

Background

The 2002 study was a joint project of the Dyson Foundation and the Marist College Institute for Public Opinion. It was conducted to better understand residents' perceptions about the present and future direction of the Hudson Valley. The goals were to assess the needs and priorities for the Hudson Valley through the opinions of both its residents and its community leaders, and to identify areas of shared and divergent perspectives. The results can be used as a tool to inform the public and those who help guide the region's future about the interests, concerns, and needs of the area. The results may also be a resource for other organizations and agencies within the region to identify and quantify community need in specific areas. Such documentation provides a value to the region and to those groups and individuals that work to make their community a better place to live and work.

Two surveys were conducted. The first survey interviewed 3,882 residents in the six counties that define the Hudson Valley region. The counties included Columbia, Dutchess, Greene, Orange, Putnam, and Ulster. The survey was designed to collect information about the 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: Columbia County included 500 interviews; Dutchess County included 830 interviews; Greene County included 502 interviews; Orange County included 1,027 interviews; Putnam County included 503 interviews; and Ulster County included 520 interviews. In order to analyze the opinions of residents of the Hudson Valley the six county samples were combined and weighted to reflect the population distribution of the entire region.

A second companion survey was conducted with 245 community leaders within the same six counties. This scientifically selected sample of community leaders included influentials and decision-makers who have a role in how community priorities are identified and realized, and how resources, goods, and services are distributed in their community. Included in the sample were government and elected officials, heads of major institutions, prominent members of the media, and business executives. The types of leaders who participated included: members of congress, state senators, county legislators and commissioners, government officials and administrators, assemblypersons, town supervisors and mayors, village trustees, common council presidents, and planning board chairs; college presidents, superintendents of schools, school principals and school board presidents; owners, publishers, editors and general managers at radio stations

and newspapers; executive directors at non-profit organizations (including the arts and culture, advocacy groups for women and minorities, children and child care, services for the disabled, disaster support, domestic violence, education, the environment, and health care); heads and high-ranking members of economic development corporations, chambers of commerce, and tourism offices; CEOs, presidents, and owners of businesses (including, but not limited to manufacturing, transportation, hospitality services and tourism, retail and wholesale sales and production, banks, credit unions, insurance, financial services, utilities, environmental, construction, real estate, and technology); community activists and members of corporate and community boards; hospital presidents and executive directors for health care providers; pastors, reverends, and rabbis.

As in the survey of residents, the design takes into account an interest in understanding the opinions of community leaders from throughout the Hudson Valley and from within each county. 24 community leaders from Columbia County were interviewed; 63 from Dutchess County; 25 from Greene County; 68 from Orange County; 22 from Putnam County; and 43 from Ulster County. The opinions of community leaders for the Hudson Valley as a whole were drawn from the combined six county samples and weighted to reflect the population distribution of the region.

	US Census Population 2000	Residents	Community Leaders
	Column %	Column %	Column %
Columbia County	6%	6%	6%
Dutchess County	28%	28%	28%
Greene County	5%	5%	5%
Orange County	34%	34%	34%
Putnam County	9%	9%	9%
Ulster County	18%	18%	18%

How to Interpret the Numbers

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 the survey of Hudson Valley residents, the results are an estimate of what would have been obtained, within a certain range, if all households throughout the six county region were interviewed. Similarly, the results for the survey of community leaders is an estimate of what would have been obtained, within a certain range, if all community leaders throughout the Hudson Valley were interviewed.

When analyzing the survey results, it should be kept in mind that in all surveys each result is an estimate of what would have been obtained had everyone in the eligible population been interviewed. For example, the difference between the

responses if all households within the Hudson Valley had been interviewed and the survey results is referred to as sampling error and is primarily based upon the number of interviews in the survey sample.

The sampling error for the results of the survey of Hudson Valley residents, 3,882 interviews, is $\pm 2\%$ for percentages near 50% at a confidence level of 95%. The sampling error for the results of the survey of Hudson Valley community leaders, 245 interviews, is $\pm 6.5\%$ for percentages near 50% at a confidence level of 95%. 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 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.

For example, 49% of Hudson Valley residents surveyed think their community spends too little money on recreational activities for children and teens. We may conclude that there is a high probability (95 times out of 100) that the average results for this question of repeated samplings of households in the Hudson Valley will fall between 47% and 51%, ($\pm 2\%$).

