final CATI/Web Approval	October 25, 2010
Sample Batch 1 Advance Letters Mailed	October 27, 2010
Sample Batch 1 Web Interview Commencement	October 28, 2010
Sample Batch 1 CATI Interview Commencement	October 29, 2010
1st Preliminary File Delivery	November 19, 2010
Sample Batch 1 Reminder Postcards Mailed	December 2, 2010
Sample Batch 1 & 2 English Mail QN Mailed	March 10-11, 2011
Sample Batch 1 & 2 Spanish Mail QN Mailed	March 9, 2011
Sample Batch 1 & 2 Vietnamese QN Mailed	March 10, 2011
Sample Batch 2 Advance Letters Mailed	January 10, 2011
Sample Batch 2 Web Interview Commencement	January 11, 2011
Sample Batch 2 CATI Interview Commencement	January 20, 2011
Sample Batch 2 Reminder Postcards Mailed	February 7, 2011
Sample Batch 3 Advance Letters Mailed	January 24, 2011
Sample Batch 3 Web Interview Commencement	January 25, 2011
Sample Batch 3 CATI Interview Commencement	February 4, 2011
Sample Batch 3 Reminder Postcards Mailed	February 21, 2011
Sample Batch 3 Mail QN Mailed	March 17-18, 2011
Field Termination	End of March
Final Data File Delivery	July 28, 2011
Final Methods Delivery	September 15, 2011

3.3 Completed Interviews

Table 8 shows the number of completions for each mode of data collection with a separate category for in-bound (toll free) telephone calls from sample members requesting to complete the survey by telephone versus outbound phone interviews where a telephone interviewer called the respondent. For the most part, questions were identical for telephone, web, and mail instruments, although there were some modifications for ease of survey completion using the mail mode. The mail survey was a condensed version of the CATI/Web instruments. The major distinction

between the telephone mode and the web and mail modes is that, in the case of the CATI interviews, a trained interviewer guided the respondent through the process, whereas the web and mail surveys were self-administered.

TABLE 8: COMPLETED INTERVIEWS BY PHONE MATCH STATUS AND MODE

	Total	With Listed Landline Telephone Number	With No Listed Landline Telephone Number
Total Interviews	**5116**	**3319**	**1797**
Phone-outbound	1811	1762	49
Phone-inbound	289	149	140
Web/Internet	1902	870	1032
Mail	1114	538	576

Although web and mail respondents were completing the questionnaires without the direct assistance of an interviewer, all correspondence with respondents included contact information for project staff who were available to assist respondents with any problems they had completing the survey. For those completing the survey on-line, there was access to both staff telephone numbers and a link for emailing for technical support.

TABLE 9: COMPLETED INTERVIEWS BY RACE

Super PUMA	Strata	Total Interviews	Black Interviews	Asian Interviews	Hispanic Interviews	Vietnamese Interviews	Other Interviews	Below Poverty Interviews
48181	Residual	79	66	0	16	0	0	25
48181	Black High	87	24	2	247	0	4	79
48181	Hispanic High	355	5	2	16	12	1	8
48181	Vietnamese High	51	6	2	10	0	3	5
48181	Asian Surname	28	10	15	5	13	1	6
48181	Vietnamese Surname	56	5	0	12	0	0	5
	TOTAL	**656**	116	21	306	25	9	128
48182	Residual	230	56	0	23	1	0	16
48182	Black High	94	19	2	106	0	3	32
48182	Hispanic High	196	27	2	16	2	0	9
48182	Asian High	32	7	3	6	1	1	4
48182	Vietnamese High	68	3	19	4	0	0	2

Super PUMA	Strata	Total Interviews	Black Interviews	Asian Interviews	Hispanic Interviews	Vietnamese Interviews	Other Interviews	Below Poverty Interviews
48182	Asian Surname	35	1	13	5	34	0	8
48182	Vietnamese Surname	69	15	4	52	0	8	11
	TOTAL	**724**	128	43	212	38	12	82
48183	Residual	248	67	4	15	0	3	18
48183	Black High	94	11	1	40	0	1	19
48183	Hispanic High	72	30	16	52	7	2	25
48183	Asian High	169	34	27	29	2	7	17
48183	Vietnamese High	179	4	45	3	1	1	3
48183	Asian Surname	62	3	24	1	30	0	6
48183	Vietnamese Surname	71	22	13	24	1	5	7
	TOTAL	**895**	171	130	164	41	19	95
48184	Residual	193	166	4	27	0	3	27
48184	Black High	226	15	2	57	1	1	18
48184	Hispanic High	94	13	2	16	1	1	6
48184	Asian High	92	3	9	5	0	2	4
48184	Vietnamese High	48	3	28	6	0	1	3
48184	Asian Surname	43	10	26	1	31	1	6
48184	Vietnamese Surname	74	32	1	31	1	6	8
	TOTAL	**770**	242	72	143	34	15	72
48185	Residual	213	8	0	8	0	0	5
48185	Black High	22	8	0	26	1	0	5
48185	Hispanic High	55	27	7	28	1	2	5
48185	Asian High	96	16	13	19	0	5	1
48185	Vietnamese High	110	2	27	3	0	0	3
48185	Asian Surname	41	1	35	4	76	0	8
48185	Vietnamese Surname	122	26	11	23	2	7	8
	TOTAL	**659**	88	93	111	80	14	35
48186	Residual	395	54	0	27	0	1	18
48186	Black High	91	9	0	80	0	3	29
48186	Hispanic High	123	10	4	14	0	2	2
48186	Asian Surname	43	4	10	2	8	1	4
48186	Vietnamese Surname	40	55	5	75	1	8	29
	TOTAL	**692**	132	19	198	9	15	82
48187	Residual	537	15	0	5	0	2	3
48187	Black High	25	5	2	8	0	0	0
48187	Hispanic High	15	7	4	6	1	0	2

Super PUMA	Strata	Total Interviews	Black Interviews	Asian Interviews	Hispanic Interviews	Vietnamese Interviews	Other Interviews	Below Poverty Interviews
48187	Asian High	34	3	19	3	1	1	0
48187	Asian Surname	39	1	27	1	33	2	4
48187	Vietnamese Surname	70	64	16	75	1	9	20
	TOTAL	**720**	95	68	98	36	14	29
	GRAND TOTAL	**5,116**	972	446	1232	263	98	523

As mentioned earlier, the HHS 2010 was administered in three languages, English, Spanish, and Vietnamese. All mailings to High Hispanic strata were provisioned with bilingual materials (English and Spanish) while all mailings to High Vietnamese and Vietnamese Surname strata were furnished English and Vietnamese materials. All Hispanic and Vietnamese strata telephone interviewing was conducted by bilingual interviewers. Any "language barriers" that were encountered in other strata were called back using bilingual interviewers.

TABLE 10: Completed Interviews by Language of Interview

Language of Interview			Total Sample
English	Spanish	Vietnamese	
4485	567	64	5116

3.4 Translation

All questionnaires were translated into both Spanish and Vietnamese to be used in all three modes of interviewing. Translations were completed by TranslationSource, a provider of translation and localization services in Houston, Texas. Translation source carries out the following procedure for all translations:

1. Review of all materials by an Account Manager/Supervisor
2. Translation and editing of documents by a professional translator
3. Review and editing of all translations by a third translator

Following the translation of all documents by TranslationSource, native speakers of Spanish and Vietnamese reviewed the instruments and suggested changes to the translations to be more consistent with colloquial usage and appropriate grammar. These changes were verified with the professional translators at TranslationSource and were incorporated into the translation as deemed appropriate.

Additionally, the HHS team contracted with native speakers of Spanish and Vietnamese and knowledgeable in public health, that performed a final revision to all the survey contact letters and questionnaires translations. All the suggestions for modifications were discussed with TranslationSource to reach an agreement upon the most appropriate translation.

3.5 Training Materials and Interviewer Training

CATI interviewers received both written materials on the survey and formal training for conducting this survey. The written materials were provided prior to the beginning of the field period and included:

1. An annotated questionnaire that contained information about the goals of the study as well as detailed explanations of why questions were being asked, the meaning and pronunciation of key terms, potential obstacles to be overcome in getting good answers to questions, and respondent problems that could be anticipated ahead of time as well as strategies for addressing them.
2. A list of frequently asked questions and the appropriate responses to those questions.
3. A script to use when leaving messages on answering machines.
4. Contact information for project personnel.

Interviewer training was conducted prior to the study pretest (described below) and immediately before the survey was officially launched. Call center supervisors and interviewers were walked through each question in the questionnaire. Interviewers were given instructions to help them maximize response rates and ensure accurate data collection. They were instructed to encourage participation by emphasizing the social importance of the project and to reassure respondents that the information they provided was confidential.

Interviewers were monitored during the first several nights of interviewing and provided feedback, where appropriate, to improve interviewer technique and clarify survey questions. The interviewer monitoring process was repeated periodically during the field period.

3.6 Telephone Mode Development

Prior to going into the field, SSRS programmed the study into a Computer Assisted Telephone Interviewing (CATI) program. The project team conducted extensive checking of the program. All skip patterns were checked through multiple runs through the CATI program, and random data were generated to confirm that all skip patterns were working correctly.

