Detailed Survey Methods

The University Center for Social and Urban Research (UCSUR) conducted the 2014 State of Aging Survey in collaboration with The University of Pittsburgh Aging Institute, University of Pittsburgh Medical Center, The Allegheny County Area Agency on Aging, and the United Way of Allegheny County. The survey was part of a larger study of the older adult population (age 55 and up) in Allegheny County that includes secondary data analyses and will serve as an update of the earlier *State of Aging and Health in Pittsburgh and Allegheny County, 2003* report. The 2014 survey asked questions about a broad range of topics related to quality of life for older adults in the region. Topics included work and retirement status, retirement confidence, sources of income, living arrangements and mobility/migration, neighborhood issues, transportation, health status, health-related quality of life, functional status and disability, health care access, health behaviors, cognitive function, anxiety and depression, social health and social support, caregiving, volunteering, service use, religiosity, internet usage, and socio-demographics. The survey was conducted by telephone using random digit dialing (RDD) sampling of both landline and cellular telephones with screening for adults age 55 and older between January and April, 2014. Details on survey methodology are provided in this appendix.

Sample design. The target population for the survey was non-institutionalized English-speaking adults age 55 and older living in Allegheny County (population of 367,764 according to the 2010 Decennial Census [out of the total county population of 1,223,348]). The goal was to complete at least 1,000 total surveys. The design involved over-sampling of African-Americans in order to allow more stable estimates for this sub-population and more reliable racial comparisons. According to the 2010 Census, African-Americans make up 9.6% of the Allegheny County population age 55 and older (non-Hispanic Whites account for 88.5%), and the goal was to complete approximately 250 (or 25%) surveys with African-Americans. Thus, a disproportionate stratified sample design, described more fully below, was used. The target of 250 African-Americans represented a compromise between maximizing the absolute number for analysis while minimizing sampling variance and "design effects" associated with over-sampling (discussed further below). Another goal of the survey was to have sufficient cases (at least 400) to conduct separate analyses of the age 55-64 sub-group. This did not require over-sampling as this younger age group makes up approximately 45% of the total population over age 55 in Allegheny County according to the 2010 Census.

RDD telephone samples of both landline and cellular telephones were drawn to conduct the survey. The sample was purchased from Survey Sampling International (SSI), one of the major survey sample providers in the U.S. Landline RDD sampling involves generation of random phone numbers in a defined geographic area - both listed and unlisted - and was the standard method for obtaining representative samples of households by telephone until the last few years. Landline RDD samples allow targeting of demographic groups – e.g., African-Americans in this survey – by linkage of area codes/telephone exchanges (first three digits of phone #) to recent Census data. The dramatic increase in the use of cell phones over the past ten years – current estimates are that approximately 39% of all U.S. households are cell only – has made incorporation of cell phones into RDD designs standard practice to avoid coverage error. However, among those 65 and older, the number is only 12.6%. Cell phones are assigned unique area code/exchanges – which represent where the phone was purchased - which allows separate sampling of landline and cell phones. Landline and cell telephone frames are defined as separate strata in these designs ("dual-frame" landline-cell phone samples). It is important to note that screening for geography of residence (in this case Allegheny County) is very important for cell phone samples, since the respondent may live in a different area than where the phone was purchased. Also, cell phone samples do not allow demographic targeting for African-Americans via Census linkage like landline samples. Some households/individuals have both landline and cell phones, and are thus contactable using either sample frame. This was measured in the survey and adjusted for through the sampling weights, which are

described below. Given that the survey targeted older adults (55+) with lower cellphone-only rates, the goal was to complete 10-15% of the interviews by cell phone.

More specifically, sampling was done from 4 mutually exclusive and exhaustive strata which cover landline and cell telephones in Allegheny County:

1. *Landline - Low density African-American* (defined as telephone exchanges with 10% or lower African-American population according to Census data)
2. *Landline - Medium density African-American* (defined as telephone exchanges with between 11 and 29% African-American population according to Census data)
3. *Landline - High density African-American* (defined as telephone exchanges with 30% or higher African-American population according to Census data)
4. *Cell Phone stratum* (does not allow racial targeting)

Within all four strata, screening was done for individuals age 55 and older. For the landline strata, households not containing anyone 55 and older were screened out as ineligible. The cell phone is considered an individual device for sampling - in that stratum, screening was done of the individual answering the phone. If under age 55, they were screened out as ineligible. Within the landline strata, a respondent was randomly selected from multiple eligible adult households (i.e., more than one 55 or older) using the most recent birthday method. No within household selection was done for the cell phone sample, which was considered an individual device. Probabilities of selection varied across strata (e.g., higher probabilities in strata 2 and 3 for African-American over-sample), which are adjusted for in the sample weights discussed below. It should be noted that two potential sources of coverage error – population members having no chance of being included in the sample - are present with this design: (1) households with no telephone service (estimated at approximately 2% of households in the U.S.), and (2) individuals who've recently moved to the region, are cellphone-only, but purchased their cell phone outside the region (no estimate available, but likely to be fairly small). Individuals who were reached who had purchased their cell phone in the region but have since moved out of the region were screened out of the survey.

