CENTER FOR THE STUDY OF
Economy & Society

Department of Sociology
327 Uris Hall
Cornell University
Ithaca, NY 14853-7601

CSES Working Paper Series

Paper #62

Michael Spiller, Annette Bernhardt,

Jason Perelshteyn, Douglas Heckathorn, et.al.

"Sampling, Fielding, and Estimation in the 2008 Unregulated Work Survey"

November 2010

Technical Report:

Sampling, Fielding, and Estimation in the 2008 Unregulated Work Survey

by Michael W. Spiller, Annette Bernhardt, Jason Perelshteyn, and Douglas Heckathorn

with Ruth Milkman, Nik Theodore, Mirabai Auer, James DeFilippis, Ana Luz González, Victor Narro, and Diana Polson

November 23, 2010

Center for the Study of Economy and Society -- Cornell University

For more information, please contact Michael W. Spiller (mws24@cornell.edu) or Annette Bernhardt (abernhardt@nelp.org).

This research was generously funded by the Ford Foundation, the John Randolph Haynes and Dora Haynes Foundation, the Joyce Foundation and the Russell Sage Foundation. The views expressed in this report are solely those of the authors.

Table of Contents

1. Introduction

Studying employers' violations of employment and labor laws is a challenging task. Employers are unlikely to admit that they are paying workers less than the minimum wage, denying workers meal breaks, or otherwise breaking the law. Businesses with the worst conditions may be operating underground and thus difficult to find. Workers who need to support their families are understandably reluctant to talk to researchers about their employers, because of possible retaliation, their immigration status, or because they are working off the books.

The result is that existing data are inadequate to assess the current state of employer compliance with U.S. workplace laws. Standard government surveys such as the decennial Census do not gather the detailed data one would need to accurately identify workplace violations. And enforcement agencies usually only document the relatively small number of cases that come before them, leaving a significant part of the labor market unmeasured. It is impossible to determine, for example, whether the cases of overtime violations brought to the U.S. Department of Labor in a given year represent a small or large share of the actual violations.[1]

The 2008 Unregulated Work Survey was intended to fill this data vacuum. We surveyed workers themselves, building on an emerging body of research that has established the viability of gathering reliable data on employment and labor law violations from workers. Specifically, in 2008 we conducted a large-scale, representative survey of 4,387 low-wage workers in the three biggest U.S. cities.

We adopted two key methodological innovations to overcome the inadequacies of previous studies. First, we used a cutting-edge sampling methodology that allowed us to reach the full range of workers in the low-wage labor market, including unauthorized immigrants and off-the-books workers. Second, we developed an extensive questionnaire that allowed us to rigorously assess whether employment and labor laws were being broken, without relying on workers' own knowledge of these laws.

These two innovations demanded extensive planning of the fielding process, vigilant monitoring of the sampling progress, detailed coding of relevant workplace laws, and the development of custom estimation procedures. This technical report provides documentation of the fielding process, sampling procedures, statistical adjustments and calculations used in the 2008 Unregulated Work Survey. For a broader overview of the study and its findings, see Bernhardt, et al. (2009).

[1] The exceptions here are: (1) the random compliance surveys conducted by the U.S. Department of Labor in 1999 (United States Department of Labor 2001), and (2) the misclassification of independent contractors, where state agencies have been able to use administrative data to robustly estimate the extent of misclassification (Carré and Wilson (2004), DeSilva et al. (2000), and United States General Accountability Office (2006)).

2. Defining the Sampling Universe

Our survey used a chain-referral sampling method called Respondent-Driven Sampling (see section 7) in order to reach low-wage workers often missed by standard sampling methods. Because these workers are difficult to locate, a sampling frame cannot be constructed before the survey begins. Instead, researchers must define rules for sample membership that are both precise and communicable to sample recruiters. Following is a description of how we developed the criteria used to determine eligibility for the survey.

In order to define the "sampling universe," we first needed to determine how to identify "low-wage workers." Out of a number of possible approaches, we decided to focus on selecting low-wage industries and occupations through which we could define sample eligibility. Because we did not want to rely on workers' calculation of their wages (which might be inaccurate or might fluctuate from workweek to workweek), we did not specify a wage cut-off for individual workers.[2] This means that workers earning any amount of money could enter the sample, as long as they worked in an industry or occupation included in our sampling frame.

We used the Census 2000 5% Public-Use Microdata Series to identify low-wage industries and occupations within each city. This was the most recent data available to us with a large enough sample size to estimate wage central tendency and dispersion at the 3-digit industry/occupation classification level. To estimate industry- and occupation-specific median wages, we first excluded any respondent who did not meet the following criteria:

1. Was age 18 or older
2. Was a member of the civilian work force
3. Was employed in Los Angeles County, Cook County (Chicago), or the five boroughs of New York City
4. Was a "front-line" worker (not a manager, supervisory worker, or technical worker)[3]
5. Was not self-employed

We then excluded from our sampling universe any industry or occupation that had fewer than 30 respondents in the Census, because it is extremely unlikely that our sample could reach enough of these rare workers to estimate industry- or occupation-specific violation rates. Next, we estimated an hourly wage for each Census sample member as yearly wage income divided by average number of hours worked per week times number of weeks worked per year. Finally, we estimated the median hourly wage for each 3-digit industry and occupation by city based on the estimated hourly wages of the individual workers.

After determining the industry- and occupation-specific median wage rates, we needed to define a cutoff to distinguish those that were "low-wage." After reviewing the relevant

[2] Note that if we censored sampling based on individual workers' wages, our industry- and occupation-specific estimates of wage violations would be biased upwards by an undeterminable amount.
[3] Excluded 3-digit Census 2000 occupation codes are: 1-244, 255-354, 370-373, 400-401, 420-421, 430-432, 470-471, 500, 600, 620, 700, 770, 900, 980, and 981.

literature, we located three potential low-wage cutoffs: 2/3 of the mean national wage (Freeman and Schettkat 2000), 2/3 of the median national wage (Organisation for Economic Co-operation and Development 1996), and 85% of the median national wage (Organisation for Economic Co-operation and Development 1994).[4] We chose to use 85% of the median city-specific wage as the cutoff for in our sampling universe.

We determined the city-specific median wages using the 2006 Current Population Survey (deflated to year 2000 dollars for use with the census data).[5] These median wages were: Chicago, $14.85; Los Angeles, $14.00; New York City $15.38 (in 2006 dollars). We then defined "low-wage industries" as those whose median wage for front-line workers was less than 85 percent of the city's median wage: Chicago, $12.62; Los Angeles, $11.90; New York City $13.07 (in 2006 dollars). Note that this procedure produced a slightly different set of industries and occupations for each city (see Table 2 below).

3. Fielding the Survey

Standard surveying techniques—phone interviews or Census-style door-to-door interviews—rarely are able to fully capture the population that we are most interested in: low-wage workers who may be hard to identify from official databases, who may be vulnerable because of their immigration status, or who are reluctant to take part in a survey because they fear retaliation from their employers. Trust is also an important issue when asking for the details about a worker's job, the wages they receive, whether they are paid off the books or not, and their personal background.