3.7 Web Mode Development

A similar procedure was used for programming and testing the web version of the program, which was also available in three languages. Unlike the CATI program, web respondents were permitted to skip questions they do not wish to answer, so missing data needed to be taken into account in checking the web program. Considerable time and effort was put into creating a web program that was aesthetically pleasing as well as allowing data entry with a minimum amount of error. Because web respondents do not have the benefit of an interviewer guiding them through the survey, it is important to provide a platform that is easy to follow.

3.8 Mail Mode Development

The hard copy version of the instrument was developed in English over a several week period and translated into both Spanish and Vietnamese. This questionnaire was limited to key questions from the main study to avoid overburdening the respondent with a large document that contained complex skip patterns. Only questions related to the adult in the household were included in the mail mode, again to avoid complexity and increase the likelihood of completing an interview. Aside from the reduction in length, the questionnaire was designed to match other

survey modes as closely as possible, particularly the layout of the web survey. Graphic design elements were incorporated into the questionnaire including a photograph of Houston on the cover page and a color background to enhance the appearance of text and check boxes and bubbles.

3.9 Pretesting

The first stage in the pretesting the CATI questionnaire involved conducting a preliminary pretest of nine CATI pretest interviews over two nights. All of the interviews ran longer than the target length of 25 minutes. This necessitated editing the questionnaire to significantly reduce the length prior to the pilot study.

Following the survey revisions, a small series of cognitive interviews were conducted where respondents were interviewed over the telephone and asked to provide feedback on questions where we had concerns about clarity and comprehensibility, as well as questions where we were asking about sensitive information that respondents might not want to discuss.

We conducted five cognitive pretest interviews. Based on the results of these interviews, it was clear that respondents were not experiencing problems understanding questions and did not feel intimidated about responding to the questions that were asked. Timers were put on all questions, and it was determined that the survey was at an appropriate length to move forward.

The next stage was a pilot study to ensure that all phases of project execution, mailing of invitations, completing interviews in multiple modes, and data processing, would work as planned. The English pilot consisted of 23 interviews completed over the telephone and eight completed on the web. The web completes were collected from September 16-22, 2010 and the CATI completes were all collected on September 22, 2010.

At all stages of pretesting and piloting, interviewers received training from project directors and supervisors in conducting the interviews and, following review of recorded pretest interviews, feedback to improve interviewing.

Both the CATI and Web programs were translated into Spanish and Vietnamese. Pilot and pretest studies were performed in all languages to determine comprehensibility and usability of the programs for English, Spanish and Vietnamese speakers. We completed ten pilot completes in Spanish and eight in Vietnamese. Of the Spanish interviews, eight were conducted over the phone and two using the web survey. Of the Vietnamese interviews, seven were conducted by phone and one on the web.

Respondents who did not respond by either phone or mail were sent a hardcopy questionnaire to use for completing the survey. This questionnaire was developed using best practices in hardcopy questionnaire design as established by Dillman in his Tailored Design Method.[3]

The hardcopy questionnaire was also piloted in English, and it was determined that it worked well and could be sent out to all non-respondents. Thirty-one mail surveys for the pilot study were sent to respondents who had requested a mail questionnaire and received seven back as completes. One respondent completed the survey online. These

[3] Dillman, D. (1999). Mail and Internet Surveys, The Tailored Design Method. New York NY: Wiley.

seven hardcopy completes were added to the final dataset since no changes to the instrument were required, and the respondents were drawn from the main study sample.

Extensive changes were made to the instrument following pretesting; however no significant changes resulted from the pilot interviews.

3.10 Incentives

In order to encourage participation in the survey, all respondents were provided a $2 cash incentive in their initial invitation letter, with the exception of the second questionnaire mailing to maintain overall project budget. For members of the AB sample without a listed phone number, an additional incentive of $20 was offered. Information on the incentives was provided in all advance letters and reminder letters and in the introduction to the survey. As mentioned earlier, respondents were offered a chance to participate in a random drawing for a $200 VISA gift card. Overall, 43% of respondents accepted and were sent the $20 incentives. Shortly after the end of field a winner to the sweepstakes was selected and successfully notified.

3.11 Call Rules for the CATI Interviews

The initial telephone interviewing included one initial call plus six callbacks. If an interview was not completed at that point, the telephone number aside for at least two weeks to "rest." After that rest period, an additional six callbacks were attempted. After another four-week rest period, the sample was dialed back three more times. Overall, households received at least 15 call attempts. To increase the probability of completing an interview, we established a differential call rule that required that call attempts be initiated at different times of day and different days of the week.

3.12 Refusal Avoidance and Conversion Strategies

With the increased popularity of telemarketing and the use of telephone answering machines and calling number identification (i.e., caller-ID), the problem of non-response has become acute in household telephone surveys. Similarly, the increasing prevalence of unsolicited advertising in the mail (i.e., junk mail) makes it more difficult to conduct surveys using only invitation letters as we are doing here with the sample without a listed telephone number. In addition to the incentives and call rules for the CATI interviews outlined above, we employed several other techniques to maximize the response rate for the survey. In the CATI interviewing, this included providing a clear and early statement that the call was not a sales call. In all three modes of the survey (telephone, web, and mail), the introduction included an explanation of the purpose of the study, the expected amount of time needed to complete the survey, and a discussion of the incentives.

In an effort to maximize the response rate in the interview phase, respondents were given every opportunity to complete the interview at their convenience. For instance, those refusing to continue at the initiation of or during the course of the telephone interview were offered the opportunity to be contacted at a more convenient time to complete the interview. They were also offered the opportunity to complete the survey on-line or to call into the 1-800 toll-free telephone number to complete the survey at their convenience. Those completing the interview on the web were able to complete the survey at their own speed and stop and re-start as needed.

A key way to increase responses rates is through the use of refusal conversions. Though all of SSRS's interviewers regularly go through "refusal aversion" training, refusals are still a regular part of survey research. SSRS used a core group of specially-trained and highly-experienced refusal conversion interviewers to call all who initially refused the survey in an attempt to persuade respondents to complete the survey.

3.13 Caller ID

A caller ID tag was included in the sample record for all respondents with a phone number. Any respondents with caller ID capabilities on their telephones received the caller ID "UT Health Survey." Although it is impossible to verify what respondents actually saw on their caller IDs, preliminary tests indicate that the caller ID was working properly. This ID was set up to decrease the likelihood that the respondent would screen out the phone calls when confronted by an unfamiliar number on the caller ID.

3.14 Completed Interviews by Telephone Status

The table below shows the number of completed interview done in households that had only a cell phone, only a landline phone, both a landline and cell phone, and the residual categories for no telephone or telephone status unknown. As expected, the proportion of completes from cell phone-only households has been increasing in each round of data collection. We completed surveys with 1,204 cell phone-only households, 3,361 landline and cell phone households, 510 landline-only households, and 41 non-telephone households.

TABLE 11: COMPLETED INTERVIEWS BY LANDLINE PHONE STATUS AND MODE

	Total	With Listed Landline Telephone Number	With No Listed Landline Telephone Number
Total Interviews	**5116**	**3319**	**1797**
Cell phone-only	1204	195	1009
Landline phone-only	510	437	73
Cell phone and landline phone	3361	2677	684
No telephone	41	10	31

3.15 Data Processing and Preparation

Data file preparation began soon after the study entered the field. CATI range and logic checks were used to check the data during the data collection process. After the first several days of data collection, all variables were checked to ensure that data are being collected according to designated skip patterns. Additional data checks were implemented as part of the data file development work, checking for consistency across variables and family members, and developing composite measures of family and household characteristics. At the conclusion of data collection, all variables were checked again to verify that the transfer of data from CATI program to SPSS datafile had been accomplished accurately. Constructed variables such as whether a respondent has health insurance were checked to ensure that data had been correctly pulled from individual items to create the composite variable.

The construction of the final public use data file required combining data from adult and child household members into common variables. Of course, this was only possible with variables that measured the same thing, such as

health insurance status or the presence of a regular healthcare provider. Once these composites were created, they were checked against the original variables to verify that data had been combined accurately.
Final checking of the datafile included checking to ensure that respondents didn't leave more than 50% of their responses blank in the online version of the study, and reviewing length of both web and CATI interviews to isolate outliers. In general, the item nonresponse was quite low; 39 out of 159 had under 1% missing values; another 58 were under 3%. While 13 variables had non-response over 10%. More detail is found in the section on non-response.

3.16 Imputations

Missing data are ubiquitous throughout social science research and can be found in almost all large survey datasets. Replacing the missing values with plausible substitutes (imputation) occurred for survey data in the United States as early as the 1930s. A wide variety of techniques have been developed since that time. Compared with earlier methods of filling in missing values, such as mean substitution and regression imputation, modern imputation methods are designed to account for the missing data mechanism and adjust for the effects of incomplete data on statistical inference. One modern method, multiple imputation (Rubin, 1976), has emerged as a general and widely used technique for analysis in the presence of missing data.

The key idea of multiple imputation (MI) is that missing values are imputed with plausible values drawn from the conditional distribution of the missing data given the observed data under a specified model. This produces a series of "complete" datasets which can then be used for analysis. For a detailed technical review of multiple imputation see Rubin (1987) and Little and Rubin (2002).

