

Many Voices One Valley 2007

How the Survey Was Conducted

The survey was conducted by the Marist College Institute for Public Opinion and funded by the Dyson Foundation.

4,320 adults eighteen years of age or older who live in the Mid-Hudson Valley in New York were interviewed in proportion to the adult population in seven counties: Columbia, Dutchess, Greene, Orange, Putnam, Sullivan, and Ulster. The survey was designed to collect information about the Mid-Hudson Valley as a region and to allow for county level analysis, as well. Representative samples were drawn from each county based on the population distribution within that county. In order to analyze the opinions of residents of the Mid-Hudson Valley, the seven county samples were combined and weighted to reflect the population distribution of the entire region.

	2007 Survey	2005 Census
County		
Columbia	4%	5%
Dutchess	27%	26%
Greene	3%	3%
Orange	34%	34%
Putnam	9%	9%
Sullivan	7%	7%
Ulster	16%	16%
Gender		
Men	49%	49%
Women	51%	51%
Race		
White non-Latino	80%	80%
African American	7%	7%
Latino	10%	10%
Income		
Less than \$15,000	10%	11%
\$15,000 to \$24,999	9%	9%
\$25,000 to \$49,999	23%	23%
\$50,000 to \$74,999	20%	20%
\$75,000 to \$99,999	14%	14%
\$100,000 or more	24%	23%
Age		
18 to 34	26%	27%
35 to 54	42%	42%
55 and over	32%	31%

	2007 Survey	2000 Census
Columbia County		
Gender		
Men	49%	50%
Women	51%	50%
Race		
White non-Latino	91%	92%
Non-white	9%	8%
Income		
Less than \$50,000	59%	59%
\$50,000 or more	41%	41%
Age		
Under 45	44%	44%
Over 45	56%	56%

	2007 Survey	2005 Census
Dutchess County		
Gender		
Men	49%	49%
Women	51%	51%
Race		
White non-Latino	80%	79%
Non-white	20%	21%
Income		
Less than \$50,000	40%	40%
\$50,000 or more	60%	60%
Age		
Under 45	45%	48%
Over 45	55%	52%
	2007 Survey	2000 Census
Greene County		
Gender		
Men	52%	52%
Women	48%	48%
Race		
White non-Latino	89%	88%
Non-white	11%	12%
Income		
Less than \$50,000	68%	66%
\$50,000 or more	32%	34%
Age		
Under 45	46%	47%
Over 45	54%	53%
	2007 Survey	2005 Census
Orange County		
Gender		
Men	50%	49%
Women	50%	51%
Race		
White non-Latino	73%	72%
Non-white	27%	28%
Income		
Less than \$50,000	40%	40%
\$50,000 or more	60%	60%
Age		
Under 45	46%	52%
Over 45	54%	48%

	2007 Survey	2005 Census
Putnam County		
Gender		
Men	49%	49%
Women	51%	51%
Race		
White non-Latino	86%	86%
Non-white	14%	14%
Income		
Less than \$50,000	23%	24%
\$50,000 or more	77%	76%
Age		
Under 45	43%	48%
Over 45	57%	52%
	2007 Survey	2005 Census
Sullivan County		
Gender		
Men	50%	50%
Women	50%	50%
Race		
White non-Latino	80%	80%
Non-white	20%	20%
Income		
Less than \$50,000	51%	52%
\$50,000 or more	49%	48%
Age		
Under 45	43%	45%
Over 45	57%	55%
	2007 Survey	2005 Census
Ulster County		
Gender		
Men	49%	49%
Women	51%	51%
Race		
White non-Latino	86%	87%
Non-white	14%	13%
Income		
Less than \$50,000	46%	49%
\$50,000 or more	54%	51%
Age		
Under 45	40%	45%
Over 45	60%	55%

Residents of Sullivan County are included in the 2007 study of the region even though they were not interviewed as part of the 2002 study. Comparisons between the 2007 and the 2002 studies may still be made to the region as a whole for two reasons. First, residents of Sullivan County generally express a similar pattern in their responses as other residents of the region. Second, Sullivan County makes up only 7% of the overall population of the region and, therefore, the survey results, as well.

The goal of a scientifically designed survey sample is to be representative of the population that is being surveyed. The results obtained from a scientific probability survey are not just answers from those individuals who responded but more importantly, because of the design and methods by which the data is collected, can be used to generalize to the population as a whole. For this survey, the results are an estimate of what would have been obtained, within a certain range, if all adults eighteen years of age or older in the Mid-Hudson region were interviewed.

How to interpret the numbers

Sampling error is the difference between the responses to the survey if all adults eighteen years of age or older who live in the Mid-Hudson Valley had been interviewed and the actual survey results. Sampling error is primarily based upon the number of interviews in the survey sample. The sampling error may be interpreted as indicating the probability (95 times out of 100) within which the results of repeated samplings, in the same time period, assuming the same sampling procedures,

could be expected to fall within a certain range.

The margin of error for the survey results of Mid-Hudson residents is $\pm 1.5\%$ for percentages near 50% at a confidence level of 95%. The sampling error diminishes slightly for questions whose results are at the extremes and the sampling error increases as the number of interviews for a particular group or sub-group within the sample declines.

