CHAPTER ONE

AN INTRODUCTION TO SURVEY METHODOLOGY

A Note to the Reader

You are about to be exposed to a system of principles called "survey methodology" for collecting information about the social and economic world. We have written this book in an attempt to describe the excitement of designing, conducting, analyzing, and evaluating sample surveys. To appreciate this fully, use the devices we have placed in each chapter to enrich your memory of the material. Throughout the book you'll see boxes with illustrations and examples of key principles, terminology notes, and highlights of classic research studies in the field. In the outside margin of each page you'll find key terms, at the point they are defined. At the end of each chapter is a set of exercises that you can use to test your understanding of that chapter's material. The best strategy is to read the text through, then, at the end of each chapter, go back, read the boxes, and review the key terms.

At 8:30 AM on the day before the first Friday of each month, a group of economists and statisticians enter a soundproof and windowless room in a building at 2 Massachusetts Avenue, NE, in Washington, DC, USA. Once those authorized are present, the room is sealed.

Those in the room are professional staff of the US Bureau of Labor Statistics (BLS), and their task is to review and approve a statistical analysis of key economic data. Indeed, they have spent the week poring over sets of numbers, comparing them, examining indicators of their qualities, looking for anomalies, and writing drafts of a press release describing the numbers. They write the press release in simple language, understandable by those who have no technical knowledge about how the numbers were produced.

At 8:00 AM the next day, a group of journalists assemble in a monitored room in the nearby main Department of Labor building, removed from any contact with the outside world. The BLS staff enter the room and then reveal the results to the journalists. The journalists immediately prepare news stories based on the briefing. At exactly 8:30 AM, they simultaneously electronically transmit their stories to their news organizations and sometimes telephone editors and producers.

The statistics revealed are the unemployment rate of the prior month and the number of jobs created in the prior month. The elaborate protections and security used prior to their release stem from the enormous impact the numbers can have

on society. Indeed, in months when the numbers signal important changes in the health of the US economy, thousands of stock market investors around the world make immediate buy and sell decisions. Within 45 minutes of the announcement, trillions of dollars can move in and out of markets around the world based on the two numbers revealed at 8:30 AM.

Both the unemployment rate and the jobs count result from statistical surveys. A household survey produces the unemployment rate; an employer survey, the jobs count. The households and employers surveyed have been carefully selected so that their answers, when summarized, reflect the answers that would be obtained if the entire population were questioned. In the surveys, thousands of individual people answer carefully phrased questions about their own or their company's attributes. In the household survey, professional interviewers ask the questions and enter the answers onto laptop computers. In the employer survey, the respondents complete a standardized questionnaire either on paper or electronically. Complex data processing steps follow the collection of the data, to assure internal integrity of the numbers.

These two numbers have such an impact because they address an important component of the health of the nation's economy, and they are credible. Macroeconomic theory and decades of empirical results demonstrate their importance. However, only when decision makers believe the numbers do they gain value. This is a book about the process of generating such numbers through statistical surveys and how survey design can affect the quality of survey statistics. In a real sense, it addresses the question of when numbers from surveys deserve to be credible and when they do not.

1.1 INTRODUCTION

This chapter is an introduction to survey methodology as a field of knowledge, as a profession, and as a science. The initial sections of the chapter define the field so that the reader can place it among others. At the end of the chapter, readers will have a sense of what survey methodology is and what survey methodologists do.

survey

A "survey" is a systematic method for gathering information from (a sample of) entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members. The word "systematic" is deliberate and meaningfully distinguishes surveys from other ways of gathering information. The phrase "(a sample of)" appears in the definition because sometimes surveys attempt to measure everyone in a population and sometimes just a sample.

statistic

descriptive statistic

analytic statistic

The quantitative descriptors are called "statistics." "Statistics" are quantitative summaries of observations on a set of elements. Some are "descriptive statistics," describing the size and distributions of various attributes in a population (e.g., the mean years of education of persons, the total number of persons in the hospital, the percentage of persons supporting the President). Others are "analytic statistics," measuring how two or more variables are related (e.g., a regression coefficient describing how much increases in income are associated with increases in years of education; a correlation between education and number of books read in the last year). That goal sets surveys apart from other efforts to describe people or events. The statistics attempt to describe basic characteristics or experiences of large and small populations in our world.

Almost every country in the world uses surveys to estimate their rate of unemployment, basic prevalence of immunization against disease, opinions about the central government, intentions to vote in an upcoming election, and people's satisfaction with services and products that they buy. Surveys are a key tool in tracking global economic trends, the rate of inflation in prices, and investments in new economic enterprises. Surveys are one of the most commonly used methods in the social sciences to understand the way societies work and to test theories of behavior. In a very real way, surveys are a crucial building block in a modern information-based society.

Although a variety of activities are called surveys, this book focuses on surveys that have the following characteristics:

1) Information is gathered primarily by asking people questions.
2) Information is collected either by having interviewers ask questions and record answers or by having people read or hear questions and record their own answers.
3) Information is collected from only a subset of the population to be described – a sample – rather than from all members.

survey methodology

error

statistical error

Since "ology" is Greek for "the study of," survey methodology is the study of survey methods. It is the study of sources of error in surveys and how to make the numbers produced by surveys as accurate as possible. Here the word, "error," refers to deviations or departures from the desired outcome. In the case of surveys, "error" is used to describe deviations from the true values applicable to the population studied. Sometimes, the phrase "statistical error" is used to differentiate this meaning from a reference to simple mistakes.

The way each of the above steps is carried out – which questions are asked, how answers are collected, and which people answer the questions – can affect the quality (or error properties) of survey results. This book will describe how to conduct surveys in the real world and how to evaluate the quality of survey results. It will describe what is known, and not known, about how to minimize error in survey statistics. Most of all, this book will attempt to distill the results of 100 years of scientific studies that have defined the theories and principles, as well as practices, of high-quality survey research.

1.2 A Brief History of Survey Research

Converse (1987) has produced an important account of the history of survey research in the United States, and we recount some of the highlights here. There are four perspectives on surveys that are worth describing: the purposes to which surveys were put, the development of question design, the development of sampling methods, and the development of data collection methods.

1.2.1 The Purposes of Surveys

census

Perhaps the earliest type of survey is the census, generally conducted by governments. Censuses are systematic efforts to count an entire population, often for

Schuman (1997) on "Poll" Versus "Survey"

What is the difference between a poll and a survey? The word "poll" is most often used for private sector opinion studies, which use many of the same design features as studies that would be called "surveys." "Poll" is rarely used to describe studies conducted in government or scientific domains. There are, however, no clear distinctions between the meanings of the two terms. Schuman notes that the two terms have different roots: "'Poll' is a four letter word, generally thought to be from an ancient Germanic term referring to 'head,' as in counting heads. The two-syllable word 'survey,' on the other hand, comes from the French *survee*, which in turn derives from Latin *super* (over) and *videre* (to look). The first is therefore an expression with appeal to a wider public, the intended consumers of results from Gallup, Harris, and other polls. The second fits the needs of academicians in university institutes who wish to emphasize the scientific or scholarly character of their work." (page 7)

purposes of taxation or political representation. In the United States, the constitution stipulates that a census must be conducted every ten years, to reapportion the House of Representatives reflecting current population residence patterns. This gives the statistics from a census great political import. Because of this, they are often politically contentious (Anderson, 1990).

A prominent early reason for surveys was to gain understanding of a social problem. Some people trace the origins of modern survey research to Charles Booth, who produced a landmark study titled Life and Labour of the People of London (1889–1903) (*http://booth.lse.ac.uk/*). As Converse recounts it, Booth spent his own money to collect voluminous data on the poor in London and the reasons they were poor. He wrote at least 17 volumes based on the data he collected. He did not use methods like the ones we use today – no well-defined sampling techniques, no standardized questions. Indeed, interviewer observation and inference produced much of the information. However, the Booth study used quantitative summaries from systematic measurements to understand a fundamental societal problem.

In contrast to studies of social problems, journalism and market research grew to use surveys to gain a systematic view of "the man on the street." A particular interest was reactions to political leaders and preferences in upcoming elections. That interest led to the development of modern public opinion polling.

In a related way, market research sought knowledge about reactions of "real" people to existing and planned products or services. As early as the 1930s there was serious research on what programs and messages delivered via the radio would be most popular. The researchers began to use surveys of broader samples to produce information more useful to commercial decision makers.

Over the early 20th century, public opinion polling and market research, sometimes done by the same companies, evolved to use mail surveys and telephone surveys. They often sampled from available lists, such as telephone, driver's license, registered voter, or magazine subscriber listings. They collected their data primarily by asking a fixed set of questions; observations by interviewers and proxy reporting of other people's situations were not part of what they needed. These features were directly tied to the most important difference between what they were doing and what those who had gone before had done: rather than collecting data about facts and objective characteristics of people, the polling and market research surveyors were interested in what people knew, felt, and thought.

The measurement of attitudes and opinions is a key foundation of the modern management philosophies that place much weight on customer satisfaction.

Customer satisfaction surveys measure expectations of purchasers about the quality of a product or service and how well their expectations were met in specific transactions. Such surveys are ubiquitous tools of management to improve the performance of their organizations.

Politicians and political strategists now believe opinion polls are critical to good decisions on campaign strategy and messages to the public about important issues. Indeed, a common criticism of modern politicians is that they rely too heavily on polling data to shape their personal opinions, choosing to reflect the public's views rather than provide leadership to the public about an issue.

1.2.2 The Development of Standardized Questioning

The interest in measuring subjective states (i.e., characteristics that cannot be observed, internalized within a person) also had the effect of focusing attention on question wording and data collection methods. When collecting factual information, researchers had not thought it important to carefully word questions. Often, interviewers were sent out with lists of objectives, such as age, occupation, education, and the interviewers would decide on how the questions would be worded. Experienced researchers often did the interviewing, with great confidence that they knew how to phrase questions to obtain good answers.

However, the market research and polling organizations were doing large numbers of interviews, using newly hired people with no special background in the social sciences. Of necessity, researchers needed to specify more carefully the information sought by the survey. Further, researchers found that small changes in wording of an attitude question sometimes had unusually large effects on the answers.

Thus, early in the development of opinion surveys, attention began to be paid to giving interviewers carefully worded questions that they were to ask exactly the same way for each interview. Also, as interviewers were used more to ask questions, it was found that how they asked questions and recorded answers could affect the results. This led eventually to researchers training and supervising interviewers more formally than earlier.

Question wording also was influenced as the academics started to pay some attention to what the commercial researchers were doing. Psychometricians, psychologists who quantify psychological states, had been interested in how to put meaningful numbers on subjective states. Measuring intelligence was the first effort in this direction. However, people such as Thurstone also worked on how to assign numbers to attitudes, feelings, and ratings (e.g., Thurstone and Chave, 1929).

For the most part, their approaches were extremely cumbersome and were used primarily when they could get captive college student volunteers to fill out lengthy, highly redundant questionnaires. Such instruments were not going to be useful for most survey interviews with representative samples; they took too long to measure one or a few attitudes. Rensis Likert in his PhD dissertation (Likert, 1932), however, demonstrated that a single, streamlined question, with a scaled set of answers, could accomplish much the same thing as a lengthy series of paired comparisons. Likert applied the work to surveys (and later founded the University of Michigan Survey Research Center in 1946).

