



Methodology Wave 3

Summer 2018 | City of Detroit

Methodology Wave 3

I. SUMMARY

Wave 3 of the Detroit Metropolitan Area Communities Study includes interviews with a representative sample of 1,216 Detroit residents. In this mixed-mode study, interviews were self-administered online and interviewer-administered in-person and on the telephone in English, Spanish, and Arabic between June 25 and September 20, 2018. One portion of respondents (n=260) were recruited from the list of 714 adults who responded to the first wave of the DMACS survey; the remainder (956) were recruited from a new randomly-selected address-based sample of Detroit households, including an oversample of households in majority-Latino U.S Census block groups. The data have been weighted in a three-step process: first, design weights were created for the new sample to compensate for the differential probability of selection in oversampled communities; second, non-response adjustments were created separately for the Wave 1 respondents and the new sample; finally, we used a raking procedure to match the total universe of respondents to the population of the City of Detroit according to U.S. Census estimates for gender, age, Hispanic ethnicity, race, and education. The margin of sampling error for a random sample survey of this size would be +/- 2.8 percentage points at the 95% confidence level; the actual margin of sampling error varies by statistic due in part to the complex sample design, and of course any survey is subject to other sources of error as well as random sampling error.

II. SAMPLE DESIGN

The target population for this study was the adult household population of the City of Detroit. The sample had two components:

- The 714 respondents to the first wave of the DMACS survey: The initial sample in this case was a simple address-based sample. In 2016 a simple random sample of addresses was drawn from a list of all household addresses in the City. A total of 3100 addresses were selected from FIPS 26/22000, City of Detroit. The sample file contained information including the tract number, block and block group, and whether the address was a single or multiple family structure. The sample provider also matched each address to the likely name and phone number of a resident when such information was available. Approximately 90% of the sample lines came with a possible name match and 68% with a possible telephone number match. All individuals who responded to the Wave 1 survey were invited to participate in the Wave 3 survey.
- A new stratified two-stage cluster sample of household addresses: For the first stage of sampling, U.S. block groups were selected with probability proportionate to estimated size within two strata. The first stratum included block groups where 70% or more of the population is Hispanic/Latino according to ACS estimates;

eight block groups were selected within this stratum. The second stratum included all other block groups within the City of Detroit; 41 block groups were selected within this stratum. The second stage of sampling consisted of selecting household addresses from a list of all households in the block group. A total of 3,852 addresses were selected at this stage.

III. QUESTIONNAIRE DEVELOPMENT, PROGRAMMING, AND TESTING

The questionnaire was developed by the DMACS Principal Investigators in partnership with the Detroit Health Department. The instrument was programmed and administered in Qualtrics, and thoroughly tested internally to ensure accuracy of question language, skip patterns, etc. The questionnaire including programming instructions was downloaded from Qualtrics, translated into Spanish and Arabic by Cetra Language Solutions and re-uploaded to Qualtrics in those languages. Respondents were able to select their language on the first screen of the survey.

IV. DATA COLLECTION

The data collection was multi-mode in recruitment and administration. All respondents who completed the survey received a \$10 post-paid incentive, mailed by check from the University of Michigan.

Mail recruitment: An initial postcard was sent to each of the sample addresses in June 2018 introducing the survey and asking respondents to look for a letter that would have their unique survey URL. A second mailing sent in July contained an invitation letter, an informed consent document with a comprehensive introduction to the project, and an individualized web address (URL) for respondents to reach the survey instrument. A third mailing contained a follow-up letter with the individualized URL.

E-mail recruitment: Existing panel respondents were contacted by e-mail, with the message including a clickable individualized URL.

Face-to-face outreach: A team of canvassers was trained on June 15, 2018. Through October 15, 2018, this group conducted face-to-face outreach and interviewing. For existing panel members, canvassers were instructed to speak only to that individual. For the new sample, canvassers asked for the youngest adult member of the household. If target individual was home, the canvasser asked them to complete the survey on the tablet at that time. They were encouraged to complete the survey themselves, but if they needed help due to vision/literacy/any other issue, the canvassers were trained to administer the survey.

If the target individual could not take the survey immediately, or if no one was home, the canvassers left behind a paper with the individualized survey link so they could take it on their own time.

Telephone outreach: Existing panelists who had provided their phone numbers were also contacted by phone. Trained interviewers provided the individualized survey link and administered the survey by phone. Interviewers made up to two contact attempts for each phone number.

