PREVALENCE **ESTIMATION**

Methods Brief

GLOBAL FUND TO **END MODERN SLAVERY**

Norad

Methods of Prevalence Estimation in Modern Slavery

An Introductory Overview

Contents

Editors

Laura Gauer Bermudez, David Okech, and Mihir Prakash

Contributing Authors

01. Household Surveys (Jacqueline Joudo Larsen, David Okech)

02. Respondent Driven Sampling/ Link-Tracing Sampling (Sheldon Zhang, Kyle Vincent)

03. Time and Location Sampling (Andee Cooper Parks, Clifford Zinnes)

04. Network Scale-Up Method (Sam Blazek, Dennis Feehan)

05. Multiple Systems Estimation (Davina Durgana, Jan Van Dijk)

06. Hybrid Approaches (Kareem Kysia, Clifford Zinnes)

Acknowledgements

This brief was produced through a collaborative editorial effort between the Global Fund to End Modern Slavery (GFEMS) and the African Programming and Research Initiative to End Slavery (APRIES) at the University of Georgia (UGA). Laura Gauer Bermudez (GFEMS), David Okech (UGA), and Mihir Prakash (GFEMS) led the editorial efforts.

The editorial team is grateful for contributions made by Jacqueline Joudo Larsen (Walk Free), David Okech (UGA), Sheldon Zhang (University of Massachussetts- Lowell), Kyle Vincent (Independent Researcher and Consultant), Andee Cooper Parks (UNICEF), Clifford Zinnes (NORC at the University of Chicago), Sam Blazek (IST Research), Dennis Freehan (University of California-Berkeley), Davina Durgana (Walk Free), Jan Van Dijk (University of Twente), and Kareem Kysia (NORC at the University of Chicago), who volunteered their time to share their knowledge and expertise on modern slavery prevalence research in the form of short chapters. We would also like to thank Natalie Wodniak (George Washington School of Public Health) who provided valuable research assistance during her tenure at GFEMS.

The editorial team would like to extend a special thanks to Jean Baderschneider (GFEMS) and Sabrina Thulander (GFEMS) for their support to the development of this publication.

The Norwegian Agency for Development Cooperation (NORAD) provided generous support to produce this document. Any views and opinions expressed by the contributing authors are those of the authors alone and do not necessarily reflect the views of funder.

INTRODUCTION

Despite being long banned and universally condemned, "slavery persists in many corners of the world today, victimizing tens of millions of people".[1] In the modern age, slavery manifests in the form of forced labor, debt bondage, forced marriage, commercial sexual exploitation, human trafficking, and other slave-like practices.[2] The commonality among these legal concepts is referred to using the umbrella term of 'modern slavery'.

Building an evidence base is critical to facilitating data-driven decisions by policy makers, legislators, and other key stakeholders. There is a strong consensus among experts and institutions, including the Global Fund to End Modern Slavery (GFEMS), the Center on Human Trafficking Research and Outreach (CenHTRO), and collaborators on this document, that evidence creation is at the heart of any strategy to address modern slavery. Yet, limited resources and a dearth of information related to what works to drastically and measurably reduce this crime against humanity makes it difficult to encourage investment in anti-slavery efforts. Thus, more evidence is needed to support resource mobilization, public policy changes, and new investments in anti-slavery programming. This evidence may come in many forms including, but not limited to, intervention effectiveness studies, policy analyses, geo-spatial mapping, vulnerability profiling, risk assessments, prevalence estimation, and meta-analyses of independent studies. This document exclusively focuses on prevalence estimation due to its complexity and the nuance with which such estimates should be undertaken and understood.

Prevalence is defined as the proportion of a population who have a specific characteristic in a given time period.[3] The term has primarily been used by public health professionals, particularly epidemiologists, as a measure to track the spread of various diseases over time. But the concept is equally applicable to measuring the scale of the modern slavery problem.

Prevalence may be represented as a count (number of individuals in a particular type of modern slavery) or as a ratio (% of individuals in a given geography or industry that have experienced modern slavery within the period under study). The presentation choice, whether it be a ratio or a count, may depend on how common the characteristic of interest is in the population.[4] Some researchers may choose to report both figures.

Beyond presentation, prevalence can also be measured in different ways. Prevalence can be measured cross-sectionally, providing an estimate of the proportion of a population in modern slavery at a specific point in time, referred to as *point prevalence*. An example of point prevalence would be the number of domestic workers in a country that self-report experiencing modern slavery at the time of interview. Another form is *period prevalence* which estimates the proportion of the population that has experienced modern slavery within a given time period. For instance, the percentage of migrant workers returning from country X that have experienced modern slavery within the last 12 months would represent period prevalence. A variant of period prevalence is *lifetime prevalence*, which uses the duration of interest as an individual's lifetime. An example of this would be the proportion of the population that, over their lifetime, has been a victim of one or more types of modern slavery at least once.

Because collecting information from every member of a target group can often be prohibitive, prevalence estimation typically involves collecting data from a *sample* of individuals that belong to the larger target population using a sample survey.[5] Ideal size of the sample and the

Traditional sampling and survey methods face limitations when attempting to gauge the prevalence of hidden victims.

sampling design (e.g., random, stratified, cluster, convenience, snowball) can vary based on the type of the study, nature of the population of interest (i.e., how visible or hidden they are), the overall population size, as well as the error tolerance of the research. The margin of error of an estimate is commonly referred to as a *confidence interval* (CI) and should always be reported alongside the statistically estimated value. Prevalence estimates, like any other estimate, are informed generalizations of the characteristics of the population, made possible under a set of empirical and theoretical assumptions.

> Prevalence is not to be confused with *incidence*, which is the rate at which new cases of individuals with the characteristic of interest emerge each month or year.

Compared to similar research in other domains, prevalence of modern slavery can require more resources and expertise to estimate. This is because the populations of interest to modern slavery researchers, such as child victims of sex trafficking, victims of debt bondage, and individuals working in labor intensive, factory or home-based operations are often very hard to reach. Victims can be hidden within communities for a variety of reasons, including employer incentives to keep workers isolated.

As such, traditional sampling and survey methods face limitations when attempting to gauge the prevalence of hidden victims, leading researchers to increasingly engage with less traditional methods of locating and counting these populations. These estimation methods, largely adapted from the public health sphere, are now experiencing wide adoption and scientific acceptance by modern slavery researchers. However, the conceptual and methodological understanding of these innovative, non-traditional methods remains out of reach for the broader non-technical community of modern slavery stakeholders.

To overcome these barriers, this document presents five of the most common methods of prevalence estimation used in modern slavery research. These methods are shared in the form of five short chapters. A sixth chapter discusses multi-method hybrid approaches, or various combinations of these five methods, that can be deployed to heighten accuracy and address limitations to the use of a singular method in a given sector or geography. Each chapter, where necessary, also discusses important factors that may lead to variations such as: the population covered, the timing of data collection, sampling design, mode of data collection, survey instruments used; operational definitions used, and statistical analysis techniques.

Each chapter provides a concise overview of the prevalence estimation technique; discusses their strengths and limitations; and includes citations for key works on each method – both within the modern slavery sector and adjacent fields. Each chapter is written by researchers known in the field, who have invested significant amounts of time towards studying and applying the method in the modern slavery context.

It is the hope of the authors that this document serves as a primer on prevalence estimation and helps facilitate informed choices on which method, or combination of methods, are best-suited for the industry and geography under study. The authors also hope that future generations of methodologists, statisticians, and data scientists working on modern slavery issues can build upon these techniques to refine, combine, and expand the ways in which to measure and understand modern slavery.

1. Council on Foreign Relations (2020). *Modern Slavery: An exploration of its root causes and the human toll.* Washington, D.C., USA. Available at: https://on.cfr.org/342c1xb Accessed on: 15th October 2020.
2. Walk Free & International Labour Office (2017). *Global estimates of modern slavery: Forced labour and forced marriage.* p.p.16. Geneva, Switzerland.
3. National Institute of Mental Health (2020). *What is Prevalence?* Available at: https://bit.ly/2Fx4w7Y Accessed on: 15th October 2020.
4. ibid.
5. While a Census gathers data, to the extent possible, on all citizens of a country, a sample survey strategically gathers information from only a representative sample of the entire population with the intent of estimating population level characteristics. A Census is typically conducted once every ten years, while the frequency of sample surveys can range from anywhere between every 1 - 5 years.

01

METHOD 01:
Household Surveying

Jacqueline Joudo Larsen and David Okech

Overview of Household Surveying

Household surveying is an important, and perhaps the most common, method used for drawing statistical inferences[6] on a population of interest. Surveys can also be used for prevalence estimation of modern slavery. The gold standard in the field of household survey methods are studies that use official statistics as a reference frame to design stratified, probabilistic, and random sampling strategies. This establishes a structure for respondent sampling that generates representative insights. The official parameters for reference can be a census or any other form of representative household survey that provides information on relevant attributes of the total population.

High level explanation of the method

To conduct household surveys, the first step is to design a sampling methodology. The methodology identifies the group of individuals that will be asked to respond to a survey with questions about their experiences, or the experiences of their household members, with respect to forced labor, forced marriage, or other dimensions of modern slavery. Where feasible, inclusion of modern slavery questions within a comprehensive national census is an excellent, and perhaps the most comprehensive approach, to survey nearly all citizens through government-sanctioned resources. However, where inclusion of a modern slavery module is not yet feasible or not sufficiently comprehensive to understand the elements of modern slavery under study, a representative household survey design is a decent option.

The term representative means that the proportions of individuals that are sampled from each sub-group of interest is the same as the proportion of that sub-group in the overall population (see: Figure 1). In other words, sampling statistics such as mean, standard deviation or proportion of people with a certain characteristic should mirror the relevant population parameters. This representativeness is based on the needs of the study and typically implies the use of criteria such as age group, ethnicity or gender to classify the population into mutually exclusive sub-groups. By ensuring that the final sample of a survey is representative, any statistical inferences that are drawn can be generalized to reflect the diversity of the overall population of interest, thereby allowing researchers to infer a lot without surveying everyone in the total population.

This idea of representativeness of the sample can also be framed in probability terms. Here, the sampling strategy provides data that, when analyzed, will provide population-level inferences if the probability of selection of an individual to be surveyed is the same as the proportion of the sub-group that they belong to in the total population. This is also called probability proportional to size (PPS) and is a commonly used technique in household surveying.

To be able to conduct a household survey that provides accurate and representative estimates, an existing reference database that provides information about the total population is necessary to develop a sampling frame. A national census, where available, or other previously conducted large-scale demographic surveys serve as a good source of information for this purpose since they provide aggregate demographic, socio-economic and other relevant characteristics of all individuals in the geographic location of study.

