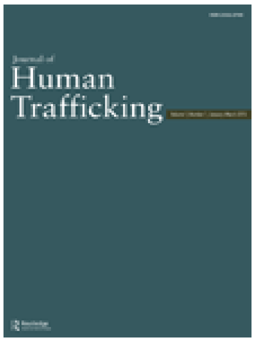

Journal of Human Trafficking

ISSN: 2332-2705 (Print) 2332-2713 (Online) Journal homepage: https://www.tandfonline.com/loi/uhmt20

How many trafficked people are there in Greater New Orleans? lessons in measurement

Kevin Bales, Laura T. Murphy & Bernard W. Silverman

To cite this article: Kevin Bales, Laura T. Murphy & Bernard W. Silverman (2019): How many trafficked people are there in Greater New Orleans? lessons in measurement, Journal of Human Trafficking, DOI: 10.1080/23322705.2019.1634936

To link to this article: https://doi.org/10.1080/23322705.2019.1634936

Published online: 17 Aug 2019.

JOURNAL OF HUMAN TRAFFICKING
https://doi.org/10.1080/23322705.2019.1634936

OPEN ACCESS Check for updates

How many trafficked people are there in Greater New Orleans? lessons in measurement

Kevin Bales [a], Laura T. Murphy [b], and Bernard W. Silverman [a]

[a]Rights Lab, University of Nottingham, Nottingham, UK; [b]Helena Kennedy Centre for International Justice, Sheffield Hallam University, Sheffield, UK

ABSTRACT
In an effort to develop a model for estimating prevalence in a city or region of the United States, this study employed Multiple Systems Estimation, a statistical approach that uses data on known cases collected from individual agencies to estimate the number not known, with the ultimate aim of estimating the prevalence of trafficking in a region. Utilizing de-identified data provided by local non-profits and law enforcement agencies, the researchers estimated the prevalence of trafficking in the New Orleans-Metairie metropolitan statistical area. This represents one of the first attempts to use Multiple Systems Estimation to quantify human trafficking in a United States context. The article provides an account of the impediments to and limitations of conducting such an estimate, given the definitional variance and political dynamics that are endemic to anti-trafficking efforts in the United States. The authors provide recommendations for data collection and prevalence analysis that could be applied in other cities or regions of the United States as well as in other similarly-resourced environments.

KEYWORDS
Human trafficking; sex trafficking; labor trafficking; measurement; multiple systems estimation; New Orleans

Researchers in the field of human trafficking have faced significant challenges in estimating the prevalence of victimization since the emergence of the field in the 1990s, and much recent research has focused on critiquing early methods and addressing their shortcomings (Brunovskis & Surtees, 2010; Goździak, 2015; Merry, 2016; Nawyn, Kavakli Birdal, & Glogower, 2013; O'Brien, 2010; Tyldum & Brunovskis, 2005). Human trafficking is a hidden crime, which for several reasons is not amenable to the most widely applied approaches to measuring crime and victimization, especially in wealthy nations such as the U.S. where trafficking is less prevalent and more hidden. Recent studies in U.S. cities or regions have tended to target specific populations with perceived vulnerabilities to trafficking, including homeless youth (Chisolm-Straker, 2017; Chisolm-Straker, Sze, Einbon, White & Stoklosa, 2018; Middleton, Gattis, Frey, & Roe-Sepowitz, 2018; Murphy, 2017), youth engaged in trading sex (Curtis, Terry, Dank, Dombrowski, & Khan, 2008; Dank et al., 2015), youth involved in the child welfare system (Fong & Berger Cardoso, 2010; Gibbs, Henninger, Tueller, & Kluckman, 2018), undocumented laborers (Zhang, Spiller, Finch, & Qin, 2014), sexual minorities (Marshall, Shannon, Kerr, Zhang, & Wood, 2010), or minor girls (Lutnick, 2016). Researchers typically do not venture estimates of prevalence from these studies, and are careful to note that any estimates that might be drawn from those studies are clearly limited to those particular populations and cannot be extrapolated to determine prevalence estimates for entire communities or the nation. Nonetheless, in the absence of more accurate estimates, journalists and scholars often inappropriately employ these limited findings as prevalence estimates.

CONTACT Laura T. Murphy ltmurphy27@gmail.com Sheffield Hallam University
Kevin Bales is Professor of Contemporary Slavery and Research Director of the Rights Lab at the University of Nottingham, UK.
Laura T. Murphy is Professor of Human Rights and Contemporary Slavery at the Helena Kennedy Centre for International Justice, Sheffield Hallam University, UK.
Bernard W. Silverman is Professor of Modern Slavery Statistics at the Rights Lab, University of Nottingham. He was formerly Chief Scientific Adviser to the UK Home Office.

Without knowing the prevalence of human trafficking, it is difficult to frame appropriate policies or effectively allocate resources. This study therefore sets out to address these research challenges by employing Multiple Systems Estimation (MSE), a statistical model that builds from lists of known cases to estimate the number not known, with the ultimate goal of estimating the total number of trafficked persons in a particular geographical location. The research detailed in this article provides an estimate, derived from MSE analysis that uses local data to determine the prevalence of trafficking in an American metropolitan area, in this case the New Orleans–Metairie Metropolitan Statistical Area (*the New Orleans MSA*), often referred to as "Greater New Orleans." Additionally, this study reports on the factors that aided and impeded an accurate and reliable MSE estimation of trafficking in the New Orleans MSA, such as definitional variance and the likelihood of political and cultural dynamics affecting the reporting of trafficking crime and victimization by different types of organizations. Based on this example, we provide some recommendations for data collection and analysis relevant to cities and regions in the United States as well as in other similarly resourced countries.