1.2.3 The Development of Sampling Methods

Early researchers, such as Booth, essentially tried to collect data on every element of a defined population. Such censuses avoided problems of errors arising from measuring just a subset of the population, but were clearly impractical for large populations. Indeed, the difficulty of analyzing complete census data led to early efforts to summarize a census by taking a sample of returns. Early efforts to sample would study a "typical" town, or they would purposively try to collect individuals to make the samples look like the population – for example, by interviewing about half men and half women, and trying to have them distributed geographically in a way that is similar to the population.

Although the theory of probability was established in the 18th century, its application to practical sample survey work was largely delayed until the 20th century. The first applications were the taking of a "1 in N" systematic selection from census returns. These were "probability samples;" that is, every record had a known nonzero chance of selection into the sample.

probability sample

A big breakthrough in sampling came from people who did research on agriculture. In order to predict crop yields, statisticians had worked out a strategy they called "area probability sampling." This is just what it sounds like: they would sample areas or plots of land and find out what farmers were doing with those plots in the spring (for example, if they were planting something on them and, if so, what) in order to project what the fall crops would look like. The same technique was developed to sample households: by drawing samples of geographic blocks in cities or tracts of land in rural areas, listing the housing units on the blocks or rural tracts, then sampling the housing units that were listed, samplers found a way to give all households, and by extension the people living in them, a chance to be sampled. The attraction of this technique included the elimination of the need for a list of all persons or all households in the population prior to drawing the sample.

area probability sample

The Depression and World War II were major stimuli for survey research. One of the earliest modern probability samples was drawn for the Monthly Survey of Unemployment, starting in December, 1939, led by a 29-year-old statistician, Morris Hansen, who later became a major figure in the field (Hansen, Hurwitz, and Madow, 1953). During the war, the federal government became interested in conducting surveys to measure people's attitudes and opinions, such as interest in buying war bonds, as well as factual information. Considerable resources were devoted to surveys during the war, and researchers who were recruited to work with the government during the war later came to play critical roles in the development of survey methods. When the war was over, methodologists understood that in order to produce good population-based statistics it was necessary to attend to three aspects of survey methodology: how questions were designed; how the data were collected, including the training of interviewers; and how samples were drawn.

Probability samples are the standard by which other samples are judged. They are routinely used by almost all government statistical agencies when data are used to provide important information for policy makers. They are used for surveys used in litigation. They are used for measurement of media audience sizes, which in turn determine advertising rates. In short, whenever large stakes ride on the value of a sample, probability sampling is generally used.

1.2.4 The Development of Data Collection Methods

The gathering of information in early surveys was only one step more organized than talking to as many people as possible about some topic. The qualitative interviews produced a set of verbal notes, and the task of summarizing them with statistics was huge. Surveys grew to be popular tools because of the evolution of methods to collect systematic data cheaply and quickly.

Mailed paper questionnaires offered very low costs for measuring literate populations. Indeed, by 1960 a formal test of census procedures based on mailed questionnaires succeeded to an extent that the 1970 census was largely a mailed questionnaire survey. Further, mailed questionnaires proved to be much cheaper than sending interviewers to visit sample cases. On the other hand, mailed surveys were subject to the vagaries of the postal service, which, even when it worked perfectly, produced survey periods that lasted months, not weeks.

With the spread of telephone service throughout the country, market researchers first saw two advantages of using the medium as a data collection tool. It was much faster than mail questionnaire surveys, and it was still cheaper than face-to-face surveys. For decades, however, the mode suffered from the clear lack of coverage of telephones among the poor and more transient members of the society. By the 1990s, however, almost all market research had moved away from the face-to-face survey and much scientific research was close behind in the abandonment of that mode. It was largely the federal government that continued to rely on face-to-face household surveys.

Like many fields of human activity, huge leaps in the efficiencies of surveys came from the invention of the computer. One of the first computers made in the United States was used in processing decennial census data. Survey researchers quickly recognized how computers could reduce the large amount of human resources needed to conduct surveys. Survey researchers first used computers to perform the analysis steps of a survey, then they began to use them to assist in checking the raw data for clerical errors, then to assist them in coding text answers, then in the data collection step itself. Now computers (from handheld devices to networked systems) are used in almost every step of survey design, data collection, and analysis. The fastest growing application is the development of Web surveys.

As these various developments evolved, the field also developed a set of performance guidelines. Empirical studies demonstrated the value of various sample designs on the quality of statistics. Interviewer training guidelines improved the standardization of interviewers. Standards about computing and reporting response rates offered the field measures useful in comparing surveys.

In the 60 years following the advent of surveys, a great deal has been learned about how to design data collection systems to improve the quality of survey statistics. However, as can be seen from this short history, the basic elements of good survey methodology were defined in the first half of the 20th century.

1.3 Some Examples of Ongoing Surveys

One way to understand the range of survey methods and the potential of surveys to provide information is to give some examples. The following is a brief descrip-

tion of six surveys. We have chosen them to use as examples throughout the book for several reasons. First, they are all ongoing surveys. They are done year after year. By definition, that means that the sponsors think that there is a continuous need for the kind of information that they provide. That also means that someone thinks they are important. These are not particularly typical surveys. They do not include public opinion, political, or market research studies. They do not include any one-time surveys, which are highly prevalent. All of these surveys are funded by government sources.

However, they do differ from one another in numerous ways. One reason we chose this set is they do give a sense of the range of topics that are addressed by surveys and the variety of survey designs that are used. They also were chosen because they provide examples of excellence in survey research. Hence, they will provide opportunities for us to discuss how different methodological problems are addressed and solved.

In the brief summaries and charts provided, we describe some of the basic characteristics of each survey:

1) their purposes
2) the populations they try to describe
3) the sources from which they draw samples
4) the design of the way they sample people
5) the use of interviewers
6) the mode of data collection
7) the use of computers in the collection of answers

Readers should think about these surveys in two different ways. First, think of them as information sources – what can we learn from them, what questions they answer, what policies and decisions they inform; in short, why they are conducted. Second, compare the design features above in order to see how different survey design features permit the surveys to achieve their different purposes.

1.3.1 The National Crime Victimization Survey

How much crime is there in the United States? Are crimes increasing in frequency or going down in frequency? Who gets victimized by crimes? Every society seeks answers to these questions. In the United States, the answers were sought through quantifying crime in a period of great public concern about organized crime. In the 1930s, the International Association of Chiefs of Police began a collection of administrative record counts. The method rested on the reporting of crimes to police in jurisdictions around the country, based on the administrative records kept by individual sheriffs, transit police, city police, and state police offices. Police chiefs had designed the record systems to have legal documentation of the circumstances of the crime, the victim of the crime, the offender, and any relevant evidence related to the crime. Individual staff members completed the paperwork that produced the administrative records. However, many crimes only come to the attention of the police if a citizen decides to report them. Often, the decision to produce a record or to label an incident was left to a relatively low-level police officer. For years, these records were the key information source on US crime.

Over the years, several weaknesses in the statistics from police records became obvious. Sometimes a new mayor, fulfilling a pledge to reduce crime, created an environment in which more police officers chose not to label an incident as a crime, thus not producing an administrative record. Further, the statistics were tainted by different jurisdictions using different definitions for crime categories. When police believed that the crime would never be solved, they encouraged the resident not to file a formal report. There was growing evidence in some jurisdictions that relations between the public and the police were poor. Fear of the police among the public led to avoidance of reporting criminal incidents. The police officers themselves carried with them attitudes toward subpopulations that led to classifying an incident as a crime for one group but not for another group. It was becoming clear that, whereas major crimes like homicide were well represented in the record systems, the records tended to miss more minor, often unreported, crimes. Some jurisdictions kept very detailed, complete records, whereas others had very shoddy systems.

Thus, over the decades many began to distrust the value of the statistics to address the simplest question: "How much crime is there in the US?" Further, the simple counts of crimes were not giving policy makers clear information about the characteristics of crimes and their victims, information helpful in considering alternative policies to reducing crime. The President's Commission on Law Enforcement and the Administration of Justice, established by President Johnson, noted in a task force report that, "If we knew more about the character of both offenders and victims, the nature of their relationships and the circumstances that create a high probability of crime conduct, it seems likely that crime prevention programs could be made much more effective" (President's Commission, 1967, as cited in Rand and Rennison, 2002).

In the late 1960s, criminologists began exploring the possibilities of using surveys to ask people directly whether they were a victim of a crime. This forced a conceptually different perspective on crime. Instead of a focus on the incident, it focused on one actor in the incident – the victim. This shift of perspective produced clear contrasts with the police reported crime record systems. Most obviously, homicide victims can't report! Victimizations of young children might not be well reported by their parents (who may not know about incidents at school), and the children may not be good respondents. Crimes against companies present problems of who can report them well. For example, if someone starts a fire that burns down an apartment building, who are the victims – the owner of the property, the renters of the apartments, the people visiting during the fire? From one perspective, all are victims, but asking all to report the arson as a victimization may complicate the counts of crimes. Victims can sometimes report that an unpleasant event occurred (e.g., someone with no right of entry entered their home). However, they cannot, as do police, gather the information that asserts intent of the offender (e.g., he was attempting a theft of a television).

On the other hand, using individual reporters, a survey can cover victimizations reported to the police and those not reported to the police. This should provide a more complete picture of crime, if there are crimes not reported to the police. Indeed, self-reported victimizations might be a wonderful addition to statistics on police-reported crimes, as a way to compare the perceived victimization in a society with the officially reported victimization status. Moreover, the survey has the advantage of utilizing standard protocols for the measurement of victimizations across the country.

sampling error

However, the notion of using a national survey to measure victimization faced other problems. Although all police agencies in the United States could be asked to report statistics from their record systems, it was financially impossible to ask all persons in the United States to report their individual victimizations. If a "representative" sample could be identified, this would make the victimization survey possible. Thus, in contrast to the record data, the survey would be subject to "sampling error" (i.e., errors in statistics because of the omission of some persons in the population).

But could and would people really report their victimizations accurately? Survey methodologists studied the reporting problem in the late 1960s and early 1970s. In methodological studies that sampled police records and then went back to victims of those crimes, they found that, in large part, they provided reports that mirrored the data in the records. However, Gottfredson and Hindelang (1977) among others noted that one problem was a tendency for persons to misdate the time of an incident. Most incidents that were important to them were reported as occurring more recently than they actually did. By and large, however, the pattern of reporting appeared to justify the expense of mounting a completely separate system of tracking crime in the country.

When all the design features were specified for the survey approach, it was clear that there were inevitable differences between the police reports and the victim reports (Rand and Rennison, 2002). The abiding strength of the survey was that it could measure crimes that are not reported to the police or for which police do not make formal documentation. However, the survey would be subject to sampling error. On the other hand, the police-reported crimes include some victims who are not residents of the country, and these would be missed in a survey of the US household population. Further, the police report statistics include homicides and arson, but do not include simple assault. Police reports exclude rapes of males; the survey could include rapes of both sexes. Some crimes have multiple victims, who could report the same incident as occurring to them (e.g., household crimes, group thefts); the survey would count them as multiple crimes, and the police reports as one incident, generally. The police report statistics depend on voluntary partnerships between the federal government and thousands of jurisdictions, but jurisdictions vary in their cooperation; the survey (using traditional methods) would suffer from some omissions of persons who are homeless or transient and from those who choose not to participate. The methodological work suggested that the survey would tend to underreport crimes where the offender is well known to the victim. There was also evidence that the survey would underreport less important crimes, apparently because they were difficult to remember. Finally, both systems have trouble with repeated incidents of the same character. In the survey method, repeated victimizations of the same type were to be counted once as "series incidents" (e.g., repeated beatings of a wife by her husband); the police-reported series would have as many reports as provided to the police. The National Crime Victimization Survey (NCVS) (Table 1.1) has its roots within the US Department of Justice in 1972. The current Bureau of Justice Statistics has the mission of collecting and disseminating statistical information on crimes, criminal offenders, victims of crimes, and the operations of justice systems at all levels of government. The Bureau of Justice Statistics contracts with the US Census Bureau to collect the data in the NCVS.