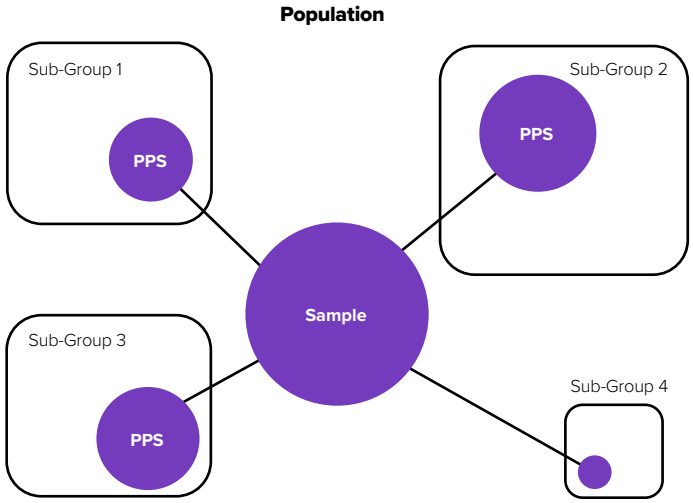


Figure 1: How Stratified Random Sampling Works

1. Each group in the total population is captured in the sample.
2. The number of sampled individuals in each group are proportional to the share of each group in the total population (probability proportionate to size - PPS).
3. Individuals in each group that are sampled have the same probability of selection (equal likelihood) as everyone else in that group. In other words, the sampling is random.

Source: Authors Illustration

In addition to developing the sampling frame, a survey questionnaire is designed. This questionnaire should go through cognitive testing to ensure that the questions that are being posed to survey respondents are sufficiently clear and the respondents are able to answer the question in a manner that is similarly clear to analysts. As with all of the estimation methods featured in this brief, pre-design work – such as formative, qualitative inquiries - can ensure cultural adaptation of the instruments and uncover hidden nuances that, if left unaddressed, might affect the validity of the research. Western-based researchers should always work closely and collaboratively with local researchers, including a thorough training of local data enumerators. Questionnaires should be translated into the appropriate local languages.

Data collection can occur face-to-face, with trained enumerators visiting each sampled household, or it can happen remotely (e.g., via telephone, SMS) when contact information for all households in the sampling frame is readily available. Traditionally, face-to-face to face surveys have been the preferred medium of enumeration for global research. However, telephone-based surveys have gained popularity particularly during periods of global health concern[7].

Conducting household surveys often requires prior government approval.[8] This is particularly relevant in the case of modern slavery research due to the sensitivities surrounding research with vulnerable populations and crimes that potentially involve the governments themselves. Protocols must be developed that maintain the safety of both respondents and enumerators including data management processes that protect respondent confidentiality as well as referral networks for connecting respondents that may require immediate assistance. Any analysis of survey data should also make sure that results are published at the aggregate level and that the identities of survey respondents is protected using data anonymization or obfuscation techniques.

Assumptions and other considerations

Enumerators that present no conflict of interest with the population that they are interviewing, and who have demonstrated a clear mastery of interviewing protocols, are critical to gathering unbiased data. When possible, enumerators who can conduct the survey in local languages should be deployed.

When asking questions on sensitive issues such as those of modern slavery, regardless of how trained an enumerator is in trust building, there can be a tendency among interviewees towards choosing a more socially desirable response rather than choices that reflect their truthful experiences. This tendency to avoid judgement by an enumerator is often referred to as *social desirability bias*. To overcome this challenge, researchers are innovating by deploying a variety of methods including computer-assisted self-interviewing (CASI), audio computer-assisted self-interview (ACASI), and computer-assisted personal interview (CAPI), which have proven effective in eliciting more truthful responses on sensitive topics. For example, a methodological experiment conducted by the Population Council in Kenya found that boys and girls report on their sexual behaviors differently between a face-to-face and an ACASI interview[9], demonstrating the ability of self-interview platforms to address social desirability bias.

Household surveys do have limitations in identifying modern slavery. For instance, victims of forced marriage will likely not be reported as such by their intimate partners and employers may not be truthful about how they treat

their domestic workers, particularly if they fear negative repercussions. In both cases, if the victim was to be interviewed directly, under the watchful eye of their abuser, they might either choose to not report on their experiences truthfully due to fear of consequences or jeopardize their safety by sharing freely about their living or work conditions.

Further, household surveys typically focus on non-institutionalized populations due to challenges with access. This means that people in institutions such as prisons, labor camps, military barracks, and orphanages are not sampled. Another limitation is that household surveys often fail to capture undocumented migrant workers whose data may not be reflected in the national census or other baseline surveys from which the sampling frame is created. Interviewing at-risk children via a household survey method can be difficult due to the presence of the parent and need for parental consent. It is, therefore, advisable to consider these limitations and any corresponding risks to study validity when designing a conventional household survey for estimating modern slavery prevalence.

Therefore, conventional probability-based household surveying remains the gold standard in global research and can be effectively utilized for inquiries on modern slavery when these limitations are accounted for within the research design.

Examples of Household Surveys

As one of the most deployed methods in social science research, many examples of household survey best practices exist. But in the context of modern slavery research, the 2017 Global Estimates of Modern Slavery (GEMS) surveys, conducted by Walk Free and the International Labor Organization (ILO), provide a good illustration of how household surveys can be used for estimating the prevalence of modern slavery. GEMS is the largest survey research program in the modern slavery field. Collectively, Walk Free and ILO have conducted nationally representative probabilistic household surveys in 48 countries using the Gallup World Poll. Gathered data is used to estimate the prevalence of modern slavery at the national level and rank countries on the Global Slavery Index.[10]

The ILO also frequently conducts National Child Labour Surveys (NCLS) that estimate the prevalence of child labor at the national level using household surveys. These surveys are conducted under the ILO's Statistical Information and Monitoring Programme on Child Labour (SIMPOC).

Reference List Of Useful Articles

Diego-Rossel, P. and J. Joudo Larsen (2018). Modelling the Risk of Modern Slavery. SSRN. Available at SSRN: https://ssrn.com/abstract=3215368 or http://dx.doi.org/10.2139/ssrn.3215368

International Labour Organization (n.d.) Statistical Information and Monitoring Programme on Child Labour (SIMPOC). Geneva, Switzerland. Available at: https://bit.ly/2K7nKTO Accessed on: 20th November 2020.

International Labour Organization & Walk Free (2017). Global Estimates of Modern Slavery. Geneva, Switzerland. Available at: https://bit.ly/3qHd9jg

International Labour Organization (2020). International Conference of Labour Statisticians (ICLS). ILOSTAT. Geneva, Switzerland. Available at: https://ilostat.ilo.org/about/standards/icls/

Joudo Larsen, J. and P. Diego-Rossell (2017). Using Surveys to Estimate the National Prevalences of Modern Slavery: Experience and Lessons Learned. CHANCE, Vol. 30(3), p.p. 30-35. DOI: 10.1080/09332480.2017.1383110

Walk Free (2018). Global Slavery Index Appendix 2: Part B: Global Slavery Index Prevalence Estimation. Minderoo Foundation. Perth, Australia. Available at: https://www.globalslaveryindex.org/resources/downloads/

6. Statistical inference is the process of using information derived from a sample to draw conclusions about what is likely true for the entire population.

7. CGIAR (n.d.) *Using mobile phones to do research in the time of COVID-19 lockdowns and beyond.* Washington, D.C., USA. Available at: https://bit.ly/3lQ1Hy0 Accessed on: 20th November 2020.

8. This difficulty has led to data coverage gaps in some key regions such as the Arab States and in Central and East Asia, where access is restricted or topics may be more sensitive.

9. For further information about this study, see: Mensch, B., Hewett, P.C., and A. Erulkar (2001). *The reporting of sensitive behavior among adolescents: A methodological experiment in Kenya.* Population Council. New York, NY, USA. Available at: https://bit.ly/3qEhDqQ

10. More information on the survey methodology and surveyed countries can be found at www.globalslaveryindex.org

02

METHOD 02:
Respondent Driven Sampling / Link Tracing Sampling

Sheldon Zhang and Kyle Vincent

Overview of Respondent Driven Sampling/Link Tracing Sampling

Respondent Driven Sampling and other link tracing sampling methods seek to derive representative samples from populations that are either hidden or highly skewed. Modern slavery is not an equal-opportunity crime that afflicts the general population evenly or by chance. Instead, victims of modern slavery *cluster* in pockets of certain social demographic strata of vulnerable populations, which by nature defy conventional social science methods to study them. In other words, probability-based sampling strategies, such as household surveys, may find few or no victims in a defined economic sector or geographical area and result in estimates which arise from a highly skewed distribution. Therefore, non-conventional methods need to be considered in the sampling design and inference strategy to obtain more accurate estimates.

High level explanation of the method

Respondent-driven sampling (RDS), developed by Heckathorn (1997, 2002), is arguably the most-known methodology for studying hidden populations. The method has been widely applied by public health researchers to study hard-to-find populations such as substance misusers or by sociologists to reach small networked individuals such as jazz musicians or sex workers.

RDS is essentially an extension, or a more structured version, of the traditional snowball sampling approach. RDS employs two main strategies to overcome some of the limitations associated with snowball sampling: (1) incentivized recruitment; and (2) structured recruitment procedure. Like traditional snowball sampling, both of these methods involve initial study recruits, also known as *'seeds,'* who fit the profile for study participation, and then refer additional people by sharing provided referral coupons within their network, who then also refer others, and so on. Each referral is called a *'wave.'* Typically, RDS involves a small number of initial seed participants and multiple waves of recruitment initiated from each seed. As the number of waves increases, according to a discrete-time Markov chain, the sample properties are believed to achieve equilibrium, i.e., becoming independent of the initial entry points.

Assumptions and other considerations

The fundamental assumption for all link-tracing strategies is that people in the target population are connected through networks that possess the qualities of interest. We know that people of similar socio-economic backgrounds and vulnerability profiles tend to know one another. For example, victims of commercial sexual exploitation (CSE) likely know other victims. As this is the case, network-based sampling strategies can then be deployed to produce estimates for these pockets or subsets of populations. Because RDS/link-tracing sampling relies on structured and incentivized recruitment, sufficient waves of recruitment and well positioned entry points are able to derive a representative sample of the "hidden" population.

By controlling the recruitment opportunities through a structured process, i.e. ensure that respondents refer other participants without any biases, diversity can be ensured and thus can be empirically verified. Volunteerism is minimized, as a dual-incentive system is believed to encourage both participation and recruitment. Such a recruitment procedure prevents researchers from deliberately seeking out particular subjects. 'Masking'[11] is minimized since researchers are not pointed in the direction of group members, but rather, respondents are recruited by group members themselves

for participation in the study. Homophily[12] is also minimized since recruitment is limited to three subjects per participant, and equilibrium can be achieved through a relatively small number of waves. The three referrals limit in the RDS method also minimizes biases that may be introduced by those with larger personal networks.

For its many benefits over other sampling methods, conventional RDS does not come without its challenges. First, it is a rather passive sampling process that may last much longer than planned, despite restricting the timeframe for recruitment and coupon redeeming activities. Second, it is difficult to balance uneven recruitment activities, with some respondents generating few waves of referrals and others appearing excessive. Third, traditional RDS also seems to be geographically constrained, with referrals and nominations unable to penetrate invisible boundaries to reach wider networks (see: Figure 2). Finally, deriving population-level estimates using RDS is not as straightforward as some other methods because link-traced individuals are recruited in the sample with unequal inclusion probabilities due to the varying network sizes and recruitment preferences. RDS inference procedures typically assume sampling follows a random walk procedure where sampling occurs with replacement (Thompson, 2020).

