UCLA

UCLA Previously Published Works

Title

Longitudinal Research at the Turn of the Century

Permalink

<https://escholarship.org/uc/item/2tk5c1sz>

Journal

Sociological Methods & Research, 36(1)

Authors

Ortiz, Vilma

Ballon, Estela Godinez

Publication Date

2007-08-01

Peer reviewed

Sociological Methods & Research

<http://smr.sagepub.com>

Longitudinal Research at the Turn of the Century: Searching for the Mexican American People

Vilma Ortiz and Estela Godinez Ballon

Sociological Methods Research 2007; 36; 112

DOI: 10.1177/0049124106296008

The online version of this article can be found at:
<http://smr.sagepub.com/cgi/content/abstract/36/1/112>

Published by:

SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Sociological Methods & Research* can be found at:

Email Alerts: <http://smr.sagepub.com/cgi/alerts>

Subscriptions: <http://smr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 19 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
<http://smr.sagepub.com/cgi/content/refs/36/1/112>

Longitudinal Research at the Turn of the Century

Searching for the Mexican American People

Vilma Ortiz

University of California, Los Angeles

Estela Godinez Ballon

California State Polytechnic University, Pomona

Longitudinal studies provide useful and unique data for social science research but are labor intensive and expensive. Key to conducting a successful longitudinal study is the ability to locate respondents in the follow-up effort. We conducted a 35-year longitudinal study among Mexican Americans in Los Angeles, California, and San Antonio, Texas. In this article, the authors describe (a) the individualized approach used in searching and (b) how recent technological developments have provided extensive sources of information that facilitated the searching process. The authors locate 79 percent of their sample. This study differs from previous longitudinal studies in that (a) the population is ethnic, urban, and working class; (b) more than 30 years had passed since the original study; (c) the authors built on a study not originally designed as a longitudinal effort; and (d) there had been no contact with respondents in the intervening period.

Keywords: *longitudinal methodology; longitudinal research; locating respondents; Mexican Americans; Hispanics*

Longitudinal studies provide much useful and unique information in social science research but are labor intensive and expensive. Most important in conducting a longitudinal study is the ability to locate respondents in the follow-up component as well as interview them. Our study is a 35-year longitudinal (as well as an intergenerational) effort with an ethnic, urban, and working-class sample—Mexican Americans in Los Angeles, California and San Antonio, Texas. Not only had our sample not been contacted in more than 30 years, but we built on a prior effort that was not originally designed as a longitudinal study.

In this article, we describe the methods used in searching. Primarily, we used an individualized approach. First, research assistants were trained to collect additional information about the respondents' life from 1965 to 1966. Second, we filled in key events in the respondents' lives from the time of the original survey to the present (primarily through public records such as property records). Third, we searched for the most recent or current information about our respondents or their relatives (frequently locating relatives led to locating respondents).

Second, we describe the sources of information that we used for searching and how these changed with time because of recent technological developments. Initially we searched sources that had been used previously in longitudinal efforts—directories and public/government records—that usually required manual and paper searching. With time, we searched sources of information available on compact disks and subsequently through the Internet.

We show that searching after 35 years is doable—we located 79 percent of respondents in the Mexican American Study Project (MASP). However, this was a labor-intensive effort, as we relied primarily on an individualized approach. We were fortunate, though, to do this project as recent technological developments produced greater access to more sources of information. Last we examine whether the located sample is selective with respect to background characteristics. We show that located respondents differ from those not located primarily because of homeownership.

Substantive Background of Project

The Mexican American People (Grebler, Moore, and Guzman 1970) provided detailed information about Mexican Americans' socioeconomic condition, social mobility, ethnic self-identification and ethnic associations,

Authors' Note: An earlier version of this article was presented at the 2000 annual meetings of the American Sociological Association. We appreciate comments from Edward Telles, Judy Seltzer, and Donald Treiman. We acknowledge grant support from the National Institute of Child Health and Development, Russell Sage Foundation, Ford Foundation, Rockefeller Foundation, Haynes Foundation, California Policy Seminar, University of California Office of the President (UCOP) Committee for Latino Research, University of California Institute for Mexico and the United States (UC MEXUS), University of California Los Angeles (UCLA) Center for American Politics and Public Policy, UCLA Institute for American Cultures, UCLA Chicano Studies Research Center, UCLA Academic Senate Grant Program, UCLA Center for Population Research, Dean of Social Science at UCLA, and Chancellor of UCLA.

family and marriage, religiosity, and political attitudes. This landmark study was based on analyses of a survey of Mexican Americans in Los Angeles and San Antonio conducted in 1965 and 1966. We conducted a follow-up survey of the original respondents in *The Mexican American People* and interviews with their adult children. Our purpose is to examine intra- and intergenerational social mobility, the intergenerational transfer of ethnic identity, and changes in gender and family attitudes among Mexican Americans in this 35-year period. We address critical issues in the sociology of race/ethnicity and assimilation, which are longitudinal and intergenerational in nature and cannot be satisfactorily addressed with the existing cross-sectional data sets that include Mexican Americans (Telles and Ortiz, 2007). Moreover, longitudinal survey samples on the general population have insufficient numbers of Mexican Americans and few questions that relate to our issues of concern.

Prior Longitudinal Studies

Studies that are designed to be longitudinal efforts from the onset have the advantage of incorporating strategies into the design that facilitate retaining respondents in the sample. These include (a) collecting full names, birthdates, home address, and phone numbers and information on backup contacts and (b) maintaining contact with respondents between follow-ups. Yet even when these strategies are used, respondents can be lost to the research team and need to be searched for. Other studies are not originally designed as longitudinal efforts; rather, circumstances provide the opportunity to recontact respondents. These studies usually require considerably more effort at searching. Our study is such an effort. Regardless of their original purpose, all longitudinal studies require searching for respondents to a greater or lesser extent. To do so successfully requires using as many sources of information as possible and doing so in a creative and persistent manner. In this article, we focus our discussion on successful searching strategies.

