

California Economic Policy Day Labor in the Golden State Web Appendix B

Description

This web appendix contains two sections. Section I provides information on the design, implementation, and findings of the National Day Labor Survey. Section II offers greater detail on the characteristics of day laborers who use worker centers and considers these characteristics in relation to the lower number of hours worked at these centers.

Contents

- I. The National Day Labor Survey, Technical Details
- II. Hours of Work and Worker Centers

Acknowledgments

I would like to thank Abel Valenzuela, Jr., at UCLA for his invaluable comments and insights into the day labor market and the University of California Los Angeles Center for the Study of Urban Poverty for providing The National Day Labor Survey. The following people provided me with valuable discussion, information, and data: Ana Luz Gonzalez (UCLA), Kristina Campbell (Mexican American Legal Defense and Education Fund), Catalina Amuedo-Dorantes (San Diego State and PPIC), Marianne Bitler (UC Irvine and PPIC), Giovanni Peri (UC Davis and PPIC), Michael Massagli (Harvard), Mónica Ramírez (ACLU), Virginia Kice (U.S. Immigration and Customs Enforcement), and seminar participants at PPIC. I would also like to thank Orley Ashenfelter (Princeton), Lynette Ubois (PPIC), Toby Ewing (New California Network), Ellen Hanak (PPIC), David Neumark (UC Irvine and PPIC), and Hans Johnson (PPIC) for particularly excellent suggestions and reviews. All errors of fact or interpretation are those of the author. Any opinions expressed in this report do not necessarily reflect the views of PPIC or the principal investigators, collaborators, or funders of the National Day Labor Survey.

**Public
Policy
Institute of
California**

Section I: The National Day Labor Survey, Technical Details¹

The NDLS is a multistage, clustered survey of day laborers conducted from July 2004 to August 2004 (Valenzuela et al., 2006). Surveying this population presents many challenges and survey design is crucial to obtaining credible results. This section of the web appendix discusses the technical details involved in the design of the NDLS, explains how the survey was conducted, and demonstrates how the population estimates for the day labor population were made.

Designing the NDLS

The NDLS data are derived from 2,660 completed surveys conducted in 36 Metropolitan Statistical Areas (MSAs) across the country. To give a higher probability of selection to cities where day laborers are likely to concentrate—those with a large total population and a large Latino population—the NDLS team implemented a disproportionate stratified sampling frame.

First, the team randomly selected 50 MSAs nationwide, based on (1) inclusion in one of four strata, and (2) within each stratum, an MSA's total population relative to the smallest MSA's total population in that stratum.

The four strata were defined as:

- (1) Small total population and small Latino population—less than 295,000 total population and less than 13,000 Latino population. This group consisted of 131 MSAs.
- (2) Small total population and large Latino population—less than 295,000 total population and 13,000 or more Latino population. This group consisted of 35 MSAs.
- (3) Large total population and small Latino population—295,000 or more total population and less than 13,000 Latino population. This group consisted of 38 MSAs. and
- (4) Large total population and large Latino population—295,000 or more total population and 13,000 or more Latino population. This group consisted of 125 MSAs.

¹ For further discussion, see Abel Valenzuela, Jr., Nik Theodore, Edwin Melendez, and Ana Luz Gonzalez, *On the Corner: Day Labor in the United States*, UCLA Center for the Study of Urban Poverty, Los Angeles, California, 2006.

A systematic procedure for selecting MSAs within each stratum insured geographic representation. MSAs were listed by Census region in rough geographical order, beginning in the northeast, proceeding south and west across the country, and terminating in the northwest.

Within each stratum, the total population of each MSA was divided by the population of the smallest MSA to determine the number of times the MSA was listed, which allowed MSAs to be selected proportionate to their size. The total number of listings was divided by the number of MSAs (n) to be selected from each stratum to determine a sampling interval (i). A random number table was used to generate an initial sample point (p), then the sampling interval was used to select the $(p + (i))$, $(p + 2(i))$, ..., $(p + n(i))$ listed areas from each stratum. In all, 50 MSAs were selected: 10 from stratum one, five from stratum two, five from stratum three, and 30 from stratum four. Because two MSAs in stratum four (New York and Los Angeles) were selected twice, this brought the total count of unique MSAs to 48.

Of these 48 MSAs, 13 were not surveyed because the research team found that they did not have a day labor population. (This included one California MSA, Yuba.) This left 35 MSAs with a confirmed day labor population.

Next, the research team collected data on a representative number of worker centers. This meant adding a supplemental list of another 11 MSAs with worker centers that had not been selected as part of the random sample, bringing the total to 46 MSAs.