Sampling error increases as the number of interviews within a sample declines. For instance, when looking at the results for the 245 community leaders that were asked the same question the sampling error increases. 48% of community leaders surveyed think their community spends too little money on recreational activities for children and teens. We may conclude that there is a high probability (95 times out of 100) that the average results for this question of repeated samplings of community leaders in the Hudson Valley will fall between 41.5% and 54.5%, ($\pm 6.5\%$).

The sampling error for the survey of residents at the county level is $\pm 4.4\%$ for Columbia County, $\pm 3.4\%$ for Dutchess County, $\pm 4.4\%$ for Greene County, $\pm 3\%$ for Orange County, $\pm 4.4\%$ for Putnam County, and $\pm 4.3\%$ for Ulster County. When assessing the results for community leaders at the county level, the findings should be considered as a cross-section of those views.

Methodology

Sample Design: Residents

As described above, the Hudson Valley is defined as the six counties of Columbia, Dutchess, Greene, Orange, Putnam, and Ulster. A stratified random digit dial (RDD) probability design was used to draw the telephone numbers for the survey of residents. RDD ensures representation of both listed and unlisted telephone numbers. The first eight digits of the sample telephone numbers (which includes the area code, telephone exchange and the first two digits of a phone number within the exchange) are selected based upon estimates of telephone households

within each exchange and the proportion of the adult population that makes up each county. Households were selected at random in proportion to the adult population in each county. Survey Sampling Inc. in Fairfield, Connecticut, provided the telephone numbers according to the above-described design.

In order to be interviewed, a household had to have a residence located within one of the six counties, and have an adult resident that was 18 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.

Sample Design: Community Leaders

As noted, the goal of the companion survey of community leaders is to understand the opinions and priorities of community leaders from throughout the Hudson Valley. The objective of the sampling design was to develop a representative list of community leaders from throughout the Hudson Valley that would also reflect the opinions and priorities of community leaders among the six counties in the region: Columbia, Dutchess, Greene, Orange, Putnam, and Ulster Counties.

Records for the community leader sample were compiled from one of three sources and divided into three groups. The first group included businesses and organizations in the six counties of the Hudson Valley with 75 or more employees. The information was collected from yellow page telephone directories and supplemented with industry specific sources including government listings, vertical files (bank records, school files, etc.), trade directories, city directories and proprietary files. This group was comprised primarily of business records. The contact name for these records was the top contact, defined as Vice President level and up. The source for this group was Survey Sampling Inc. in Fairfield, Connecticut.

To supplement this initial list, research assistants at MIPO compiled the second group for the community leader sample. To ensure the list was as current and up-to-date as possible, contact names and phone numbers were collected from the Internet. The types of websites that were drawn from were: local, county, and state government sites; school and college sites; chambers of commerce, economic development and tourism sites; media sources and regional directory sites; non-profit organization and place of worship listing sites; hospitals and health care provider sites; and search engine directories for all of the above. This group was primarily government, media and non-profit records.

Records from these first two groups included a contact name, organization or company name, job title, phone number, and county code. Additional codes were added to mark the source of the record (either SSI or MIPO) and the type of record (government, business, media, or non-profit). The first group included 589 records and the second group contained 283 records. The two groups were combined into one sample file and duplicates were removed. If duplicate phone

numbers had multiple contact names, the record with the highest-ranking individual was chosen. The resulting file was sorted by source of the record, type of record, and county. It was then randomized again.

Replicate numbers were assigned to each record in order to stratify by source of the record, type of record, and county. Replicates are systematically nth-selected subsets of the sample itself. They provide a way of insuring that a representative sample is being selected and that interviews are distributed across the entire geography at any one time. During data collection, replicates were released one replicate at a time. Using quota controls, counts for source of the record, type of record, and county were carefully monitored to insure that each category was neither over- or under-represented in the collected data.

In order to provide a sample of community leaders of sufficient size and breadth, a third source of information was used. Any study of community leadership must also include individuals that participate in community decision-making but who may not hold a formal position. They may be active through membership in community organizations, corporate boards of directors, or organize less visible community activities. To address this concern multiplicity sampling was added to the sample design.

Multiplicity sampling, also referred to as snowball sampling, involves collecting additional names during the survey process from survey respondents, the individuals that are interviewed.* It is a method to expand the number of people in the sample of the target population that is being studied. These models are often used for studying groups with rare characteristics and/or low incidence in the population. Respondents are asked to suggest other members in their network who are part of the target population that is being studied. At the end of each interview, respondents are invited to identify other people who are like themselves on some specific criteria. For this study, individuals who were interviewed in the community leaders' survey were asked the following question at the end of their interview:

“Finally, this is a survey of people who are seen as influential in this community. Is there someone else in this community that you can recommend we speak with to get their opinion?”