In light of these challenges, we adopted an innovative sampling method that operates through respondents' own social networks. All of the workers in the low-wage worker population have friends, family, or co-workers that they come into regular contact with and rely on for support; thus our approach relied on a system in which survey respondents recruited people they already knew into the survey, a recruitment technique known as chain-referral sampling.

The best known sampling method using this form of recruitment is snowball sampling, an approach that yields only convenience samples which are not representative of the target population. Snowball sampling cannot replicate the desirable properties of probability sampling methods that allow one to make inferences about the population based on sample data. This method therefore would not have fulfilled the aims of our study.

To overcome this limitation, we adopted a newer form of chain-referral sampling, developed by co-author Douglas Heckathorn in the late 1990s (Heckathorn 1997, 2007). This method was subsequently further developed in collaboration with other scholars. Called *Respondent-Driven Sampling* (RDS), it is based on a mathematical model of the social networks that connect survey

[4] For our purposes, the relevant wage is the city-specific mean/median, rather than the national mean/median.
[5] John Schmitt of the Center for Economic and Policy Research generously provided these estimates.

respondents. Since some individuals or groups tend to have more social connections than others, they are more likely to be recruited into a survey. To make the results of an RDS-based survey representative of the whole population (and not of just workers with large social networks), we weighted our data based on respondents' social network size—that is, based on their probability of being captured by our survey technique—as well as other features of the network which can affect the sampling process.

In addition, RDS features an important difference from snowball and other traditional chain-referral methods: it employs a dual-incentive structure. This approach involves remunerating respondents not only for the time they spend responding to the survey, but also for each eligible population member they recruit into the survey. To increase the breadth of the social network captured by the sample (and to prevent a cottage industry of survey recruitment), the number of recruitments that each respondent can make is limited through a coupon-based quota system.

Our RDS survey began with an initial set of population members to be surveyed, which we located through our contacts in each city. These "seeds" were then given a fixed number of uniquely numbered dollar-bill sized coupons to pass on to other eligible population members. These recruits then brought the coupons to one of several survey sites, where the number on the coupon was recorded, the recruit was surveyed, and then the respondent was given a fixed number of coupons with which to recruit other eligible workers.[6] This process was repeated over a period of several months, yielding large numbers of respondents in each city (see Table 1). As the recruitment progressed, the sample became increasingly diverse, eventually becoming independent of the initial sample of "seeds."

Respondent-Driven Sampling requires that easily accessible interview sites be available to all potential respondents. Interviewing workers at their worksites was not an option because workers' survey responses might have been affected by the presence of co-workers and supervisors. Because respondents must feel at ease both going to and from the site and about revealing potentially sensitive information during an interview, our study needed sites that were in safe and accessible locations and were "neutral" from the workers' point-of-view. Our team identified diverse sites such as community colleges, churches, social service agencies and community-based organizations, at which respondents were interviewed.[7]

---

[6] The number of coupons given to respondents varied over the course of the survey; on average, respondents recruited two other workers into the sample.

[7] We are extremely grateful to the organizations that allowed us to use their space to conduct interviews. Please see the Acknowledgements in our city-specific data reports for a full list.

Table 1 Summary of Survey Fielding

	Chicago	New York City	Los Angeles
Fielding period	January—June 2008	March—August 2008	April—August 2008
Number of sites	6	5	7
Number of interviewers, translators and researchers on staff	18	22	22
Monetary incentive for being surveyed	$30	$50	$30
Number of valid surveys completed	1,140	1,432	1,815

As shown in Table 1, Chicago established 6 interview sites, New York City 5, and Los Angeles 7. The number of surveys completed per day at the sites varied both between sites and over time, as it was up to respondents to choose the site at which they were interviewed. Additionally, not all sites were open for the duration of the survey period.

Because respondents could choose among multiple sites, it was imperative that each city maintain continuously updated records of who had been surveyed. Additionally, scheduling the interviews became extremely important as the number of respondents being surveyed increased (and translators' schedules filled up). These data were managed in real time using web-based databases.

The interview process proceeded as follows:

1. A potential respondent was given a recruitment coupon by a current respondent (also called a "recruiter").[8]
2. The potential respondent called a phone number printed on the coupon to schedule an interview.
    a. The scheduler recorded the respondent's unique coupon number as identification.
    b. The scheduler confirmed that the unique coupon number had not been used, notifying the potential respondent that his/her coupon was ineligible if it had been used.
    c. The scheduler asked the respondent a brief series of questions to "pre-screen" ineligible respondents.
    d. The scheduler and respondent determined an appropriate interview site and time, and the interview was scheduled.
3. The respondent arrived at the site for his/her interview.
    a. A survey team member recorded the respondent's unique coupon number.

[8] As noted above, a diverse set of initial survey members were located by project staff through city contacts.

    b. The team member confirmed in the database that the coupon number had not been previously used.
4. The respondent began the interview.
    a. The interviewer first administered a detailed module of screening questions in order to confirm the respondent's survey eligibility. If the respondent was ineligible, the survey was terminated.
    b. The interview was administered using a Computer Assisted Personal Interview (CAPI) program developed by the Survey Research Center at the University of California, Berkeley, with all responses entered by the interviewer (surveys were conducted in English, Spanish, Russian, Polish, Bengali, Hindi, Urdu, Mandarin, Cantonese, Korean, Portuguese, French and Haitian Creole).
5. The respondent completed the interview.
6. The respondent was asked to be a recruiter. If interested, the respondent was given instructions for recruiting eligible population members known to him/her into the survey, given coupons with which to recruit, and notified about the additional recruiting incentives.
7. The respondent was remunerated for his/her time taking the survey.

As noted above, respondents were given unique paper coupons to distribute to eligible population members known to them. Along with the coupons, respondents received an instruction sheet listing the survey eligibility requirements, as well as drawings and descriptions of the eligible jobs (all outreach materials were translated into multiple languages).

After recruiting population members into the sample, respondents collected remuneration for recruitment in the following fashion:

1. Respondents who had recruited called the telephone number printed on their eligibility sheet.
2. A survey team member checked the database to determine whether any of the respondent's coupons had been successfully redeemed (respondents were not remunerated for recruiting ineligible persons), and informed the respondent about how much he/she was owed.
3. On pre-specified days, recruiters could come to interview sites to collect the money owed to them.

We needed a way to effectively communicate to recruiters which workers were eligible to be recruited into the sample, so we collated the sampling frame list of industries and occupations into a communicable list. This process involved many rounds of discussion, as there is no clear-cut approach to reducing such a detailed list into a manageable number of jobs. The final list is comprised of industry descriptions (such as "maids and housekeeping cleaners"), with occupations specified within some industries. The list of eligible jobs is presented in Table 2.