V. WEIGHTING

Statistical weighting to control for the impacts of the sample design and non-response was performed in three stages:

- Design weight: A design weight of 1 was assigned to existing panel members, who were drawn from a simple random sample of the City. For newly-sampled households, respondents were assigned a design weight that was equal to the inverse of the selection probability (the number of households in the city as estimated by the Census, divided by the number of sampled households in the block group). This was then divided by a constant to adjust the scale of the weights to a mean of 1.
- Non-response weight: Non-response weights were calculated separately for the returning panelists and the new Wave 3 sample members. For the returning panelists, steps in generating this weight included:
 - Factor analysis on 13 block-group variables from the 2011-2015 American Community Survey to reduce the number of potential predictors;
 - Multiple imputation by chained equations to impute 25 datasets with complete W1 data for all respondents;
 - Examination of the bivariate relationships between Wave 3 response and potential predictors, including ACS data, Wave 1 responses, and paradata from Waves 1 and 3;
 - Running a response propensity model on all 25 imputed datasets. This model was an unweighted logit model using limited set of predictors (those where $p < .1$ in the bivariate relationship to W3 response).
 - Smoothing the weights generated by creating quintile groupings of the inverse of the predicted probability of response

For the new Wave 3 sample members, the process was very similar; though the potential predictors available were fewer, and without prior data, the multiple imputation phase was not necessary. In addition,

- Factor analysis was conducted on 15 block-group variables from the 2012-2016 American Community Survey rather than the 13 from the earlier wave of the ACS that were attached to Wave 1;
- The selected response propensity model was a weighted logit model that included ACS factor scores as predictors and random effects for block group. This model was selected because as selected as the preferred model because (a) it includes the design weights for selecting the new wave 3 sample, (b) the random effects for block groups (our primary sampling units) were significant, and (c) it produced the least amount of variance in predicted probabilities of the potential models tested.
- **Post-stratification weight:** after multiplying the design weight by the non-response weight, an additional post-stratification weight was developed to calibrate the demographic distribution of respondents to the target population of the City of Detroit. We first used multiple imputation to create ten datasets that were complete for all respondents for the variables used in raking. In order to preserve the correlations between these variables and other survey outcomes, a larger set of variables was imputed, including income, length of residence at current address and length of residence in Detroit, number of places R has lived in last five years, home ownership, whether R ever experienced homelessness, marital status, internet access at home, neighborhood satisfaction, views on community assets (Q6a-k), views on priorities to improve public health (Q7a-q), fear of crime, support from social networks, whether Rs neighborhood has name, primary source of health care, insurance status, affiliation with community associations, ability to pay for current care or health emergency, attendance of religious services, and political ideology. The predictors for these imputation models included ACS factor scores (see step 3a) and other wave 3 variables for which there were no missing data. This weight was developed with an iterative proportional fitting (raking)¹ procedure (using the “ipfraking” package in Stata 15)² and includes adjustments for age, gender, race, Hispanic ethnicity, and education to match the American Community Survey (ACS) 2012-2016 estimates for the population 18 and older in the City of Detroit. Weights were trimmed to a maximum value of 4.

VI. DISPOSITION AND RESPONSE RATE

The response rate to the Wave 3 survey was 32.2%, calculated using AAPOR Response Rate 3. The final dispositions are reported in Table 1.

Table 1: Dispositions and Response Rate

¹ For more information, see “Raking” in Lavrakas, P. J. (2008). Encyclopedia of survey research methods Thousand Oaks, CA: SAGE Publications Ltd.

² For more information on the “ipfraking” package, see Kolenikov, S. (2014). “Calibrating survey data using iterative proportional fitting (raking).” The Stata Journal 14, pp. 22-59.

Category	Returning panelists	New W3 sample	Total
Interview (Category 1)			
Complete	260	956	1216
Partial	4	23	27
Eligible, non-interview (Category 2)			
Refusal and break off	46	649	695
Non-contact	410	1135	1545
Other	0	32	32
Unknown eligibility, non-interview (Category 3)			
Unknown if housing unit/unknown if occupied	0	263	263
Not eligible (Category 4)			
Vacant housing unit	0	71	71
Non-residence	0	3	3
AAPOR Response Rate 3:			
$(I)/((I+P) + (R+NC+O) + e(UH+UO))$	36.1%	31.3%	32.2%