This report was prepared by Abt Global, LLC using federal funding provided by the Bureau of Justice Statistics.

Document Title: Human Trafficking Prevalence Estimation Feasibility Study

Author(s): Shelby Hickman, PhD, Abt Global, LLC
Charlotte Lopez-Jauffret, PhD, formerly of Abt Global, LLC
Stephanie Fahy, PhD, Abt Global, LLC
Vanessa Masick, Abt Global, LLC

Document No.: NCJ 309951

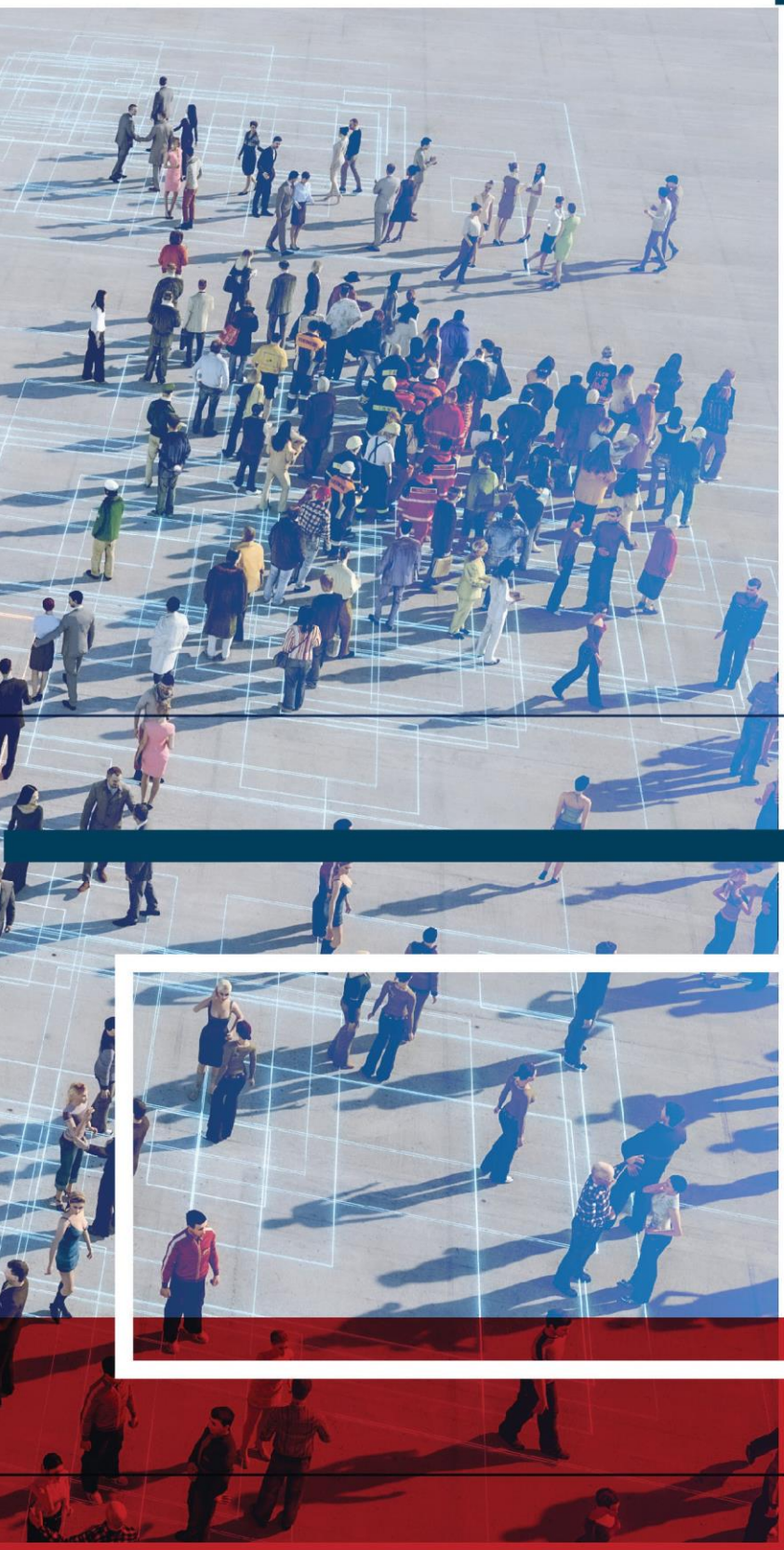
Publication Date: January 2025

Award No.: This project was supported by Cooperative Agreement No. 2019-85 CX-K002

Abstract:
Abt Global (Abt) conducted a study for the Bureau of Justice Statistics to identify and recommend methods for estimating the prevalence of human trafficking in the United States. This study supports BJS's mission to collect, analyze, publish, and disseminate information on crime, criminal offenders, victims of crime, and the operation of justice. As part of the study, Abt completed an environmental scan and held an Advisory Panel Meeting on Human Trafficking Estimation with BJS that included experts invited to serve on the Advisory Panel.

Findings from the environmental scan and the expert panel meeting were used to develop three recommendations for a pilot study to test the feasibility of estimating the prevalence of human trafficking in the United States. Capture-recapture is the preferred method since it has the strongest capability for estimating the prevalence of hidden populations, including human trafficking. Vincent link-trace sampling is the second recommendation, followed by random sampling.

Disclaimer
The Bureau of Justice Statistics funded this third-party report. It is not a BJS report and does not release official government statistics. The report is released to help inform interested parties of the research or analysis contained within and to encourage discussion. BJS has performed a limited review of the report. Any statistics included in this report are not official BJS statistics unless they have been previously published in a BJS report. Any analysis, conclusions, or opinions expressed herein are those of the authors and do not necessarily represent the views, opinions, or policies of the Bureau of Justice Statistics or the U.S. Department of Justice.



Human Trafficking Prevalence Estimation Feasibility Study

Final Report

January 2025

Submitted to:
Bureau of Justice Statistics
999 North Capitol Street NW
Washington, DC 20002

Submitted by:
Abt Global LLC
6130 Executive Boulevard
Rockville, MD 20852

Authors
Shelby Hickman, PhD
Charlotte Lopez-Jauffret, PhD
formerly of Abt Global
Stephanie Fahy, PhD
Vanessa Masick

Acronyms

BJS	Bureau of Justice Statistics
DMST	Domestic Minor Sex Trafficking
ILO	International Labor Organization
IRB	Institutional Review Board
LTS	Link-Tracing Sampling
NHTH	National Human Trafficking Hotline
MSA	Metropolitan Statistical Area
MSE	Multiple Systems Estimation
NSUM	Network Scale Up Method
PPS	Probability Proportional to Size
PRIF	Prevalence Reduction Innovation Forum
RDS	Respondent-Driven Sampling
TLS	Time-Location Sampling
VLTS	Vincent Link Tracing Sampling

CONTENTS

Acronyms

Executive Summary 1

Volume I:
Findings from the Environmental Scan 2

1. Introduction 3
   1.1 Overview of Existing Prevalence Estimates and Limitations 3
   1.2 Study Objective 4

2. Research Methods 5
   2.1 Identification of Studies for Review 5
   2.2 Abstract Review 5
   2.3 Full Text Review 6

3. Findings 7
   3.1 Results by Method 13
      3.1.1 Random Sampling (and its Variations) 14
      3.1.2 Time-Location Sampling 15
      3.1.3 Network Scale Up Method 18
      3.1.4 Convenience Sampling (and its Variations) 19
      3.1.5 Multiple Systems Estimation and Capture-Recapture 21
      3.1.6 Respondent-Driven Sampling 24
      3.1.7 Vincent Link-Tracing Sampling 26

4. Conclusion 28
   4.1 Balancing Precision and Generalizability 28
   4.2 Aligning Methods to the Data Sources Available 29
   4.3 Aligning Methods to the Characteristics of the Population of Interest 30

Volume II:
BJS Advisory Panel Meeting on Human Trafficking Estimation
Summary 31

1. Introduction 32
   1.1 Participants 32
   1.2 Summary of Key Decision Points from Meeting 33

2. Key Considerations and Discussion Points by Agenda Item 34
   2.1 National Versus Subnational Estimates 34
   2.2 How to Break Down Human Trafficking into Estimable Components 34
      2.2.1 Forms of Trafficking and Exploitation 35
      2.2.2 Geography 35
      2.2.3 Data Sources 36
   2.3 Methodology 38
      2.3.1 Sex Trafficking 38
      2.3.2 Labor Trafficking 38
      2.3.3 Considerations for Methodology Selection 39
   2.4 Summary of Discussion Points 40

Volume III:
Recommendations for a National Estimate Pilot Study 41

1. Introduction 42
1.1 Overview 42
1.2 Cognitive Interviews 45
1.3 Ethical Considerations 45

2. Recommendations for Pilot Study 46
2.1 Recommendation 1: Capture-Recapture (Direct Observation) 46
2.1.1 Overview 46
2.1.2 Strengths of Capture-Recapture 46
2.1.3 Design Overview 47
2.2 Recommendation 2: Vincent Link-Trace Sampling (VLTS) Survey 52
2.2.1 Overview 52
2.2.2 Strengths 52
2.2.3 Population 52
2.2.4 Time Frame 53
2.2.5 Sampling Process 53
2.2.6 Sample Size 53
2.2.7 Recruitment of Participants 53
2.2.8 Cost 54
2.2.9 Limitations 54
2.3 Recommendation 3: Random Sampling Survey 55
2.3.1 Overview 55
2.3.2 Sampling Frame and Population 55
2.3.3 Time Frame 56
2.3.4 Cost 56
2.3.5 Strengths 56
2.3.6 Limitations 56
2.3.7 Formative Assessment 56

Bibliography 57

Executive Summary

Abt Global (Abt) conducted a study for the Bureau of Justice Statistics (BJS) to identify and recommend methods for estimating the prevalence of human trafficking in the United States (U.S.). This study supports BJS's mission to collect, analyze, publish, and disseminate information on crime, criminal offenders, victims of crime, and the operation of justice. As part of the study, Abt completed an environmental scan and held an Advisory Panel Meeting on Human Trafficking Estimation with BJS in February 2024 that included experts invited to serve on the Advisory Panel.

The environmental scan yielded 24 research articles that provided or discussed prevalence of any form of human trafficking spanning geographies inside and outside of the U.S. Between 2017 and 2023, seven articles included random sampling to estimate trafficking prevalence, five used link-tracing, three used convenience sampling, four used capture-recapture/multiple systems estimation, and four used respondent-driven sampling. A summary of each sampling method is included in the findings for the environmental scan in Volume I, including strengths and limitations for each type of method. Additionally, key recommendations (balancing precision and generalizability, aligning methods to data sources available, and aligning methods to the characteristics of the population of interest) are included from the environmental scan together with considerations for how BJS can develop a methodology that can produce an unbiased prevalence estimate of human trafficking in the U.S. The Advisory Panel Meeting on Human Trafficking Estimation occurred on February 28, 2024, and included subject matter experts, BJS and National Institute of Justice representatives, and Abt Global study team members. The meeting resulted in key decision points for estimating human trafficking victimization in the country, including considerations around trafficking type, geography, methodology, and data access and quality. Possible priority areas for trafficking type and trafficking subpopulations are presented in the meeting summary included in this report.

Findings from the environmental scan and the expert panel meeting were used to develop three recommendations for a pilot study to test the feasibility of estimating the prevalence of human trafficking in the U.S. Capture-recapture is the preferred method since it has the strongest capability for estimating the prevalence of hidden populations, including human trafficking. Vincent link-trace sampling is the second recommendation, followed by random sampling.

This report is arranged in three volumes: 1) Findings from the Environmental Scan; 2) BJS Advisory Panel Meeting on Human Trafficking Estimation Meeting Summary; and 3) Recommendations for a National Estimate Pilot Study.

Volume I: Findings from the Environmental Scan

1. Introduction

Abt Global (Abt) conducted an environmental scan of international and domestic studies that estimated the size of the human trafficking victim population in the U.S., including the benefits and limitations of each method and the scenarios in which each method would be used. This volume describes the results of the environmental scan and an initial assessment of the strengths and limitations of the methodologies reviewed.

1.1 Overview of Existing Prevalence Estimates and Limitations

Gaining further knowledge and quantifying the prevalence and scale of human trafficking victimization in the U.S. is critical to inform policy and practice, help government and non-governmental organizations assess their resource needs, identify areas for improvement, inform program strategies, and measure change over time (Office for Victims of Crime, 2021; United Nations Office on Drugs and Crime, 2022). Unfortunately, there are few rigorous prevalence estimates of human trafficking in the U.S., and none that estimate prevalence across different trafficking types, economic sectors, and demographic groups (Dank et al. 2021; Owens and Farrell 2014; Wright et al. 2021; Zhang et al. 2019).

Accordingly, estimates of human trafficking in the U.S. vary widely. For example, DeliverFund reports that between 15,000 and 50,000 women and children are victims of sex trafficking each year (DeliverFund, 2024), while a U.S. Department of Health and Human Services study estimates the number to be between 250,000 and 325,000. Estes and Weiner produced a study in 2001 reporting that more than 300,000 children were at risk of commercial sexual exploitation each year—a study that was shared widely among the media and public, but which was later criticized for a number of methodological problems, including artificially inflating the number of youths in different at-risk categories via double counting (Stransky & Finkelhor, 2008). Large variations in estimates of human trafficking undermine the credibility of any individual estimate.

When it comes to estimating labor trafficking, there is even less consensus on a national prevalence estimate. To date, only two studies in the U.S. have used a rigorous sampling approach when estimating labor trafficking prevalence (Zhang 2012; Zhang et al. 2019), and both focused on migrant workers in agricultural settings. A recent exploratory study of U.S. citizens, however, found that labor trafficking impacts a much broader range of industries and individuals than migrant workers in the agricultural sector (Dank et al. 2021).

Although its primary purpose is not to serve as a data collection mechanism, the National Human Trafficking Hotline (NHTH) provides the largest known dataset

on sex and labor trafficking victimization in the U.S. In 2021—the most recent year for which published data are available—the NHTH handled 7,499 situations of sex trafficking, 1,066 situations of labor trafficking, and 400 situations of sex and labor trafficking. A "trafficking situation" can refer to one or many potential victims. Although the NHTH provides useful information that can help shed light on the nature of human trafficking across different populations, the data collected are not, by design, generalizable (White, 2020).

When looking to international studies to identify how other countries estimate the prevalence of human trafficking, one of the most trusted sources is the International Labour Organization (ILO), which produces global estimates of modern slavery each year. The ILO includes forced labor, a component of human trafficking, and forced marriage in their definition of modern slavery. Their estimate of forced labor comprises forced labor in the private economy, forced sexual exploitation of adults, commercial sexual exploitation of children, and state-imposed forced labor. While the ILO does not provide a specific estimate for the U.S., they estimate that 3.6 million people in the Americas are victims of forced labor (International Labour Organization et al., 2022). The authors acknowledge limitations with the ILO methodology, however, that may impact the reliability of the estimates. For example, the set of surveyed countries used to produce the estimates is treated as a random sample, despite the countries not being selected at random.

1.2 Study Objective

This volume of the report provides findings from a systematic literature review of sampling methods that have been used to generate prevalence estimates of human trafficking victimization across different types of exploitation, population, and geographic area. It also describes the strengths and limitations of each method, including the forms of trafficking for which each method is best suited. Findings from this environmental scan were used to inform recommendations for a pilot study (Volume III).

Per the Trafficking of Victims Protection Act of 2000, human trafficking is defined as, "a crime involving exploitation of a person for labor, services, or commercial sex."

2. Research Methods

2.1 Identification of Studies for Review

The Abt team conducted a literature review to identify studies that estimate the prevalence of human trafficking, using the Elton B. Stephens Company, PubMed, and ScienceDirect search engines to facilitate our search. Since approaches to estimating the prevalence of hidden populations have advanced over time, the search was limited to articles published in the past six years in an effort to review the most recent methodologies. The team also limited the scope of the search to journal articles, research, and technical reports published between 2017 and 2023[1]. There were no limiting criteria for country of publication or population. The team used the following search terms on titles and abstracts:

- "Scoping review" OR Methodology OR Measurement OR Estimation OR Prevalence
- Trafficking OR Exploitation OR "modern slavery"
- Human OR Sex OR Lab*r
- NOT Gene Sequencing OR Endangered Animals

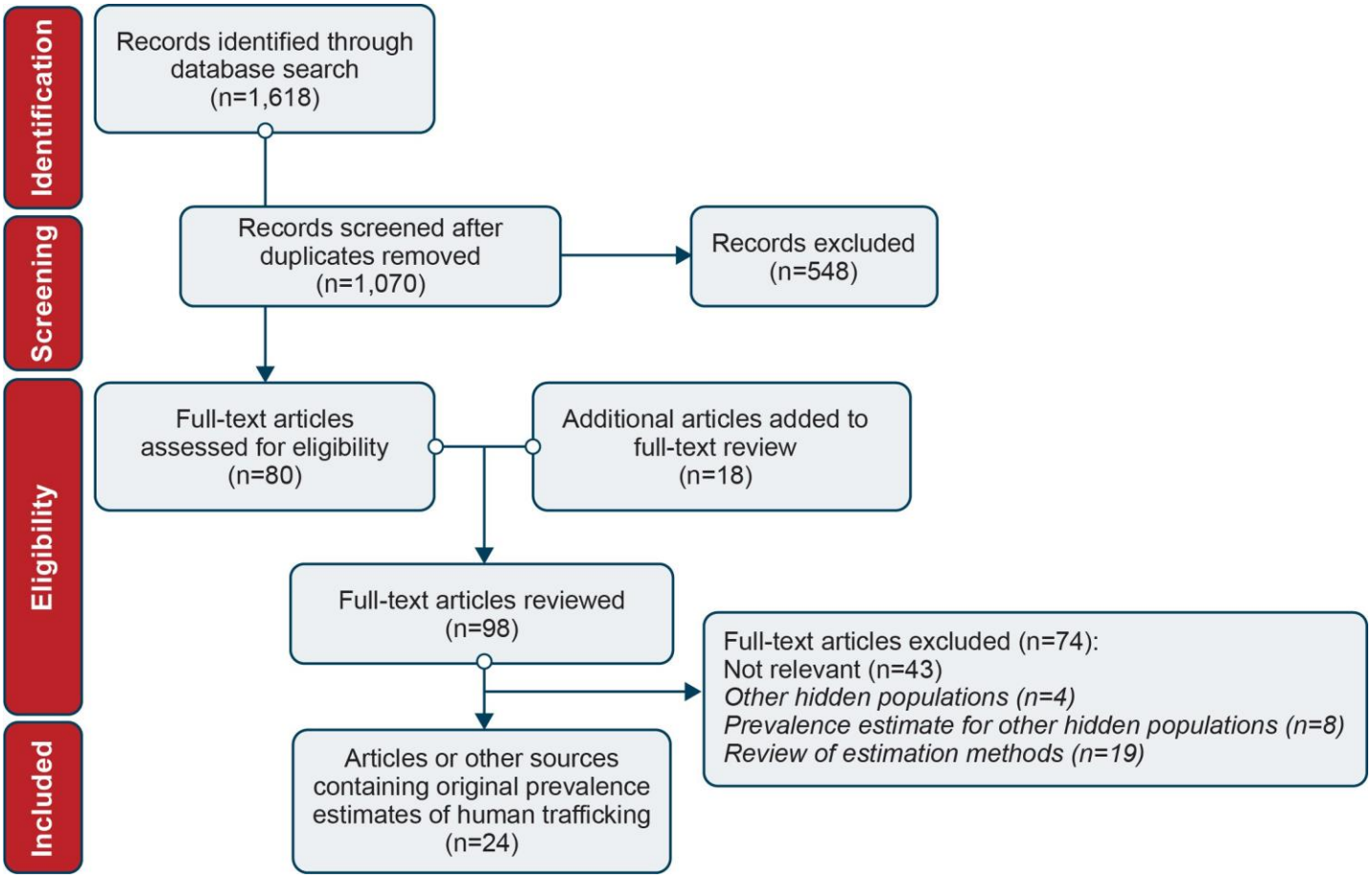
Once duplicates were removed, the team had 1,070 unique results from this search (**Exhibit 1**).

2.2 Abstract Review

The abstracts from the results of the literature search were uploaded to Abstrackr, an online tool that helps users conduct systematic literature reviews. The team included four members who reviewed 267 abstracts to identify which articles should be included in the full text review. At this stage, abstracts were excluded ("screened out") if they were not related to prevalence estimation(s), or if the prevalence estimation was for an irrelevant subject (e.g., animal hormones, antibiotic resistance). The team identified 80 articles to include after reviewing the abstracts.