Table 2 Low-Wage Industries in Each City's Sampling Universe

	Chicago	LA	NYC
Security services industry: security guards only	x	x	x
Elementary and secondary schools: teacher's assistants only	x	x	x
Private households: child care, cleaning, and home health care workers only	x	x	x
Home health care industry: home health care workers only	x	x	x
Child care industry: child care workers and teacher's assistants only	x	x	x
Restaurants and food services industries: all non-office front-line workers	x	x	x
Barber shops, beauty salons, and nail salons: front-line workers only	x	x	x
Residential construction industry: basic trades occupations	x	x	x
Landscaping services: grounds maintenance workers only	x	x	x
Service stations: attendants only	x	x	x
Dry-cleaning and laundry services: laundry and dry-cleaning workers only	x	x	x
Food manufacturing: non-office and non-sales front-line workers	x	x	x
Textile manufacturing: non-office and non-sales front-line workers	x	x	x
Furniture manufacturing: non-office and non-sales front-line workers	x	x	x
Retail industry: front-line and sales workers only	x	x	x
Banks: tellers only	x	x	x
Building services: janitors and building cleaners only*		x	x
Car washes: car wash workers only		x	x
Automotive repair and maintenance: non-office and non-sales workers only		x	x
Parking garages: attendants only		x	x
Warehousing workers		x	
Hotels: housekeepers only		x	
Couriers and messengers industry: couriers and messengers only			x

**For NYC, this category only included workers in small commercial and residential buildings*

Because some potential respondents had limited literacy levels, the printed list of eligible jobs that was given to each recruiter also included graphical depictions of each job. A sample drawing is found in Figure 1.

Figure 1 Sample Drawings Included in List of Eligible Jobs Distributed to Recruiters.

4. Adjustments in the Field

Because there was no prior information available to us regarding low-wage worker social networks, we were careful to closely track the progress of recruitment. For the duration of the fielding, we generated weekly reports on the number of surveys completed at each site, the racial and ethnic make-up of those who had been surveyed to date, the industry/occupation composition of those who had been surveyed to date, and the extent to which respondents were recruiting co-workers. All of these measures were used to detect network affiliation patterns that could generate bias in our sample.

So that we would have a rough guide of our population's characteristics, we estimated demographic and industry/occupation characteristics using the 2006 American Community Survey (which was the most recent data available in early 2008). If we detected affiliation patterns that significantly diverged from our rough population characteristics, we issued additional coupons to recruiters who had previously recruited across race/ethnicity or nationality. We implemented this strategy based on our week-by-week tracking of sampling progress.

Because the survey remuneration provided badly needed income for a population receiving low wages, we were aware of the possibility of respondents "scamming" the survey by recruiting themselves and interviewing at more than one site. The continuously updated coupon and scheduling databases allowed us to stay ahead of this potential problem, but we knew it was possible for respondents to evade our detection since we were not uniquely identifying them. However, having the same staff members rotate among interview sites largely prevented "self-recruiting" from occurring because respondents were recognized, and most scams were obvious and easy for staff to detect. The scammers we encountered were almost all caught before or during the interview (before they could be issued coupons with which to recruit).

5. The Survey Instrument

The 2008 Unregulated Work Survey is unique in that it measures a range of violations of employment and labor law, using an original battery of detailed, in-depth questions. Interviews typically lasted between 60 and 90 minutes.

The survey instrument was designed to gather information which would allow us to detect violations of laws guaranteeing the minimum wage and overtime pay; full and timely payment of wages owed; provision of legally required meal and rest breaks; protection against retaliation by employers for complaints about working conditions or attempting to organize; and access to workers' compensation in the case of an on-the-job injury. Due to time and measurement constraints, however, we were not able to measure violations of health and safety, family and medical leave, and most anti-discrimination laws, although these too are critical worker protections.

The questionnaire did *not* rely on workers having any direct knowledge about their rights under employment and labor law, or about whether they had experienced a workplace violation. Instead, our strategy was to gather raw "inputs" from workers—the necessary data about their hours, earnings and working conditions, as well as relevant employer actions. We then used these data to determine whether or not a law had been violated.

For example, we did not ask workers whether they were being paid the minimum wage. Instead, we gathered day-by-day data on exactly how many hours the respondent worked the week before the survey interview was administered, the amount of money he or she received, whether the employer made any deductions (e.g. for uniforms or meals), and whether the respondent worked off the clock. We then calculated the worker's effective hourly wage, and determined whether or not it was below the minimum wage. This approach—gathering raw data and then calculating whether a workplace violation occurred—was used for the majority of the measures that we report.

Finally, in calculating the various violation measures, we were careful never to double-count. For example, if a respondent worked five overtime hours but was not paid for those hours, we recorded an overtime violation; once these five hours were "tagged" as unpaid, they did not contribute to any other violation (for example, they could not also trigger a minimum wage violation).

Ten percent of respondents in the sample reported holding more than one job during the previous work week. For all violation measures in the previous work week, we estimated violation rates across all jobs. However, because of the survey's length, we only gathered detailed information on the primary employer (the one for whom a respondent worked the most hours during the previous week). Therefore, when we correlate job and employer characteristics with violations measured during the previous week, we are correlating violation rates measured across all jobs with the characteristics of the primary job only.

The majority of the survey involved questions about the last full week (Monday to Sunday) before the respondent was surveyed. However, we gathered information pertaining to rarer violations about the previous year or three years. Table 3 contains a detailed summary of the survey questionnaire's topics, along with the time frame to which each topic's questions referred.

Table 3 Contents of the Survey Instrument

Section	Time frame
1: Job Characteristics Number of employers last week For main employer last week: Occupation & industry of main employer Union membership Job and industry tenure Employer characteristics (e.g. size) Health insurance , vacation days and sick days Raises Use of labor market intermediaries	Last week
2: Hours Worked Last Week (for up to 5 employers) Hours worked last week Meal and rest breaks Time worked before and/or after official start time	Last week
3: Earnings Last Week (for up to 5 employers) Earnings last week (including tips, commission, bonuses) Whether earnings were pre- or post-tax withholdings Method of payment Deductions Overtime pay	Last week
4: Employment and Violations Over Last 12 Months Length of time working in city Number of employers Part-time or full-time status Weeks/months unemployed last year Instances of wage theft, late payment, bounced check, payment in goods	Past 12 months
5: Retaliation and Right to Organize Complaints to employer Employer retaliation Union organizing Verbal abuse & other types of abuse	Past 12 months
6: Workers' Compensation On-the-job injuries Employer's response to injury Workers' compensation filing Who paid for medical attention	Last 3 years
7: Knowledge of Workers' Rights Minimum wage Overtime wage Rights of undocumented workers	Current
8: Demographics Gender, age, race/ethnicity, marital status, number of children Education and training, criminal record Country of origin, immigration status, English proficiency	Current
9: Social Network Questions Estimate of the size of respondent's social networks in each industry/occupation	Current

6. Cleaning the Data

During administration of the computerized questionnaire, interview staff were able to enter comments, thoughts, or extra information by typing into an open-ended "text box" on the screen. This information was valuable because it provided additional detail about respondents' answers to survey questions and allowed qualifying information to be entered if respondents were confused by a question or wanted to give an especially complex response.

