

Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes

Measuring Human Trafficking Prevalence in Construction: A Field Test of Multiple Estimation Methods

OMB Information Collection Request
New Collection

Supporting Statement Part B

MARCH 2022

Submitted By:
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

4th Floor, Mary E. Switzer Building
330 C Street, SW
Washington, D.C. 20201

Project Officer:
Mary G. Mueggenborg
Kelly Jedd McKenzie

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

Part B

B1. Objectives

Study Objectives

The Office of Planning, Research, and Evaluation (OPRE), in collaboration with the Office on Trafficking in Persons (OTIP), within the Administration of Children and Families (ACF) at the U.S. Department of Health and Human Services (HHS) seeks Office of Management and Budget (OMB) approval for a one-time survey of construction workers in one U.S. geographic location. The goal of this study is to advance knowledge of promising methods for estimating human trafficking prevalence. Study objectives include to: (1) estimate the prevalence of labor trafficking among construction workers using two estimation strategies; (2) describe the nature and type of exploitation experienced by construction workers; and (3) explore potential risk and protective factors associated with risk of trafficking victimization. The survey will be administered to construction workers in Houston, Texas. Data will be used to inform future prevalence estimation research as well as policy, programs, and practices intended to prevent, identify, and respond to labor trafficking in the construction industry.

Generalizability of Results

This study is intended to present two internally valid prevalence estimates of labor trafficking victimization in the construction industry in one U.S. geographic location, not to promote statistical generalization to other sites or industries. Despite not being generalizable, the information will be helpful to inform future prevalence estimation research as well as policy, programs, and practices intended to prevent, identify, and respond to labor trafficking in the construction industry.

Appropriateness of Study Design and Methods for Planned Uses

The study design selected is intended to identify a representative sample of construction workers, which is important for estimating prevalence. Surveying construction workers about their experiences in the industry will allow us to assess whether, and what types, of labor exploitation or trafficking they have experienced as well explore potential risk and protective factors. The survey sample will be representative of workers in only the construction industry in Houston. Study findings should not be extrapolated to other industries or geographic locations. This context, and the limits to representativeness, will be clearly documented in the text of all written materials associated with this study.

As noted in Supporting Statement A, this information is not intended to be used as the principal basis for public policy decisions and is not expected to meet the threshold of influential or highly influential scientific information.

B2. Methods and Design

Target Population

We plan to recruit current workers from construction sites and current/former workers through workers' social networks. The research team will use probability (i.e., time-location sampling) and network referral sampling to identify a representative sample of workers to participate in the survey.

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

We will survey up to 4,200 individuals who have worked in construction in the past 24 months in the study site.

Sampling and Site Selection

Construction Site Selection. A systematic approach will be used to select construction sites to include in the time-location sample. A list of current, permitted construction sites will be requested via a Texas Public Information Act request. A random selection process will be used to select permitted construction sites for inclusion in the sample. Because this approach excludes unpermitted construction sites, the initial sample will be supplemented with any unpermitted construction sites identified within the same Census block as a selected permitted site. The field team will have tablets with a mapping application that shows all of the selected permitted sites overlaid with Census block boundaries. The field team will physically walk the Census block to identify additional construction sites. The mapping application will allow the field team to mark additional sites and add them to the sample.

There is no site selection for the network sample.

Construction Worker Sample Selection. Two approaches will be used to sample individuals who have worked in construction in the past 24 months for this study: probability sample (time-location) and network sample.

A **time location sampling** (TLS) design will be used to recruit current workers from construction sites. Potential time intervals for recruitment will be added to the list of construction sites. The final result will comprise the sampling frame. Upon selecting a set of construction sites and times for recruitment, a systematic sampling design will be used to invite a set of individuals from the study population to participate in the survey. The spacing out of invitations is used to strategically avoid any clustering effects that may arise with selecting individuals that tend to congregate at the site. TLS will allow the team to estimate both the size and prevalence of the study and target populations.

Sample size calculations for the probability/TLS sample are based on the formulas presented by Levy and Lemeshow (2011). Calculations based on a one-stage cluster sampling design are used to determine a ballpark number of construction sites and individuals within construction sites to observe. For these calculations, prevalence is set to a conservative value of 20%. The team has decided to tentatively recruit from twenty construction workers per sampled site. Assuming a moderately sized intracluster correlation of 0.25, and setting the level of confidence to 95%, we seek a sample size that will result in an estimate that is within 20% of the population prevalence; for example, if prevalence is 20% then 95 times out of 100 the estimate should be within $0.20 * 0.20 = 0.04$ or 4 percentage points of the true value. With these sample size calculation parameters, it is found that a total 92 construction sites must be observed, and a total of $92 * 20 = 1,840$ construction workers must be interviewed.