The team then had to identify sites and select people to survey. To do so, all day labor worker sites within the 46 MSAs were identified, and workers were randomly selected for the survey. These steps are described in more detail below.

Identifying Sites

From November 2003 to March 2004 the research team identified hiring sites and day labor populations in cities in each of the MSAs selected in the sample. Several procedures were

used to identify day labor hiring sites. First, a variety of local groups were contacted, including community-based organizations, advocacy groups, churches, home improvement stores, police departments, city planning departments, and merchants. The research team also conducted Internet searches (i.e., looked at newspapers, websites, and articles) to identify as many sites as possible within each MSA. More than half of the hiring sites were identified using this method. Finally, new sites were identified using a “referral” system that in many ways resembles snowball sampling. Day laborers were approached at different sites and asked where else they go in search of work, and workers at those new sites were asked the same question. This method led to the identification of the remainder of the sites.

In preparation for selecting survey participants, all known sites were visited, except for some sites in Los Angeles, New York, and Orange County—these were not contacted because of budget restrictions. (This aspect of the sample design is also incorporated into the survey weights.) Despite multiple visits to the identified day labor hiring site, no day laborers were found in 10 MSAs, including one MSA in inland California (Fresno). This brought the total number of MSAs surveyed to 36.

Selecting Day Laborers

Finally, the research team used information from the site identification research described above (i.e., field notes and counts of day laborers) and relied on a scouting exercise conducted in May 2004, to establish “selection” counts for each site. Selection counts were based on the size of the hiring site (total number of day laborers) before the actual survey itself, which was conducted in July and August 2004.

Selection counts operated in the following manner. At 6:30 in the morning, the research team arrived at a given site and counted all workers. The count was repeated every hour until 10:30 a.m. and included day laborers who arrived after the initial count had been made. Included

in the count was a general description of each worker (usually based on physical features and attire). After all the workers had been counted, a simple random sampling procedure was administered whereby potential participants were identified. Each worker who fell within the selection count (a random number) was approached and asked to participate in the survey. Workers were randomly selected at 264 hiring sites in 139 municipalities in 20 states and the District of Columbia.

A total of 2,660 surveys were completed, 753 in California, in 36 MSAs. The majority of the interviews were administered in Spanish and all were conducted face to face. The survey was undertaken during a continuous seven-week period (the last week of June to mid-August 2004). Each interview included more than 100 questions and took approximately 35 minutes to complete, and each participant received \$10 and a certificate of participation.

Survey Participation Issues

A scientific study of day laborers—a highly mobile, highly visible, yet largely unknown population—requires creative research approaches. Before implementing the survey, nine research team members were sent to scout all of the MSAs that had conflicting or “thin” information about day laborers. As a result of this exercise, the team was able to confirm the site identification research and eliminate sites that did not have a day labor population. Sending researchers to scout proved to be quite fruitful as they were also able to establish a relationship with key community leaders in each area. In addition, the research team members sent letters to community organizers and community-based organizations informing them of the date and time the survey would take place.

The study was surprisingly well received by most workers. Most, if not all, were extremely cooperative and many were excited to participate in a major research study. When available, community organizers would come to the site the day of the survey to help advertise

the study and establish trust with workers. The refusal rate (day laborers unwilling to participate in the survey) was considerably low (21%), particularly in light of what most survey experts regard as a difficult population to approach and convince to participate in a research study.

Analyzing the Results

Under a complex sample survey design, it is necessary to use weights that reflect the design whenever subsets of respondents with differing probabilities of selection are combined into a single group for analysis. In the NDLS, the probability of selection varied across sites within MSAs, and the probability of an MSA being included differed across the four strata and the supplemental list described above. Therefore, analysis of the NDLS that combines respondents from different sites within MSAs and from different strata should use design-based weights. The weights used in “Day Labor in the Golden State” are based on the entire set of sampled MSAs (random and nonrandom) and are representative at the MSA level. The estimates obtained from the NDLS are adjusted by stratum, cluster, and individual sampling weight using the *svy* commands in STATA. The derived estimates are MSA-level estimates.

California MSAs

The California data represent day laborers at 92 hiring sites in six MSAs—Los Angeles, Oakland, Orange, San Diego, San Francisco, and San Jose. Given the location of these MSAs, it is important to note that any statistical analysis of the NDLS California sample represents characteristics of day laborers in the large coastal metropolitan areas of the state—which may differ from those in California’s inland regions.

In accordance with the original sample design, the first column of Table 1 lists the California MSAs eligible for sampling, grouped by stratum. The sample design resulted in eight of 25 California MSAs being eligible for the NLDS. The only MSA from stratum one, Redding, was not included. Yuba City is the only California MSA from stratum two. There were no

California MSAs in stratum three. Four MSAs were sampled from stratum four. Finally, three California MSAs are from the supplemental list.