The Challenge

Measuring human trafficking in relatively wealthy, so-called "developed" nations is particularly challenging. A lack of awareness that trafficking affects people in the global north makes it difficult for people to recognize forced labor in their midst. Furthermore, since trafficking is relatively rare in these countries, people randomly selected for a crime prevalence survey are unlikely to have experienced (or know someone who has experienced) trafficking. Lastly, in countries where law enforcement is robust, trafficking is even more likely to be hidden in the shadows.

In the United States, most crimes are estimated through two specific mechanisms: the National Crime Victimization Survey (NCVS) and the FBI's Uniform Crime Reporting Program (UCR). The NCVS is a nationally-representative sample survey of more than 135,000 households. Conducted by the Bureau of Justice Statistics, it is used to determine the "frequency, character, and consequences of criminal victimization" for several specific crimes – rape, sexual assault, robbery, aggravated and simple assault, household burglary, theft, and motor vehicle theft. Human trafficking victimization is not measured, though cases of sex trafficking may be identified as sexual assault or rape under these limiting circumstances (Bureau of Justice Statistics, 2017). While human trafficking victimization is not being tracked by the NCVS, the UCR is collecting data on the human trafficking crimes committed. The most comprehensive aspect of the FBI's UCR is the National Incident-Based Reporting System (NIBRS), which collects data from over 18,000 law enforcement agencies on each "crime incident" (and all offenses related to that incident) that participating agencies have investigated. In 2008, the U.S. Congress mandated that human trafficking be integrated into the UCR, and beginning in January 2013, two new crimes were introduced to their data collection efforts: "human trafficking/commercial sex act" and "human trafficking/involuntary servitude" (FBI, 2014). Since that time, law enforcement agencies have been able to report their investigations of both labor and sex trafficking to this collaborative data collection effort, though because it is still based on voluntary submission of data, those data sets remain incomplete (Newton, Mulcahy, & Martin, 2008).

With the advent of the collection of human trafficking data through the UCR, it might seem that a robust environment is being built for determining the prevalence of human trafficking in the United States. However, the law enforcement data collected by the UCR is significantly underreported for several reasons. The UCR does not collect data on the number of victims encountered; it only records those cases in which law enforcement identify a perpetrator who is suspected of having committed a crime, which of course does not include all victimization. And, as Farrell and Reichert (2017) argue, perhaps more concerning is the reality that law enforcement data has been riddled with false negatives and false positives, misclassification, and double counting. Law enforcement agents often misidentify trafficking as another crime with which they have more familiarity, typically prostitution or human smuggling, and sometimes the victims are misidentified as offenders themselves. Furthermore, federal law enforcement data often reports number of investigations, rather than number of victims or perpetrators, further obfuscating the prevalence of trafficking crime and victimization. Duplicate reporting by multiple agencies collaborating on an investigation can, conversely, falsely inflate the numbers reported (Farrell & Reichert, 2017). The result is data that is inadequate to the task of prevalence estimation.

It is worth noting that there have been other estimation techniques applied to the estimation of human trafficking in the United States in recent years. For example, Abt Associates researchers used an oversampling survey strategy to estimate a Minnesota county prevalence of human trafficking victims (Shively et al., 2019). This extension of traditional complex survey sampling is another viable option for estimating hidden population size, including human trafficking. Likewise, in 2004, the Human Rights Center at Berkeley and the non-governmental organization (NGO) Free the Slaves published an estimate of forced labor in the U.S. obtained by triangulating a number of sources: a survey of newspaper articles reporting incidents of forced labor between January 1998 and December 2003, a telephone survey of forty-nine service providers that had worked with or were knowledgeable of forced labor cases, and a review of reports published by the U.S. government regarding the number of forced labor cases it has investigated and prosecuted (Fletcher, Bales, & Stover, 2005).

Some U.S. state legislatures have recently begun enacting laws that require state agencies to collect data on human trafficking victimization from both law enforcement and social service providers. The U.S. Department of Health and Human Services (Office on Trafficking in Persons, 2017) reported that as of July 2017, 30 states have legislation mandating, or strategic plans in place regarding, the collection of data on human trafficking. These new initiatives allow data on both law enforcement identified crimes, and on reported victimization to be represented. This new frontier in data collection is very promising, but at this time, there is little in the way of research based on that data, perhaps because data collection at the state-level is far from complete. Though U.S. state authorities sometimes collect data from a wider variety of sources, including social service agencies, that diversity can introduce many of the same issues and additional significant problems with a lack of shared definitions and a lack of adherence to the legal definitions that might serve as a unifying criteria for inclusion (Dempsey, 2017). Many states have focused their legislative and data collection efforts primarily on domestic minor sex trafficking of girls, which reduces the number of adult, male, and labor trafficking cases that are integral to a complete portrait of trafficking in a region.