1 The initial literature search was conducted in September of 2023.

Exhibit 1. Study Inclusion Diagram



2.3 Full Text Review

The Abt team created a workbook to organize the articles included in the full-text review and extracted information from each article on the following variables: Journal, Year of Publication, Abstract, Data Source, Data Collection Time Frame, Sample Size, Sampling Method, Crime Estimated or Reviewed (type of trafficking), Population Estimated or Reviewed (adult victims, child victims, adult and child victims), Industry (for labor trafficking only), Country (of population being estimated), Locality (of population being estimated), and Prevalence Calculated. At this stage, the team only retained articles that included original human trafficking prevalence estimates for the analysis.

Excluded articles included those that focused on other hidden populations (e.g., illicit drug users), contained results that were not prevalence estimates, and which were general methodology reviews. During this process, the team also added some articles to the full text review that were not part of the initial literature search, but which were either sent to the project team by the client or identified by the Abt team during examinations of specific sampling methodologies and during the planning of an expert panel convening. The team included a total of 24 articles in the final pool of articles for analysis.

3. Findings

The environmental scan yielded 24 articles that provided or discussed prevalence of any form of human trafficking. Of these articles, 29 percent were published in 2021 (n=7) and 33 percent in 2023 (n=8). The remaining 38 percent of articles were published between 2017 and 2020. With over half of the articles published within the last two years, this environmental scan offers insights into the most current prevalence estimation methods for human trafficking.

Environmental Scan Articles by Year of Publication

Year of Publication	Number of Articles
2017	3
2018	1
2019	4
2020	1
2021	7
2022	0
2023	8

The scan yielded equal numbers of studies estimating the prevalence of sex and labor trafficking (10 articles on sex trafficking studies and 10 on labor trafficking studies). Only 17 percent of the articles (n=4) included studies involving both sex and labor trafficking populations. The victim populations studied fell into three main categories: adult and child victims (n=8), adult victims only (n=7), and child victims only (n=8). Only one article focused on perpetrators of human trafficking.

Five articles focused on the U.S., while 19 focused on other countries and three included more than one country. Of the articles that focused on the U.S., the states included California, Ohio, Georgia (Atlanta), and the Midwest. **Exhibit 2** provides an overall summary of findings from the environmental scan, detailing all 24 articles included in the analysis.

Exhibit 2. Environmental Scan Summary

Year	Author	Title	Trafficking Type	Population	Country	Data Source	Sample Size	Sampling Method	Prevalence Estimated
2017	Mak, J.; Abramsky, T; Sijapati, B; Kiss, L; Zimmerman, C	What is the prevalence of and associations with forced labour experiences among male migrants from Dolakha, Nepal? Findings from a cross-sectional study of returnee migrants	Labor	Adult victims	Nepal	Survey (primary)	159	Census	73% of returned migrants experienced forced labor.
2017	O'Brien, J. E.; Li, W; Givens, A; Leibowitz, G. S.	Domestic minor sex trafficking among adjudicated male youth: prevalence and links to treatment	Sex	Child victims (male)	United States (Midwest)	Survey (secondary)—cross-sectional data from adjudicated youth in two states	671	Convenience sampling	10% of adjudicated young men reported they experienced domestic minor sex trafficking.
2017	Cruyff, M.; van Dijk, J.; & van der Heijden, P G. M.	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	Sex and Labor	Adult and child victims	Netherlands	Administrative data—(National Police, Royal Border Police, Regional Coordination Officed, Labor Inspectorates, etc.)	8,234	Multiple systems estimation	Total estimates of the population of victims of human trafficking: 2010=6,000 2011=9,400 2012=9,500 2013=8,000 2014=7,800 2015=6,600
2018	Grosso, A.; Busch, S.; Mothopeng, T.; Sweitzer, S.; Nkonyana, J.; Mpooa, N.; Taruberekera, N.; Baral, S.	HIV risks and needs related to the Sustainable Development Goals among female sex workers who were commercially sexually exploited as children in Lesotho	Sex	Child victims (female)	Lesotho	Survey (primary)	710	Respondent-driven	20% of participants were sexually exploited as children.
2019	Khatab, K.; Raheem, M. A.; Sartorius, B.; Ismail, M.	Prevalence and risk factors for child labour and violence against children in Egypt using Bayesian geospatial modelling with multiple imputation	Labor	Child victims	Egypt	Survey (secondary), including the 2014 Egypt Demographic and Health Survey from the Ministry of Health and Population	20,560	Random sampling	31.6% of children ages 5–10 are working, while 44.7% of children ages 5–10 were engaged in hazardous work.

Year	Author	Title	Trafficking Type	Population	Country	Data Source	Sample Size	Sampling Method	Prevalence Estimated
2019	Zhang, S. X.; Dank, M.; Vincent, K.; Narayanan, P.; Bharadwaj, S.; Balasubramaniam, S. M.	Victims without a Voice: Measuring Worst Forms of Child Labor in the Indian State of Bihar	Labor	Child victims	India	Survey (primary), and Qualitative	1,032 from 15 cities; 562 from Patna	Random sampling, link-tracing	Prevalence of child labor among child workers in Bihar is 61%.
2019	Urada, L. A.; Rusakova, M.; Odinokova, V.; Tsuyuki, K.; Raj, A.; Silverman, J.G.	Sexual Exploitation as a Minor, Violence, and HIV/STI Risk among Women Trading Sex in St. Petersburg and Orenburg, Russia	Sex	Adult and child victims	Russia	Survey (primary)	896	Time-location	11% reported being under the age of 18 at the time of entering the sex trade. Of these, 43% indicated they were forced into it. Overall, 57% of women who entered the sex trade as an adult said they felt forced.
2019	Anderson, V. R.; Kulig, T.C.; Sullivan, C. J.	Estimating the Prevalence of Human Trafficking in Ohio, 2014-2016	Sex	Adult and child victims	United States (Ohio)	Administrative data, including local child welfare, legal services, and law enforcement	486	Census	Prevalence of being at-risk for human trafficking victimization is estimated to be 2,250, with a 95% confidence interval of between 2,158 and 2,345. When excluding juvenile justice identified risk, prevalence is 987 with a 95% confidence interval between 926 and 1,050.

Year	Author	Title	Trafficking Type	Population	Country	Data Source	Sample Size	Sampling Method	Prevalence Estimated
2020	Sharifi Far, S.; King, R.; Bird, S.; Overstall, A.; Worthington, H.; & Jewell, N.	Multiple systems estimation for modern slavery: Robustness of list omission and combination	Sex and Labor	Adult and child victims	United Kingdom and Romania	Administrative data from the UK, including from local authorities, non-governmental organizations, police force, government organizations, and the general public data from Romania, including: police, International Organization for Migration, non-governmental organizations, foreign authorities, and other sources	UK: 2,744 Romania: 879	Multiple systems estimation	In the UK, prevalence was estimated to be 11,313 with a 95% confidence interval between 9,750 and 12,876. In Romania, prevalence was estimated to be 921 with a 95% confidence interval between 879 and 993.
2021	Saewyc, E. M.; Shankar, S.; Pearce, L.A.; Smith, A.	Challenging the Stereotypes: Unexpected Features of Sexual Exploitation among Homeless and Street-Involved Boys in Western Canada	Sex	Child victims (male)	Canada	Survey (secondary)—Canada Homeless and Street-Involved Youth Health Survey	2006: n=762 2014: n=681	Convenience sampling	One in four homeless or street-involved male youth reported sexual exploitation.
2021	Wright, E. R.; LaBoy, A.; Tsuker-man, K.; Forge, N.; Ruel, E.; Shelby, R.; Hig-bee, M.; Webb, Z.; Turner-Harper, M.; Darkwa, A.; Wallace, C.	The Prevalence and Correlates of Labor and Sex Trafficking in a Community Sample of Youth Experiencing Homelessness in Metro-Atlanta	Sex and Labor	Adult and child victims	United States (Atlanta)	Survey (primary)	564	Convenience sampling	16% of runaway or homeless youth reported commercial sexual exploita-tion. 29% reported coerced labor.

Year	Author	Title	Trafficking Type	Population	Country	Data Source	Sample Size	Sampling Method	Prevalence Estimated
2021	Stockl, H.; Fabbri, C.; Cook, H.; Galez-Davis, C.; Grant, N.; Lo, Y.; Kiss, L.; Zimmerman, C.	Human trafficking and violence: Findings from the largest global dataset of trafficking survivors	Sex and Labor	Adult and child victims	Global	Administrative data from the International Organization for Migration Victim of Trafficking Database	10,369	Census	The prevalence of reported violence during human trafficking included: 54% physical and/or sexual violence; 50% physical violence; and 15% sexual violence, with 25% of women reporting sexual violence.
2021	Giommoni, L. and Ikwu, R.	Identifying human trafficking indicators in the UK online sex market	Sex	Adult and child victims	130 countries	Website data	25,056	Census	59% of advertisements had one indicator of human trafficking. 21% had two indicators of human trafficking. 19% had no indicators.
2021	Li, R.; Tobey, M.; Mayorga, M. E.; Caltagirone, S.; Ozaltin, O. Y.	Detecting Human Trafficking: Automated Classification of Online Customer Reviews of Massage Businesses	Sex	Perpetrators	United States (California)	Website data	430,682	Census	10% of 1,735 online reviews of massage businesses were labeled illicit.
2021	Vincent, K.; Zhang, S.X.; Dank, M.	Searching for Sex Trafficking Victims: Using a Novel link-tracking Method Among Commercial Sex Workers in Muzaffarpur, India	Sex	Adult victims	India	Survey (primary)	317	Respondent-driven, link-tracing	Preliminarily estimated 1,195 victims. Using link-tracing, estimated 2,816 sex workers and of those, 340 as victims of sex trafficking.
2021	Abusaleh, K.; Islam, M. R.; Ali. M. M.; Khan, M. A.; Shahinuzzaman, M.; Haque, M. I.	Prevalence of Economic Exploitations and Their Determinants Among Child Labourers in Dhaka City, Bangladesh: A Mixed-Method Study	Labor	Child victims	Bangladesh	Survey (primary), and qualitative (interviews)	Survey: 324 Interviews: 15	Probability sampling (survey), convenience sampling (interviews)	The survey revealed 97% of child laborers were economically exploited.
2023	Dank, M.	Estimating the Prevalence of Forced Labor Among Domestic Workers in Tanzania	Labor	Adult victims (female)	Tanzania, Zanzibar	Survey (primary)	Stratified simple random sampling: 1,052; link-tracing: 788	Random sampling, link-tracing	Overall, the prevalence for stratified simple random sampling is 69% and for link-tracing is 59%.

Year	Author	Title	Trafficking Type	Population	Country	Data Source	Sample Size	Sampling Method	Prevalence Estimated
2023	Zhang, S.; Vincent, K.; Cole, K.; Rydberg, J.; Thompson, S.; & Dank, M.	Domestic Servitude in Tunis, Tunisia: Findings from Two Sampling Methods and Policy Implications (Prevalence Reduction Innovation Forum Learning Series)	Labor	Adult victims (female)	Tunisia	Survey (primary)	365 sites 1,029 (time-location); 923 (capture-recapture)	Time-location, capture-recapture	For time-location sampling, 25% meet criteria for trafficking in persons. For capture-recapture, 26% meet criteria for trafficking in persons.
2023	Anderson, E. and Kendall, C.	Estimating the Prevalence of Sex Trafficking of Children in Recife, Brazil	Sex	Child victims (female)	Brazil	Survey (primary)	Respondent-driven sampling: 602; network scale up: 1,002	Respondent -driven, network scale-up	Respondent-driven sampling: 22,600; network scale up: 19,700.
2023	Babiarz, K.	Estimating the Prevalence of Human Trafficking in Brazilian Agriculture	Labor	Adult and child victims	Brazil	Survey (primary)	Population Representative Survey: 8,814; network scale up: 687	Random sampling, network scale-up	Overall, the prevalence of human trafficking for agricultural workers per thousand for the network scale up method was 26.19.
2023	Branscum, C. and Richards, T. N.	The Wrong Rite of Passage: Comparing Sex Trafficking to Other Types of Child Maltreatment Using the National Child Abuse and Neglect Data System (NCANDS)	Sex	Child victims	United States	Administrative data	4,255,946	Census	Child sex trafficking comprised 0.18% of all substantiated reports of child maltreatment
2023	Sahai, R.	Estimating the Prevalence of Forced Labor Among Domestic Workers in Casablanca, Morocco	Labor	Adult victims (female)	Morocco	Survey (primary)	Capture-recapture: 1,067 link-tracing: 989	Capture-recapture, link-tracing	Overall, the prevalence for capture-recapture is 29% and for link-tracing 32%.
2023	Robinson, C.	Estimating the Prevalence of Forced Labor in the Brick Kiln Industry in Pakistan	Labor	Adult victims	Pakistan	Survey (primary), and qualitative (interviews)	800	Respondent-driven, random sampling	Overall prevalence for respondent-driven sampling is 34% with 18% probability. Proportional to size.
2023	Dank, M.	Estimating the Prevalence of Forced Labor in the Fishing Industry in Costa Rica	Labor	Adult victims	Costa Rica	Survey (primary)	Population representative survey: 1,017 link-tracing: 1,009	Random sampling, link-tracing	Overall, the prevalence for probability proportional to size is 22% and for link-tracing 23%.

3.1 Results by Method

The Abt team examined each sampling method to document the number of articles identified for each method by year of publication, type of trafficking, and trafficking population. Definitions of each method appear in the following sections. Between 2017 and 2023, seven articles included random sampling to estimate trafficking prevalence, five used link-tracing, three used convenience sampling, four used capture-recapture/multiple systems estimation, and four used respondent-driven sampling. Though used in earlier years, studies using convenience sampling were not included in articles published in 2022 and 2023. In 2023, most articles included prevalence estimates from random sampling followed by link-tracing (**Exhibit 3**).

Exhibit 3. Number of Sampling Methods by Year of Publication

Sampling Method	Year							Total*
	2017	2018	2019	2020	2021	2022	2023	
Capture-recapture/multiple systems estimation	1	-	-	1	-	-	2	4
Census	1	-	1	-	3	-	1	6
Convenience	1	-	-	-	2	-	-	3
Link-tracing	-	-	1	-	1	-	3	5
Network scale up	-	-	-	-	-	-	2	2
Random	-	-	2	-	1	-	4	7
Respondent-driven	-	1	-	-	1	-	2	4
Time-location	-	-	1	-	-	-	1	2

*Articles which indicated more than one sampling method were counted in each sampling method category. Nine articles indicated more than one sampling method.

Looking at the type of trafficking by sampling method (**Exhibit 4**), the team found that most of the labor trafficking only articles (N=7) used random sampling followed by link-tracing (N=4). Articles focused only on sex trafficking were more likely to include census (N=4) and respondent-driven sampling methods (N=3). Capture-recapture/multiple systems estimation methods were included in two labor trafficking articles only, and two articles focused on sex and labor trafficking.

Most of the articles on labor trafficking only focused on adult victims (N=11) using random sampling and link-tracing sampling. The majority of articles that estimate prevalence of sex trafficking focus on children (N=6) and use convenience and respondent-driven sampling methods.

Exhibit 4. Number of Sampling Methods by Trafficking Type

Sampling Method	Trafficking Type			Number of Articles*	Percentage of Environmental Scan (n=24 articles)
	Labor	Sex	Sex & Labor		
Capture-recapture/multiple systems estimation	2	-	2	4	17%
Census	1	4	1	6	25%
Convenience	-	2	1	3	13%
Link-tracing	4	1	-	5	21%
Network scale up	1	1	-	2	8%
Random	7	-	-	7	29%
Respondent-driven	1	3	-	4	17%
Time-location	1	1	-	2	8%

*Articles which indicated more than one sampling method were counted in each sampling method category. Nine articles indicated more than one sampling method.

The sections below break out sampling methods into four major categories: 1) traditional probability-based sampling methods (random sampling and its variations); 2) probability-based sampling adaptations for hard-to-reach populations (network scale up method and time-location); 3) traditional non-probability-based sampling methods (convenience samples); and 4) non-probability sampling methods adapted for hard-to-reach populations (capture-recapture, respondent-driven sampling, link-tracing). Each method includes a table with their respective strengths and limitations.

3.1.1 Random Sampling (and its Variations)

Probability sampling, or random sampling, is considered the gold standard for generating unbiased estimates that are generalizable to a larger population because it gives all individuals within a sampling frame an equal chance of being surveyed. This is important because the estimators used to generate prevalence estimates (and to carry out most frequentist statistical procedures) rely on an assumption that the sample is randomly drawn (Angrist & Pischke, 2009). The assumptions required for random sampling are simple and easy to interpret, but also shed light into the limitations of random sampling for measuring human trafficking victimization. First, random sampling assumes that everyone in the target population has an equal chance of being selected in the sample. Second, it assumes that researchers can contact or access any member of the population if selected.

Researchers have used probability sampling methods to calculate trafficking prevalence estimates, including taking a random sample of households from a government census to identify respondents who may be victims of trafficking (Youle & Long, 2020). Although this approach enables generalization to a larger population, using random sampling increases the likelihood of undercounting or underreporting trafficking victims since they are part of a hidden or hard-to-reach population. Some researchers have suggested that probability-based sampling is more feasible when estimating the prevalence of labor trafficking victims, particularly for studies examining industries that are visible to the public such as construction (Barrick & Pfeffer, 2021). In fact, of the seven studies in our review that used random sampling, all examined just labor trafficking. Three studies included estimates for adult victims, three studies included estimates for child victims, and one included estimates for adult and child victims.

One study by Khatab and colleagues (2019) estimated the prevalence of child labor in Egypt using data from the 2014 Egypt Demographic and Health Survey, which uses a random sample to collect data on child labor in Egypt. The study found that 31.6 percent of children ages 5-10 were working, and of those 44.7 percent were engaged in hazardous work. Abusaleh and colleagues (2021) also conducted a survey that used a systematic random sampling to estimate the prevalence of exploited child laborers in Dhaka City, Bangladesh. The survey revealed 96.6 percent of child laborers were economically exploited, earning below minimum wage, and 86 percent of child laborers worked long hours with no breaks. Other studies (Dank, 2023a; Zhang et al., 2019) used random sampling in combination with link-tracing (described later in this report) to estimate the prevalence of adult and child labor trafficking victims. Dank (2023) estimated the prevalence of forced labor among domestic workers in Tanzania. Using stratified simple random sampling, they found that the overall rate of domestic servitude was estimated at 68.5 percent, and the overall rate of domestic servitude using link-tracing was estimated at 59.1 percent. Zhang and colleagues (2019) used systematic random sampling and link-tracing sampling to estimate the prevalence of child labor trafficking in Bihar, India, and found the overall prevalence of the most psychologically and physically detrimental forms of child labor among the child worker population in Bihar was 61 percent, and that 88.5 percent of the work was illegal, including violations of age and hours caps. The authors indicated that they used weighting to combine the two samples, so estimates are a combination of random and link-tracing sampling. **Exhibit 5** describes the high-level strengths and weaknesses of random sampling for estimating the prevalence of human trafficking victimization.

Exhibit 5. Random Sampling Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Assumptions are simple, and if met, random sampling produces an unbiased estimate generalizable to the population from which the sampling frame was drawn.	Assumes that everyone in the target population has an equal chance of being selected in the sample, which narrows parameters around to whom the estimate is generalizable. May undercount actual victim population. Assumes that researchers can contact or access any member of the population if selected, which is often not the case with trafficking victims who may be fearful to engage with researchers, or are socially, linguistically, or geographically isolated.
Breadth of research	Probability sampling is considered the gold standard for sampling methods due to its ability to produce unbiased estimates. It has been used in prior human trafficking prevalence studies.	Research limited to industries/populations where data are available to construct a sampling frame from which to randomly sample (e.g., national survey of child laborers).
Data	Many cities and states in the U.S. have sufficient data on households, residences, employers, to construct a sampling frame from which to randomly select a sample. Breadth of examples of national surveys using probability sampling, like the BJS National Crime Victimization Survey.	Requires data sources to construct sampling frame that capture victims of human trafficking. Often, victim residences/locations are unknown, and researchers must rely on a best guess which can lead to undercounting. May be more feasible for labor trafficking estimation, where work sites are identifiable through official sources.
Who can be counted	Can sample individuals or employers. Works well with general population samples that can be accessible through Census data or included in household- or school-based samples.	Will not capture anyone who is not included in the initial sampling frame. Trafficking victims that are isolated, living in informal arrangements, not attending school, or who intentionally avoid contact with authorities/researchers may not be included in the sample or agree to participate.
Forms of trafficking it can estimate	Has been used for both sex and labor trafficking, but may be more feasible for labor trafficking, particularly for industries for which data on work sites are available/ known (and thus researchers can have confidence about where workers may live), and multi-stage samples can be used where employers are the primary sampling unit.	Employers may deny access to employees or provide a limited or biased sample of employees (e.g., legally authorized workers).
Consideration of lived experiences of victims and other interested parties	Opportunity to include the victim's voice via face-to-face survey collection.	Participants must be willing to answer survey questions related to victimization.
Cost	Although none of the random sampling articles reviewed used this approach, if items are added to an existing probability survey it could be a low-cost option,	Random sampling is costly due to large sample size required, and field work associated with data collection.