A **network sample** will also be used to sample individuals who have worked in construction in the past 24 months. Individuals in the network sample will be recruited to participate in the study by a peer who has already completed the survey. The initial sample (also known as “seeds”) will be strategically subsampled from the TLS design as well as those identified by local partners and NGOs. Based on observations and insights gathered from a formative assessment, either a respondent driven sampling (RDS) (Heckathorn and Cameron, 2017) or Vincent Link-Tracing Sampling (VLTS) (Vincent and Thompson, 2017; Vincent, 2019) design will be used. Though both RDS and VLTS rely on a peer-recruitment strategy, RDS is based on a restricted branching process that gives rise to a tree-like network structure

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

whereas VLTS is a free-branching process that gives rise to a web-like networked sample. The VLTS method is based on selecting a moderately-sized initial sample of an array of individuals from the study population. Selecting such a diverse sample can assist with mitigating biases that are introduced with the selection of conspicuous seeds, as commonly seen in RDS. Further, when efforts are placed at the onset of the study and for initial sample selection this can reduce the required sample size to reach a desired level of precision. VLTS will be used if it is determined to be feasible through the formative assessment.

The focus of sample size calculations for the network sample is based on estimation of population size since it is anticipated that this design will reach/access individuals not attached to a formal list of construction sites. Such estimators bare a strong resemblance to the Lincoln-Petersen estimation (Chapman, 1951), and therefore sample size specifications are guided by a simulation study based on this estimator. The corresponding variance estimator for the Lincoln-Petersen estimator is presented in Seber (1970), and the calculations for the anticipated margin-of-error make use of this estimator. It is noted here with importance that, as shown in Vincent (2019) and Vincent and Thompson (2017): (1) with a stratified setup one can expect efficiency gains of at least 25% over the margin-of-error based on these crude approximations, and (2) the Rao-Blackwellized versions of these estimators are likely to give rise to substantial gains in improvement in terms of the margin-of-error, and the magnitude of improvement is likely to be in the vicinity of one-half. Simulation studies are conducted based on a range of population sizes oriented about the reported census numbers and which have been inflated to account for unreported numbers. It is found that an initial sample of size 600 with 1800 individuals added via link-tracing can be expected to give to reasonable margin-of-errors.

B3. Design of Data Collection Instruments

Development of Data Collection Instrument

One instrument will be used for this data collection (see Instrument: Construction Worker Survey). The following table presents the questions, constructs, data sources, and instrument module/sections—by research question (RQ).

Table 1. Research Questions by Instrument Section

Instrument Section
RQ1. How do the number and characteristics of construction workers who self-reported exploitation and trafficking experiences compare by prevalence estimation strategy?
Section 2- Work Experiences Section 3- Work Experiences of Friends Section 5- Sampling Method
RQ2. What is the nature and type of exploitation experienced by construction workers?
Section 2- Work Experiences

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

Instrument Section
RQ3. What are potential risk and protective factors associated with trafficking victimization?
Section 1- Professional Background
Section 4- Personal Demographics

The survey instrument was developed to address the three main research questions. The questions used to estimate the prevalence and describe the nature of labor exploitation and trafficking were adapted from a set of items developed by the Prevalence Reduction Innovation Forum, a collaboration funded by the U.S. Department of State’s Office to Monitor and Combat Trafficking in Persons, to encourage standardization of measurement in prevalence research. The use of those items will allow the results from this data collection to be included in a meta-analysis of other federally funded human trafficking prevalence studies. Questions on potential risk and protective factors included professional background and personal demographics and were developed based on a review of published labor trafficking studies (Barrick et al., 2013; Zhang et al., 2014, 2019). The entire survey was assessed at grade level 7.2. It will be pre-tested with no more than 9 individuals who have worked in construction or experienced labor exploitation in the spring 2022.

B4. Collection of Data and Quality Control

The contractor, RTI International, will collect all data for this study. One-time surveys will be conducted with sampled construction workers.