Furthermore, without a standardized, de-identified mandatory reporting mechanism (such as electronic collection mandated by law), states are collecting data only from agencies and organizations that voluntarily provide their client data, which is something that many organizations fear will violate their internal confidentiality agreements with their clients or funding agencies. For example, 42 U.S. Code § 13925 prohibits agencies that receive federal grant funding for victims of domestic violence, dating violence, sexual assault, or stalking to disclose certain identifiable information, including full name, address, contact information, identification numbers, or in conjunction with one of the aforementioned, other information such as date of birth, race or religion. In the state of Louisiana, this provision caused some organizations to refuse to participate in state-mandated reporting (Louisiana Department of Children and Family Services, 2018). Federal agencies confirmed that this was a problem in other states as well. Limitations such as these are hindering efforts to produce local estimates of human trafficking in the United States.

The Dark Figure

These challenges are formidable, but they do not present impenetrable barriers to reliable estimates of trafficking in the United States. Instead, these challenges point to the variation of “dark figures” within the measurement of crime and victimization. The frequency of crime reported by a representative victimization survey is typically compared to official crime reports and statistics.

That comparison yields a “dark figure,” the difference between the reported crimes and the actual crime rates uncovered by a victimization survey. There is an accepted “rule of dark figures” which states that the more serious the crime, the smaller the “dark figure” will be. Hence, virtually all murders are reported to the authorities, so the difference between official murder rates and actual murder rates will be small. By contrast, in the U.S. only about a third of all property crimes are reported (Morgan & Kena, 2018).

The “rule of dark figures,” however, is confounded by other factors. Rape and other forms of sexual assault are serious violent crimes that reverse the “rule.” In the United States, according to data drawn from the NCVS, between 2006 and 2010, an average of 65% of sexual assaults went unreported to law enforcement. In some years, the under-reporting rate was as high as 77%. No other serious violent crime measured by the NCVS approaches these rates of under-reporting (Langton, Berzofsky, Krebs, & Smiley-McDonald, 2012). The reasons for this under-reporting are well-known: the failure of law enforcement and the courts to effectively support victims of rape, stigma linked to sexual assault, a sense of shame felt by victims, reluctance to identify as a victim, fear of perpetrators, and inconsistent definitions of “rape” and “sexual assault” held by survey designers and survey subjects.

Similar practical, psychological, and societal barriers often hinder the reporting of human trafficking victimization to authorities. Identification of oneself as a victim of trafficking requires individuals to overcome a barrage of psychological barriers, including fear of traffickers, reluctance to engage with law enforcement, fear of deportation, stigmatization, post-traumatic stress disorder, fear of re-traumatization, reluctance to testify, lack of feasible alternatives to their exploitative situations, and shame, all of which have been richly documented in the last fifteen years (Clawson & Dutch, 2008; Farrell, McDevitt, & Fahy, 2008, 2010; Love, Hussemann, Yu, McCoy, & Owens, 2018). Moreover, victims of human trafficking are often secluded by traffickers, may move often, and sometimes struggle to communicate through language barriers or because of surveillance. Such hidden and transient populations are notoriously difficult to measure because they are not reached by surveys or captured in crime statistics. Even more confounding, trafficked individuals may not know that they would technically be considered the victim of a crime at all, and they may be reluctant to identify as such if it will mean stigmatization or disclosure of something they otherwise want to keep secret (Tyldum, 2010). This is in part due to the complex, counter-intuitive use of the term “trafficking” to describe forced labor when most people understand it to mean movement of a commodity. American law complicates the issue even more, as it includes as criminal trafficking any commercial exchange of sex with a minor, even when there are no third party perpetrators (colloquially known as “pimps”) involved and even when youth describe their experiences as voluntary. Additionally, the American cultural motif of “personal responsibility” encourages people to take responsibility for their life experiences and avoid blaming others, even when they have been victimized or exploited (Murphy, 2017). This means that many people who would fit the legal definition of trafficking do not readily recognize themselves as such. Add to that the fact that survivors sometimes encounter inappropriate responses from law enforcement and the courts and do not trust that their own ideals of justice will be served, and there is a veritable storm of factors that lead to under-reporting (Love et al., 2018).

Human trafficking victimization suffers from yet another specific hindrance to reporting. Crime survey questions typically characterize crimes as a discrete, and presumably short-term, event and assume that victims are able to report the occurrence of their victimization as having occurred at a specific time and place in the past, even if the crime was not reported to the authorities. However, the crime of trafficking presents a special challenge to estimation because of its indeterminate duration. As Datta and Bales (2013) note of contemporary forms of slavery (and can be applied to trafficking):

> Enslavement as a crime is more a process than an event; it is an open-ended victimization. At the initiation of the crime of enslavement, it is difficult to predict how long the victimization will last in that it is ultimately limited only by the lifespan of the victim. Additionally, throughout the indeterminate period of victimization, and unlike the victims of

> most crimes, the slavery victim is almost always held *incommunicado*, unable to call for help, and unavailable to be contacted so that their experience may be recorded and counted and compared to official reports of crime.

It is worth noting that even when those who have been victims of human trafficking crime gain freedom, they also, like those who have been sexually assaulted, feel stigma and shame and are consequently less likely to report their victimization. For these reasons calculating the "dark figure" of trafficking crime is not possible using typical statistical methods, very often because the available data as it is currently being collected by governmental and other agencies fail to meet necessary standards of comparability. Multiple Systems Estimation provides an avenue for overcoming these shortcomings in the estimation of human trafficking.