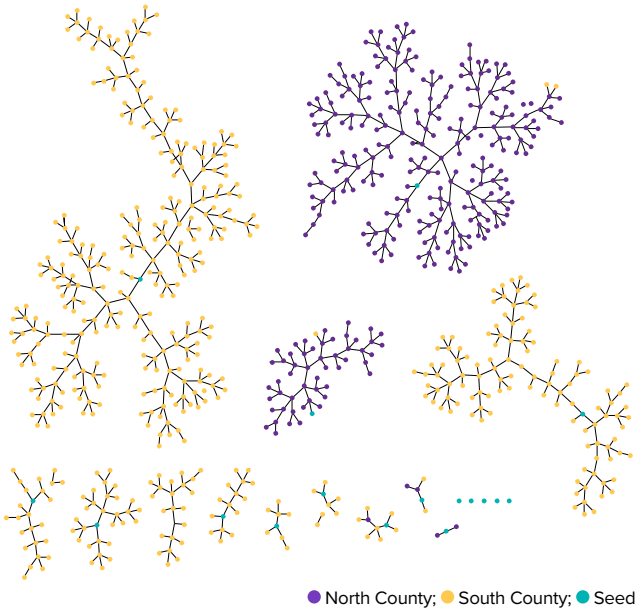


Figure 2: Referral Chains by South/North County

Source: Zhang et al. (2014)

Note: The figure visualizes the network of referrals by seeds in San Diego's North County (Red) and South County (Yellow). Recruitment waves remained mostly inside the same geographical area, unable to breach into different parts of the county.

Recently, Canadian statistician Kyle Vincent devised a sampling strategy – the Vincent Link Tracing Sampling (VLTS) – to overcome some of the major shortcomings of traditional RDS designs. VLTS retains the features of conventional sampling designs but follows the traditional RDS recruitment process (Vincent & Thompson, 2017). Briefly, VLTS draws on the benefit of any existing sampling frames such as those designed using the census for household surveys (comprehensive or partial) to develop a generously sized initial sample of seeds. The initial selection of the seeds can be based on conventional sampling designs such as simple random sampling, stratified random sampling, or systematic sampling that are based on available, albeit imperfect, sampling frames. Individuals are then added at each wave for two to three waves via subsampling from personal networks of the initial seeds. Further, links within the final sample, not just the network size, are explored to identify overlaps between individual networks with the assistance of sophisticated inference techniques that accompany this estimation strategy.

The VLTS procedure exploits the ability of observing adjacent (neighboring) units of sampled individuals once a unit of high interest has been found. The procedure has the ability to retain the attractive features of conventional sampling strategies, such as the ability to obtain unbiased estimators, and control for final sample sizes. Upon selection of the initial sample, one can develop referrals or nominations where social networks can be mapped to identify overlaps, thus adaptively building up the final sample. A Rao-Blackwell inference strategy outlined by Vincent and Thompson (2017) and Vincent (2019) can incorporate the members selected through link-tracing into the inference procedure without introducing any bias to population size and other population quantity estimators, while reducing the variance of the resulting estimators.

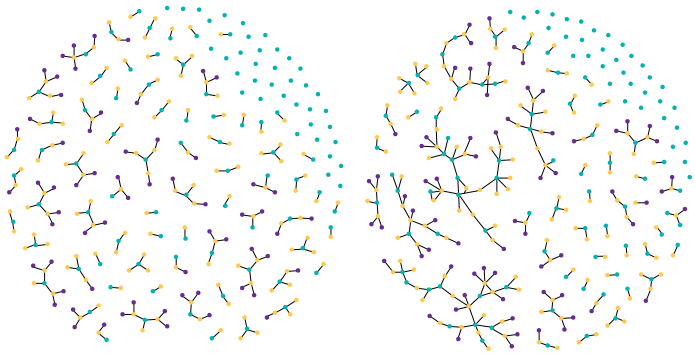
The VLTS uses all available resources, in the form of auxiliary information based on knowledge of the study population, for the start of data collection in order to obtain as representative a sample as possible early on in the study. Hence, one need not be concerned with obtaining a large number of additional waves, in contrast to conventional RDS, before unbiased estimation strategies can be applied. Additional linking, which extends beyond what is solely observed through coupon referrals, allows for more sophisticated network analyses and efficient population size estimation procedures derived from mark-recapture procedures. The initial sample, or the "seeds", comprise the first wave. Sampling continues in this pattern until a fixed number of waves are reached, where all links are traced from each wave.

VLTS has several clear advantages over the traditional RDS. First, it takes advantage of all auxiliary data sources to derive

Figure 3: Illustration of VLTS applied to a hidden population of sex workers in Muzzafarpur, India. RDS-traced links (left) and ex-post linking (right)

Source: Vincent, Zhang & Dank, 2019.

Note: The illustration on the left is the network sample generated by tracing links. The illustration on the right is the fully observed networked sample based on identifying nominations within the final sample.



a generous list of "entry points" for link-tracing. The idea of obtaining a widely distributed list of seeds is intended to achieve what most sampling methodologists tend to agree that it is always preferable to sample wide as opposed to sample deep. Second, VTLS typically sets the limit on the recruitment process to 2-3 waves, thus significantly increasing the efficiency by reducing the uncertainty of prolonged and undetermined numbers of waves of referrals. Third, by identifying the overlaps between respondents' networks, the time-tested mark-recapture techniques can be easily deployed to estimate population characteristics.

In short, link-tracing sampling strategies offer a rigorous alternative to conventional probability-based sampling when no sampling frames are available. While traditional RDS as originally developed by Hechathorn (1997) has been around for a few decades, some of its limitations have become apparent. VLTS, a much more recent development, offers attractive features that can improve RDS and produce more robust estimations.

Examples of RDS/Link Tracing

RDS has been regularly applied in prevalence estimation of human trafficking, within limited or specific geographical scopes. For instance, Curtis et al. (2008) applied RDS to estimate the prevalence of the commercial sexual exploitation of children (CSEC) in New York City. Jordan et al. (2020) applied RDS to estimate the CSEC population in Nepal. Zhang et al. (2014) used RDS to estimate the prevalence of labor trafficking among undocumented workers in San Diego County in the United States.

Reference List Of Useful Articles

Frank, O. and T. Snijders (1994). *Estimating the Size of Hidden Populations Using Snowball Sampling.* Journal of Official Statistics. Vol. 10, p.p. 53-67. Available at: https://bit.ly/2VW9PIO

Heckathorn, D. D. (1997). *Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations.* Social Problems. Vol. 44(2), p.p. 174-199. DOI: 10.2307/3096941

Heckathorn, D. D. (2002). *Respondent-Driven Sampling II: Deriving Valid Population Estimates from Chain-Referral Samples.* Social Problems Vol. 49, p.p. 11-34. DOI: 10.1525/sp.2002.49.1.11

Jordan, L., Chui, C.H., Larmar, S., & P. O'Leary (2020). *Overcoming methodological challenges in prevalence studies in developing contexts with vulnerable children.* International Social Work. Vol. 63(3), p.p. 371-385.

Seber, G. A. F. (1970). *The Effects of Trap Response on Tag Recapture Estimates. Biometrics, International Biometric Society.* Vol. 26, p.p. 13-22.

Thompson, S. (2020). *New Estimates for Network Sampling.* DOI: arXiv:2002.01350

Vincent, K. and S. Thompson(2017). *Estimating Population Size with Link-Tracing Sampling.* Journal of the American Statistical Association. Vol. 112, p.p. 1286-1295. DOI: 10.1080/01621459.2016.1212712

Vincent, K. (2019). *Recent Advances in Estimating Population Size with Link-Tracing Sampling.* DOI: arXiv:1709.07556v1

Vincent, K., Zhang, S. X., and M. Dank (2019). *Searching for Sex Trafficking Victims: Using a Novel Link-Tracing Method Among Commercial Sex Workers in Muzaffarpur, India.* Crime and Delinquency. DOI: 10.1177/0011128719890265

Zhang, S. X., Dank, M., Vincent, K., Narayanan, P., Bharadwaj, S. and S.M. Balasubramaniam (2019). *Victims without a voice: Measuring worst forms of child labor in the Indian State of Bihar.* Victims & Offenders, Vol.14(7), p.p. 832-858. DOI: 10.1080/15564886.2019.1660289

Zhang, S.X., Spiller, M.W., Finch, B.K. and Y. Qin (2014). *Estimating labor trafficking among unauthorized migrant workers in San Diego.* The Annals of the American Academy of Political and Social Science Vol. 653, p.p. 65-86. DOI: 10.1177/0002716213519237

11. Masking refers to respondents hiding valuable information or not responding truthfully. This can happen during surveys and interviews of victims of modern slavery due to the lack of trust between the victim and the enumerator.

12. Homophily refers to the tendency of individuals to associate with others with similar attributes (e.g., income and education levels, ethnicities or profession).

03

METHOD 03:
Time and Location Sampling

Andee Cooper Parks and Clifford Zinnes

Overview of Time-Location Sampling

Time-Location Sampling (TLS), or sometimes called time-space sampling and venue-based sampling, is a method originally rooted in public health research to study hard-to-reach populations. The method has been adapted for the anti-trafficking and modern slavery research in recent years. It is a probability-based method used to generate representative samples of location-based populations, reaching individual respondents at specific locations and times where they gather. TLS is typically employed in circumstances where a sampling frame of the target individuals is difficult to formulate or largely hidden and unknown. The main limitation of the method is the potential lack of access to certain locations (at the appropriate time/day) as well as the target population at the location. The findings may be limited to more accessible venues or more visible or active target population members.

High level explanation of the method

TLS was developed in the 1990s for use with collecting data on travelers to particular locations, such as museums or polling sites) and populations especially at risk of HIV/AIDS, such as men who have sex with men and injection drug users. The justification for the design of the method is that a sampling frame cannot be constructed for individuals, but one common characteristic among the population's patterns is that they can be found at specific, identifiable locations.

The method generally involves three stages:

1. Conducting formative research to identify the locations and days and times of visitation

2. Constructing the sample of locations, which includes mapping and often a two-stage sampling process

3. Collecting the data

The *formative stage* includes collecting qualitative data from key informants as well as any secondary data analysis that might help identify demographic or other characteristics about the total population.

The *sample construction stage* involves detailed mapping of the locations and the distribution of visitation frequency over the days of the week and times of day of the target population. This could entail headcount observation or informal interviews with venue bystanders (e.g. community members, or other workers) and other key stakeholders. This initial verification is key to ensuring that the target population is present, or visits happen, and to determining how many target-population members visit out of all visitors – and during which days/times. Knowledge of visitation habits are necessary for planning efficient enumeration and for ensuring representative sampling. If the formative research suggests that behaviors of interest may vary by strata of the target population visiting certain locations or only at particular times or days (e.g. younger girls only show up after 10pm, specific ethnicities are only present on Saturdays, male- vs. female-preferred venues, or even customer demographic differences), then this must be incorporated into the sampling design and plan.