Estimating the Day Labor Population

By design, day labor markets that may lie in geographic areas wholly outside MSA boundaries were not included in the NDLS. For the areas that were included, the day laborer counts at sampled sites were used to generate estimates of the day labor population for the nation and for California.

To build a national estimate, observed counts within each sampled MSA were used to impute values for sites where observations were not made at the time of the survey. The sum of counts for each MSA then contributed to the stratum-level population estimate according to a weight inversely associated with the size of the contributing MSA. Thus the largest MSAs are self-representing, whereas smaller MSAs represent numerous smaller MSAs that were not actually sampled. The national estimate of 117,647 day laborers in NDLS by Valenzuela et al. (2006) depends on the stratum-level estimates given in the second to last column of Table 1.

Table B.1: California Cities Eligible for and Selected for NDLS, and Estimated Day Labor Population in the United States and California

Strata	California MSAs Eligible for Selection	Randomly Selected? In NDLS?		Population Estimates	
				United States	California
Stratum 1				2,215	19
	Redding				
Stratum 2				5,755	1,028
	Chico—Paradise				
	Yolo				
	Yuba City	Y	N		
	Merced				
	San Luis Obispo—Atascadero—Paso Robles				
	Santa Cruz—Watsonville				
Stratum 3				5,690	-
	None				
Stratum 4				92,566	33,255
	Bakersfield				
	Fresno	Y	N		
	Los Angeles—Long Beach	Y	Y		
	Modesto				
	Orange County	Y	Y		
	Riverside—San Bernardino				
	Sacramento				
	San Jose	Y	Y		
	Santa Barbara—Santa Maria—Lompoc				
	Santa Rosa				
	Stockton—Lodi				
	Vallejo—Fairfield—Napa				
	Ventura				
	Visalia—Tulare—Porterville				
Supplemental				11,421	5,996
	Oakland	N	Y		
	San Diego	N	Y		
	San Francisco	N	Y		
Total		5 MSAs	6 MSAs	117,647	40,298

Within the national sample, California respondents represent probability samples of workers in each selected MSA. These MSAs cover a large percentage of the California metropolitan area population—62 percent of the state’s total population and 64 percent of the state’s Hispanic/Latino population. Although the survey does not sample all geographic areas throughout California, a combination of California and national data can be used to estimate the total size of the day labor population in the state.² For strata one and two—for which there are no site counts in California MSAs—the simple average for MSAs in the same stratum within the national sample is multiplied by the number of times the city was listed (as described above) to derive the estimated day labor population. (As noted, there are no California cities in stratum three.) For California cities in stratum four, the imputed average for cities without any count is obtained by multiplying the per-stratum average number of day laborers in states other than California multiplied by the number of listings for that MSA. This imputed number is added to the observed counts of sampled California cities. The last column of Table 1 gives the estimated population for each stratum within California and shows that there were over 40,000 day laborers in California at the time of the survey.

² Michael P. Massagli, statistical consultant to the NDLS team, provided these calculations.

Section II: Hours of Work and Worker Centers

Day laborers in worker centers earn \$65 less per week than those at informal sites, primarily because they work about four hours less. Is the lower intensity of work a result of participating at worker centers or a result of the type of workers who use each site? An analysis of worker characteristics found that those using workers centers are slightly older, more likely to be married or living with a partner, and less likely to hold a regular job than those using informal sites. They were also more likely to be based in Los Angeles than in other metropolitan areas. The complete list of characteristics is presented in Table 2.