The text box data entered by interviewers turned out to be very useful in the data cleaning process. Because much of the text box data included information asked elsewhere in the survey, we used it as a "check" on the measures programmed by survey staff. Additionally, situations in which our calculations were at odds with a respondent's answers were resolved by examining the extra information entered by the interviewers.

Overall, our use of text box data affected the data for 1517 (34.6 percent) respondents from our total sample, with a median number of 2 and a mean of 3.8 variable changes per respondent (out of approximately 850 variables for the entire survey). Data for 154 respondents were invalidated and excluded from the analysis. The principal reasons for exclusion were that the respondent was shown to fall outside the study's sampling frame or was attempting to be re-interviewed after already participating in the survey.

In addition to the text box data, we also coded each respondent's industry and occupation for his/her primary job (the job at which he/she worked the most hours during the previous week). Our industry/occupation questions and coding procedures were taken directly from the 2000 Census.

7. Post-Sampling Estimation and Data Adjustments

RDS is unique in its approach to estimating population parameters from chain-referral data. It uses information collected during the sampling process to quantify features of the social network connecting respondents, and then uses these features to make inferences about population composition (Salganik and Heckathorn 2004). The two types of information are (1) each respondent's personal network size (the number of eligible population members whom he/she might recruit or be recruited by), and (2) the proportion of network ties that are between (as opposed to within) different groups in the population, as revealed by patterns of cross-group recruitment.

The first piece of information was gathered in our survey through a series of questions used to elicit how many population members of each eligible job type were known to the respondent. These questions asked respondents about people that they know and had seen in the past six months, including family, friends, coworkers, neighbors, and acquaintances. Interviewers specified that the people needed to be age 18 or older, not self-employed, and working in the county or city, and to make sure to include people who work off the books or lack immigration papers.

Respondents were then asked about how many people they knew who worked in each of the sampling frame's job categories, with periodic reminders about the eligibility criteria listed above. The job categories were randomly ordered for each respondent in order to prevent bias due to question fatigue. The approach of asking about smaller (but exhaustive) groups of population members instead of simply asking "How many population members do you know?" is known as the "scale-up method" and is believed by researchers to be more reliable than the one-shot alternative (Killworth *et al* 1998). RDS estimates depend on the relative, not on the absolute, size of respondents' networks, a factor that further increases the validity and reliability of estimates based on network sizes (see Wejnert 2009).

The second piece of information needed for RDS estimation is the proportion of network ties that are between members of different groups. Data to estimate this was collected in our survey by keeping track of who recruited whom into the survey. For example, if a male respondent recruited another male respondent into the survey, we would label it an M->M tie. Alternatively, if a male respondent recruited a female respondent into the sample, we would label it an M->F tie. By classifying all recruitments in this fashion, we can estimate the proportion of all recruitments that were M->F or F->M, which is the proportion of cross-cutting ties that we need for the RDS estimator.

Once these two pieces of information are in hand, we can produce RDS estimates of the population proportion of each group. The mechanics of this estimation procedure have undergone multiple stages of development, and are too detailed to be fully described in this document.[9] The point estimator used to produce our data report estimates is the "dual-component" estimator described in Heckathorn (2007).[10] Variance estimation for RDS proportion estimates is very complex. While there have been recent developments of analytical variance estimators (Volz and Heckathorn 2008), our data report relied on the bootstrapping procedure proposed in Heckathorn (2002) and Salganik (2006).

One feature of the RDS methodology is the ability to conduct detailed tracking of recruitment patterns throughout the entire sampling period, in order to identify and adjust for deviations from pure random recruitment from respondents' social networks. For example, recruitment might be driven by strong social identities, such as race, ethnicity or age, so that respondents recruit disproportionately within their own group.

[9] Please see Heckathorn (1997), Heckathorn (2002), Salganik and Heckathorn (2004), Salganik (2006), Heckathorn (2007), and Volz and Heckathorn (2008) for a full lineage of RDS point and variance estimation.
[10] The RDS estimation software "RDS Analysis Tool v6.1" is available at www.respondentdrivensampling.org.

The RDS methodology anticipates that personal networks are not randomly distributed, and therefore adjusts for small to moderate levels of network clustering (people having ties to others like them), in the form of post-sampling weights. For example, if the sample contained more members of a given group than would be expected under purely random sampling, then cases in that group are given less weight in analyses of the data. However, if network clustering becomes pronounced on one or more dimensions, then it is necessary to use additional, external sources of data in order to weight the final sample to be representative of the intended population.

In our study, we identified high levels of non-random recruitment among several racial/ethnic groups (the specific groups varied by city), as well as between US-born and foreign-born workers. (We did not find high levels of non-random recruitment on other dimensions, such as the workers' industry and occupation, employer, or most important, the experience of workplace violations). That meant that RDS generated representative samples within the various race/ethnic/nativity groups, but not across the sampling universe as a whole—in effect, our study generated multiple sub-samples.

To address this problem, we generated RDS violation rate estimates within each sub-sample, and then recombined these group-specific violation rates using a weighting system based on estimates of the relative sizes of the race/ethnic/nativity groups drawn from other survey data. This is an approach that has been used in another RDS study, the CDC's National HIV Behavioral Surveillance (NHBS) program (Lansky et al, 2009).

In order to implement this approach, we followed standard survey protocol by incorporating information produced by larger, less variable surveys. In this case, we drew on the race and nationality distributions provided by the 2007 American Community Survey (ACS) administered by the U.S. Census Bureau.

However, standard surveys like the ACS are known to undercount different population groups, notably unauthorized immigrants. To address this problem, we adjusted the ACS population distributions based on estimates of the number of undocumented workers in each of our cities broken down by country of origin and broad occupation category (these estimates were generously provided by Jeffrey Passel of the Pew Center for Hispanic Research). Based on those estimates, we added 10 percent of the (estimated) total number of undocumented workers in each city to the ACS population distributions, based on the undercount estimates of Hoefer, Rytina and Baker (2008).

We also compared the job-specific wage distributions contained in our survey to those of front-line workers in the ACS. This comparison revealed that the sample wage distributions had significantly higher concentrations of low-wage earners than the ACS wage distributions – which was to be expected, given that our survey sampled hard-to-reach workers (such as undocumented workers) that the ACS tends to under-sample. We adjusted our ACS population estimates by (1) estimating the proportion of the ACS industry-specific wage distribution that we captured in our sample and (2) excluding the ACS members whose wages were in the far-right tail of our sample's corresponding wage distribution. Specifically, we calculated the ratio between our sample's median wage and the ACS median wage for each job. We then took the reciprocal of this ratio and used it as the percentage cut-off above which our sample did not reach. For example, if our sample's median hourly wage for a job was $10 per hour, and the median hourly wage for the same job in the ACS was $15 per hour, we estimate that we were sampling the bottom $10/$15=2/3 of the job's wage distribution. We then excluded the top one-third of that job's wage earners from our ACS population estimates.[11]

To incorporate this supplemental information, we calculated RDS proportion and variance estimates within each of the racial/ethnic and nationality groups.[12] We then combined these group-specific point and variance estimates using the adjusted ACS population distributions described above. We combined the estimates using the following formula for the point estimates:

$$P_A = \frac{\sum_{i=1}^{n} N_i P_i}{\sum_{i=1}^{n} N_i},$$

where P_A is the aggregated proportion estimate, P_i is the proportion estimate for group i, and N_i is the number of members in population group i. We combined the variance estimates using the following formula:

$$v(P_A) = \frac{\sum_{i=1} N_i^2 s_i^2}{\left(\sum_{i=1}^{n} N_i\right)^2},$$

where $v(P_a)$ is the variance of the aggregated proportion estimate, s_i^2 is the variance of the proportion estimate for group i, and N_i is the number of members in population group i. For our survey, the i groups are the races/ethnicities/nationalities within which we estimated RDS proportions. Because our sample's composition varied by city, the number and composition of groups also varied by city.