The key to successful longitudinal efforts is persistence. In a number of earlier studies, done in the 1960s and 1970s, respondents were successfully recontacted. Some of these had response rates of more than 90 percent (see Eckland 1968 for a list of 10 longitudinal studies done in the early 1960s; as well as Clarridge, Sheehy, and Hauser 1978; Crider, Willits, and Bealer 1971-1972; Freedman, Thornton, and Camburn 1980; McAllister, Butler, and Goe 1973; McAllister, Goe, and Butler 1973). To locate respondents in the sample, these studies used phone directories and postal services (such

as mail forwarding), searched public records (such as voter registration, property, motor vehicle), visited original addresses, and contacted others related to the original respondent (family, employers, etc.). One example from the sociological literature—the Wisconsin Longitudinal Study, which followed Wisconsin high school seniors in 1957 for more than 20 years—had a recontact rate of more than 95 percent (Clarridge et al. 1978).

Similarly among studies designed as longitudinal efforts, persistence is the key strategy to successfully keeping respondents in the sample. For instance, the National Longitudinal Survey of Youth, which began in 1979 and continued to survey respondents annually, had a low overall attrition rate. This was accomplished with significant efforts at recontacting respondents even if they had missed intervening years, and 10 percent of the sample was regularly missing in each wave (Falaris and Peters 1998; MaCurdy, Mroz, and Gritz 1998). The High School and Beyond study put considerable effort into recontacting respondents, and 78 percent of respondents were present in all four waves between 1980 and 1986 (Burham and Lee 1998). The Survey of Income and Program Participation reinterviews respondents eight times in a two-year period. Seventy-one percent of the 1984 sample and 73 percent of the 1990 sample were present for all eight waves (Zabel 1998). The British Household Panel Survey interviewed 78 percent of their sample in all four waves between 1991 and 1994 (Laurie, Smith, and Scott 1999).

On the other hand, even studies originally designed as longitudinal studies can lose a significant portion of their sample if they do not invest considerable efforts at keeping respondents. For instance, the National Longitudinal Survey of Young Men and Young Women, begun in 1967, did not attempt to recontact respondents who missed two consecutive years, and consequently they had a lower recontact rate after a number of years (Falaris and Peters 1998). Also, little effort was put into recontacting respondents who missed one follow-up with the Panel Study of Income Dynamics. Consequently, 49 percent of respondents in the 1968 sample was present in 1989 (Fitzpatrick, Gottschalk, and Moffit 1998a), and 52 percent of the children of the original respondents (who were followed as part of the research design) were present in 1989 (Fitzpatrick, Gottschalk, and Moffit 1998b).

More recent longitudinal efforts have had the benefit of using sources of information that have developed with recent technological developments. Ours is such a study. When we began our study and reviewed the literature on longitudinal studies, there were no other reports of studies using computerized or Internet sources of information. Only after we completed our data

collection and were writing up the results of our search process did we find other studies that used some similar sources of information (Hampson et al. 2001; Lyons et al. 2004). Hampson et al.'s (2001) study, in particular, had many parallels to our study—they searched for 2,000 students between 1998 and 2000 who were first interviewed in elementary school in the 1960s on two Hawaiian islands. Similar to our study, their study was not originally designed as a longitudinal study, the follow-up took place many years after the initial study, there was little contact information available, and there had been no contact with respondents in the intervening period. Despite these obstacles, Hampton et al. found 70 percent of the original sample.

In this article, we describe the methods used in searching for an ethnic, urban, and working-class sample that had not been contacted in more than 30 years, and we document our success at locating respondents. Our follow-up effort is also unique in that (a) the sample is ethnic, urban, and working class rather than a cross-section of the general population; (b) the 35-year period between the original study and the follow-up effort is longer than in most longitudinal efforts; (c) there was no contact with respondents in the intervening period; and (d) the study was not originally designed as a longitudinal effort. This meant that although we had names and addresses of the original respondents, there was little other identifying information (as typically collected in longitudinal studies) to help in the search process. We did not have names or addresses for contact persons, names of spouse/children, exact date of birth, or exact spelling of names (and were missing names in a small number of cases). These characteristics would suggest that it would be extremely difficult to locate respondents in the MASP. Nevertheless, we were confident that we could overcome these barriers by combining strategies employed in earlier longitudinal studies with strategies made available by recent technological developments.

1965–1966 MASP Survey and Sample

Prior to pursuing a longitudinal (and intergenerational) study with the 1965–1966 MASP, we ensured that the sampling strategy of the original survey was well designed. The Mexican American population in Los Angeles County, California, and the city of San Antonio, Texas, where this study was conducted, represented 25 percent of all Mexican Americans in the Southwest and 37 percent of all Mexican Americans in urban areas at the time (Grebler et al. 1970:295). These two urban areas were

selected because they represented the two largest concentrations of Mexican Americans at the time and because they were substantially diverse.

The original survey included a sample of 973 households in Los Angeles County and 603 households in San Antonio. The larger sample of Mexican Americans in Los Angeles reflected the larger overall population in that county as compared to San Antonio. To draw the sample, 1960 census information was used to select tracts—the primary sampling units—and blocks—the secondary sampling units. Respondents were randomly selected within census tracts stratified by the proportion Mexican American (i.e., high density, medium density, or low density). Households, the tertiary sampling units, were chosen by (a) screening for self-identified Mexican American households from among all households in high-density tracts or (b) selecting from among Spanish surname households using reverse directories in Los Angeles and city directories in San Antonio in low- and medium-density tracts. A comparison to 1960 census distributions indicates that the Los Angeles and San Antonio samples were representative of the population with respect to education, occupational status, and income (Grebler et al. 1970:302–5). Thus, this sample is well suited for a 30-year follow-up.

Mexican American households were those in which either head or spouse of head identified as Mexican American or Spanish. Persons identifying as “Spanish” were included in the sample because that label was frequently used at the time for the Mexican-origin population given the stigma associated with the label *Mexican*.

Because households were selected on the basis of Spanish surname in low- and medium-density tracts, this may have led to undersampling of the small portion of the Mexican origin in these communities that did not have Spanish surnames. This bias was more likely for women because they are more likely to have acquired a non-Spanish surname from intermarriage. Those with non-Spanish surnames in largely Anglo neighborhoods may be among the most assimilated elements of the population (because they have intermarried or are descendants of an intermarriage and reside in largely Anglo neighborhoods). Despite this limitation, Spanish surname was the best proxy for Mexican American at the time that this sample was drawn (Grebler et al. 1970:Appendix A).

