

READ.ME

**LPS PUBLIC USE FILE:
THE 1989 LEGALIZED POPULATION SURVEY (LPS1) AND
THE 1992 LEGALIZED POPULATION FOLLOW-UP SURVEY (LPS2)**

The LPS public use data file contains information about a sample of 6,193 illegal aliens who were living in the United States in 1986/87 when they sought legal permanent residence through the Immigration Reform and Control Act of 1986. Survey data was collected from the entire group in 1989, and again (for 4,012 of these respondents) in 1992.

There are three data files in LPS.ALL:

- README.pdf Technical documentation
- CBK.pdf: The codebook for the LPS matched file
- DATA.tpt.zip: The LPS data tape

INTRODUCTION

The Immigration Reform and Control Act of 1986 (IRCA) established two separate programs through which illegally resident aliens could obtain lawful permanent resident (LPR) status in the United States. The largest of these programs, described in section 245A of the Immigration and Nationality Act (INA), enabled about 1.6 million aliens, illegally resident since the end of 1981, to become lawful permanent residents. A national sample of these legalized aliens provided the information contained the LPS public use data file. Although IRCA's other legalization program for Special Agricultural Workers (SAWs) granted an additional 1.3 million farm workers LPR status, there are no similar data for the SAW population.

The "pre-1982" legalization program consisted of two phases. During phase I, about 1.8 million applicants qualified for temporary legal residence. Those who did so were given an additional 18 months to prepare their phase II applications, leading to lawful permanent residence. One stipulation of this second phase was satisfaction of English language and American civics requirements, either through participation in special classes or by passing appropriate examinations. About 1.6 million temporary resident aliens successfully completed phase II.

IRCA mandated two Administration reports to the Congress assessing the characteristics of the legalized population before, at the time of, and about five years after legalization. These reports were to focus on section 245A beneficiaries. The basis for these reports was a two-wave, nationwide, longitudinal survey of legalized aliens.

The first wave of the legalized population survey (LPS1) was sponsored by the Immigration and Naturalization Service (INS) in 1989. In 1992, the Department of Labor sponsored a second wave (LPS2), co-sponsored by the Departments of Agriculture and Commerce.

Both waves were conducted by a private research firm, Westat, Inc., of Rockville, MD. Each involved primarily face-to-face interviews with a national sample of legalization applicants, in the language of their choice. LPS1 took place between February and June of 1989, shortly after respondents had received temporary legal residence. It focused on their characteristics and experiences prior to and at the time of phase I application. A total of 6,193 respondents participated in LPS1.

Roughly three years later, after most applicants had completed phase II of this process, the second wave went into the field. Between April and September of 1992, 4,012 LPS1 respondents were reinterviewed. All of these respondents were by that time lawful permanent residents. LPS2 examined the effects of legalization on their lives during the four to five years since they first applied for lawful residence.

The response rates of the two surveys were high: 83 and 82 percent, respectively. Refusals were rare. Most non-response resulted from interviewers' inability to locate respondents.

DATA AND DOCUMENTATION

File Structure. This public use file includes merged LPS1 and LPS2 records for individual respondents, together with certain information drawn from the Legalization Application Processing System (LAPS) at the time of each survey. There are four sections to the individual record, ordered as follows: LPS1 responses, data from the January 1989 LAPS file, LPS2 responses, and data from the January 1992 LAPS file. The fields for LPS2 and final LAPS variables are blank in records of those only interviewed in 1989.

Documentation. *Please read this documentation and annotations in the codebook before processing data.* Although many questions are quoted verbatim in the codebook, it is advisable to verify precise wording and skip patterns using the original questionnaire. The LPS1 and LPS2 questionnaires can be obtained from Shirley Smith at the Department of Labor's Bureau of International Labor Affairs by calling (202) 693-4915.

LPS variable-naming conventions. Naming conventions have been established to help identify the source of specific variables.

- Variable names usually include the number of the corresponding survey question.

- Variables ending in "X" come from LPS1 or the 1989 LAPS file. LPS2 variables have no similar ending. Thus, C1X comes from question C1 of LPS1; variable C1 comes from question C1 of LPS2.
- A trailer of "_L" denotes information from the Legalization Application Processing System (LAPS), 1992.
- While there are some identical questions in the two surveys, questionnaire layout and length differed. Corresponding question numbers do not imply corresponding content.