[11] More complex adjustments than simply truncating the wage distribution were tested, but these did not yield significantly different results.
[12] Because our adjustments were based on race/nativity groups, we were unable to run estimates involving race or legal status separately by race/nativity group.

Using this approach, we minimized the effects of network segregation on our estimates while maintaining the statistical representativeness of the population that motivated the survey.

In Table 4, we summarize our estimates of the number of workers in each city that our sample represents—altogether, about 1.64 million workers, which we estimate represent roughly 31 percent of the front-line workers, and 15 percent of all workers, across the three cities.

Table 4 Low-Wage Worker Population Estimates for Surveyed Cities

	Three cities combined	Chicago	New York City	Los Angeles
Estimated number of front-line workers in low-wage industries	1,640,747	310,205	586,322	744,220
Percentage of all front-line workers	31.4	25.1	31.0	34.4
Percentage of all workers	15.1	12.2	14.2	17.0

8. Hourly Wage calculation

In order to measure pay and working-time violations in the workplace, we collected extensive data on the respondents' most recent work week. As noted above, we did not assume that respondents knew any labor laws or whether those laws had been violated. Instead, we used their detailed accounts about the time spent at work and pay received for their labor to calculate their actual hourly wage during the previous work week.

For each shift during the previous work week, respondents reported the time when it started, when it ended, time given for each meal and rest break and whether those breaks were paid or unpaid. For the entire week, they also reported the amount of time worked off the clock (not during official shift hours) and whether or not they were paid for that time. We used this information to calculate the exact number of hours for which the worker was legally entitled to be compensated. The total hours worked during the week is the sum of all shift hours, paid meal and rest breaks, and paid off-the-clock time (unpaid off-the-clock time is a violation which we report separately from minimum wage violations, so it does not contribute to our overall wage calculation).

We then calculated the respondents' pay for the previous work week. We collected information on the type of pay arrangements respondents had with their employers (by hour, day, week, job, piece, etc.) and the frequency with which they received payments (daily, weekly, bi-weekly, monthly). We then asked about respondents' most recent pay amount or expected pay amount, which would include compensation for all work performed during the week. Taking into account pay arrangement and payment frequency, we calculated the amount of money that respondents received or were expecting to receive for that week. For example, a bi-weekly paycheck would be divided by number of hours worked during the two-week pay period, or a piece worker's wage would be calculated as number of pieces completed in the previous work week multiplied by the pay rate for each piece.

We made a number of additional adjustments to the hourly wage calculation based on pay-related information that respondents provided.

9. Tax Estimation

We asked respondents whether the pay they reported was pre- or post-tax, and 23% indicated that at least one part of their wage was post-tax. Since employment and labor law is concerned with pre-tax pay, we had to convert post-tax into pre-tax pay by estimating the amount of taxes each respondent paid. To calculate the likely tax rate, we used the same approach as the IRS: first, we estimated respondents' yearly income by multiplying their weekly income by 51. We then used each respondent's reported number of children, marital status, and yearly income to place the respondent into an appropriate tax bracket, thus yielding a tax withholding rate. We included FICA withholdings for all workers, regardless of their household size or structure. We then adjusted respondents' post-tax income to reflect our estimate of their pre-tax income.

For single workers with or without children, this estimation procedure was fairly straightforward: they were treated as single filers, or heads of household (with deductions for children), respectively. However, the estimation procedure became more complex for married respondents with children because we were forced to make assumptions about how these respondents would have divided their children among their W-4 withholding declarations.

In making these assumptions, we relied on discussions with experts in how low-income individuals or households typically file their taxes. For each assumption, we chose the most "conservative" option available. For example, there is little research documenting how spouses typically divide child deductions between their withholding rates. Instead of assuming that each spouse claimed a deduction for each child (which would mean the family claimed two deductions for each child if both parents work), we divided the number of children between spouses.[13]

Conservative choices like this mean that any error induced by our assumptions resulted in an *overestimate* of a respondent's tax rate, and a corresponding adjustment to weekly wages that was too large – resulting in an *under-estimate* of minimum wage violations.

However, this strategy had the opposite effect on overtime violation rates for a very small number of respondents who reported a specific dollar amount of overtime pay (as compared to reporting "time-and-a-half" for overtime pay). Because these respondents' base hourly wage

[13] On IRS form W-4, individuals may legally claim as many children for deductions as they would like because they will end up paying the correct amount of taxes in April no matter how much money was or was not deducted from each paycheck. Our goal in tax estimation was to replicate the approach taken by most low-wage tax filers.

was inflated by the above tax estimation, it would appear that their overtime pay was less than time-and-a-half even though it may actually have been time-and-a-half. Given that only 3 percent of respondents who worked overtime hours reported a specific dollar overtime pay amount and reported post-tax earnings, our calculation was that this exceedingly minimal inflation of overtime violation rates was justified by the guarantee of a conservative minimum wage violation rate.

We obtained tax rate schedules from the Internal Revenue Service, the New York State Department of Taxation and Finance, New York City Department of Finance, the State of California Franchise Tax Board, and the Illinois Department of Revenue.

10. Deduction adjustment

Workers also reported the amount of money employers deducted from their paychecks (if any). If the deductions were legal – for the benefit of the employee – we added the amount back into the calculated wage.[14]

11. Overtime adjustment

Based on state laws and the total number of hours a respondent worked in the previous work week, we were able to determine the type and the number of overtime hours that a respondent worked in the week. We used each state's overtime criteria to determine which hours should be counted as overtime and the legally required rate of pay for those hours. The most common type of overtime is an over-40-hour week, which requires that a worker receive one and one-half times his/her regular hourly wage (time-and-a-half) for all hours worked beyond 40 in a given week.

Given the complexity of the legal overtime requirements, we were careful not to double-count or undercount any overtime that respondents should have received. Because both California and New York have laws governing both weekly and daily overtime, and any given work hour can only contribute to one type of overtime, we assigned each respondent the overtime combination which would result in a maximum amount of pay. This approach is the standard way of calculating legally required overtime pay when estimating lost wages in wage theft cases.