Data collection and response rates. Data were collected in UCSUR's computer-assisted telephone interviewing (CATI) lab by trained interviewers between January 17 and April 9, 2014. CATI involves programming the survey instrument using standard software that displays the questions on the interviewer's computer screen in proper order with automated skip patterns. Data are entered into the system as answers are provided. UCSUR telephone interviewers receive rigorous training in survey interviewing techniques and are continuously monitored during data collection for quality control purposes. Each sampled telephone number was called up to six times on different days of the week at different times of the day with evenings and weekends emphasized to maximize probability of contact. Once an individual or household was contacted, as many calls as necessary were made to either complete the survey or obtain a refusal. The CATI system also automates call scheduling and callbacks. A total of 1,049 interviews were completed, including 254 African Americans, thus approaching our 25% target. Over 70% (185, 72.8%) of the African-American respondents came from the high density landline stratum, with another 19.3% (49) from the medium density landline stratum. Interviews took an average of about 60 minutes to complete. A $15 debit card was offered as an incentive for completing the survey. The completed surveys included 397 with 55-64 year olds, 359 with ages 65-74, and 288 with ages 75 and older (5 respondents refused to report an exact age). Thus, the older age group was somewhat over-represented in the total sample (62% versus 55% of the population). In terms of landline versus cell phone interviews, 1,003 of the interviews were conducted on landlines (96%) and only 46 on cell phones (4%, including 13 African-Americans), well below the original 10-15% cell target. Early in the field period, we found that the cell sample was costly and inefficient, with each completed interview taking 2 to 3 times as much interviewer effort (approximately 8 hours) to complete. Each completed interview

took an average of 63 dialings for landline sample versus 264 dialings for each completed cell interview. We believe this was due to (1) the need to screen for those 55 and older and (2) the length of the interview (60 minutes is much longer than recommended for cell surveys). Due to practical and cost considerations, we conducted the remainder of the interviews on landline phones. Prior to conducting any cell phone interview, the respondent was asked to confirm that they were in a safe place (i.e., not driving) to answer the questions. However, the survey did not require that the cell phone respondents be at home while answering the questions.

Landline response rate. Of the initial 16,991 landline phone numbers put into the system, 4,339 were determined to be non-households (businesses, other institution, disconnected, non-working numbers), and thus out of scope. Of the remaining 12,652 numbers, the following outcomes were achieved: 829 were screened not eligible (i.e., no one 55 or over in household); 1003 completed interviews; 272 eligible respondents refused to participate; 187 eligible respondents could never be reached to complete the interview, 4,465 known households were not screened; and we were unable to complete screening interviews at 5,896 numbers due to multiple no answers/answering devices/busy signals (i.e., household status unknown). Using the proportion of households found among phone numbers where household status was determined as the "e" multiplier for the 5,896 unknown household status numbers (e = .586), we calculated an AAPOR #3 landline screening rate of 22.4%. The interview completion rate [1003/(1003+272 + 187)] was 68.6%, for an overall *landline AAPOR RR3 (response rate) of 15.4%* (.224 X .686).

Cell response rate. Of the initial 3,093 cell phone numbers put into the system, 516 were determined to be non-households (businesses, other institution, disconnected, non-working numbers), and thus out of scope. Of the remaining 2,577 numbers, the following outcomes were achieved: 398 were screened not eligible (i.e., not 55 or over); 46 completed interviews; 20 eligible respondents refused to participate; 8 eligible respondents could never be reached to complete the interview, 618 refused to be screened; and we were unable to complete screening interviews at 1,487 numbers due to multiple no answers/answering devices/busy signals (i.e., individual eligibility status unknown). Using the proportion of potentially eligible individuals found among the remaining phone numbers as the "e" multiplier for the 1,487 unknown status numbers (e = .428), we calculated an AAPOR #3 cell phone screening rate of 27.3%. The interview completion rate [46/(46+20+8)] was 62.2%, for an overall *cell phone AAPOR RR3 (response rate) of 17.0%* (.273 X .622).

Overall response rate. Of the initial 20,084 phone numbers put into the system, 4,855 were determined to be non-households (businesses, other institution, disconnected, non-working numbers), and thus out of scope. Of the remaining 15,229 numbers, the following outcomes were achieved: 1,227 were screened not eligible (i.e., no one 55 or over); 1049 completed interviews; 292 eligible respondents refused to participate; 195 eligible respondents could never be reached to complete the interview, 5,083 known households/individuals were not screened; and we were unable to complete screening interviews at 7,383 numbers due to multiple no answers/answering devices/busy signals (i.e., status unknown). Using the proportion of households/individuals found among phone numbers where status was determined as the "e" multipliers for the 7,383 unknown status numbers, we calculated an AAPOR #3 overall screening rate of 23.1%. The interview completion rate [1049/(1049+292 + 195)] was 68.3%, for an *overall AAPOR RR3 of 15.8%* (.231 X .683).