The NCVS asks people to report the crimes they have experienced in the 6 months preceding the interview. If they asked people to report for 12 months, the

Table 1.1.

Sponsor
Collector
Purpose
Year Start
Target Po
Sampling I
Sample D
Sample Si
Use of Inte
Mode of A
Computer
Reporting
Time Dime
Frequency
Interviews of Survey
Levels of C
Web Link

researchers c
cient use of i
drop-off in th
that happen
crimes even
reporting wo
or better yet,
fewer and fe

Table 1.1. Example Survey: National Crime Victimization Survey (NCVS)

Sponsor	US Bureau of Justice Statistics
Collector	US Census Bureau
Purpose	Main objectives are to: • develop detailed information about the victims and consequences of crime • estimate the number and types of crimes not reported to the police • provide uniform measures of selected types of crimes • permit comparisons over time and by types of areas
Year Started	1973 (previously called the National Crime Survey, 1973–1992)
Target Population	Adults and children 12 or older, civilian and noninstitutionalized
Sampling Frame	US households, enumerated through counties, blocks, listed addresses, lists of members of the household
Sample Design	Multistage, stratified, clustered area probability sample, with sample units rotating in and out of the sample over three years
Sample Size	About 42,000 households – 76,000 persons
Use of Interviewer	Interviewer administered
Mode of Administration	Face-to-face and telephone interviews
Computer Assistance	Paper questionnaire for 70% of the interviews, both face-to-face and telephone interviews; computer assistance for 30% of the interviews
Reporting Unit	Each person age 12 or older in household reports for self
Time Dimension	Ongoing rotating panel survey of addresses
Frequency	Monthly data collection
Interviews per Round of Survey	Sampled housing units are interviewed every six months over the course of three years
Levels of Observation	Victimization incident, person, household
Web Link	*http://www.ojp.usdoj.gov/bjs/cvict.htm*

researchers could learn about more events per interview; it would be a more efficient use of interview time. However, early studies showed that there is a marked drop-off in the accuracy of reporting when people are asked to remember events that happen more than 6 months in the past. In fact, there is underreporting of crimes even when the questions ask about the past 6 months. The accuracy of reporting would be higher if the questions asked about only one or two months – or better yet, only a week or two. However, as the reporting period gets shorter, fewer and fewer people have anything to report, so more and more interviews

provide minimal information about victimization. The designers of the survey chose 6 months as a reasonable point at which to balance reporting accuracy and the productivity of interviews.

The sample for the NCVS is drawn in successive stages, with the goal of giving every person 12 years old and older a known chance of selection, and thereby producing a way to represent all age eligible persons in the United States. (The jargon for this is a "multistage, stratified clustered area probability sample," which will be described in Chapter 4.) The sample is restricted to persons who are household members, excluding the homeless, those in institutions, and in group quarters. (The survey methodologists judged that the cost of covering these subpopulations would be prohibitively expensive, detracting from the larger goals for the survey.) The sample is clustered into hundreds of different sample areas (usually counties or groups of counties) and the sample design is repeated samples of households from those same areas over the years of the study. The clustering is introduced to save money by permitting the hiring of a relatively small group of interviewers to train and supervise, who travel out to each sample household to visit the members and conduct interviews. Further, to save money all persons 12 years and over in the household are interviewed; thus, one sample household might produce one interview or many.

A further way to reduce costs of the survey is to repeatedly measure the same address. When the design randomly identifies a household to fall into the NCVS sample, the interviewer requests that, in addition to the first interview, the household be willing to be visited again six months later, and then again and again, for a total of seven interviews over a three-year period. In addition to saving money, this produces higher-quality estimates of change in victimization rates over years. This design is called a "rotating panel design" because each month, different people are being interviewed for the first time, the second time, the third time, the fourth time, the fifth time, the sixth time, and the last (seventh) time. Thus, the sample is changing each month but overlaps with samples taken six months previously.

rotating panel design

Each year, the NCVS collects interviews from about 42,000 households containing more than 76,000 persons. About 92% of the sample households provide one or more interviews; overall, about 87% of the persons eligible within the sample households provide an interview. The interviews contain questions about the frequency, characteristics, and consequences of criminal victimizations the households may have experienced in the previous six months. The interview covers incidents of household victimization and personal victimization – rape, sexual assault, robbery, assault, theft, household burglary, and motor vehicle theft. An interviewer visits those households and asks those who live there about crimes they have experienced over the past six months. One person in the household acts as the informant for all property crimes (like burglary, vandalism, etc.); each person then reports for him- or herself about personal crimes (e.g., assault, theft of personal items). The interview is conducted in person in the first wave; subsequent waves attempt whenever possible to use telephone interviewers calling from two different centralized call centers. Thus, the NCVS statistics are based on a mix of telephone (60%) and face-to-face interviews (40%).

The questionnaire asks the respondent to remember back over the past six months to report any crimes that might have occurred. For example, the questions in the box ask about thefts.

I'm going to read some examples that will give you an idea of the kinds of crimes this study covers.

As I go through them, tell me if any of these happened to you in the last 6 months, that is since _ _(MONTH) _ _(DAY), 20_ _.

Was something belonging to YOU stolen, such as:

a) Things that you carry, like luggage, a wallet, purse, briefcase, book
b) Clothing, jewelry, or calculator
c) Bicycle or sports equipment
d) Things in your home – like a TV, stereo, or tools
e) Things from a vehicle, such as a package, groceries, camera, or cassette tapes

If the respondent answers "yes, someone stole my bicycle," then the interviewer records that and later in the interview asks questions about details of the incident. Figure 1.1 shows the kind of statistics that can be computed from the NCVS. The percentages of households reporting one or more crimes of three different types (property crimes, vandalism, and violent crime) are displayed for the years 1994–2000. Note that the percentages are declining, showing a reduced frequency of crime over the late 1990s. Policy makers watch these numbers closely as an indirect way to assess the impact of their crime-fighting programs. However, other research notes that when the nation's economy is strong, with low unemployment, crime rates tend to decline.

"Crime at Lowest Point in 25 Years, Fed Says" reads the headline on CNN on December 27, 1998. "Fewer people in the United States were the victims of crimes last year than at any time since 1973, the Justice Department reported Sunday," reads the first line. Later in the story, "President Bill Clinton applauded the new crime figures Sunday. They 'again show that our strategy of more police, stricter gun laws, and better crime prevention is working,' he said in a statement." These findings result from the NCVS. The statement attributing drops in crime to policies promoted by the current administration is common, but generally without strong empirical support. (It is very difficult, given available information, to link changes in crime rates to policies implemented.)

NCVS data have informed a wide audience concerned with crime and crime prevention. Researchers at academic, government, private, and nonprofit research institutions use NCVS data to prepare reports, policy recommendations, scholarly publications, testimony before Congress, and documentation for use in courts (US Bureau of Justice Statistics, 1994). Community groups and government agencies use the data to develop neighborhood watch and victim assistance and compensation programs. Law enforcement agencies use NCVS findings for training. The data appear in public service announcements on crime prevention and crime documentaries. Finally, print and broadcast media regularly cite NCVS findings when reporting on a host of crime-related topics.

1.3.2 The National Survey on Drug Use and Health

What percentage of people in the United States use illicit drugs? Are rates of usage higher among the poor and less educated than among others? How does

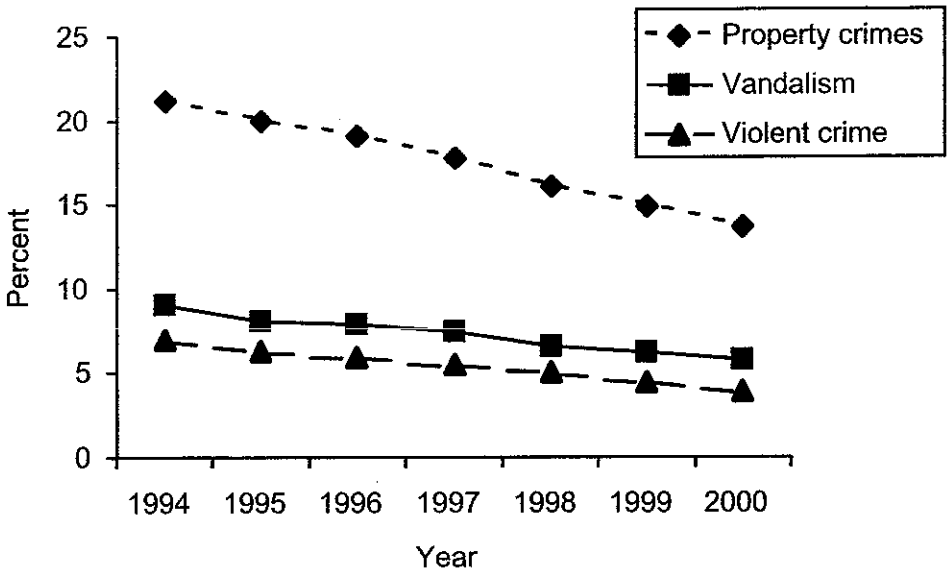

Figure 1.1 Percentage of US households experiencing a crime, by type of crime by year, based on the 1994–2000 National Crime Victimization Survey. (Source: www.ojp.usdoj.gov/bjs/)

drug use change as people get older? Are rates of usage changing over time? Do different groups tend to use different drugs? Is use of alcoholic beverages related to drug use? Are the rates of usage changing over time? Do different states have different drug usage patterns? Each year the National Survey on Drug Use and Health (NSDUH) draws samples of households from each state of the union. Interviewers visit each sampled home and ask questions of each sample person about their background and other relatively nonsensitive information. To collect data on drug usage, the interviewers provide a laptop computer to the respondent, who dons headphones to listen to a voice recording of the questions and keys in the answers on the laptop keyboard. Every year, the NSDUH produces estimates of rates of usage of several different kinds of drugs. The data are used to inform US federal government drug policy, aimed at reducing demand and supply of illegal drugs.

The first national household survey attempt to measure drug use was conducted in 1971. The survey oversampled persons 12 to 34 years old in order to have enough cases to make separate estimates of that age group because it was suspected to have the highest prevalence of drug use. There was early concern about the US population's willingness to be interviewed about drug use. The designers were concerned about response rates. In the early rounds, personal visit interviewers contacted sample households and asked questions of sample persons. When the interviewer came to sensitive questions about drug use, the interviewer noted that the data would be held confidential, and switched the questioning from oral answers to a self-administered mode, with respondents writing down their answers on an answer sheet, placing the completed form in a sealed envelope and accompanying the interviewer to mail it.