3.1.2 Time-Location Sampling

Time-location sampling (TLS), also referred to as time-space sampling or venue-based sampling, is used to collect data from hard-to-reach populations, such as people at risk of HIV; individuals who identify as gay, lesbian, bisexual, transgender, questioning and/or queer; and, more recently, victims of human trafficking. TLS involves randomly selecting locations from an exhaustive list of all relevant venues, randomly selecting the days and times to recruit participants at each of the randomized venues, then randomly selecting study participants from the randomized venues at the randomized days and times. Given the importance of documenting all relevant venues, TLS teams often include an ethnographic component in which the research teams work closely with local stakeholders—including service providers and individuals with lived experience—to get a complete list of venues. TLS approximates cluster random sampling, but instead of randomly selecting individuals, it begins with randomly selecting venues. This is helpful because no census or register typically exists to capture all individuals at risk for human trafficking in a given area, as would be needed for traditional random sampling.

The assumptions required for TLS to approximate cluster random sampling are that everyone in a venue either has equal chances of being surveyed or known chances of being surveyed that can be modeled. Since field teams must administer surveys in person, studies often limit data collection to daylight hours for the safety of data collectors, which is an often-cited limitation of this method.

Two studies in our review used TLS. One study examined labor trafficking and one study examined sex trafficking. The labor trafficking study focused on adult victims, and the sex trafficking study focused on adult and child victims.

Zhang (2023) conducted a study designed to estimate the prevalence of domestic servitude among women in Tunis, Tunisia. The author used TLS, which included identifying all venues where domestic workers were expected to pass through or congregate. The sample included 6,281 domestic workers—from this sample the author estimated about 25 percent of the domestic worker population in Tunis met the criteria for human trafficking. The author acknowledged that foreign migrant workers were not captured well using this sampling method, in part because recruitment was limited to respondents who were in public venues during daytime hours.

Another study by Urada and colleagues (2019) recruited Russian women in the sex trade using TLS to estimate the prevalence of child sexual exploitation among women engaged in sex work in St. Petersburg and Orenburg, Russia. For this study, time-location clusters of street sex trade (with different times/days across locations) were identified; however, all invited women trading sex were invited to participate in the study versus being randomly selected. Eleven percent of the women reported being under the age of 18 when they entered the sex trade (making them victims of child sexual exploitation), and 43 percent of these women reported they were forced into sex work. The authors note that generalizability may be limited by the time and location of the data collected at two study sites within two Russian cities. Additionally, as noted above, participants were not randomly selected from the randomized venues at the randomized days/times.

Although TLS has been used in public health research (such as for populations at high risk of infectious disease), it has not been as widespread in trafficking research. It is, however, a promising approach, since it provides access to groups of trafficked individuals who may not be captured using traditional sampling methods (Schroeder et al., 2022). This approach can be resource and time intensive since it requires substantial groundwork to identify known venues. The strengths and limitations of time-location sampling studies for measuring human trafficking victimization are provided in **Exhibit 6**.

Exhibit 6. Time-location Sampling Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Assumes representation by approximating random cluster sampling by sampling sufficient places and days/times. Can produce a large sample and generalizable prevalence estimate.	Assumes the population of interest is the population present/passing through selected venues. Requires that everyone has equal likelihood of being surveyed in the venue, or that likelihoods are known and can be modeled. The probability sample generated in TLS is visits, not individuals.
Breadth of research	Time-location sampling has been used in public health research, including studies that examine individuals at risk of HIV.	Has not been as widely used for trafficking victims compared with other hidden populations. Other hidden populations researched may have stronger likelihood of attending the same venues (e.g., health clinics for HIV positive individuals, harm reduction centers for injection drug users, LGBTQ community centers for LGBTQ individuals).
Data	Does not require a complete list/census to form sampling frame of individuals at risk of human trafficking.	Requires a complete map of the universe of venue-day-time.
Who can be counted	Works well with hidden or hard-to-reach populations, including trafficking victims who cannot be observed by traditional sampling methods, but who can be observed in public locations.	By design, biased towards those who attend venues at certain times, leaving out individuals who do not or rarely attend venues. For safety purposes, often excludes venues frequented at night or individuals who only attend venues at night.
Forms of trafficking it can estimate	Has been used to estimate both sex and labor trafficking. May be more feasible for labor trafficking, particularly for industries that are outside or visible to the public.	Less feasible for forms of human trafficking for which victims do not congregate in public venues such as rural agricultural work.
Consideration of lived experiences of victims and other interested parties	Requires formative ethnographic style research within the community while working with community partners to develop map of the universe of venue-day-times.	Participants must be willing to answer survey questions related to victimization.
Cost	Focusing on venues where victims are known to congregate, which is often less costly, allows for estimation with a smaller sample size relative to probability sampling for households, for example.	Ethnographic research required to identify complete list of venue-day-time can be time intensive and costly.

3.1.3 Network Scale-Up Method

The network scale-up method (NSUM) allows researchers to estimate the prevalence of hidden populations using traditional sampling techniques without directly surveying or interviewing members of the target population (Schroeder et al., 2022). The basis of NSUM is that the likelihood of knowing someone in a subpopulation (e.g., victims of human trafficking) is the size of that subpopulation divided by the population size. For example, if one individual knows 100 people and one victim of human trafficking, then one percent of the overall population from which that individual was randomly drawn are considered to be victims of human trafficking.

To use this information, NSUM requires that researchers determine each individual's personal network size (i.e., degree). This is done by asking respondents how many people they know who belong to objectively measured subpopulations (e.g., number of twins, number of people with Type II diabetes). Each individual should know their degree times the number of people actually in the subpopulation. Researchers use this information (degree, known subpopulations) to estimate the size of the hidden population via a scale-up estimator.

There are three key assumptions in NSUM, and the main criticisms of this method tie to common violations of these assumptions. The first is that social connections are formed at random. NSUM, therefore, does not consider that people make friends with those similar to themselves. This category of bias is referred to as a *barrier effect*. A second assumption is that respondents know which of their contacts belong to the hidden subgroup of interest. This is a key limitation of the method. Although people are typically willing to share characteristics like whether they are a twin, or even if they have type II diabetes, with their friends, they are much less likely to openly share that they are being exploited. Respondents, therefore, may not know that their connections are victims of human trafficking, referred to as *transmission bias*. Finally, NSUM requires participants to accurately recall the size of their personal network and their characteristics. In reality, however, people tend to underestimate their total connections in larger subgroups and overestimate their contacts in more niche subgroups, referred to as *recall bias*. Another limitation of NSUM, unrelated to assumptions underlying the method, is that there are limited practical applications for the results. Since the reports are made indirectly (respondents reporting about their social networks), NSUM surveys do not get information about victim characteristics that would be useful for informing policy and practice. On the other hand, because NSUM only asks how many victims of human trafficking a respondent knows, NSUM does not require victims to disclose their unique individual experiences to researchers. Even if a respondent includes themselves in their known tally of victims, they do not have to share that information with the researcher.

Two articles in our review used NSUM to generate estimates of human trafficking population size (Anderson & Kendall, 2023; Babiarz, 2023). Both were part of the Prevalence Reduction Innovation Forum (PRIF), a partnership between the U.S. State Department and the University of Georgia with the goal of improving measurement and reducing the prevalence of human trafficking. One used NSUM to estimate the prevalence of labor trafficking in Brazilian agriculture. This study first performed a population representative household survey, then sampled a subset of 10 percent of respondents to answer NSUM survey questions. A strength of this study is that the researchers administered the representative household survey in person, allowing them to make the determination about trafficking victimization in real time. This enabled researchers to ask the participants who were flagged as likely victims about how much they shared their experiences with other people in their network. This allowed them to understand and model transmission of information and to see how visible trafficking experiences are within the network. In this study, researchers found that the prevalence estimated using NSUM—26.19 per thousand—was in line with the prevalence calculated in the household survey for the moderate threshold, at 27.7 per thousand.

The other study in our review that used NSUM—also part of the PRIF—measured the prevalence of child sex trafficking in Recife, Brazil. This study surveyed adult female survivors about their experiences as children. Using NSUM and respondent-driven sampling (RDS), the study found that the NSUM median estimate (19,700 victims) was lower than the RDS estimate of (22,600). Using both ranges (from NSUM and RDS) they found similar prevalence estimates: about 16.7 percent of young women ages 18–21 at time of the study were child sex trafficking victims.

A third study (Sahai, 2023), measuring the prevalence of forced labor among domestic workers in Morocco, attempted to use NSUM but ultimately found that it was not feasible due to the length of the survey instrument that would be required. A summary of the strengths and limitations of NSUM studies for measuring human trafficking victimization is provided in **Exhibit 7**.

Exhibit 7. Network Scale-up Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Newer statistical adjustments can reduce some biases associated with violations of assumptions.	Barrier effects (social connections not formed at random); transmission effects (people may not know that their connections are victims of human trafficking); recall bias (people are bad at remembering the number of connections they have in different groups).
Breadth of research	Method first used in 1989, significant advancements since that time in improving estimators.	Fewer applications to measuring human trafficking than methods that have been around for similar lengths of time like multiple systems estimation (MSE) and RDS. "Network Scale-up [is] too new to know if [it is] good for estimating human trafficking" (Barrick and Pfeffer, 2021).
Data	Since NSUM doesn't ask respondents to identify individuals in the target subpopulation, may improve honesty of responses.	Requires population-level information on known characteristics, which is not readily available in all countries. Requires relatively long survey instrument to get all needed information, not practical in some contexts (e.g., Sahai 2023).
Who can be counted	Includes individuals and their networks for anyone who will be captured through traditional sampling design.	Similar to limitations of traditional sampling designs; may not include institutionalized, incarcerated, homeless individuals. May underestimate prevalence because completely isolated individuals would likely not be surveyed or known to survey respondents.
Forms of trafficking it can estimate	Can and has been used for any form of human trafficking. Double scale-up method which leverages multiple indicators of trafficking can be used to account for people not knowing full details about their connections.	NSUM surveys do not get information about victim characteristics that would be useful for informing policy and practice. Information that social connections share about exploitation experiences may not be detailed enough for respondent to know what form of human trafficking their connection experienced.
Cost	Items, in theory, can be added to any regular survey with a conventional sampling approach, which may provide significant cost savings relative to other methods.	Requires a relatively long list of survey items which may make it undesirable as an add-on to existing surveys, requiring full cost of original survey collection.
Consideration of lived experiences of victims and other interested parties	Relies on indirect reporting, so victims do not have to self-disclose their experiences.	In general, does not intentionally engage with victims, service providers, or other individuals with lived experiences.

3.1.4 Convenience Sampling (and its Variations)

Studies examining hard-to-reach populations, including human trafficking, may use non-randomized sampling strategies. Convenience sampling is one of the most common types of non-probability sampling techniques and is often used when there are budget constraints, or when random sampling is not feasible practically or due to budget. Convenience samples are not representative of a broader population and are therefore not generalizable (Andrade, 2021).

Three studies in the review used convenience sampling, two of which examined sex trafficking and one that examined sex and labor trafficking. Two studies included estimates for child victims only and one estimate included adult and child victims.

A study by O'Brien and colleagues in 2017 surveyed adjudicated male youth housed in juvenile detention facilities in two American states to understand the prevalence of domestic minor sex trafficking (DMST) among this unique population. The study found that 10 percent of youth surveyed had experienced DMST prior to their arrest. The authors acknowledge that, while this study is one of the few to investigate DMST among adjudicated youth generally and male youth specifically, the findings cannot be generalized to populations of adjudicated youth in other states. Additionally, this type of facility-based sample may result in overrepresentation of the more salient forms of exploitation for detained youth (i.e., sexual exploitation within facilities; Gerassi et al., 2017).

Saewyc and colleagues (2021) conducted a study designed to understand the prevalence and experiences of sexual exploitation among adolescent boys who participated in the 2006 and 2014 Homeless and Street-Involved Youth surveys in Canada. These surveys used a convenience sample of homeless and street-involved youth in multiple cities and small towns across Canada. The study found that just over one in four boys age 17 and under reported trading sex for money, which in Canada is a form of sexual exploitation, and one in 10 youth who were 18 years or older reported first trading sex at age 17 or younger. The authors note that there is uncertainty about the true homeless and street-involved population. Without an accurate denominator for the target sample, it is hard to estimate the extent to which the sample is representative, so prevalence estimates should be treated with caution.

Another study reported estimates of sex and labor trafficking of youth experiencing homelessness in the metro-Atlanta area. Wright and colleagues (Wright et al., 2021) identified time periods and locations most likely to be successful in recruiting runaway and homeless youth from which to draw the sample and conduct a cross-sectional survey. Findings revealed that 15.6 percent of youth reported they were victims of commercial sexual exploitation while homeless, and 29.3 percent indicated they had been victims of coerced labor. Although these findings suggest that young people experiencing homelessness in the metro-Atlanta area are at a relatively high risk for trafficking, these findings are limited to the metro-Atlanta area and are not generalizable to a larger population.

Other studies included in our review analyzed data from existing databases and are considered population-based convenience studies, or census studies. Although there is no sampling involved since the whole population of eligible subjects in the database is studied, it is still a form of convenience sampling with limited external validity because the database was conveniently available and only the subjects eligible to belong in the database were studied. Therefore, findings are limited to the population in the database and may not be generalized beyond the group included in that database (Andrade, 2021).

Six studies in our review examined data using a census approach. One study examined labor trafficking, four examined sex trafficking, and one examined sex and labor trafficking. One study included estimates for adult victims, one study included estimates for child victims, three studies included adult and child victims, and one study for perpetrators.

Stockl and colleagues (2021) examined the prevalence of reported violence during human trafficking using the IOM Victim of Trafficking Database, which is the largest available international database that collects information on individual victims of trafficking. They found that 54 percent of victims reported physical and/or sexual violence during trafficking, 50 percent reported physical violence, and 15 percent reported sexual violence. Physical and sexual violence was significantly higher among women and girls, and violence was frequently associated with trafficking into manufacturing, agriculture, and begging. It is important to note that the Victim of Trafficking Database is a case management database and is not designed to systematically collect survey data. Additionally, it is only representative of known trafficking victims.

Another study, by Anderson and colleagues (2019), used administrative data collected from 12 state and local agencies to develop statewide estimates of known trafficking victims and individuals at risk for trafficking in Ohio. The overall estimate was 484 known victims. For those individuals who were classified as at-risk for human trafficking, the expected counts were 2,250 and 987 for the samples including and excluding the juvenile justice-identified risk, respectively. The authors acknowledge, however, that it is likely there are more known victims and at-risk individuals in Ohio than were identified from these sources. **Exhibit 8** provides a review of the strengths and limitations of convenience sampling studies for measuring human trafficking victimization.

Exhibit 8. Convenience Sampling Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	No assumption requirements.	The sample is not representative of a broader population and has limited utility of findings for policy or practice.
Breadth of research	Convenience sampling has been used extensively for human trafficking studies.	While findings are useful when no other estimates are available, they are generally not regarded as rigorous or reliable estimates.
Data	Convenience samples typically do not suffer from low response rates, which is a common challenge for probability sampling techniques.	People with strong views or opinions are more likely to participate in studies, therefore, convenience sampling has a higher risk of bias. This gives an unfair advantage or voice to certain members of a population likely to participate in the study. Researcher views or opinions could also influence the sample, resulting in researcher or selection bias.
Who can be counted	For researchers with strong ties to victim communities, these relationships can be leveraged to build trust.	As the name implies, only individuals who are easily accessible to researchers are counted in convenience samples. For original data collection, this often excludes victims in remote areas or who are socially isolated and do not interact with service providers or other public-facing agencies. For analyses of secondary datasets, this often means detailed information about exploitation is not available.
Forms of trafficking it can estimate	None noted.	Cannot estimate for any population; can only describe/report observed frequencies for the sample itself because most conventional statistical inference methods rely on assumption of random sampling.
Consideration of lived experiences of victims and other interested parties	For original data collection such as qualitative research or snowball sampling for surveys, enables victims to share their experiences. Information gleaned can be useful for designing more rigorous studies.	In analyses of secondary datasets, does not engage individuals with lived experience. For original data collection, participants must be willing to answer survey questions related to victimization.
Cost	This type of sampling is usually quick and inexpensive (i.e., convenient) relative to more rigorous sampling designs.	Some convenience sampling methods, like snowball sampling may require incentives.

3.1.5 Multiple Systems Estimation and Capture-Recapture

Multiple systems estimation (MSE) and capture-recapture allows researchers to make inferences about hidden populations using a convenience, or non-random, sample. This makes MSE an appealing option for researchers studying hidden populations, like human trafficking victims. The basic concept of MSE is taken from a broader category of methods called capture-recapture (or "mark-recapture") that were originally applied in ecological studies to measure the prevalence of wildlife populations. We describe MSE and capture-recapture methods together because they rely on the same underlying statistical assumptions and analysis. However, while MSE and capture-recapture rely on the same underlying concept, there are some differences. MSE typically uses lists from different organizations, service providers, or government entities, while capture-recapture, when applied to studying human trafficking, typically involves surveying people in person in key locations where they expect victims may be present and measuring overlap in who is identified in person (whereas MSE measures overlap in individuals on different lists).

There are two assumptions for using MSE and capture-recapture to generate an unbiased estimate: 1) mutual independence, which means that the probability of inclusion in one dataset (a list for MSE, being surveyed at a given time point for capture-recapture) is independent of inclusion in another dataset; and 2) homogeneous inclusion probability, which means that, for at least one of the data collection points (list or survey), the probability of being observed is the same for all members of the population of interest. The primary criticisms of MSE/capture-recapture stem from

the impracticality of these assumptions for populations of human trafficking victims. Mutual independence, for example, is rare, as different victim service providers and law enforcement agencies often refer to one another. Further, some individuals are more likely to seek help or come into contact with authorities than others. Additionally, some agencies may only accept certain groups of victims (e.g., women-only shelters).

Of the 24 articles in the environmental scan which produced original prevalence estimates of human trafficking, two used MSE and two used capture-recapture methods. Both capture-recapture studies in our review estimated the prevalence of labor trafficking, specifically domestic servitude of adult female victims. This is not surprising given that capture-recapture relies on identifying people in public spaces, which is more common for victims of labor trafficking in some sectors like domestic work and construction (compared to sex trafficking or more isolated industries like agriculture or fishing). Zhang et al (2023) used capture-recapture to estimate the prevalence of labor trafficking among domestic workers in Tunisia. They estimated that 26.4 percent of domestic workers met the criteria of human trafficking. Sahai (2023) also used capture-recapture to estimate labor trafficking among domestic workers in Morocco and estimated that 28.7 percent of domestic workers met the criteria for human trafficking. A strength of these studies is the narrow focus on one form of exploitation and one subgroup of victims. This reduces the number of covariates needed to generate the MSE estimate because they do not need to create strata or control for sex and/or age. It also reduces the likelihood of parameter redundancy or the inability to generate maximum likelihood estimators, both of which can lead to unstable estimates and the risk of overestimating population size and/or having very large standard errors. Capture-recapture can yield reliable estimates of the prevalence of domestic servitude but may not be appropriate, or may require much larger sample sizes, to be applied to broader categories of human trafficking. Preliminary findings from these studies were presented at the PRIF Results from the Field meeting. The final findings have not published, so details on how the authors account for heterogeneity in individuals' probability of being captured in each location is not clear. This, however, can be controlled for in the modeling process.