Information on other household members. Both LPS surveys collected several pieces of information about other persons residing with the respondent.

This is not a hierarchical file. The characteristics of these other individuals appear as variables within the respondent record. Related questions about a single individual are often grouped together. These sets are then repeated for successive individuals. In such cases, the variable names identify both the source question and the individual member. For instance, the LPS2 record includes the following variables in this order: PERSON01, E1601, E1701, PERSON02, E1602, E1702, etc.

LPS1 households contained a maximum of 26 individuals, including the respondent. The corresponding LPS2 maximum was 18. In this flattened file, each record contains data fields to accommodate the maximum number of household members. If the data do not exist, the corresponding fields are blank.

LPS1 Universe of Analysis. The total universe for LPS1 was 1,491,897 applicants. Excluded from this number were aliens whose legalization applications had already been denied or terminated, as well as those who had not reached the age of 18 by January 31, 1989.

LPS1 Sample Design. The first legalization survey utilized a two-stage stratified cluster design. Legalization offices (LOs) were selected for inclusion prior to selection of individual applicants. Non-Mexicans were sampled at about 2.6 times the rate of the Mexican-born to ensure adequate sample size.

The initial LPS1 sampling frame was drawn from a preliminary INS Legalization Application Processing System file listing 107 legalization offices (LOs) and about 1.4 million applicants above the age of 17. The 20 largest LOs were included in LPS1 with certainty. Thirteen other LOs were excluded, either due to location (9 were outside the coterminous United States) or because of minimal case load (4 processed fewer than 400 applications). These omitted LOs accounted for less than one percent of total applications. The remaining 74 legalization offices were sorted by INS Region and proportion of applicants

from Mexico. Another 20 LOs were then systematically sampled, taking account of the number and nationality mix of the applicants processed.

Sub-samples of applicants were then drawn within each sampled LO. This occurred twice, first before the survey entered the field and again in April 1989, after additional applicants had been processed. Records of all applicants in the selected LOs were extracted from the LAPS file. Each case was assigned a measure of size (MOS), i.e. a conditional probability of selection that was directly proportional to the sampling fraction for their nationality and inversely proportional to the selection probability of their LO. Next, the extract file was sorted in descending order by LO size, country of citizenship, sex and year of birth. Finally, a sample of applicants was drawn with probabilities proportionate to the measure of size. This yielded a self-weighting sample within two nationality groups, Mexican and non-Mexican.

LPS2 Universe of Analysis. LPS2 respondents represent just 1,294,562 of the 1,763,434 persons legalized under section 245A of the Immigration and Nationality Act. Excluded from the LPS2 target population were those whose age or immigration status disqualified them from the earlier survey, as well as LPS1-eligibles who had died, permanently left the contiguous 48 states, or failed to receive lawful permanent residence by January 1992.

LPS2 Sample Design. The LPS2 sample was derived from that of LPS1, as follows. The above-listed criteria were applied to an updated LAPS file, to ascertain LPS1 respondents' eligibility for reinterview. This search disqualified 502 individuals, of whom two who were known to have died and 500 had not been granted LPR status.

Resource limitations made it necessary to exclude another 691 of the remaining 5,691 LPS1 respondents. They were randomly selected for exclusion. Westat statisticians then examined the gender, nationality and age distributions of the retained and excluded samples to ensure that differences were not statistically significant.

Table 1 tracks what happened to the original 6,193 LPS1 respondents. Some of the 5,000 individuals who passed the LAPS eligibility screen were later disqualified for reinterview because of additional information gathered in the field. The manner in which interviews were conducted, as well as field and telephone center response rates, are also shown.

WEIGHTING THE LPS FILE

Weighting data from this sample of legalization applicants yields *estimates* of population totals and proportions for the

corresponding years. Since sample sizes differed in 1989 and 1992, weighting factors for the two periods must adjust for this disparity. Your analytical design will dictate which weights to apply. The following discussion of survey design may be helpful in making that determination.