Table 1.2. Example Survey: National Survey of Drug Use and Health (NSDUH)

Sponsor	Substance Abuse and Mental Health Services Administration (SAMHSA)
Collector	RTI International
Purpose	Main objectives are to: • provide estimates of rates of use, number of users, and other measures related to illicit drug, alcohol, and tobacco use at the state and national level • improve the nation's understanding of substance abuse • measure the nation's progress in reducing substance abuse
Year Started	1971 (formerly named National Household Survey on Drug Abuse)
Target Population	Noninstitutionalized adults and children 12 or older in the United States
Sampling Frame	US households, enumerated through US counties, blocks, and lists of members of the households
Sample Design	Multistage, stratified clustered area probability sample within each state
Sample Size	136,349 housing units; 68,126 persons (2002 NSDUH)
Use of Interviewer	Interviewer-administered, with some self-administered questionnaire sections for sensitive questions
Mode of Administration	Face-to-face interview in respondent's home, with portions completed by respondent alone
Computer Assistance	Computer-assisted personal interview (CAPI), with audio computer-assisted self-interview (ACASI) component
Reporting Unit	Each person age 12 or older in household reports for self
Time Dimension	Repeated cross-sectional survey
Frequency	Conducted annually
Interviews per Round of Survey	One
Levels of Observation	Person, household
Web Link	*http://www.samhsa.gov/hhsurvey/hhsurvey.html*

This first survey set many of the ongoing features of the design. Like the NCVS, the NSDUH targets the household population and noninstitutional group quarters (e.g., shelters, rooming houses, dormitories). It covers persons who are civilian residents of US housing units and 12 years of age or older. Over the years, the sample has grown and has recently been redesigned to provide independent samples of each of the 50 states, with about 70,000 persons measured each year.

repeated cross-section design

This permits the survey to provide separate estimates of drug use for eight of the largest states each year and for the other states through combining data over several years. The design also oversamples youths and young adults, so that each state's sample is distributed equally among three age groups (12 to 17 years, 18 to 25 years, and 26 years or older). In contrast to the NCVS, each sample person is interviewed only once. A new sample is drawn each year, and estimates for the year are based on that new sample. (The label for this is a "repeated cross-section design.") About 91 percent of households provide the screening interview (that measures household composition), and 79 percent of the selected individuals, the full interview.

Although the survey is sponsored by a federal government agency, the Substance Abuse and Mental Health Services Administration, it is conducted under contract by RTI International. This structure is common for large household surveys in the United States but is rarer in other countries. The United States developed over its history an ongoing partnership between survey research organizations in the private sector (both commercial, nonprofit, and academic) and the federal and state governments, so that much of the governmental survey information is collected, processed, and analyzed by nongovernmental staff.

The survey covers a wide set of drugs (e.g., alcohol, prescription drugs). The designers periodically update its measurement procedures, with special focus on how it could get more accurate self-reports of drug use. The interviewers now use laptop computers (the process is called "computer-assisted personal interviewing" or CAPI) to display questions and store answers. It also uses audio computer-assisted self-interviewing (ACASI), which has the respondent listen to questions via earphones attached to a laptop computer, see the questions displayed, and enter his/her responses using the keyboard. Increases in reporting of drug use appear to be produced by this technique versus having an interviewer administer the survey [e.g., 2.3 times the number of persons reporting cocaine use (Turner, Lessler, and Gfroerer, 1992, p. 299)]. The survey has increasingly tried to help guide policy on drug treatment and prevention.

By creating statistics from self-reports of the same drugs each year, NSDUH can provide the United States with net changes over time in drug use. Much of the methodological research of the NSDUH has been focused on the fact that drug use tends to be underreported. By comparing statistics over time, using the same method, the hope is that any underreporting problems remain relatively consistent over time and thus the change from year to year is an accurate gauge of differences over time. For example, Figure 1.2 shows that there are small increases between 1999 and 2001 in the use of marijuana, psychotherapeutic drugs, and hallucinogens. Such increases could be cause for reassessment of programs attempting to reduce the supply of drugs and to treat those abusing the drugs.

"Clinton Proposes Holding Tobacco Industry Accountable for Teen Smoking," is the headline in the CNN story of February 4, 2000. The first line reads, "To crack down on smoking by minors, President Bill Clinton wants to make the tobacco industry pay $3000 for every smoker under the age of 18, and also proposes raising the tax on cigarettes by another 25 cents per pack." Later, "The administration says there are currently 4.1 million teen smokers, according to the latest National Household Survey of Drug Abuse conducted by the Department of Health and Human Services." Policy makers watch the NSDUH as evidence of success or failure of current antidrug policies. Activist governments tend to propose new programs when the data suggest failure.

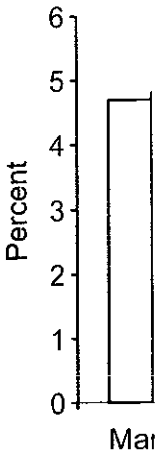

Figure 1.2

1.3.3 The

Are people o
ning to make
Do they think
more optimis

In 1946,
their views of
ful informati
from the basi
ior on consur
chases and s
Since that tir
ducted an ong
by the federa
Leading Eco
the Federal R
country.

Each mo
numbers, loc
those in the h
costs, only p
telephones te
Chapter 3).]
samples telep
resulting sam
mercial, fax,

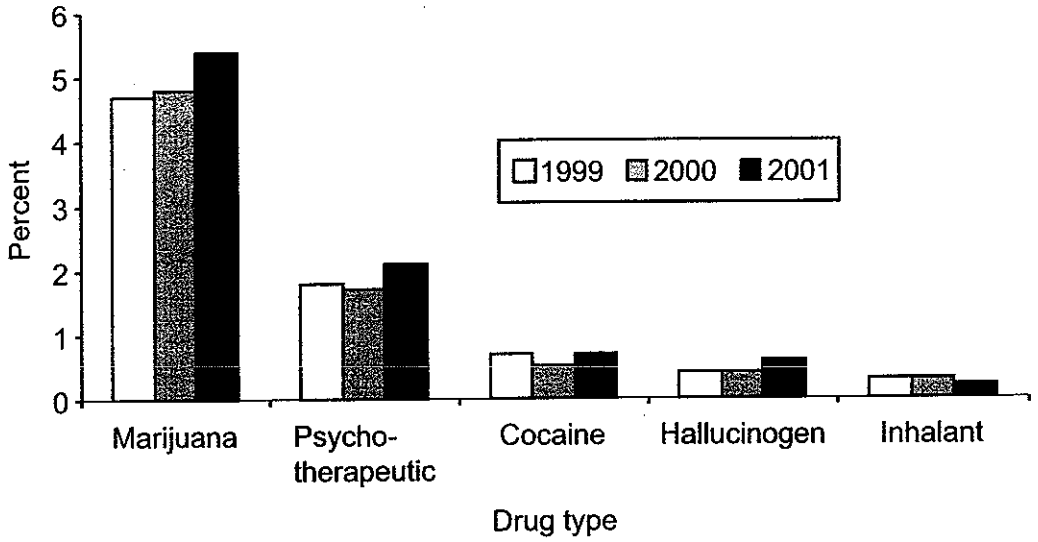

Figure 1.2 Percentage of persons reporting illicit drug use in past month, by drug type by year, 1999–2001. (Source: NSDUH.)

1.3.3 The Survey of Consumers

Are people optimistic or pessimistic about their financial future? Are they planning to make major purchases (e.g., automobiles, refrigerators) in the near future? Do they think they are better off now than they were a few years ago? Are the rich more optimistic than the poor? Are levels of optimism changing over time?

In 1946, the economist George Katona discovered that asking people about their views of their personal and the nation's economic outlook could provide useful information about the future of the entire nation's economy. This stemmed from the basic finding that people's attitudes about the future affect their behavior on consumer purchases and savings. Further, the individual decisions on purchases and savings are an important component of the health of the economy. Since that time, the University of Michigan Survey Research Center has conducted an ongoing survey of consumer attitudes. The survey is not financed solely by the federal government (despite the fact that its statistics form part of the US Leading Economic Indicators), but by a consortium of commercial entities and the Federal Reserve Board, the organization that makes monetary policy for the country.

Each month, the Survey of Consumers (SOC) telephones a sample of phone numbers, locates the household numbers among them, and selects one adult from those in the households. The target is the full household population, but to reduce costs, only persons in households with telephones are eligible. [Those without telephones tend to be poorer, to live in rural areas, and to be more transient (see Chapter 3).] The sample is drawn using a "random digit dialing" design, which samples telephone numbers in working area codes and prefixes. (Not all of the resulting sample numbers are household numbers; some are nonworking, commercial, fax, or modem numbers.) Each month about 500 interviews are con-

random digit dialing

Table 1.3. Example Survey: Surveys of Consumers (SOC)

Survey Name	Survey of Consumers
Sponsor	University of Michigan
Collector	Survey Research Center, University of Michigan
Purpose	Main objectives are to: • measure changes in consumer attitudes and expectations • understand why such changes occur • evaluate how they relate to consumer decisions to save, borrow, or make discretionary changes
Year Started	1946
Target Population	Noninstitutionalized adults in the coterminous United States (omits Hawaii and Alaska)
Sampling Frame	Coterminous US telephone households, through lists of working area codes and exchanges
Sample Design	List-assisted random digit dial sample, randomly selected adult
Sample Size	500 adults
Use of Interviewer	Interviewer administered
Mode of Administration	Telephone interview
Computer Assistance	Computer-assisted telephone interviewing (CATI)
Reporting Unit	Randomly selected adult
Time Dimension	Two-wave panel of persons
Frequency	Conducted monthly
Interviews per Round of Survey	Two; reinterview conducted six months after initial interview on subset of wave 1 respondents
Levels of Observation	Person, household
Web Link	*http://athena.sca.isr.umich.edu/*

ducted. Currently about 60% of selected adults in the households provide an interview. When a sample household is contacted, interviewers attempt to take an interview and then call again 6 months later for a followup interview. Thus, the SOC, like the NCVS, is a rotating panel design. After the data are collected, there are statistical adjustments made to the estimates in an attempt to repair the omission of nontelephone households and nonresponse. Each month, statistics about consumer confidence are computed, using a mix of cases from the first-wave interview and from the second-wave interview.

The interview contains about 50 questions about personal finance, business conditions, and buying conditions. The telephone interviewers working in a centralized telephone interviewing facility use a computer-assisted interviewing sys-

tem, which uses desktop computers to display the questions, accept responses, and route the interviewer to the next appropriate question.

Each month, the Consumer Sentiment Index is published, which has been found predictive of future national economic growth. This indicator, combined with a set of others, forms the Index of Leading Economic Indicators, which is used by economic forecasters in advising on macroeconomic policy and investment strategies. It is interesting and awe-inspiring to many that simple questions on a survey can predict important shifts in something as complex as the US economy. For example, Figure 1.3 shows a plot of two series of statistics. One is the answer to the question, "How about people out of work during the coming 12 months – do you think that there will be more unemployment than now, about the same, or less?" The figure compares an index based on the answers to this question (see left y axis) to the actual annual change in the unemployment rate (right y axis). The consumer expectation index is the dark line and the actual percentage point change in the unemployment rate is the light line. Note that consumer expectations have consistently anticipated changes in the unemployment rate months in advance. It has been shown that consumer expectations contain predictive information about future changes in unemployment that are not captured by other economic information (Curtin, 2003).

