PART I

Perspectives on Telephone Survey Methodology

COPYRIGHTED MATERIAL

CHAPTER 1

Telephone Survey Methods: Adapting to Change

Clyde Tucker
U.S. Bureau of Labor Statistics, USA

James M. Lepkowski
University of Michigan, and Joint Program in Survey Methodology, University of Maryland, USA

1.1 INTRODUCTION

In 1987, the First International Conference on Telephone Survey Methodology was held in Charlotte, NC. The conference generated a widely read book on telephone survey methodology (Groves et al., 1988). Although that book continues to be a standard reference for many professionals, the rapid changes in telecommunications and in telephone survey methodology over the past 15 years make the volume increasingly dated. Considerable research has occurred since 1987, including myriad advances in telephone sampling in response to changes in the telecommunication system.

The goal of the Second International Conference on Telephone Survey Methodology was to once again bring together survey researchers and practitioners concerned with telephone survey methodology and practice in order to stimulate research papers that (1) contribute to the science of measuring and/or reducing errors attributable to telephone survey design, (2) provide documentation of current practices, and (3) stimulate new ideas for further research and development. This volume presents invited papers from the conference.

This chapter provides a brief introduction to the field, where it is today, and where it might be going. It begins by reviewing where the field stood at the time of

Advances in Telephone Survey Methodology, Edited by James M. Lepkowski, Clyde Tucker, J. Michael Brick, Edith de Leeuw, Lilli Japec, Paul J. Lavrakas, Michael W. Link, and Roberta L. Sangster
Copyright © 2008 John Wiley & Sons, Inc.

the 1987 conference and goes on to detail changes that have taken place since that time. Besides discussing the rapid changes in telecommunications and the social and political environments over the past two decades, the chapter considers ways telephone survey methodologists have adapted to these changes and what further adaptations may be needed in the future. The final section provides a brief overview of the contents of the volume.

1.2 THE CHANGING ENVIRONMENT

1.2.1 The Picture in 1987

Survey research began in the 1930s with the use of quota samples (Elinson, 1992). The controversy over the use of quota samples versus probability sampling lasted until the early 1950s. By that time, academicians and government statisticians had convinced most in the survey industry that probability sampling was a necessary ingredient in the proper conduct of surveys (Frankovic, 1992). But during all those years and through most of the 1960s, much of the survey research was conducted either by mail or through personal visits to households. The telephone was used largely for follow-up purposes. Certainly, the most important national surveys in the United States (e.g., the Gallup Poll, the Current Population Survey, and the National Election Study) were conducted face-to-face.

By the late 1960s, however, the costs of personal visits were escalating while, at the same time, the proportion of households with telephones had grown to close to 90 percent, both in North America and Europe. Furthermore, the decline in response rates in face-to-face surveys in especially the commercial sector made the possibility of using the telephone as a collection mode more attractive (Nathan, 2001). Concerns about the methodological shortcomings of telephone surveys were satisfied by the results of several studies conducted in the 1960s and 1970s (Hochstim, 1967; Sudman and Bradburn, 1974; Rogers, 1976; Groves and Kahn, 1979). Survey organizations began relying more and more on telephones for conducting surveys once random digit dialing (RDD) was introduced (Cooper, 1964; Nathan, 2001), even though the problem of locating residential numbers among the universe of possible numbers was daunting. That problem was solved by Warren Mitofsky and Joseph Waksberg with the invention of the two-stage Mitofsky–Waksberg telephone sampling methodology that took advantage of the fact that residential numbers tended to be clustered in 100-banks (Mitofsky, 1970; Waksberg, 1978), an approach also suggested but not fully developed by Glasser and Metzger (1972) and Danbury (1975). Thus, by the 1980s, telephone surveys were a part of standard survey practice; however, along with the growing reliance on the telephone survey came a number of methodological problems that had to be addressed.

It is in this environment that the first International Conference on Telephone Survey Methodology was held. One focus of that conference was telephone coverage, both in the United States (Thornberry and Massey, 1988) and in other countries (Steel and Boal, 1988; Trewin and Lee, 1988), and the potential for biased

estimates as a result of ignoring nontelephone households, defined only as those without landline service. At the time, this was largely assumed to be a fixed state—a household always had service or it never did. A great deal of interest about within household sampling also existed (Whitmore et al., 1988; Oldendick et al., 1988; Maklan and Waksberg, 1988).

As today, there also was the focus on sample designs for telephone surveys, but the range of discussion was much more restricted. Several papers at the conference discussed refinements of the Mitofsky–Waksberg method (Burkheimer and Levinsohn, 1988; Alexander, 1988; Mason and Immerman, 1988). Treatment of list-assisted designs was mostly limited to methods for accessing individual-listed phone numbers directly, although there was some mention of the type of list-assisted design we know today (Groves and Lepkowski, 1986; Lepkowski, 1988). Dual-frame and mixed-mode designs were covered (Nathan and Eliav, 1988; Lepkowski, 1988; Sirken and Casady, 1988), but these studies dealt only with combining telephone and address frames in the context of telephone or personal visit surveys. Issues surrounding variance estimation, survey costs, and weighting were also addressed (Massey and Botman, 1988; Sirken and Casady, 1988; Mason and Immerman, 1988; Mohadjer, 1988).

One of the most important topics at the 1987 conference was computer-assisted telephone interviewing (CATI), which was relatively new at the time. As might be expected, one study compared the results from CATI and paper surveys (Catlin and Ingram, 1988). The construction of CATI questionnaires was the focus of several papers (Futterman, 1988; House and Nicholls, 1988). Designing and using CATI systems were also addressed (Sharp and Palit, 1988; Baker and Lefes, 1988; Weeks, 1988).

The administration of telephone surveys, a topic still of interest today, was also important in 1987. Of particular interest were best practices in the administration of centralized CATI centers (Whitmore et al., 1988; Berry and O'Rourke, 1988; Bass and Tortora, 1988). Other topics in this area included the performance of telephone interviewers (Pannekoek, 1988; Oksenberg and Cannell, 1988) and the optimal calling strategies (Kulka and Weeks, 1988; Alexander, 1988).

Finally, two areas of research that have grown greatly in importance over the past two decades were covered in 1987—nonresponse and measurement error. You will note that much of this research revolved around the comparison of telephone and face-to-face surveys. Certainly, this research was sparked by the growing concern over the rising costs of in-person visits. Could telephone surveys be a viable alternative?

Unit nonresponse rates were already considered a problem in 1987, and Groves and Lyberg (1988) made it clear that the situation was likely to get only worse. Refusal rates in the United States, Canada, and Britain were reported in two studies that examined the differences between face-to-face surveys and telephone surveys (Collins et al., 1988; Drew et at., 1988). A rather limited amount of research examined possible reasons for nonresponse in telephone surveys, but few definitive results were obtained. Only the effects of interviewers on nonresponse were notable. Collins and his colleagues found differences in interviewer response rates, and Oksenberg

and Cannell (1988) found that response rates differed according to interviewer vocal characteristics. Interestingly, both studies found that the interviewer's ability to project confidence and competence coincided with higher response rates.

As for measurement error, a number of papers investigated indicators of data quality. Most of them looked at differences in estimates by mode, usually face-to-face and telephone. One paper was a meta-analysis of a number of comparisons between face-to-face and telephone surveys (de Leeuw and van der Zouwen, 1988). Nathan and Eliav (1988) looked at the consistency of reporting in a panel survey depending on the mode (telephone or face-to-face), and Kormendi (1988) examined the quality of income reporting in the two same modes. Sykes and Collins (1988) reported differences in estimates, again for telephone and face-to-face surveys, for a sensitive topic—close-ended, open-ended, and scale questions concerning alcohol consumption. Bishop et al. (1988) examined differences in data quality (the effects of question order and wording) in a telephone survey compared to a self-administered one. Just looking at data quality within a telephone survey, Stokes and Yeh (1988) evaluated the effects of interviewers on survey estimates.