The interview was given alternately to males and females that were heads or spouses of the head. In households not consisting of husband and wife, the head was interviewed. In the final sample, women were over-represented because more single-headed households were headed by women.

Our Longitudinal and Intergenerational Samples

Our research design involves re-interviewing the original respondents, who were ages 18 to 50 in 1965–1966 and are in their early 50s to early 80s today. We omitted respondents who were older than 50 from 1965 to 1966 because of the high probability that they may have died. From Life Tables, we estimated that 77 percent of those older than 50 have died 30 years later, whereas 19 percent of the 18- to 50-year-old respondents would have died by 1995 (based on the 1980 Male and Female Total Population Life Tables for the United States, which approximates that for Mexican Americans [Frisbie 1993]). The 18- to 50-year-olds composed 81 percent of the original sample in Los Angeles, thereby reducing the target sample size in Los Angeles to 792 from the original 973, and 67 percent of the San Antonio sample, thereby reducing the target sample size to 401 from the original 601. Even among the 18- to 50-year-olds in 1965–1966, some would have died by the time of our follow-up survey. In these cases, we searched for family members and conducted an interview with a surviving family member about the respondent (referred to as informant interviews).

Second, we interviewed two of the living children of original respondents that were age zero to 18 in 1965 and would be in their early 30s to early 50s at the time of the follow-up. These children were selected so that we can examine the effect of parental characteristics and attitudes at the time when children were actually young on future child outcomes. We excluded children born after the 1965–1966 survey because parents may have changed after having children. Where there were three or more of these living children, two children were randomly selected. If there were only one or two living children between age zero and 15, they were all selected for the interview.¹

The focus on two cities helps make this study more feasible because of the reduced costs of searching and conducting face-to-face interviews in primarily two places. It is not uncommon for longitudinal studies to be limited to one or two geographic areas. (For instance, the Wisconsin Longitudinal Study was drawn from a sample of Wisconsin high school seniors in 1957 [Clarridge et al. 1978].) The Mexican American Study sample is not representative of the entire Mexican American population today because of significant immigration since 1965. Nevertheless, it is representative of long-term residents and U.S.-born Mexican Americans in two major urban areas. Moreover, the intragenerational and intergenerational

nature of the study provides unique data for addressing important sociological questions.

Search Procedures

Throughout the process, we used an individualized approach to searching for respondents. Research assistants were trained to handle the search process on a case by case basis. We trained research assistants to take note of the information in the original survey and think creatively about the person's life—"What is this person likely to have done given their circumstances in 1965–1966, or where are they likely to have moved (e.g., moved back to their state of birth)?" A key initial strategy was to collect additional information about the respondents' life from 1965 to 1966. We did this by examining city directories, phone directories, and reverse directories from the 1960s and 1970s. Second, we attempted to fill in what had happened to the respondents since that time with respect to marriage, divorce, death, and/or selling or buying property (events that are documented in public records). Third, we searched for the most recent or current information possible about our respondents or their relatives. We frequently located respondents by first finding relatives who provided us with contact information for the respondent.

At every point in the search process, we attempted to link individuals listed in the various databases to the person who lived at 1965–1966 address. Sometimes we could do this by using databases that maintain address histories (e.g., Trans Union). But mostly, it was by speaking directly with the persons we identified that we were able to verify whether this was our respondent. Occasionally we identified individuals that appeared to be our respondent because they closely matched them on numerous characteristics (such as gender, age, name, and marital status), but the individual was not the respondent (i.e., never lived at the 1965–1966 address). In cases where the original respondent had died, we searched for a family member so that we could conduct an informant interview.

We held team meetings every one to two weeks during which we reviewed cases. In the process, we discussed strategies for finding cases. Occasionally, we reassigned cases to other research assistants to provide a fresh perspective on searching. The project manager (second author on this article) reviewed cases and provided feedback to research assistants about possible avenues to pursue.

Another factor that facilitated the search process was the very dedicated group of research assistants that we employed. The research assistants we

hired were Mexican origin or Hispanic. Thus, they could speak Spanish, were familiar with Spanish names, were familiar with the communities where the original study had taken place, and could approach respondents in culturally sensitive ways. The fact that the entire team was Hispanic helped us deal with misspellings of names, a common problem—in longitudinal studies—and we regularly developed strategies to deal with this in team meetings. As part of our broader discussions, we explained the relevance of the original study and its important contributions to our historical and contemporary understanding of the Mexican American community. Our research assistants understood the uniqueness of the follow-up study and that it could make a significant contribution if we located most of the respondents. Although we cannot quantify the effect of employing research assistants who felt a personal investment in the project, we believe that it was significant.

Most of the searching for the Los Angeles sample was conducted prior to the interview phase of the project (because we did not have sufficient funding at the time). We found that when the interviewing phase began, we needed to continue searching because a small number of respondents, who we had previously located, had moved.

Sources of Information for Searching by Stage of Search Process

First Stage in Los Angeles

One of the very first strategies that we used was to visit original addresses. We located some respondents, almost all who were homeowners in 1965, at the original addresses. However, most respondents had moved, so we quickly shifted to other sources of information. Although we added new sources of information on a continuous basis, in retrospect, we can group our search strategy into three major stages that used different types of information. In Table 1, we list type of information used and the name of the source, sorted by the year we began using the source.