Recruitment. Participants will be recruited through two methods, depending on which sample they are in. Individuals in the time-location sample will be approached while they are entering or departing a construction site. Upon selecting a set of construction sites and times for observation, a systematic sampling design will be used to invite a set of individuals working at the site to participate in the survey. A field recruiter will approach the selected worker, briefly introduce the survey – including a brief overview of the content, purpose, and the token of appreciation – and hand the worker a card with information on accessing the survey (e.g., QR code that takes the participant directly to the web survey) or a tablet to take the survey on site. Individuals in the network sample will be recruited to participate in the study by a peer who has already completed the survey. The peer will either provide the worker with information on accessing the survey (e.g., unique link that takes the participant directly to the web survey) or provide the individual’s phone number or email address. If a phone number or email address is provided, RTI will send the individual an invitation to take the survey by text or email. The link will include unique identifiers to track which peer provided it; this connection is needed for the analysis and to properly compensate the referrer (i.e., the peer gets compensated for each referral). When a participant initially accesses the survey, they will be presented with the informed consent form, which will be both visible on screen to read and presented via audio. At the end of the informed consent, the participant will have the option to select a button consenting to participate in the short survey or to decline participation. If they agree to participate, they will be redirected to start the survey. If they decline to participate, they will receive a message thanking them for their time.

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

Data Collection Mode. The survey is web-based and will be self-administered via a cell phone, tablet, or laptop. Respondents may use their personal device to self-administer the survey; however, some respondents may not have access to a device or may prefer not to use their personal device. Two options will be available for respondents to use a study tablet if they do not have access to or prefer not to use their personal device. First, the field recruiters for the TLS sample will have tablets available while recruiting so that individuals can immediately complete the survey. Tablets will also be made available at 1-3 community agencies or organizations in the study site for members of the network sample. Participating agencies or organizations may include unions, workers' rights organizations, victim service providers, or similar that have agreed to provide access to both the tablet and a private space for respondents to complete the survey.

Data Quality. Computer-assisted data collection improves survey data quality by eliminating routing errors, implementing logical range checks, and increasing response rates. Even with this technology, however, the quality of the data gathered depends largely on the abilities of the field staff and proper execution of the established field procedures. Field recruiters will attend an in-person or virtual training focused on the study procedures and survey administration. During training, each field recruiter's performance will be evaluated, and additional training provided as necessary to ensure that each has the skills required for the study. The final component of training will involve certification of staff in key areas of field performance, including answering questions about the study and gaining cooperation.

Data Monitoring. During data collection, data quality monitoring activities will be employed on the study, including regular monitoring of field workers, and data frequency reviews (e.g., weekly review of number of respondents). RTI will provide training to field workers and a field manager will monitor field workers' performance as part of their routine quality assurance monitoring. The RTI field manager will oversee field workers through monitoring of select survey recruitment efforts to confirm protocol compliance and proper coding of responses.

B5. Response Rates and Potential Nonresponse Bias

Response Rates

There is no prior research using TLS or network sampling among construction workers upon which to estimate a response rate. Moreover, response rates are typically not reported for studies using these sampling methods because they are not observed by the research team. For example, among human trafficking prevalence studies using network sampling (Curtis et al., 2008; Zhang et al., 2014; Dank et al., 2019; Vincent et al., 2019; Zhang et al.; 2019b) and TLS (Zhang et al. 2019), none report response rates. For network sampling, recruitment relies on study participants inviting their peers to participate. Because the respondents are only surveyed once, the research team does not have information on the number of individuals that each respondent attempted to recruit and how many accepted. With respect to TLS, a short screener is used to determine if a respondent is part of the study population (i.e., construction worker). Oftentimes, a refusal to participate in the survey will come before any information is obtained so response rates would only apply to the general population. With suitable tokens of

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

appreciation built into the study, it is assumed that response rates will be low but will not affect the difference between the observed and unobserved part of the study population.

Response rates will be calculated differently for the TLS and network samples. With respect to the **TLS strategy**, unit non-response will be addressed with a logistic regression model as follows. The field team will be instructed to systematically approach every k^{th} encountered individual at the visited site, screen for eligibility (i.e., construction worker), and then provide eligible individuals a card with a unique code and link to the survey. When an individual logs on to the survey, their demographic information will be recorded as part of the survey questionnaire along with the site/venue and day/time they were recruited. The site/venue will be built into the unique code on the invitation card. The field team will keep a rough tally of the number of refusals for each site/venue and day/time of recruitment. However, this response rate will only be an estimate of the general population since some refusals may not have been eligible to participate. Note that after an individual refuses an invitation card (either before or after screening), the field team will be instructed to invite the next individual they encounter.