Multiple Systems Estimation (MSE)

Multiple systems estimation (MSE), an extension of the mark-recapture approach, has been used in criminology, human rights, and public health research, to estimate hard-to-reach populations, hidden populations, and victims of murder and mass atrocities. A survey of MSE and relevant software are provided by Baillargeon and Rivest (2007) and some interesting historical perspectives are given by Goudie and Goudie (2014). A mark-recapture approach was used in the 2011 population census of the United Kingdom to estimate the undercount of the total population of the country. An elementary description distributed to the general public was provided by Benton (2010). Though only recently applied to human trafficking and contemporary forms of slavery, MSE has been applied to various questions of public policy sensitivity, for example to estimate the numbers of casualties in armed conflicts (Manrique-Vallier, Price, & Gohdes, 2013), and numbers of drug users and heroin-related death rates (King, Bird, Overstall, Hay, & Hutchinson, 2013). It is especially well-suited for addressing the particularly hard to ascertain dark figures characteristic of human trafficking. MSE was first used to address human trafficking in 2014 in the United Kingdom (Bales, Hesketh, & Silverman, 2015) generating an estimate of 10,000 to 13,000 victims within the country at a time when there were 2,744 known and recorded cases. The research reported in this paper is among the first uses of MSE to generate a scientific estimate of the number of trafficked people in cities or regions of the United States (see also Farrell et al., 2019). The nature of the data required development of novel statistical methodology, building on previous approaches, as described further below.

In practice, multiple systems estimation is the generalization to multiple lists, in this case lists of victims/survivors of human trafficking, of the classical mark-recapture method for estimating a population size. For example, if there are lists from eight different agencies or entities serving victims, which we might call A, B, C, D, E, F, G, H, then there are 256 possible combinations of the presence or absence of any individual victim/survivor from these lists. Of these 256 combinations, 255 include presence on at least one list. The 256th combination, absence from all observed lists, corresponds to the dark figure which we are trying to estimate.

An individual survivor might be only on list A or B or C, or they may be on A&B, A&C, or A&B&C, etc. The combination of lists on which a particular individual appears is called their "capture history." From our data on the capture history of individual cases, we compile the total observed in each of the 255 combinations that involve presence on at least one list. The only combination we do not observe is the number of individuals who are not on any list. MSE uses the observed combination totals (255 of them if there are 8 lists) to estimate the unobserved dark figure. In most applications, the data are coalesced into fewer than 8 lists, but the principle is the same. Unlike most statistical methods for which duplication is anathema, MSE actually takes advantage of multiple reports of the same crime. MSE can be used for nations, as shown in the UK, or other areas or government entities with boundaries such as states, provinces, or cities.

Application of MSE to the New Orleans-Metairie Metropolitan Statistical Area

While we approached several communities with the opportunity to conduct an MSE for human trafficking, the New Orleans MSA was the only one where it was deemed by local agencies to be feasible at the time to collect the necessary data. Their ability to readily participate was facilitated by a culture of data collection and reporting that the community developed through collaborative anti-trafficking efforts. In 2014, the state legislature of Louisiana introduced into law (Louisiana Revised Statutes, 2014) a requirement that all social service organizations who serve victims of human trafficking provide the Department of Children and Family Services (DCFS) with data regarding the number of victims served and the services rendered, and that DCFS submit an annual statistical report with their findings. Every year since, New Orleans MSA organizations are responsible for a significant proportion of the data collected for the report. The New Orleans MSA is the most populous urban area in the state of Louisiana and thus likely to have proportionally high rates of trafficking, but there are a number of other factors that make it a robust environment for data collection on human trafficking.

In 2014, Loyola University's Modern Slavery Research Project collaborated with Covenant House New Orleans (Murphy, Taylor, & Bolden, 2015) to study the prevalence of trafficking among the homeless youth sheltered in their emergency, transitional, and long-term housing programs, which provided the city with a base-line assessment of how the crime was affecting that particular demographic. Using the findings from that study to make the case for requesting funding for a community-wide response to the demonstrated existence of trafficking in the region, the city won the Department of Justice/Office of Victims of Crime Enhanced Collaborative Model (ECM) grant, which established a collaborative, regional task force. The ECM grant has among its priorities the collection and reporting of data and a fundamental requirement for data-based change over the course of the grant. The hiring of coordinators, by the subsequently formed Greater New Orleans Human Trafficking Task Force, who were trained in human trafficking research methods and monitoring and evaluation allowed for the training of grantees and partners in data reporting. The inclusion of a university researcher (co-author Murphy) as a dedicated participant in the task force, as both evaluator and developer of trainings based on evaluation findings, reflected the task force's commitment to data collection and data-informed change, and it also built trust for research and data reporting within the community.