Many algorithms have been proposed to impute missing values, but two approaches have been widely adopted and are available in the statistical packages commonly used by social science researchers. The first approach is based on Markov Chain Monte Carlo (MCMC) methods and the second on chained equations. The MCMC approach uses a "normal" statistical model that assumes that the missing values follow a MAR pattern and all the variables in the model are continuous with a multivariate normal distribution (Rubin, 1987; Schafer & Olsen, 1998). Categorical variables can be included as sets of dummy variables and ordinal variables are treated as continuous. The "normal" assumption has been found to be robust even when many of the variables are not continuous or do not have a multivariate normal distribution (Lee, 2010; Schafer & Olsen, 1998). The first widely used implementation of this approach was in the public domain NORM software program (http://www.stat.psu.edu/~jls/misoftwa.html). It has also been implemented in the SAS MI and Stata MI procedures.

The chained equations approach (also referred to as Fully Conditional Specification, or FCS) imputes missing values by iteratively fitting a set of regression equations where each variable is successively treated as the outcome variable and regressed on all other variables in the model. The set of regression equations is used to predict values, random error components are added to the values, and the values are substituted for the values that were missing. Each successive iteration uses the imputed values from the previous iteration in its equations. In this approach, the chained regression models can be tailored to correspond to the level of measurement of the variable. For example, binary variables are estimated using logistic regression, categorical variables with three or more categories by multinomial regression, and ordered categorical variables by ordinal regression.

Instead, all variables can also be treated as continuous, in which case the imputed estimates would approximate those obtained with the "normal" model. The most widely used implementations of the approach are the ICE

procedure in Stata, IMPUTE in the IVEware statistical package available for free download from the University of Michigan Center for Survey Research website (http://www.isr.umich.edu/src/smp/ive/), and the MI module available as an extra cost option in recent versions of SPSS(PASW). Although each of these procedures uses a chained equation approach, the algorithms used and the options available are slightly different.

While MI is new to some social scientists, it is well grounded in a statistical literature dating back to Rubin's seminal paper in 1976. Bayesian theory underlies the MI procedure which allows it to be useful in making inferences in small samples even when the proportion of missing values is large (Allison, 2001; Little & Rubin, 2002). A review of the literature shows it is a widely accepted technique (Graham, 2009; Raghunathan, 2004; Schafer & Graham, 2002).

Several advantages of MI make it a preferable strategy among missing data methodologists. MI provides the researcher a complete data matrix ready to be analyzed. A complete imputed dataset is advantageous because it may reduce missing data bias, improve statistical power, and lead to analysis with consistent results (Kenward & Carpenter, 2007).

MI can be applied very generally to large datasets with complex patterns of missingness among the covariates. MI can have a mixed vector of nominal and interval-level variables. Some imputation techniques, such as "hot-deck" methods, require collapsing categories within variables; this reduces the measure's variance and explanatory power (Marker et al., 2002). It is relatively simple to accommodate restrictions on the values to be imputed, such as imputing values where skip patterns were present or questions were inapplicable. It is also possible to impose logical or consistency bounds, so that the imputed values are consistent with values and distributions of the observed data (Yucel et al., 2008).

MI provides a convenient route for incorporating a considerable amount of information in the model for missingness. Joint relationships among multiple variables in the dataset are estimated, which allows the preservation of a large number of associations (Collins et al., 2001; Rubin, 1987). This improves the efficiency of the imputation model. It is also possible to incorporate information on survey design features, such as survey mode or data on the sampling frame, into the imputation model (Reiter et al., 2006). This combination of advantages is not present with other strategies for dealing with missing data such as complete case analysis, Heckman selection correction (Heckman, 1979; Puhani, 2000), and weighting procedures (Robins et al., 1995; Scharfstein et al., 1999).

Imputation Method and Results

When a "Don't know" or "Refusal" was obtained directly from a respondent for any item, these responses were treated as missing data. The levels of missingness ranged from approximately .01 to 30 percent. Missing data were imputed for 159 variables in this dataset. Details on the missing values for each variable are included in Table 12.

Items were only imputed when at least 3 other measured variables were statistically significant ($p<.001$) predictors of the observed responses, and at least one other variable had a correlation of .20 or higher with the observed responses. This step was taken to be consistent with multiple imputation methodology research that highlights the importance of good auxiliary information in the performance of an imputation model (Collins, Schafer & Kam, 2001). For some variables, there was insufficient predictor information to impute the missing values.

Overall, the patterns of missing values found in these data were typical of RDD surveys on health-related topics. Sensitive questions, such as those asking about financial information, elicited the highest levels of non-response. Missing data were imputed for 159 variables with missingness ranging from .1 to 30.1 percent, shown in Table 12. The vast majority (96%) of respondents who skipped more than one question showed unique missing data patterns. For respondents with more than one missing value, no more than 25 people showed the same pattern of nonresponse. All imputation models assumed (necessarily) that the missing values were missing at random (MAR) (Rubin, 1985). Each imputation model contained a series of correlated auxiliary predictors that were believed to be related to both the likelihood of missingness and to the observed responses, a step which makes the MAR assumption plausible.

The three imputation approaches used most often by social science researchers are the normal-Markov chain Monte Carlo procedures (as implemented in SAS MI and Stata MI) and the chained-equation procedure (as implemented in Stata ICE and SPSS MI). Recent simulation studies (Lee, 2010) find that the MCMC and chained-equation multiple imputation approaches yield similar results. The missing data here were imputed in the ICE application implemented in Stata (Royston, 2005). ICE imputes missing values by iteratively fitting a set of regression equations in which each variable is successively treated as the outcome variable and regressed on all other variables in the model. This set of regression equations is used to predict values including random error components, which are then substituted for the values that were missing. Each successive iteration utilizes the imputed values from the previous one in its equations.

The regression models for many of the imputed values were tailored to correspond to the level of measurement of the outcome variable. For example, binary outcomes were estimated using logistic regression, categorical variables with three or more categories by multinomial regression, and ordered categorical variables by ordinal regression. For continuous variables, or ordered categorical variables with more than 10 categories, a “fully normal” (FN) model was employed which used linear regression in the prediction equations. The result of the FN model is that imputed values do not directly correspond to the researcher’s original level of measurement. For example, income could have been originally measured in thousand-dollar increments, but the imputed values could take on finer gradation (e.g., 25,231.56). To solve this problem, many imputed values were rounded and ranged to be consistent with the original level of measurement. Methodologists have noted concerns about the potential bias of this strategy (Horton, Lipsitz and Parzen, 2003). Problems are most likely to occur, however, in data with much higher levels of missingess than was observed here. Additionally, rounding and ranging appears to be the most practical strategy for researchers who are not methodologists to find the data usable (Johnson and Young, 2009; Johnson and Young, 2011).

For each variable that was imputed, a corresponding “flag” was created to indicate whether a particular value was imputed. The flag variables are coded as “1” if the variable was imputed and “0” if not. Each flag is named with the convention “flag” and the original variable name. For example, the variable qnp11 was imputed; the corresponding imputation indicator is named flagqnp11.

Many of the questions in these data were contingent questions – being asked only if a particular answer had been received to a prior question or set of questions. This poses a bit of a dilemma in missing value imputation. For example, question P10 asks “In the past 12 months have you seen your doctor or other professional, for problems with your mental health, emotions, or nerves, or use of alcohol and drugs?” If the respondent answers “Yes”, question P11 is asked, “Did you seek help for your mental or emotional health or for an alcohol or drug problem or for both?” If the respondent did not answer P10, their data would be imputed. If a “Yes” response was imputed, then it might appear that this respondent had missing data for question P11; since he or she said “Yes” to P10, question P11 “should” have been asked. One strategy to solve this dilemma is to impute values for P11 based on the imputed values for P10. Some people object to this method, however, because it requires data

to be imputed for respondents who were never asked a question, and whether or not the question was asked was not missing completely at random (Rubin 1985; Graham et al., 2006). In the imputation strategy used here, values were imputed only when a "Don't know" or "Refusal" was obtained directly from a respondent. If a respondent failed to answer P10, a value was imputed, but P11 was not imputed for this respondent even if the value imputed for P10 was a "Yes". This strategy is consistent with the idea that imputed values are not intended to be the true value that a respondent would have given, but instead act as a plausible substitute that facilitates statistical analysis when complete cases are required (Acock, 2005; Allison, 2001).