If a respondent worked overtime hours in the previous work week, he/she was asked about the amount of pay he/she received for those hours. By subtracting overtime hours from total hours and overtime pay from total pay, we separated overtime pay violations from minimum wage and other pay violations.

[14] See section 13g below for a list of illegal deductions.

12. Overtime imputation

During the data cleaning phase we noticed that an unusually large number of hourly workers (N = 207) reported not receiving any pay at all for their overtime hours. Since such egregious overtime violations are unlikely to be this common among hourly workers (who are paid by the hour, with a standard set rate, often with a paycheck), we hypothesized that they might have misunderstood the question. Fortunately, since we asked hourly workers for both their agreed-upon hourly wage and their total weekly earnings, we had the means to impute their likely overtime rate.

To understand how each respondent might have interpreted the question, we calculated three hypothetical overtime pay scenarios based on their stated hourly wage. The three hypothetical scenarios are that a worker was paid:

1. Legal overtime pay (time-and-a-half or double-time where required)
2. Straight overtime pay (the respondent's normal hourly wage)
3. No overtime pay (no pay at all)

First, we calculated what each respondent's hypothetical weekly earnings would have been if he/she had been paid according to each pay scenario (i.e., time-and-a-half, straight pay, or no pay at all). We then compared these hypothetical weekly earnings to the actual weekly earnings reported by the workers. Specifically, we divided each hypothetical earnings amount by actual weekly earnings, thus producing three ratios per respondent. These ratios allowed us to see which hypothetical value was closest to the actual pay and its scaled distance from it.

Using these ratios, we then assigned each respondent to a category indicating the likely overtime pay scenario – choosing whichever hypothetical scenario produced a weekly earnings amount closest to respondent's actual weekly earnings.

Based on the procedure, cases were assigned to one of the three categories: Full overtime (11%), Some overtime (24%), or No overtime (65%). That is, 35 percent of cases were reassigned from the "No Overtime" category to either "Full" or "Some" overtime; the rest remained unchanged.

13. Legal Criteria

This section contains detailed descriptions of the legal regulations we used to determine whether a workplace violation occurred or not. Staff lawyers at the National Employment Law Project and legal services lawyers in all three states compiled the workplace laws in effect during the fielding period for each of the three cities by examining the relevant legal statutes. In general, state and city-specific laws supersede federal laws if they set higher standards, so we compiled laws at both levels in order to determine which was currently enforced in each city.

13a. Minimum Wage

Minimum Wage Law Criteria

Following the wage adjustments described above, each respondent's calculated hourly wage was compared to each state's minimum wage threshold. For tipped workers, the total hourly wage included both a base hourly wage and an hourly tip rate (except for California where the base wage must equal at least the full minimum wage, regardless of tips earned).

State minimum wage rates at the time of the survey

- Illinois - $7.50
- California - $8.00
- New York - $7.15

Tipped Minimum Wage Law Criteria

Tipped workers' pay was also checked to determine whether it passes each state's minimum base wage criteria (except for California, where the minimum wage does not differ for tipped workers). The following are the specific criteria we used to determine a tipped minimum wage violation:

New York

- Minimum hourly base wage for hotel food service workers, identified as working within *Traveler accommodation* industry and in *Food preparation and serving* occupations:
    - $4.60
- Minimum hourly base wage for food service workers in restaurant industry, identified as working within *Restaurants and other food services* or within *Drinking places, alcohol beverages* industries and in *Food preparation and serving* occupations BUT NOT working as a *First-line supervisors/managers of food preparation and serving worker*:
    - $ 4.60
- Minimum hourly base wage for non-food service workers in restaurant industry, identified as working within *Restaurants and other food services* or *Drinking places, alcohol beverages* industries BUT NOT in *Food preparation and serving* occupations:
    - $4.85, when hourly tip rate is >= $2.30
    - $5.55, when hourly tip rate is >= $1.60 and < $2.30
    - $7.15, when hourly tip rate is < $1.60
- Minimum hourly base wage for all other industries and occupations:
    - $5.40, when hourly tip rate is >= $1.75
    - $6.05, when hourly tip rate is >= $1.10 and <$ 1.75
    - $ 7.15, when hourly tip rate is < $1.10

Illinois

- Minimum base wage threshold is 60% of regular minimum wage:
    - $4.50

Minimum Wage Law Exemptions

New York

- Exempt from state minimum wage and tip credit laws but covered by federal law (the Fair Labor Standards Act, or FLSA) laws are public sector workers – those employed by city, state or federal government.
    - Their minimum wage threshold is $5.85 prior to July 24, 2008 and $6.55 after July 24, 2008.

- Exempt from both state and federal minimum wage laws are live-in home health care workers not employed by home health care industry, identified as living in client's home and working as *Personal and Home Care Aide* or as *[all other] Personal care and service worker* or as *Nursing, Psychiatric, and Home Health Aide* BUT NOT within *Health Care* industry.

Illinois

- Exempt from state minimum wage and tip credit laws but covered by FLSA minimum wage are live-in AND live-out domestic workers, identified as working as *Child Care Worker* or as *Maids and Housekeeping Cleaner* and within *Private household* industry.
    - Their minimum wage threshold is $5.85 prior to July 24, 2008 and $6.55 after July 24, 2008.

- Exempt from state minimum wage and FLSA minimum wage are home health care workers (*Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*) who are employed by a private household (*Private household* industry).

- Exempt from state minimum wage, but covered by FLSA, are workers who work for an employer with less than 4 employees.

13b. Overtime

Based on reported hours, we determined which overtime protection workers were eligible for and we made sure overtime hours were not double-counted. A worker could not have both daily and weekly (over 40 hours per week) overtime, because those overtime hours may only contribute to either one or the other type of overtime.

Overtime Law Criteria

All three states

- Pay for work beyond 40 hours in a week must be compensated with 1-½ times the normal hourly wage

California

- Pay for work between 8 and 12 hours in a day must be compensated with 1 ½ times normal hourly wage (only when workweek is less than 40 hours, otherwise all hours are counted as either weekly or daily overtime hours, depending on whichever is bigger)
- Pay for work beyond 12 hours in a day must be compensated with 2 times normal hourly wage
- Pay for work for the first 8 hours on the 7th consecutive day of work must be compensated with 1 ½ times normal hourly wage (again to the exclusion of possible weekly overtime hours)
- Pay for work beyond 8 hours on the 7th consecutive day of work must be compensated with 2 times normal hourly wage

New York

- An additional payment (beyond normal wages) of 1 hour of minimum wage ($7.15) for each shift lasting over 10 hours[15]

The calculation of how much a worker is owed for overtime hours is different for tipped workers. Their extra overtime pay (i.e. the 'half' in time-and-a-half pay or the 'one' of double pay) is multiplied by the state's minimum wage, rather than their base hourly wage.