Base weights, LPS1: Base weights for LPS1 (BASEWTX) are equal to the reciprocal of individual's probability of selection for that study, p_i , adjusted for ineligibility and case nonresponse. Initial weights were assigned on the basis of nationality. All Mexican applicants (whether or not they responded to the survey) were assigned an initial base weight of 292.13, while all non-Mexican applicants received an initial weight of 100.40. These corresponded to their final overall selection probabilities of 0.0034 and 0.0091, respectively. The legalized population was broken into 30 weighting classes, i.e. two sexes, four age groups, and three geographic regions as well as two time periods of sample selection for total males and total females). After the sample was drawn from the Legalization Application Processing System file, adjustments for ineligibility and nonresponse were applied to these 30 classes.¹

By design, each weight was equal to the number of legalized aliens represented by the respondent. The sum of the LPS1 weights equalled the total number of legalized aliens in the target population, 1,491,897.

Base weights, LPS2: LPS2 used 24 weighting classes which were identical to the LPS1 categories, except that no distinction was made by period of LPS1 sample selection. Age was computed as of the same date used in the LPS1 calculations.

Because the LPS2 sample was somewhat smaller than that of LPS1, it was necessary to inflate LPS1 base weights to ensure that the LPS2 sample represented the entire target population.

A weighting factor, F_i , for a case in weighting class i was calculated using the following formula:

$$F_i = \frac{S_{i;T}}{S_{i;R} + S_{i;I}}$$

where, for the i th class,

$S_{i;T}$ is the sum of the LPS1 individual weights for Total (see Exhibit B)

$S_{i;R}$ is the sum of the LPS1 individual weights for Respondents;
and

¹See U.S. Department of Justice, INS, Immigration Reform and Control Act: Report on the Legalized Alien Population (1992), Appendix B, for a detailed discussion of LPS1 weights.

$S_{i,I}$ is the sum of the LPS1 individual weights for Ineligibles.

For the i th weighting class, the LPS1 base weight, BASEWTX, was inflated to a corresponding LPS2 weight, BASEWT, as follows:

$$\text{BASEWT} = \text{BASEWTX} * F_i$$

Family weights, LPS1: In studying family characteristics, it is important that each family count only once, regardless of the number of respondents living in that unit. In the following discussion, let:

k_1 = the number of family members on the sampling frame
 k_2 = the number of family members interviewed

A family with more than one member on the sampling frame ($k_1 > 1$) had a greater probability of being sampled than did a unit with just one member represented. Consequently, the former should have a smaller family weight. Within a family, the probability of selection was about the same for each family member, except in the few instances where there were both Mexican and non-Mexican family members. Thus, in LPS1, the probability of selecting a family, p_f , was related to the probability of selecting an individual, p_i , through the formula $p_f = (1 - [1 - p_i]^{k_1})$, which is approximately equal to $p_i k_1$.

Furthermore, when $k_2 > 1$ of these k_1 members appeared on the survey file, all k_2 of the respondents were included in the analysis and each of their weights were further divided by k_2 . This effectively caused the average family response to be used for purposes of estimation. The family weight was calculated using the following formula:

$$\text{Family weight} = \frac{1}{p_f * k_2} \text{ approx equal to } \frac{\text{Base weight}}{k_1 * k_2}$$

Family weights, LPS2: By extension, the following formula was used to derive LPS2 family weights from LPS2 base weights:

$$\text{FAMWT} = \frac{1}{(1 - [1 - \{1/\text{BASEWT}\}]^{k_1}) * k_2^2}$$

Household weights, LPS2: In examining household attributes, each unit must count only once, regardless of the number of respondents in the household. The first survey did not explore the legalization status of unrelated household members. Hence, there was no need for separate household weights for LPS1-- family weights eliminated double-counting. However, LPS2 did examine this issue, increasing the probability that unrelated

household members might also be found on the sampling frame.

Let:

k_3 = the number of household members on the sampling frame

k_4 = the number of household members interviewed

The following formula was used to derive LPS2 household weights to account for this possibility:

$$\text{HHDWT} = \frac{1}{(1 - [1 - \{1/\text{BASEWT}\}]^{k_3}) * k^4}$$

Spousal weights, LPS2: In a few instances, respondents' spouses were themselves LPS2 respondents. Since both reported information on the same family's behavior, their responses must be weighted to reflect this fact. Let:

k_5 = the number of spouses in the sampling frame

k_6 = the number of spouses interviewed

The LPS2 spousal weights are derived as follows:

$$\text{FAMWT} = \frac{1}{(1 - [1 - \{1/\text{BASEWT}\}]^{k_5}) * k^6}$$

Replicate weights, LPS1 and LPS2: The replicate weights for each survey can be used to determine the variance of survey estimates. These weights may be used with the WESTVAR program, available through Westat, Inc. of Rockville, Md. The WESTVAR software is available at:

[HTTP://WWW.WESTAT.COM](http://WWW.WESTAT.COM)

WESTVAR user support can be obtained at the following e-mail address:

WESTVAR@WESTAT.COM

For further information, contact Mike Brick, of Westat, Inc., at (301) 251-1500.