TABLE 12. TOTAL MISSING VALUES FOR EACH IMPUTED QUESTION

Variable Name	Missing Values	Total Valid Respondents	Missing Percent
gender	3	5,116	0.1%
qnh3	1	1,403	0.1%
qnp10	6	4,002	0.1%
qnm9	2	1,244	0.2%
qnr5	8	4,002	0.2%
qnr4a	2	868	0.2%
qnp16	12	4,002	0.3%
qnr12	5	1,389	0.4%
qnu5	20	5,116	0.4%
qnq2	6	1,346	0.4%
qnp9	19	4,002	0.5%
qnr7	19	4,002	0.5%
qnu11	20	4,002	0.5%
qno5a	7	1,378	0.5%
qnv3	7	1,378	0.5%
qngh1	27	5,116	0.5%
qnpp4	6	1,105	0.5%
qnu3a	29	5,116	0.6%
qno1	8	1,378	0.6%
qnn1	32	5,116	0.6%
qnw2	25	3,952	0.6%
qnl7	3	449	0.7%
qnu1	35	5,116	0.7%
qnn2	28	4,002	0.7%
qnr2	13	1,803	0.7%
qna1	37	5,116	0.7%
qnl6	24	3,245	0.7%
qnb1	40	5,116	0.8%
qnr11	11	1,389	0.8%
qnq1	16	1,958	0.8%
qnp13	3	357	0.8%
qnr1	43	5,116	0.8%
qne1	44	5,116	0.9%
qno5c	12	1,378	0.9%
qnh1	45	5,116	0.9%
qnl2	7	771	0.9%

Variable Name	Missing Values	Total Valid Respondents	Missing Percent
qnp17a	3	327	0.9%
qni1	47	5,116	0.9%
qnv11	13	1,378	0.9%
qnq4	31	3,242	1.0%
qns10	49	5,116	1.0%
qnf1	50	5,116	1.0%
qnl12	51	5,116	1.0%
qnq5	23	2,304	1.0%
qno4	14	1,378	1.0%
qno5b	15	1,378	1.1%
qnu7	56	5,116	1.1%
qnb1a	57	5,116	1.1%
qno5d	16	1,378	1.2%
qnp1	62	5,116	1.2%
qnm4	17	1,378	1.2%
qnu4	60	4,568	1.3%
qnpp2	44	3,308	1.3%
qnp4	74	5,116	1.4%
qnl5	58	4,002	1.4%
qnv13	21	1,378	1.5%
qnp17c	5	327	1.5%
qnn12d	80	5,116	1.6%
qnn12a	82	5,116	1.6%
qnn9	84	5,116	1.6%
qnn12b	86	5,116	1.7%
qnp2	88	5,116	1.7%
qnu3	91	5,116	1.8%
qng1a	92	5,116	1.8%
qnp6	93	5,116	1.8%
qnpp1	62	3,308	1.9%
qnp3	97	5,116	1.9%
qnpp3	63	3,308	1.9%
qnn7	63	3,238	1.9%
qnq14	46	2,344	2.0%
qny2	103	5,116	2.0%
qnq10	37	1,807	2.0%
qnn10	107	5,116	2.1%
qnt4	107	5,116	2.1%
qnu2	110	5,116	2.2%
qnv16	30	1,378	2.2%
qneh2	96	4,396	2.2%
qna3_01	112	5,116	2.2%
qna3_02	112	5,116	2.2%
qna3_03	112	5,116	2.2%
qna3_07	112	5,116	2.2%
qnv21	20	911	2.2%

Variable Name	Missing Values	Total Valid Respondents	Missing Percent
qnp5	113	5,116	2.2%
qnn12c	114	5,116	2.2%
qng1b	115	5,116	2.2%
qnm1g	31	1,378	2.2%
qnq11	23	998	2.3%
qnn6a	17	735	2.3%
qnm1a	32	1,378	2.3%
qng1c	119	5,116	2.3%
qnr8	59	2,501	2.4%
qnl14	125	5,116	2.4%
qnv20	23	911	2.5%
qnl13	29	1,129	2.6%
qnw3	2	73	2.7%
qnp17b	9	327	2.8%
qnr6	12	430	2.8%
qny1	151	5,116	3.0%
qnr3	15	483	3.1%
qneh1	137	4,396	3.1%
qny3	160	5,116	3.1%
qnemp6	96	3,062	3.1%
qnv18	44	1,378	3.2%
qnemp1	168	5,116	3.3%
qnl1a	172	5,116	3.4%
qnp17d	11	327	3.4%
qnpp9	36	1,036	3.5%
qnm1b	48	1,378	3.5%
qngh5	180	5,116	3.5%
qnm2	5	138	3.6%
qnm15	6	159	3.8%
qnu7a	197	5,116	3.9%
qnm1f	54	1,378	3.9%
qnm1i	54	1,378	3.9%
qnp11	14	357	3.9%
qnn6	201	5,116	3.9%
qnu9b	201	5,116	3.9%
qnu9g	206	5,116	4.0%
qnm1h	56	1,378	4.1%
qnq19	165	4,002	4.1%
qnm1j	58	1,377	4.2%
qnt14	219	5,116	4.3%
qnu9h	228	5,116	4.5%
qnt2	229	5,116	4.5%
qngh2	234	5,116	4.6%
qnv17	64	1,378	4.6%
qnemp5	146	3,062	4.8%
qnl10	47	965	4.9%

Variable Name	Missing Values	Total Valid Respondents	Missing Percent
qnl15	37	754	4.9%
qnemp7	156	3,062	5.1%
qnu9e	268	5,116	5.2%
qnt1	269	5,116	5.3%
qnu9d	305	5,116	6.0%
qnl9	70	1,143	6.1%
qnr10	89	1,433	6.2%
qnt16	319	5,116	6.2%
qnp12	24	357	6.7%
qnr9	72	1,068	6.7%
qnp14	9	122	7.4%
qnv19	71	911	7.8%
qnemp8	254	3,062	8.3%
incom11	341	4,002	8.5%
qnr4	14	163	8.6%
qnl1f	440	5,116	8.6%
qnemp2a	21	243	8.6%
qnu7b	503	5,116	9.8%
qnl1b	528	5,116	10.3%
qnl1g	547	5,116	10.7%
qnl1i	588	5,116	11.5%
qnl1h	594	5,116	11.6%
qnu9c	669	5,116	13.1%
incom1	640	3,062	20.9%
qnp8	769	3,679	20.9%
qnp7	837	3,679	22.8%
qnt5	1,174	5,116	22.9%
qnl1j	1,180	5,116	23.1%
qnpp5	446	1,922	23.2%
qnt3	1,227	5,116	24.0%
incom5	1,540	5,116	30.1%

4. RESPONSE

4.1 Overview

Response rates are one method used to assess the quality of a survey, as they provide a measure of how successfully the survey obtained responses from the sample. The American Association of Public Opinion Research (AAPOR) has established standardized methods for calculating response rates (AAPOR, 2008). This survey uses AAPOR's response rate definition RR4, with an AAPOR-approved alternative method of addressing ineligible households.

4.2 Defining the Response Rate

SSRS calculates response rates in accordance to AAPOR RR3 calculations. However, the AAPOR Standard Definitions manual does not provide explicit guidelines for ABS designs, nor does it provide more than general guidance for screener surveys.

Screener Studies

Generally, screener surveys are different than general population surveys in that there are two levels of eligibility: household and screener. That is, a sample record is "household eligible" if it is determined that the record reaches a valid household. Screener eligible refers to whether known household-eligible records are eligible to in fact complete the full survey. In the case of the Health of Houston survey, screener eligibility refers to whether a household has a member under the age of 65, for those surveys in which such criteria are mandatory. As well, households must not be vacation homes and must reside within the geographic target area of the study.

The standard AAPOR RR3 formula is as follows:

$$\frac{I}{I + R + NR + [UNR + UR]e}$$

Where:
I: Completed Interview
R: Known Eligible Refusal/ Breakoff
NR: Known Eligible Non-Respondent
UR: Household, Unknown if Screener Eligible
UNR: Unknown if Household
e: Estimated Percent of Eligibility

At issue with this calculation for screener surveys is that it does not distinguish the two separate eligibility requirements: UNR and UR and both multiplied by an overall "e" that incorporates any and all eligibility criteria. An alternative RR4 calculation utilized by a large number of health researchers and academicians simply divides "e" into two separate numbers, one for household eligibility and one for screener eligibility:

$$\frac{I}{I + R + NR + [(UNR)e2 + (UR)]e1}$$

Where:
e2 = Estimated Percent of Household Eligibility
e1 = Estimated Percent of Screener Eligibility

"E" calculations are completed via the standard "proportional representation" method dictated by AAPOR. In short, e2 is all identified household / (all identified households + all identified non-households) and e1 = all identified households eligible to do the full survey / (all identified households known to be eligible to do the full survey + all identified households know to not be eligible to do the full survey).

ABS Studies

ABS studies are particularly challenging for response rate calculations given that they are typical multi-modal. That is, while the frame is address-based, the method of interviewing is often web and/or telephone as well as address. Therefore, the question is how to treat telephone dispositions when the frame is based on address. Prior studies (specifically, 2010 Massachusetts Health Insurance Survey) show that over 95% of the time, completed interviews via phone were completed by a person at the address sampled. Given this high "hit rate," it is our opinion that all sample records determined by phone to be an occupied household should be considered a successful match between phone and address. This is important, because it therefore means any eligible refusal should in fact be treated as an eligible refusal (meaning, we assume that the phone matched the address and therefore it is a refusal from a valid sample record). That said, any non-working, fax, and business disposition is, by its nature, proven to be an unsuccessful match between phone and address (if it were a successful match, after all, we would not have reached a non-working number!). Therefore, any such records should not be treated as ineligible, but in fact UNR, a sample record for which household eligibility has not yet been established.