The two MSE studies examined both sex and labor trafficking. In a 2017 study, Cruyff and colleagues examined any human trafficking victimization among adults and children in the Netherlands between 2010–2015. They estimated that between 6,000 and 9,500 individuals were victims of human trafficking each year over the study period. A key strength of this study is that it was one of the first MSE studies to incorporate covariates for year, age, gender, and type of exploitation. Further, because another MSE study existed that covered one of the same years, the authors were able to demonstrate that including covariates led to a higher estimate of victimization compared to the study that did not include covariates. This was feasible in part because of the thorough recordkeeping by various government agencies in the Netherlands.

Sharifi Far and colleagues conducted a simulation study in 2020 using data from two prior studies in the United Kingdom and Romania, which had little overlap between the administrative datasets used. Having little overlap, or sparse MSE, is a key shortcoming of this method because the underlying concept relies on using the overlap between lists to measure where the datasets do not overlap to get a final prevalence estimate. This study tested two methods to handle sparse MSE, including combining datasets and leaving one dataset out. Based on this approach, they estimated 11,313 victims in the UK sample and 921 victims in the Romania sample, and in the article, described how different methodological choices can affect estimates.

A key assumption of MSE is that lists are independent, meaning that the likelihood of being on one list does not influence the likelihood of being on another list. In the Cruyff et al study (2017), the authors identified positive correlations between some of their lists. When positive correlation is present it likely indicates the presence of referral mechanisms, for example, the police refer victims to non-governmental organizations as standard practice. The presence of referral mechanisms violates the assumption of independence between lists. Because the authors tested for positive correlation, they were able to adjust for this in their modeling.

One limitation of the Cruyff et al study (2017), which is also a concern for any MSE study, is that external factors affected the administrative data used. In this case, the definition of human trafficking used in the Netherlands changed during the study period, and the Dutch police changed who they considered to be a potential victim (to be more conservative) which the authors noticed was associated with fluctuations in the total victims

reported. Another limitation, present in both studies and also a common limitation of MSE studies in general, is that administrative lists identify potential victims who may not be actual victims.

Exhibit 9 provides high-level strengths and weaknesses of MSE and capture-recapture methods as whole. Since these methods rely on the same statistical assumptions, they have similar strengths and weaknesses

Exhibit 9. Multiple Systems Estimation and Capture-Recapture Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Recent statistical advances allow for flexibility in model assumptions by modeling list dependence and capture probabilities (for MSE) or probability of being surveyed in a given location (capture-recapture).	The standard capture-recapture model has unrealistic assumptions that cannot be easily verified (e.g., mutual independence, homogenous inclusion probabilities).
Breadth of research	One of the most mature methodologies for estimating hidden population size.	None noted.
Data	Data already exist and does not require new data collection (MSE). New advances greatly improve success of record linking with incomplete data (MSE).	Requires relatively good administrative data and record keeping and at least three lists with victim information, including identifiers and covariates (MSE). External factors can affect administrative data (changes to victim definition, policy priorities) (MSE). Relies on providers to be knowledgeable about human trafficking victimization, ideally use validated instruments to screen, and record sufficient covariate information for analysis (MSE). Capture-recapture using survey data collection requires skilled field data collectors who can approach and survey individuals in public locations. Requires detailed record keeping to simultaneously protect participant confidentiality and document which participants are observed in more than one time point.
Who can be counted	Leverages the ability of service providers to build rapport and trust with victims that may not have been willing to engage with researchers.	Samples are biased toward people who are likely to engage with service providers/police/etc. Lists rely on willingness of victims to share their personal experiences with service providers.
Forms of trafficking it can estimate	Works best when fewer covariates are needed (i.e., to measure one form of exploitation among one population), but with a large enough sample that models can account for different forms of exploitation and other strata.	Sex and labor trafficking occur in different sectors and geographic areas. They may not get overlap between service provider lists if attempting to measure both.
Consideration of lived experiences of victims and other interested parties	MSE does not require victims to share personal details with researchers.	In general, does not engage with victims, service providers, or other individuals with lived experiences. Participants must be willing to answer questions related to victimization for survey-based capture-recapture studies.
Cost	"MSE is probably the least expensive method to produce prevalence estimation of forced labor" (Zhang and Larsen 2021). If multiple institutions in an area are already set up to detect and support victims of trafficking, MSE is a feasible and low-cost option. Capture-recapture using survey data collection is a high-cost effort.	In the U.S., there are no national referral mechanisms for victims of human trafficking which means researchers must undergo significant effort to identify and link sufficient administrative datasets.
Other	MSE is a good option when researchers are interested in repeating estimates to identify trends.	Data may not be able to be published for replication of estimates due to the possibility of deductive identification of victims.

3.1.6 Respondent-Driven Sampling

Respondent-Driven Sampling (RDS) is a network-based sampling method frequently used to estimate the prevalence of small, hidden populations, or populations that generally do not wish to participate in research (Heckathorn, 1997). At a 2019 National Academy of Sciences workshop, RDS was highlighted as a best practice for researchers attempting to estimate the prevalence of human trafficking in the U.S. population (White, 2020). RDS combines features from network analysis and snowball sampling but has methodological advantages over simple snowball or convenience samples that allow it to approximate a random sampling design (Heckathorn, 1997). Importantly, RDS allows researchers to draw a statistically representative sample and consequently produce unbiased population estimates. RDS has five key assumptions: 1) seeds (the first round of participants that begin the recruitment chains) and the individuals they recruit are known to each other and recognize each other as members of the target population (reciprocity); 2) smaller networks associated with each seed are linked to a single network; 3) sampling occurs with replacement so that successive recruitments do not deplete the pool of potential future recruits; 4) respondents accurately report the size of their networks; and 5) respondents are randomly selected from their recruiters' network list. Given these assumptions, using RDS is typically only appropriate among groups with clear, identifiable social connections (which often means within one industry and/or geographic area).

Four studies in our review used RDS—one examined labor trafficking and three examined sex trafficking. Two studies included estimates for adult victims only and two included child victims only.

In one of the RDS studies, Grosso and colleagues (2018) estimated the prevalence of childhood sexual trafficking among a sample of adults by asking them to report on experiences they had as children. A key limitation of this study is potential issues in recall. The authors found that an estimated 20 percent of participants were sexually exploited as children. Vincent and colleagues (2021) used a hybrid of RDS and link-tracing to estimate the prevalence of trafficking among sex workers in the city of Muzaffarpur, India. They found that an estimated 12.1 percent of sex workers were victims of sex trafficking. Thus, while they generated a reliable estimate that accounted for many potential biases in RDS studies, the prevalence estimate only applied to the sex worker population in one city. Extending RDS methodologies to a larger geographic area or population may not be feasible. While cost and resource constraints are the more obvious limitations to obtaining a generalizable estimate, a larger concern relates to a key underlying assumption of RDS—that all members of the population belong to one latent social network. This assumption is difficult to verify, but in narrow applications of RDS such as Vincent and colleagues (2021), it can be justified without testing.

Robinson (2023) used RDS to estimate the prevalence of forced labor among adult male and female victims in the brick kiln industry in Pakistan. This study estimated that 33.7 percent of brick kiln workers were victims of human trafficking. This study was part of the PRIF and also used probability proportional to size (PPS) to estimate the prevalence. Interestingly, they found that the RDS estimate at all thresholds of forced labor victimization was higher than with PPS (overall 33.7 percent for RDS versus 18.1 percent for PPS). A similar pattern was found in Anderson and Kendall (2023), who examined the prevalence of sex trafficking among children in one city in Brazil (Recife) using RDS and NSUM. In this study, the authors found that the RDS estimates were higher (RDS identified 22,600 victims while NSUM identified 19,700 victims). The authors are not sure how to explain the difference in these estimates aside from variation in sampling, but more information may be available once these studies are formally published.

One challenge with RDS studies is calculating the target sample size. To obtain an estimate of the sample needed to generate a prevalence estimate, researchers have to include an estimate of the variance of the key study variable (human trafficking victimization). However, since the prevalence of victimization is unknown, it is difficult to know what value to use. In Grosso et al (2018), they used the prevalence of HIV for this input but did not provide much explanation as to why this is a suitable proxy. However, they also recruited more than the target sample size so perhaps this was not a concern in this study. Many RDS studies choose to follow guidance from the World Health Organization (2013) which includes using a variance of .5, the largest possible variance for a binary variable.

Another challenge with RDS studies, particularly as it compares to studies that use administrative data, is that they require experienced, high-quality field teams who can approach strangers using information from referral coupons and survey administration. This type of field work requires persistence, good people skills, and high attention to detail to ensure that teams record all required network characteristics from each respondent. A final limitation of RDS is that, due to the design beginning with only a few seeds and relying on depth of social networks to reach sample size, research teams may not know if they will obtain the needed sample size until relatively late in the fielding process. This means that they may incur substantial costs only to realize that they will not obtain the sample they need to generate a reliable estimate.

A summary of the strengths and limitations of RDS studies for measuring human trafficking victimization is provided in **Exhibit 10**.

Exhibit 10. Respondent-Driven Sampling Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Assumptions are relatively easy to meet in study groups with identifiable social connections (e.g., employees in same industry in same geographic area).	Assumption that all individuals belong to one latent social group is less likely to hold up the more broadly defined the population becomes.
Breadth of research	RDS has been used for more than ten years to measure hidden populations.	Limited evidence on how representative RDS estimates actually are.
Data	Leverages participants to contact and recruit their peers rather than researchers having to do cold introduction and recruitment. Many RDS studies fail to achieve the target sample size and have to add new seeds. RDS is a good option when no sampling frame is available from which to identify respondents.	Hinges on the research team's ability to recruit the most appropriate seeds. Requires a good field team to access participants, build trust, and gain buy-in for participants to recruit from their network. Respondent behavior, including how they choose to form social connections, identify those eligible for coupons/referrals, and count their network size, can influence the validity of estimates. Seeds are most likely to give coupons to people they know the best, which may not be the people who best meet the definition for inclusion.
Who can be counted	Appropriate for victims of any form of exploitation if there is reasonable expectation that victims are socially networked.	Socially isolated people will be left out. Those with safety concerns may not participate or be selected out, leading to a less-representative estimate.
Forms of trafficking it can estimate	Has been used to estimate sex and labor trafficking populations. Most appropriate when applied to smaller geographic areas.	Less appropriate for large geographic areas, industries where there is little expectation of social networking among victims within industry (e.g., victims of online sexual exploitation).
Consideration of lived experiences of victims and other interested parties	Helps include the participant voice in research because researchers must interact substantially with key stakeholders.	Participants must be willing to answer survey questions related to victimization.
Cost	None noted.	Relatively cost intensive due to a dual incentive system and need for field team. May need larger sample size than initially estimated, making it harder to budget. Requires an initial leap of faith—researchers don't know if they will obtain the needed sample until late into recruitment.

3.1.7 Vincent Link-Tracing Sampling

Vincent link-tracing sampling (VLTS) uses the same recruitment process and underlying logic as RDS—that when individuals in the population of interest are difficult to identify, researchers can use social connections from a few members to get a large enough sample size. The main difference between VLTS and RDS is that VLTS begins with a larger, more representative sample of initial seeds and uses only two or three waves of sampling. VLTS also typically begins with analysis of administrative data to identify geographies from which to select seeds—more akin to a conventional sampling process—whereas RDS begins with a true convenience sample of a small number of seeds that are expected to yield a deep and diverse social network. In VLTS, researchers then link network characteristics of participants within each wave and across all respondents in the final sample, using respondent information and administrative data. As with RDS, VLTS can yield a final sample that approximates a random sample, allowing for inferences to be made about prevalence.

Five studies in the environmental scan used VLTS—four studied labor trafficking and one sex trafficking. All examined adult populations except for one labor trafficking study that focused on children. The sex trafficking article by Vincent and colleagues (2021) was also described in the RDS section above and found that an estimated 12.1 percent of sex workers in one city in India were victims of sex trafficking.

Dank (2023a) used VLTS to estimate the prevalence of forced labor among adult female domestic workers in Tanzania. One challenge encountered by researchers was that it was difficult to implement VLTS with this population because they are mobile, and VLTS requires tracking and tracing links within each wave and among the final sample. This resulted in an estimate with a large range (2,000 to 16,000) for this study. This method is not ideal when there are many isolated networks that do not connect together—thus, VLTS is not ideal for mobile populations (likely the case for forms of labor exploitation that involve migration).

Sahai (2023) used VLTS to estimate the prevalence of trafficking among female domestic workers ages 16 and older in Casablanca, Morocco. This study, which was part of the PRIF, also used capture-recapture methodology to estimate the prevalence of trafficking among this population, and it compares the results of the two methods. The researchers identified the 93 seeds for the VLTS estimate during the capture-recapture survey process. The VLTS method estimated a smaller population overall compared to capture-recapture, but a higher rate of trafficking within the identified population compared to capture-recapture, which calculated a larger population with a smaller rate of trafficking. For the moderate threshold of victimization, capture-recapture generated a 27 percent prevalence of forced labor violations while VLTS generated a 32 percent prevalence of forced labor violations.

Another study led by Dank (Dank, 2023b), used VLTS to estimate forced labor among adult male and female fishers in Costa Rica. In this study, researchers found that the fishing communities were tightly knit and well connected to one another. They used 98 seeds and a final sample of 1,009 individuals, which generated an estimate of 8,029 individuals working in the fishing industry. This was much less than with proportional probability sampling, which estimated an estimate of 26,295 individuals. Both methodologies led to similar rates of forced labor—VLTS estimated 23.3 percent in the moderate threshold and PPS estimated 20.2 percent. The VLTS sample was also more diverse than the household survey sample, with 12.4 percent of respondents not from Costa Rica compared with 5.1 percent. Overall, however, few non-Costa Ricans were captured in either sample.

A key limitation of VLTS studies is that the most isolated, and thus perhaps the most vulnerable populations, may not be reached. A summary of the strengths and limitations of VLTS studies for measuring human trafficking victimization is provided in **Exhibit 11**. Since both RDS and VLTS share similar underlying concepts, many of the strengths and limitations are similar.

Exhibit 11. Vincent Link-Tracing Sampling Strengths and Limitations for Human Trafficking Estimates

	Strengths	Limitations
Assumptions	Main assumption of network connectedness is that it is relatively easy to meet in study groups with identifiable social connections (e.g., employees in same industry in the same geographic area). Yields a more efficient population estimation via linking networks of the final sample, not just connections within each wave.	Assumption that all individuals belong to one latent social network so estimates may only apply to networked/connected members of the population.
Breadth of research	Based on adaptive sampling techniques (which includes RDS) which have been extensively applied to human trafficking estimation. VLTS validated through empirical simulation of observed data for identifying HIV/AIDs in one study.	Newer methodology and fewer applications to human trafficking relative to other methods in the review.
Data	Leverages participants to contact and recruit their peers rather than researchers having to do cold introduction and recruitment. Leverages auxiliary information, not just information recorded about individual's network characteristics, allowing unbiased estimates after only two or three waves. Researchers have more control over recruitment than in RDS because they start with a larger sample of seeds. By sampling wide rather than deep, sample will be more diverse than in RDS.	Requires a good field team to access participants, build trust, and gain buy-in for participants to recruit from their network. Respondent behavior, including how they choose to form social connections, identify those eligible for coupons/referrals, and count their network size, can influence the validity of estimates. Seeds are most likely to give coupons to people they know the best, which may not be the people who best meet the definition for inclusion.
Who can be counted	Works well with populations that are highly networked. Can be used to capture victims in larger geographic areas relative to RDS.	Due to the assumption that all members of the subpopulation have social connections, this method will not work well for socially isolated victims. Those with safety concerns may not participate or may select out, leading to a less representative estimate. Not suitable with highly mobile populations, as they may have fewer local social connections.
Forms of trafficking it can estimate	Predominantly used to estimate labor trafficking but may be appropriate to estimate sex trafficking if the argument can be made that individuals are tightly networked.	Trafficking victims must be socially networked to be counted.
Consideration of lived experiences of victims and other interested parties	Helps include the participant voice in research because researchers must interact substantially with key stakeholders.	Participants must be willing to answer survey questions related to victimization.
Cost	Due to fewer waves of recruitment (relative to RDS) there is less risk that the final sample will not reach equilibrium (which is only known in RDS after five or more waves of recruitment).	As with RDS, relatively cost intensive due to a dual incentive system, need for experienced field team, and potentially large sample size.

4. Conclusion

Estimating the prevalence of human trafficking in the U.S. is a challenging undertaking. Developing a methodology (or set of methodologies) that can produce an unbiased prevalence estimate requires: 1) balancing precision and generalizability; 2) aligning methods to the data sources available; and 3) aligning methods to the characteristics of the population of interest (to the extent these characteristics are known). **Exhibit 12** provides a high-level summary of the key strengths and limitations of the methods reviewed in this environmental scan.

4.1 Balancing Precision and Generalizability

A key theme identified during the environmental scan is that there is a tradeoff between an estimate's precision and its generalizability. The more narrowly defined the sample becomes, the more reliable the estimate will be. This is particularly true for network-based methods like RDS and VLTS, because they rely on an assumption that individuals within the target population belong to a connected social network. This assumption is more valid when considering, for example, domestic workers within one city versus domestic workers in an entire country. The utility of an estimate, however, also declines as the population it can be generalized to shrinks. If BJS aims to estimate the prevalence of human trafficking for the nation, understanding the prevalence for one industry in one city may not be sufficient, even if the estimate is precise.

A key consideration for BJS is therefore to determine how to classify human trafficking into measurable components. While most methods can be applied to sex or labor trafficking, it is less feasible (and less common) to measure both in the same study because victims of these distinct forms of exploitation often do not belong to the same larger population.

For sex trafficking, separating adult and child victims is one avenue worth exploring. One of the commonly cited challenges to estimating the prevalence of human trafficking in the U.S. is that practitioners (e.g., law enforcement, service providers, district attorneys) and victims themselves have a hard time defining what constitutes human trafficking. This can be an issue for studies that leverage administrative data, and those that rely on respondents to be aware of which of their contacts may be a victim of human trafficking. Because children cannot consent to any form of sex work, child sex trafficking may be a subpopulation for

which data are more readily available. Additionally, this is a subpopulation that receives substantial public attention and attention from researchers who have published wide-ranging estimates of this crime, highlighting the need for a rigorous, reliable estimate. MSE may be even more efficient for estimating prevalence among child populations because children may not be required to consent/assent to having their information included in lists that may be used for MSE studies (due to mandatory reporting laws). Additionally, one study in the scan surveyed adults and asked them to report retrospectively about their experiences as children, which may be a good option given the potential ethical concerns of surveying children.

When examining labor trafficking, keeping estimates within industry and geography is a common tactic; however, to achieve coverage of the entire country in this way would be costly and likely not feasible. When considering what industries to prioritize, those where employees are more socially networked and more likely to make public appearances, or in industries that employ a greater share of workers, are areas for consideration. The NHTH, for example, provides data on victims who call into the hotline and can shed light into the industries and geographies where labor trafficking may be more prevalent. This data can potentially serve as a useful starting point when attempting to prioritize industries and geographies for estimation.

4.2 Aligning Methods to the Data Sources Available

In general, studies in the environmental scan are sorted into those that use surveys and those that use existing data sources. A key benefit of survey research is that researchers can tailor the questions and sampling strategy to the research questions of interest, rather than having to use data designed with an entirely different purpose in mind to meet the needs of the study. Regardless of whether a survey-based or administrative data-based method is used, however, researchers are likely to need to leverage existing data sources. Probability-based sampling, NSUM, and VLTS all require a list defining the population from which to construct a sampling frame, and methods like MSE need administrative data more directly to construct the database from which to carry out the analysis.

One factor that facilitates the process of estimating the prevalence of human trafficking in the U.S. is that the U.S. has relatively good administrative record keeping, especially at the state and federal level. One way that BJS could support prevalence estimation in the future may be to identify data sources that could be used for MSE studies and help to enhance these data collections or the way they are published to be more useful for researchers. For example, federal data from agencies including the U.S. Marshals, the Executive Office of U.S. attorneys, and the U.S. Sentencing Commission provide data on known human trafficking cases and defendants. These data are typically not linked to any victim characteristics, but BJS may be able to use presentence reports and other documents collected by law enforcement or the U.S. Sentencing Commission to produce deidentified victim data, or data on the location of crimes, to facilitate MSE efforts.