Financial investment firms and stock markets watch the consumer confidence statistics. They appear to believe that they are credible predictors of later behaviors. On June 15, 2002, The New York Times, ran an article entitled, "Falling Consumer Confidence Sends the Dow Lower," describing sentiments of those investing in the US stock exchanges, which led to a short-term decline in stock prices because of increased selling. To illustrate the complexity of the information impact on the markets, the same article lists a car bomb in Pakistan as another possible cause of the drop. There are many sources of information that affect financial decision makers' behaviors. It is clear that one of them is the ongoing measure of consumer confidence.

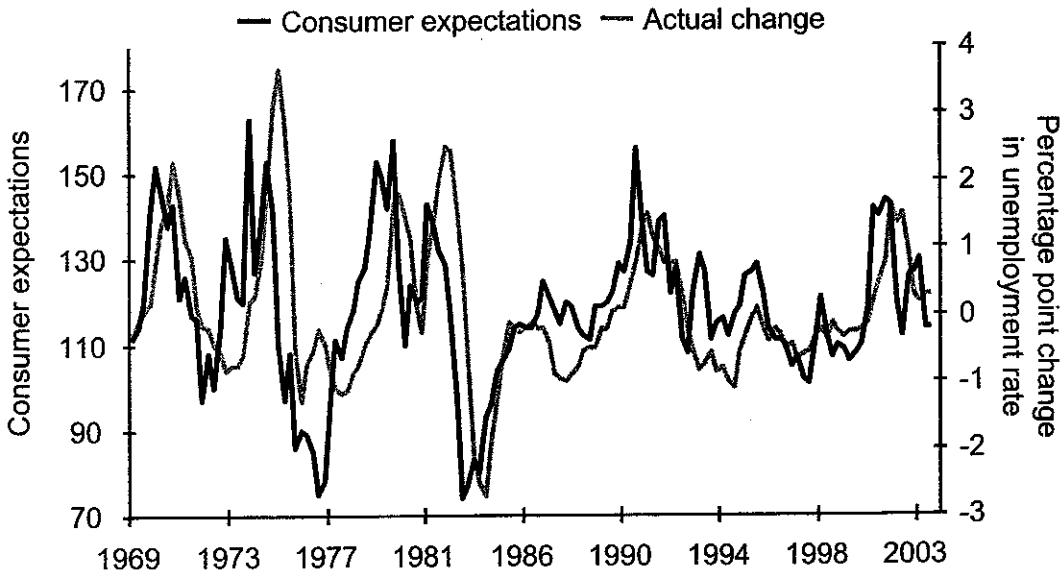

Figure 1.3 Consumer unemployment expectations and actual change in the US unemployment rate, 1969–2003. (Source: Curtin, 2003.)

1.3.4 The National Assessment of Educational Progress

What level of skills do elementary school children have in basic mathematics, reading, and writing? Do some schools foster more learning than others? Do children from lower-income families and racial/ethnic subgroups perform more or less well? Are there areas of the country or states where the children's achievement on mathematics and verbal tasks differ? How does the United States compare to other countries? Are achievement levels changing over time?

For many decades, there have been examinations that different school districts have given to their students from time to time. State boards of education often link state funding to the testing, so that whole states might administer the same test. Indeed, over the past few decades, politicians, with the support of parents, have urged more accountability by public schools, and this has led to widespread testing.

Unfortunately, despite the prevalence of testing, there was too little uniformity of the assessments. Each state, and sometimes each school district, makes independent decisions about which tests to use. Hence, the comparability of assessments over areas is low. Furthermore, not all schools use the standardized assessments. Often, districts with low levels of funding omit the assessments. Hence, the answers to the questions above are not coming systematically from administrative procedures used in the schools around the country. (Note how this resembles the situation with the NCVS and police reports of crime.)

The spur to uniform assessments of educational performance coincided with an increasing role of the federal government in education during the 1960s (Vinovskis, 1998). But creating a statistical survey to provide uniform statistics was politically controversial. This is a great example of how survey statistics can become so important that they themselves become a political issue.

First, some states did not want to be compared to others on assessment scores; hence, there was more support for a national survey limited to national estimates than one providing individual estimates for each state. Second, there was opposition to making the assessment a tool to provide individual student evaluations, for fear that the statistical purposes of the survey could be undermined by special efforts of schools, parents, and students to perform so well on the assessments that they would not reflect the national performance. There was also political opposition arguing that the money spent on assessments might be better spent on improving education itself. Finally, there were political ideologies that clashed. Some believed strongly in local control over schools and saw a national assessment of education as a threat that the federal government sought control over education.

Despite all these issues, the US Department of Education launched the National Assessment of Educational Progress (NAEP) in 1969 to produce national assessment summaries, and state-based samples (permitting accurate estimates for each state) were added in 1990. The NAEP is really a collection of surveys with three separate assessments: the "main national" (which provides the annual national figures), "main state" (which provides the state estimates), and "trend" (which is used for assessing change over time). Each of these assessments consists of four components: Elementary and Secondary School Students Survey, School Characteristics and Policies Survey, Teacher Survey, and Students with Disabilities or Limited English Proficiency (SD/LEP) Survey (for the main NAEP) or Excluded Student Survey (for the trend NAEP). In 1985, the Young

Table 1.4. Example Survey: National Assessment of Educational Progress (NAEP)

Sponsor	National Center for Education Statistics, US Department of Education
Collector	Westat
Purpose	Main objectives are to: • assess the academic performance of fourth, eighth, and twelfth graders in a range of subjects • reflect current educational and assessment practices • measure change over time
Year Started	1969
Target Population	National NAEP – schoolchildren in grades 4, 8, and 12 State NAEP – schoolchildren in grades 4 and 8
Sampling Frame	US elementary and secondary school children, through US counties or groups of counties, listed schools and students within schools
Sample Design	Multistage, stratified clustered area probability sample of primary sampling units (PSUs); sample of schools within PSU drawn, classrooms of students
Sample Size	2,000 schools and 100,000 students (National NAEP) 100 schools and 2,500 students per subject grade (State NAEP sample size per state) [typical sample sizes]
Use of Interviewer	None; self-administered background questionnaires completed by students, teachers, and principals; cognitive assessments completed by students; proctored by survey administrators
Mode of Administration	Paper-and-pencil self-administered questionnaires and cognitive assessment instruments
Computer Assistance	None
Reporting Unit	Students, teachers, school principals
Time Dimension	Repeated cross-sectional survey
Frequency	Conducted annually
Interviews per Round of Survey	One
Levels of Observation	Student, class, school
Web Link	*http://nces.ed.gov/nationsreportcard/sitemap.asp*

Adult Literacy Study was also conducted nationally as part of NAEP, under a grant to the Educational Testing Service and Response Analysis Corporation; this study assessed the literacy skills of persons 21- to 25-year-old. In addition, NAEP

conducts a high school transcript study. In this book, we will concentrate our remarks on the main national NAEP.

The National Center for Education Statistics, a federal government statistical agency, sponsors NAEP, but the data are collected through a contract, like NSDUH. The Educational Testing Service, a company that conducts the Scholastic Aptitude Test (SAT) and other standardized tests, develops the assessments. Westat, a survey firm, designs and selects the sample, and conducts the assessments in sample schools. NCS Pearson, another testing and educational assessment firm, scores the assessments.

The national NAEP selects students enrolled in schools in the 50 states and the District of Columbia in grades 4, 8, and 12. The sample design is clustered into primary sampling units (counties or groups of counties), just like NCVS and NSDUH (this is a cost-saving effort for obtaining cooperation from school administrators and for the administration of the assessments). After the stage of sampling areas, schools are selected. Within each sampled school, students from the appropriate grades are selected directly, so that they are spread over different classes in the school. NAEP measures different subjects over the years. Further, the design assigns different assessments to different students in each selected school. The content of the tests vary randomly over students, while measuring the same subject.

Each assessment is built around a well-specified and evaluated conceptual framework. These frameworks define the set of knowledge that underlies various levels of sophistication of the student's understanding. A complicated and lengthy consensus process involving teachers, curriculum experts, parents, school administrators, and the general public produces the framework. Then experts in the field write and pretest individual questions representing the components of the framework. In contrast to the surveys reviewed above, there are multiple questions measuring the different components within the framework. For example, in the mathematics framework, there would be different types of mathematical knowledge within the framework (e.g., for 4th grade, addition, subtraction, multiplication, division).

The NAEP is used as an important indicator of the performance of the nation's schools. Funding advocates for educational initiatives use the information to justify levels and targets of funds. NAEP provides results regarding subject-matter achievement, instructional experiences, and school environment for populations of students (e.g., fourth graders) and subgroups of those populations (e.g., female students, Hispanic students). The results are highly watched statistics within the political and policy domains. Figure 1.4 presents the average scores (scaled so that 500 is a perfect score) on the mathematics assessment for high school seniors by type of school. This shows a very common finding: that students in private schools tend to attain higher scores, that Catholic schools are in the middle, and that public schools have the lowest average scores. Much of this is related to what types of students attend these schools. There in no uniformly increasing trend in the scores of students; for example, the year 2000 scores appear to decline except among the Catholic school students. However, all were higher than their corresponding scores in 1990. When declines in the scores occur, there is usually some public discussion and policy debate about the alternative causes of the decline and the desirability of change in educational policies.

A CNN story of June 20, 2003 led with "Are Little Kids Smarter than High Schoolers?" The story was:

*The
rar

Figure 1.4

"Fc
ski
rep
gra
ana
cie
36
gra
is t
pus
cen
fro
our
ing
dau
the
rea

These a
there are gro
of 1998, bas

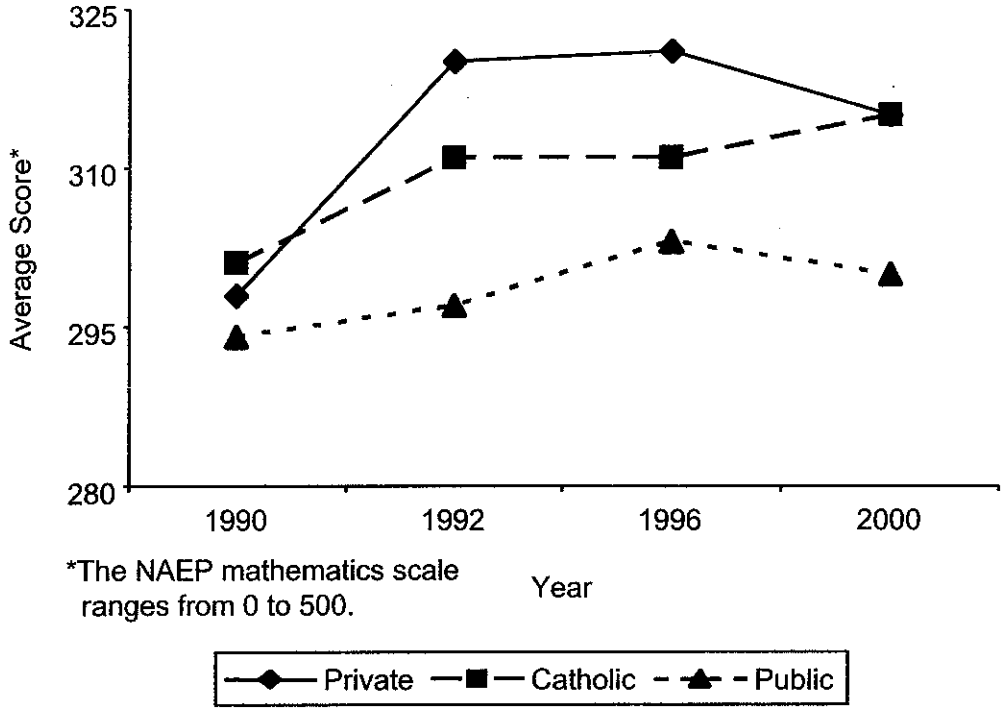

Figure 1.4 Average scale scores on grade 12 mathematics assessment, by year by type of school. (Source: US Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1996, and 2000 Mathematics Assessments.)