A logistic regression model will be formulated based on the indicator of non-response or response for each individual i , y_i say where $y_i=0, 1$ respectively for non-response and response. Such values will be regressed against their observed demographics X_i , as above, mathematically denoted as

$$p_i = \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)} \text{ where } p_i \text{ is individual } i\text{'s probability of accepting the invitation, i.e. } y_i=1. \text{ The}$$

regression parameters will be approximated with a maximum likelihood estimation approach and the selection probabilities will be approximated through this mathematical expression.

Non-response may also occur at the site/venue level, either due to it being out of operation or refusal by the site supervisor upon request for entry. With respect to the former, if it is suspected that a moderate number of sites will be out of operation at the time of the study then sample selection will incorporate an additional level of randomization through a Bernoulli sample design. Essentially, each site sampled from the frame will be visited with a probability, α say, as this will allow for the unbiased incorporation of any newly found sites not listed on the frame. For sites that are out of operation, they are considered part of the sample but simply receive a zero count for their corresponding operations.

With respect to the latter, a logistic regression model similar to that outlined for the unit level response/non-response will be posited and based on site-level observations such as area of operation, approximate number of workers, type of construction project (i.e. business or residential), day/time of observation, etc.

The **network sampling design** will commence with an initial sample (also known as “seeds”) of individuals strategically subsampled from the TLS design as well as those not typically accessible through the TLS design but accessible through the local partners and NGOs. This will ensure the sample is comprised of individuals with an array of demographic profiles. The branching design will then be based entirely on peer recruitment. At the time of interview, individuals will be asked for demographic information of their nominees so that a similar logistic regression model to that outlined above can be formulated to model recruitment from a respondent’s personal network.

NonResponse

Nonresponse is expected to be minimized by the use of a one-time self-administered survey and limiting the survey to 30 minutes in length, at maximum, and utilize skip patterns that ensure participants only

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

receive questions that are relevant to them. Tokens of appreciation will also be used to mitigate item-level non-response. For cases where a small to moderate amount of item non-response is encountered, the data will first be checked to ensure it satisfies the missing at random assumption (MAR) (Rubin, 1976); MAR suggests that the survey responses are missing in a systematic pattern that could result in bias in the estimates, but whose pattern is likely to be explained/predicted with demographic information and other survey responses. Hence, a regression model will then be formulated based on the observed demographics and survey variables which correlate well with the variable under consideration, and a multiple imputation procedure (van Buuren, 2012) will be used to properly address such missing entries. This procedure in itself does not require weight adjustments since the estimation procedure will incorporate the imputed values and therefore be based on a collection of fully observed data sets.

An alternative procedure will be considered and which is not based on imputation. The procedure entails either rescaling or recalibrating the weights for those whose entries have been observed. Recalibrating the weights may be considered appropriate for cases where the calibration variables sufficiently predict the missing item entries.

For cases where a large percentage of observations are missing for a specific item, the research team will report on this occurrence and then discuss the option of either removing the item from the analysis entirely or to research a suitable method for addressing the missing entries.

B6. Production of Estimates and Projections

Estimates produced by this work will be prepared for internal use by ACF and external release by the agency in the form of reports, presentations, and/or other publications. While we intend for the study sample to be representative of construction workers in Houston, reports and other dissemination products will note that the sample is not nationally representative and data should not be used to generate broader population estimates.

B7. Data Handling and Analysis

Data Handling

RTI staff will thoroughly test the instruments to ensure that the surveys are programmed into Voxco Web accurately and in a manner that minimizes error. This includes automated and manual checks on the programming of each system to ensure that skip patterns are properly programmed, and response options are within allowable ranges. RTI staff will attempt to test all possible response categories for each question in order to identify embedded logic errors, as well as obvious skip problems. Several analysts may test the program simultaneously to identify problems quickly, and to double check the comprehensiveness of the testing.

The surveys will be hosted in RTI's National Institute of Standards and Technology (NIST) Moderate network and as data are collected, they will be stored in SQL server databases within RTI's NIST Moderate network.

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

Data Analysis

Along with the list of core research questions, **Table 2** lists the research questions. The same analytic strategies, which are detailed below, are used to answer each of these questions.