4.3 Aligning Methods to the Characteristics of the Population of Interest

As shown in **Exhibit 12**, some methods are better suited for different target populations. For example, capture-recapture, MSE, random/probability sampling, and time-location sampling can capture individuals who are not socially networked, while link-tracing and RDS are better suited to capture individuals who have social networks but may not be visible in public spaces or identified easily by service providers. Exploratory and qualitative research that sheds light onto the characteristics of known victims can help when determining which intersections of industries, victim subpopulations, and geographies may fall into these different categorizations. Across all the methods reviewed, it seems that none are suitable for identifying the most at-risk victims of trafficking—those who are completely geographically and/or socially isolated.

Exhibit 12. Use Cases by Sampling Method for Human Trafficking Estimates

	Ability to Produce Unbiased Estimates (in Theory)	Ability to Produce Unbiased Estimates (in Practice)	Can Account for Individuals with Limited/No Social Network	Can Account for Individuals Who Are Not "Visible" in Public Spaces or by Providers
Capture-Recapture/Multiple Systems Estimation	✓	✓	✓	●
Convenience (and its variations such as Census)	●	●	●	●
Link-Tracing	✓	✓	●	✓
Network Scale Up	✓	●	●	✓
Random	✓	✓	✓	✓
Respondent-Driven	✓	✓	●	✓
Time-Location	✓	●	✓	●

The Abt team examined these areas and others during an expert panel meeting on February 28, 2024, and throughout their review and analysis. Findings from this scan, the expert panel, additional outreach with expert panel members, and conversations with BJS, inform our proposed recommendations.

Volume II:
Advisory Panel Meeting on Human Trafficking Estimation Summary

1. Introduction

On February 28, 2024, the Bureau of Justice Statistics Advisory Panel Meeting on Human Trafficking Estimation was held at the Office of Justice Programs in Washington, DC. The meeting included experts invited to serve on the Advisory Panel, representatives from the Bureau of Justice Statistics (BJS) and the National Institute of Justice (NIJ), and Abt Global (Abt) study team members. The purpose of the meeting was to generate a discussion that would support BJS's overall objective to produce estimates of the prevalence of human trafficking in the U.S. by developing a framework that BJS can use when designing a pilot study. This volume of the report provides a summary of the meeting submitted to BJS.

1.1 Participants

Participants included experts making up the Advisory Panel, as well as staff from BJS, Abt, and NIJ. **Exhibit 13** details our expert panelists.

Exhibit 13. Expert Panel Members

Name	Title/Affiliation
David Banks, Ph.D.	Professor, Duke University
Meredith Dank, Ph.D.	Research Professor, New York University Marron Institute of Urban Management
Davina Durgana, Ph.D.	Visiting Professor, University of Oxford Quantitative Research Lead, Minderoo Foundation
Leah Meyer, MPP	AI Policy Lead, *Independent Researcher*
David Okech, Ph.D.	Professor, University of Georgia Director, Center on Human Trafficking Research & Innovation
Sheldon Zhang, Ph.D.	Professor, University of Massachusetts Lowell

1.2 Summary of Key Decision Points from Meeting

Trafficking Type

Framing the data collection effort as a labor survey and measuring exploitation in various work categories has benefits for response rate, policy implications, and generalizability.

Geography

There is a better chance of success if BJS focuses on metropolitan statistical areas (MSA) because this geographic scale bridges the gap between national and local levels. An added benefit is that focusing on MSAs or dense population centers gives better coverage of the national population overall.

Methodology

BJS should identify the population that it wants the estimate to be generalizable to prior to selecting a methodology for prevalence estimation. Some methods are better suited if focusing on localized or industry specific estimates (e.g., VLTS, RDS) but not suitable for broader efforts. Capture-recapture (with survey) is a good alternative but only for visible industries and/or victims.

Data Access and Quality

Concerns about administrative data quality, storage, and accessibility limit confidence in possible MSE studies. However, the group acknowledged that the U.S. has robust data collection systems that are underutilized and may have potential for use in MSE.

Repeatability vs. Coverage

Repeating estimates over time is an important value-add that BJS is uniquely suited to provide. Every 3-5 years would be sufficient. However, trafficking hot spots move over time so half of the group would favor breadth of coverage (additional subpopulations/areas) over repeatability.

2. Key Considerations and Discussion Points by Agenda Item

2.1 National Versus Subnational Estimates

Participants discussed the utility and advantage of BJS pursuing national versus subnational estimates; the conversation showed that the experts were mixed in opinion.

> *"A national estimate is too distracting."*

National estimates fail to capture nuances of human trafficking victimization within local contexts. They also may not be as useful at local levels to inform law enforcement and service provider programs and policies. Subnational estimates are more useful for actionable responses, which was of greater interest to some participants.

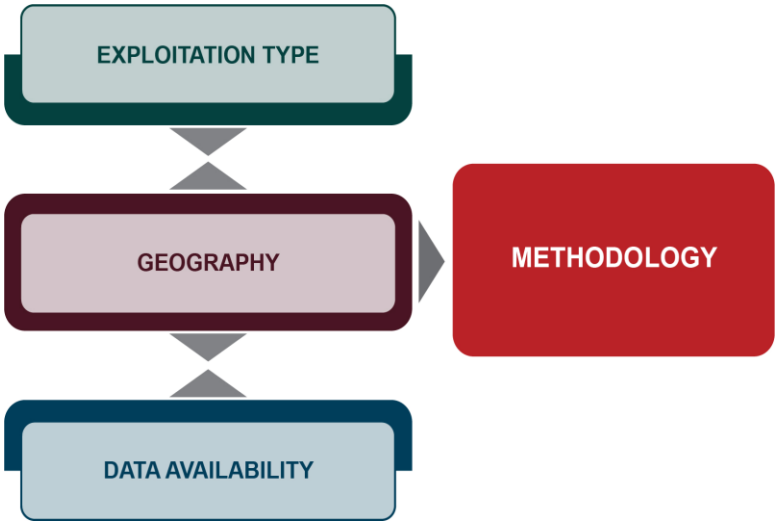
> *"National estimates are happening with or without us, we should focus on creating a reliable one."*

The U.S. is well suited for a national estimate because of high data availability and clear laws related to human trafficking. Further, when discussing the value-add of BJS providing an estimate versus local research efforts, experts agreed that BJS is uniquely suited to provide a national estimate, despite the challenges with this approach. Still, the realities of human trafficking can vary significantly across different regions and communities. Because of this, not all methodologies are effective or applicable across different populations.

2.2 How to Break Down Human Trafficking into Estimable Components

During the discussion, experts explained that there is an important interplay between types of exploitation, geographic area of focus, and data availability (**Exhibit 14**). BJS must identify the population of interest prior to selecting the appropriate methodology. This decision should be based on BJS's policies, priorities, and anticipated uses for the estimate, as well as what is known about human trafficking in the U.S. and where the country is likely to be most successful in identifying victims. Key considerations for each category are discussed below.

Exhibit 14. Key Considerations When Selecting a Methodology for Population Size Estimation



2.2.1 Forms of Trafficking and Exploitation

Participants discussed key considerations related to selecting a form of trafficking and exploitation to measure.

Key Considerations: Forms of Trafficking and Exploitation

- Labor trafficking is the most pressing area of focus.
- It is important to measure the spectrum of exploitation and to draw the threshold for trafficking out from there.
- In line with international efforts, sex work could be included as a category of work in a labor survey. This would not account for all forms of sexual exploitation/trafficking.
- It is important to focus on known covariates for exploitation, including industries with a low barrier to entry and populations with limited economic mobility.
- The group acknowledged the importance of growing rates of sexual exploitation online, which would likely need to be a separate effort.

Discussion

Participants agreed that focusing on labor trafficking is a more favorable starting point than sex trafficking since labor laws are more clearly defined than laws related to sexual exploitation, which change frequently. Labor trafficking is also under-researched and underreported compared to sex trafficking, despite growing literature aimed to measure the prevalence of labor trafficking largely outside of the U.S. context.

Within labor trafficking, experts recommended examining different industries (e.g., agriculture, domestic servitude, construction) due to population differences and geographic considerations within each industry. Additionally, there are within-industry complexities that are important to study and account for in methodological approach.

Participants also discussed the need to consider online sexual exploitation. They discussed some challenges with this approach, such as not knowing where someone is being exploited (i.e., difference between geographic location of the online advertisement versus physical location of the victim). Additionally, the participants discussed the difference between using the internet for advertisement of in-person sexual exploitation (e.g., back-page ads) and exploitation that takes place on the internet or virtually.

There was a short discussion about the value of asking open-ended questions about experiences with exploitation and trafficking to gain insight into forms of trafficking that may not be expected within the study design.

2.2.2 Geography

Participants discussed key considerations related to selecting a geographic area for a prevalence estimation study.

Key Considerations: Geography

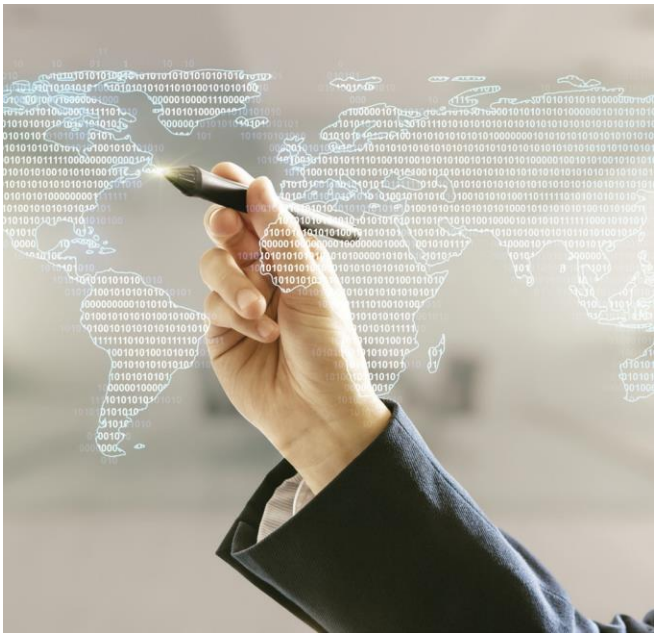
- Focusing on densely populated areas like metropolitan statistical areas, which can provide a better chance at identifying victims and allow for greater generalizability.
- Identifying correlates of human trafficking to identify geographies (e.g., proximity to international borders, foster care system size)
- Taking into account that trafficking is not stagnant and hot spots may change over time.
- Identifying places where barriers to entry are removed (e.g., Nevada has legalized sex work).
- Noting that safe haven cities may be of interest due to higher populations of foreign nationals.

Discussion

There was some overlap with the earlier discussion about national versus subnational estimates. Participants provided criteria to consider when selecting geographic areas to prioritize. In addition to the points identified in the box above, participants discussed the benefits of two distinct approaches for selecting geographic areas:

- The geography-first approach focuses on the most populous regions in the U.S. This approach allows for the potential extrapolation of a national estimate.
- The industry-first approach focuses on industries particularly related to labor trafficking. This approach aims to select the most appropriate region based on the industry, such as agriculture. This approach would lend itself better to a subnational estimate.

Participants concluded that the chosen approach to identifying geographic areas will ultimately depend on the specific objectives of the study, as each approach is distinct to the type of trafficking and generalizability to other regions/domains.



2.2.3 Data Sources

Participants discussed different data sources that can be used to estimate the prevalence of human trafficking or aid in study design and planning.

Key Considerations: Data Sources

- Phone or online based surveys suitable for quick and easy" national estimates.
- Overall preference for original survey data collection.
- Abundance of data sources that may be helpful for constructing sampling frames, such as:
  - Visa lottery trend data.
  - Various administrative (non criminal) data sources to generate sampling frame.
  - Corporate supply chain.
  - Cyber footprint (especially for sex trafficking).
  - Spatial temporal mapping.
- Although the U.S. collects robust administrative data, experts were hesitant about MSE studies due to a lack of consistency in definitions, data entry, and data storage. Despite these concerns, possible sources for a MSE study include:
  - Healthcare data.
  - Enforcement data and data from Enhanced Collaborative Model grantees.
  - Service provider data (but need to be mindful about referral mechanisms between enforcement and service providers).

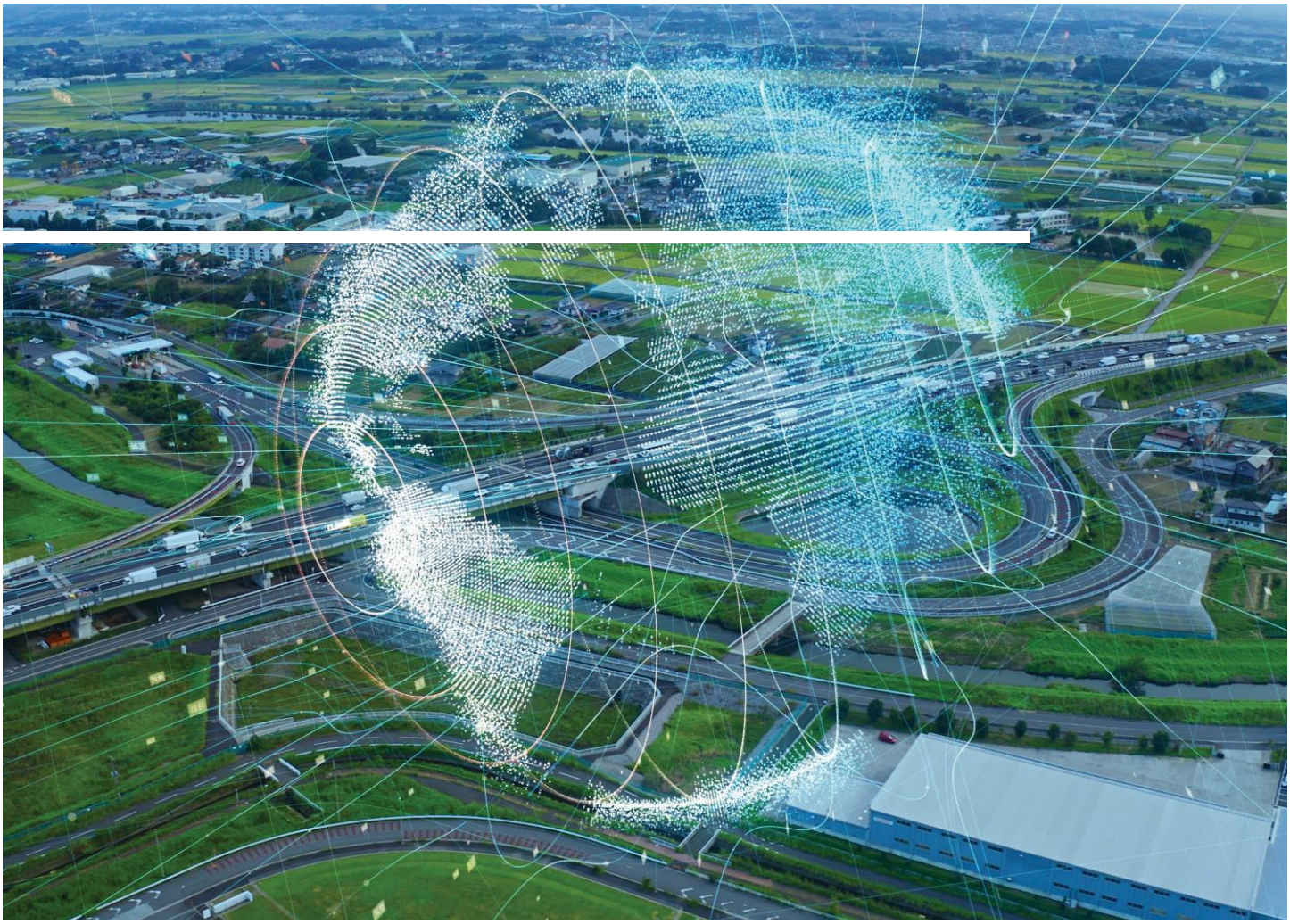
Discussion

Participants discussed issues related to the data governance policies of collecting agencies. For example, the quality of data collected, definitions, and quality control vary significantly between agencies, limiting the use of data by service providers and law enforcement. Overall, participants did not view the FBI's Uniform Crime Reporting program's National Incident Based Reporting System (NIBRS) as a reliable source of truth for capturing human trafficking incidents or arrests because officers may classify an arrest based on the charge given the current information they are provided. Without training specific to human trafficking, officers may code the arrest for a different charge like abduction or solicitation.

Some sources recommended for consideration were:

- Healthcare data.
- Enhanced collaborative model for human trafficking data.
- Internal law enforcement incident tracking systems.
- Hotline data.
- Public awareness campaigns.
- Immigration data (e.g., visa lottery).

Recommendations for BJS to make existing data more accessible were also provided. Participants also recommended creating a clearinghouse for federal data where data could be deidentified and anonymized to be made publicly available or available for restricted use by researchers.



2.3 Methodology

Participants discussed the pros and cons of different methodologies, separately for sex and labor trafficking estimation and are provided in **Exhibit 15.**

Exhibit 15. Expert Feedback on Different Prevalence Estimation Methodologies

	Expert Feedback
MSE and capture-recapture	• MSE/capture-recapture are fairly robust. That would be a good choice for an estimation pilot.
Link-tracing sampling (LTS)/Vincent link-tracing sampling (VLTS) and respondent-driven sampling (RDS)	• In places where sex work is legalized, barriers to entry are removed. In this case, LTS or RDS could be feasible for sexual exploitation. • VLTS would be good use nationally. However, the Prevalence Reduction Innovation Forum (PRIF) has led to more questions than answers (e.g., methods, feasibility, geography). Assumptions about links have been wrong, and people are not as linked as they thought. • VLTS/LTS is sensitive to the model. It does not work well in situations where there are no commonalities across networks. It might work better for domestic sex trafficking. • There are some industries where laborers/crews might shift around and know each other (e.g., construction) but in other industries teams can be isolated (e.g., agriculture).
Network scale-up method (NSUM)	• Concerns about NSUM, especially for sex trafficking. It is an unrealistic estimate of the social network.
Random sampling	• Probability sampling is the gold standard. • Probability sampling might be the best of bad options when victims do not know one another/ have connected social network. • Probability sampling is expensive and impractical. • Probability sampling would work well for visible industries.
Economic approach	• If the cost of sex is low there might be a high demand, which could be a trafficking proxy.

2.3.1 Sex Trafficking

Participants reported that methodological preferences to assess sex trafficking depend on the victim population. In particular, they pointed out important differences between internationally and domestically trafficked victims. Internationally trafficked victims may have strong networks tied to individuals from their home country, but not to victims from and living in the U.S. Participants mentioned that if the aim is to capture internationally trafficked victims then probability sampling is likely the best option. For victims who are from and living in the U.S. (and who are more likely to know other victims who are from and living in the U.S.), capture-recapture and link-tracing were preferred.

However, while the assumption has been that sex trafficking victims in the U.S. make up strong networks, participants noted that they are often not highly linked, limiting the overall use of link-tracing. Link-tracing is also highly sensitive and does not work effectively when there are little or no commonalities across networks.

Participants agreed NSUM was the least preferred method for sex trafficking estimation and labor trafficking.

Participants also discussed the possibility of using an economic approach to labor trafficking based on cost of sex, for example, as a proxy for supply/demand balance. This discussion highlighted space for potential new methodologies or combined methodological approaches.

2.3.2 Labor Trafficking

Participants were split between preferring probability sampling and capture-recapture if the industry of focus is a visible industry. Participants agreed that network-based methods like link-tracing or RDS were preferrable for less visible industries.

2.3.3 Considerations for Methodology Selection

Participants documented the following questions that they believed BJS should critically examine when selecting methodologies:

- Do the victims know each other? When identifying a population, consider how particular types of trafficking networks operate and the likelihood that victims within those networks interact both with each other and with service providers/government agencies/anyone else who is the intended recipient of the survey.
- How practical and cost-effective is this method? The best method in theory may not work well in the field.
- What is the target population and geography? The U.S. population is not homogenous—some populations may be overrepresented depending on the method and the source.
- How visible is the type of trafficking that is the focus of the study? This is important to consider while thinking about the limitations and assumptions of different methodologies.