CAUTIONARY NOTES

Estimates of change. LPS2 reinterviewed only 4,012 of the original 6,193 LPS1 respondents. Estimates of change over the survey interval can thus be derived either by using a) the full sample at each date, or b) the restricted sample of 4,012 for whom two records are available, i.e. LPS2 respondents.

Using the full sample for each period retains the larger number of observations for 1989. However, this approach yields synthetic estimates that may be influenced by changes in sample composition, e.g., the loss of certain out-migrants. If the full samples are used, findings for each period should be inflated by the corresponding period weights.

The Second Report on the Legalized Population based its estimates of change on the smaller, continuous sample, LPS2 respondents. LPS2 weights were applied to data for both time periods.

Geographic detail and geographic mobility. The legalization surveys were designed to obtain reliable national estimates for the legalized population. Users should be aware that they were not designed to provide robust State or local estimates. *Sub-national estimates should be treated with caution.* Except in California, samples are very small and sampling variance is likely to be high.

By the same token, these data are not an appropriate basis for the study of geographic mobility of legalized aliens. Efforts to relocate LPS1 respondents in 1992 were most concentrated in the State where they resided three years before. Thus out-migrants may be underrepresented in the second survey.

Educational attainment. Grade levels and educational degrees in the United States do not always correspond to those of source countries. As a result of such disparities and the limited schooling most respondents had had, information on this topic may be inconsistent between the two surveys.

Receipt of financial assistance. The LPS1 survey used a screening question (G1X) to identify respondents who did not receive financial assistance. They were skipped past the section of the questionnaire dealing with social services. There is reason to believe that many respondents misunderstood question G1X, answering "no" when they should have said "yes." This would have resulted in understatement of social service usage in 1989. Consequently, *estimates of change that use LPS1 data as a baseline may exaggerate true increases in social service usage following legalization.* LPS2 explored the probable magnitude of this bias. Compare responses to question F1 (a repeat of G1X), and F2 (the LPS2 social service screen).

Legal status of family members. Although both files include

information on the immigration status of family members, these data were derived differently during the two time periods. The 1989 survey did not ask about status directly. Rather, it was inferred by comparing LPS1 survey data with information from the LAPS file. Family members were classified as either legalized aliens, citizens, or probable illegal aliens.

Resident status was explored more directly in LPS2 (questions E1301 through E1318). Respondents were read and given a handcard listing six categories, which exhausted all legal possibilities. Two residual categories, "some other status" and "don't know," may identify certain illegal aliens. *Care should be exercised in interpreting these data.* Some in the "don't know" category are legal residents whose status the respondent really did not know. Likewise, because status is a sensitive issue, some of those classified in legal categories may actually be undocumented.

Table 1. Relationship Between LPS1 and LPS2 Samples, Manner of Contact with LPS2 Respondents, and Disposition of Cases