If the response was yes, respondents were asked to provide a contact name, company or organization name, job title, and phone number. Respondents could

* Multiplicity sampling generally refers to when original respondents are allowed to provide names. Snowball sampling generally allows multiplicity derived respondents to provide names in addition to the original respondents. For more information regarding these sampling methods see “Sampling Racial and Ethnic Minorities,” William D. Kalsbeek, Director, Survey Research Unit, University of North Carolina; “Telephone Screening for Rare Characteristics Using Multiplicity Counting Rules,” Jane W. Bergsten and Stephanie Pierson, Research Triangle Institute; “On Finding and Interviewing the Needles in a Haystack: The Use of Multiplicity Sampling,” George S. Rothbart, Michelle Fine, and Seymour Sudman, *Public Opinion Quarterly* Vol. 46:408-421.

nominate up to two individuals. 60% of respondents provided at least one name and 25% offered two names. If respondents did not have a phone number, research assistants used phone directories and the Internet to provide the needed data. Based on researched information, codes for source, type, and county were added to each record. Added names were cross-referenced with the original two groups for any duplication. If there was a duplicate phone number in one of the two original groups, the record with the highest-ranking individual was chosen. Replicates were then added to each record. Multiplicity records were estimated to provide one-third of the sample, and replicates were assigned with that proportion as a guide. New records were added to the phone sample on a daily basis.

Data Collection

The questionnaires and the telephone samples were programmed for computer assisted telephone interviewing (CATI). Interviewing was conducted from a centralized telephone facility using trained interviewers who were specifically briefed on this study. Polling supervisors regularly monitored, evaluated, and provided feedback to the interviewing staff.

93 pretest interviews for the survey of Hudson Valley residents were administered by telephone on September 27th, 2001. Interviewers contacted households between 5:00 p.m. and 9:45 p.m. As a result of the pretest, the questionnaire was updated and revised.

Interviews for the survey of Hudson Valley residents were administered by telephone from October 5th through December 10th, 2001. Residents had a choice to complete the survey in English or Spanish. For the survey of residents, interviewers attempted to contact households between 5:00 p.m. and 9:45 p.m. on weeknights (except Friday) and from 3:00 p.m. to 9:00 p.m. on Sunday. Callbacks were also conducted between 9:00 a.m. and 5:00 p.m. on weekdays. Up to 16 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.

Interviews for the survey of Hudson Valley community leaders were conducted by telephone from November 15th through December 19th, 2001. For the survey of community leaders, interviewers attempted to contact these individuals on weekdays from 9:00 a.m. through 5:00 p.m. unless an appointment for another time was scheduled. Attempts to contact an individual were continued until the total number of interviews from a county was complete. This included callbacks to initial unsuccessful contacts and scheduled appointments. A toll free number was

also used so that community leaders could call back the survey center at their convenience. 30% of the community leaders who completed a survey interview contacted 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 are computed using Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys published by the American Association of Public Opinion Research, 2001.

- The response rate is the number of completed interviews divided by the number of households in the sample.
- The cooperation rate is the proportion of all cases interviewed of all eligible households ever contacted.
- The refusal rate is the proportion of all cases in which a respondent refuses to be interviewed or breaks-off an interview.
- The contact rate is the proportion of all cases in which some household member was reached.

	Residents						
	Hudson Valley	Columbia	Dutchess	Greene	Orange	Putnam	Ulster
	%	%	%	%	%	%	%
Response Rate	73%	71%	71%	75%	73%	78%	70%
Cooperation Rate	80%	75%	78%	82%	81%	83%	79%
Refusal Rate	7%	11%	7%	6%	5%	7%	6%
Contact Rate	96%	100%	93%	100%	96%	99%	93%

	Community Leaders						
	Hudson Valley	Columbia	Dutchess	Greene	Orange	Putnam	Ulster
	%	%	%	%	%	%	%
Response Rate	61%	73%	56%	61%	73%	58%	52%
Cooperation Rate	97%	100%	100%	100%	95%	100%	96%
Refusal Rate	1%				2%		2%
Contact Rate	63%	73%	56%	61%	77%	58%	54%