The *data collection stage* involves either simple random sampling or stratified sampling of the locations (venue-day-time units or VDTs) where the probability of selection is proportional to the number of target population members at each location (probability proportional to size or PPS). The data are collected either through observation or through interviews with target

population members. In the former case, enumeration should include both the number of target population members and of all visitors. In the latter case, the first module of the data collection instrument (or survey questionnaire) would normally include a screening question to identify eligible respondents who would then complete the rest of the survey. In either case it is important to collect data on mobility and venue visitation habits of the respondent since these affect the probability of inclusion in the sample and, therefore, calculation of sampling weights Data on any security or safety concerns that might affect responses, including refusals and incomplete participation, should also be collected.

The TLS estimation method has several limitations. First, if the enumerators cannot access the venues selected for sampling on all necessary days and times or cannot access particular target population members that are present at the venue, then the findings will be skewed to only those that are accessible (for VDTs) or observable. If the venue itself is not accessible, then another venue should be selected from the same stratum of the sampling frame. If certain locations are not identifiable or approachable at the mapping stage, then the original sample may be biased – and, possibly, in unknown ways – towards more visible or accessible locations. If this occurs, there are two options. One is to qualify the findings as only being representative of a subset of the target population who visits the known and accessible locations in the sampling frame. If some characteristics of these missing VDTs are (roughly) known, then the other option is to apply multivariate probability models to adjust the sampling weights.

Second, TLS excludes any target population members who do not visit any identifiable locations at any time. Therefore, this method generally does not work well for street-based target populations, i.e., where individuals are not tied to specific locations but accessible anywhere, or for "private network" phenomena (e.g. exclusive access based on relationships of trust or private networks such as online commercial sexual exploitation of children).[13]

Third, while the results of TLS are reproducible over time, the researcher should use additional information and instrument questions to determine whether the phenomena under investigation remain time-invariant with respect to location. For example, it is challenging to rule out displacement as a cause of prevalence reduction in a specified area within mapped locations. Combining TLS with respondent-driven sampling (RDS) or capture-recapture methods can mitigate some of the method's limitations and estimate the size of a trafficked population in a given area.[14] Proper planning and validating of the sampling strategy and attention to logistical, regulatory, and ethical considerations are critical to success in implementation and unbiased prevalence estimates.

Assumptions and Other Considerations

TLS is relevant for hard-to-find populations like victims of modern slavery, but only if they can be found at locations within the sampling frame and exhibit stable visitation habits. In order to provide accurate, generalizable estimates with TLS, there are some important considerations that are *specific* to TLS to ensure accurate sampling, planning, and data collection. First, the data collectors need to be well trained on identifying the target population members and experienced with safety and security protocols when entering venues containing criminal practices. Second, re-validating that the sampled locations or target population members meet the eligibility criteria may be necessary in longitudinal studies since locations, characteristics, and other ways of the trade may have changed.

TLS is relevant for hard-to-find populations like victims of modern slavery, but only if they can be found at locations within the sampling frame and exhibit stable visitation habits.

Understanding patterns of social, economic, and legal influences is also key to ensuring accurate collection and comparable re-collections over time (e.g. holidays, pay days, seasonality, and law enforcement routines). As with all data collection modalities that entail a real-time interaction with victims, researchers will need to consider compliance with relevant ethics or any mandatory reporting requirements and ensure a response protocol for dire victim circumstances and psychosocial support for data collectors.

Examples of TLS Use

TLS has been utilized to estimate child sex trafficking, commercial sexual exploitation of children (CSEC), migrants, factory workers, and other most-at-risk populations (MARPs) such as men who have sex with men and injection drug users. While studies have been conducted all over the world using this method in HIV/AIDS-related research, so far this method has only been applied in the Philippines, Cambodia, and the Dominican Republic in the context of trafficking.

Reference List Of Useful Articles

Child-Trafficking-Related Studies

International Justice Mission (2015). *How to Measure the Prevalence of Child Sex Trafficking Expert Paper*. Washington, D.C., USA. Available at: https://bit.ly/37OmMnx

International Justice Mission (2013). *Commercial Sexual Exploitation of Children in Cambodia: A venue-based application of time-space sampling to measure prevalence in Phnom Penh, Siem Reap and Sihanoukville.* Washington, D.C., USA. Available at: https://bit.ly/33Qer1m

International Justice Mission (2016). *Child Sex Trafficking in Metro Manila: Using time-space sampling to measure prevalence of child sex trafficking in Metro Manila, the Philippines.* Washington, D.C., USA. Available at: https://bit.ly/3lXBei6

International Justice Mission (2016). *Child Sex Trafficking in Angeles City: Using time-space sampling to measure prevalence of child sex trafficking in Angeles City and Mabalacat in the Philippines.* Washington, D.C., USA. Available at: https://bit.ly/2K2yW3Q

International Justice Mission (2014). *Commercial Sexual Exploitation of Children in the Dominican Republic.* Washington, D.C., USA. Available at: https://bit.ly/37IMauM

Fedina, L. and B. R. DeForge (2017). *Estimating the Trafficked Population: Public-Health Research Methodologies May Be the Answer.* Journal of Human Trafficking. Vol. 3:1, p.p. 21-38. DOI: 10.1080/23322705.2017.1280316

A Sample of HIV/AIDS-Related Studies

Karon, J. M., and C. Wejnert (2012). *Statistical methods for the analysis of time-location sampling data.* Journal of Urban Health Vol. 89,p.p. 565–586. Available at: https://bit.ly/2K5QZ9b

Kendall, C., Kerr, L. R. F. S., Gondim, R. C., Werneck, G., Macena, R. H. M. M., Pontes, M. K., Johnston, L.G., Sabin, K. and W. McFarland (2008). *An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil.* AIDS and Behavior. Vol. 12(1), p.p. 97–104. DOI: 10.1080/23322705.2017.1280316

MacKellar, D. A., Gallagher, K. M., Finlayson, T., Sanchez, T., Lansky, A., and P. S. Sullivan (2007). *Surveillance of HIV risk and prevention behaviors of men who have sex with men – A national application of venue-based, time-space sampling.* Public Health Reports. Vol. 122(1), p.p. 39–47. DOI: 10.1177/00333549071220S107

Marpsat, M. and N. Razafindratsima (2010). *Survey methods for hard-to-reach populations: Introduction to the special issue.* Methodological Innovations Online. Vol. 5, p.p. 3–16. DOI: 10.4256/mio.2010.0014

Semaan, S. (2010). *Time-space sampling and respondent-driven sampling with hard-to-reach populations.* Methodological Innovations Online. Vol. 5, p.p. 60–75. DOI: 10.4256/mio.2010.0019

Fritz, K., McFarland, W., Wyrod, R., Chasakara, C., Makumbe, K., Chirowodza, A., Mashoko, C., Kellogg, T., and G. Woelk (2011). *Evaluation of a peer network-based sexual risk reduction intervention for men in beer halls in Zimbabwe: results from a randomized controlled trial.* AIDS and behavior. Vol. 15(8), p.p. 1732–1744. DOI: 10.1007/s10461-011-9922-1

Li, A., Varangrat, A., Wimonsate, W., Chemnasiri, T., Sinthuwattanawibool, C., & Phanuphak, P., Jommaroeng, R., Vermund, S. and F. van Griensven (2008). *Sexual Behavior and Risk Factors for HIV Infection Among Homosexual and Bisexual Men in Thailand.* AIDS and behavior. Vol. 13(2), p.p. 318-27. DOI: 10.1007/s10461-008-9448-3

Newman, P. A., Lee, S. J., Roungprakhon, S., and S. Tepjan (2012). *Demographic and behavioral correlates of HIV risk among men and transgender women recruited from gay entertainment venues and community-based organizations in Thailand: implications for HIV prevention.* Prevention science: the official journal of the Society for Prevention Research, Vol. 13(5), p.p. 483–492. DOI: 10.1007/s11121-012-0275-4

Parsons, J. T., Grov, C. and B.C. Kelly (2008). *Comparing the Effectiveness of Two Forms of Time-Space Sampling to Identify Club Drug-Using Young Adults.* Journal of drug issues. Vol. 38(4), p.p. 1061–1081.

A Sample of TLS Studies in Other Sectors

Ferreira, L.O.C., de Oliveira, E.S., Raymond, H.F., Chen, S.Y. and W. McFarland (2008). *Use of Time-location Sampling for Systematic Behavioral Surveillance of Truck Drivers in Brazil.* AIDS Behavioral. Vol. 12(4), p.p. 32–38. DOI: 10.1007/s10461-008-9386-0

Chew Ng. R. A., Muth S. Q. and C. L. Auerswald (2013). *Impact of social network characteristics on shelter use among street youth in San Francisco.* Journal of Adolescent Health. Vol. 53(3), p.p. 381–386. DOI: 10.1016/j.jadohealth.2013.04.018

McKenzie, D.J. and J. Mistiaen (2013). *Surveying Migrant Households: A Comparison Of Census-Based, Snowball, and Intercept Point Surveys.* Policy Research Working Paper No. 4419. World Bank, Washington, DC. Available at: https://openknowledge.worldbank.org/handle/10986/7634

Remafedi, G., Jurek, A. M. and J. M. Oakes (2008). *Sexual identity and tobacco use in a venue-based sample of adolescents and young adults.* American Journal of Preventive Medicine. Vol. 35(6), p.p. S463–S470. DOI: 10.1016/j.amepre.2008.09.002.

13. However, if the respondent both visits a venue as well as engages in street-based or a private-network activities then the inclusion in the data collection instrument of questions about the activity frequencies across venues and non-venues can potentially correct for this threat to internal validity.

14. For a discussion of hybrid design strategies see the Method 6: Hybrid Approaches chapter.

04

METHOD 04:
Network Scale-Up Method

Sam Blazek and Dennis Feehan

Overview of the Network Scale-Up Method

The network scale-up method (NSUM) is a method of estimating the size of hidden or hard-to-reach populations, such as victims of modern slavery, using measures of the size of personal networks of individuals in the general population. For example, if a respondent to a household survey knows 100 people, two of whom are survivors of human trafficking, it may be surmised that 2% of the total population are survivors of human trafficking. NSUM expands the logic behind this example to a larger sample pool and it adds features to try to account for both individual- and population-level biases in estimates.

High level explanation of the method

Though new to modern slavery research, NSUM has been widely used in public health literature to estimate the size of key populations at risk to HIV, such as number of intravenous drug users, men who have sex with men (MSM), and sex workers. Other previous applications of NSUM for hidden population size estimation include measuring the number of Americans directly and indirectly affected by the September 11, 2001 terrorist attacks,[15] estimating the prevalence of homelessness and sexual assault in the United States,[16] and estimating death rates in countries that lack vital registration systems.[17]

NSUM uses a sample of the general population, which could either be obtained through a standard household survey or a telephone-based survey. The sampled participants are asked a series of *"how many X do you know"* questions to gather data on their social network size and composition, where *X* corresponds to several sub-populations of known and unknown size (McCarthy et al., 2001). Groups for which sub-population sizes are known through measurements from other sources, such as administrative records or a population census, are used as reference groups. Unknown groups correspond to the target hidden populations of interest (e.g., individuals enduring forced labor conditions). The network data from all survey respondents is scaled up to estimate the sizes of the hidden populations overall. This can either be done using a simple design-based estimation or by applying statistical models to adjust for any differences between polled groups in the NSUM survey instrument (i.e., reference groups and hidden populations). In either case, NSUM does not require respondents to indicate who in their personal network belongs to which sub-population, allowing respondents to answer truthfully without repercussion or counting for variance in the size and structure of individual networks.