| Disposition | LPS2 Contact Status | | | Percent of LPS2 Eligibles |
|--------------------------------|---------------------|------------|--------------|---------------------------|
| | Field | Phone | Total | |
| LPS1 Sample | NA | NA | 6,193 | NA |
| Not Sampled for LPS2 | NA | NA | 1,193 | NA |
| Ineligible in January 1992 | NA | NA | 502 | NA |
| Not in LAPS file | NA | NA | 3 | NA |
| Known dead | NA | NA | 2 | NA |
| No decision | NA | NA | 471 | NA |
| Awaiting information | NA | NA | 18 | NA |
| Denied | NA | NA | 4 | NA |
| Terminated | NA | NA | 5 | NA |
| Others Excluded from Sample | NA | NA | 691 | NA |
| Sampled for LPS2 | 4,714 | 286 | 5,000 | NA |
| Later Found to be Ineligible | 96 | 2 | 98 | NA |
| Left U.S. Permanently | 69 | 1 | 70 | NA |
| Known to be dead | 27 | 1 | 28 | NA |
| Believed to be Eligible | 4,618 | 284 | 4,902 | 100% |
| LPS2 Respondents | 3,851 | 161 | 4,012 | 82% |
| Response rate | 83% | 56% | 82% | 82% |
| Completed in person | 3,423 | NA | 3,423 | 70% |
| Completed by telephone | 417 | 161 | 578 | 12% |
| Completed by proxy | 11 | NA | 11 | 0% |
| Non-Respondents | 767 | 123 | 890 | 18% |
| Couldn't be located | 574 | 96 | 670 | 14% |
| Refusal, breakoff | 97 | 6 | 103 | 2% |
| Not at home | 33 | 15 | 48 | 1% |
| Left U.S. temporarily | 21 | 2 | 23 | 0% |
| Too ill to interview | 17 | - | 17 | 0% |
| Moved too far, no phone | 16 | - | 16 | 0% |
| Postal problem | 8 | - | 8 | 0% |
| Language problem | 1 | 4 | 5 | 0% |

NA Not available.
 - Not applicable

Table 2. LPS2 Respondents, Unweighted and Weighted Counts

| Characteristic | Unweighted Sample | Weighted Legalized Adults | Characteristic | Unweighted Sample | Weighted Legalized Adults |
|-------------------------------------|-------------------|---------------------------|---|-------------------|---------------------------|
| Total | 4,012 | 1,294,562 | Total | 4,012 | 1,294,562 |
| Sex | | | Labor Force Status | | |
| Male | 2,165 | 724,099 | Employed | 3,034 | 969,480 |
| Female | 1,846 | 570,049 | Unemployed | 247 | 83,199 |
| Region of Origin | | | Not in Labor Force | 730 | 241,468 |
| Mexico | 1,937 | 893,035 | Firm Size (number of workers employed) | | |
| Central America | 1,029 | 189,453 | 1 employee | 337 | 91,870 |
| Other Western Hemisphere | 446 | 92,500 | 2-4 | 367 | 110,439 |
| Canada and Other | 598 | 118,951 | 5-9 | 345 | 113,178 |
| Period of Arrival | | | 10-49 | 844 | 290,099 |
| Before 1975 | 778 | 284,665 | 50-99 | 373 | 123,996 |
| 1975-1979 | 1,493 | 497,096 | 100-499 | 479 | 150,936 |
| 1980 or later | 1,722 | 505,896 | 500 or more | 245 | 74,789 |
| Age at Arrival | | | Occupation | | |
| 17 or less | 1,027 | 394,381 | Prof., Tech., and Mgr. | 305 | 77,382 |
| 18 or more | 2,965 | 893,115 | Sales | 186 | 51,270 |
| Years of Schooling Completed | | | Administrative support Services | 236 | 67,768 |
| 0 to 6 | 1,582 | 597,322 | Private household | 201 | 51,912 |
| 7 to 11 | 1,011 | 327,527 | Food | 298 | 100,703 |
| 12 or more | 1,418 | 369,298 | Other services | 371 | 108,622 |
| Ability to Speak English | | | Farm | 129 | 54,160 |
| Not at all | 745 | 283,035 | Craft, repair | 417 | 141,567 |
| Not very well | 1,597 | 540,834 | Operator, transport | 657 | 228,245 |
| Very well, well | 1,658 | 467,176 | Labor | 233 | 87,849 |
| Place of Residence | | | Industry | | |
| California | 2,178 | 749,710 | Agriculture | 116 | 49,339 |
| Texas | 464 | 165,985 | Construction | 226 | 81,824 |
| New York | 428 | 94,221 | Manufacturing | | |
| Other | 941 | 284,231 | Durable goods | 308 | 111,557 |
| Age in 1992 | | | Nondurable goods | 366 | 125,720 |
| 20-29 | 772 | 285,788 | Wholesale, Retail Trade | | |
| 30-39 | 1,771 | 580,753 | Restaurant | 300 | 100,365 |
| 40 or more | 1,442 | 421,368 | Other | 474 | 152,432 |
| Family Income in 1991 | | | Services | | |
| \$11,999 or less | 785 | 256,241 | Business | 224 | 65,435 |
| \$12,000-\$19,999 | 925 | 313,647 | Private household | 218 | 58,340 |
| \$20,000-\$29,999 | 903 | 298,529 | Other Services | 511 | 146,397 |
| \$30,000 or more | 1,018 | 308,640 | Other | 288 | 77,682 |
| | | | Class of Work | | |
| | | | Wage and Salary | 2,596 | 849,690 |
| | | | Government employees | 129 | 34,280 |
| | | | Self-employed | 268 | 73,375 |
| | | | Unpaid family worker | 24 | 6,896 |

Some categories may not add to total N due to missing data for some respondents.