By definition then, a large percent of sample records will end up as UNR. Among unmatched sample, there will be completed interviews, then a few break-offs via the Internet, and returned mail that will be dispositioned as bad addresses. The vast majority of unmatched sample, however, will be considered a "no answer," given that invitations to participate were mailed, without any response whatsoever. And as mentioned, within matched sample, all non-working/fax/business telephone dispositions for which there is no more "important" web or mail dispositions (like a completed interview) will be considered UNR as well. It is critical that e2 is handled with care, therefore, since it affects so many records. Currently there are two modes of thought on e2 for ABS studies. The first is to allow proportional representation to run its course, which often leads e2 to range from .5 to .8. However, others argue that the USPS would not spend so much energy trying to deliver mail to such a large percent of bad addresses, on a daily basis. And furthermore, co-listing efforts in the past have found that over 90 percent of all No Answer addresses in fact are valid households. As such, e2 should not be determined by proportional representation, but rather should be forcibly set to be in the .9 range. We believe the latter of these two schools of thought holds more weight, though the institutionalization of a .9 e2 can easily drop response rate 15 percent or more. Given this fact, our advice is to analyze the dispositions on a case-by-case basis, as non-eligible addresses naturally vary from location to location. For example, Cape Cod has a high degree of non-eligible address due to

vacation homes; New Orleans contains a significantly larger share of vacant households, etc. In short, each location may have its own story, leading to a geographic-specific e2 calculation. We used proportional representation for e2 during the study.

ABS Dispositions

Telephone studies are often difficult because there are so many interim dispositions that then have to be converted back to final dispositions. For example, a refusal that turns into a No Answer should be moved back into refusal at the end of a study. ABS studies that utilize telephone interviewing require the same steps. Simply put, any record that is dispositioned as a R or NR at one point and then moves to NR and UNR should move back to R and NR. In telephone surveys, any record that moved from R or NR to non-working/fax/business would remain non-working/fax/business under the assumption that the sampled phone number is no longer reaching a valid household. In an ABS survey, therefore, such a record would become UNR.

It is important to note that CATI systems come with their own dispositions that are not always in-line with AAPOR dispositions. In fact, many AAPOR dispositions, particularly in screener studies, have to be generated based on not just CATI dispositions but raw data from the questionnaire as well. Dispositions used for the study, matched up with the CATI dispositions as provided by the CfMC CATI software are presented in Table 13:

TABLE 13: DISPOSITION DEFINITIONS

AAPOR Dispositions	CATI Disposition	ABS Disposition	Description
Completed (Full) Interview	Complete	I	R completes interview
Refusal unknown eligibility	E, V, or M on dispo screen w/o enough data to determine screener eligibility	UR	R refuses to be interviewed before answering screener questions
Ref Eligible	E, V or M on dispo screen with enough data to determine screener eligibility	R	R refuses to be interviewed before answering screener questions
Answering Machine	O/Q on dispo screen w/o enough data to determine screener eligibility	UNR	Answering machine picks up
AM known eligible	O/Q on dispo screen with enough data to determine screener eligibility	NR	Answering machine picks up, screening was completed during a prior call
Physically/Mentally Unable	L on dispo screen w/o enough data to determine screener eligibility	UNR	Hard of hearing/mentally disabled such that a conducting a survey is not possible
Physically/Mentally Unable Eligible	L on dispo screen with enough data to determine screener eligibility	NR	Hard of hearing/mentally disabled such that a conducting a survey is not possible, nevertheless screening data was acquired
Spanish/Vietnamese	H/J on dispo screen w/o enough data to determine screener eligibility	UNR	Household only speaks Spanish or Vietnamese, screener not completed
Spanish/Vietnamese Eligible	H/J on dispo screen with enough data to determine screener eligibility	NR	Household only speaks Spanish or Vietnamese, screener completed
Lang Unable	D on dispo screen w/o enough data to determine screener eligibility	UNR	Household does not speak English/ Spanish/Vietnamese, no screener data
Lang Unable Eligible	D on dispo screen with enough data to determine screener eligibility	NR	Household does not speak English/ Spanish/Vietnamese, but screener completed
Other	F or G on dispo screen w/o enough data to determine screener eligibility	UNR	Duplicate/already completed interview/not available for duration
Other Eligible	F or G on dispo screen with enough data to determine screener eligibility	NR	Duplicate/already completed interview/not available for duration screener completed
Busy	S on dispo screen w/o enough data to determine screener eligibility	UNR	Busy tone
Busy Eligible	S on dispo screen with enough data to determine	NR	Busy tone, prior call attained screener data

	screener eligibility		
No Answer	R on dispo screen w/o enough data to determine screener eligibility	UNR	No answer
No Answer Eligible	R on dispo screen with enough data to determine screener eligibility	NR	No answer, prior call attained screener data
Hard Solicitor Block	P on dispo screen w/o enough data to determine screener eligibility	UNR	Privacy manager/call zapper blocks the call
Fax Data	C on dispo screen w/o enough data to determine screener eligibility	UNR	Fax, data tone, or modem
Not Working/Other Tech Problem	A on dispo screen	UNR	No dial tone, not working automated message
Cell Phone	K on dispo screen	UNR	Dialed to a cell phone (note we will not have these…we will conduct the interview)
Business	B on dispo screen	UNR	Answering machine or person indicates a business
Screener Ineligible: Terminate	Determined by data	Not in Calculation (except to determine e1)	Screener data completed and household is not eligible to do the survey
Call Back	T or U on dispo screen w/o enough data to determine screener eligibility	UR	Respondent asked to be called back at a later time
Call Back Eligible	T or U on dispo screen with enough data to determine screener eligibility	R	Respondent asked to be called back at a later time, but got through the screener
Web Suspend	Based on data from web survey, w/o enough data to determine screener eligibility	UR	Respondent started web survey but did not complete screener data
Web Suspend Eligible	Based on data from web survey, with enough data to determine screener eligibility	R	Respondent started web survey and did complete screener data

4.3 Final Response Rates

Final response rates are summarized in Table 14 and 15. The response rate for the study was 28.9%.

TABLE 14: RESPONSE RATES BY SUPERPUMA

Disposition	48181	48182	48183	48184	48185	48186	48187	Total
Interview	656	724	895	770	659	692	720	**5,116**
Refusal unknown eligibility	339	398	442	403	430	361	410	**2,783**
Refusal Eligible	66	65	90	91	58	67	72	**509**
Answering Machine	191	327	367	307	398	337	448	**2,375**
AM known eligible	13	26	24	19	16	27	24	**149**
Physically/Mentally Unable	6	8	7	12	8	6	13	**60**
Physically/Mentally Unable Eligible	1	3	0	5	1	3	3	**16**
Spanish	5	0	4	3	2	0	1	**15**
Spanish Eligible	0	0	0	0	0	0	0	**0**
Lang Unable	33	55	98	52	83	26	53	**400**
Lang Unable Eligible	1	2	1	4	3	0	0	**11**
Other	40	56	52	52	60	46	63	**369**
Other Eligible	15	13	21	20	18	21	16	**124**
Busy with AM or Hhold	48	28	24	27	14	37	14	**192**
Busy Eligible	1	0	0	1	0	0	0	**2**
No Answer with AM or Hhold	1256	1280	1387	1063	1277	1287	1267	**8,817**
No Answer Eligible	22	17	19	12	7	17	13	**107**
Mail Undeliverable	873	881	1787	821	570	821	610	**6,363**
Hard Solicitor Block	4	7	8	8	5	6	4	**42**
Hard Solicitor Block Unknown Hhold or Elig	0	0	0	0	0	0	0	**0**
Fax Data	98	119	106	119	117	136	116	**811**
Non-Working	370	408	434	416	427	403	431	**2,889**
Cell Phone	11	7	1	5	1	0	2	**27**
Business	23	37	52	29	47	31	48	**267**
Ineligible (terminate)	62	59	77	81	144	61	102	**586**
Call Back	81	59	69	60	37	47	29	**382**
Call Back Eligible	15	23	15	19	7	13	5	**97**
Web Suspend	0	1	2	0	2	2	2	**9**
Web Suspend Eligible	3	5	5	2	5	2	2	**24**
Total Sample	4233	4608	5987	4401	4396	4449	4468	**32,542**
AAPOR Response Rate 4	28.9%	27.9%	34.0%	30.0%	25.7%	27.6%	26.3%	**28.9%**
Refusal Rate	16.9%	17.1%	19.4%	18.3%	16.6%	16.3%	16.2%	**17.4%**
Cooperation Rate	63.3%	62.3%	64.0%	62.5%	61.0%	63.1%	62.1%	**62.6%**