1.2.2 Changes in Technology

Clearly, numerous challenges to conducting telephone surveys existed in 1987; however, those challenges may seem relatively small compared to the problems faced today. In 1987, the growth in the number of telephone service providers was yet to occur. The expansion in the number of area codes, leading to a dilution in the concentration of residential numbers among all available telephone numbers, was just in the planning stages. New technologies that coincided with the economic growth during the 1990s, such as answering machines, caller ID, and mobile phones, were not on the market. Computers were not yet in enough households to be considered as a vehicle for the administration of surveys. The public's concerns about privacy and confidentiality, while certainly present, had not reached a critical level. Important changes in the demographics of the U.S. population, such as the increased immigration of Hispanics in the 1990s, had not happened.

One of the most important developments since 1987 has been the rapid changes in telephony (Tucker et al., 2002). The number of area codes, and, thus, the total number of telephone numbers in the North American system, has almost doubled. The number of valid prefixes increased by 75 percent between 1988 and 1997, and today there are 90 percent more available telephone numbers. In contrast, the number of households has increased only a bit over 10 percent. As a result, the proportion of all telephone numbers assigned to a residential unit has dropped from about 0.25 to not more than 0.12. Figure 1.1 shows the relative change in the proportion of "active banks," those with one or more listed telephone numbers, since the late 1980s. Active banks have increased, but they now are a smaller percentage of all banks. A further complication is that it became more difficult to determine if numbers were residential because telephone business offices were less forthcoming with information and often inaccurate in their determinations (Shapiro et al., 1995). Screening numbers based on tritones was also problematic particularly in the west (Rizzo et al., 1995).

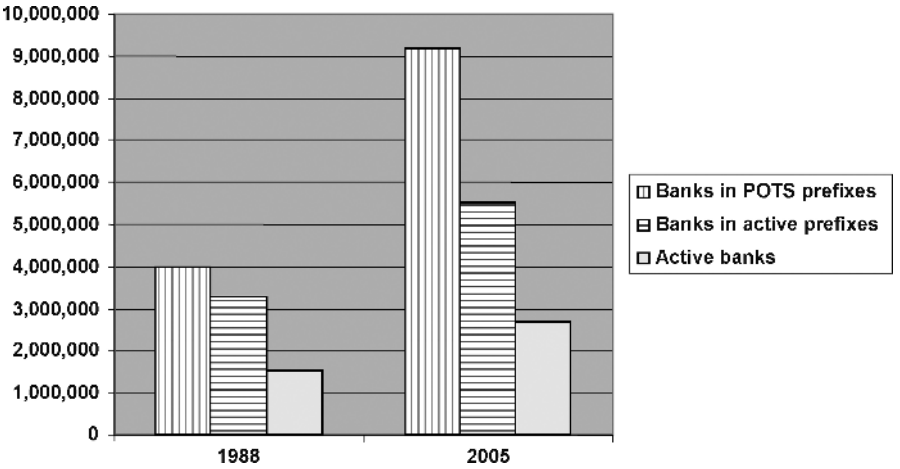

Figure 1.1. Number of 100-banks of various types in the United States, 1988 and 2005. (*Source*: Survey Sampling, Inc.)

There was also evidence that telephone companies appeared to be less systematic in the assignment of residential numbers across 100-banks. While the number of residences grew by just over 10 percent, the number of 100-banks with residential numbers has increased by over 50 percent. This increase in residential banks has resulted in a decline in the proportion of the numbers in a listed 100-bank that are residences. Figure 1.2 illustrates this change just for listed numbers. The decline has been from percentages in the low to middle 50s in 1990 to percentages in the upper 30s today. While the proportion of unlisted numbers is now approaching 30 percent in the United States and even higher in the United Kingdom (Collins, 1999), and largest in urban areas, this change cannot explain that much of the decline in the density of residential numbers in 100-banks.

In addition, there has been substantial growth in the number of households with multiple lines. Second lines dedicated to computers, fax machines, and home businesses have made it more difficult to distinguish noncontacts from nonworking numbers. Finally, there has been an increase in the assignment of whole prefixes to a single business customer. The identification of business numbers and the separation of those numbers from residential ones have become more problematic.

Accompanying these massive changes has been the amazing growth in telephone technology, and this has been a worldwide phenomenon. Besides fax machines and computers, call-screening devices have become commonplace in most homes. The first development in this area was the answering machine, and its presence in households grew dramatically during the 1990s (Nathan, 2001). In a recent Pew Research Center study (2006), almost 80 percent of U.S. households reported having either voice mail or an answering machine. Answering machines do have the advantage of potentially identifying residential numbers and allowing the interviewer to leave a message

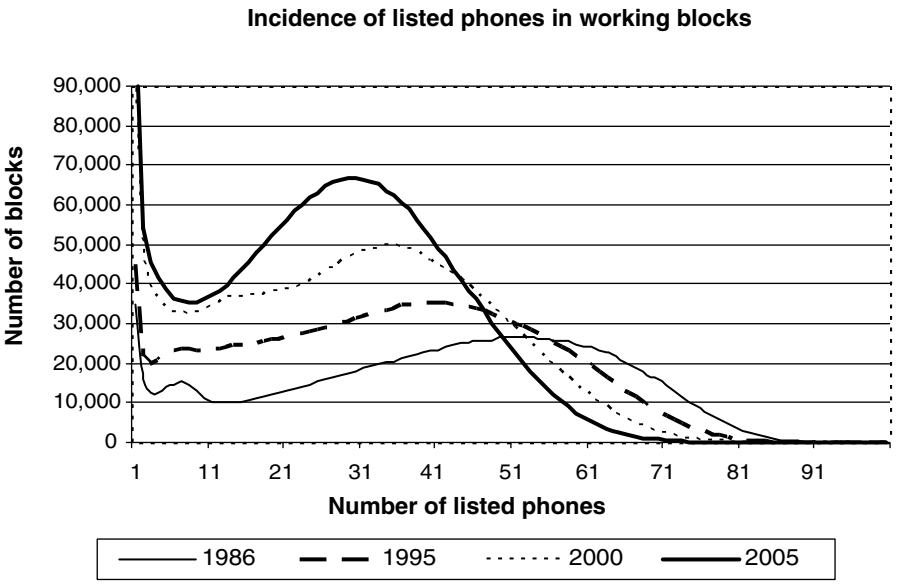

Figure 1.2. Frequency of 100-bank blocks by number of listed telephone numbers in the block. (*Source*: Survey Sampling, Inc.)

that distinguishes them from telemarketers, but they can also be used to screen out unwanted calls. With the advent of caller identification, or caller ID (one half of U.S. households now have it, according to the Pew Research Center (2006)), potential respondents could screen calls without waiting for a message to be left. Of course, households in countries such as Finland, which have heavy mobile-phone penetration (98 percent of adults have them (Kuusela et al., 2007, Chapter 4 in this volume), will have voice mail and caller ID. In fact, in Italy, the prevalence of answering machines has decreased due to the switch to mobile phones (Callegaro et al., 2007), and the proportion of households with answering machines has not changed over the past 6 years in France (Nathan, 2001; Vanheuverzwyn and Dudoignon, 2006).

While answering machines could be viewed as a device to increase the ability to communicate with the outside world, it is difficult to see how caller ID would do that. Seveal studies (Piazza, 1993; Xu et al., 1993; Oldendick and Link, 1994; Tuckel and O'Neill, 1996; Link and Oldendick, 1999) actually found that these new technologies were not having an appreciable effect on respondent cooperation in the mid-1990s, but, of course, they were unable to ascertain the effects from those potential respondents who do not answer the telephone. Now, according to the Pew study (2006), over 40 percent of U.S. households use caller ID, voice mail, or answering machines to screen calls.