---

[15] Courts have differed in their interpretations of the state's daily overtime requirement. Some courts, and the New York State Department of Labor, have determined that the daily overtime requirement mainly applies to workers who are paid the minimum wage – "phasing out" this requirement for workers paid more than minimum wage. *See, e.g.*, *Espinosa v. Delgado Travel Agency*, 2007 U.S. Dist. LEXIS 44844 (E.D.N.Y. 2007). Other courts have found that the daily overtime requirement provides all workers – including those who earn more than minimum wage – an extra hour of pay at the minimum wage rate. *See, e.g.*, *Yang v. ACBL Corp.*, 2005 U.S. Dist. LEXIS 31567, at *11 n.10 (S.D.N.Y. 2005). We have adopted the latter interpretation, which is more generous for low-wage workers. The state Department of Labor is expected to revise the daily overtime requirement for hotel and restaurant workers in the near future, adopting the more generous interpretation. *See* Order of Commissioner of Labor M. Patricia Smith on the Report and Recommendations of the 2009 Wage Board at p.3 (2009), http://www.labor.state.ny.us/workerprotection/laborstandards/PDFs/Order_of_Commissioner_in_response_to_Wage_Board_11-5-09.pdf.

Overtime Law Exemptions

New York

- "Weaker" overtime for live-in domestic workers, identified as living in employer's home and working as *Child Care Worker* or as *Maids and Housekeeping Cleaner* and within *Private household* industry:
    - Overtime pay must be no less than 1-½ times the state minimum wage for work beyond 44 hours in a workweek

- "Weaker" overtime for home health care workers
    - For live-out home health care workers (*Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*) who are employed by an agency or by a private household (*Private household* industry or within *Home health care services* industry), overtime pay must be 1.5x the state minimum wage over 40 hours a week.

    - For live-in home health workers (*Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*) who are employed by an agency (*Home health care services* industry), overtime pay must be no less than 1-½ times the state minimum wage for work beyond 44 hours in a workweek

- Exempt from New York State's daily overtime requirement ($7.15 for days longer than 10 hours) are public sector workers, but they still qualify for over 40-hour workweek overtime under FLSA.

- Exempt from federal or state overtime are live-in home health care workers, identified as living in client's home and working as *Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker* and within *Private household* industry.

California

- Exempt from California's daily or 7th day overtime requirement are public sector janitors – those employed by city, state or federal government and working as *Janitor and*

- Exempt from California's daily or 7th day overtime requirement are live-out nannies, identified as working as *Child Care Worker* and within *Private household* industry, but they still qualify for over 40-hour workweek overtime under FLSA.

- Exempt from federal or state overtime are live-in private household child care workers, identified as working as *Child Care Worker* in the *private household*.

- Exempt from state overtime, but covered by federal overtime, are live-out private household child care workers, identified as working as *Child Care Worker* in the *private household*.

- Exempt from federal or state overtime are private household/agency home health care workers, identified as working as *Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*.

Illinois

- Exempt from federal or state overtime are live-in domestic workers, identified as living in client's home and working as *Child Care Worker* or as *Maid and Housekeeping Cleaner* and within *Private household* industry.

- Exempt from state OT, but covered by FLSA are live-out domestic workers, identified as working as *Child Care Worker* or as *Maid and Housekeeping Cleaner* and within *Private household* industry.

- Exempt from federal and state OT are home health care workers (*Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*) who are employed by a private household (*Private household* industry or within *Home health care services* industry).

- Exempt from state OT, but covered by FLSA, are workers who work for an employer with less than 4 employees.

13c. Break Violations

Based on reported shift times, we were able to determine the number and duration of meal or rest breaks a worker should have received under law. Meal breaks may be paid or unpaid, but rest breaks (required only in California) must be paid, otherwise they do not qualify as bona fide rest breaks.

Break Law Criteria

New York

- For every 6 hour shift, a 30 minute meal break is required.
- For every 10 hour shift, where shift starts between 1PM and 6AM, a 45 minute meal break is required (workers get 45 minutes for dinner during evening and night shifts).
- For every 10 hour shift, where shift starts before 11AM and ends after 7PM, a 75 minute meal break is required (workers get 30 minutes for lunch plus 45 minutes for dinner during such shifts).
- For every 8 hour shift, a 60 minute meal break is required for factory workers only, identified as working within *Manufacturing* industry.

Illinois

- For every 7.5 hour shift, a 20 minute meal break is required.
- For every 7 hour shift, a 30 minute meal break is required for hotel attendants only, identified as working within *Traveler accommodation* or within *Recreational vehicle parks and camps, and rooming and boarding houses* industry.

California

- For every 5 hour shift, a 30 minute meal break is required.
- For every 4 hour shift, a 10 minute rest break is required.

Break Law Exemptions

California

- No meal or rest breaks are required for public sector janitors – those employed by city, state or federal government and working as *Janitor and Building Cleaner* or as *Baggage Porter, Bellhop, and Concierge*.
- No meal or rest breaks are required for nannies and home health care workers, identified as working as *Child Care Worker* or as *Nursing, Psychiatric, and Home Health Aide* or as *Personal and Home Care Aide* or as *[all other] Personal care and service worker*.

13d. Workers' Compensation

Respondents were asked to report any serious work-related injury (one requiring medical attention) they experienced during the last three years. We then asked a series of detailed questions about their and their employer's subsequent actions *for the most recent injury only*.

We identified a Workers Compensation violation based on employers' reactions to employees' injuries. Following is the list of employer actions considered illegal under state law.

Worker's Compensation Law Criteria

All three states

- Employer threatened to fire or deport worker if Workers Compensation claim was filed.
- Employer knew about worker's injury but told worker not to file for Workers Compensation.
- Employer fired worker shortly after injury.[16]
- Worker paid out-of-pocket all or part of medical bills after Workers Compensation claim was accepted.

California

- Worker informed employer about the injury but employer did not tell worker to file for Workers Compensation.
- Employer refused to help worker with the injury, saying that someone else is to blame.

13e. Retaliation

If a respondent reported making a complaint to his or her employer or supervisor about working conditions in the past twelve months, we asked about how and why he or she complained and about the employer's reaction *for the most recent complaint only*. We also asked respondents whether they tried to form a union and their employer's reaction. Following the criteria we used to determine whether an employer's reaction to either of these actions would be considered illegal. States do not differ in how they determine illegal retaliation.

Retaliation Law Criteria

Workers who complained <u>alone</u> are at risk for illegal retaliation only if they complained for one of the following reasons:

- Paid below minimum wage
- Not paid for all hours worked
- Forced to work off the clock
- Not paid for overtime
- Not paid on time
- Improper deductions from paycheck
- Dangerous working conditions

[16] Depending on the state in question, firing an injured worker is not necessarily illegal, in and of itself. But retaliating against an injured worker for filing a workers' compensation claim, or attempting to file one, is illegal. In analyzing this series of questions, we interpreted respondents' reports of being fired shortly after an injury as retaliatory acts, based on both the wording of the questions and on previous research and fieldwork (see, for example, Bernhardt, McGrath and DeFilippis (2007)).