In the first stage of searching, we relied heavily on public sources of information, which, for the most part, required travel to government offices. The most important sources of information during this period were voter registration, property, death, and marriage records, all available through the county offices (located in Norwalk, California, approximately 30 miles from the University of California, Los Angeles [UCLA]). These records served to provide more recent addresses (particularly voter

Table 1
Stages and Sources of Information in the Search Process in Los Angeles and San Antonio

Year	Type of Source	Name/Description of Source	How Accessed
First stage in Los Angeles			
1994	Phone directory	Current phone books	Phone company
1994	Phone directory	Current directory assistance	Phone company
1994	Reverse directory	Current reverse directories	Owned by UCLA's Survey Research Center
1994	City directory	LA 1960s–1970s city directories	Searched at LA Public Library
1994	Reverse directory	LA 1960s–1970s reverse directories	Searched at LA Public Library
1994	Voter registration	LA County Voter Registration	Searched at county offices
1994	Property records	LA County property	Searched at county offices
1994	Death records	LA County death	Searched at county offices
1994	Marriage records	LA County marriage	Searched at county offices
1994	Driver's license	California DMV	Request by mail from DMV
1994	Property records	LA County property	Purchase CD from TRW
1994	National Credit Headers	Equifax National Credit Bureau	Request by mail from Equifax
1994	National Credit Headers	TransUnion National Credit Bureau	Modem connection to TransUnion
Second stage in Los Angeles			
1995	Death records	Nationwide Death Index	Purchase CD from CSRA
1995	Phone directory	National select phone	Purchase CD from ProCD
1996	Voter registration	Texas/Colorado voters (1995)	Purchase CD from CSRA
1996	Driver's license	Texas driver's licenses (1996)	Purchase CD from CSRA
1997	Multiple sources	CA date of birth reference index	Purchase CD from CSRA
1997	Voter registration	CA statewide voters	Purchase CD from Deepdata
1997	Driver's license	Online CA DMV	Modem connection to DMV
First stage in San Antonio			
1998	City directory	SA 1960s–1970s city directories	Searched at SA public library
1998	Reverse directory	SA 1960s–1970s reverse directories	Searched at SA public library

(continued)

Table 1 (continued)

Year	Type of Source	Name/Description of Source	How Accessed
1998	Phone directory	SA 1960s–1970s telephone directories	Searched at SA public library
1998	Voter registration	SA voter registration	Searched at county offices
1998	Property records	SA property	Searched at county offices
Last stage in Los Angeles and San Antonio			
1997	National Credit Headers		
1997	Phone directory	Missing links	Internet: Purchase search CDB Infotek
1997	Phone directory	Switchboard	Internet: Free search
1997	Phone directory	Anywho	Internet: Free search
1997	Phone directory	Infospace	Internet: Free search
1997	Phone directory	411	Internet: Free search
1997	Phone directory	Bigfoot	Internet: Free search
1998	Multiple sources	Wingate National Peoplefinder	Internet: Purchase search from Merlin
1998	Multiple sources	TraceWizard national residential locator	Internet: Purchase search from Merlin
1998	Death records	Social security death index	Internet: Free search
1998	Property records	Bexar appraisal district property records	Internet: Free search
1998	Property records	CA statewide property (1997)	Purchase CD from Merlin
1998	Phone directory	Texas Peoplefinder (1997/1998)	Purchase CD from Merlin
1998	Property records	Texas property (July 1997)	Purchase CD from Merlin
1998	Marriage records	Texas marriage (1997)	Purchase CD from Info Depot
1998	Divorce records	Texas divorce (1996)	Purchase CD from Info Depot
2000	National Credit Headers	National people locator	Internet: Purchase search from Merlin
2000	Phone directory and other	Residential locator	Internet: Purchase search from Merlin
2000	Property records	National real property	Internet: Purchase search from Merlin
2000	Property records	National property deed transfers	Internet: Purchase search from Merlin
2000	Death records	Social security death index	Internet: Purchase search from Merlin
2000	Legal records	National bankruptcies, judgments & liens	Internet: Purchase search from Merlin
2000	Business information	Fictitious business names	Internet: Purchase search from Merlin

Note: UCLA = University of California, Los Angeles; LA = Los Angeles; SA = San Antonio; DMV = Department of Motor Vehicles; TRW = TRW Inc.; CD = compact disc; CSRA = Cambridge Statistical Research Associates, Inc.; ProCD = ProCD, Inc.; CA = California.

registration and property) and information on family members that assist in locating the respondent (namely death and marriage). Also, voter registration, death, and marriage records all provide dates of birth as well as dates of death or marriage. To access these records, research assistants searched manually through original record books (for instance, pre-1950 marriage records or pre-1965 property deeds were not computerized), scanned microfiche, and/or used the agencies' public-use computers. We also accessed computerized data sources of death and marriage records at the local Family History Center of the Church of Jesus Christ of Latter Day Saints (located in West Los Angeles, close to UCLA). These computerized databases were useful for knowing that a death or marriage had occurred but did not provide all the information available on the actual death or marriage certificate, so we followed up by searching for the actual certificates at the county offices.² To a lesser extent, we used divorce records available through the court system (these were hard to locate but when available provided a wealth of information).

Phone directories and directory assistance were important sources of information, although it was frequently difficult to use them. The Los Angeles area has many phone books (one for each incorporated city or area and several for the city of Los Angeles), and it was nearly impossible to acquire a complete set for our project offices. We had difficulties with local directory assistance when we did not know the respondent's current city of residence (these could be relatively small entities considering the multicity nature of the greater Los Angeles area). At the Los Angeles Public Library, we accessed historical sources from the 1960s and 1970s—phone directories, city directories from the 1960s for parts of Los Angeles, and reverse directories starting in 1972. These sources of information served to add more information than that collected in the 1965 survey, such as names of spouse and children.

Searching through county records and address/phone directories would frequently produce a list of possible individuals who might be our respondents. Of course, more common names produced longer lists of possible individuals. In this situation, we would call (when we had phone numbers) or mail letters to the list of possible individuals. When we had no response, we mailed at least two follow-up letters and subsequently visited the address. We used calling and sending letters throughout the search process.

Occasionally we were able to use a targeted search process for a group of respondents. For instance, early in the search process, we searched for

16 respondents who resided in a housing project. None of the respondents was still at the original address. We contacted the management office of the housing project, and the resident manager was quite cooperative (he was actually familiar with the original Grebler project and book). We were able to access the forms that residents completed when they moved away from the projects. Although many of the forms did not list a forwarding address, many contained information (such as a relative's name or social security number) that helped us locate these respondents. This was useful as most of these respondents were nonmarried women, an especially difficult group to search for because of remarriage and name changes. We eventually located 15 of the 16 respondents.

Throughout the search process (including during the second and last stages described later), we continued doing field visits when deemed necessary. For respondents we knew were no longer at the original address but for whom we could find no leads, we visited the 1965–1966 address to find long-term neighbors who might provide additional information about the respondent (names of spouse or children) or might know where the respondent had moved to. In cases where we suspected that we had located the respondent but were unsuccessful at establishing contact (no phone number, no response to multiple letters, no response to phone messages), we would visit the home.