Table 2. Research Questions

RQ1: How do the number and characteristics of construction workers who self-reported exploitation and trafficking experiences compare by prevalence estimation strategy?	
RQ1a	How many and what percentage of construction workers in the study site have experienced labor exploitation and trafficking?
RQ2: What is the nature and type of exploitation experienced by construction workers?	
RQ2a	What types of exploitation were experienced by construction workers during recruitment (e.g., coercion, deception, fees)?
RQ2b	What types of exploitative employment practices and penalties were experienced by construction workers (e.g., pay withheld, made to perform additional services or responsibilities, unpaid overtime)?
RQ2c	How do employers treat the personal life and property of construction workers (e.g., control over a meaningful part of someone’s life, confiscation of mobile phones)?
RQ2d	What types of degrading conditions were experienced by construction workers?
RQ2e	To what extent was the freedom of movement or communication of construction workers restricted (e.g., identification documents confiscated, surveillance and monitoring) ?
RQ2f	What forms of debt or dependency were experienced by construction workers?
RQ2g	What forms of violence (or threats of violence) were experienced by construction workers?
RQ3: What are the potential risk and protective factors associated with trafficking victimization?	
RQ3a	What individual-level factors (e.g., gender, national origin, English proficiency) differentiate construction workers who report trafficking experiences from other workers?
RQ3b	What employment characteristics (e.g., construction work during natural disaster recovery, type of construction work, length of employment in construction, methods for finding work in construction) differentiate construction workers who report trafficking experiences from other workers?

Each of the research questions in Table 2, along with population size estimation, will be investigated and answered with each of the inference approaches that correspond to the TLS and network sampling design, and the majority of the statistical analyses will be conducted in the R programming language where data cleaning will be based on logical checks and outlier detection procedures. For the TLS design, linkages across observations made at different construction/venue sites and/or dates will be facilitated

Alternative Supporting Statement for Information Collections Designed for Research, Public Health Surveillance, and Program Evaluation Purposes

with the 'RecordLinkage' package (Borg, A. and Sariyar, 2016) as this is required for identifying duplicate entries and Multiple Systems Estimations (MSE) estimating procedures. The procedures detailed in Leon et al. (2015) and Chan et al. (2020) will be used for population size estimation. Sample weights will be evaluated based on selection probabilities corresponding to the sample frame which is the Cartesian product of the venues and dates. Statistical packages like 'survey' (Lumley, 2010) will be used to develop sampling weights for the TLS sample, which in turn will be used to calculate point estimates, bootstrap-based variance estimates, and confidence interval calculations to estimate survey variables such as those which correspond to the research questions.

For RQ3a, weighted regression analyses will be used to gain insight into how demographic profiles relate to the corresponding observed characteristics.

Data Use

Dissemination of survey findings may include research briefs, summary reports, and peer-reviewed journal manuscripts. Priority audiences for these products will include researchers focused on estimating the prevalence of human trafficking, labor trafficking service providers, and the construction industry. ACF and RTI are committed to transparency in research. Sharing access to the survey data set will support further analyses that can inform policy and practice related to labor trafficking. We will archive the final, deidentified survey data set, likely with the National Archive of Criminal Justice Data (NACJD), following their requirements for ensuring privacy. RTI will prepare supporting materials to contextualize and assist in interpretation of the data and as required by NACJD, including documentation of sampling methods, response rates, population of inference, construction of analytic variables, and construction and appropriate use of survey weights, along with a variable list and codebook. Archiving data and supporting information will support future contributions to our understanding of labor trafficking victimization, including analyses by researchers and within other federally-funded studies.

B8. Contact Persons

Mary Mueggenborg
Senior Social Science Research Analyst
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
330 C Street SW, 4th floor
Washington, DC 20201
mary.mueggenborg@acf.hhs.gov | 202.401.5689

Kelly Jedd McKenzie
Senior Social Science Research Analyst
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
330 C Street SW, 4th floor
Washington, DC 20201

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

kelly.mckenzie@acf.hhs.gov | 202.245.0976

Kelle Barrick
Senior Research Criminologist
Division of Applied Justice Research
RTI International
3040 E Cornwallis Rd 27709
Research Triangle Park, NC
kbarrick@rti.org

Rebecca Pfeffer
Research Criminologist
Division of Applied Justice Research
RTI International
3040 E Cornwallis Rd 27709
Research Triangle Park, NC
rpfeffer@rti.org

Attachments

Instrument 1: Construction Worker Survey
Appendix A: Recruitment Script for Construction Workers
Appendix B: Consent Form
Appendix C: Public Comments
Appendix D: Public Responses

References

Barrick, K., Lattimore, P. K., Pitts, W. J., & Zhang, S. (2013, August). *Indicators of labor trafficking among North Carolina migrant farmworkers* (Report No. 0212466). Prepared for National Institute of Justice.