Perhaps, the invention that will prove most disruptive to traditional telephone survey methodology is the mobile telephone. Until now, most telephone survey samples have been drawn only from banks of landline numbers. However, the widespread use of mobile technology, particularly the growing number of households with only

Table 1.1. Percent Distribution of Households by Telephone Status in the United States, 2004 (Current Population Survey)

Telephone status	Percent
Mobile and landline	46.4
Landline only	42.2
Mobile only	6.0
No telephone	5.4

mobile service, makes this methodology problematic. Table 1.1 gives the estimate of mobile-phone households in the United States in 2004 based on questions asked in a special supplement of the current population survey (Tucker et al., 2007). Figure 1.3 contains these numbers for the United States as well as selected European nations (Callegaro et al., 2005). Note that, while the United States had fewer than 6 percent mobile-only households, France and Italy had over 15 percent, and more than 33 percent of Finnish households were mobile-only.[1] Those individuals living

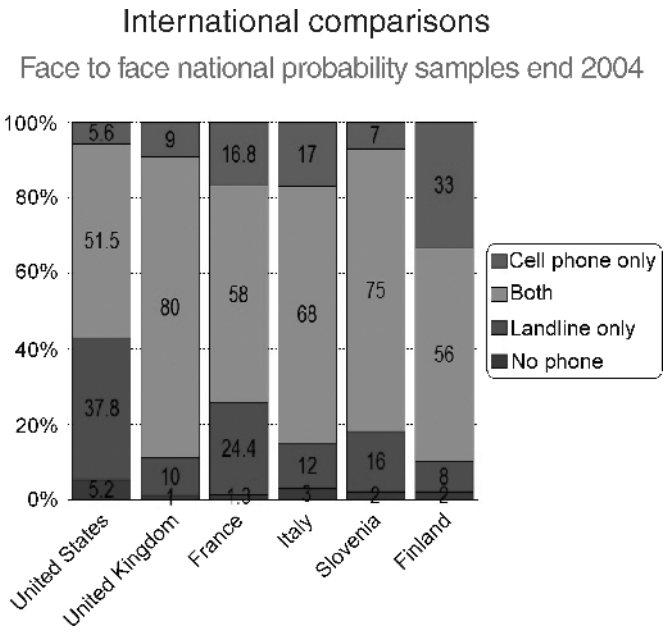

Figure 1.3. Distribution of mobile-only, landline only, mobile and landline, and phoneless households for six selected countries, 2004.

[1]Projections for the United States are, by the year 2007, the portion of mobile-only households will surpass the 10 percent level, and the percentage will be much larger for certain subpopulations, such as young, single adults.

in mobile-only households are more likely to be unemployed, have inadequate health care, and engage in risky behaviors compared to those in other households (Tucker et al., 2007; Blumberg, 2005).

Another concern is that half or more of the households in these countries had both mobile and landline phones. Those with both that rely mostly on their mobile phones might also be underrepresented in telephone surveys. Tucker et al., (2007) found that those households with both types of services receiving over half of their incoming calls on a mobile phone (a third of households with both) were more likely to be urban and younger.

Other technological developments could prove problematic to the conduct of telephone surveys (Piekarski, 2005). The Pew study (2006) reported that almost 20 percent of U.S. households have some form of electronic call blocking. Another new service, number portability, may undermine the ability to geographically target telephone samples in the years to come. Not only can a potential respondent move across the country and keep the same number, but also the number can be transferred to a wireless device without the caller being aware of that fact. Number portability has the additional troublesome effect of altering the probability of selection of a household without the respondent's knowledge. Call forwarding also presents the problem of conducting interviews on mobile phones, even when a landline has been called. Finally, with the rapid growth of home computers will come the switch to voice over Internet protocol (VoIP). Steeh and Piekarski (2007, Chapter 20 in this volume) report that it is estimated that up to 75 percent of worldwide voice traffic could be handled by VoIP by 2007. VoIP could increase the uncertainty of the location of particular telephone numbers, and VoIP service, of course, will be affected by electrical outages.

The advances in technology have had an additional effect on telephone surveys. With rapid computerization, the creation of the Internet and digital technology, new modes of survey administration have become available. As of fall 2003, the U.S. Department of Commerce (2004) reported that almost 62 percent of households had a computer and about 55 percent of households had Internet access at home (20 percent had broadband). The Harris Poll Online (2006) recently reported that over 60 percent of households now have Internet access at home. The web survey is the most widely studied alternative to the telephone survey, either as a stand-alone method or as one alternative in a multimode survey that also includes telephone administration. Many web surveys, especially marketing surveys, have employed nonprobability samples (Fischbacher et al., 1999; Poynter, 2000). Some establishment surveys early on used probability-based web designs, at least in the multimode context (Nusser and Thompson, 1998; Clayton and Werking, 1998; Tedesco et al., 1999). The same was true for some household surveys (Couper, 2000b). More recently, internet households have been recruited using probability-based methods, most notably RDD (Couper, 2000b). In some cases, the survey is administered only to those with Internet access, destroying the representativeness of the sample (Flemming and Sonner, 1999). Others attempt to remain representative by providing recruited respondents not already online with Internet access (Rivers, 2000; Couper, 2000b). Two problems arise in these latter surveys. Often the recruitment rate for what are ongoing Internet panels is quite

low. In a 2003 report, Knowledge Networks (Pineau and Slotwiner, 2003) indicated that, while coverage of the U.S. population was initially 96 percent for its panels, only 37 percent of the households contacted initially agreed to participate. In addition, the recruitment process and the installation of Internet equipment can prove costly when compared to telephone surveys with response rates at the same level or higher. Amortization across the Internet panel is possible, but then there is the matter of attrition.

As summarized by Nathan (2001), other advances in electronic communication have found their way into telephone surveys. Telephone surveys can now be conducted without live interviewers (computer-assisted self-interviewing or CASI) using interactive voice recognition (IVR) and touchtone data entry (TDE). These methods, especially TDE, were tested for both establishment and household surveys at the U.S. Bureau of Labor Statistics (Werking et al., 1988; Clayton and Winter, 1992; McKay et al., 1994). Turner et al. (1998) tested the use of telephone CASI for measuring sensitive items.

1.2.3 Changes in the Sociopolitical Climate

At least in the United States, significant societal changes have accompanied the technological changes over the past two decades. Some of these changes could pose problems for those engaging in telephone survey research by increasing nonresponse and/or exacerbating measurement error. Table 1.2 provides information on changes that might be related to respondent cooperation. In some cases, actually little has changed. The total hours worked is about the same; although, given the change in household size, the hours per household member has gone up. This might be explained by the fact that the unemployment rate in 2004 was 0.7 percent lower than in 1987. Hours spent at home is a little less now, but hours spent in leisure is

Table 1.2. Change in Select Characteristics of Households and Persons in Households Effecting Respondent Cooperation, United States (American Time Use Survey)

	Annual average	
Characteristic	1987	2004
Total hours worked in household	126.32	125.10
Percent households with only 1 member	7.82	10.66
Percent households with Hispanic reference person	7.06	11.09
Mean household size	2.64	2.52
	1985	**2003**
Hours spent at home per day	6:38	6:16
Hours spent in leisure time per day	4:33	4:59
	1987	**2004**
Percent married, but only male working	21.3	16.9
Percent of families with only male head and in labor force	13.5	17.6
Percent of expenditures spent on eating out	37.2	39.5

higher. There is little increase in the amount of eating out, at least in terms of the percentage of food expenditure. What has changed is the characteristics of households themselves. The percentage of one-person households (with the average size of the household declining) has increased and so has the percentage of Hispanic households. The percentage of households where the woman stays at home has declined considerably, continuing a trend begun in the 1950s that accelerated in the 1970s and 1980s. Note also the growth in the number of households headed by only one worker. These changes may have led to increased time pressures within the American household and could explain the increases in noncontact rates as well as the decline in cooperation rates in telephone surveys over the past two decades. The larger percentage of Hispanic households may be just one indicator of the growing diversity among households making standardized survey procedures less effective. An important question is what is happening in other countries with respect to demographic change?