> "Fourth graders are showing they are better readers, while the skills of 12th graders are declining, the government said in a report Thursday. ... Overall, less than one-third of fourth graders and eighth graders showed they could understand and analyze challenging material. That skill level, defined as proficient, is the focal point of the test. Among high school seniors, 36 percent hit the mark. Four years ago, 29 percent of the fourth grader were proficient. That increased to 31 percent in 2002. It is those younger students who are at the center of a national push to improve basic education. Among the seniors, the percentage of those who reached the highest skill level dropped from 40 percent. ... 'There are no scientific answers as to why our high school seniors have performed so poorly on this reading assessment, but we're still searching for solutions to these daunting challenges,' said Education Secretary Rod Paige. 'At the same time, we know what works to teach youngsters to read, and we know that all children can learn.'"

These are complex results and require more analysis to disentangle whether there are group differences between the 12th graders of 2002 and the 12th graders of 1998, based on experiences they had in grammar school, or whether there are

contemporaneous education policies producing the results. The absence of firm data on the causes produces many alternative speculations from policy makers about the alternative solutions. Despite the complexity of disentangling causes of the assessment levels, understanding these phenomena requires the kind of uniform measurement that NAEP offers.

1.3.5 The Behavioral Risk Factor Surveillance System

How many people exercise, smoke cigarettes, or wear seatbelts? Do the states within the United States vary on these phenomena? How do such health-related behaviors vary across the states? Do people exhibit more or fewer healthy behaviors as they get older? Is there change over time in health behaviors that are associated with public health education programs in the states?

Since 1965, the National Center for Health Statistics had been providing annual US survey estimates on a wide variety of health behaviors and conditions (e.g., self-reported health conditions, frequency of visits to doctors, self-reported exercise, and other risk-related behaviors). The survey estimates, coupled with biomedical research, clearly indicated that individual personal behaviors affected premature morbidity and mortality. Much public health policy and oversight rests at a state level, however, and there were no comparable set of statistics at the state level. State health agencies have a lead role in determining resource allocation to reduce health risks linked to behaviors.

Since the early 1980s, the Behavioral Risk Factor Surveillance System (BRFSS) coordinated by the US Centers for Disease Control and Prevention has provided state-level survey estimates of key health factors. In contrast to all the surveys described above, BRFSS involves a partnership of individual states with assistance from the Federal Centers for Disease Control. The states determine the questions and conduct the survey. The Federal Centers for Disease Control and Prevention coordinate the definition of a core set of questions, develop standards for the survey data collection, perform the postdata collection processing of the data, and distribute a national combined data set. The core questions ask about current health-related perceptions, conditions, and behaviors (e.g., health status, health insurance, diabetes, tobacco use, selected cancer screening procedures, and HIV/AIDS risks) and questions on demographic characteristics.

Like the SOC, BRFSS uses random-digit-dialed samples of the telephone household population. Unlike SOC, each state draws an independent sample, and arranges for the data collection via computer-assisted telephone interviewing. Most states uses commercial or academic survey organizations under contract. Some collect the data themselves. The sample sizes vary across the states. In each sample household, an adult (18 years old or older) is selected to be interviewed.

The survey publishes annual estimates of rates of smoking and other risky health behaviors for each of the states. The estimates are used as social indicators of the health status of the population and guide government health policies that might affect health-related behavior. For example, the three Figure 1.5 maps show the dramatic increase in obesity in the US in the 1990s. The statistic is presented separately for each state, using a different shading to denote the outcome. In 1994, there were no states with more than 20% of the adult population who were obese; that is, with a body mass index equal to 30 or more. By 2001, over half of the states had that prevalence of obesity. By tracking trends like this at a state level,

Table 1.5. Example Survey: Behavioral Risk Factor Surveillance System (BRFSS)

Sponsor	US Centers for Disease Control and Prevention
Collector	Varies by state – for the 1998 BRFSS, 21 state or jurisdictional health departments collected their own data, whereas 31 used outside contractors
Purpose	The main objectives of the BRFSS are to: 1) collect uniform, state-specific data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases in the adult population 2) enable comparisons between states and derive national-level conclusions 3) identify trends over time 4) allow states to address questions of local interest 5) permit states to readily address urgent and emerging health issues through addition of topical question modules
Year Started	1984
Target Population	US adult household population
Sampling Frame	US telephone households, through lists of working area codes and exchanges, then lists of household members
Sample Design	Varies by state. For the 1999 BRFSS, all but three states (Alaska, California, and Hawaii) used probability designs
Sample Size	Average state sample size of 3,075 (1999 BRFSS)
Use of Interviewer	Interviewer-administered
Mode of Administration	Telephone interview
Computer Assistance	Computer-assisted telephone interview (CATI) in 50 areas; two areas use paper-and-pencil interviews (PAPI)
Reporting Unit	Randomly selected adult
Time Dimension	Repeated cross-sectional survey
Frequency	Conducted annually
Interviews per Round of Survey	One
Levels of Observation	Person
Web Link	*http://www.cdc.gov/nccdphp/brfss*

state public health officials can both compare their state to others and compare trends over time in their state. States use BRFSS data to establish and track State health objectives, plan health programs, or implement a broad array of disease prevention activities.

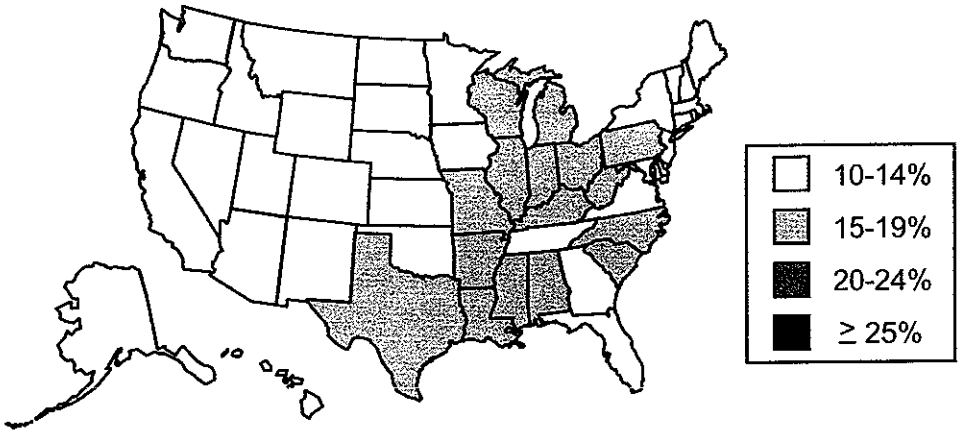

Figure 1.5a Percentage of state adults who are obese (body mass index ≥ 30) by state, 1994, BRFSS. (Source: BRFSS, CDC.)

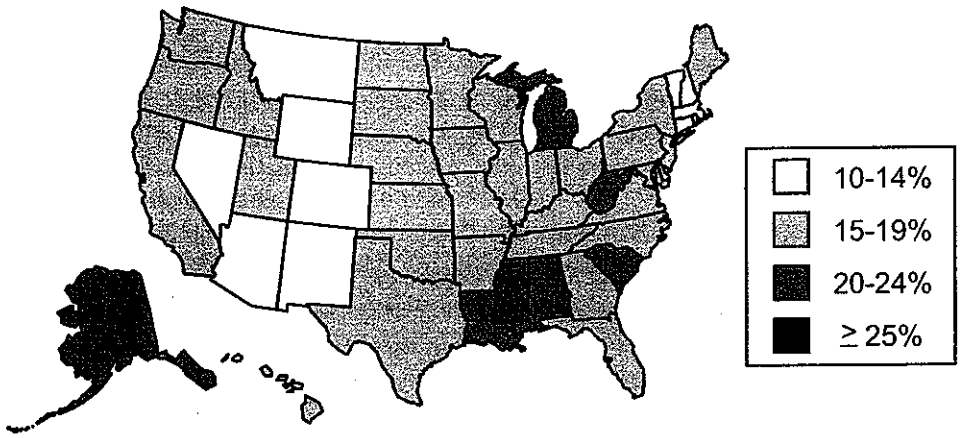

Figure 1.5b Percentage of state adults who are obese (body mass index ≥ 30) by state, 1998, BRFSS. (Source: Mokdad et al. *Journal of the American Medical Association* 1999;282:16.)

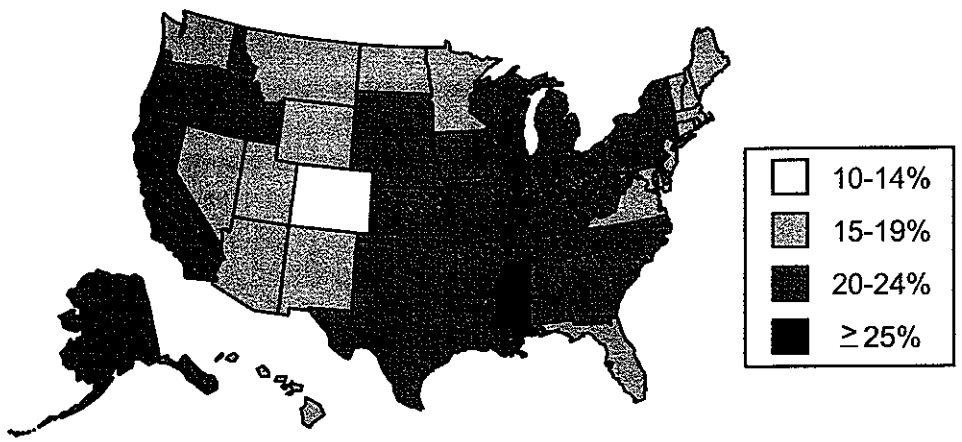

Figure 1.5c Percentage of state adults who are obese (body mass index ≥ 30) by state, 2001, BRFSS. (Source: Mokdad et al. *Journal of the American Medical Association* 2003;289:1.)

The BRFSS data quickly enter the national debate. Although there were other indicators of obesity problems in the US population, the BRFSS helped present a local perspective to the problem. The March 6, 2003 Washington Post ran the story, "When Spotlight Shines, San Antonio Citizens Seem to Swell to Fill It – City in Texas Called Obesity Capital of US" detailing the data. Policy officials acted. On May 14, 2003, the same paper ran a story entitled, "Health Costs of Obesity Near Those of Smoking. HHS Secretary Presses Fast-Food Industry." Then on July 2, 2003, the Washington Post ran an article titled, "Slimming Down Oreos: Kraft Plans to Make Its Food Products Healthier," describing plans to reduce portion size, alter content, and change marketing strategies within US school cafeterias.

Survey data like those of BRFSS can make a difference when they measure important issues and have the credibility to gain the attention of decision makers.

1.3.6 The Current Employment Statistics Program

How many jobs were created in the US economy in the last month? Are some industries changing their employment counts more rapidly than others? Which industries are growing; which are declining? Are large employers producing most of the employment change or does the dynamic nature of the job market lie among new, small firms? Are there regional differences in growth or decline in jobs?