To carry out the MSE-based estimation, the researchers invited twelve law enforcement agencies and social service providers to participate in the study. Participating organizations were selected based on their engagement in the local anti-trafficking task force, funding for assisting victims of human trafficking, and self-identification as agencies engaged in working with survivors of human trafficking. Eight of those invited agreed to participate, with one non-response and three law enforcement agencies unable to participate due to denial of permission from superiors. The eight participating agencies represented a multi-disciplinary array of law enforcement, social service providers, housing providers, and legal assistance providers, which altogether shared information on 185 individuals who had been deemed "confirmed victims of human trafficking" by the organizations during the 2016 calendar year. Organizations were guaranteed the highest level of confidentiality, so it is not possible to provide the specific names of the organizations or associate their names with the data sets they provided, but the advantage of the MSE methodology is that such specificity is not required to estimate prevalence.

Participating organizations reported on both labor and sex trafficking, including adult and minor, foreign national and domestic victims. The majority of the victims served by these organizations were female U.S. citizen sex trafficking victims, but the other demographics were also represented in the complete data set. Agencies shared raw data (including initials, date of birth, and type of trafficking) that the majority of them had already prepared for the DCFS report regarding victims of trafficking. Those data were provided directly to the local research partner (Murphy), either in writing or in person, depending on organizational confidentiality protocols and internal institutional ethical review. The local research partner processed and anonymized the data to create a simple

Table 1. Breakdown of the observed cases according to their occurrence on various lists.

Cases observed on only one list		Cases observed on exactly two lists		Cases observed on exactly three lists	
Lists	Number	Lists	Number	Lists	Number
A	25	A & C	1	A & C & G	1
B	5	A & D	2	A & D & E	1
C	70	A & E	1		
D	33	B & F	1		
E	6	C & D	2		
F	6	C & E	1		
G	6	C & G	1		
H	21	D & E	2		
		E & H	1		

quantitative table as given in Table 1. No other person (in particular the statistician conducting the analysis) received any more information than the anonymized table.

Some of the individual victims were known to more than one organization, as shown in Table 1; these overlaps give the data needed to carry out multiple systems estimation to obtain an estimate of the total number of those trafficked/enslaved, including those who do not appear on any observed list. The analysis gave a point estimate of 997 victims of trafficking, with a 95% confidence interval of 645 to 1618 victims.

We recommend rounding to the nearest 50 to avoid suggesting any spurious accuracy, thereby placing the point estimate of the total population of trafficked people in the New Orleans MSA in 2016 at 1000 with a confidence interval from 650 to 1600. As can be seen this is not an enumeration of trafficked persons in the New Orleans MSA; rather, it is an estimate taking account of the "dark figure" built upon the number of uncounted and unrecorded crimes and victims within the geographical area, as well as the number that are known. But to sum up, our estimate using this method is that *there were between 650 and 1600 trafficked people in the New Orleans MSA in this time period.*

The Statistical Methodology

The model fitted to the data was a Poisson log-linear model as set out by Cormack (1989) for the number of cases occurring on any particular combination of lists. For any collection S of lists, let N_S be the number of cases that appear on the lists in S and no others. We then model N_S as Poisson(λ_S), where

$$\log \lambda_S = \mu + \sum\nolimits_{i \in S} \alpha_i + \sum\nolimits_{i,j \in S} \beta_{ij}$$

For example, the number of cases only on list 2 has expected number $\exp(\mu + \alpha_2)$and the number on lists 3 and 4 but no others has expectation $\exp(\mu + \alpha_3 + \alpha_4 + \beta_{34})$. The expected value of the dark figure is $\exp(\mu)$, since this corresponds to the case where the collection S is empty, so the estimate and confidence interval for μ give the estimate and confidence interval for the dark figure.

The data in Table 1 demonstrate a typical feature of human trafficking MSE data, not accounted for in existing methodology and software, that there are pairs of lists which have no observed cases in common. Indeed this is the case for 18 of the 28 possible pairs of lists. If the β_{ij} parameter for such a pair of lists is included in the model, then its maximum likelihood estimator is $-\infty$. Motivated by the New Orleans area data, Chan, Silverman, and Vincent (2019a) have developed theory and software to account for this aspect. It includes consideration of its impact on the choice of which interaction terms β_{ij} are included in the model, whether or not any cases are actually observed in the overlap between lists i and j.

The methodology is implemented and documented within the R package SparseMSE (Chan, Silverman, & Vincent, 2019b). When applied to the data in Table 1, the method shows that there is not sufficient statistical evidence to include any of the β_{ij} in the model, and that the data are well fitted by a model where $\log \lambda_S = \mu + \sum_{i \in S} \alpha_i$. In simple terms, the 18 zero overlaps are well within what would be expected by chance if the eight lists were thought of as independent "captures" of the population, with capture probabilities specific to the individual lists.

The SparseMSE package then uses standard Generalized Linear Model methodology (the built-in R program glm) to obtain point estimates and confidence intervals for each of the parameters, in particular for the parameter μ, which then give, by an exponential transformation, an estimate and confidence interval for the dark figure itself. Adding the 185 cases actually observed gives the estimate and confidence interval for the total population.

Because some of the list totals are so small, a five-list version of the data was also produced by combining the four smallest lists BEFG into a single list. Applying the methodology then gives a point estimate of 1034 and a confidence interval of (658, 1709); again, there is no evidence that the interaction terms should be included in the model. It is reassuring that the five-list and eight-list analyses give such similar results. Further details are given in Chan et al. (2019a), which allows the analysis we have presented to be fully reproduced.