2.4 Summary of Discussion Points

Exhibit 16. Key Areas of Consideration

Trafficking Type	Subpopulation	Geography	Data Sources

• Within labor, consider:
  ○ industries with low barriers to entry.
  ○ (agriculture) which crops require manual labor vs. automation.
  ○ industries and geographic locations where there are clearly defined labor laws.
• Within labor and sex, consider:
  ○ the complexity of the network.
  ○ calling it a "labor survey."
	• Limited economic mobility or opportunity.
	• Consider online networks or platforms as a subpopulation.
	• Children vs. adults.
	• Domestic vs. internationally trafficked.
	• Language families.
		• Rural may be more difficult.
		• Major metropolitan areas preferred.
		• Identify correlates of human trafficking to identify geographies (e.g., international borders, foster care system size).
		• Consider that trafficking is not stagnant.
		• Places where barriers to entry are removed (legalized sex work).
		• Safe haven cities.
			• Phone or online based surveys for national estimates.
			• Visa lottery trend data.
			• Original collection/surveys.
			• Various administrative (non-criminal) data sources to generate sampling frame.
			• Healthcare data.
			• Enforcement data.
			• Corporate supply chain.
			• Cyber footprint (especially for sex trafficking).
			• Enhanced Collaborative Model recipients.
			• Spatial temporal mapping.

Possible Priority Areas

Trafficking Type	Subpopulation	Geography	Data Sources
• Janitorial services. • Landscaping. • Construction. • Online sex work.	• Migrant workers (expand to anyone migrating to the U.S. within country and international). • J1 visa holders. • Au pairs (if interested in domestic servitude). • People with disabilities. • Men/women. • Race/ethnicity. • Runaway/homeless young people. • LGBTQIA+.	• Large U.S. cities like New York, Boston, Los Angeles.	• Availability of data sources can guide geography selection.

Volume III: Recommendations for a National Estimate Pilot Study

1. Introduction

In this volume, we describe three options for a pilot study to test the feasibility of estimating the prevalence of human trafficking. We present recommendations in order of preference, with the first recommendation being the strongest recommendation and the third being the recommendation with the most significant limitations. These recommendations are informed by the environmental scan, the expert panel meeting held in February 2024, and key informant interviews.

1.1 Overview

Across all recommendations, the Abt team focuses on estimating the prevalence of **labor trafficking and exploitation.** Our environmental scan and conversations with key informants suggest that labor trafficking is more prevalent than other forms of human trafficking, making it a critical area to study. Indeed, approximately half of the estimated 27.6 million victims of labor trafficking worldwide are exploited in countries with developed economies, such as the U.S. (ILO, 2022). As evidence of the prevalence of labor trafficking in the U.S., labor trafficking victims make up over 74% of all trafficking visas issued to victims of human trafficking to date, approximately 1,360 annually (Polaris, 2021a). Another advantage in investigating the prevalence of labor trafficking over other forms of human trafficking is that labor laws are more clearly defined and change less frequently relative to the myriad of laws that can be applied to sexual exploitation. Further, labor exploitation has received less attention as compared to sexual exploitation in U.S.-based research. We designed our recommendations to be responsive to this research gap and expert panelists' guidance. Our recommendations are also responsive to government representatives and senators' expressed interest in investigating the prevalence of human trafficking—including labor trafficking—in the U.S., laying the groundwork for national estimates (Put Trafficking Victims First Act, 2024). The proposed approaches will contribute to a more comprehensive understanding of the scope, nature, and entry points into victimization of human trafficking in the U.S.

Building on this foundation, the study also prioritized methods that, if implemented at scale, could generate a **national estimate of labor trafficking victimization.** This strategic approach ensures that insights gained from the pilot study can be effectively translated into a robust and representative assessment. While there are limitations to estimating prevalence at a national level, the Bureau of Justice Statistics (BJS) is uniquely situated to conduct this work, which has become increasingly necessary to address congressional mandates. The expert panel explicitly identified BJS as a key stakeholder to lead this study and recommended selecting metropolitan statistical areas (MSA) for a national estimate. This approach, however, could limit generalizability as it would, by design,

exclude rural and tribal communities. Key informants expressed that local estimates, rather than national ones, could have more practical implications for law enforcement and victim service providers, given the legal structure of the U.S. The Abt team has therefore provided one recommendation that would use local estimates to build a weighted national estimate, one recommendation that would build a local estimate, and one recommendation for a national estimate.

The team also prioritized methodologies that would **allow BJS to produce detailed information about victims and consequences of exploitation/ trafficking to inform policy and practice**. Accordingly, our recommendations include methodologies that require original, primary data collection with potential victims of human trafficking. These recommendations intentionally engage individuals who have been exploited, enabling insight into the complex pathways that may lead to trafficking victimization.

Finally, the Abt team chose methods that could **permit comparison of trends in victimization over time**, allowing for an examination of the effects of policies or programs that aim to reduce human trafficking. The team's recommended methods also allow for repeatability of data collection and analysis techniques, enabling research that validates findings and builds upon them in future research endeavors. These methodological efforts will provide a foundation that is grounded in research and expert opinions to encourage further research for other industries. **Exhibit 17** provides a summary comparison of the three recommendations.

The first recommendation for a pilot study is a direct observation using capture-recapture, which offers the highest rigor but is also the costliest option, estimated at $1 million-$1.5 million. The second recommendation is a Vincent Link Tracing survey with an estimated cost ranging from $750,000- $1 million. Lastly, we recommend a random sampling survey which has a highly variable cost depending on the length of the survey and number of questions or options included. Some of our consulted experts suggested that costs for random sampling could be as low as $100,000, making it a more budget-friendly option with trade-offs in terms of representativeness. All of these cost estimates relate to the aforementioned pilot study. A full implementation study would be more expensive.

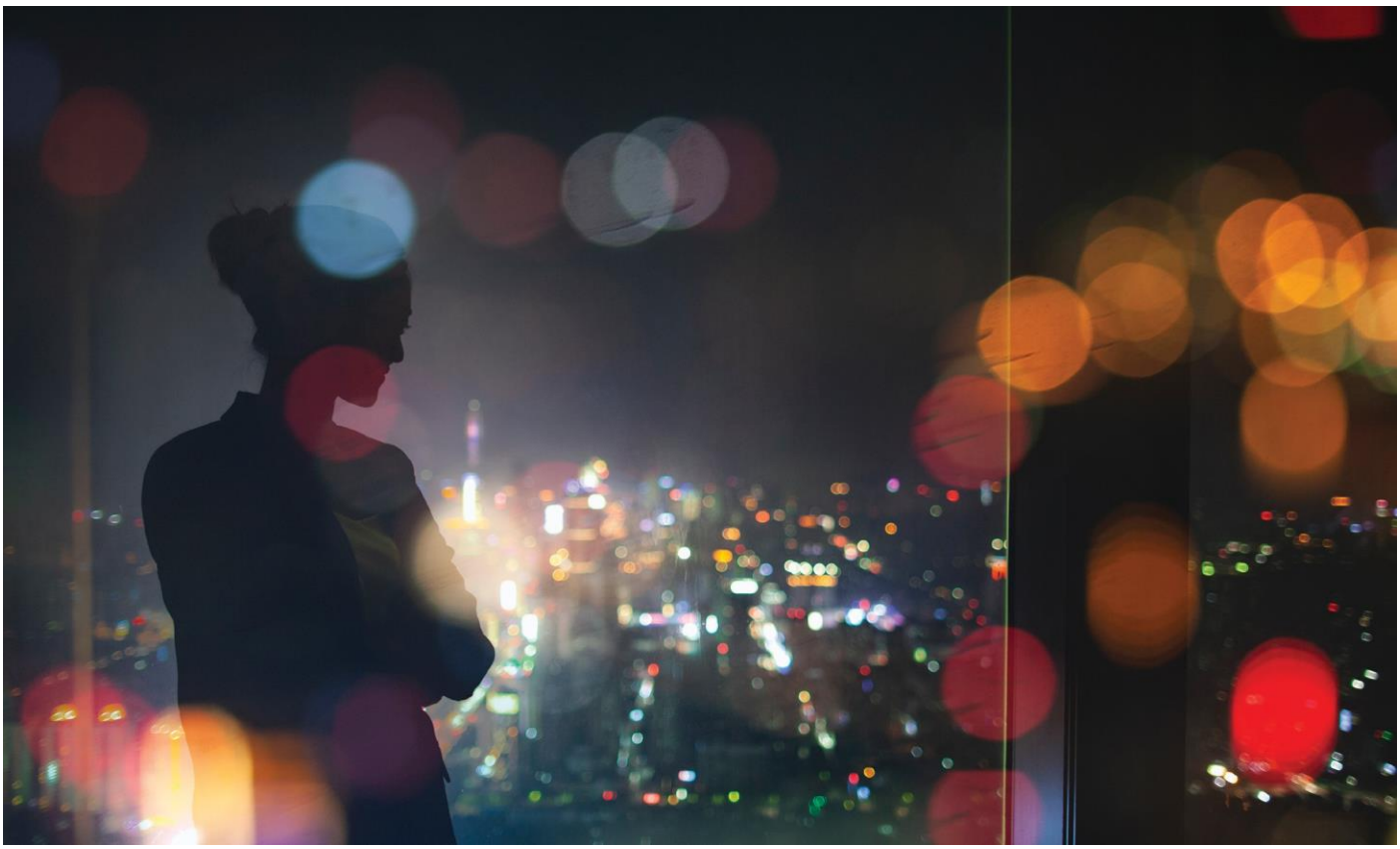


Exhibit 17. Summary Comparison of the Recommended Methodologies

Methodology	Cost	Scope of Estimate	Data Collection Method	Strengths	Potential Challenges
Capture-recapture	Estimated $1 million-$1.5 million	• Use local estimates to build a weighted national estimate	• In person	• Reliable method for studying hidden populations and can be used when a sampling frame is not available for random sampling. • Does not rely on social networking or connectedness, increasing the feasibility of studying labor exploitation across industries. • Allows for examination of a spectrum of labor exploitation experiences. • Data can be used as one data source in a multiple system estimation (MSE) study.	• Obtaining permission from local public transit authorities (or any venues) may be burdensome. • Need to scout venues for the right location that balances high traffic flow and staff safety. • Recapturing may be difficult if individuals select out because they already participated. • Must account for safety of potential victims who may be traveling with their traffickers.
Vincent link-tracing sampling (VLTS)	Estimated $750,000- $1 million	• Local estimate only	• Hybrid: In person/ phone & online	• Innovative method for studying hidden populations and can be used when a sampling frame is not available for random sampling. • Allows for examination of a spectrum of labor exploitation experiences with a given industry.	• Requires substantial input from the community to identify and recruit initial seeds. • Requires respondents to be knowledgeable about the victimization experiences (their own and their social connections). • In-person data collection can be costly; data collection via phone or web may provide cost savings. • Limits study to one industry and geographic area, limiting generalizability.
Random sampling/ proportional to size sampling	Estimated $100,000, for one to five questions added to a pre-existing online national panel	• National estimate	• Online	• Organizations that provide online panel data already have potential participants on hand for the study. • If done online using a pre- existing panel, represents substantial cost savings of in-person data collection methods.	• There is not a suitable sampling frame from which to draw a random sample of labor trafficking victims. • Given the low prevalence of human trafficking victimization in the general population, would require a large sample size. • Sampling frames used by panel data collection companies like those in our recommendations may not be appropriate for the population of interest, representing a possible threat to internal validity of the estimate.

1.2 Cognitive Interviews

Across all recommendations, we suggest conducting up to three rounds of cognitive interviews to inform and refine the survey instrument used. The research team should conduct cognitive interviews in English, Spanish, and additional languages identified during formative research in the selected locality. The research team can use cognitive interviews to improve question wording, ordering, and format, and to make any significant changes in question instructions and/or definitions as well as any changes in consent wording. For example, this process can help evaluate participant comprehension of data collection instructions, including explanations of any unusual procedures used with capture-recapture or Vincent link-tracing sampling (VLTS). This could include, for example, sending a text link to participants' cell phones during the initial data collection so the survey team can then use the location of the cell phone to identify if they were in the recapture area when the link was sent (capture-recapture). Additionally, findings from the cognitive interviews can help understand participants' comprehension of and willingness to consent. Cognitive interviews can also be used to adapt or refine questions that will be used to understand if participants meet thresholds of human trafficking, such as those identified by the Prevalence Reduction Innovation Forum (PRIF).

1.3 Ethical Considerations

All recommendations involve asking individuals about their experiences with labor trafficking victimization. This presents inherent ethical risks that the study team must account for in consultation with their Institutional Review Board (IRB) for human subject protection. For recommendations which involve in-person data collection, there are additional safety and ethical considerations related to the safety of participants and field team staff. The study team and the IRB should work together to put into place a safety plan that complies with all mandated reporting laws and protects respondents from potentially being exposed to their traffickers for participating.

Respondent safety. Whether in person or online, potential participants could experience harm from participating in a survey about their victimization experiences. There are a few ways to proactively avoid this outcome, which we describe below.

- Any study should have an informed consent process that makes it clear what participation entails, what kinds of questions will be asked, and recommends that the participant take the survey in a private space. The survey should also include informed consent language, along with instructions that participants can stop the survey at any time or choose not to answer questions.
- The study team should program online/mobile surveys with easy escape shortcuts that allow respondents to quickly exit the screen and prevent the web browser from repopulating from browser history.
- The study team should provide national and local resources (e.g., number to the National Human Trafficking Hotline, Victim Resource Center) to study participants in the event they want to report behaviors and/or seek help for their victimization. Resources may also include mental health resources for questions that are anxiety-provoking or may potentially retraumatize victims.
- Incentives, while encouraging participation, should not be coercive (e.g., they should be appropriate to compensate individuals for their time but not of a magnitude that compels people to participate who otherwise would not).
- For the capture-recapture recommendation, an additional ethical consideration is the use of cell phones to identify participants' locations (for recapture). Participants should understand that this is happening, and field staff should take steps to protect participant confidentiality, ensuring that participants' locations are not disclosed to anyone outside of the research team at any point during and/or after the study is complete.

Field team safety. For in-person data collection, it is critical that the field team is safe from harm. This may include precautions like setting up tables near security cameras and/or in visible spaces with a high flow of traffic. The team should plan to provide virtual incentives so that enumerators are not targeted because they are known to have gift cards on hand. Across all recommendations, any field staff involved in data collection should complete training that includes survey procedures, ethical protocols, and risks and field procedures specific to the study.

2. Recommendations for Pilot Study

The following sections detail our three recommendations for conducting a prevalence estimation study of labor trafficking in the U.S. The Abt team based our recommendations on rigorous research methods, our environmental scan, and discussions with key informants. We have carefully considered various approaches and evaluated their feasibility and cost to provide the most suitable options.

2.1 Recommendation 1: Capture-Recapture (Direct Observation)

2.1.1 Overview

The goal of the pilot is to test the feasibility of capture-recapture in one city with high public transit usage (greater than or equal to five percent of the population). There are two primary objectives for the pilot study:

1. **Assess the feasibility of capture-recapture:** Ability to recapture members of the target population.
2. **Assess the ability of the methodology to detect victims:** Ability to detect labor trafficking exploitation and victimization.

Following the description of the pilot methodology, the Abt team provides a high-level overview of how this method could be scaled up to generate a national estimate of labor trafficking among adults who use public transit.

2.1.2 Strengths of Capture-Recapture

Capture-recapture is one of the most mature methodologies for estimating the prevalence of hidden populations. Below are the method's key strengths that led to its inclusion as the top recommendation.

Capture-recapture is a reliable method for studying hidden populations and can be used when a sampling frame is not available for random sampling. Capture-recapture allows researchers to make inferences about hidden populations using a convenience, or non-random, sample. The underlying concept behind capture-recapture is that researchers identify all known members of a population in a given location and time, then repeat this process at another time point and use the overlap between lists of identified victims to estimate the total who were not captured at either data collection period.

Capture-recapture is appropriate for measuring victims who are able to access public spaces. Capture-recapture relies on identifying people in public spaces, which is common for victims of labor trafficking in some sectors like construction, hospitality, and janitorial services.

Capture-recapture allows for examination of a spectrum of labor exploitation experiences. For example, PRIF (2020), in collaboration with the U.S. State Department, identified 38 indicators of trafficking that the PRIF grantees used to measure trafficking and exploitation in their studies. The researchers used

the indicators to form thresholds for the likelihood of a survey respondent being a victim of trafficking or exploitation. The higher threshold was for clearly defined cases of trafficking, whereas the lower threshold was for cases of exploitation that could not be clearly defined as trafficking.

Capture-recapture does not rely on social networking or connectedness, increasing the feasibility of studying labor exploitation across industries. Other rigorous methodologies for prevalence estimation, such as respondent-driven sampling (RDS) and VLTS, rely on social networks and referral chains among individuals to generate the sample. Social networks, however, rarely cross multiple industries, making these methods ideal for studying one industry at a time. Additionally, social networks are typically limited in their geographic spread, although the widespread use of social media has increased social networking beyond traditional geographic boundaries.

Data from the capture-recapture study can be used as one data source in a multiple-system estimation study. Capture-recapture is an umbrella of methodologies that include direct observation (such as is being proposed) to survey individuals at multiple time points. Another option is multiple-system estimation (MSE), which uses administrative data to identify individuals who appear on multiple lists to estimate the number who do not appear on any list, thereby estimating the population size. MSE is a cost-efficient and rigorous way to measure the prevalence of hidden populations and has been applied to study human trafficking victimization (Cruff et al., 2017; Sharifi et al., 2020).

During the expert panel meeting, however, researchers raised concerns about using MSE in the U.S. Namely, the U.S. does not have a national referral mechanism for human trafficking victims that can serve as a reliable data source, as in some countries. Given this, existing data from victim service providers would have to be checked by researchers for quality; consistency in definitions of victimization; and availability of key covariates like age, race, and gender, as well as data for matching across lists like full name and date of birth. In addition, law enforcement data on incidents and arrests do not reliably flag human trafficking cases, and law enforcement officers typically code the arrest charge based on what is clear when they arrive at the scene, leaving human trafficking undercounted in arrests.

Recognizing these limitations, BJS could make the dataset generated from a capture-recapture study available for researchers (restricted access) interested in conducting an MSE study, allowing for the field to generate additional prevalence estimates. If BJS or other government agencies such as the Office for Victims of Crime, Bureau of Justice Assistance, Office of Community Oriented Policing Services, or others implement quality control efforts and systematize the ways that victimization data are collected by service providers and law enforcement, the methodology could easily be converted to MSE at a much lower cost on a larger scale but may require initial investment.

2.1.3 Design Overview

Pilot Study

For the pilot, the Abt team recommends keeping the design simple, allowing it to function as a proof of concept for capture-recapture for identifying labor trafficking victims. Accordingly, the design and logistical considerations are selected to: 1) reduce heterogeneity in capture probabilities (and therefore reduce number of parameters needed to estimate after data is collected and the total required sample size), and 2) to keep costs as low as possible.

Location

The Abt team recommends selecting one city within a metropolitan statistical area that has a public transit ridership greater than or equal to five percent. The Washington DC metro area could be a good place for the pilot given BJS's physical presence in the area, high public transit usage and diversity in economic industries, racial/demographic population, and percentage of foreign-born residents. The below table provides a snapshot of the U.S. Census Bureau estimates for the DMV.

2022: ACS 5-Year Estimates for the Washington-Arlington-Alexandria, DC-VA-MD-WV Metro Area	Percentage	Estimate
Total Population		6,346,083
Workers who used public transportation *(S0801)*	8.8%	296,103
Industry (civilian employed population 16 years and older) *(DP03)*		3,382,923
Agriculture, forestry, fishing and hunting, and mining	0.3%	11,325
Construction	6.6%	223,750
Manufacturing	3.0%	100,240
Wholesale trade	1.1%	36.404
Retail trade	7.8%	262,788
Transportation, warehousing, utilities	4.1%	140,374
Information	2.5%	83,065
Finance and insurance, real estate, rental and leasing	6.1%	205,179
Professional, scientific, management, administrative, waste management services	22.1%	747,356
Educational services, health care, social assistance	19.5%	658,259
Arts, entertainment, recreation, accommodation, food services	7.8%	262,709
Public administration	13%	438,856
Other services	6.3%	212,618
Race (alone or in combination with one or more other races) *(DP05)*		
White	55.4%	3,514,312
Black or African American	27.6%	1,753,220
American Indian and Alaska Native	1.4%	89,094
Asian	12.6%	798,129
Native Hawaiian and Other Pacific Islander	0.2%	14,524
Other	11.9%	755,978
Hispanic or Latino (of any race) *(DP05)*	16.5%	1,050,132
Foreign born residents *(B05012)*	23.2%	1,469,883

Note: Variable names are included in parentheses.
Source: United States Census Bureau, 2022

Sample Size

Calculating the appropriate sample size for a capture-recapture study is multi-faceted. In our conversations with expert panel members and statisticians experienced in capture-recapture, we heard that formal sample size calculations should be interpreted with caution because they rely on assumptions and estimation. For example, to calculate estimated sample size, we have to input the total abundance in the population and the probability of being captured, two numbers for which we do not have reliable estimates. Instead, our key informants suggested using a sample size as large as budget constraints would allow. One panel member suggested, regardless of population size, 1,000-1,200 as a typical rule of thumb. In one of the studies we reviewed, a study aiming to estimate the prevalence of labor trafficking in domestic servitude among women in Tunis, Tunisia, the researchers sampled 923 respondents from 365 venues (Zhang, 2023). In another study, aiming to estimate the prevalence of forced labor among domestic workers in Casablanca, Morocco, the researchers sampled 1,067 respondents from 203 venues (Sahai, 2023).