The Current Employment Statistics (CES) program of the US Bureau of Labor Statistics is a focus of the story that opens this chapter. CES is one of two parallel surveys that attempt to provide monthly estimates of the employment situation. The CES is a survey of employers that asks for six different pieces of information – the total number of employees on the payroll, number of women employees, number of production workers, the total payroll of the production workers, the total hours worked by production workers, and the total number of overtime hours by production workers. The Current Population Survey (CPS) is a household survey that asks residents whether they are employed or looking for work. That survey produces a monthly rate statistic – the unemployment rate.

The sample of the CES is large – over 160,000 employers per month. The sample is drawn from a list of active employers who have registered their company with the state unemployment insurance agency. The sample design gives large employers very high probabilities of selection (some are permanently in the sample) and small employers small probabilities. Once an employer falls in the sample, it is repeatedly measured over months and years. Smaller employers rotate out of the sample and are "replaced" by other smaller employers; larger employers tend to stay in the sample for a longer time. CES is conducted as a partnership between the federal Bureau of Labor Statistics and the state employment security agencies.

The CES differs from the other surveys described above in that it uses many different methods of collecting the data simultaneously. Indeed, an employer is given great freedom to choose the method of delivering the answers to the six questions each month. Paper forms can be completed and mailed or faxed back to the Bureau of Labor Statistics. The respondent can use touchtone data entry, whereby the employer uses a telephone keypad to enter the digits of the answers in response to a recorded voice. The respondent can enter the data on a special

Table 1.6. Example Survey: Current Employment Statistics (CES)

Sponsor	US Bureau of Labor Statistics, US Department of Labor
Collector	US Bureau of Labor Statistics, State Employment Security Agencies
Purpose	The main objective of the CES is to produce monthly estimates of employment, hours, and earnings for the nation, states, and major metropolitan areas.
Year Started	1939
Target Population	US employers
Sampling Frame	US employers filing unemployment insurance tax records with State Employment Security Agencies
Sample Design	Originally quota sample; then a probability sample was fully implemented in 2002
Sample Size	About 160,000 business establishments
Use of Interviewer	Most self-administered; about 25% of respondents interviewed by telephone
Mode of Administration	Many sample units (about 30%) use touchtone data entry (TDE) to complete the interview by phone keypad, with prerecorded questions read to the respondent and answers entered by pressing touchtone phone buttons. Other modes used include mail, fax, Web entry, electronic data interchange, and computer-assisted telephone interviewing (CATI).
Computer Assistance	Touchtone data entry (TDE), electronic data interchange (EDI), Web entry, computer-assisted telephone interviewing (CATI)
Reporting Unit	Contact person at establishment
Time Dimension	Longitudinal panel survey of employers
Frequency	Conducted monthly
Interviews per Round of Survey	Variable
Levels of Observation	Establishment, employer
Web Link	*http://www.bls.gov/ceshome.htm*

secure Web page. The respondent can send electronic records directly to BLS for processing. The respondent can orally deliver the data in response to a telephone interviewer. These different methods, while each requiring some cost to develop, appear to fit different employers; the overall cooperation rate with the survey can be increased by allowing the sample employers to use a method easiest for them.

To illustrate how CES tracks the US economy, Figure 1.6 shows the total number of jobs in nonfarm employers between 1941 and 2001. Looking at the trend over so many decades, it is easy to see periods of no or slow growth in jobs

and periods of rapid growth. For example, the recessions of the early 1980 and early 1990s are followed by periods of high growth in number of jobs. The leveling off of the growth during the early 2000 recession is also obvious.

Statistics like these are important because they are used by the Federal Reserve, the White House, and by Congress to consider changes in economic policy. It is common that periods of recession contain many statements by politicians that the economy must be stimulated in order to create jobs for the country. Fairly or unfairly, political regimes rise and fall because of the values of such statistics.

1.3.7 What Can We Learn From the Six Example Surveys?

All of these are large-scale, national, repeated surveys. Five have lengthy histories – the Current Employment Statistics Program dates back to 1939, the Survey of Consumers originated in 1946, NAEP began in 1969, and the NCVS and NSDUH were started in the early 1970s. All of these are used by government agencies and social scientists as indicators of the status of society or the economy.

The design features of the surveys are tailored to fit their purposes. The target populations for these surveys are varied – for example, the NCVS aims to describe adults and children age 12 or older, whereas the CES survey has a target population of employers. The NCVS, NSDUH, and NAEP surveys all draw samples of households in two steps – first sampling geographic areas, then sampling from created lists of households within those areas. The Survey of Consumers and BRFSS draw samples of telephone numbers from all possible numbers within working area codes and exchanges. Then the numbers are screened to find those associated with households.

The example surveys use varying data collection methods. The NAEP survey does not use interviewers: respondents complete self-administered forms and assessments. The NCVS, BRFSS, and SOC all rely on interviewers to collect

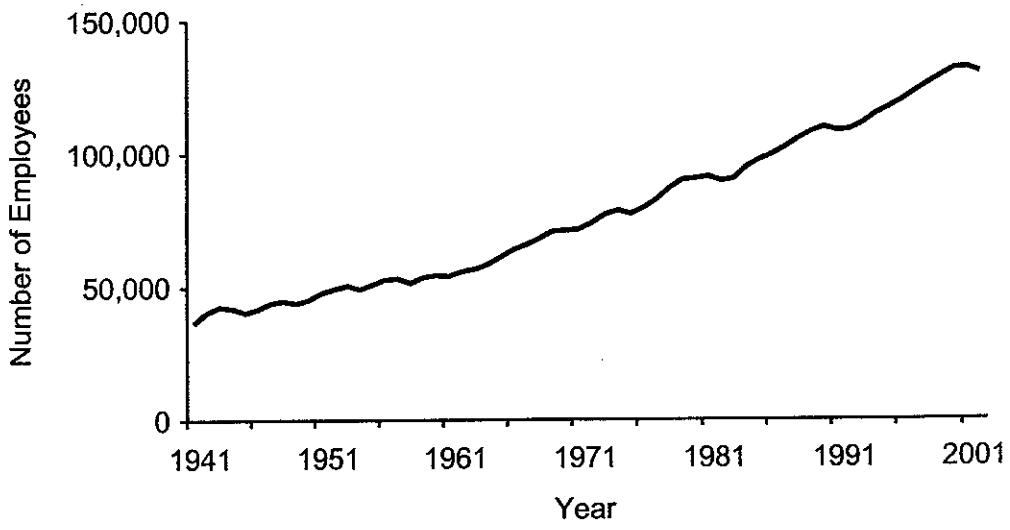

Figure 1.6 Number of employees of all nonfarm employers in thousands, annual estimates 1941–2002, Current Employment Statistics. (Source: www.data.bls.gov)

answers. The NSDUH uses both techniques, with some questions self-administered and others asked by interviewers. Another dimension on which the surveys vary is the use of computer assistance. The NAEP uses paper forms, followed by electronic processing of the forms. The CES uses many different types of computer assistance, including touchtone data entry, electronic data interchange, and computer-assisted telephone interviewing.

All of the surveys are ongoing in nature. This is because they all are designed to measure change in the phenomena they study. All of the designs supply estimates of change at the national level, measuring how population means and totals have changed over time. Those that collect data from the same people more than once, such as the CES, SOC, and the NCVS, can measure changes in individuals or employers, as well as overall changes in the population.

Many of the qualities and costs of a survey are determined by these kinds of design features. The first questions to ask about a new survey are:

1) What is the target population (whom is it studying)?
2) What is the sampling frame (how do they identify the people who have a chance to be included in the survey)?
3) What is the sample design (how do they select the respondents)?
4) What is the mode of data collection (how do they collect data)?
5) Is it an ongoing survey or a one-time survey?

Reread the tables that accompany Sections 1.3.1–1.3.6 to get an overview of the purpose and design of each example of survey. Throughout this book, we will refer to these surveys to illustrate specific methodological issues.

1.4 What is Survey Methodology?

Survey methodology seeks to identify principles about the design, collection, processing, and analysis of surveys that are linked to the cost and quality of survey estimates. This means that the field focuses on improving quality within cost constraints, or, alternatively, reducing costs for some fixed level of quality. "Quality" is defined within a framework labeled the total survey error paradigm (which will be discussed further in Chapter 2). Survey methodology is both a scientific field and a profession.

Within the scientific side of surveys, the achievement of high-quality survey results requires applying principles from several traditional academic disciplines. Mathematics, especially the principles of probabilities or chance events, is key to knowing the relative frequency of various outcomes. A whole subfield of statistics is devoted to the principles of sampling and inference from sample results to population results. Thus, the elements of sampling and analysis are historically grounded in the mathematical sciences.

However, because human beings often are involved either as interviewers or respondents, a variety of principles from the social science disciplines also apply to surveys. When we turn to the study of how data collection protocols affect survey estimates, the psychologists have been the leading sources of knowledge. Furthermore, social psychology provides the framework for understanding how

interviewer behaviors may influence the activities of respondents, both when they are recruited as respondents and during the survey interview. When it comes to understanding question design, cognitive psychology offers principles regarding how memories are formed, how they are structured, and what devices are helpful to the recall of memories relevant to the answers to survey questions. Sociology and anthropology offer principles of social stratification and cultural diversity, which inform the nature of reactions of subpopulations to requests for survey-based measurement or to particular questions. Computer science provides principles of database design, file processing, data security, and human-computer interaction.

Because survey methodology has this inherently multidisciplinary nature, it has only recently developed as a unified field. The scientific side of survey methodology was practiced largely outside traditional academic disciplines. For example, the important developments in probability sampling theory applied to surveys largely took place in large government survey organizations in the 1930s and 1940s. The major early texts in sampling were written by scientists with feet firmly planted in applied survey environments such as the US Bureau of the Census (e.g., Hansen, Hurwitz, and Madow, *Sample Survey Methods and Theory*, 1953; Deming, *Some Theory of Sampling*, 1950). Many of the early scientific contributions regarding survey data collection came from staff in government agencies conducting surveys of the military and citizenry during World War II. Major early texts in interviewing similarly were written by scientists in academic survey organizations, grappling with practical survey problems (Kahn and Cannell, *The Dynamics of Interviewing*, 1958; Hyman, *Interviewing in Social Research*, 1954).

It is consistent that the research literature or survey methodology is widely scattered. There is a concentration of survey methods in statistical journals: *Journal of the American Statistical Association* (Applications section), *Journal of Official Statistics*, and *Survey Methodology*. An important outlet for reports of survey methods studies is the *Proceedings of the Survey Research Methods Section of the American Statistical Association*. However, an interdisciplinary journal, *Public Opinion Quarterly*, is another major outlet for survey methods papers, and there are many papers on survey methods in journals of the various academic disciplines and applied fields such as health, criminology, education, and market research.

The field also has active professional associations, which act as gathering places for both scientists and professionals. There are four professional associations for which survey methodology is a central focus. The American Statistical Association has a large membership in the Survey Research Methods Section. The American Association for Public Opinion Research (and its international counterpart, the World Association for Public Opinion Research) is an association of commercial, academic, and government survey researchers. The International Association for Survey Statisticians is a component of the International Statistical Institute. The commercial survey organizations have created a trade association, the Council of American Survey Research Organizations, which promotes activities of benefit to the members. All of these organizations, including the Survey Research Methods Section of ASA, are interdisciplinary, including members with a wide array of formal training backgrounds in mathematical, social, or various applied sciences.