TABLE C2

Critical Items

A number of questions, necessary for analysis, were identified as critical items. If these items were not completed, an attempt was made to recontact the respondent and obtain the necessary information. The critical items included the following:

LPS1 CRITICAL ITEMS SUBJECT TO TELEPHONE VERIFICATION

| | |
|-----------|---|
| A1AGEX | Age |
| A2YEARX | Year in which applied for temporary residence |
| A3X | Sex |
| B1CSPECX | Language (specific) |
| B2CX | Read instructions or recipies in English |
| C3YEARX | Year entered US for more than 3 days |
| C5YEARX | Year first entered US and stayed for more than 3 days |
| C6YEARX | Year last entered US and stayed for more than 3 days |
| C10ISPECX | Reason for staying in US (specify) |
| D11X | Industry of employment, 1st US job |
| D12X | Occupation of employment, 1st US job |
| D16X | Initial weekly earnings, 1st US job |
| D19X | Industry, job with most hours, 1st year in US |
| D20X | Occupation, job with most hours, 1st year in US |
| D22X | Initial weekly earnings-job w. most hours, 1st yr in US |
| D23SPEC | What doing most of week bef. applied for temp residence |
| (specify) | |
| D24X | Any work at all, week before application? |
| D25X | Temp absent or on layoff week before application? |
| D26X | Why absent fr work, week before application. |
| D32AMONX | Month began job at which wked week bef application |
| D32AYEARX | Year began job at which wked week bef application |
| D32BMONX | Month ended job at which wked week bef application |
| D32BYEARX | Year ended job at which wked week bef application |
| D33X | Industry, job week bef application |
| D34X | Occupation, job week bef application |
| D37X | Usual weekly earnings bef ded., job week bef applic. |
| D40X | Industry, 2nd job week bef application |
| D41X | Occupation, 2nd job week bef application |
| D43X | Usual weekly earnings, 2nd job week bef applic. |
| E1X | Marital status when applied for temp residence |
| E2X | Number parents living with resp when applied |
| E3X | Number children living with resp when applied |
| E4X | Number bros/sisters(and bro/sist in-law) living with respondent when applied |
| E5X | Number grandparents living with resp when applied |
| E6X | Number grandchildren living with resp when applied |
| E701X | |
| -E1125X | Relationship, ed. attain., and wk status and immigration status for all family members at time of application FOR ALL FAMILY MEMBERS |
| E14X | Number of family relatives elsewhere in US when applied for temporary residence |
| F1X | Whether now pregnant |
| F2X | Whether have seen a doctor abt pregnancy |
| F3AX | Preg care pd for by insurance through work |
| F3BX | Preg care pd for by insurance from ins company |
| F3CX | Preg care pd for by resp or spouse/partner |
| F3DX | Preg care pd for by parents or relatives |
| F3EX | Preg care pd for by Medicaid or Medical Assistance |
| F3FX | Preg care pd for by other (specify) |
| G3GX | Preg care free of charge |
| F6AX | Any US-born child need extra hosp. stay at birth? |
| F6BX | How many US-born children needed extra hosp stay? |
| F7X | Usual source of health care before applied for temp res |
| F11AX | Personal health coverage by Medicaid or state gov't medical assistance |
| F11BX | Personal health coverage by Medicare |
| F11CX | Personal health coverage by health ins through work |
| F11DX | Personal health coverage purchased directly from company |
| F12AREASX | Reasons for not being covered by private health ins |
| F12ARANKX | and rank of importance for each reason |
| through | . |
| F12HREASX | . |
| F12HRANKX | . |
| G1X | Whether any family member in household received financial assistance or food stamps from public or gov't agency at time applied for temp res. |
| I3X | Total family income in 1987 |