The attitudes of respondents (and of governments), especially in the areas of privacy, confidentiality, and respondent burden, have also changed over time. Singer and Presser (2007, Chapter 21 in this volume) present evidence in this volume of these attitudinal shifts based on a review of a number of studies involving mail, face-to-face, and telephone surveys. One note of caution is that studies of privacy, confidentiality, and respondent burden are plagued by the lack of information from nonrespondents, and, as we will see later, nonresponse in telephone surveys has been increasing. Singer and Presser found that the public's growing concerns about privacy and confidentiality coincided with the increased alienation and mistrust of government that began during the Vietnam War and Watergate (Westin, 1967). (For information on the explicit connection between political events and response rates, see Harris-Kojetin and Tucker (1999).) In more recent years, these concerns may have intensified with the advent of more sophisticated methods for monitoring the actions of the individual that came about as the result of computerization. The easy access to personal information, especially in the commercial sector, may have fueled the concern about privacy and confidentiality. Also, some technologies may result in more sensitivity to privacy concerns. Calling on a mobile phone may be considered more of an invasion of privacy than calling on a landline.

In general, privacy concerns in the United States have increased over the years. A good part of that increase, however, occurred prior to 1987. This is not surprising given the links to Vietnam and Watergate. Figure 1.4 shows the changes in trust of the Federal government from the National Election Studies since 1958. Note that a steady decline took place from 1964 to 1980. Although trust increased somewhat after that, it has never reached the levels recorded in the early 1960s, and, in fact, 1994 was a low point. That was the year the Republicans had the "Contract with America" and took control of the House of Representatives. There does seem to be more concern on the part of the U.S. public about privacy and confidentiality relative to the government than to business. In other countries, the findings on concerns about privacy and confidentiality vary widely by country. Some countries had increases in concern across time, and others had decreases. To the extent that privacy

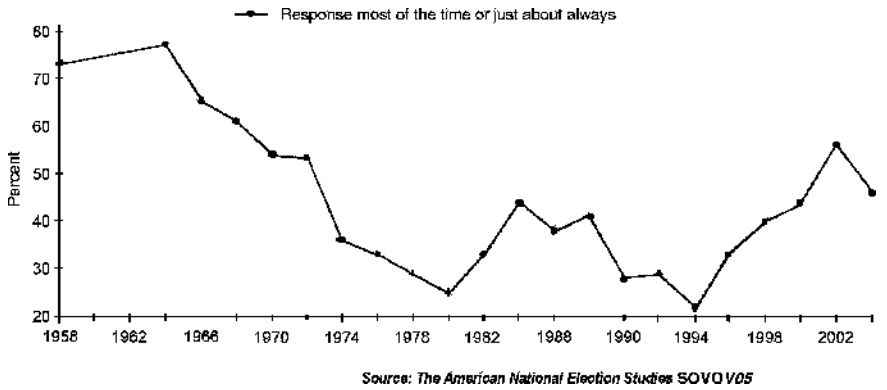

Figure 1.4. Trend in trust in U.S. federal government, American National Election Studies, 1958–2003.

concerns (and survey participation) are related to alienation and mistrust, perhaps, a more telling indicator is the decline over time in election turnout in several democracies. Figure 1.5 shows these trends, compiled largely by the International Institute for Democracy and Electoral Assistance, for selected democracies. In most of the countries, there has been a decline in turnout in general elections, with only the Netherlands and the United States countering these trends. The declines (particularly in Japan, Canada, Italy, and the United Kingdom) tend to be greater than in earlier periods.

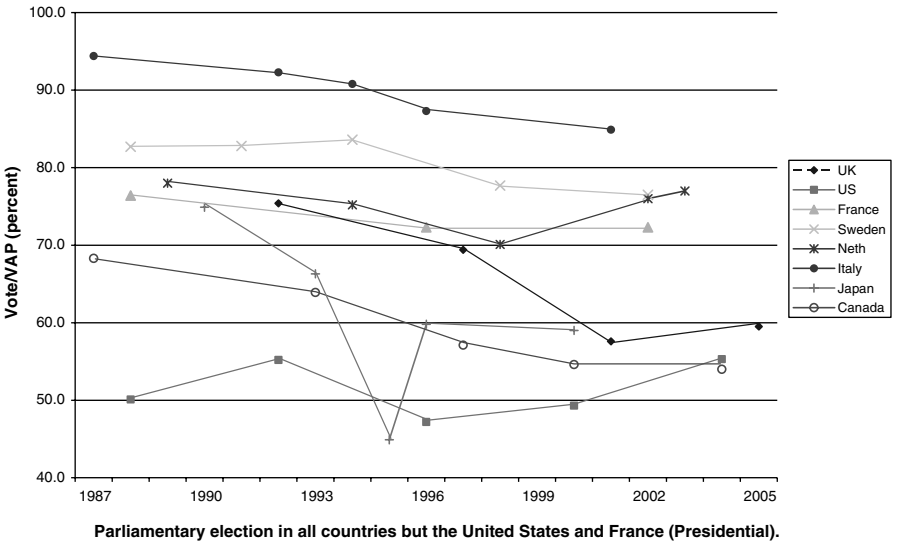

Figure 1.5. Turnout in presidential or parliamentary elections in eight selected countries, 1987–2005. Tucker, C., *Telephone Survey Methods: Adapting to Change*; © Institute for Democracy and Electoral Assistance, 2006.

Singer and Presser report that feelings about privacy and confidentiality seem to have relatively small, but detrimental effects, on a potential respondent's willingness to participate in a survey, and this effect is accentuated in the case of item nonresponse for sensitive items. But it does seem clear that the participation of those concerned about privacy and confidentiality is lower than that for those who are not. Unfortunately, assurances of confidentiality on the part of the survey organization can have either a positive or negative effect on response rates.

In terms of respondent burden, one clear finding from Singer and Presser's work is that the public either is or perceives it is receiving more survey requests today than ever before. The relationship between actual respondent burden (measured in terms of length of the interview) is less clear; perhaps, because refusals often come before the potential respondent knows the length of the survey (DeMaio, 1980). In contrast, the length of the interview could be related to the quality of the respondent's answers. Tucker et al., (1991) found this to be the case for telephone interviews collecting information about consumer expenditures. Finally, perceived burden can be much different from actual burden, and actual burden can depend on how hard a respondent is willing to work at the survey task. Tucker (1992) found that younger respondents to a consumer expenditure survey reported lower burden than older respondents, but the expenditure reports received from the older respondents were of a higher quality than those from younger respondents.

Whether or not potential survey respondents are more concerned about privacy and confidentiality today than in the past, it certainly is true that governments are. In 1991, the U.S. Congress passed the Telephone Consumer Protection Act (TCPA), and the Federal Communications Commission (FCC) issued a directive the next year restricting the use of automatic and predictive dialers (Dautch, 2005). The FCC rules apply only to telemarketers except in the case of mobile phones, where survey researchers also are prohibited from using autodialers. In 2003, the FCC established the Do Not Call (DNC) registry. Some states already had such lists. Telemarketers, but not legitimate survey researchers, were prohibited from calling the numbers placed on the list by individual consumers. The European Union (EU) passed the Directive on Data Protection in 1995 to protect the confidentiality of an individual respondent's survey data. Britain's Communications Act of 2003 provides legal recourse for those disturbed by unwanted calls, and Canada introduced similar legislation in 2004 (Singer and Presser, 2007, Chapter 21 in this volume).