Survey methodology is also a profession, a set of occupational classes devoted to the design, collection, processing, and analysis of survey data. Throughout the world, there are specialists in survey methods found in academia, governments, and commerce. In the United States, the academic sector contains survey researchers who use the method to investigate discipline-based questions in sociology, political science, public health, communications studies, psychology, criminology, economics, transportation studies, gerontology, and many other fields. On campuses in the United States, there are over 100 survey research centers that have permanent staffs, providing survey capabilities for faculty and staff on their campus. The federal government plays an even larger role than academia in collecting and commissioning surveys. The Bureau of the Census, the Department of Agriculture, and the Bureau of Labor Statistics all collect survey data themselves, and over 60 other agencies commission numerous surveys. Moreover, in the United States, the survey efforts of the private commercial sector, including polls, political surveys, and market research, are several times larger than the combined effort of the agencies of the federal government.

Because the field of survey methodology did not develop within a single academic discipline, education in the field historically was rather haphazard. Professionals in the field are not certified, although from time to time discussions regarding the desirability of certification are held. Formal training in survey and marketing research techniques has existed for some years. Training in survey methods can be found in undergraduate and graduate departments in all the social sciences and in professional schools of all kinds. For those who specialize in survey methodology, almost invariably an apprenticeship in an organization that conducts surveys is part of the training process. Practical experience in solving the real challenges that designing and executing surveys entail is a complement to formal classroom training. One way to think about the purpose of this book is that it is designed to provide the intellectual foundation needed to understand and address the problems that survey research projects pose.

1.5 THE CHALLENGE OF SURVEY METHODOLOGY

Surveys are not the only way to collect information about large populations. There is no "right" way. Administrative record systems of governments and businesses sometimes offer all the necessary data for good decision making for those covered by the systems. Qualitative investigations, involving trained ethnographers or sociologists, offer the possibility of rich encounters with deep understanding of the perspectives of the subjects. Observations of behaviors of persons can yield quantitative information about the frequency of events in public spaces. Randomized experiments in controlled settings can answer important questions about whether various stimuli cause behaviors.

However, in administrative record systems, the researchers have little control over the measurement, and their results may be tainted by data of poor quality. Ethnographic investigations often use small groups of informants, limiting the studies' ability to describe large populations. Observations of persons are limited to a tiny fraction of all human behaviors. Randomized experiments face challenges of applicability in the real world.

Conversely, surveys often limit their measures to those that can be standardized and repeated over large numbers of persons. Surveys are conducted in the uncontrolled settings of the real world and can be affected by those settings. Surveys gain their inferential power from the ability to measure groups of persons that form a microcosm of large populations, but rarely achieve perfection on this dimension. Part of the task of a survey methodologist is making a large set of decisions about thousands of individual features of a survey in order to improve it. Some of the most important decisions include:

1) How will the potential sample members be identified and selected?
2) What approach will be taken to contact those sampled, and how much effort will be devoted to trying to collect data from those who are hard to reach or reluctant to respond?
3) How much effort will be devoted to evaluating and testing questions that are asked?
4) What mode will be used to pose questions and collect answers from respondents?
5) (If interviewers are involved) How much effort will be devoted to training and supervising interviewers?
6) How much effort will be devoted to checking the data files for accuracy and internal consistency?
7) What approaches will be used to adjust the survey estimates to correct for errors that can be identified?

Each of these decisions has the potential to affect the quality of estimates that emerge from a survey. Often, though not always, there are cost implications of how these decisions are made; making the decision that involves more effort or that has a better chance of minimizing error in the survey often costs more money.

There is a methodological research literature that provides the intellectual basis for understanding the implications of these various decisions on data quality. One of the most important goals of this book is to document what currently is known, and in some cases not known, about how these decisions affect data quality and the credibility of the data.

The second goal of this book is to convey an understanding of the concept of total survey error and how to use it. Essentially, all surveys involve some kinds of compromises with the ideal protocol. Some of those compromises are based on cost. Researchers have to decide how much to invest in each of the components of a survey listed above. When there is a finite budget, researchers often have to weigh the merits of spending more on one aspect of the survey, while reducing costs on another. For example, a researcher might decide to increase the size of the survey sample and offset those added costs by spending less on efforts to maximize the rate of response.

One challenge for the survey methodologist is to figure out how best to use the available resources – how to balance the investments in each of the components of a survey to maximize the value of the data that will result. Recognizing that each aspect of a survey has the potential to affect the results, the survey methodologist takes a total survey error approach. Rather than focusing on just one or a few of the elements of a survey, all the elements are considered as a

whole. A survey is no better than the worst aspect of its design and execution. The total survey error approach means taking that broad perspective and ensuring that no feature of the survey is so poorly designed and executed that it undermines the ability of the survey to accomplish its goals.

In a similar way, in some cases there are only imperfect solutions to survey design problems. All the approaches available to solve a particular problem may have pluses and minuses. The survey methodologist has to decide which of a set of imperfect options is best. Again, the total survey error approach is to consider the various ways that the options will affect the quality of the resulting data and choose the one that, on balance, will produce the most valuable data.

Survey methodology is about having the knowledge to make these tradeoff decisions appropriately, with as much understanding as possible of the implications of the decisions. When a well-trained methodologist makes these decisions, it is with a total survey error perspective, considering all the implications of the decisions that are at stake and how they will affect the final results.

1.6 ABOUT THIS BOOK

This book describes how the details of surveys affect the quality of their results. The second chapter is a discussion of what we mean by "error" in surveys. It turns out that this is a potentially confusing topic, because there is more than one kind of error that affects survey estimates. Different disciplines also use different words to describe the same or similar concepts. The focus of all the subsequent chapters is on how to minimize error in surveys. Understanding what is meant by error is an essential first step to understanding the chapters that follow.

The subsequent chapters take on different aspects of the design and execution of surveys. They lay out the decisions that face researchers, why they are potentially important, and what our science tells us about the implications of the available choices. Which choice is best often depends on the survey's goals and other features of the design. In a few cases, there is a clear best practice, an approach to carrying out surveys that consistently has been found to be critical to creating credible data. The chapters attempt to give readers an understanding of how these choices should be made. Throughout the book, we mix practical concrete examples (often drawn from the six example surveys described previously) with the general discussion of knowledge and principles, so readers can understand how the generalizations apply to real decisions.

The last two chapters cut across the others. In Chapter 11, we discuss issues related to the ethical conduct of survey research. In the final chapter, we answer a set of common questions that survey methodologists are asked. When readers have completed the book, we expect that they will have a solid foundation on which to build through reading, additional course work, and practical experience in order to become survey methodologists.

ABOUT THIS

KEYWORDS

analytic statist
area probabili
census
descriptive sta
error
probability sa
random-digit

FOR MORE

National Crin

Taylor, Bruce
Survey R
Measuren
Statistical
Florida. (

Pastore, Ann
Justice St.

National Surv

Turner, C. F.,
Drug Use

Gfroerer, J., E
National
SMA 03-
Administ

Surveys of Co

Curtin, Richa
sca.isr.um

Curtin, Richa
www.sca.

National Asse

Vinovskis, Ma
and Evol
US Depa

KEYWORDS

analytic statistic
area probability sample
census
descriptive statistic
error
probability sample
random-digit dialing
repeated cross-section design
rotating panel design
sampling error
statistic
statistical error
survey
survey methodology

FOR MORE IN-DEPTH READING

National Crime Victimization Survey

Taylor, Bruce M. and Rand, Michael R., "*The National Crime Victimization Survey Redesign: New Understandings of Victimization Dynamics and Measurement*," Paper prepared for presentation at the 1995 American Statistical Association Annual Meeting, August 13–17, 1995 in Orlando, Florida. (*http://www.ojp.usdoj.gov/bjs/ncvsrd96.txt*)

Pastore, Ann L. and Maguire, Kathleen (eds.) (2003), *Sourcebook of Criminal Justice Statistics* [Online]. Available: *http://www.albany.edu/sourcebook/.*

National Survey of Drug Use and Health

Turner, C. F., Lessler, J. T., and Gfroerer, J. C. (1992), *Survey Measurement of Drug Use*, Washington, DC: National Institute on Drug Abuse

Gfroerer, J., Eyerman, J., and Chromy, J. (eds.) (2002), *Redesigning an Ongoing National Household Survey: Methodological Issues*, DHHS Publication No. SMA 03-3768. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Statistics.

Surveys of Consumers

Curtin, Richard T. (2003), *Surveys of Consumers: Sample Design, http://www.sca.isr.umich.edu/documents.php?c=i.*

Curtin, Richard T. (2003), *Surveys of Consumers: Survey Description, http://www.sca.isr.umich.edu/documents.php?c=i.*

National Assessment of Educational Progress

Vinovskis, Maris A. (1998), *Overseeing the Nation's Report Card: The Creation and Evolution of National Assessment Governing Board*, Washington, DC: US Department of Education.

US Department of Education. National Center for Education Statistics. *NCES Handbook of Survey Methods,* NCES 2003–603, by Lori Thurgood, Elizabeth Walter, George Carter, Susan Henn, Gary Huang, Daniel Nooter, Wray Smith, R. William Cash, and Sameena Salvucci. Project Officers, Marilyn Seastrom, Tai Phan, and Michael Cohen. Washington, DC: 2003.

Behavioral Risk Factor Surveillance System

Centers for Disease Control (2003), *BRFSS User's Guide, http://www.cdc.gov/brfss/usersguide.htm.*

Centers for Disease Control (2003), *BRFSS Questionnaires, http://www.cdc.gov/brfss/brfsques.htm.*

Current Employment Statistics

US Bureau of Labor Statistics (2003), *BLS Handbook of Methods, http://www.bls.gov/opub/hom/home.htm.*

US Bureau of Labor Statistics, *Monthly Labor Review, http://www.bls.gov/opub/mlr/mlrhome.htm.* (Web site contains frequent articles of relevance to CES.)

EXERCISES

1) Find the Web page for the National Crime Victimization Survey (first go to *http://www.ojp.usdoj.gov/* and look for the Bureau of Justice Statistics). Find the questionnaire of the NCVS and identify what types of victimizations are collected from the household respondent and what types of crimes are collected on each self-respondent.

2) Go to the BRFSS Web site (*http://www.cdc.gov/brfss/*) and complete the training regimen that interviewers are offered on the Web site.

3) Read the most recent annual report from the NSDUH (*http://www.samhsa.gov/oas/nhsda.htm*). For what drugs did the prevalence of use increase over the recent years?

4) Go to the technical notes for the Current Employment Statistics program (*http://www.bls.gov/web/cestn1.htm*). What is the current distribution of respondents by the various methods of returning the data to BLS?

5) Read the latest press release on the consumer confidence index from the Survey of Consumers Website (*http://www.sca.isr.umich.edu/main.php*), then go to a news service Web site to search for "consumer confidence." Compare the news treatment of the statistics to the report from the survey. What is the same; what is different?

6) Go th
(*http:/*
read it
the ch
data th
survey

6) Go the Website of the National Assessment of Educational Progress (*http://nces.ed.gov/nationsreportcard/*). Find the latest assessment report; read it to find comparisons with the past report. Try to find explanations for the changes that were observed. Do the authors cite reasons from the survey data themselves or do they cite possible external causes, not measured in the survey?