Some sources of information did not work out to help us locate respondents. Birth records were unavailable (would have been useful for identifying children's names) and immigration records were too difficult to access and connect to our respondents.

From the very beginning of the search process, we had access to some computerized databases. For example, we accessed national credit headers in TransUnion via a modem connection as early as 1994 (part of the first stage). To search through TransUnion (which stores current and two prior addresses) required both a name and address. Frequently, our search did not yield positive results because the 1965–1966 address was too old to match one of the addresses in TransUnion. TransUnion was more helpful once we were further into the search and had obtained more recent addresses from other sources. Also, TransUnion required a last name and at least a first initial. Because occasionally first names were missing or incorrectly recorded for our respondents, this limited the flexibility of our search. Our research assistants learned to handle this restriction by repeatedly searching each letter of the alphabet as the first initial (A* . . . Z*) paired with the last name and address. This was a creative but time-consuming way of searching a database.

Second Stage in Los Angeles

As is evident from the previous section, using our initial sources of information was labor intensive, so we continued to identify alternative sources of information for searching. Our work coincided with the development of compact disks, which have made available large amounts of data on small mediums. These sources, listed on Table 1, mainly cover the period between 1995 and 1997 in our search for the Los Angeles sample. Eventually we were able to purchase compact disks with property records in California, voter registration records in California, a national death index, and a date of birth reference. These data sources were available through commercial companies, such as TRW (compiled property records). National phone directories also became available on compact disks. We used databases compiled from credit headers—these are the name/address/phone information taken from applications for credit and other financial sources (without providing any of the credit/financial information). We accessed these via modem from our project office (available from Trans-Union). Also we were able to access Department of Motor Vehicles (DMV) records by special permission (DMV records are not public in California as an antistalking measure). These were available via special computer connection from the offices of the Survey Research Center (SRC) at UCLA. Having access to these sources of information shifted most of the searching activities to our project offices.

In the second stage, one of the most important changes was the increased flexibility available in databases. Missing Links (a database of credit headers available through the Internet in 1997) increased our flexibility by allowing searching with just a name. Although certainly more flexible, searching with just a name frequently produced too many hits (especially for respondents with common names).

First Stage in San Antonio

We searched for the San Antonio sample after most of the searching for the Los Angeles sample was completed.³ Therefore we were able to draw on our experiences in Los Angeles to develop procedures for searching in San Antonio. The early search process in San Antonio involved many of the same kinds of sources used in Los Angeles, such as phone directors, reverse directories, and historical city directories. This information was available at the San Antonio Public Library and was easier to access than comparable information for Los Angeles (fewer and more complete volumes). However,

because searching began later in San Antonio, we were able to use many of the sources used only in the last stage in Los Angeles from the beginning of the search process in San Antonio. Moreover, public records were generally easier to access in San Antonio than in Los Angeles. For instance, Texas motor vehicle records were public, and Bexar County property records were available through the Internet (described in more detail later). One exception was death records. We were still able to know if someone had died from National Death Index, but obtaining a copy of the death certificate was considerably more difficult in Texas than in California. This is unfortunate because death certificates provide name and contact information of a family member. Although we did most of the searching for San Antonio from our UCLA project offices, our research assistants traveled to San Antonio several times. During visits, the research assistants would visit homes of respondents or family members to establish contact or would conduct research that could not be done long distance, such as visit the public library to search historical directories.

Last Stage in Los Angeles and San Antonio

The third stage of our search process coincides with the availability of information through the Internet. There is considerable overlap between the second stage and the last stage in that much of the data sources we began to use in the second stage we continued to use when they became available through the Internet.

The availability of information through the Internet had many benefits. First, the Internet made it easier to identify and purchase sources of information on compact disks, such as California Property, Texas People Finder, and Texas Property. Second, national phone directories are available (frequently free of charge) on the Internet. Third, the Internet introduced sources of information that allowed for more powerful and flexible searching strategies (see discussion below). Last, the Internet introduced data sets that contained large amounts of data merged from several sources. For example, although earlier we used Missing Links, which provided contact information from credit headers, later on we used Merlin Flat Rate, which combined information from directory assistance, property information, and directories of business owners. Also in the last stage, the National People Locator, a database of credit headers available through the Internet, allowed us to search using various combinations of items. For instance, we searched on last name and city, first name and birth month and year (especially useful for women who may have married), address

and city, and street name and zip code. Although we were fortunate to use credit header information throughout the life of the project, the ease with which we searched these databases increased significantly over time, making our searches more efficient and successful.

Many of the same sources of information that we used in Los Angeles were available for San Antonio, and fortunately, many were easier to access and more flexible to use by the time we searched for the San Antonio sample. For instance, we gained access to California DMV records through a special application process, which was time consuming; in contrast, the Texas records are public and readily available. City directories from the 1960s, which are available for only some communities in Los Angeles, are available for all of San Antonio in one volume. Moreover, throughout the search process in San Antonio, we benefited from using the Internet and large databases that were only available to us in the last stage of our search in Los Angeles. For instance, we accessed property records for Bexar County (where San Antonio is located) through the Internet.

Searching for respondents is labor intensive and therefore expensive. Although we did not have a precise estimate of the costs of searching, we estimate that it was approximately \$300 per respondent. This is as much as the cost of interviewing, thus doubling the costs of conducting a survey. Fortunately, we had a grant from the National Institute of Child Health and Development that allowed us to put a lot of effort into the search process.