Survey Sites

As a first step, we recommend identifying victim service providers and/or advocacy groups within the Washington DC metropolitan area who serve labor trafficking victims. We then recommend scheduling semi-structured interviews with providers to understand the transportation patterns of labor trafficking victims or potential victims, specifically how victims may be part of multiple types of industries travel to and from work. Questions would also be designed to understand the prevalence of victims using public transportation (rail and bus), if/how it varies by type of industry, the most common routes taken by victims, and days/times that we would be most likely to encounter victims at rail and/or bus stops. This information would be synthesized and used in the selection of survey sites.

Selection of survey sites is conducted using a similar methodology as time location sampling. This involves mapping out all venues (rail and bus stops) and hourly increments of the day, and days of the week, then randomly selecting within each category to identify venue-day-times for sampling. Since our sampling frame for venues are rail and bus systems, we must stratify randomization with rail and bus lines to ensure we do not over or under sample stops within a given route.

Getting permissions for a study in public transit venues. While steps to get access to conduct research in local transit venues may vary by city, we have outlined general steps that are likely applicable across a wide jurisdiction, with specific examples for Washington Metropolitan Area Transit Authority (WMATA).

1. Identify local regulations related to permits and exceptions to the permit requirement (WMATA).
2. Obtain IRB and/or OMB approval, or a letter from the IRB that you will obtain approval prior to beginning research.
3. Submit the permit application or exception application per guidance in local regulation (WMATA application).

Note that some cities may have fees associated with the permit application.

Requesting data from local transit authorities will require a different process. Many local departments of transportation regularly carry out research about their ridership and demand for public transit in general. WMATA's Office of Customer Research, for example, conducted a study as recently as 2022 on capacity and reliability. Most of WMATA's studies are publicly available and linked online. DC's open data portal also has a section on transportation data: https://projects.ddot.dc.gov/. Other jurisdictions may require filing a formal data-sharing request which may introduce delays.

We recommend randomly selecting times separately from venue-days, then surveying in all venues at the same time of day. We further recommend conducting recaptures for each venue on the same day of the week and time of day as the original capture rather than re-randomizing. This will limit variation in capture probabilities across venues and allow BJS to determine if recapture is possible with fewer variables affecting data collection. See **Exhibit 18** for an example data collection schedule for one enumeration team.

Exhibit 18. Example Schedule for One Field Enumeration Team During One Week

	Monday	Wednesday	Thursday	Saturday
Time	5am-9am	5am-9am	5am-9am	5am-9am
Stop 1	✓			
Stop 2				✓
Stop 3			✓	
Stop 4		✓		

There are a few key considerations when carrying out site selection. First, field interviewers (i.e., enumerators) typically require at least four-hour increments of scheduling. Therefore, the smallest unit of randomization within hours of the day are four-hour blocks. Second, the universe of days, times, and stops should be informed by qualitative research in the selected city. In particular, we want to select a combination of venue-day-times for which we can be reasonably confident that 80% or more of our target population (i.e., working adults that use public transit) pass through each day. This may require eliminating stops with less ridership, and eliminating days of the week, and hours of the day where we expect minimal ridership. We anticipate that the hours of 4am-8am or 6pm-10pm, for example, may be best suited for data collection. Third, when selecting venue-day-times and when generating a staffing plan, we must plan for the safety of data collectors, which may mean taking into consideration daylight hours given the time of year, not carrying physical incentives like gift cards, working in groups, and setting up near surveillance/cameras. Logistical considerations are discussed in more detail further in this recommendation.

Capture and Recapture Event Definition

We recommend using calendar weeks as capture events with one to two weeks between each event. Thus, if someone is captured on week one at stop one and again on week one at stop seven, that would count as one capture. If someone is captured on week one at stop one and week two at stop seven, that would count as one capture and one recapture (see **Exhibit 19**).

Exhibit 19. Example Capture-Recapture

Participant A captures	Monday 4/1 6am, stop 1	Friday 4/5 8am stop 7	**One capture, no recaptures**
Participant B captures	Monday 4/1 6am, stop 1	Monday 4/15 8am stop 7	**One capture, one recapture**

Time Frame

We recommend an eight-week fielding period between March-May. There are two population assumptions to use when generating prevalence estimates from a capture-recapture study: open population and closed population. Closed population estimates assume no migration in or out of the population, including no births and no deaths. Open population estimates allow for migration into and out of the population but require estimates of a variety of parameters in order to account for this flow. A short fielding period allows us to use closed population estimates for our prevalence estimate which introduces the least amount of bias. Ultimately, the bias introduced by guesswork for the parameters for open populations tends to drive higher bias than minor violations of the closed population assumption. Further, there are design characteristics that can help to reduce violations of closed population assumptions. First, keeping data collection as short as possible (ideally 8 weeks but no more than 12 weeks) can reduce the likelihood of massive influx or outflux of the adult working population in a city. There is also a balance, however, as we want to field for enough time for each individual in the population to have an equal likelihood of being captured for both the initial capture period and the recapture period. Second, selecting a time of the year when industries of interest are in season with more temperate climate to reduce the burden of spending time outside, and when there are no major events that may affect demand for workers in our population of interest (e.g., sporting events, political summits, natural disasters). **Exhibit 20** below describes overlap of seasonality for four key industries of interest.

Exhibit 20. Industry and Seasonality

	Hospitality	Construction	Janitorial Services	Commercial Landscaping
Spring	Busy	Busy	Little seasonality	Peak
Summer	Peak	Peak	Little seasonality	Busy
Fall	Busy	Busy	Little seasonality	Peak
Winter	Slowest	Slowest	Little seasonality	Slowest

Logistics

In-person data collection involves significant coordination of logistics, from where to set up, to how to approach participants. Decisions related to data collection procedures offer opportunities to strengthen our methodological approach. We present key considerations for in-person data collection below.

Screen out workers in low-risk industries. This methodology is best able to capture labor trafficking and exploitation among adults working in industries that are visible to the public. Industries that meet this criterion are known to have high risk for labor trafficking and exploitation and include hospitality, construction, janitorial services, and commercial landscaping. The survey itself will be designed and marketed as a general labor survey; however, to conserve costs associated with data collection, we recommend developing a brief screener that rules out industries with very low risks for trafficking.

People using public transit are often in a hurry. We do not anticipate that enumerators will be able to hold respondents for more than three minutes. Accordingly, we recommend that enumerators administer a brief screener that consists of name, date of birth, cell phone number, and industry in which they work. If they screen into the survey, the enumerator will send a link to their cell phone with the full survey. This should take about 10-15 seconds from when the link is pushed out. Enumerators will confirm in person that participants receive the link to ensure we have the right cell phone number. This will require that data collectors have devices with data cards, and ideally, also a Wi-Fi hotspot in case there is sparse cellular data connection available. There are multiple benefits to this strategy:

- Respects time constraints for commuters.
- Ensures we have a cell phone number so we can follow up if they don't complete the survey within a day to minimize non-response.

- Will provide sample information that allows us to match them as a known respondent for capture-recapture calculations.
- Allows us to use a longer survey instrument than would be feasible if it needed to be collected face-to-face.
- If potential victims are in the company of an employer, they have the option to not complete the survey or answer questions in the company of the employer.

Enumerators should scout the venue ahead of time to identify the best place to set up. Public transit stops often have multiple points of entry (for rail) and different entry points depending on the direction of the transit (e.g., bus stops on opposite sides of the road). We recommend that the researchers provide detailed guidance to enumerators on where to set up based on time of day/direction of the main flow of traffic, overall safety, shelter from weather, and ability to connect to cellular data and/or Wi-Fi. Wherever enumerators set up for the first data collection, they should be sure to set up in the same area for the subsequent recapture data collection.

Survey translations and computer-assisted audio for survey administration. It is critical to translate the survey into multiple languages. Abt's survey team, for example, can program questionnaires in 42 different languages, including the 12 most common languages used in metro areas, which allows the survey to reach more respondents. The use of Computer Assisted Personal Interviewing (CAPI) would allow users to hear questions read aloud on their mobile device in their language of choice. This can be helpful for participants who may have challenges with literacy and is shown to improve survey item comprehension.

Incentives and "trap happy" and "trap shy" respondents. Studies involving survey data collection often achieve better response rates when incentives are offered for participation (Gajic et al., 2012; Goritz, 2006). In this study, we must also consider that human behavior to either seek out or actively avoid participation can bias our findings. Thus, our plan for incentives is designed to attract "trap shy" respondents (those who would go out of their way to not engage with enumerators) and protect against "trap happy" respondents (those who change their behavior to try to get an incentive). Accordingly, we recommend three methods for incentives. First, we recommend that enumeration teams set up tables with bottled water or juice and a snack appropriate for the time of day (e.g., bagels/donuts/muffins, if morning, and granola bars/chips/ cookies, if afternoon). This will increase the likelihood that individuals pause near the enumerators to allow the enumerators to engage them in the screening process. We recommend attaching a small card to the snack container explaining that we are aiming to talk to people multiple times, and therefore the respondent should not avoid the team if they see them again. Second, we recommend providing a small (around $10), virtual gift card for individuals that complete the survey. We do not recommend physical incentives such as in person physical gift cards as this may pose a security threat to the data collection team. Finally, we recommend a slightly higher incentive for participation in the recapture screener (around $15).

Make recapture as simple as possible. Getting in-person participation in a survey one time is challenging. Attempting to pause participants at a subsequent time will be even more challenging. We recommend implementing procedures to reduce the burden of recapture for participants. This includes:

1. Handing out laminated QR codes that participants can present to enumerators if they see them again. The enumerator will scan the code and the participant can carry on without reentering data.
2. Sending a text link to participants during the data collection period for the recapture for that venue. The survey team can then use the location of the cell phone to identify if they were in the recapture area when the link was sent.

Cost

Capture-recapture through direct observation is a cost-intensive study. One of our expert panel members estimated that it is at least 20% more expensive than conducting a probability proportional to size (PPS) study. Our estimate for the pilot study is between $1-1.5 million. This includes field work needed to map out the sampling frame to identify suitable venues, days, and times to sample. It also includes costs associated with hiring an experienced field data collection team. Enumerators for this effort need to have previous experience with data collection and receive more training than a standard study related to safety, ethical considerations, and other procedures

for interacting with potential participants. Field data collection and management also incurs substantial costs, as the sample needs to be monitored closely during short fielding periods. Finally, capture-recapture studies require incentives for participation in the initial survey and for checking in for the recapture.

Limitations

Using this capture-recapture methodological approach has several limitations but each are readily addressed in this proposed collection method. First, the proposed approach would only be generalizable to public transit users in urban areas. This would also exclude rural community members. The population of interest, however, does not need to be generalizable to a broader population. Our aim is to understand labor trafficking specifically, and public transportation creates a grounded approach to generate relevant estimations. Further, the proposed study would limit to only capturing the prevalence of labor trafficking. This is an intentional decision based on the research objectives and knowledge that sex and labor trafficking are distinctly separate in both legal and research contexts. This intentional separation is also widely accepted in research and confirmed in our environmental scan.

Specific measurements of labor trafficking are also considered a limitation with prevalence research. For example, our expert panelists noted that the measures within the PRIF survey may not fully capture prevalence. This limitation, however, can be easily remediated with supplemental data such as offering follow-up cognitive interviews for respondents that opt-in the survey allowing for a comprehensive quantitative and qualitative approach in prevalence estimation efforts.

Additionally, the survey would only include adult participants and thus limit generalizability to only adults. Adding questions regarding children that participants work alongside, however, would help gather preliminary information about children in general terms and provide insight into respondents' experiences (if any) working alongside children. Despite these limitations, our recommendation to use capture-recapture as a prevalence estimation method serves as a clear starting point to investigate the prevalence of labor trafficking while clearly addressing and acknowledging methodological limitations.

2.2 Recommendation 2: Vincent Link-Trace Sampling (VLTS) Survey

2.2.1 Overview

VLTS uses a network- based recruitment method that relies on chain or peer referrals among individuals who make up hard-to-reach populations. This method relies on an assumption that individuals within the target population belong to a connected social network, making it most appropriate for studying human trafficking within one industry and geographic area; preferably among a population that is somewhat stable (i.e., no significant migration in and out of the network). The goal of the proposed VLTS approach is to assess the prevalence of labor trafficking within two social networked industries—hospitality and restaurant/food services—in one geographic area in the U.S.

2.2.2 Strengths

A key strength of VLTS is that it can be used for hidden or hard-to-reach populations, and it can yield a final sample that approximates a random sample, allowing for inferences to be made about prevalence. It also yields a more efficient population estimation than other network-based methods, like RDS, by linking network characteristics across all respondents in the final sample, and not just within each wave. Additionally, this approach can be used to capture victims in larger geographic areas relative to RDS, and due to fewer waves or recruitment (relative to RDS) there is less risk that the final sample will not reach equilibrium.

2.2.3 Population

We recommend using VLTS to estimate the prevalence of labor trafficking among individuals working in restaurant/food services and hospitality in a geographic area located in the U.S. Findings from our environmental scan revealed four labor trafficking studies used link-tracing to estimate victim prevalence. Labor trafficking victims working in these specific industries are well suited for VLTS since they are part of tight social networks and are generally not captured in research that uses more conventional sampling strategies.

2.2.4 Time Frame

We recommend an eight-week data collection period, which is consistent with studies included in the environmental scan that used VLTS. We also propose fielding the survey outside of the "off season" in the selected locality. For many areas, this means avoiding winter months.

2.2.5 Sampling Process

VLTS typically begins with analysis of administrative data to identify geographies from which to select seeds. The National Human Trafficking Hotline (NHTH) provides the largest known data set in sex and labor trafficking in the U.S. Additionally, the NHTH collects information about the type of labor trafficking, including restaurant/food service and hospitality. In 2021—the most recent year for which published data is available—the NHTH reported 1,066 situations of labor trafficking. Restaurant/food service ranked fourth (n=60) out of venues for labor trafficking and hospitality sixth (n=36). The NHTH further breaks down types of labor trafficking by state. We recommend selecting a state with a large number of labor trafficking cases that include restaurant/food service and hospitality venues. Additionally, we recommend looking at states that include cities that typically attract tourism, such as New York and California, as the tourism industry includes other factors that make people vulnerable to human trafficking. Other factors that increase the risk for human trafficking include proximity to international borders, large immigrant populations, being home to numerous ports and airports, and having large economies that include industries that attract forced labor.

2.2.6 Sample Size

The five studies included in our environmental scan that used link-tracing ranged from between 93 to 147 initial seeds with two waves of recruitment. Total sample sizes for the five studies ranged from between 317 (111 initial seeds) to 1,009 (98 initial seeds). We propose using a similar number of initial seeds (between 90-150) with two waves. We expect to reach equilibrium with a sample of between 800 and 1350 respondents (the point at which successive samples/waves no longer mirror initial samples).

2.2.7 Recruitment of Participants

Once a state has been identified, we recommend identifying victim service providers or advocacy groups that can help identify and recruit seeds within the state. For example, in California, there are a number

of organizations that provide services and support to trafficking victims, such as the Coalition to Abolish Slavery & Trafficking, which is based in Los Angeles. New York is home to Safe Horizon, an anti-trafficking program dedicated to supporting survivors of labor and sex trafficking in New York City. Once providers are identified, we recommend conducting semi-structured interviews with them to understand the prevalence of labor trafficking in restaurants and hospitality venues, characteristics of trafficking victims working in those industries, if and how victims are recruited, and any information related to victim networks specific to restaurant and hospitality venues. This information will be used to identify and recruit initial seeds.

We recommend starting with a large number of seeds (between 90-150) using two waves of recruitment, similar to the studies included in our environmental scan. Field staff would consult with service providers and advocacy groups to identify labor trafficking victims they have served who worked in restaurants and hospitality venues. Each seed making up Wave 1 would complete an online survey and receive three coupons that they can use to recruit peers that would make up Wave 2. In line with other network-based surveys in the U.S., we propose offering incentives to participants for participation and for recruiting their peers to participate in the survey. We recommend consulting with victim service providers and advocacy groups regarding the maximum total compensation that would be considered motivating without being coercive.

2.2.8 Cost

VLTS is typically less expensive than capture-recapture or PPS. We estimate a cost of $750,000-$1 million for a VLTS study. However, there is a moderate cost due to the dual incentive system (incentives for participation and for referring other participants) and the cost for an experienced field team to complete the formative assessment and data collection component. There is a potential for cost savings by doing a web or phone survey versus in- person data collection.

2.2.9 Limitations

As mentioned above, a key limitation of VLTS is that it is limited to one geography and one industry or a set of industries that are socially networked together. Another limitation of VLTS studies is that since it relies on social networking, the most isolated, and thus perhaps the most vulnerable populations, may not be reached. Additionally, this method does not work with highly mobile populations since they may have fewer social connections. This approach also requires an experienced field team to build trust and gain buy-in for participants to recruit from their network, which can be costly. Finally, respondent behavior, including how they choose to form their social connections, identify those eligible for coupons/referrals, and count their network size, can influence the validity of estimates.

Based on the challenges mentioned above, we present considerations for implementation below:

Select industry-geography combinations where victims are tightly networked: VLTS relies on an assumption that individuals within the target population belong to a connected social network. Therefore, it is critical that the trafficking industry and the geographic location selected for study include populations that are tightly networked. This requires a considerable amount of time up front to identify and examine administrative data and/or previous research to first identify the appropriate trafficking industry and initial seeds for study. A more costly alternative is combining VLTS with other methods, such as capture-recapture or probability proportionate to size sampling, to yield a sample of seeds for the first wave. Since VLTS is also limited to one geography, it will be important to identify a geographic location that is conducive to the type of industry being examined. The location should also include other factors that make them vulnerable to human trafficking, such as proximity to international borders, large immigrant populations, being home to numerous ports and airports, and having large economies that include industries that attract forced labor.

Hire experienced field enumerators and provide additional training specific to working with victims of human trafficking: Once the industry and geographic location are selected, the next step is to identify appropriate victim service providers within the location from which to recruit victims. This requires an experienced field team who can identify and secure participation from victim service providers for the purpose of recruiting participants. Field staff also need to build trust and gain buy-in from participants to complete a survey and to recruit from their network, while prioritizing participant safety. All staff should be trained prior to data collection on ethical protocols designed to protect participants and should understand the risks associated with the study. Field staff should be trained to conduct cognitive interviews

to understand participants' comprehension of data collection instructions and their comprehension of and willingness to consent. It is also important that staff protect participant confidentiality and adhere to all other IRB procedures, including compliance with all mandated reporting laws and protocols designed to protect respondents from potentially being exposed to their traffickers for participating. Additionally, since participants are responding to questions that are related to victimization, field staff should be prepared for situations where a participant is retraumatized or may need additional support, such as the NHTH number or mental health resources.

Closely monitor participant recruitment and survey completion: Since VLTS usually involves no more than three waves and a relatively large number of seeds in the first wave, the research team will need to closely monitor the number of completed surveys from seeds in the initial wave to identify any problems or red flags and to provide incentives. This will be cost-intensive due to the increased staff hours needed to monitor the number of completed surveys, send reminders to participants who have not completed the survey, and keep track of participants who are eligible for incentives. It will be important for field staff to closely monitor how participants choose to form social connections, including understanding the characteristics of individuals who are recruited by participants to ensure they are eligible for coupons/referrals. For example, seeds are most likely to give coupons to people they know the best, which may not be the people who best meet the definition for inclusion.