Scientific and Policy Implications of Findings

There are three reasons why we believe this use of MSE is important. First, MSE is currently the only method of estimating the scale of human trafficking in developed countries that does not rely on extrapolation from secondary sources but on actual reports of confirmed victims. Second, the multiple systems approach has also proved to be fruitful when applied to other types of human rights violations and difficult-to-reach populations. And, third, if multiple systems estimation is tested and refined and can be verified and found to be useful, then public and policy responses to human trafficking might be placed on a firmer footing, as has already been the case in the United Kingdom, where the multiple systems estimate played an important role in the launch of national law in the Modern Slavery Act 2015. More reliable and realistic estimation can aid in the development of appropriate policies and the effective allocation of resources to the control or eradication of human trafficking and modern slavery.

Challenges in the Use of MSE

This research illuminated two key challenges to achieving an accurate estimation. Neither of these challenges is statistical, though they can be thought of as methodological. The first has to do with definitional variation, likely endemic to the United States context. The researchers originally proposed a study of the prevalence of "slavery" in the New Orleans MSA, but because the U.S. definition of "human trafficking" is the dominant paradigm for understanding forced labor, and because that definition includes minor engagement in commercial sex without evidence of force, fraud, or coercion, the data the agencies were collecting typically did not differentiate between those who were forced laborers in the sex industry and those who were identified as sex trafficking victims by dint of their age. Furthermore, there is some community and academic opposition to the term "slavery" as it is applied in a U.S. context because of the racialized history of slavery and its potential for obfuscating legacies of racial violence. This meant that "slavery" was not an appropriate operationalizable category in the available dataset. Thus, the focus of the study was to determine the prevalence of trafficking, using the data on "confirmed victims of human trafficking," which conformed to the specific requirements of state-level reporting through DCFS.

However, follow-up interviews with several of the participants in this study revealed that organizations were using at least three different definitions that suggested variations in the understanding of both "confirmed" and "trafficking." Several of the organizations were using the legal definition of trafficking in the United States, which included all those who had been induced to work through force, fraud, or coercion, and all those who had engaged in a commercial sex act under the age of 18, which was the shared definition the researchers employed at the outset of the study. At least one of the organizations, however, was also counting adult-age sex workers who were engaged with a third-party controller (or "pimp"), even when the person did not confirm that force, fraud, or coercion initiated or maintained the relationship (in this case, we asked the organization to re-sort their data to indicate cases of force, fraud, and coercion, and they complied). One organization counted people working in the broader sex trade, including as exotic dancers, who preferred other work, regretted their decisions, or felt they had few options, but who were paid while working and were able to leave the job thus demonstrating a level of personal agency and a lack of coercion not typically associated with "trafficking." In this case, the definitions of force, fraud, and coercion were used more broadly than in other organizations. Differing definitions of what constituted labor trafficking also arose within the community during the process of the study. One organization interpreted the law to suggest that people can be confirmed as victims of labor trafficking only if they have been transported into the country as a result of the trafficking victimization; the inducement of labor through force, fraud, or coercion outside the migration experience was not identified as trafficking. One organization counted forced engagement in the drug trade as labor trafficking, but others did not. A former task force coordinator indicated that there was inadequate definitional guidance or mandates provided by federal grant liaisons to reconcile the different definitions used by individual agencies within the task force, which created disparities in their data that remained unresolved.

The second challenge was related to the variation in goals and approaches within the different groups from which the lists of cases were collected. Many of the groups' human trafficking programs were dedicated to only one form of trafficking – either sex or immigrant forced labor. Only two of the eight organizations reported engaging with both labor and sex trafficking victims, one of which only reported labor trafficked clients who had also been trafficked for sex. One organization worked exclusively with labor trafficked clients, due to funding regulations. The other five had identified only sex trafficking victims. This imbalance is due, in part, to funding and mission requirements, but it also reveals a fairly widespread dynamic in the United States that focuses resources for human trafficking primarily on sex trafficking and siloes labor from sex trafficking. Only 20 of the 185 victims identified by the 8 organizations had been trafficked for labor. Thus, we were likely unable to capture the full extent of labor trafficking in the New Orleans MSA.

The collection and processing of these data also revealed several unpredicted findings. It was expected that all (or at least the majority of) minors who were listed as victims of trafficking within the New Orleans MSA would have been duplicated on the list of the organization that is assigned to handle mandatory reporting of child abuse in the city. In fact, there were very few overlaps, though many of the victims were minors. This pointed to a lack of awareness of prescribed mandatory reporting protocols within the city among agencies supporting victims of human trafficking. As a result of this ancillary finding, additional mandatory reporting instructions were subsequently included in the task force workshops and training. In addition, though the overlaps between the lists indicate few shared clients among the organizations, it also suggests minimal referrals were made between service providers participating in the collaborative network of the task force in the first year of its existence. In part as a result of this finding, the task force is developing an online services referral network and has increased networking opportunities across the community of anti-trafficking organizations. In these ways the MSE process and methods exposed significant gaps in policy and practice that, if corrected, will improve services to victims/survivors.