Although concerns about confidentiality, privacy, and respondent burden have grown over the years (along with the decline in trust in government), there is no clearly documented correlation at the microlevel between these trends and the increases in nonresponse in telephone surveys, just as it is difficult to establish a one-to-one connection between demographic changes and changes in survey participation. In contrast, the trends do coincide over the past generation, at least in the United States, and may not bode well for the future of telephone surveys.

1.2.4 Problems Resulting from Changes

One of the more important issues resulting from the growth in the universe of telephone numbers was the increasing inefficiency of RDD designs. It became more

difficult to find residential numbers using the old methods. The problem was exacerbated by the unsystematic way numbers were now being assigned. The likelihood of locating a residential number using the Mitofsky–Waksberg design declined from just over 60 percent to less than 40 percent (Steeh and Piekarski, 2007, Chapter 20 in this volume). This situation posed a problem for production telephone surveys. Either the number of 100-banks needed or the number of numbers *(k)* selected per bank would have to be increased. In either case, more would have to be spent to get a given effective sample size. If the *k* were increased, the intraclass correlation would be larger, resulting in a greater variance, and the number of calls to reach *k* households would rise. In addition, a number of survey organizations were finding the Mitofsky–Waksberg method cumbersome to manage, and because it was a clustered design, the variances were greater than desired.

With the changes in technology came the increase in the number of telephone lines into the household. Besides dedicated fax and computer lines, there were now mobile phone numbers and, with the growth of telecommuting, probably more lines in the home dedicated to business. Associated with these developments was the increase in the difficulty of determining the number of lines on which respondents could actually be reached by survey organizations. Furthermore, because of dedicated computer and fax lines as well as caller ID, it was not an easy task to identify the reasons for ring-no-answers (RNAs). At least, with answering machines or voice mail, the ability to locate a residence may have reduced the number of RNAs, but this would not necessarily lead to respondent cooperation.

The explosion in technology also gave rise to the consideration of multimode surveys. In particular, the combination of a telephone survey and a Web survey seemed attractive. Weeks (1992) discussed alternative modes that had been made possible through technological advances. One problem with multimode surveys, however, is that they open up the possibility of mode effects on estimates. Secondly, it was not clear if the design of CATI surveys could be easily transferred to the Web, or whether new design principles would have to be applied. Furthermore, the Web realistically could be used only as an alternative mode for those households with Internet capabilities.

The growth in the mobile-only population posed a coverage problem, but conducting a survey over mobile phones was problematic. How easy would it be to contact residential mobile phone owners? What about the problem of finding enough mobile-only households? Given the increasing current noncontact rates, should households with both landlines and mobile service also be candidates for a mobile phone survey? Could mobile phone respondents be convinced to do the survey, especially in the United States, where the mobile phone owner pays for the call? There were also ethical issues with which to contend (e.g., the appropriateness of conducting a survey while the respondent is driving). Perhaps, the most troublesome problem was a conceptual one. Does conducting a survey on mobile phones imply changing the sampled unit from the household to the individual?

The changes in society (both technological and social) over the past 20 years made it increasingly difficult to maintain response rates in telephone surveys with the passage of time. This situation created an even greater threat of nonresponse

bias in survey estimates. Battaglia et al., (2007, Chapter 24 in this volume) illustrate the decline in response rates with examples from telephone surveys ranging from the Survey of Consumer Attitudes (SCA) (Curtin et al., 2005) to the state-based Behavioral Risk Factors Surveillance System (BRFSS). Recently, Curtin et al., (2005) showed that the overall response rate in the SCA declined considerably from 1997 to 2003 at the average annual rate of 1.5 percentage points to 48.0 percent. The National Household Education Survey (2004) reported a decline in the response rate from 72.5 percent in 1999 to 62.4 percent in 2003, an annual rate of decline of 2.5 percentage points. The U.S. Centers for Disease Control and Prevention (2003) reported that the BRFSS indicated a decline in the median response rate for the 50 states from 68.4 percent in 1995 to 53.2 percent in 2003 (an average decline of 1.9 percentage points per year). The RDD component of the National Survey of America's Families (2003) reported a decline in the overall response rate from 65.1 percent in 1997 to 62.4 percent in 1999 and to 55.1 percent in 2002 among the surveys of children, and 61.8, 59.4, and 51.9 percent among the adult surveys in 1997, 1999, and 2002, respectively. Finally (Holbrook et al., 2007, Chapter 23 in this volume), in their review of surveys from 1996 through 2003 conducted by 14 private U.S. survey organizations, found a strong negative correlation between year of the survey and response rate. They also reported that contact rates declined more than refusal rates. The same was true for face-to-face U.S. government surveys (Atrostic et al., 2001). de Leeuw and de Heer (2002) found that this trend toward increasing nonresponse rates held when looking across a number of countries.

1.3 ADAPTING TO THE CHANGING ENVIRONMENT

1.3.1 Adapting to Changing Technology

In the late 1980s, survey methodologists began searching for alternatives to the Mitofsky–Waksberg methodology. One such attempt was a dual-frame design that combined estimates from a sample of listed numbers with an RDD sample (Groves and Lepkowski, 1986). At the same time that there was growing dissatisfaction with the Mitofsky–Waksberg approach, several companies were developing sophisticated methods of processing files of residential listings, including Donnelly Marketing Information Systems. Although the listed telephone number frame itself was not suitable for direct sampling of telephone numbers because a substantial share of telephone households did not appear in the frame, by sampling numbers from 100-banks that contained listed telephone numbers, efficiencies obtained in the second stage of the Mitofsky–Waksberg could nearly be achieved. Sample selection could be simple-random, or stratified-random, selection of telephone numbers from across 100-banks containing one or more listed telephone numbers. The loss in precision due to cluster sampling was eliminated, and samples generated from list-assisted methods were less cumbersome to implement than the two-stage cluster method of Mitofsky–Waksberg. The residential hit rate for these designs was usually above 50 percent.

List-assisted methods were examined by Casady and Lepkowski (1993), who laid out the statistical theory underlying the design and presented empirical analysis on the properties of stratified list-assisted design options. Brick et al. (1995) showed that the potential bias resulting from the loss of residential telephones in 100-banks without listed numbers was small. Government agencies, academic survey organizations, and private survey firms subsequently adopted list-assisted designs.

As time went on, two commercial firms, Survey Sampling International and the Marketing Systems Group, appended information from other data sources to the records of listed numbers to make tailored list-assisted designs available along with automated sample delivery systems. Although the residential hit rate began to decline in the late 1990s, recent developments in methods to purge ineligible numbers from samples has brought the hit rates back to near the levels of the original Mitofsky–Waksberg design. The major question remaining was how many listed numbers should a 100-bank contain before being included in the sampling frame. Survey organizations today vary as to the minimum number of listed numbers required for a bank to enter the frame. This number is usually between 1 and 3, but with the better purging techniques, it may be possible in the future to actually include banks with no listed numbers without substantially adding to screening costs. To evaluate the efficiency of the list-assisted designs in the context of the growth in the size of the telephone number universe, Tucker et al., (2002) repeated Casady and Lepkowski's earlier analysis. They found that the list-assisted designs, while somewhat less efficient in an absolute sense, were relatively more efficient than the Mitofsky–Waksberg design than they had been 10 years earlier.

Another area of research has to do with selection of the respondent within the household. Binson et al. (2000) examined the effectiveness of using the Kish, Next-Birthday, or Last-Birthday method for selecting a respondent. As DeMaio (1980) found, most refusals occur before the respondent selection stage. It appears that the Kish method, requiring all household members be listed prior to respondent selection, causes the greatest problem. However, this problem seems to be an interviewer problem more than a respondent problem. In contrast, the birthday method was more prone to error (Lavrakas et al., 1992; Lavrakas et al., 2000; Lind et al., 2000). Rizzo et al. (2004a) proposed a more streamlined version of the Kish method that produced accurate results.