	Sample Size	Margin of Error
Mid-Hudson	4320	$\pm 1.5\%$
Columbia	504	$\pm 4.5\%$
Dutchess	827	$\pm 3.5\%$
Greene	503	$\pm 4.5\%$
Orange	956	$\pm 3.5\%$
Putnam	501	$\pm 4.5\%$
Sullivan	502	$\pm 4.5\%$
Ulster	527	$\pm 4.5\%$

For example, 52% of Mid-Hudson Valley residents surveyed think their community spends too little money on public transportation. We may conclude that there is a high probability, 95 times out of 100, that the average results for this question of repeated samplings in the region will fall between 53.5% and 50.5%, $\pm 1.5\%$. All results are rounded to the next whole number.

In evaluating the information detailed in this report, several points should be noted. First, the analysis often focuses on differences among subgroups rather than commonalities. Therefore, the charts depicted in the report do not include all factors but rather those where differences are evident. Tables in the appendix may be referenced for further detail among

subgroups and other key segmenting variables.

Second, some questions in the survey were only asked of a subgroup of those interviewed depending on an answer to a previous question. For instance, residents were asked if they were currently employed. If they were not, they were not asked subsequent questions about their job such as commute time. Questions that were not asked of all residents are identified.

Finally, totals in tables may not add to 100% due to rounding.

Methodology: sample design

A stratified random digit dial (RDD) probability design was used to draw the telephone numbers for the survey. RDD ensures representation of both listed and unlisted telephone numbers. Telephone numbers were selected based upon a list of telephone exchanges from throughout the seven counties. The exchanges were selected to ensure that each county was represented in proportion to the adult population. The telephone numbers were obtained from Survey Sampling International in Fairfield, Connecticut. The sample file was electronically matched after selection to the yellow pages business directory and screened for business and/or disconnected numbers.

In order to participate in the survey a household had to have a residence located within one of the seven counties and have an adult resident that was eighteen years of age or older. A member within each household was then selected to be

interviewed through the use of a random household selection method. In order to be interviewed, a respondent needed to be at least eighteen years of age or older and be a resident of the Mid-Hudson Valley.

Methodology: data collection

The questionnaire and the telephone sample were programmed for computer assisted telephone interviewing (CATI). Interviewing was conducted from a centralized telephone facility at the Marist College Institute for Public Opinion using trained interviewers who were specifically briefed on this study. The interviews were conducted in either English or Spanish.

There were a total of 151 interviewers and supervisors who worked on this study. On average, the supervisor to interviewer ratio was seven to one. Polling supervisors regularly monitored, evaluated, and provided feedback to the interviewing staff throughout the data collection period. The average length of the survey was nineteen minutes.

Eighty-seven pretest interviews for the survey were administered by telephone on April 9th, 2007. As a result of the pretest, the questionnaire was updated and revised. Due to the limited nature of the changes, pretest interviews were included as part of the final dataset.

Additional interviews were conducted April 10th through June 19th, 2007. Interviewers contacted households between 5:15 p.m. and 9:30 p.m. Callbacks were also conducted between 9:00 a.m. and 5:00 p.m. on weekdays. Up to eighteen

attempts to contact a household were made before a telephone number was eliminated from the sample. Callbacks were staggered over times of day and days of the week to maximize the chances of making contact with a potential respondent. Suspended interviews and refusals were re-contacted at least once in order to attempt to convert them to a completed interview. This included callbacks to initial unsuccessful contacts, refusal conversion, and scheduled appointments. A toll free number was used so that respondents could call back the survey center at their convenience to complete an interview. 2%, or seventy-three respondents, completed a survey interview by contacting the survey center using the toll free number.

Information collected from survey participants is both confidential and anonymous. Personal identifying information is removed from files after the integrity of the data has been verified.

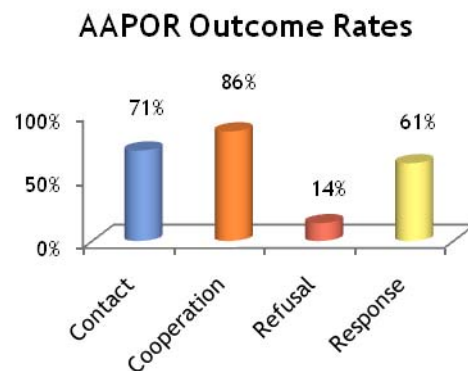
Outcome rates

Outcome rates were computed using Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys published by the American Association of Public Opinion Research (AAPOR), 2006.¹

- The contact rate is the proportion of all cases in which some responsible household member is reached.
- The cooperation rate is the proportion of all cases that complete an interview of all eligible households that are contacted.

¹ This report may be found at: http://www.aapor.org/uploads/standarddefs_4.pdf

- The refusal rate is the proportion of all cases in which a household member or the respondent refuses to be interviewed, or breaks-off an interview, of all potentially eligible cases.
- The response rate is the number of complete interviews with eligible adults divided by the number of eligible households in the sample. The response rate is calculated by multiplying the contact rate by the cooperation rate.



For this survey, the contact rate is 71%, meaning that for about seven in ten phone records it was established that the phone number was a valid household phone number. The remaining 29% of records include phone numbers that were not verifiable, those that were attempted multiple times at a maximum of eighteen calls without confirmation that there was an adult age eighteen or older living in the household, or bad numbers such as fax or data lines, disconnected or non-working, or business numbers.

There was a very high cooperation rate, 86%. This means that for those records where contact was made with an adult age eighteen or older in the household, there was a very high chance of completing a survey. This was accompanied by a refusal rate of 14%. In order to compute the response rate, the contact rate, 71%, is multiplied by the cooperation rate, 86%. The total response rate for the survey is 61%.