Recommendations

How might we address these challenges to measuring prevalence in the United States and other similar countries? The New Orleans case can be instructive in several ways. In the New Orleans MSA, the federal funding requirement for data reporting and data-informed change provided impetus for more robust data collection as well as incentives for analysis of the trends. Questions regarding what constituted trafficking for the purposes of reporting remained, however. All U.S. federal agencies (and we note that this would apply to other countries as well) that require data collection should provide more detailed guidance and technical assistance regarding the precise definition of trafficking victimization. Furthermore, U.S. federal and other agencies should provide mandatory data collection training to facilitate improved reporting for all funded human trafficking programs. Routine checks should be integrated into the system to ensure that state and local groups and agencies are correctly reporting.

Second, the institution of state-level data collection on human trafficking in Louisiana facilitated prevalence research and meant that data was readily available and prepared for analysis. States and other government entities should mandate both data collection and training on how to collect and analyze it, so that the response to human trafficking can be geared to the reality on the ground. Standardized state-mandated electronic reporting systems for human trafficking for all law enforcement agencies that aligns with national uniform crime reporting protocols will further facilitate the estimation of the prevalence of trafficking. All state and federal crime and victimization reporting programs should include trafficking – at the very least including human trafficking/involuntary servitude and human trafficking/commercial sex, but ideally should include additional capacity to mark multiple crimes at one time, so that the complex terrain of trafficking can be understood.

Third, the example of New Orleans provides further encouragement that a data collection and sharing culture can be built by anti-trafficking collaboratives over time. Providing training on data collection, employing skilled data analysts, collaborating with local researchers, and engaging in data-informed change in strategic planning allowed New Orleans area agencies to be prepared to analyze their data, report findings, and use the data for growth and improvement of victim services. Community anti-trafficking task forces should engage skilled analysts and researchers to ensure that their strategic planning is guided by sound research and evaluation methodologies.

Researchers should be aware that definitional variation still complicates statistical findings about prevalence, whether those variations are internal to a task force, reported at the state or federal level, or operationalized by external researchers. At the outset of task force (or other community-based response) development, a clear definition, adhering to state and national definitions, should be established to ensure comparability of data. Once a single operational definition is determined and applied, additional studies could use multiple methods to study the same population, improving the reliability of estimates. It is also important to note that researchers seeking to use MSE for the estimation of trafficking prevalence will face difficult political and social challenges. Concerns about confidentiality and unwillingness to share data with "competitor" organizations may hinder cooperation among relevant agencies. Likewise, front-line service providers sometimes justifiably have little time, attention, or staff to devote to data collection while their clients are in situations of extreme vulnerability and need. Researchers using state-or local-level data should take precautions regarding these limitations.

Conclusion

We have demonstrated the use of Multiple Systems Estimation as a way to discover with greater reliability the prevalence of trafficking victimization within a major city in the United States. In doing so, we discovered challenges to using this method of estimation that were not so much statistical as methodological, social, political, and, in a word, human. It is important to illuminate these challenges in that the response to trafficking crime, not just in the United States but globally, is hampered by a lack of reliable measurement, which is in turn impeded by these challenges.

Perhaps the most damaging to the measurement process, and politically daunting, is the challenge of how trafficking is defined legally and conceptually by different agencies and non-governmental organizations, and operationally by social scientists. There exists more than one legal definition within the national laws of the United States depending on which sub-type of trafficking is being considered. Additionally, some states have enacted legislation with further variations of definition. NGOs construct their own definitions based on which constituencies of need they seek to support, on their ideological approaches to the social problems they face, or in response to donors or other actors. Social scientists, who may be working for governments or NGOs or some other entity, will seek to construct an operational definition that allows a specific activity of interest to be delineated and studied. In the ongoing circles of influence, politicians may craft new legal definitions based not so much on jurisprudence as the influences and arguments of interest groups, including but not limited to those just mentioned. The result is a hodge-podge of definitions at different levels of governance or civil society provision, many of them contradictory. While the victim of an intentional killing might fall within a short list of possible and relatively clear-cut and mutually exclusive categories, the victims of human trafficking might be classified as victims of a wide variety of crimes, or be seen and treated as not victims at all. They may even be considered perpetrators themselves. Definitional confusion leads to confusion in the provision of legal protections and support by NGOs and others. It also makes reliable estimation of the size of the problem, which is essential to appropriate social and governmental response, extremely difficult. There have been a number of efforts, notably the promulgation of the Bellagio-Harvard Guidelines (Research Network on the Legal Parameters of Slavery, 2012), to address this definitional confusion. However, the trend seems to be a continued broadening of definitions, some of which seem to deny that the phenomenon of trafficking and slavery exists at all, and none of which have drawn upon the knowledge of the lived experience of being trafficked and enslaved available from the many survivors of this crime (Nicholson, Dang, & Trodd, 2018).

A second key challenge is statistical. As already noted, it is typical of data collected regarding human trafficking for multiple systems estimation that not every pair of lists contains cases in common. The standard MSE approaches do not deal properly with this case, and so one of the authors has developed an appropriate algorithm (Chan et al., 2019a,b.). In addition, Silverman (2018) shows that some methods do not give very stable results on real modern slavery or human trafficking datasets. There is a need for a larger corpus of publicly available datasets specifically obtained in the modern slavery and human trafficking contexts to facilitate the development of robust standard statistical methods, and it is hoped that this paper will encourage both the wider application of the approach and the publication of data in easily accessible form.