2.3 Recommendation 3: Random Sampling Survey

2.3.1 Overview

Random sampling, or probability sampling, is considered a methodological gold standard to create generalizable estimates. This recommendation provides an option to obtain a national estimate at a lower cost relative to the capture-recapture recommendation. Random sampling assumes that every member of a population has an equal chance of being selected, minimizing bias and allowing research to produce more generalizable results. However, measuring human trafficking using random sampling may increase the chances of undercounting or underreporting, given that victims are typically not captured in traditional sampling frames. Despite these risks, random sampling approaches allow for minimizing bias by removing the preselection of where individuals or groups are more likely to be and ensures validity of statistics gathered by determining margins of error and confidence levels of their estimates.

Using random sampling as a prevalence estimation approach, we propose partnering with current random sampling platforms to conduct web-based surveys with human trafficking estimation measures.

2.3.2 Sampling Frame and Population

Random sampling allows individuals within a sampling frame to have an equal likelihood of being selected to participate in the survey. For either of our recommendations using a random sampling approach, the sampling frame includes addresses based on U.S. Postal Service (USPS) records. One advantage of address-based sampling is that households without phone or computer access may be included. A downside to this is that many people who we may want to capture could be excluded, including: individuals with unstable housing, individuals living in non-residential buildings, individuals housed with their traffickers, those renting a room (and therefore not listed as a separate address), are just a few examples.

Companies that specialize in random sample surveys and polls recruit nationally representative, address-based samples to participate in a range of surveys. The companies provide participants with internet access and pay them to complete any surveys they opt into. Each panel member is paid based on the number of surveys they qualify for and complete. Data collected from these panels are weighted to account for non-responses and the demographic makeup of the target population, in this case the entire U.S. Based on prior experience fielding national surveys, conversations with expert panelists, and a power analysis with a rough estimate for design effect, we anticipate a minimum required sample size of 2,730 individuals. Total sample size will vary based on a specific companies' weighting structure and design effect estimates for their panel.

IPSOS, for example, uses an address-based sampling spanning across the U.S. Addresses are gathered through the USPS, covering almost 100% of the U.S. population. The addresses are maintained regardless of their phone status, which is uniquely distinct compared to other random household address sampling. Surveys are then repeated in waves furthering randomization.

For IPSOS, respondents are already part of panels where they can select to opt in to taking surveys and are compensated according to the number of surveys completed. Specific survey questions can either be added to existing survey dissemination efforts, or entire new surveys can be collected and administered. Survey questions added on must be approved and accepted by the platform selected. These options also exist using Gallup. Each of these platforms have been used for the collection and dissemination of prevalence estimation of human trafficking, however neither have been used in the U.S. context alone.

2.3.3 Time Frame

The time frame for random sampling is largely dictated by the sample size, and because there are not population assumptions, there is no set limit of time the survey needs to be live. Typically, from start to finish, a survey or polling company can collect the data in under one year.

2.3.4 Cost

The cost of conducting a random sampling household survey can vary depending on several factors, such as the desired sample size and the length and complexity of the survey. It could be cost-effective to add a small number of questions to an existing online panel survey. Depending on their complexity, one to five questions could be added to a national panel for an estimated one hundred thousand dollars. While members of our expert panel recommended exploring the feasibility of adding questions to existing federal data collection efforts, doing so would result in a cost closer to the other two options.

2.3.5 Strengths

The underlying assumptions of random sampling, if met, are straightforward and generalizable to the population from which the sample is drawn from. Further, random sampling has been extensively studied and successfully employed in human trafficking estimation for both sex and labor trafficking. Sampling frames can be easily defined due to accessibility of data such as census information. Finally, this methodology offers a lower-cost option compared to the other proposed methodologies, with the possibility to add to existing random sampling approaches or using third parties to assist in data collection. These collection methods also allow for the inclusion of minors (13+), which are another underrepresented group of individuals in human trafficking research.

2.3.6 Limitations

The ability to generalize from random sampling is based on the assumption that everyone in the target population has an equal chance of being selected; however, victims of human trafficking may be hesitant to engage, potentially leading to an undercount of actual victims. Further, using address-based approaches may not guarantee that victims will be the one(s) completing the survey. With this approach, it is hard to know who the respondent is. All that is known is that they are a resident or occupant of a particular randomly selected address. Compounding this issue, if narrow search parameters are used in an effort to represent victims of human trafficking more accurately, or focus on one geographic space, victims may be further underrepresented. In the context of labor trafficking specifically, random sampling could only be narrowed by geography and not by industry. Random sampling would not allow for specific industries to be targeted unless they are known industries with a particular area.

2.3.7 Formative Assessment

Based on the recommendation received from the expert panelists, the Abt team proposes disseminating a household survey using RDS as a proposed option despite its limitations. An RDS survey, regardless of its flaws, would still serve as a practical starting point for addressing a portion of the known research gap, rather than continued discussion of the gap without an initial attempt in prevalence estimation. Further, a household survey would not only offer a lower cost solution but could also build a foundation to gather valuable data and insights that can be further refined and built upon in future research efforts.

Bibliography

Abusaleh, K., Islam, M. R., Ali, M. M., Khan, M. A., Shahinuzzaman, M., & Haque, M. I. (2022). Prevalence of economic exploitations and their determinants among child labourers in Dhaka City, Bangladesh: a mixed-method study. Child Indicators Research, 1-20.

Anderson, E., & Kendall, C. (2023). Estimating the Prevalence of Sex Trafficking in Recife, Brazil. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Anderson, V. R., Kulig, T. C., & Sullivan, C. J. (2019). Estimating the prevalence of human trafficking in Ohio, 2014–2016. American Journal of Public Health, 109(10), 1396-1399.

Andrade, C. (2021). The Inconvenient Truth About Convenience and Purposive Samples. Indian Journal of Psychological Medicine, 1(43), 86–88. https://doi.org/10.1177/0253717620977000

Angrist, J. D., & Pischke, J.-S. (2009). Mostly Harmless Econometrics. Princeton University Press.

Babiarz, K. (2023). Estimating the Prevalence of Human Trafficking in Brazilian Agriculture. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Bales, K., Murphy, L. T., & Silverman, B. W. (2020). How many trafficked people are there in Greater New Orleans? Lessons in measurement. Journal of Human Trafficking, 6(4), 375-387.

Banks, D., & Mokel, E. (2023). Multiple Systems Estimation and Human Trafficking. CHANCE, 36(4), 12-14.

Banks, D. L., & Hooten, M. B. (2021). Statistical challenges in agent-based modeling. The American Statistician, 75(3), 235-242.

Barrick, K., & Pfeffer, R. (2021). Advances in Measurement: A Scoping Review of Prior Human Trafficking Prevalence Studies and Recommendations for Future Research. Journal of Human Trafficking. https://doi.org/10.1080/23322705.2021.1984721

Bird, S. M., & King, R. (2018). Multiple systems estimation (or capture-recapture estimation) to inform public policy. Annual Review of Statistics and Its Application, 5, 95-118.

Branscum, C., & Richards, T. N. (2023). The wrong rite of passage: comparing sex trafficking to other types of child maltreatment using the National Child Abuse and Neglect Data System (NCANDS). Victims & Offenders, 18(3), 399-422.

Chan, L., Silverman, B. W., & Vincent, K. (2021). Multiple systems estimation for sparse capture data: Inferential challenges when there are nonoverlapping lists. Journal of the American Statistical Association, 116(535), 1297-1306.

Chohaney, M. L. (2016). Minor and adult domestic sex trafficking risk factors in Ohio. Journal of the Society for Social Work and Research, 7(1), 117-141.

Cruyff, M., Van Dijk, J., & van der Heijden, P. G. (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, 30(3), 41-49.

Cruyff, M., Overstall, A., Papathomas, M., & McCrea, R. (2021). Multiple system estimation of victims of human trafficking: Model assessment and selection. Crime & Delinquency, 67(13-14), 2237-2253.

Dank, M. (2023a). Estimating the Prevalence of Forced Labor Among Domestic Workers in Tanzania. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Dank, M. (2023b). Estimating the Prevalence of Forced Labor in the Fishing Industry in Costa Rica. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Dank, M., Farrell, A., Zhang, S., Hughes, A., Abeyta, S., Fanarraga, I., ... & Solis, V. O. (2021). An Exploratory Study of Labor Trafficking Among U.S. Citizen Victims. CrimRxiv.

DeliverFund. (2024). Facts About Human Trafficking in the U.S. DeliverFund.

Durgana, D., & van Dijk, J. (2021). Measuring the hidden numbers of human trafficking through multiple systems estimation: Lessons learned and challenges outstanding. Crime & Delinquency, 67(13-14), 2188-2212.

Farrell, A., Owens, C., & McDevitt, J. (2014). New laws but few cases: Understanding the challenges to the investigation and prosecution of human trafficking cases. Crime, law and social change, 61, 139-168.

Gajic, A., Cameron, D., & Hurley, J. (2012). The cost-effectiveness of cash versus lottery incentives for a web-based survey, stated-preference community survey. The European Journal of Health Economics, 13, 789–799. https://doi.org/10.1007/s10198-011-0332-0

Gerassi, L., Edmond, T., & Nichols, A. (2017). Design strategies from sexual exploitation and sex work studies among women and girls: Methodological considerations in a hidden and vulnerable population. Action Res (Lond), 2(15), 161–176. https://doi.org/10.1177/1476750316630387

Giommoni, L., & Ikwu, R. (2021). Identifying human trafficking indicators in the UK online sex market. Trends in Organized Crime, 1-24.

Goritz, A. (2006). Incentives in Web Studies: Methodological Issues and a Review. International Journal of Internet Science, 1(1), 58–70.

Grosso, A., Busch, S., Mothopeng, T., Sweitzer, S., Nkonyana, J., Mpooa, N., ... & Baral, S. (2018). HIV risks and needs related to the Sustainable Development Goals among female sex workers who were commercially sexually exploited as children in Lesotho. Journal of the International AIDS Society, 21, e25042.

Heckathorn, D. D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. Social Problems, 44(2), 174–199. https://doi.org/10.2307/3096941

International Labour Organization, Walk Free, & International Organization for Migration. (2022). Global Estimates of Modern Slavery: Forced Labour and Forced Marriage.

Johansen R (2019) 'UNODC's use of Multiple Systems Estimation (MSE) to assist countries in measuring human trafficking and reporting on SDG indicator 16.2.2' bit.ly/Johansen-UNODC-MSE-2019

Khatab, K., Raheem, M. A., Sartorius, B., & Ismail, M. (2019). Prevalence and risk factors for child labour and violence against children in Egypt using Bayesian geospatial modelling with multiple imputation. PLoS one, 14(5), e0212715.

King R, Brooks SP. On the Bayesian analysis of population size. Biometrika. 2001a;88:317–336.

King R, Bird SM, Brooks SP, Hutchinson SJ, Hay G. Prior information in behavioural capture-recapture methods: demographic influences on drug injectors' propensity to be listed in data sources and their drugs-related mortality. Am J of Ep. 2005;162:1–10.

King R, Bird SM, Overstall A, Hay G, Hutchinson SJ. Estimating prevalence of injecting drug users and associated heroin-related death-rates in England using regional data and incorporating prior information. J R Stat Soc A. 2014;177:1–28.

Li, R., Tobey, M., Mayorga, M. E., Caltagirone, S., & Özaltın, O. Y. (2023). Detecting human trafficking: Automated classification of online customer reviews of massage businesses. Manufacturing & Service Operations Management, 25(3), 1051-1065.

Mak, J., Abramsky, T., Sijapati, B., Kiss, L., & Zimmerman, C. (2017). What is the prevalence of and associations with forced labour experiences among male migrants from Dolakha, Nepal? Findings from a cross-sectional study of returnee migrants. BMJ open, 7(8), e015835.

O'Brien, J. E., Li, W., Givens, A., & Leibowitz, G. S. (2017). Domestic minor sex trafficking among adjudicated male youth: Prevalence and links to treatment. Children and Youth Services Review, 82, 392-399.

Office for Victims of Crime. (2021). Collecting Data to Better Understand Human Trafficking. Office for Victims of Crime, Office of Justice Programs. https://htcbc.ovc.ojp.gov/sites/g/files/xyckuh311/files/media/document/Collecting_Human_Trafficking_Data_508c.pdf

Prevalence Reduction Innovation Forum. (2020). Human Trafficking Statistical Definitions.

Put Trafficking Victims First Act of 2024, S. 4007, 118th Cong. (2024). https://www.congress.gov/bill/118th-congress/senate-bill/4007/text

Robinson, C. (2023). Estimating the Prevalence of Forced Labor in the Brick Kiln Industry in Pakistan. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Saewyc, E. M., Shankar, S., Pearce, L. A., & Smith, A. (2021). Challenging the stereotypes: unexpected features of sexual exploitation among homeless and street-involved boys in Western Canada. International journal of environmental research and public health, 18(11), 5898.

Sahai, R. (2023). Estimating the Prevalence of Forced Labor Among Domestic Workers in Casablanca, Morocco. Results from the Field: Prevalence Reduction Innovation Forum. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Schroeder, E., Edgemon, T. G., Aletraris, L., Kagotho, N., Clay-Warner, J., & Okech, D. (2022). A Review of Prevalence Estimation Methods for Human Trafficking Populations. Public Health Reports, 137(1), 465–525. https://doi.org/10.1177/00333549211044010

Shahen, Md. A. (2021). The Prevalence of Child Vulnerability and Its Nature: An Analytical Study on Bangladesh Perspective. International Journal of Qualitative Research, 1(1), 10–23. https://doi.org/10.47540/ijqr.v1i1.294

Sharifi Far, S., King, R., Bird, S., Overstall, A., Worthington, H., & Jewell, N. (2021). Multiple systems estimation for modern slavery: robustness of list omission and combination. Crime & Delinquency, 67(13-14), 2213-2236.

Stöckl, H., Fabbri, C., Cook, H., Galez-Davis, C., Grant, N., Lo, Y., ... & Zimmerman, C. (2021). Human trafficking and violence: Findings from the largest global dataset of trafficking survivors. Journal of migration and health, 4, 100073.

Stransky, M., & Finkelhor, D. (2008). How many juveniles are involved in prostitution in the U.S. Durham, NH: Crimes Against Children Research Center, University of New Hampshire.

Trafficking Victims Protection Act of 2000 (Pub. L. 106–386, div. A, §102, Oct. 28, 2000, 114 Stat. 1466.)

United States Census Bureau. (2022). 2022: ACS 5-Year Estimates Data Profiles. https://data.census.gov/table/ACSDP5Y2022.DP03?g=040XX00US11&tid=ACSDP5Y2022.DP03

Urada, L. A., Rusakova, M., Odinokova, V., Tsuyuki, K., Raj, A., & Silverman, J. G. (2019). Sexual exploitation as a minor, violence, and HIV/STI risk among women trading sex in St. Petersburg and Orenburg, Russia. International journal of environmental research and public health, 16(22), 4343.

U.S. Bureau of Labor Statistics. (2024). Employed Persons in Washington-Arlington-Alexandria, DC-VA-MD-WV (MSA) [LAUMT114790000000005] [dataset]. retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/LAUMT114790000000005

United Nations Office on Drugs and Crime. (2022). Monitoring Human Trafficking Prevalence Through Multiple Systems Estimation. United Nations Office on Drugs and Crime. https://www.unodc.org/documents/data-and-analysis/tip/2022/MSE_TIP_UNODC_ENG.pdf

van der Heijden, P. G., de Vries, I., Böhning, D., & Cruyff, M. (2015). Researching hidden populations: approaches to and methodologies for generating data on trafficking in persons.

Vincent, K., Zhang, S. X., & Dank, M. (2021). Searching for sex trafficking victims: Using a novel link-tracing method among commercial sex workers in Muzaffarpur, India. Crime & Delinquency, 67(13-14), 2254-2277.

World Health Organization. (2013). Introduction to HIV/AIDS and sexually transmitted infection surveillance: Module 4: Introduction to respondent-driven sampling (No. WHO-EM/STD/134/E).

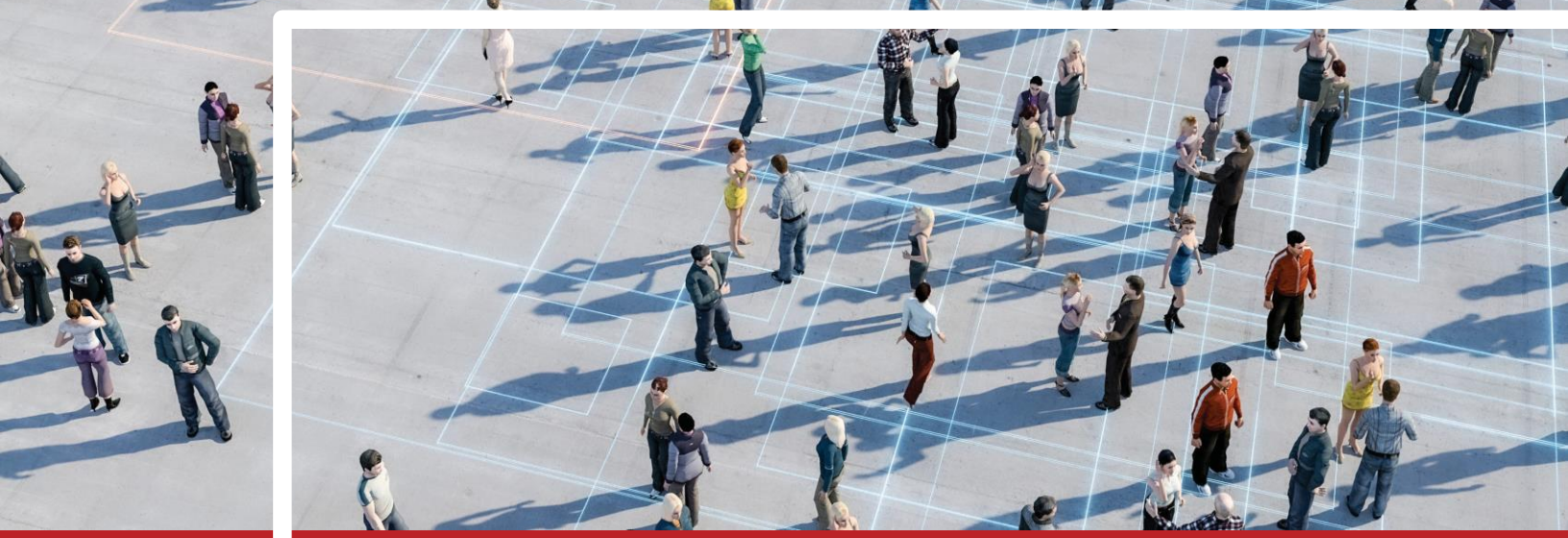
White, J. (2020). Estimating the Prevalence of Human Trafficking in the United States: Considerations and Complexities. Committee on National Statistics and Committee on Population, Division of Behavioral and Social Sciences and Education. The National Academies of Sciences, Engineering, and Medicine, Washington, DC. https://doi.org/10.17226/25614

Wright, E. R., LaBoy, A., Tsukerman, K., Forge, N., Ruel, E., Shelby, R., Higbee, M., Webb, Z., Turner-Harper, M., Darkwa, A., & Wallace, C. (2021). The Prevalence and Correlates of Labor and Sex Trafficking in a Community Sample of Youth Experiencing Homelessness in Metro-Atlanta. Soc. Sci., 10(2), 32. https://doi.org/10.3390/socsci10020032

Youle, J., & Long, A. (2020). The same ruler for everyone: Improving trafficking estimates.

Young, L.J., Young, J.H. (1998). Capture—Recapture: Open Populations. In: Statistical Ecology. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-2829-3_10

Zhang, S. X. (2012). Measuring labor trafficking: a research note. Crime, Law and Social Change, 58(4), 469-482.
Zhang, S. (2023). Domestic Servitude in Tunis, Tunisia: Findings from Two Sampling Methods and Policy Implications. Results from the Field: Prevalence Reduction Innovation Forum Learning Series. https://www.youtube.com/watch?v=PyjTBHnxHew&list=PLuOvNvQoTvTCOnG6Ui6o_GNew59qozNux&index=3

Zhang, S. X., Dank, M., Vincent, K., Narayanan, P., Bharadwaj, S., & Balasubramaniam, S. M. (2019). Victims without a Voice: Measuring Worst Forms of Child Labor in the Indian State of Bihar. Victims & Offenders, 14(7), 832–858. https://doi.org/10.1080/15564886.2019.1660289

Zhang, S. X., & Larsen, J. J. (2021). Estimating the size of the human trafficking problem: MSE and other strategies. Crime & Delinquency, 67(13-14), 2169-2187.