Locating by Source

Table 2 presents the sources (grouped into broad categories) that were used to locate respondents. This analysis is limited to the Los Angeles sample, where we used the greatest number of sources to search for respondents and to cases that were successfully located. The first column presents the percentage that used the type of source at any point in the search process for cases that were successfully located. Because multiple sources could be used to locate a respondent, the percentages in the first column add up to more than 100 percent. We see that voter registration and property records were the most useful in locating respondents. Second to that were personal contacts and credit headers. What made these sources useful is that they provided several pieces of information on which we could match to our respondents and provided additional information that we could use for additional searching. For example, property records frequently provided the name of the spouse or a forwarding address. Voter

Table 2
Sources Used to Locate Respondents in Los Angeles^a

Source	Source Is Used in Search to Locate Respondent (%)	Source Provides First Lead in Locating Respondent (%)
Personal contacts ^b	28	8
Directories	14	5
Historical directories	9	8
Voter registration	34	21
Property	33	22
Marriage	14	12
Death	10	6
Driver's license	11	5
Other public ^c	5	3
Credit header ^d	23	4
Internet sources	7	5
Total		100

a. Thirty-nine percent of the sample was located with one source, 36 percent with two sources, 15 percent with three sources, and 10 percent with four or more sources.

b. Personal contacts include visiting original addresses and contacting neighbors, landlords, family coworkers, church, and school.

c. Other public includes birth, divorce, naturalization, and housing records.

d. Credit headers from TransUnion via modem.

registration records provided place of birth. The second column presents the source that provided the first lead in cases where we successfully located the respondent. Again we note that voter registration and property records were the most useful as a first source. But there are some interesting differences between the source that helped in general and those that provided the first lead. Personal contacts and credit headers were not as useful in providing the first lead as they were in providing leads in general. In contrast, marriage records frequently provided a first lead but were not as useful in general.

Location Rate

As can be seen from Table 3, our overall location rate is 79 percent. This is impressive considering the ethnic, urban, and working-class composition of our sample and the length of time elapsed and lack of contact since the first survey. As we anticipated, the search process in San Antonio

Table 3
Percentage Located for Combined Sample and by City

	Percentage Located	Size of Sample Used in Search
Combined	79	1,195
Los Angeles	78	794
San Antonio	82	401

Note: Chi-square = 2.52, not significant.

was more successful than the process in Los Angeles—82 percent located in San Antonio and 78 percent in Los Angeles—although this difference is not statistically significant (see Table 3). One reason is that San Antonio was a smaller city to begin with and has grown considerably less than Los Angeles in the intervening period. Second, San Antonio has had less immigration, so we were searching for our respondents in a smaller Latino population. Third, we found that the San Antonio sample is more stable—more respondents were at the original address or in the same community. Last, a higher percentage of respondents were homeowners in San Antonio than in Los Angeles (64 percent vs. 43 percent, respectively), and as we show later, homeowners are much easier to locate than renters.

Availability of Names and Searching

Some of the information on names/addresses in the 1965–1966 questionnaires was incomplete. As can be seen from Table 4, 83 respondents were missing first names and 23 respondents were missing both first and last names.⁴ We nevertheless attempted to search for these respondents. We used the historical reverse directories to find names, because reverse directories are sorted by address (not names). With names from reverse directories, we could find more information about the respondents in city directories (sorted by name but including more information, such as spouse's name). With this additional information, we were fairly successful in locating respondents for whom we had only last names (67 percent), although we located more respondents (81 percent) among those for whom we had both first and last names. And we even located respondents for whom we were missing both first and last names (43 percent); obviously it was much harder to locate respondents who were missing both names than those missing first names.

Table 4
Percentage Located by Available Name and Common Name

	Percentage Located	Number
Available name		
Complete name	81	1,089
Only last name	67	83
No name	43	23
Chi-square = 26.03, $p < .001$		
Common name		
Uncommon (found 1 to 9 times)	84	427
Common (found 10 or more times)	73	378
Not found in directory	80	390
Chi-square = 14.54, $p < .001$		

The other factor that affected our ability to locate respondents was whether respondents had a common name. To show this empirically, we compared respondents with more common names to those with less common names. We searched each respondent's name in a national name/address/phone directory (SelectPhone on compact disk) and coded the number of times that the respondent's name was listed. Respondents whose name appeared 10 or more times were more difficult to locate (73 percent were located) as compared to those whose name appeared fewer than 10 times (83 percent were located). Some respondents were not listed in the directory at all, and the rate at which we located them falls between those with common names and those with uncommon names (80 percent). This latter group may be a combination of various subgroups—one subgroup may be those with very uncommon names that do not appear in this directory, whereas another subgroup may be those with more common names that were incorrectly recorded.

Factors Associated With Being Located

Previous research has frequently shown the significant differences between located and nonlocated respondents on important characteristics, and we must consider whether the same is true for our study. First, we examined key 1965–1966 characteristics, comparing located respondents to those not located in a bivariate fashion. Subsequently, we did a multivariate analysis where we predicted location controlling for the interrelationships

Table 5
Odds of Locating Respondents by City

1965–1966 Characteristics	Los Angeles		San Antonio	
	Unadjusted	Adjusted	Unadjusted	Adjusted
Demographic characteristics				
Age	1.01	0.98	1.01	1.00
Male	1.32	1.13	1.18	0.90
Married	1.94***	1.19	3.89***	2.74*
Immigration/assimilation characteristics				
Born in United States	2.12***	1.44	1.04	0.89
English interview	2.47***	1.51	1.57	1.11
Socioeconomic characteristics				
Education	1.10***	1.03	1.11**	1.02
Family income (thousands)	1.13***	1.03	1.39***	1.20*
Homeowner	4.14***	3.68***	3.73*	2.63***
Sample size	794		401	

* $p < .05$. ** $p < .01$. *** $p < .001$.

among these key characteristics. Table 5 represents results from the logistic analysis for the Los Angeles and San Antonio samples (the appendix presents means for the total, located, and interviewed samples). The first column presents the odds ratio unadjusted for any other factors, and the second column presents odds ratios controlling for the other factors listed on Table 5.

The unadjusted differences in the Los Angeles sample indicate those located were more likely to be married, born in the United States, interviewed in English, better educated, of higher income, and homeowners. The largest difference is with respect to homeownership. Because these background characteristics are highly correlated, a multivariate analysis allows us to see which factors most strongly predict location. We see that in the Los Angeles sample, the effects of marriage, place of birth, language, education, and income are reduced and are no longer significant in the multivariate analysis. The one factor that shows a large and significant effect in both the unadjusted and adjusted models is homeownership. Respondents who were homeowners in 1965 were almost four times more likely to be located 35 years later than renters.