Coverage problems also have been a subject of study since 1987, particularly the coverage of nontelephone households. Weighting for undercoverage did not always solve the problem (Massey and Botman, 1988; Brick et al., 1992). Keeter (1995) suggested using households with intermittent telephone service to adjust for the coverage of nontelephone households. Using panel surveys, he showed that the characteristics of these "transient" households were much closer to those for nontelephone households than to those of households that always had telephone service. Brick et al. (1996b) found that collecting information on past breaks in telephone service in a cross-sectional survey could also be used to improve coverage in the same way Keeter had described. Brick et al. also used sociodemographic characteristics to improve the estimates.

As discussed in the previous section, in order to take advantage of the new technologies developed in the 1990s, survey methodologists began experimenting with multimode designs. Even with the growing problems with telephone surveys, it is likely to be retained as an important mode in multimode designs. Besides their work on combining directory listings and RDD, Groves and Lepkowski began looking at mixed-mode designs in the 1980s (Groves and Lepkowski, 1985). So did Whitmore et al. (1985). In 1997, a study of the combination of RDD and area sampling was reported by Waksberg et al. (1997) and Cunningham et al. (1997) in which the non-telephone households found in the area sample were provided with mobile phones for completing the CATI interview. Both the British Labor Force Survey (Wilson et al., 1988) and the one in the United States (current population survey) now use a mixed-mode design involving in-person interviews in the first interview and telephone interviewing in later months (U.S. Department of Labor and U.S. Department of Commerce, 2002). Pairing Internet surveys with telephone surveys is now a growing possibility (de Leeuw 2005). Biemer and Lyberg (2003) report that mixed-mode surveys are quite common today, and they have been treated in recent textbooks (Czaja and Blair, 1996; Groves et al., 2004).

There are several reasons for considering the use of a mixed-mode design. One is that coverage problems with each mode can be compensated for using the other mode, particularly when surveying special populations (Waksberg et al., 1997; Cunningham et al., 1997; Nathan, 2001; Edwards et al., 2002; Parackal, 2003; Srinath et al., 2004). As de Leeuw (2005) reports, one of the most important reasons for using a multimode design is to reduce nonresponse (Poe et al., 1990; Fowler et al., 2002). This is particularly the case when the modes are used sequentially (Jackson and Boyle, 1991; Japec, 1995; Alexander and Wetrogan, 2000; Grembowski and Philips, 2005; Dillman and Christian, 2005). The use of a Web survey in conjunction with a telephone survey is attractive because of the relatively low cost of the Web alternative (Couper, 2000; Dillman, 2000). One particular advantage that the Web survey might have over a face-to-face or telephone survey is eliciting more truthful answers regarding sensitive behaviors. Like mail surveys, the respondent is not reporting these behaviors directly to an interviewer (Aquilino and Losciuto, 1990; Presser, 1990; de Leeuw, 1992; Tourangeau et al., 1997; Fowler et al., 1998; Currivan et al., 2004). Sometimes face-to-face surveys may elicit more truthful answers than telephone surveys, but the evidence is mixed (Aquilino, 1994; Aquilino and Wright, 1996; Couper et al., 2003; Holbrook et al., 2003). de Leeuw (2005) did find a combination of the Web and telephone modes to have become more popular in market surveys.

Of course, as mentioned earlier, mixed-mode designs imply the possibility of mode effects in estimates. de Leeuw (1992) attributes mode differences to three possible sources—how a medium is used, how the information is transmitted, and the presence or absence of the interviewer. de Leeuw found that these factors appeared to have only small effects on estimates. Dillman et al., (1995) reported little in the way of primacy and recency effects when comparing mail and telephone surveys. When looking at the differences between web and telephone surveys, de Leeuw (2005) found some evidence of more extreme answers over the telephone (Oosterveld

and Willems, 2003; Dillman and Christian, 2005) but with few other differences (Oosterveld and Willems, 2003). It is possible that mode effects are quite complicated and not fully understood at this point (Biemer, 2001; Voogt and Saris, 2005). Unfortunately, the truth of the matter probably lies in the interactions between the mode, the interviewer, the respondent, and the survey content and presentation.

Although not a mixed-mode design, given the growth of mobile phone ownership (especially mobile-only households), another design that must be mentioned is a dual-frame design using both landline and mobile numbers. This type of design is very new and little is known about it. After several studies, Steeh (2004a) concluded that conducting a survey with mobile phone respondents can be challenging and require a different approach than used in traditional landline surveys. Brick et al. (2006) found that producing an unbiased composite estimate using this type of design was extremely difficult.

Over the past 20 years, some work has also proceeded on weighting and estimation issues in telephone surveys. One of the more important areas has been how to produce estimates for dual-frame and mixed-mode designs. Casady et al. (1981), Groves and Lepkowski (1985), and Kalton and Anderson (1986) had done work in this area prior to the 1987 conference. Skinner and Rao (Skinner, 1991; Skinner and Rao, 1996) extended the methodology by incorporating auxiliary data, like Brick et al. (1996), to better inform the compositing of the estimates. Variance estimation for these types of composite estimates was studied by Lohr and Rao (2000).

The other aspect of weighting that has been explored to some extent since 1987 is weighting for the number of phone lines in the household. Obviously, this task has been complicated by technological innovations. To determine the correct number of lines, the researcher must eliminate dedicated fax and computer lines as well as exclude mobile phone lines (if only sampling landlines). The problem has been exacerbated by the fact more workers are now telecommuting. Fortunately, with the growth of DSL and cable, the confusion over the number of lines has been reduced to some extent (Piekarski, 2005). Lavrakas (1993) and Czaja and Blair (1996) discuss procedures for determining the correct number of lines and using the information in weighting.

The other area of development over the past 20 years, and one that will have more prominence in the next section, is the cognitive aspects of survey methodology (CASM) movement. CASM received little, if any, attention at the 1987 conference even though the first CASM seminar was held in 1983 (Jabine et al., 1984). Since the 1987 conference, the study of the role of cognition in surveys has exploded. A second CASM seminar was held in the 1990s (Sirken et al., 2000), and the subject of the last international conference on survey methodology was devoted to questionnaire design and testing (Presser et al., 2004). Several papers in that conference specifically addressed questionnaire design for different modes of administration, including CATI (Tarnai and Moore, 2004). Cognitive methods also play a role in usability testing of CATI instruments (Presser and Blair, 1994; Schaeffer and Maynard, 1996; Hansen and Couper, 2004). Of course, there is now a rich literature on questionnaire issues and pretesting for all types of surveys (Tourangeau et al., 2000; Groves et al.,

2004; Conrad and Schober, 2005). Many of these studies look specifically at questionnaire design issues in telephone surveys that sometimes consider mode effects (Dillman and Tarnai, 1991; Dovidio and Fazio, 1992; Fowler, 1992; Conrad and Schober, 2000; Dashen and Fricker, 2001).

1.3.2 Adapting to the Changing Sociopolitical Climate

Certainly sociopolitical changes have occurred, at least in the United States, over the past 20 years. The characteristics of American households have changed, fairly dramatically in some ways; and concerns about privacy and confidentiality seem to receive a lot more attention now than 20 years ago. To the extent that these changes have affected telephone surveys, they have made it more difficult to successfully conduct these surveys. This fact is most clearly seen in the greater effort needed to locate and gain the cooperation of respondents (Curtin et al., 2000; Brick et al., 2003). This is reflected in the search for new methods to gain cooperation, the increasing costs of survey projects, and the greater focus on response rates and their calculation (Groves, 1989; McCarty, 2003; AAPOR, 2004).