- Discrimination
- No breaks, or not enough breaks
- Paid less than owed

Workers who complained about working conditions as a group are at risk for illegal retaliation no matter the specific reason for the complaint.

The following employer responses to either individual or group complaints are illegal:

- Employer threatened to fire worker or co-workers
- Employer threatened to call immigration
- Employer fired worker or co-workers
- Employer suspended worker or co-workers
- Employer cut back worker's or co-workers' hours
- Employer cut worker's or co-workers' pay
- Employer gave worker or co-workers worse work assignment
- Employer abused or harassed worker or co-workers
- Employer denied work tools
- Employer increased workload

The following employer responses to union organizing efforts are illegal:

- Employer threatened to fire worker or co-workers
- Employer threatened to call immigration
- Employer fired worker or co-workers
- Employer suspended worker or co-workers
- Employer cut back worker's or co-workers' hours
- Employer cut worker's or co-workers' pay
- Employer gave worker or co-workers worse work assignment
- Employer abused or harassed worker or co-workers
- Employer tried to discourage workers
- Employer increased workload

Retaliation Law Exemptions

- If a worker is exempt from any minimum wage or overtime violation (see above), then the worker is also exempt from retaliation violations when making a complaint individually.
- Domestic workers are not covered by right to organize laws – those not employed by city, state or federal government and not employed by a health/domestic care agency and working in a *Private Household* as a *Child Care Worker* or as a *Nursing, Psychiatric, and Home Health Aide* or as a *Personal and Home Care Aide* or as a *Maid and Housekeeping Cleaner.*

13f. Verbal abuse

We asked respondents if they had experienced regular and repeated verbal abuse from an employer or supervisor during the last 12 months. Verbal abuse on the following bases is illegal in all three states:

- Race, Religion, Gender, Sexual Orientation, National Origin, Age, and Disability

13g. Deductions

We asked respondents if they had or expect to have any money deducted from their pay for last week. If so, we inquired about the reasons for deductions.

Deduction Law Criteria

Following are reasons for which it is illegal to deduct any money from employee pay:

- Damage to or loss of work-related tools or materials
- Work-related tools, materials or transportation
- Miscellaneous unlawful reasons (e.g. punishment)
- Uniforms
- Any reason that is not communicated to the employee

Deduction Law Exemptions

California

- Workers whose uniforms could be used as regular clothes, such as nurse uniforms. Healthcare workers are exempt if uniform costs were deducted, identified as *Health Care*.
- Californians who earn more than double the minimum wage may be required to pay for their own tools and tool accessories.

13h. Complete exemptions

These workers are likely considered independent contractors under the law, so we have made them exempt from all violations reported in this study:

- Own-home childcare worker, identified as working as a *Child Care Worker* within the *Child day care services* industry, and not working in a day care center.

13i. Other reported pay violations

Working off-the-clock

- Violation counted if a worker reported working before the beginning and/or after the end of a shift and was not paid at all for that time.

Paid in tips only

- Violation counted if a worker reported receiving tips but received no other wage from the employer.

Tip theft

- Violation counted if a worker reported that an employer or supervisor collected all or any portion of the tips given to workers at the establishment.

Late pay

- Violation counted if a worker reported that an employer was late in paying the wages (later than the regular or agreed-upon day).

No pay stub

- Violation counted if a worker reported that he or she did not receive a document listing pay and deductions.

14. Bibliography

Bernhardt, Annette, Ruth Milkman, Nik Theodore, Douglas Heckathorn, Mirabai Auer, James DeFilippis, Ana Luz Gonzalez, Victor Narro, Jason Perelshteyn, Diana Polson, and Michael Spiller. 2009. *Broken Laws, Unprotected Workers: Violations of Employment and Labor Laws in America's Cities.* Found at: http://www.unprotectedworkers.org.

Bernhardt, Annette, Siobhán McGrath and James DeFilippis. 2007. *Unregulated Work in the Global City: Employment and Labor Law Violations in New York City.* New York: Brennan Center for Justice at New York University School of Law.

Carré, Francoise and Randall Wilson. 2004. *The Social and Economic Costs of Employee Misclassification in Construction*. Boston: Construction Policy Research Center, Harvard Law School Labor and Worklife Program and Harvard School of Public Health.

DeSilva, Lalith, Adrian Millett, Dominic Rotondi and William Sullivan. 2000. *Independent Contractors: Prevalence and Implications for Unemployment Insurance Programs*. Report prepared for the United States Department of Labor. Rockville, MD: Planmatics. Found at: http://wdr.doleta.gov/owsdrr/00-5/00-5.pdf.

Freeman, Richard B. and Ronald Schettkat. 2000. "Low Wage Services: Interpreting the US-German Difference," National Bureau of Economic Research Working Paper No. 7611, 2000. Found at: http://www.nber.org/papers/w7611.

Heckathorn, Douglas D. 1997. “Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations,” *Social Problems* 44(2):174-199.

—. 2002. “Respondent-Driven Sampling II: Deriving Valid Population Estimates from Chain-Referral Samples of Hidden Populations.” *Social Problems* 49(1):11-34.

—. 2007. “Extensions of Respondent-Driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment,” *Sociological Methodology* 37(1): 151-207.

Killworth, P.D.; C. McCarty; H.R. Bernard; G.A. Shelley, and E.C. Johnsen. 1998. “Estimation of seroprevalence, rape and homelessness in the U.S. using a social network approach,” *Evaluation Review* 22: 289–308.

Lansky, A; A Drake; and HT Pham. 2009. “HIV-Associated Behaviors among Injecting-Drug Users--23 Cities, United States, May 2005-February 2006.” Morbidity and Mortality Weekly Report, April 10, 2009 / 58(13);329-332.

Organisation for Economic Co-operation and Development. 1994. *The OECD Jobs Study: Evidence and Explanations*. Paris: Organisation for Economic Cooperation and Development.

—. 1996. *OECD Employment Outlook: July 1996*. Paris: Organisation for Economic Co-operation and Development.

Salganik, Matthew J. 2006. "Variance estimation, design effects, and sample size calculations for respondent-driven sampling," *Journal of Urban Health*, 83:98-111.

Salganik, Matthew J. and Douglas D. Heckathorn. 2004. "Sampling and estimation in hidden populations using respondent-driven sampling," Sociological Methodology, 34:193-239.

United States Department of Labor. 2001. *1999-2000 Report on Low-Wage Initiatives*. Washington, D.C.: United States Department of Labor Employment Standards Administration Wage and Hour Division.

United States General Accountability Office. 2006. *Employment Arrangements: Improved Outreach Could Help Ensure Proper Worker Classification*. Report to the Ranking Minority Member, Committee on Health, Education, Labor, and Pensions, U.S. Senate. Washington, D.C.: United States General Accountability Office.

Volz, Erik and Douglas D. Heckathorn. 2008. “Probability Based Estimation Theory for Respondent Driven Sampling,” Journal of Official Statistics, 24(1): 79-97.

Wejnert, Cyprian. 2009. “An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and Out-of-Equilibrium Data,” Sociological Methodology, 39(1):73-116.