As also can be seen on Table 5 for the San Antonio sample, there are significant differences by marital status, education, income, and homeownership. Moreover, we observe that most of the unadjusted differences

are reduced in the multivariate model. The married difference is reduced but is still statistically significant. The education difference is reduced and no longer statistically significant. The income difference is reduced but still statistically significant. Homeownership continues to be a large and statistically significant effect in the multivariate analysis. The homeowner effect is smaller in the San Antonio than in the Los Angeles sample but still sizeable.

One reason that homeownership has such a strong effect is that it represents stability—homeowners move less than renters. Also, homeowners continue to be homeowners even when they sell. Third, transfers in property leave a paper trail, making it easier to locate respondents in public records. Our findings are consistent with previous studies on differences between located and nonlocated respondents—showing that homeownership has a strong effect on being located (Laurie et al. 1999; Lin, Schaeffer, and Seltzer 1999). Overall, we find it is easier to locate respondents who have deeper roots, more family connections, and more stable lives.

Interviewing

The Los Angeles sample interviews were conducted by the Survey Research Center of UCLA's Institute for Social Science Research and the San Antonio sample by a Texas-based firm, NuStats Research and Consulting, in Austin. Interviews were primarily conducted in respondents' homes. Contact information was forwarded from the UCLA project staff to SRC and NuStats field staff. First, the original respondent was reinterviewed. Subsequent to that interview, the randomly selected children of the respondent were contacted for an interview. Respondents were paid an incentive of \$20. In cases where respondents had moved away from Los Angeles or San Antonio, telephone interviews were conducted (22 percent of the interviews with the Los Angeles sample were done by phone). For respondents who had died or were incapacitated, we did an informant interview with a family member, usually the surviving spouse but also adult children or another close family member (31 percent of the reinterviews were done by an informant).

With the Los Angeles sample, we interviewed 71 percent of those located, resulting in an overall response rate of 55 percent (71 percent multiplied by 78 percent located). With the San Antonio sample, we interviewed 76 percent of those located, resulting in a response rate of 62 percent. Although we interviewed a smaller percentage of respondents than

we had hoped, our response rate compares favorably to the one other study we identified that was similar to our study. Hampson et al. (2001) report interviewing 62 percent of those located, thus making their response rate 47 percent (62 percent multiplied by 75 percent located, which we cited earlier).

As a means of understanding the impact of this nonresponse on the second wave of interviews, we compared the 1965–1966 demographic characteristics of those we interviewed to those we did not among located respondents. We compared respondents with respect to age, gender, marital status, place of birth, language of interview, education, income, and homeownership. We found that the 1965–1966 characteristics of interviewed respondents did not differ from those not interviewed among located respondents.⁵ This is consistent with Lin et al.'s (1999) findings that there are few significant differences between those interviewed and those not interviewed (although they reported large differences between those located and those not located, as we found in our study).

The main reason that located respondents were not interviewed was refusal to be interviewed.⁶ For the most part, we found that respondents were receptive, as well as surprised, when we located them. Most did not remember being interviewed more than 35 years earlier, not surprising considering that this was a one-time event that occurred so long ago. Many, however, were curious that we located them after so many years, especially as they did not recall participating in the original study. Some were suspicious or concerned that we were able to do so. And it is these respondents that were more likely to refuse to be interviewed. We suspect that our study faced similar factors that affect other survey efforts today: that the general public has been overwhelmed by telemarketers, market research, polling, and a variety of commercial interests. Additionally, because our original respondents are now primarily elderly, we thought age might be a factor.⁷ However, we did not find a significant effect of age on being interviewed in regression analysis, although bivariate tables show that older respondents were slightly more likely to not be interviewed.

Our staff employed extraordinary methods to convert the refusals into interviews. Project staff that searched for respondents made sure to note any resistance to our calling in our initial contact with them, and this information was passed to the interviewing staff at SRC and NuStats. Refusing or reluctant respondent cases were assigned to skilled supervisors/interviewers trained specifically for refusal conversion. Interviewers were trained to be considerate of respondents' concerns (whether it be

issues of time, family illness, suspicion). Special letters were sent to respondents to assure them of our legitimacy and to attempt to gain their confidence and trust, addressing their individual concerns when voiced. If after numerous attempts we had not secured an interview, we let cases sit and then followed-up at a later time. In some instances, refusals were converted when respondents would consent to a telephone interview but refused in-person contact. The telephone version was originally developed for use with respondents residing at a distance too far from Los Angeles or San Antonio to be visited in person.

Sample Selection Bias

Based on the differences between located respondents and not located ones, we must consider ways of dealing with sample selection bias in our future substantive analysis. Because we have information about the full sample in 1965, we can address some bias in estimating 2000 outcomes by controlling for 1965 variables (Winship and Mare 1992). A common method for controlling sample selection bias is to regress a binary variable for whether we secured an interview in the follow-up on a set of potential explanatory variables from the original study. However, this and other corrections to bias may not be reliable, as statisticians have been unable to specify a good model to correct for selection. We recognize that there is little agreement about whether to correct and, when correcting, which procedure is most appropriate. Adjustment procedures, such as Heckman's (1979) estimator, may also worsen data as often as they correct it (Stolzenberg and Relles 1997). By comparing adjusted and unadjusted models in our future analysis, we may find that estimates of slopes based on these methods are not different from those based on unadjusted regressions, suggesting that sample selection bias is not a problem.

Summary and Conclusions

We searched for a sample of Mexican-origin respondents who resided in Los Angeles and San Antonio in 1965 through 1966 after 35 years. We located 79 percent of respondents across the two cities by employing an individualized approach in searching for respondents. We used sources of information for locating respondents that had been used in previous longitudinal studies (mostly public records from government offices) and found these useful but insufficient. We benefited from recent technological

developments that provide large data sources of individuals on compact disk and through the Internet.

Appendix

1965–1966 Characteristics for Total, Located, and Interviewed Samples in Los Angeles and San Antonio

1965–1966 Characteristics	Los Angeles			San Antonio		
	Total	Located	Interviewed	Total	Located	Interviewed
Demographic characteristics						
Age (in years)	34.8	34.9	34.5	35.8	35.9	35.6
Male (%)	43	44	44	42	43	42
Married (%)	80	82	83	86	90	90
Immigration/assimilation characteristics						
Born in United States (%)	73	77	76	90	90	89
English interview (%)	58	63	61	68	70	70
Socioeconomic characteristics						
Education (in years)	9.1	9.4	9.4	8.0	8.2	8.2
Family income	\$7,055	\$7,351	\$7,210	\$4,550	\$4,813	\$4,797
Homeowner (%)	43	50	48	65	71	70
Sample size	794	616	436	401	327	250

Notes

1. Because we sampled from living children, the issue of children dying was mostly irrelevant. Although, we did have one case in which the child died between the time we did the interview with the original respondent and did the child selection, and we contacted the child to do the interview. We did not select another child for interviewing.