The changing character of the American family, especially its growing diversity in several respects, has played a part in some survey researchers call for a move away from the standardized survey advocated by Fowler and Mangione (1990) toward tailoring procedures to the respondent (Tucker, 1992; Groves and Couper, 1998). This would include altering the survey introduction or methods for refusal conversion depending on the respondent's characteristics, and these efforts have also been driven by the application of psychology to the survey situation (Dijkstra and van der Zouwen, 1987; van der Zouwen et al., 1991; Morton-Williams, 1993; Groves and McGonagle, 2001; Mayer and O'Brien, 2001; Dijkstra and Smit, 2002; O'Brien et al., 2002; Shuttles et al., 2002). Others have argued that the interview should actually have more of the characteristics of a conversation (Schaeffer, 1991; Conrad and Schober, 2000). Besides improving response rates, this approach could reduce measurement errors.

Clearly, the move away from standardization is influenced by psychological theory. This involves not only cognitive but also social psychology. For example, Groves et al. (1992) argued that the theory of reciprocity and compliance had much to offer in explaining the willingness of respondents to agree to the survey request, and the same was true with the leverage-saliency theory proposed by Groves et al. (2000). Certainly, the recent turn to the use of monetary incentives (itself, an increase in the level of effort and, of course, cost) follows from the notions of reciprocity laid out by Groves et al. On the contrary the effects of incentives are quite complicated. First of all, the promise of an incentive in RDD surveys is less effective than one delivered in person or in the mail (prepaid) (Cantor et al., 2007, Chapter 22 in this volume). Furthermore, while incentives generally will increase response rates, their effects tend to decrease with increased level of effort and the size of the incentive (Singer et al., 1999b; Brick et al., 2005c). A number of factors besides the incentive will affect the decision to participate (Groves and Couper, 1998). For instance,

the sponsor of the survey can have a decided effect on the power of a monetary incentive (Rizzo et al., 2004; Curtin et al., 2005). Incentives also tend to have a greater effect when the response rate is low (Keeter et al., 2004). Finally, there is the use of the incentive just for refusal conversion. While cost effective, there is a question of ethics regarding providing incentives only to initial refusals (Singer et al., 1999a).

One area of effort in the battle to improve respondent cooperation is the use of advance letters, and these letters could be tailored to respondents if auxiliary informations were available (Goldstein and Jennings, 2002). Of course, cognitive theories of information processing could be applied to the development of advance letters too (Dillman, 1978; Dillman et al., 2002). In telephone surveys, the use of advance letters is aided by the increased ability to match telephone numbers to addresses (Brick et al., 2003b). The effects of advance letters in telephone surveys have been mixed (Traugott et al., 1987; Singer et al., 2000). Goldstein and Jennings (2002) do point out, however, that using only one version of an advance letter may limit its utility in terms of increasing response rates. Heerwegh (2005) did find that the personalization of an advance letter did improve cooperation, although Traugott et al. (1987) did not find that to be the case in an RDD survey. Link and Mokdad (2005a) recently cautioned that, while advance letters could improve response rate, this could be at the expense of increased bias.

Leaving voice messages and text messages (on mobile phones) can also be used in a manner similar to advance letters. While Xu et al. (1993) found that leaving messages increased response rates, Baumgartner (1990) saw no differences in the type of message left. In terms of the effectiveness of caller ID for reducing nonresponse, a lot depends on what information is scheduled, and that is controlled by the telephone service provider (Link and Oldendick, 1999). Yuan et al., (2005) found text messaging had no effect on nonresponse in a mobile phone survey, but it is now questionable whether or not text messaging by a survey organization is actually legal.

More operational solutions to overcoming nonresponse, especially with respect to resolving noncontacts, are call scheduling, extending the number of call attempts, and using information gathered by interviewers on previous calls to the number. Most studies of effective call scheduling find the best time for reaching a household is weekday evenings (Weeks et al., 1987). Perhaps, the best information on optimal call scheduling has been gathered by researchers working on the National Immunization Survey, and that data suggest the previous finding to be true (Dennis et al., 1999). Yuan et al. (2005) found that greater success on reaching respondents during the daytime is obtained with mobile phones as compared to landlines. Most studies on the optimal number of call attempts are old. Burke et al. (1981) reported that the optimal number of call attempts appeared to be in the range of five to seven. With the increasing problem of noncontacts, however, there has been growing concern over the residency rates among these numbers known as ring-no-answers or RNAs. Brick and Broene (1997) attempted to determine residency rates by calling telephone business offices, but, as Shapiro et al. (1995) found, business office reports are often error prone. Another method is to call a number many times (Keeter and Miller, 1998), but this is not practical in the

standard survey environment. A new method for estimating residency rates was proposed recently by Brick et al. (2002). This method takes advantage of work done on survival analysis (see also Sangster and Meekins, 2004). Another line of inquiry used to reduce both noncontacts and refusals is taking advantage of information from previous calls to maximize the chance of contact and aid refusal conversion (Groves and Couper, 1998; Purdon et al., 1999; Dixon and Figueroa, 2003; Bates, 2003; Sangster and Meekins, 2004).

When all of these other methods fail, weighting adjustments for nonresponse can be made, and, as reported by Montaquila et al. (2007, Chapter 25 in this volume) and Biemer and Link (2007, Chapter 26 in this volume), this is not such a bad strategy. Unless a telephone survey is confined to local areas, where nonresponse weighting adjustments may use auxiliary information, such as exchange, that is available for all sample units, poststratification involving raking to known population totals for demographic groups is done (Brick and Kalton, 1996; Groves and Couper, 1998; Bethlehem, 2002; Groves et al., 2004). It is assumed that these groups (or weighting cells) are homogeneous with respect to the characteristics of interest (Holt and Smith, 1979). Bethlehem (2002) also discusses the use of a general regression estimator based on available data when nonresponse is present, and Gelman and Carlin (2002) illustrate the use of inverse-probability weights.

Another approach to nonresponse adjustment first proposed by Heckman (1979) shows much promise—response propensity weighting. Using information about the number of callbacks required to complete an interview, Pothoff et al. (1993) developed weights that increase as the respondent proves harder to get. Thus, those responding only after many callbacks will have greater influence on estimates than those responding with few callbacks. Groves and Couper (1998) employed a similar approach in their work on face-to-face household surveys. It is also possible to construct separate propensity models for different subgroups of respondents. The use of response propensity models can provide corrections of nonresponse bias that improve upon traditional weighting methods (Little, 1986; Kalton and Maligalig, 1991; Bethlehem, 2002).

All of the above methods were designed to reduce the potential for nonresponse bias in estimates resulting from the failure to gain cooperation from a significant portion of eligible households in an RDD survey. However, with the growing problem of nonresponse in RDD surveys, survey researchers have also turned their concern to exactly how much bias does arise from nonresponse. Recent studies by Keeter et al., (2000) and Curtin et al. (2000) suggest that level of effort does not have that much effect on estimates and that estimates remain stable across a wide range of response rates. In contrast, it is interesting that even estimates for demographic variables did not vary much across this range (Montaquila et al., 2007, Chapter 25 in this volume). Furthermore, a substantial level of nonresponse (at least 30 percent) remained. The fact is that the relationship between nonresponse and bias is likely to be a complicated one that can differ both according to the characteristic of interest as well as by subpopulation (Tucker et al., 2005).

1.4 LOOKING TO THE FUTURE

1.4.1 Technological Change

Technological change is likely to continue at a rapid pace. More and more households will move to VoIP or wireless broadband in the coming years (Steeh and Piekarski, 2007, Chapter 20 in this volume). Number portability will be an increasing problem; but, eventually, individuals could have unique communication numbers (like social security numbers), which might make telephone sampling easier in the long run (Nathan, 2001). Another possibility is that a single device will receive multiple types of communications (Ranta-aho and Leppinen, 1997; Baker, 1998). Certainly, computers and mobile phones can be used to conduct face-to-face interviews over the telephone soon. Of course, other technologies yet unimagined could have even greater ramifications for telephone surveying.