Borg, A. and Sariyar, M. (2016). RecordLinkage: Record Linkage in R.

Chan, L., Silverman, B. W., and Vincent, K. (2020). Multiple systems estimation for sparse capture data: Inferential challenges when there are non-overlapping lists. *Journal of the American Statistical Association*. DOI: [10.1080/01621459.2019.1708748](https://doi.org/10.1080/01621459.2019.1708748)

Chapman, D. (1951). Some properties of the hypergeometric distribution with applications to zoological sample census. *University of California Publications in Statistics*. **1**, p. 131-160.

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

Curtis, R., Terry, K., Dank, M., Dombrowski, K., & Khan, B. (2008). The CSEC Population in New York City: Size, Characteristics, and Needs. Center for Court Innovation.

Dank, M., Vincent, K., Hughes, A., Dhungel, N., Gurung, S. & Jackson, O. (2019). Prevalence of Minors in Kathmandu's Adult Entertainment Sector. Freedom Fund.

Heckathorn, D. D., and Cameron, C. J. (2017). Network sampling: From snowball and multiplicity to respondent-driven sampling. *Annual Review of Sociology* **43**, p. 101-119.

Leon, L., Jauffret-Roustide, M., and Le Strat, Y. (2015). Design-based inference in time location sampling. *Biostatistics* **16**, 3, p. 565-579.

Levy, P. and Lemeshow, S. (2011). *Sampling of Populations: Methods and Applications*. Wiley.
Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons, Ltd.

McCarty, C., Killworth, P., Bernard, H., Johnsen, E. and Shelley, G. (2001). Comparing two methods for estimating network size. *Human Organization* **60**, 1, p.28-38.

Rubin, D. B. Inference and Missing Data. *Biometrika*, *63*, 581-592.

Seber, G. A. F. (1970). The effects of trap response on tag recapture estimates. *Biometrics*. **26**, p. 13-22.

Jonsson J, Stein M, Johansson G, Bodin T, Strömdahl S (2019) A performance assessment of web-based respondent driven sampling among workers with precarious employment in Sweden. *PLoS ONE* *14*(1): e0210183. <https://doi.org/10.1371/journal.pone.0210183>

Thompson, S. (2020). New estimates for network sampling. *arXiv preprint 2002.01350*.

van Buuren, S. (2012). *Flexible Imputation of Missing Data*. Chapman and Hall/CRC Press. Taylor & Group. F.

Vincent, K. (2019) Recent Advances on Estimating Population Size with Link-Tracing Sampling. *arXiv preprint: 1709.07556*

Vincent, K. and Thompson, S. (2016). Estimating Population Size with Link-Tracing Sampling. *Journal of the American Statistical Association*, *112* (519), 1286-1295.

Vincent, K. and Thompson, S. (2020). Estimating the Size and Distribution of Networked Populations with Snowball Sampling. *Journal of Survey Statistics and Methodology* (To Appear).

Vincent, K., Zhang, S.X., & Dank, M. (2019). Searching for sex trafficking victims: Using a novel link-tracing method among commercial sex workers in Muzaffarpur, India. *Crime & Delinquency*, <https://doi.org/10.1177/0011128719890265>

Zhang, S.X., Barrick, K., Evans, B., Weber, R., McMichael, J., Mosquin, P., Vincent, K., & Ramirez, D. (2019a). Labor Trafficking in North Carolina: A Statewide Survey Using Multistage Sampling. Final Report submitted to the U.S. Department of Justice.

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

Zhang, S.X., Dank, M., Vincent, K., Narayanan, P., Bharadwaj, S., Balasubramaniam, S.M. (2019b). Victims without a voice: Measuring worst forms of child labor in the Indian State of Bihar. *Victims & Offenders*, 14(7), 832-858.

Zhang, S.X., Spiller, M.W., Finch, B.K., & Qin, Y. (2014). Estimating labor trafficking among unauthorized migrant workers in San Diego. *The Annals of the American Academy of Political and Social Science*, 653, 65-86.