2. The availability of death certificates helped considerably in being able to locate the family of original respondents who had died. The death certificate lists the name and contact information of a family member (as the person who provided information about the deceased).

3. There were several reasons why we conducted the search for the Los Angeles and San Antonio samples at separate points in time: We had funding for the Los Angeles portion of the study first; given our location at the University of California, Los Angeles (UCLA), it was easier to search for the Los Angeles sample; and other than location, we expected the San Antonio search to be easier and wanted to conduct the more difficult task first.

4. In the San Antonio sample, only one respondent was missing both last name and first name.

5. Analysis is available from authors on request.

6. A few of the Los Angeles respondents were lost between the time they were located and the time we went back to interview them. This happened with the Los Angeles sample because of the lag time between searching and interviewing in Los Angeles (in San Antonio, we interviewed quickly on locating respondents). We searched for respondents again in these cases, but some were never located again. However, only a small number of respondents were lost.

7. There has been an ongoing public service campaign aimed at Latino elderly to be careful of solicitors, telemarketers, and any stranger coming to the door or calling.

References

- Burham, David and Valerie Lee. 1998. "Effects of Monotone and Nonmonotone Attrition on Parameter Estimates in Regression Models with Educational Data: Demographic Effects on Achievement, Aspirations and Attitudes." *Journal of Human Resources* 33 (2): 555-74.
- Clarridge, Brian, Linda Sheehy, and Taissa Hauser. 1978. "Tracing Members of a Panel: A 17-Year Follow-Up." *Sociological Methodology* 9:185-203.
- Crider, Donald, Fern Willits, and Robert Bealer. 1971-1972. "Panel Studies: Some Practical Problems." *Sociological Methods and Research* 2 (1): 3-19.
- Eckland, Bruce. 1968. "Retrieving Mobile Cases in Longitudinal Surveys." *Public Opinion Quarterly* 32:51-64.
- Falaris, Evangelos and Elizabeth Peters. 1998. "Survey Attrition and Schooling Choices." *Journal of Human Resources* 33 (2): 531-54.
- Fitzpatrick, John, Peter Gottschalk, and Robert Moffit. 1998a. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33 (2): 251-99.
- Fitzpatrick, John, Peter Gottschalk, and Robert Moffit. 1998b. "The Impact of Attrition in the Panel Study of Income Dynamics on Intergenerational Analysis." *Journal of Human Resources* 33 (2): 300-44.
- Freedman, Deborah, Arland Thornton, and Donald Camburn. 1980. "Maintaining Response Rates in Longitudinal Studies." *Sociological Methods & Research* 9 (1): 87-98.
- Frisbie, W. Parker. 1993. "Preliminary Mortality Estimates for Mexican Americans in Bexar County: 1910-1980." Unpublished manuscript. Department of Sociology, University of Texas at Austin.
- Grebler, Leo, Joan Moore, and Ralph Guzman. 1970. *The Mexican American People*. New York: Free Press.
- Hampson, Sarah, Joan Dubanoski, Winter Hamada, Anthony Marsella, Jennifer Matsukawa, Edward Suarez, and Lewis Goldberg. 2001. "Who Are They Now? Locating Former Elementary-School Students after Nearly 40 Years for a Longitudinal Study of Personality and Health." *Journal of Research in Personality* 35:375-87.
- Heckman, John J. 1979. "Sample Selection Bias as Specification Error." *Econometrica* 47:153-61.
- Laurie, Heather, Rachel Smith, and Lynne Scott. 1999. "Strategies for Reducing Non-response in a Longitudinal Panel Survey." *Journal of Official Statistics* 15 (2): 269-82.
- Lin, I-Fen, Nora Schaeffer, and Judith Seltzer. 1999. "Causes and Effects of Nonparticipation in a Child Support Survey." *Journal of Official Statistics* 15 (2): 143-66.

- Lyons, Karen, Julie Carter, Emily Carter, Kirsten Rush, Barbara Stewart, and Patricia Archbold. 2004. "Locating and Retaining Research Participants for Follow-Up Studies." *Research in Nursing and Health* 27:63-8.
- MaCurdy, Thomas, Thomas Mroz, and Mark Gritz. 1998. "An Evaluation of the National Longitudinal Survey of Youth." *Journal of Human Resources* 33 (2): 345-436.
- McAllister, Ronald, Edgard Butler, and Steven Goe. 1973. "Evolution of a Strategy for the Retrieval of Cases in Longitudinal Survey Research." *Sociology and Social Research* 58 (October): 37-47.
- McAllister, Ronald, Steven Goe, and Edgard Butler. 1973. "Tracking Respondents in Longitudinal Surveys: Some Preliminary Considerations." *Public Opinion Quarterly* 37:413-6.
- Stolzenberg, Ross and D. Relles. 1997. "Tools for Intuition about Sample Selection Bias and Its Correction." *American Sociological Review* 62 (3): 494-507.
- Telles, Edward and Vilma Ortiz. 2007. *Racialized Ethnicity: Mexican Americans and the Persistence of Ethnic Boundaries*. Unpublished manuscript.
- Winship, Christopher and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18:327-50.
- Zabel, Jeffrey. 1998. "An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior." *Journal of Human Resources* 33 (2): 479-506.

Vilma Ortiz is associate professor of Sociology at University of California, Los Angeles. Her research addresses theoretical issues on racial/ethnic stratification and social inequality. For more than 20 years, she has studied the socioeconomic experiences of Latinos in the United States, focusing both on specific Latino groups as well as comparative studies with other racial/ethnic groups.

Estela Godinez Ballon is assistant professor in the Liberal Studies Department at California State Polytechnic University, Pomona. Her main fields of interest include sociology of education, tracking and ability grouping, language in schools, Chicana/o Studies, and teacher education.