Although it is likely that the percentage of mobile-only households will reach a plateau, there is no indication that plateau is near-at-hand. It could be as large as 15 percent by the time of the 2008 Presidential election, and, as long as the billing algorithm remains the same, mobile telephone users will be reluctant to pay to do survey interviews. Furthermore, the way households use their mobile phones may change over time. It is unclear what other changes will occur in the telephone system in the coming years, and how telephone sampling might be affected. While the total size of the system should not grow as rapidly in the coming years as it has in the last decade, the allocation of numbers to different types of 100-banks shown in Table 1.2 may continue to change. Assuming change in the telephone system slows down, the listing rates (and, presumably, the unlisted numbers too) within 100-banks should increase. With the increasing densities would come an increase in efficiency for all designs. Of course, this assumes that numbers will be assigned as they have been in the past.

1.4.2 Climate Change

Technological change, of course, is not the central problem. The continued feasibility of telephone surveys depends little on how we locate respondents and what communication devices we use and much more on the respondent's willingness to cooperate. The growing concerns about privacy and confidentiality could cause response rates to deteriorate even further.

Perhaps of even greater concern might be the changing face of society. In many developed countries, household size is shrinking, making it more difficult to contact respondents. Greater mobility is also contributing to a decline in contactability. Family structure is becoming more complex, so our telephone interviewers confront a larger variety of situations once they reach a household. They must be more flexible than in the past in the ways they approach the respondent. Added to this is the growing diversity within western societies, leading to language and cultural barriers that will make surveying, in general, more problematic.

1.4.3 Research

In order to accommodate the changes that are taking place now and that will likely accelerate in the future, telephone survey methodologists must undertake a concerted research program. Given that the telephone mode is likely to be paired with other modes in the future, experimental analyses of mode effects on a large scale will be needed; these analyses should go beyond just measuring differences in mode to consider response error. Without large-scale experiments, we will not be able to uncover the interactions between mode, question, and respondent characteristics.

If future telephone surveys will be conducted on both landline and mobile phones, frequent benchmarks of the telephone service populations should be taken for use in weighting to avoid over representing mobile-only households. These are the households most likely to answer their mobile phones (see Brick et al., 2006.). The only source of these benchmarks at the present comes from the National Health Interview Survey (Blumberg et al., 2007, Chapter 3 in this volume). The growth in the use of mobile phones also presents us with an important conceptual issue. Do we now abandon household surveys in favor of personal surveys? This issue must be confronted at some point.

We do not yet know how the public will respond to all of these new telecommunication choices. Certainly, there will be an interaction between survey respondent characteristics and the use of telecommunication services. We need to understand these interactions in order to develop efficient telephone survey designs. Unfortunately, given the dynamics of an ever-changing technological world, the picture is unlikely to remain constant very long.

Between the technological changes and sociopolitical climate changes, the hope that response rates will improve or even plateau seems unrealistic at this time. First, we need to make greater efforts to estimate the costs and benefits of extra levels of effort to reduce nonresponse (see Groves, 1989; Montaquila et al., 2007, Chapter 25 in this volume). To do this, we will need more detailed information on field activities and their costs. Will incentives offer a realistic solution? How much would they need to be, and what would be an effective means of delivery? Current research on the use of incentives has provided mixed results, and promises to pay do not work as well as cash-in-hand (Singer et al., 1999b; Brick et al., 2005c; Cantor et al., 2007, Chapter 22 in this volume). Secondly, we need a better understanding of the relationship between nonresponse and bias, especially when the nonresponse rates are extremely high. This work will require investigating estimates for a variety of characteristics at subpopulation levels (see Tucker et al., 2005.) Finally, there needs to be a serious discussion about guidelines for evaluating the validity of results from telephone surveys.

1.5 OVERVIEW OF VOLUME

The 25 remaining chapters in this volume are grouped into four sections: Sampling and Estimation, Data Collection, Operations, and Nonresponse. Each section had a team of two editors who selected the papers, with advice from other editors, for the

section and then worked with the authors to prepare presentations for the conference and to prepare final manuscripts for the volume. Several editors read and commented on chapters in sections besides their own. In addition, the editors requested and received editorial assistance from colleagues in the field who read and provided comments on chapters written by the editors themselves. We thank Don Dillman, Marty Frankel, Scott Keeter, and Lars Lyberg for their expert review of chapters presented in this volume.

The Sampling and Estimation section contains seven chapters, beginning with an overview of past and recent developments in sample design and estimation for telephone surveys (Kalsbeek and Agans, Chapter 2). Blumberg and colleagues present in Chapter 3 recent findings on the growing mobile-only population in the United States, and Kuusela, Callegaro, and Vehovar review the international situation in Chapter 4. Flores Cervantes and Kalton review the use of telephone sampling methods in rare population settings in Chapter 5, Tortora, Groves, and Peytcheva discuss the potential uses of multiplicity sampling methods in surveys of mobile-only households in Chapter 6, and Brick and Lepkowski examine the nature of and findings about error properties of mixed mode and multiple frame surveys with telephones as a mode or frame in Chapter 7. Lee and Valliant's, Chapter 8, concludes the section with discussion of propensity weighting methods.

Section 3 on Data Collection contains six chapters on a wide range of data collection issues in telephone surveys. Japec examines the important role interviewers play in telephone surveys in Chapter 9. Conrad, Schober, and Dijkstra in Chapter 10 discuss the kinds of indicators that arise in telephone interviews that respondents are having difficulty in understanding the questions being posed. Harkness and colleagues review in Chapter 11 another communication problem in telephone surveys, translating questions to multiple languages. Mode and format issues for questions are examined in Chapter 12 by Christian, Dillman, and Smyth; visual elements in questionnaire design for presentation to interviewers are discussed by Edwards, Schneider, and Brick in Chapter 13; and mode effects in the Canadian Community Health Survey are reviewed in Chapter 14 by Beland and St. Pierre.

Section 4 also has six chapters, opening with a review of issues in establishing a new computer-assisted telephone interviewing call center by Kelly and colleagues in Chapter 15. Hansen examines features of sample management systems in telephone surveys in Chapter 16, while Tarnai and Moore present a study of interviewer performance and productivity assessment by telephone survey organizations in Chapter 17. Groves and colleagues reexamine the relationship between interviewer voice characteristics and telephone survey response in Chapter 18. The section concludes with Steve and colleagues examination of new measures to monitor and evaluate telephone interviewer performance in Chapter 19 and Steeh and Piekarski's review of the impact of mobile and emerging voice over Internet protocol, or VOIP, technology on telephone survey design in Chapter 20.

The concluding section's six chapters examine a range of issues with respect to nonresponse. Singer and Presser examine the role that attitudes about privacy and confidentiality among the public may have on propensity to respond in Chapter 21. Cantor, O'Hare, and O'Connor conduct a review of findings about the use of

incentives in improving survey response rates in Chapter 22. Holbrook, Krosnick, and Pfent present findings from a survey of research organizations that examines the relationship between procedures used in surveys and response rates in Chapter 23. Battaglia and colleagues examine trends in response rates and where they may be going in the future in Chapter 24. Montaquila and colleagues discuss models for survey nonresponse and provide a theoretical framework for the bias possibly introduced by nonresponse in Chapter 25. The final chapter by Biemer and Link examines an important area of current research, the extent to which early cooperators with telephone surveys differ from those who respond later, and whether models can be developed to adjust early cooperator data to account for late cooperators.

This volume is dedicated to the memory of Joseph Waksberg and Warren Mitofsky, two gentle men who mentored many of us as we became interested in and began working in the field of surveys and telephone survey methodology. Joseph Waksberg's death days before the conference was on the minds of all participants as they talked about topics he would have enjoyed discussing with us. Warren Mitofsky's death as the final versions of chapters were being submitted in fall 2006 left all who work in this area in sorrow at the loss of yet another colleague who had contributed so much, and friend who would be sorely missed. We all do and will miss their continued guidance and contributions to the field, and their enthusiasm and leadership.