Although these may seem low, they are comparable to current response rates obtained by similar survey organizations using similar methods in other studies. However, the completion rates (close to 70%) among those screened eligible are fairly high, showing that most of the non-response was due to an inability to successfully screen households and individuals. The rates are also higher than response rates obtained using standard 3-4 day political polling methodology, which process two to three times as many

telephone numbers to complete the same number of surveys using limited callbacks. We also applied standard weighting techniques to the data in an attempt to partially adjust for non-response (see below).

Statistical weighting. A two-step statistical weighting process was used in which each completed case was adjusted for (1) probability of selection, and (2) post-stratification on sex, age, race, education, employment status, and living arrangements. These steps are described in more detail here.

1. *Probability of selection weight*. This contained 3 components: (a) initial probability of selection of the telephone number, which varied across strata; (b) # adults in the household (landline sample only; weight equal to number of adults); and (c) telephone ownership status - those reachable by both cell and landline are given a weight of 0.5 at this stage given twice the probability of selection; those with cell only or landline only are given a weight of 1. The initial probability of selection (step a) was highest for the high density African-American landline stratum used for over-sampling (these cases were down-weighted), while the probabilities of selection were lowest for the cell phone stratum (these cases were up-weighted). The medium density (slight down-weighting) and low density (slight up-weighting) African-American landline strata had intermediate probabilities of selection and thus were assigned less extreme initial weights.
2. *Post-stratification raking adjustment*. To further adjust for survey non-response, we used an iterative proportional fitting "raking" algorithm ("ipfweight") using the STATA statistical package. The algorithm performs a stepwise adjustment of survey sampling weights (starting with the probability of selection weight from step 1) to achieve known population distributions on the variables included. In this case, the 2012 American Community Survey (ACS) data for the 55 and older population of Allegheny County were used to obtain the population totals. Separate raking models were run for the 55-64, and 65 and older age groups, as most of the survey estimates are reported separately for those groups. The 65 and older group raking model included sex, age (65-74, 75+), education (high school graduate or less, some college, bachelor's degree or higher), race (non-Hispanic White, other), currently employed (yes, no), and living alone (vs. with others). The 55-64 age group raking model included the same variables minus age. Raking involves an attempt to adjust the weights in order to make the survey distributions on the included variables "mirror" the population to the greatest extent possible. That is, the final weighted marginal distributions on sex, age, education, race, employment status, and living arrangements closely match the marginal distributions from the 2012 Allegheny County ACS estimates for the 55 and older non-institutionalized population. This standard survey methodology is a way to statistically adjust survey estimates in order to increase accuracy and reduce bias due to differential non-response across demographic sub-groups.

The probability of selection and post-stratification weights were combined to produce the final case weight, which is used for all of the estimates in this report. Upper values of the final weight variable were trimmed to a maximum value of 5 and lower values to a minimum of .20 to reduce variance. The 10^{th} and 90^{th} percentiles were 0.20 and 2.48 respectively. The 25^{th} percentile of the weighting variable was 0.21 and the 75^{th} percentile was 1.24.

Precision of the survey estimates. Approximate 95% confidence interval margins of error for the survey estimates for the entire sample, age 55-64, age 65 and over, and African-American sub-samples are presented in the table below. Since the sample design was not a simple random sample – we used disproportionate stratified sampling requiring sample weights – the complex sample design results in a loss of precision (i.e., "design effects"). The design effects in the survey (which are calculated separately for each survey estimate) ranged from 1.2 to 1.8, with the typical design effect being about 1.5. This is an estimate of the variance of our complex sample design parameter estimates to the variance we would have obtained from a simple random sample of the same size. This typical design effect is taken into account

for the following margin of error estimates. The table shows margins of error for different proportions. Note that error is greatest for proportions at .5 and gets smaller as the estimate approaches 0.

Sub-Group	**Sample size**	**95% C.I. margin of error (+ / -) (adjusted for design effect = 1.5)**		
		p = .50	**p = .30**	**p = .10**
Allegheny county age 55+	1,049	4.5%	4.2%	2.7%
Age 55-64	397	7.4%	6.8%	4.4%
Age 65 and over	652	5.7%	5.3%	3.4%
African-Americans (any age)	254	9.2%	8.4%	5.5%

However, it is important to note that surveys are subject to additional non-sampling errors, including those due to non-response and measurement errors, which are not accounted for in the margin of error estimates. These should be taken into account when interpreting these and any other survey data.