TABLE 15: RESPONSE RATES BY STRATA

Disposition	AA High	Hispanic High	Vietnamese High	Asian High	Asian Surname	Vietnamese Surname	Residual
Interview	1896	639	911	423	456	291	500
Refusal unknown eligibility	1027	359	325	239	289	149	395
Refusal Eligible	200	93	50	37	53	30	46
Answering Machine	1006	221	273	239	206	144	286
AM known eligible	59	18	22	12	11	12	15
Physically/Mentally Unable	24	14	5	2	7	0	8
Physically/Mentally Unable Eligible	6	6	1	0	2	0	1
Spanish	1	0	0	0	2	0	12
Spanish Eligible	0	0	0	0	0	0	0
Lang Unable	72	37	14	27	55	76	119
Lang Unable Eligible	1	3	0	1	3	1	2
Other	159	43	42	34	35	22	34
Other Eligible	39	20	24	12	9	9	11
Busy with AM or Hhold	42	44	70	4	8	7	17
Busy Eligible	0	1	1	0	0	0	0
No Answer with AM or Hhold	2957	1054	1630	636	882	561	1098
No Answer Eligible	32	18	37	2	7	4	7
Mail Undeliverable	2255	1111	1323	651	670	148	205
Hard Solicitor Block	15	6	8	6	2	2	3
Hard Solicitor Block Unknown Hhold or Elig	0	0	0	0	0	0	0
Fax Data	290	122	128	45	74	42	110
Non-Working	1003	418	540	253	328	130	217
Cell Phone	2	1	21	0	0	2	1
Business	113	34	19	22	35	17	27
Ineligible (terminate)	212	82	66	73	40	43	70
Call Back	93	42	132	22	36	16	41
Call Back Eligible	14	17	39	1	15	5	6
Web Suspend	5	0	0	2	0	1	1
Web Suspend Eligible	13	2	5	1	1	1	1
Total Sample	11536	4405	5686	2744	3226	1713	3233
AAPOR Response Rate 4	29.8%	30.3%	32.2%	32.5%	26.7%	25.8%	20.4%
Refusal Rate	18.0%	20.1%	12.8%	19.2%	19.0%	16.5%	16.4%
Cooperation Rate	62.5%	60.4%	71.9%	63.3%	58.6%	64.1%	55.5%

There has been some debate over whether ABS designs are comparable to RDD studies with regard to response rate. In general, ABS studies will often attain response rates some 10 to 15 percentage points lower than comparable RDD studies (see as one comparison public release documents from the Massachusetts Health Insurance Surveys of 2009-present). It is difficult, however, to make a claim that the lower response rate is primarily due to lower overall response, due to substantial difference is the characteristics of the two samples. RDD studies tend to attain yields of 20 or even 30:1 (that is, 30 sample records needed to attain a single interview); ABS studies typically range from 4 to 8:1. This is because ABS sample is far more likely to reach valid households. RDD sample has a high prevalence of ineligibility (fax machines, non-working numbers, businesses, etc.), whereas such records are rare in ABS studies (as discussed earlier, even if one is to encounter a non-working number, such a record is not considered ineligible since eligibility can only be determined based on the actual mailing address, not the matched phone number). RDD sample as well has a large percent of "unknown" sample (no answers, busies); ABS however is mostly unknown in that a large percent of sample records are sent an invitation and a response is never attained. Yet these two types of unknown dispositions are dramatically different. RDD unknowns are more often than not non-working numbers that ring to a no answer. In ABS studies we assume that the vast majority are eligible households. This is evidenced by RDD studies were "e" is often under 10% while "e" can range as high as 90% in ABS studies. In the end, while certainly some proportion of the typically lower ABS response rate is due lower response in unmatched sample where an outbound call attempt is impossible, it is also the case the some of the difference is an artifact of the calculation itself and the quite difference characteristics of the samples.

5. SURVEY WEIGHTS AND VARIANCE ESTIMATION

5.1 Survey Weights

Survey data are weighted to adjust for differential sampling probabilities, to reduce any biases that may arise because of differences between respondents and non-respondents (i.e., nonresponse bias), and to address gaps in coverage in the survey frame (i.e., coverage bias). Survey weights, when properly applied in surveys can reduce the effect of nonresponse and coverage gaps on the reliability of the survey results (Keeter et al. 2000, Groves 2006).

We constructed analytical survey weights for the HHS 2010 using standard procedures. That is, separate weights are created for all persons and for the target-person in the household. The weights can be used to produce adult- and child-level population estimates as well as estimates of the total population in Houston.

5.2 Constructing the Base Weights

The first step in the weighting process for each sample is to create a base (design) weight for each completed survey. That household weight is used to construct weights for each person in the household and for the target-person in each household.

Base Weight Adjustment #1—Stratification Correction: We first adjusted the base weight so that all the households are adjusted for differential probabilities of selection. This adjustment corrects for the over-sampling of addresses in some strata in comparison to others:

$$(f=i_{interviews}/N_{frame})$$

Where:
f = the household probability of selection based weight
n = the number of interviews by strata
N = the household counts by strata

TABLE 16: STRATIFICATION WEIGHTS BY SUPERPUMA AND STRATA

Super PUMA	Strata	Percent of Households	Allocation of Interviews	Stratification Weight
48181	Residual	2.0%	2.0%	1.29
48181	Black High	2.1%	1.7%	1.24
48181	Hispanic High	7.7%	6.7%	1.11
48181	Vietnamese High	0.1%	1.0%	0.08
48181	Asian Surname	0.1%	0.6%	0.12
48181	Viet. Surname	0.1%	0.9%	0.06
48182	Residual	6.6%	4.9%	1.47
48182	Black High	1.8%	1.8%	0.97
48182	Hispanic High	4.8%	4.5%	1.26
48182	Asian High	0.2%	1.4%	0.12
48182	Vietnamese High	0.3%	0.8%	0.47
48182	Asian Surname	0.1%	0.8%	0.17
48182	Viet. Surname	0.2%	1.0%	0.13
48183	Residual	9.9%	5.2%	2.04
48183	Black High	2.2%	1.6%	1.22
48183	Hispanic High	2.0%	1.3%	1.43
48183	Asian High	2.4%	3.0%	0.68
48183	Vietnamese High	1.0%	3.4%	0.31
48183	Asian Surname	0.3%	1.0%	0.21
48183	Viet. Surname	0.2%	1.3%	0.18
48184	Residual	6.5%	4.0%	1.73
48184	Black High	4.4%	3.6%	0.99
48184	Hispanic High	2.5%	1.9%	1.36
48184	Asian High	0.8%	0.8%	0.84
48184	Vietnamese High	0.2%	2.2%	0.13
48184	Asian Surname	0.2%	1.0%	0.20
48184	Viet. Surname	0.2%	0.9%	1.29
48185	Residual	10.7%	4.5%	0.15
48185	Black High	0.1%	0.3%	2.57
48185	Hispanic High	0.9%	0.8%	0.22
48185	Asian High	1.0%	2.6%	0.88
48185	Viet. High	0.4%	1.6%	0.47
48185	Asian Surname	0.2%	1.0%	0.23
48185	Viet. Surname	0.5%	2.5%	0.30
48186	Residual	8.5%	6.5%	0.19
48186	Black High	1.5%	1.7%	1.11
48186	Hispanic High	2.3%	2.0%	0.86
48186	Asian Surname	0.1%	0.8%	0.95
48186	Viet. Surname	0.1%	0.7%	0.09
48187	Residual	13.7%	11.3%	0.08
48187	Black High	0.4%	0.4%	1.31
48187	Hispanic High	0.1%	0.3%	0.87
48187	Asian High	0.2%	0.8%	0.20
48187	Asian Surname	0.2%	0.6%	0.25
48187	Viet. Surname	0.2%	1.1%	0.20

This step has the additional feature of correcting for non-response[4] as well, since the percent of interviews, rather than total sample, in each strata, is matched to the percent of households in each strata. Therefore,

[4] Nonresponse creates biases survey estimates because the characteristics of those interviewed differ from those who were not interviewed. The size of the bias is based on this difference and the response rate (Groves, 1989). Non-response adjustments are designed to reduce this bias. A weighting class adjustment (Brick and Kalton, 1996) method is the type of nonresponse adjustment procedure typically used in most survey research, and is utilized here as a function of the stratification adjustment as described in the body of the text.

non-response and stratification are corrected in one step (compared to matching sample to households, and then correcting the number of interviews to sample as a separate non-response correction).

This correction was used for both target and child weights.
Base Weight Adjustment #2—Number of Persons Correction: As well, a number of persons adjustment was made, such that households with 1 member received a base weight correction of 1, and upward so that 3 members received a base weight correction of 3 (*t*). This correction was capped at 4 to prevent large weights.

Base Weight Final: The final base weight is a simple product of the stratification and persons in the household adjustments,

$$B_i=(f_i*t_i)^{-1}$$

Where:
B = the final base weight
f = the household stratification weight
t = the number of household members

This correction was used for both target and child weights.

5.3 Constructing the Adult and Child Weights

To create a weight for each selected adult respondent and selected child in an interviewed household, we started with the base weight and then post-stratified so that our weighted sample population totals equaled population control totals based on data for Houston. Specifically, we aligned the sample to current Census population estimates for Houston for age, race/ethnicity by education (children were by race only), gender, homeownership (adults only), and race/SuperPuma. The demographic information and homeownership percentage data came from the American Community Survey, 2009, while the total household and population counts were based on the 2010 U.S. Census. We examined the distribution of the resulting person weights and determined that there was not need to implement trimming rules.

We utilized an iterative proportionate fitting (IPF) procedure to create the post-stratification weights. IPF is a now-ubiquitous sample balancing routine originally developed by W. Edwards Deming and Frederick F. Stephan to adjust samples in economic and social surveys on selected demographic characteristics against data obtained from the U.S. Census. The theory behind IPF is explained in Deming's book Statistical Adjustment of Data (1943), available in reprint from Dover Publications. Details on the Deming-Stephan method are spelled out in Chapter VII, "Adjusting to Marginal Totals" (Werner, 2004). IPF ("raking") uses least-squares curve fitting algorithms to obtain a unique weight for each case that minimizes the root mean square error (RMSE) across multiple dimensions simultaneously. Then it applies these weights to the data and repeats the procedure using the newly obtained marginal counts to obtain yet another set of weights. This process is repeated for a specified number of iterations or until the difference in the RMSE between successive steps becomes less than a specific minimum value. This study employed an IPF procedure using the statistical software, QBAL. QBAL not only is the "industry standard" software for sample balancing post-

stratification but also allows for the application of a pre-existing base weight to the input data for the sample balancing process.