Table B.2: Characteristics of Day Laborers at Worker Centers and Informal Sites

	Hours Worked in Day Labor (Including 0)						Positive Hours Worked in Day Labor	
	Total		Worker Center		Informal Site		Mean	Std.Err
	Mean	Std.Err	Mean	Std.Err	Mean	Std.Err		
Hours worked in day labor	17.081	-0.879	14.535	-1.321	18.584	-0.942	22.775	0.974
Day labor worker center	0.371	-0.081	1.000		0.000		0.358	0.083
Age	34.470	-0.599	37.151	-0.538	32.888	-0.641	33.971	0.611
Day labor months of experience	37.408	-3.245	35.171	-5.977	38.729	-3.789	36.010	3.422
No schooling	0.226	-0.020	0.229	-0.032	0.224	-0.027	0.244	0.026
Elementary/junior high	0.607	-0.027	0.588	-0.029	0.619	-0.038	0.587	0.035
High school/General Equivalency Diploma (GED)	0.112	-0.016	0.116	-0.026	0.110	-0.020	0.112	0.019
Some college	0.055	-0.011	0.068	-0.019	0.047	-0.012	0.057	0.014
Married/living with partner	0.421	-0.026	0.355	-0.042	0.460	-0.029	0.439	0.024
English speaking ability	0.243	-0.026	0.286	-0.050	0.217	-0.027	0.226	0.031
Born in Mexico	0.706	-0.038	0.712	-0.064	0.703	-0.049	0.698	0.047
Undocumented	0.808	-0.020	0.772	-0.029	0.830	-0.025	0.823	0.022
In good health	0.657	-0.024	0.650	-0.045	0.660	-0.028	0.662	0.027
Minimum wage to work in day labor	9.239	-0.184	8.999	-0.292	9.381	-0.235	9.356	0.207
Hours worked in regular job	5.145	-0.746	3.794	-0.963	5.942	-0.933	3.605	0.635
Los Angeles MSA	0.529	-0.047	0.758	-0.085	0.394	-0.079	0.541	0.045
Orange County MSA	0.057	-0.011	0.112	-0.058	0.189	-0.040	0.052	0.011
San Jose MSA	0.161	-0.025	0.061	-0.042	0.188	-0.040	0.170	0.024
Oakland MSA	0.054	-0.008	0.025	-0.023	0.075	-0.018	0.060	0.010
San Diego MSA	0.060	-0.019	0.011	-0.012	0.079	-0.015	0.055	0.019
San Francisco MSA	0.141	-0.022	0.034	-0.027	0.075	-0.031	0.122	0.023
Sample size	620		161		459		457	

Notes: All estimates are weighted. See Table 2 of "Day Labor in the Golden State" for sample definition.

Table 3 presents the results of regressions of total hours worked per week on whether or not a day laborer participates in a worker center. The first two columns report results for the sample of all day laborers, including those who worked zero hours in the previous week, and the second two columns report results only for those workers who obtained some day labor work in the previous week. In each case, the first column shows the simple difference between hiring sites in hours worked, and the second shows the adjusted difference, controlling for worker characteristics and regional controls. Regardless of definition, and with or without controls, the mean difference in Table 3 between informal sites and worker centers is four hours (as seen in the coefficient on the “Day labor worker center” variable in the first row).

Table B.3: Determinants of Hours Worked per Week, by Worker Center Status

	Hours Worked in Day Labor (Including 0)				Positive Hours Worked in Day Labor			
	Est.	Std.Err	Est.	Std.Err	Est.	Std.Err	Est.	Std.Err
Day labor worker center	-4.049	(1.623) ***	-4.468	(1.759) ***	-4.024	(1.580) ***	-4.132	(1.925) ***
Age			-0.110	(0.073)			-0.059	(0.082)
Day labor months of experience			-0.012	(0.011)			-0.018	(0.016)
Elementary/junior high			-7.836	(3.622) ***			-7.795	(3.154) ***
High school/GED			-9.370	(3.602) ***			-8.337	(3.323) ***
Some college			-5.070	(4.231)			-4.809	(4.133)
Married/living with partner			2.437	(1.424) *			1.787	(1.532)
English speaking ability			-0.402	(1.507)			1.931	(1.746)
Born in Mexico			-0.715	(1.668)			-0.658	(1.211)
Undocumented			1.280	(1.910)			0.195	(2.309)
In good health			4.415	(1.361) ***			4.557	(1.413) ***
Minimum wage to work in day labor			-0.123	(0.113)			-0.346	(0.100) ***
Hours worked in regular job			-0.180	(0.081) ***			-0.018	(0.084)
Orange County MSA			-3.115	(2.992)			0.435	(2.248)
San Jose MSA			0.428	(1.793)			-1.775	(2.443)
Oakland MSA			2.080	(2.438)			-1.034	(3.301)
San Diego MSA			-1.300	(2.048)			0.655	(2.704)
San Francisco MSA			-4.340	(1.776) ***			2.653	(2.220)
Constant	18.584	(0.942) ***	29.407	(5.503) ***	24.132	(1.141) ***	33.512	(5.523) ***
Sample size	620		620		461	509	461	
R-squared	0.022		0.1041		0.0112		0.1527	

Source: Authors' calculations using NDLS.

Notes: All estimates are weighted. See Table 2 of “Day Labor in the Golden State” for sample definition. Excluded regional category is Los Angeles. Excluded educational category is no school. Excluded country of origin category is non-Mexican-born. ***, **, and * indicate that the difference in means between worker centers and informal sites is statistically significant at the 99 percent, 95 percent, and 90 percent levels, respectively.