NSUM has unique benefits in estimating the number of individuals who are members of stigmatized and hidden groups. For one, it assists in reducing social desirability bias, because it does not target individuals who are members of these groups and does not ask respondents to expose their own membership in such groups. Furthermore, because NSUM does not require knowledge of the identities of a respondent's personal network members, respondents may feel more comfortable divulging the fact that they know members of stigmatized or hidden groups because respondents can be assured that they will not need to provide further information. Lastly, investigations that rely solely on community-driven methods, such as RDS, may miss isolated members of a hidden population, who may be common in some cases. By contrast, NSUM estimates that are obtained from general population survey response data can learn about hidden population members, regardless of how isolated or connected they are to one another.

An NSUM module can be added to a standard household or telephone survey, meaning that in many settings it may be possible to partner with ongoing social, health, or economic surveys to produce estimated numbers of victims of modern slavery. And, by designing the survey instrument carefully, it may be possible to provide estimates at different geographical scales (e.g., national, regional, provincial) from a single survey.

Assumptions and Other Considerations

The foundational implementation of NSUM by Killworth et al. (1998) bases its estimates on three assumptions:

1. Social ties are randomly formed between people in different groups, and people in different groups have, on average, personal networks of the same size;

2. Respondents are aware of their personal network members' characteristics and conditions; and

3. Respondents are willing and able to provide accurate information about their personal network members.

As with any model that relies on multiple assumptions and generalizations, several forms of bias may affect the quality of the data set and the resulting estimate. For instance, the Killworth model assumes that a theoretical respondent to an NSUM study would have a heterogeneous network, composed of an array of groups produced by random mixing. However, in reality, most networks are not the result of perfectly random mixing. Instead, community homophily is common and therefore *barrier effects*[18] exist.

In addition, respondents may be unaware of or unwilling to identify members of their personal network who are part of a hidden population, making NSUM prone to *transmission bias*[19] and *response bias*[20] respectively. Furthermore, NSUM may suffer from *recall bias*[21] as respondents may overestimate or underestimate the size of sub-populations in their networks. According to Maltiel et al. (2015), respondents tend to underestimate the number of people they know in larger groups because they forget some of these contacts, and to overestimate the number of people they know in small or unusual groups. A study that relies on NSUM estimates should carefully design and transparently describe its mitigation strategies for such biases. Some commonly used mitigation strategies for biases as well as other key considerations are shared below.

Transmission Bias

Several NSUM estimators exist that account for transmission bias. Some implementations of NSUM parameterize this bias explicitly, treating it as the result of an underlying distribution that may be quantified if an appropriate distribution is known or can be assumed. For example, Maltiel et al. (2015) introduced a Bayesian model that assumes all known populations have no transmission bias. However, in their study, Maltiel et al. find that there is little information in NSUM data to directly inform estimates of transmission bias – typically, if transmission bias exists, some external source of information about it is required.

The network data from all survey respondents is scaled up to estimate the sizes of the hidden populations overall.

Other approaches, such as the generalized scale-up estimator of Feehan and Salganik (2016) suggest combining NSUM data with additional data collected from a sample of hidden population members, such as might be obtained through RDS or TLS. This data collected from the hidden population members can be used to estimate and adjust for transmission bias in scale-up reports. Thus, this approach uses direct engagement with the hidden population in order to determine hidden population members' visibility. However, this additional data may be difficult for many researchers to collect in practice and may also introduce new biases related to the subjective perception of visibility. On the other hand, this conceptualization reframes transmission bias as a problem that can be solved through additional data collection.

Another approach to reducing transmission bias is to strategically choose the relationship that respondents are asked to report about. For example, rather than asking respondents about people they know, reports may be more accurate when respondents are asked about people they shared a meal with, people they work with, people who are their neighbors, or people they are connected to through some other social relationship (Feehan and Salganik, 2016). Previous work has shown that survey experiments can be used to test out survey questions based on different social relationships, providing a way to iteratively identify more promising approaches over time.

Response Bias

If response bias can be reasonably assumed to be present among one or more reference groups as well, then an NSUM model can 'learn' from the apparent response bias and better account for it in the hidden population estimation.

It may also be possible to mitigate response bias by strategically choosing the relationship respondents are asked to report about.

Recall Bias

Recall bias may be avoided through systematically defining network ties and reference groups so respondents are less likely to forget or inaccurately estimate connections to a particular group and can accurately recall information about their connections with ease. In most studies, this is achieved by targeting relatively small reference groups and hidden populations. For example, if a respondent knows between 15 and 20 members of a group, they may struggle to remember all of these connections; however, if a respondent is likely to only know 1 or 2 members of a group, they can more easily count their connections. For this reason, it is recommended that researchers performing NSUM evaluate whether the hidden populations under study are likely to be appropriately sized to minimize recall bias. If so, the researchers should also select reference groups that are generally within an order of magnitude of the anticipated size range of the hidden population. It may also be possible to minimize recall bias by strategically choosing the relationship respondents are asked to report about.

Barrier Effects

An NSUM model that treats the estimated personal network sizes of respondents as fixed can lead to an underestimate. Maltiel et al. developed a hierarchical model that can be used to improve uncertainty estimates by allowing them to reflect the impact of barrier effects. Other respondent-specific weighting approaches exist, including incorporating sampling weights to apply an overall adjustment to response data based on external data sources.

Definition of Network Ties

When asked about 'how many individuals do you know,' a respondent can interpret knowing' in many ways. For instance, an NSUM survey may ask respondents to consider their next-door neighbors or Facebook friends. In theory, a stricter or looser definition does not affect the estimate. Instead, it will merely result in upward or downward adjustments to the scale-up factor that determines the relationship between the personal network and full population sizes. However, in practice, certain definitions may produce more or less accurate responses from respondents. For example, many individuals in areas of lower internet connectivity may not have Facebook accounts and would be excluded from estimates that ask respondents specifically about their Facebook friends.

Network ties definitions may also have their own set of biases. For example, people may be willing to report on a survey that they have acquaintances who are known to inject drugs but they may not be very well informed about whether or not their acquaintances inject drugs. On the other hand, people may be aware of whether or not their close friends inject drugs, but they may be reluctant to admit that their close friends inject drugs on a survey. Empirical work, both qualitative and quantitative, can help researchers better understand which tie definitions are most promising for a given hidden population. The key considerations in selecting a definition are that the definition must be consistently applied by all respondents and should minimize the addition of external biases such as barrier effects.

Examples of NSUM Use in Modern Slavery Research

Applying NSUM to modern slavery research has only just begun. A recent study conducted by IST Research and Kantar used a combination of NSUM and RDS to estimate the total population of sex workers and child sex trafficking (CST) victims in Maharashtra, India. Additionally, a research design document authored by NORC at the University of Chicago has outlined the use of NSUM for the estimation of modern slavery across three sectors in South Asia – apparel, overseas labor recruitment, and commercial sexual exploitation.

Reference List Of Useful Articles

Blazek, S., Campbell, A., Shah, M., Ravindran, S. and S. Lyons (2020). *Estimating the Prevalence of Child Sex Trafficking in Maharashtra, India.* IST Research. Fredericksburg, VA, USA.

Crawford, F., E. Keaveney, K. Kysia, S. Islam, T. Mittelberg, S. Blazek and C. Zinnes (2019). *Prevalence Design Report:* Chapters 3 - 5. NORC at the University of Chicago. Chicago, IL, USA (Unpublished deliverable funded by the Prevalence Estimation Research Program of the Global Fund to End Modern Slavery).

Ezoe S., Morooka T., Noda T., Sabin M.L. and S. Koike (2012). *Population size estimation of men who have sex with men through the network scale-up method in Japan.* PLoS One.Vol. 7(1). DOI: 10.1371/journal.pone.0031184

Feehan, D. M., and M. J. Salganik (2014). *Network reporting package (Version 0.1.1).* R Foundation for Statistical Computing. Available at: https://bit.ly/3owayXi

Feehan, D. M., and M. J. Salganik (2016). *Generalizing the Network Scale-Up Method: a new estimator for the size of hidden populations.* Sociological Methodology. Vol. 46(1), p.p. 153–186. DOI: 10.1177/0081175016665425

Guo, W., Bao, S., Lin, W., Wu, G., Zhang, W., Hladik, W., Abdul-Quader, A., Bulterys, M., Fuller, S. and L. Wang (2013). *Estimating the size of HIV key affected populations in Chongqing, China, using the Network Scale Up Method.* PloS one. DOI: 10.1371/journal.pone.0071796

Jing, L., Lu, Q., Cui, Y., Yu, H. and T. Wang (2018). *Combining the randomized response technique and the network scale-up method to estimate the female sex worker population size: an exploratory study.* Public Health. Vol. 160, p.p. 81-86. DOI: 10.1016/j.puhe.2018.04.002

Kadushin, C., Killworth, P. D., Bernard, H. R. and A. A. Beveridge (2006). *Scale-up methods as applied to estimates of heroin use.* Journal of Drug Issues. Vol. 6, p.p. 417-440. DOI: 10.1177/002204260603600209

Killworth, P. D., McCarty, C., Bernard, H. R., Shelley, G. A. and E. C. Johnsen (1998). *Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach.* Evaluation Review. Vol. 22(2), p.p. 289–308. DOI: 10.1177/0193841X9802200205

Lewis, K., Gonzalez, M. and J. Kaufman (2012). *Social selection and peer influence in an online social network.* Proceedings of the National Academy of Sciences. Vol. 109(1), p.p. 68–72. DOI: 10.1073/pnas.1109739109

Maltiel, R., Raftery, A. E., McCormick, T. H. and A. J. Baraff (2015). *Estimating population size using the Network Scale Up Method.* Annals of Applied Statistics. Vol. 9(3), p.p. 1247–1277. DOI: 10.1214/15-AOAS827

McCarty, C., Killworth, P. D., Bernard, H. R., Johnsen, E. C. and G. A. Shelley (2001). *Comparing two methods for estimating network size.* Human Organization. Vol. 60, p.p. 28–39. DOI: 10.17730/humo.60.1.efx5t9gjtgmga73y

Salganik, M. J., Fazito, D., Bertoni, N., Abdo, A. H., Mello, M. B. and F. I. Bastos (2011). *Assessing network scale-up estimates for groups most at risk of HIV/AIDS: Evidence from a multiple-method study of heavy drug users in Curitiba, Brazil.* American Journal of Epidemiology. Vol. 174(10), p.p. 1190–1196. DOI: 10.1093/aje/kwr246

Shelley, G. A., Bernard, H. R., Killworth, P., Johnsen, E. and C. McCarty (1995). *Who knows your HIV status? What HIV+ patients and their network members know about each other.* Social Networks. Vol. 17(3–4), p.p. 189–217. DOI: 10.1016/0378-8733(95)00262-M

Shelley, G. A., Killworth, P. D., Bernard, H. R., McCarty, C., Johnsen, E. C. and R. E. Rice (2006) *Who knows your HIV status II? Information propagation within social networks of seropositive people.* Human Organization. Vol. 65(4), p.p. 430–444. DOI: 10.17730/humo.65.4.08mwg9d3nfy8w9th