Below are the control totals used and frequencies of the data, before and after the post-stratification routine. Note the adjustment of the control targets to account for missing data in the sample, a standard method of dealing with missing data for weighting procedures:

TABLES 17 – 19: RESPONDENT (ADULTS) WEIGHT TARGETS AND OUTCOMES

Gender	Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
Male	49.54%	49.5%	36.6%	49.5%
Female	50.46%	50.5%	63.4%	50.5%
Total	100.00%	100.00%	100.00%	100.00%
Home Ownership				
Rent	36.11%	35.7%	33.4%	35.3%
Own	63.89%	63.2%	66.6%	62.5%
Total	100.00%	100.0%	100.0%	100.0%
Age				
18 thru 24	12.91%	12.9%	4.6%	12.9%
25 thru 34	21.41%	21.3%	16.8%	21.3%
35 thru 44	21.73%	21.7%	19.8%	21.7%
45 thru 54	20.03%	20.0%	21.3%	20.0%
55 thru 64	13.36%	13.3%	20.8%	13.3%
65 thru 96	10.55%	10.5%	16.5%	10.5%
Total	100.00%	99.7%	99.7%	99.7%
DK/R		0.3%	0.3%	0.3%

Race by Education		Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
No H.S. Diploma	White/Other	2.45%	2.4%	1.6%	2.4%
	Black	2.65%	2.6%	1.5%	2.6%
	Asian	0.27%	0.3%	0.3%	0.3%
	Vietnamese	0.32%	0.3%	0.6%	0.3%
	Hispanic	15.75%	15.4%	7.5%	15.3%
H.S. Diploma	White/Other	9.66%	9.5%	6.7%	9.5%
	Black	6.40%	6.3%	4.8%	6.3%
	Asian	0.62%	0.6%	0.8%	0.6%
	Vietnamese	0.55%	0.5%	1.2%	0.5%
	Hispanic	11.17%	11.0%	6.5%	11.0%
Some College	White/Other	12.29%	12.0%	8.8%	12.2%
	Black	6.06%	5.9%	5.5%	5.9%
	Asian	0.86%	0.8%	1.3%	0.8%

Race by Education (cont'd)		Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
	Vietnamese	0.34%	0.3%	0.9%	0.3%
	Hispanic	6.08%	6.0%	4.2%	5.9%
College Degree	White/Other	15.95%	15.6%	24.6%	15.6%
	Black	2.92%	2.9%	7.3%	2.9%
	Asian	2.20%	2.2%	7.0%	2.2%
	Vietnamese	0.45%	0.4%	2.0%	0.4%
	Hispanic	3.01%	3.0%	5.0%	3.0%
	Total	100.00%	98.0%	98.0%	98.0%
	DK/Ref		1.97%	2.0%	2.0%

Race by SuperPUMA		Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
48181	White/Other	2.59%	2.5%	3.6%	2.6%
	Black	2.45%	2.4%	2.3%	2.4%
	Asian	0.20%	0.2%	0.9%	0.2%
	Hispanic	8.18%	8.0%	5.9%	8.1%
48182	White/Other	4.77%	4.7%	5.8%	4.7%
	Black	2.53%	2.5%	2.5%	2.5%
	Asian	0.57%	0.6%	1.7%	0.6%
	Hispanic	5.88%	5.8%	4.1%	5.8%
48183	White/Other	5.67%	5.6%	7.4%	5.6%
	Black	3.10%	3.0%	3.4%	3.1%
	Asian	1.32%	1.3%	3.4%	1.3%
	Hispanic	4.44%	4.4%	3.1%	4.4%
48184	White/Other	4.70%	4.6%	5.4%	4.7%
	Black	3.47%	3.4%	4.7%	3.4%
	Asian	0.72%	0.7%	2.1%	0.7%
	Hispanic	4.18%	4.1%	2.7%	4.1%
48185	White/Other	6.58%	6.5%	5.5%	6.5%
	Black	2.01%	2.0%	1.7%	2.0%
	Asian	1.64%	1.6%	3.5%	1.6%
	Hispanic	4.60%	4.5%	2.1%	4.6%
48186	White/Other	6.29%	6.2%	6.3%	6.2%
	Black	2.57%	2.5%	2.6%	2.5%
	Asian	0.34%	0.3%	0.6%	0.3%
	Hispanic	5.34%	5.2%	3.8%	5.3%
48187	White/Other	9.73%	9.5%	8.1%	9.8%
	Black	1.91%	1.9%	1.9%	1.9%
	Asian	0.81%	0.8%	2.2%	0.8%
	Hispanic	3.39%	3.3%	1.8%	3.4%
	Total	100.00%	98.9%	99.1%	99.1%
	DK/Ref		0.9%	0.9%	0.9%

TABLES 20 – 21: CHILD WEIGHT TARGETS AND OUTCOMES

Gender	Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
Male	51.14%	50.9%	36.6%	50.9%
Female	48.86%	48.7%	63.4%	48.7%
Total	100.00%	99.6%	100.00%	99.6%
DK/R		0.4%	0.4%	
Age				
0 thru 5	36.34%	36.2%	33.5%	36.2%
6 thru 12	37.68%	37.5%	37.0%	37.5%
13 thru 17	25.98%	25.9%	29.1%	25.9%
Total	100.00%	99.6%	99.6%	99.6%
DK/R		0.4%	0.4%	0.4%
Race/Ethnicity				
White/Other	27.20%	26.5%	33.5%	26.5%
Black	19.78%	19.2%	17.9%	19.2%
Asian	3.21%	3.1%	8.8%	3.2%
Vietnamese	1.57%	1.5%	7.3%	1.5%
Hispanic	48.24%	46.9%	29.8%	46.9%
Total	100.00%	97.3%	97.3%	97.3%
DK/R		2.7%	2.7%	2.7%

Race by SuperPUMA		Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
48181	White/Other	1.22%	1.2%	2.4%	1.2%
	Black	2.24%	2.2%	1.6%	2.2%
	Asian	0.06%	0.1%	0.9%	0.1%
	Hispanic	9.87%	9.6%	6.3%	9.6%
48182	White/Other	1.85%	1.8%	4.0%	1.8%
	Black	2.14%	2.1%	2.0%	2.1%
	Asian	0.33%	0.3%	1.5%	0.3%
	Hispanic	8.31%	8.1%	5.5%	8.1%
48183	White/Other	2.43%	2.4%	4.1%	2.4%
	Black	2.89%	2.8%	3.4%	2.8%
	Asian	0.71%	0.7%	2.8%	0.7%
	Hispanic	4.55%	4.4%	4.4%	4.4%
48184	White/Other	3.19%	3.1%	3.8%	3.1%
	Black	3.97%	3.9%	3.8%	3.9%
	Asian	0.87%	0.8%	2.3%	0.8%
	Hispanic	5.54%	5.4%	3.9%	5.4%
48185	White/Other	5.03%	4.9%	5.1%	4.9%
	Black	2.45%	2.4%	2.0%	2.4%
	Asian	1.54%	1.5%	5.2%	1.5%

Race by SuperPUMA (cont'd)		Population Proportion	Adjusted Percent	Pre-Rake Sample	Post-Rake Sample
	Hispanic	6.33%	6.2%	2.4%	6.2%
48186	White/Other	4.58%	4.5%	5.6%	4.5%
	Black	2.99%	2.9%	2.7%	2.9%
	Asian	0.26%	0.3%	0.8%	0.3%
	Hispanic	8.85%	8.6%	5.2%	8.7%
48187	White/Other	8.89%	8.7%	8.6%	8.7%
	Black	3.09%	3.0%	2.4%	3.0%
	Asian	1.02%	1.0%	2.5%	1.0%
	Hispanic	4.80%	4.7%	2.1%	4.7%
	Total	100.00%	97.3%	98.3%	97.3%
	DK/Ref		2.7%	2.7%	2.7%

5.4 Variance Estimation and the Average Design Effect

Complex survey designs and post-data collection statistical adjustments affect variance estimates and, as a result, tests of significance and confidence intervals. Variance estimates derived from standard statistical software packages that assume simple random sampling are generally too low, which leads significance levels to be overstated and confidence intervals to be too narrow.

The impact of the survey design on variance estimates is measured by the design effect. The design effect describes the variance of the sample estimate for the survey relative to the variance of an estimate based on a hypothetical random sample of the same size. In situations where statistical software packages assume a simple random sample, the adjusted standard error of a statistic should be calculated by multiplying by the design effect. Each variable will have its own design effect. Average design effects are summarized below. For respondents (adults), the average design effect for estimates for the target person in the household is 2.78.

Consideration was made with regard to trimming the final weight. While the ratio between largest and smallest weight is somewhat significant, the number of cases with very large and small weights are small. A trimming procedure to reduce the ratio from largest to smallest weight to 20 resulted in an overall design effect of 2.4. Given that this reduction is not great and the desire to have a single final weight with as little bias as possible, the decision was made to not trim to final weight.