Shokoohi, M., Baneshi, M. R., and A. A. Haghdoost (2012). *Size Estimation of Groups at High Risk of HIV/AIDS using Network Scale Up in Kerman, Iran.* International Journal of Preventive Medicine. Vol. 3(7), p.p. 471–476. Available at: https://bit.ly/2VQ00Gc

Sulaberidze, L., Mirzazadeh, A., Chikovani, I., Shengelia, N., Tsereteli, N. and G. Gotsadze (2016). *Population size estimation of men who have sex with men in Tbilisi, Georgia: Multiple methods and triangulation of findings.* PLoS One. Vol. 11(2). DOI: 10.1371/journal.pone.0147413

Teo, A. K. J., Prem, K., Chen, M. I., Roellin, A., Wong, M. L., La, H. H. and A. R. Cook (2019). *Estimating the size of key populations for HIV in Singapore using the Network Scale Up Method.* Sexually Transmitted Infections. Vol. 95(8), p.p. 602–607. DOI: 10.1136/sextrans-2018-053747

Zheng, T. Z., Salganik, M. J. and A. Gelman (2006). *How many people do you know in prison? Estimating overdispersion in count data to estimate social structure in networks.* Journal of the American Statistical Association. Vol. 101(474), p.p. 409–423. DOI: 10.1198/016214505000001168

15. Bernard, H. R., Killworth, P. D., Johnsen, E. C., Shelley, G. A., and C. McCarty (2001). Estimating the ripple effect of a disaster. Connections. Vol. 24(2), p.p. 18–22. Available at: https://bit.ly/3qFHfU0

16. Killworth, P. D., McCarty, C., Bernard, H. R., Shelley, G. A., and E. C. Johnsen (1998). Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. Evaluation review. Vol. 22(2), p.p. 289–308. DOI: 10.1177/0193841X9802200205.

17. Feehan, D.M., Mahy, M. and M. J. Salganik (2017). The Network Survival Method for Estimating Adult Mortality: Evidence from a Survey Experiment in Rwanda. Demography. Vol. 54, p.p. 1503–1528. DOI: 10.1007/s13524-017-0594-y.

18. While models based on NSUM assume personal networks are randomly mixed and heterogeneous, individual personal networks tend to be homogeneous (i.e., people tend to know other people like themselves). Additionally, hidden populations such as sex workers or victims of sex trafficking are often not randomly distributed in the general population, meaning some respondents have a higher propensity to know people who are sex workers or sex trafficking victims. If the sample systematically excludes those more likely to know members of the hidden population, an underestimate would result. However, if the sampling frame systematically excludes those less likely to know members of the hidden populations, an overestimate would result. The likely magnitudes of these problems are unknown. Thus, it is critical to obtain a representative sample of personal networks from the study geography.

19. Transmission bias occurs when respondents may know people who are members of a hidden population, but not realize they are members of the hidden population. Because the populations under study are often stigmatized, failure to account for this lack of knowledge may lead to an underestimate of unknown size.

20. Response bias occurs when respondents know people who are members of a hidden population, but they are unwilling to provide this information because of the possible stigma involved. This phenomenon could result in an underestimate of that sub-population's size.

21. Recall bias is an error resulting from the participants' inaccurate recollection of their social network.

05

METHOD 05:
Multiple Systems Estimation/ Capture Recapture

Davina Durgana and Jan Van Dijk

Overview of Multiple Systems Estimation/Capture Recapture

Multiple Systems Estimation (MSE) is a statistical technique that uses the comparison of concurrent and identifiable victim administrative lists that are typically held and maintained by various national government offices and agencies to produce prevalence estimates for modern slavery. MSE builds upon the classic capture-recapture method and applies it to lists of victim data to estimate the total population of slavery victims. Estimates are based on how often certain victims appear on more than one list within a certain time period. This method is most applicable in the context of countries with advanced statistical systems where victim lists in the form of administrative data are more common.

High level explanation of the method

Conducting household surveys in high-income countries can be costly and have limitations in the identification of modern slavery prevalence. MSE is a viable and cost-effective alternative when seeking to quantify modern slavery at national and sub-national levels using only readily available administrative data that is maintained by official sources.

MSE is a valuable option for prevalence estimation for several reasons. First, it allows governments to take leadership positions in the measurement and analysis of modern slavery prevalence based on existing administrative data. Second, most estimates of modern slavery are static and provide a prevalence estimate at a point in time, when the study was conducted. MSE, on the other hand, can be dynamic and updated from year to year with relative ease. Third, as an important component of advanced prevalence estimation techniques, MSE provides necessary insights into how capable national systems are in identifying victims of modern slavery for particular demographics and types of exploitation.

By applying probability theory to known values of individuals that appear in more than one list, MSE estimates the overall victim population size. To illustrate this with a simple capture-recapture example, let's assume there are two lists of victims. List A has five victims and list B has ten victims, and 2 victims are recorded in both lists A and B. If the total victim population is N, then the probability of a victim being in list A is 5/N and the probability of a victim being in list B is 10/N. The probability of being in both lists is 2/N. Then if we estimate that the total victim population N equals the number of victims in List A (5 victims) times the number of victims in List B (10 victims) then divided by the number of victims on both lists which is 2 (2 victims). (List A * List B)/ Number of Victims on both Lists = Total Estimated Population or N. In this example, we can then estimate that N = (5*10)/2 = 25 or N = 25. This is the same simple ratio that underpins the Petersen-Lincoln population estimator. The preference in conducting MSE is to have a minimum of three distinct lists. Figure 4 visualizes how MSE uses list overlaps to estimate total population size.

The process of MSE is equally about the exercise of gathering and sharing administrative data that has been matched for victims as it is about empowering national governments to take ownership over their responsibility to measure prevalence. These exercises are also important in order to build the foundations for these collected data to be improved

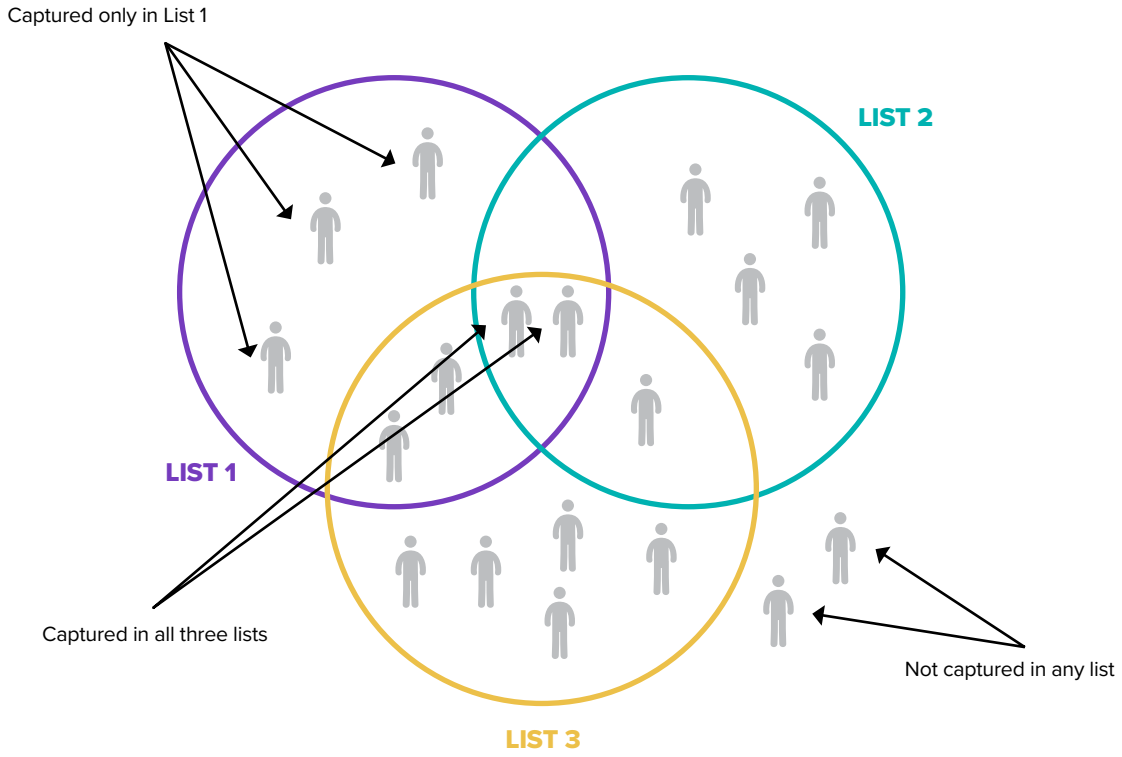


Figure 4: How MSE works

Source: Adapted from illustration by Jana Asher, Human Rights Data Analysis Group

Note: MSE uses probability theory to estimate the size of the total population of victims using known values of the number of individuals that have been captured in one or more administrative lists.[22]

upon and for these estimates to be re-analyzed with more and better data. Experiences in several countries suggest that using data from an organization that maintains a victim administrative list for MSE entices other organizations to share their data as well, which results in increasing the accuracy of estimates over time.

Assumptions and Other Considerations

In order to conduct MSE based on a country's administrative data on victims of modern slavery, the following conditions must be met:

1. **The administrative data on victims of modern slavery must be organized into at minimum three distinct lists originating from different agencies.** For example, one list of victims can be from law enforcement, another list of victims from social services, and another list of victims from immigration assistance. It is important that each list has equal potential for victims to be listed, and thus appear multiple times across lists, but there should not be a formal mechanism that refers 100% of victims directly from one organization to another and subsequently results in two identical lists, which would invalidate one of those lists for use in MSE. There also must be some overlap across the lists.

2. **There must be a unique identifier for the victims of modern slavery on every list** such that individuals that feature on multiple lists can be identified. Examples of unique identifiers are social security numbers, national identity numbers or any unique combination of birth date, industry of exploitation, gender, and sometimes initials among other identifiers that allow for reasonable determination of how frequently the same victims appear on multiple victim databases.

3. **Victim information should include their key attributes along with their unique identifier.** Some examples of these key attributes are industry of exploitation, age group of the victims, gender of the victims, nationality of the victims,[23] or any other demographic information that might be available to develop covariates that enable country-level sub-population estimates.

The process of MSE is equally about the exercise of gathering and sharing administrative data that has been matched for victims as it is about empowering national governments to take ownership over their responsibility to measure prevalence.

These requirements mean that MSE has largely been successful as an estimation method in countries with advanced statistical systems. Even among countries with more advanced data infrastructure there are still many countries have significant gaps in their collection and curation of modern slavery data. This can either be due to data ecosystem deficiencies or due to national legal contexts that do not yet identify and recognize issues such as forced marriage or other components of modern slavery as illegal. As such, MSE can only be as strong as the data that are integrated into these estimates. As advocacy efforts as well as use of administrative data for modern slavery research increase across the globe, MSE will become more viable for a wider range of governments and estimates derived from MSE will continue to be refined for greater prevalence accuracy.

Examples of where the method has been used

MSE is an emerging area of research that requires significant investment to further develop the method and to overcome obstacles that are encountered in the commission of these estimates in its early stages.