At a final design effect of 2.78, margins of error for the full sample are inflated by 1.67 (the square root of the design effect). Therefore, a perfectly self-weighting sample (a sample that requires no weight at all) would attain margins of error of +/- 1.335% for an estimated percentage level of 50% at 95% confidence. The study in turn attained a margin of error of +/-2.225%.

TABLE 22: RESPONDENT (ADULTS) DESIGN EFFECTS

Race/Ethnicity	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
White Non-Hispanic	37.6%	1.1%	35.6%	39.7%	2.5	2,068
Black Non-Hispanic	17.4%	0.8%	15.8%	19.0%	2.3	966
Hispanic	36.8%	1.1%	34.6%	39.0%	2.8	1,246
Asian Non-Hispanic	4.9%	0.4%	4.2%	5.7%	1.5	704
Other Non-Hispanic	2.6%	0.5%	1.8%	3.8%	4.8	100
Don't know	0.2%	0.1%	0.1%	0.5%	2.3	6
Refused	0.5%	0.1%	0.3%	0.8%	1.9	26
Education						
No H.S. Diploma	21.3%	1.6%	18.3%	24.6%	3.1	215
H.S. Diploma	28.4%	1.7%	25.1%	31.8%	2.9	401
Some College	25.6%	1.6%	22.7%	28.8%	2.6	401
College Degree	24.7%	1.2%	22.5%	27.0%	1.4	985
DK/Ref	1.20%	0.20%	0.80%	1.80%	2.7	57
Age						
18 thru 24	13.0%	1.1%	10.9%	15.3%	5.7	237
25 thru 34	21.3%	1.0%	19.4%	23.3%	3.0	860
35 thru 44	21.8%	0.9%	20.0%	23.6%	2.6	1,018
45 thru 54	19.9%	0.9%	18.3%	21.6%	2.4	1,082
55 thru 64	13.3%	0.6%	12.1%	14.6%	1.7	1,062
65 thru 96	10.5%	0.5%	9.5%	11.5%	1.4	840
DK/Ref	0.3%	0.1%	0.1%	0.7%	3.3	17
Gender						
Male	49.5%	1.2%	47.2%	51.8%	2.8	1,874
Female	50.5%	1.2%	48.2%	52.7%	2.8	3,239
DK/Ref	0.0%	0.0%	0.0%	0.2%	1.1	3
Home Ownership						
Rent	62.5%	1.1%	60.3%	64.7%	2.8	3,344
Own	32.4%	1.1%	30.3%	34.6%	2.8	1,540
Other Arrangements	2.90%	0.40%	2.20%	3.80%	3.0	122
DK	1.70%	0.30%	1.10%	2.40%	3.2	65
Ref	0.50%	0.10%	0.40%	0.80%	1.3	45
SuperPUMA						
48181	13.3%	0.5%	12.3%	14.4%	1.2	656
48182	13.8%	0.5%	12.8%	14.9%	1.2	724
48183	14.6%	0.6%	13.5%	15.8%	1.5	895
48184	13.0%	0.6%	11.9%	14.1%	1.4	770
48185	14.9%	0.7%	13.5%	16.4%	2.1	659
48186	14.5%	0.5%	13.5%	15.6%	1.2	692
48187	15.9%	0.6%	14.7%	17.1%	1.4	720

TABLE 23: CHILD DESIGN EFFECTS

Race/Ethnicity	Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
			Lower	Upper		
White Non-Hispanic	22.9%	1.2%	20.6%	25.5%	1.2	387
Black Non-Hispanic	17.6%	1.3%	15.2%	20.4%	1.7	239
Hispanic	52.7%	1.7%	49.5%	56.0%	1.5	506
Asian Non-Hispanic	2.6%	0.4%	1.9%	3.6%	0.9	105
Other Non-Hispanic	1.6%	0.6%	0.7%	3.5%	3.5	23
Asian Vietnamese	1.7%	0.3%	1.2%	2.3%	0.6	112
Don't know	0.2%	0.2%	0.1%	1.0%	1.7	2
Refused	0.5%	0.3%	0.2%	1.6%	2.2	4
Gender						
Male	50.9%	1.9%	47.3%	54.5%	1.9	679
Female	48.7%	1.9%	45.1%	52.3%	1.9	694
DK/Ref	0.4%	0.2%	0.1%	1.1%	1.4	5
SuperPUMA						
48181	13.7%	0.6%	12.4%	15.0%	0.5	161
48182	12.3%	0.7%	11.1%	13.7%	0.6	182
48183	10.6%	0.5%	9.7%	11.6%	0.4	208
48184	13.2%	0.6%	12.2%	14.4%	0.4	192
48185	15.6%	0.9%	14.0%	17.5%	0.8	210
48186	16.6%	0.7%	15.4%	18.0%	0.4	203
48187	17.9%	0.8%	16.4%	19.5%	0.6	222

Variance estimation procedures have been developed for most standard software packages to account for complex survey designs. We provide a replicate stratum (strata) on the survey data files that can be used with the appropriate weight variable to obtain corrected standard errors using a Taylor series approximation (or other related linearization method). Users interested in using a linearization method can choose to use SUDAAN, the "SVY" commands in Stata, the "PROC SURVEYMEANS" and "PROC SURVEYREG" commands in SAS, or the "CSELECT" complex samples procedures in the SPSS complex samples module.

6. REFERENCES

Acock, A. C. 2005. "Working With Missing Values." *Journal of Marriage and Family* 67:1012-1028.

Allison, P. D. 2001. Missing Data. Thousand Oaks, CA: Sage.

Blumberg, S.J., Luke J.V., Ganesh N., et al. (2011). Wireless substitution: State-level estimates from the National Health Interview Survey, January 2007–June 2010. National Health Statistics Reports; no 39. Hyattsville, MD: National Center for Health Statistics.

Brick, J.M., and Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, 5, 215-238.

Collins, L. M., J. L. Schafer, and C. M. Kam. 2001. "A Comparison of Inclusive and Restrictive Strategies in Modern Missing Data Procedures." *Psychological Methods* 6:330-351.

Deming, W.E., and Stephan, F. F. (1943) Statistical Adjustment of Data. Dover Publications.

Graham, J.W. and B.J. Taylor, A.E. Olchowski, & P.E. Cumsille. 2006. Planned Missing Data Designs In Psychological Research. *Psychological Methods* 11:323-343.

Groves, R. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70 (5): 646-675.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47*, 153-161.

Horton, N. J., S. R. Lipsitz, and M. Parzen. 2003. "A Potential for Bias When Rounding in Multiple Imputation." *The American Statistician* 57:229-233.

Johnson, D. R. and R. Young. 2009. “Improving the Utility of Imputed Values in Survey Datasets.” In JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Johnson, D. R. and R. Young. 2011. “Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations.” *Journal of Marriage and Family*, Forthcoming.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., and Presser, S. (2000) Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly*, 64 (2): 125-148.

Kenward, M. G., & Carpenter, J. (2007). Multiple Imputation: Current Perspectives. *Statistical Methods in Medical Research, 16*, 199-218.

Lee, K. J. 2010. Multiple Imputation for Missing Data: Fully Conditional Specification versus Multivariate Normal Imputation. *American Journal of Epidemiology* 171:1-9.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis With Missing data*. Hoboken: John Wiley & Sons, Inc.

Marker, D. A., Judkins, D. R., & Winglee, M. (2002). Large Scale Imputation for Complex Surveys. *Survey Non-Response*. New York: Wiley.

Puhani, P. A. (2000). The Heckman Correction for Sample Selection and its Critique. *Journal of Economic Surveys, 14*(1), 53-67.

Raghunathan, T. E. (2004). What Do We Do With Missing Data? Some Options for Analysis of Incomplete Data. *Annual Review of Public Health, 25*, 99-117.

Reiter, J. P., Raghunathan, T. E., & Kinney, S. (2006). The Importance of Modeling the Sampling Design in Multiple Imputation for Missing Data. *Survey Methodology, 32*(2), 143-149.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1995). Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association, 90*(106-129).

Royston, P. 2005. "Multiple Imputation of Missing Values." *Stata Journal* 4:227-241.

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika, 36*, 581-592.

Rubin, D. B. 1987. Multiple Imputation for Nonresponse in Surveys. New York: Chichester.

Schafer, J. L., & Graham, J. W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods, 7*(2), 147-177.

Schafer, J. L., & Olsen, M. K. (1998). Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. *Multivariate Behavioral Research, 33*, 545-571.

Scharfstein, D. O., Rotnitzky, A., & Robins, J. M. (1999). Adjusting for Nonignorable Drop-Out Using Semi-Parametric Nonresponse Models (with comments). *Journal of the American Statistical Association, 94, 1096-1146.*

The American Association for Public Opinion Research. 2008. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 6th edition. AAPOR.

Werner, J. (2004). QBAL: A Professional Sample Balancing Program, User's Guide Revision: 04.1.27. Jan Werner Data Processing.

Yucel, R. M., He, Y., & Zaslavsky, A. M. (2008). Using Calibration to Improve Rounding in Imputation. *The American Statistician*, 62(2), 125-129.