Some of the most noteworthy examples of the applications of MSE to estimate modern slavery prevalence include studies conducted by the UN Office on Drugs and Crime and Walk Free in Ireland, Romania, Serbia, and Slovakia. In 2014, the United Kingdom also used MSE to estimate the prevalence of human trafficking in the country.

Reference list of useful articles

Anderson, V.R., Kulig, T.C. and C. J. Sullivan (2019). *Estimating the Prevalence of Human Trafficking in Ohio, 2014-2016.* American Journal of Public Health. Vol. 109(10), p.p. 1396-1399. DOI: 10.2105/AJPH.2019.305203

Bales, K., Hesketh, O. and B. W. Silverman (2015). *Modern Slavery in the UK: How many victims?* Significance Vol. 12, p.p. 16-21B. DOI: 10.1111/j.1740-9713.2015.00824.x

Bales, K.O., Murphy, L.T. and B. W. Silverman(2020). *How many trafficked people are there in Greater New Orleans: Lessons in Measurement.* Journal of Human Trafficking. Vol. 6(4), p.p. 375-385. DOI: 10.1080/23322705.2019.1634936

Bird, S. M. and R. King (2018). *Multiple Systems Estimation (or Capture-Recapture Estimation) to Inform Public Policy.* Annual Review of Statistics and its Application. Vol. 5, p.p. 95–118. DOI: 10.1146/annurev-statistics-031017-100641

Chan, L., Silverman, B. W., and K. Vincent (2020). *Multiple Systems Estimation for Sparse Capture Data: Inferential Challenges When There Are Non-overlapping Lists.* Journal of the American Statistical Association. Online. DOI: 10.1080/01621459.2019.1708748

Cruyff, M., van Dijk, J.J.M., P. van der Heijden (2017). *The Challenge of Counting Victims of Human Trafficking, Not on the record: A multiple systems estimation of the numbers of human trafficking victims in the Netherlands in 2010–2015 by year, age, gender, and type of exploitation.* CHANCE. Vol. 30(3), p.p. 41–49. DOI: 10.1080/09332480.2017.1383113

van Dijk, J.J.M. and C. H. D. Steinmetz (1979). *The RDC victimization surveys 1973-1978/De WODC-slachtofferenquetes 1973-1978.* Ministry of Justice/Ministerie van Justitie/Staatsuitgeverij. Available at: https://bit.ly/3qEqdpq

van Dijk, J.J.M. (2008). *The World of Crime: breaking the silence on the problems of security, justice and development across the world.* Sage Publications. DOI: 10.4135/9781483329789

UNODC (2016). Research Brief. *Multiple Systems Estimation for estimating the number of victims of human trafficking across the world.* United Nations. Vienna, Austria. Available at: https://bit.ly/2LbgdUg

Durgana, D. P. and P. L. Zador (2017). *Fighting Slavery through Statistics: A Discussion of Five Promising Methods to Estimate Prevalence in the United*

States. CHANCE. Vol. 30(3). DOI: 10.1080/09332480.2017.1383114

Committee on National Statistics and Committee on Population (2019). *Estimating the Prevalence of Human Trafficking in the United States: Considerations and Complexities: Proceedings of a Workshop - in Brief.* The National Academies Press. Available at: https://bit.ly/2LkUZ6D

Durgana, D. P. and J.J.M. van Dijk (Forthcoming). *Multiple Systems Estimation in Developed Countries: Best Practices and Expected Deviation across Countries of Origin and Destination.*

Eurostat (2015). *Trafficking in human beings.* European Union. Luxembourg. Available at: https://bit.ly/2Itd7tX

Farrell, A., Dank, M., Kafafian, M., Lockwood, S., Pfeffer, R., Hughes, A. and K. Vincent (2019). *Capturing Human Trafficking Victimization through Crime Reporting.* NCJRS, Office of Justice Programs. Available at: https://bit.ly/3lX2bCo

International Labour Office (2009). *Operational indicators of trafficking in human beings: results from a Delphi survey implemented by the ILO and the European Commission.* ILO. Geneva, Switzerland. Available at: https://bit.ly/3mXavDH

Lyneham, S., Dowling, C. and S. Bricknell (2019). *Estimating the dark figure of human trafficking and slavery victimisation in Australia.* Statistics Bulletin 16. Australian Institute of Criminology. Australian Government. Available at: https://bit.ly/3lTbHq5

Lynch, J.P. and L. A. Addington (Eds.) (2007). *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and the UCR.* Cambridge Studies in Criminology. Cambridge University Press. ISBN-10: 0719572053

Penney, T.L. (2014). *Dark Figure of Crime (Problems of Estimation).* Encyclopedia on Criminology and Criminal Justice. DOI: 10.1002/9781118517383.WBECCJ248

Rojas, D. Q., Dufraix Tapia, R. D. and R. R. Rodríguez (2019). *Human Trafficking Cases in Chile: Challenges for Reducing the "Dark Figure."* The Palgrave International Handbook of Human Trafficking, p.p. 1151-1164. Springer.

Silverman, B. (2014). *Modern Slavery: an application of Multiple Systems Estimation.* Home Office. United Kingdom. Available at: https://bit.ly/3gtgUUM

Spierenburg, P. (2016). The rise of Criminology in its historical context, Knepper, P. and A. Johansen (Eds). The Oxford Handbook on the History of Crime and Criminal Justice, Chapter 19, p.p. 373- 395. DOI: 10.1093/oxfordhb/9780199352333.013.20

UNODC (2015). Global Report on Trafficking in Persons 2014. United Nations. New York, USA. Available at: https://bit.ly/36XAJQY

UNODC and National Rapporteur on Trafficking in Human Beings and Sexual Violence against Children. (2017). *Monitoring Target 16.2 of the United Nations Sustainable Development Goals: A multiple systems estimation of the numbers of presumed human trafficking victims in the Netherlands in 2010-2015 by year, age, gender, form of exploitation and nationality.* Research Brief. UNODC Research. Vienna, Austria. Available at: https://bit.ly/3lVsusQ

UNODC and Walk Free Foundation (2018). *Monitoring Target 16.2 of the United Nations Sustainable Development Goals: multiple systems estimation of the numbers of presumed victims of trafficking in persons. Ireland.* Research Brief. UNODC Research. Vienna, Austria. Available at: https://bit.ly/3n01fOS

UNODC and Walk Free Foundation (2018). *Monitoring Target 16.2 of the United Nations Sustainable Development Goals: multiple systems estimation of the numbers of presumed victims of trafficking in persons. Romania.* Research Brief. UNODC Research. Vienna, Austria. Available at: https://bit.ly/3qDChar

UNODC and Walk Free Foundation (2018). *Monitoring Target 16.2 of the United Nations Sustainable Development Goals: multiple systems estimation of the numbers of presumed victims of trafficking in persons. Serbia.* Research Brief. UNODC Research. Vienna, Austria. Available at: https://bit.ly/37MplRt

22. For further details, see: Green, A.H. (2013) Multiple Systems Estimation: Stratification and Estimation. Available at: https://bit.ly/3lV4nKz

23. Designations as simple as citizens or foreign nationals can also provide researchers with valuable insights.

06

METHOD 06:
Hybrid Approaches

Kareem Kysia, Erika Keaveney and Clifford Zinnes

Overview of Hybrid Approaches

Policymakers, donors, and stakeholders have come to recognize that it is difficult to develop cost-effective policy responses to eradicate modern slavery without having an accurate and reliable measure of its prevalence within populations. Researchers have, therefore, been developing and refining sampling methodologies and *ex post* analyses. Among these are, as discussed in previous chapters, stratified/multistage probability sampling proportionate to size (S-PPS),[24] response-driven sampling (RDS), time-location sampling (TLS), network scale-up methodologies (NSUM, M-NSUM, and G-NSUM), and multiple systems estimation (MSE), which includes capture/recapture techniques (CRT). Although these methods offer promising ways to estimate prevalence, each method has strengths and weaknesses. While it is often possible to implement several approaches at once and tell a story to triangulate the resulting estimates, this chapter summarizes when and how these methods can be *combined* under one empirical strategy to reach a better prevalence estimate for a given target population.

High-level explanation of the method

There are many opportunities to combine sampling methods to mitigate a weakness of one's preferred approach by drawing on the strength of another approach. The following table suggests some of the potential combinations that can be used in modern slavery research.

Table 1: Potential method combinations that can be used for modern slavery research.

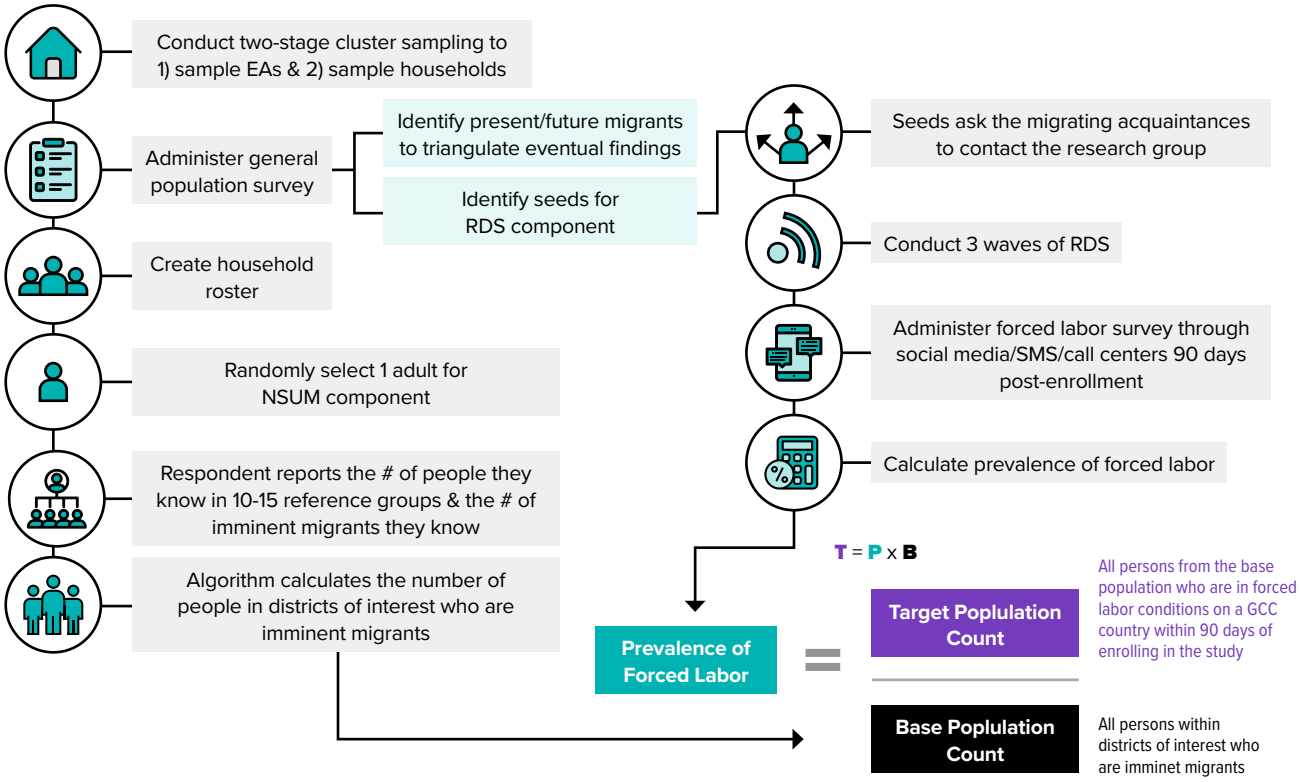
Sampling method	Sampling method			
	S-PPS	NSUM	RDS	TLS
S-PPS				
NSUM	X			
RDS	X	X		
TLS		X	X	
MSE	X		X	X

Source: Authors' Illustration

Each combination has its merits and demerits, applications and complementarities, which would require a longer article to cover adequately. However, a high-level overview of some of the most widely used combinations is feasible to present in consolidated form and are noted below.

RDS and an S-PPS-based population survey.
There are two central yet related problems facing RDS. First, a source of initial "seeds" is required with which to initiate the network chain-linking. Second, the equilibrium prevalence rate, to which RDS converges, is not necessarily statistically representative. These problems may be simultaneously addressed by conducting a modest-sized S-PPS-based population survey (e.g., as would be administered for NSUM or TLS). Then, a randomly selected set of the latter survey's respondents can be used as seeds to start the RDS sampling process. Since by design the sample drawn for the population survey was representative, any random sub-sample of it will also be representative (mitigating the issue of non-representativeness within an RDS approach). There is no need to wait for RDS waves to converge and the sampling process can use other factors (e.g., cost or time) to decide on the optimal composition of seeds and referrals to include in the RDS sample.[25] This methodology is being utilized for a GFEMS CSEC-prevalence study in Bangladesh and one in India.[26]

Figure 5. Implementation of a hybrid approach to migrant forced labor



Source: Crawford, F., Keaveney, E., Kysia, K., Islam, S., Mittelberg, T., Blazek, S. and C. Zinnes (2019). "Prevalence Design Report", Chapter 3, Figure 1, NORC at the University of Chicago (Unpublished deliverable funded by the Prevalence Estimation Research Program of the Global Fund to End Modern Slavery).

NSUM and RDS: Example 1

NSUM, though widely used in practice, may result in several types of biases due to the assumptions required to use it. These assumptions include: (1) social ties are formed completely at random (i.e., random mixing); (2) respondents are perfectly aware of the characteristics of their acquaintances ("alters" in network-speak); and (3) respondents are able to provide accurate answers to survey questions about their personal networks.[27] While the details of how to mitigate these threats to internal validity are beyond the present scope, the reasons behind reducing these biases are: (1) to draw on the fact that acquaintance "visibility" is symmetric (if X knows Y is in forced labor then Y knows that X knows this); and (2) to strategically select the relational reference populations (the so called "probe alters") required by NSUM.[28]

Once threat mitigation has been addressed through the research design, an NSUM survey can be used to infer the frame population's share of all its acquaintances that are in the hidden population (as traditionally done in scale-up studies) and questions can be added to the RDS survey to infer the hidden population's beliefs about the share of their acquaintances in the frame population rate that is aware of their membership in the hidden population. This permits the calculation of a new, unbiased estimator — the generalized scale-up estimator or G-NSUM — for hidden population size.

RDS and NSUM: Example 2

As we see above, choice of method depends on whether one wants to know the scale or the impact of the issue being studied. What if you want *both?* By combining count- and share-specific approaches (i.e., RDS and NSUM) in innovative ways the researcher can often cost-effectively estimate both the count and the share of those in the hidden population of interest.

TLS and RDS

One of the challenges to the new sampling approaches, described in this document, is that an initial set of respondents (seeds) is required. For example, RDS may be the technique of choice, but where does the researcher find the requisite set of seeds (or migrants to track)? This problem can be solved by first administering a TLS population-based survey in locations known to harbor or be

visited by the target population. A subset (or all) members of the target population found (i.e., satisfying the identification criteria) within the resulting sample can then be used for the RDS seeds.

Assumptions and other considerations

Each prevalence calculation carries its own informational constraints, cost implications to achieve a given statistical precision, threats of biases, degree of respondent acceptance, and level of credibility. Furthermore, the relative importance of these aspects is dependent on the form of modern slavery targeted for measurement. For example, the degree of a group's visibility ("hiddenness") greatly influences the appropriate choice of sampling methodology: survivors of commercial sexual exploitation of children (CSEC) are more hidden and socially stigmatized than survivors of a forced-labor situation. The relative "hiddenness" of each group is essential for identifying the appropriate prevalence estimation methodology. NSUM may be preferable for the former since it is less incriminating. If the target population is believed to be a very small share of the base population, RDS or MSE will likely be much less costly than S-PPS (with RDS also reducing the time required relative to S-PPS). Likewise, when there is no meaningful sampling frame of the targeted (victim) population, this may be overcome using RDS or MSE.

The choice of method also depends on whether one is interested in a prevalence count or a share. A prevalence count refers to the number of units experiencing a condition in the base population at a given time and location (e.g., the number of migrants in a certain location and time under forced labor conditions). A prevalence ratio (rate or share) refers to the proportion of the base population experiencing a condition at or over a given time and location (e.g., the proportion of all migrants at a certain location and time under forced labor conditions). On the one hand it is good to know that there are, for example, 10,000 individuals under forced-labor conditions at a given time and place. However, it is equally important to know if that means 2% of the base population or 50% of the base population are affected in order to know the scale and impact of the issue being studied. Some prevalence methods are appropriate for counts (NSUM for example) and others more appropriate for ratios (RDS for example). However, by combining the two approaches in innovative ways the researcher can cost-effectively estimate both a count and a ratio of the population of interest.

Reference List of Articles Utilizing Hybrid Methods

TLS and RDS

Kendall, C., Kerr, L. R. F. S., Gondim, R. C., Werneck, G. L., Macena, R. H. M., Pontes, M. K., Johnston, L. G., Sabin, K. and W. McFarland (2008). *An Empirical Comparison of Respondent-driven Sampling, Time Location Sampling, and Snowball Sampling for Behavioral Surveillance in Men Who Have Sex with Men, Fortaleza, Brazil.* Aids and Behavior. Vol. 12(97). DOI: 10.1007/s10461-008-9390-4

Crawford, F., Keaveney, E., Kysia, K., Islam, S., Mittelberg, T., Blazek, S. and C. Zinnes (2019). *"Prevalence Design Report", Chapter 4.* NORC at the University of Chicago (Unpublished deliverable funded by the Prevalence Estimation Research Program of the Global Fund to End Modern Slavery).

The cited study does not have any publicly available documentation since it is currently ongoing. Details on the research methodology and findings will be published in the near future.

Raymond, H.F., Chen, Y.H. and W. McFarland (2019). *'Starfish Sampling': a Novel, Hybrid Approach to Recruiting Hidden Populations.* Journal of Urban Health. Vol. 96, p.p.55–62. DOI: 10.1007/s11524-018-0316-9

NSUM and RDS (and G-NSUM)

Crawford, F., Keaveney, E., Kysia, K., Islam, S., Mittelberg, T., Blazek, S. and C. Zinnes (2019). *"Prevalence Design Report", Chapter 5.* NORC at the University of Chicago (Unpublished deliverable funded by the Prevalence Estimation Research Program of the Global Fund to End Modern Slavery).

The cited study does not have any publicly available documentation since it is currently ongoing. Details on the research methodology and findings will be published in the near future.

TLS and NSUM

Crawford, F., Keaveney, E., Kysia, K., Islam, S., Mittelberg, T., Blazek, S. and C. Zinnes (2019). *"Prevalence Design Report", Chapter 4.* NORC at the University of Chicago (Unpublished deliverable funded by the Prevalence Estimation Research Program of the Global Fund to End Modern Slavery).

The cited study does not have any publicly available documentation since it is currently ongoing. Details on the research methodology and findings will be published in the near future.

25. Perhaps, paradoxically, the more RDS rounds that are administered, the greater the risk of moving further *away* from a representative sample.

26. NORC (2019, forthcoming). "Prevalence Design Report",2nd draft, Global Fund to End Modern Slavery, 19 November.

27. Feehan, D. M. and M. J. Salganik (2016). *Generalizing the Network Scale-Up Method: A New Estimator for the Size of Hidden Populations.* Sociological Methodology. Vol. 46(1), p.p. 153-186. DOI: 10.1177/0081175016665425

28. Specifically, the probe-alter acquaintance types should be chosen so that the rate at which the hidden population is visible to the probe alters be the same as the rate at which the hid-den population is visible to those selected for NSUM interviews (the "frame population").

CONCLUDING THOUGHTS

The purpose of this document is to provide an introductory overview of the various methods of prevalence estimation that are being used by experts working in the modern slavery field. More than ten expert researchers and methodologists have contributed a synthesis of their knowledge and experience in the form of six brief chapters that have covered household surveys, respondent driven sampling or link tracing sampling, time and location sampling, network scale-up method, multiple systems estimation, and hybrid methods to estimating the prevalence of modern slavery. Within each chapter, the experts have also shared a series of reference publications that would provide readers with deeper knowledge of these methods.

It is our hope that this document offers continued utility as a quick reference guide to readers - from donors to policy makers, researchers to practitioners, and other communities that are actively engaging on the issue of modern slavery. Nonetheless, there are some key considerations that readers should be mindful of when using this knowledge.

First, the non-traditional methods that are presented in this document are attractive for their ability to reach hidden or hard-to-reach populations, such as victims of modern slavery. However, each of these methods are based on a series of empirical and logical assumptions, that should be scrutinized and well understood prior to making decisions on research design. Any statistical estimates derived from one of these methods should also be accompanied by relevant caveats in order to avoid situations where users draw inaccurate conclusions.

Second, it is worth noting that modern slavery is a new field of study and is constantly evolving. This means that, unlike public health where the definitions of key concepts and their associated methods of measurement have become consistent and standardized over a century of practice, definitions and measurement methods in modern slavery are still making progress towards building the necessary global consensus. It is, therefore, strongly advised to be mindful of conceptual and definitional differences when designing research, conducting analyses, and communicating results.

And third, prevalence estimation is a critical step towards building the necessary evidence base that facilitates better policy formation and resource mobilization. It is worth recognizing, however, that prevalence estimation is one of a variety of research methods and techniques that generate evidence in the modern slavery space. There are a rich variety of additional quantitative and qualitative methods that are also important to facilitate understanding and spur action towards the elimination of this human rights abuse. It is the compilation of multiple, complementary research efforts that is key to building comprehensive narratives on the state of modern slavery and how to remedy it.

With a new generation of social scientists, statisticians, data scientists, and practitioners joining the efforts to end modern slavery, the arsenal of research techniques and data collection and analysis tools for prevalence estimation continues to evolve. Modern computing-based social media analytics and other technologies such as satellite and drone imagery are some examples of new and cost-effective sources of data on modern slavery. Nonetheless, the value of careful research design and informed methodological choices remains timeless. It is in this spirit of sharing the knowledge of current best-practices and inviting collaborations to pave the future of modern slavery research that the authors share this document.

GLOBAL
FUND
TO
END
MODERN
SLAVERY