Acknowledgments

This work was supported by the Arts and Humanities Research Council and the Economic and Social Research Council [grant number ES/P001491/1] grant Modern Slavery: Meaning and Measurement (PaCCS Transnational Organised Crime, University of Nottingham, 2016-18). Additional support was provided by the British Academy Visiting Fellows Program and the National Humanities Center (U.S.).

ORCID

Kevin Bales http://orcid.org/0000-0003-1516-7148
Laura T. Murphy http://orcid.org/0000-0002-1220-9931
Bernard W. Silverman http://orcid.org/0000-0002-4059-2376

References

Baillargeon, S., & Rivest, L.-P. (2007). Rcapture: Loglinear models for capture-recapture in R. *Journal of Statistical Software*, *19*(5), 1–31. doi:10.18637/jss.v019.i05

Bales, K., Hesketh, O., & Silverman, B. W. (2015). Modern slavery in the UK: How many victims? *Significance*, *12*(3), 16–21. doi:10.1111/j.1740-9713.2015.00824.x

Benton, P. (2010). *Trout, catfish and roach: The beginner's guide to census population estimates*. London, UK: Office for National Statistics.

Brunovskis, A., & Surtees, R. (2010). Untold stories: Biases and selection effects in research with victims of trafficking for sexual exploitation. *International Migration*, *48*(4), 1–37. doi:10.1111/imig.2010.48.issue-4

Bureau of Justice Statistics. (2017). Data collection: National Crime Victimization Survey (NCVS). Retrieved from https://www.bjs.gov/index.cfm?ty=dcdetail&iid=245

Chan, L., Silverman, B. W., & Vincent, K. (2019a). Multiple systems estimation for sparse capture data: Inferential challenges when there are non-overlapping lists. arXiv:1902.05156 [stat.ME]. Retrieved from https://arxiv.org/abs/1902.05156

Chan, L., Silverman, B. W., & Vincent, K. (2019b). *SparseMSE: Multiple systems estimation for sparse capture data*. [R package]. Retrieved from https://cran.r-project.org/web/packages/SparseMSE

Chisolm-Straker, M. (2017). *Recognizing human trafficking among homeless youth*. Newark: Covenant House New Jersey.

Chisolm-Straker, M., Sze, J., Einbond, J., White, J., & Stoklosa, H. (2018). A supportive adult may be the difference in homeless youth not being trafficked. *Children and Youth Services Review*, *91*, 115–120. doi:10.1016/j.childyouth.2018.06.003

Clawson, H., & Dutch, N. (2008). *Identifying victims of human trafficking: Inherent challenges and promising strategies from the field*. ASPE, U. S. Department of Health and Human Services. Retrieved from https://aspe.hhs.gov/report/identifying-victims-human-trafficking-inherent-challenges-and-promising-strategies-field

Cormack, R. M. (1989). Log-linear models for capture-recapture. *Biometrics*, *45*, 395–413. doi:10.2307/2531485

Curtis, R., Terry, K., Dank, M., Dombrowski, K., & Khan, B. (2008). The commercial sexual exploitation of children in New York City: Volume one: The CSEC population in New York City: Size, characteristics and needs. Report submitted to the National Institute of Justice. New York, NY: Center for Court Innovation and John Jay College of Criminal Justice.

Dank, M., Yahner, J., Madden, K., Bañuelos, I., Yu, L., Ritchie, A., … Conner, B. (2015). *Surviving the streets of New York: Experiences of LGBTQ youth, YMSM, and YWSW engaged in survival sex*. Washington, DC: Urban Institute.

Datta, M. N., & Bales, K. (2013). Slavery in Europe: Part 1, estimating the dark figure. *Human Rights Quarterly*, *35*, 817–829. doi:10.1353/hrq.2013.0051

Dempsey, M. M. (2017). What counts as trafficking for sexual exploitation? How legal methods can improve empirical research. *Journal of Human Trafficking*, *3*, 61–80. doi:10.1080/23322705.2017.1280325

Farrell, A., Dank, M., Kafafian, M., Lockwood, S., Pfeffer, R., Hughes, A., & Vincent, K. (2019). *Capturing human trafficking victimization through crime reporting*. Report on award number 2015-VF-GX-0105 for the National Institute of Justice. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/252520.pdf

Farrell, A., McDevitt, J., & Fahy, S. (2008). *Understanding and improving law enforcement responses to human trafficking*. Report for the National Institute of Justice. Institute on Race and Justice, Northeastern University.

Farrell, A., McDevitt, J., & Fahy, S. (2010). Where are all the victims? Understanding the determinants of official identification of human trafficking incidents. *Criminology & Public Policy*, *9*, 201–233. doi:10.1111/j.1745-9133.2010.00621.x

Farrell, A., & Reichert, J. (2017). Using US law-enforcement data: Promise and limits in measuring human trafficking. *Journal of Human Trafficking*, *3*, 39–60. doi:10.1080/23322705.2017.1280324

Federal Bureau of Investigation (2014). Human Trafficking in the Uniform Crime Reporting (UCR) Program. Retrieved from https://ucr.fbi.gov/human-trafficking

Fletcher, L. E., Bales, K., & Stover, E. (2005). Hidden slaves: Forced labor in the United States. *Berkeley Journal of International Law*, *23*, 47–111.

Fong, R., & Berger Cardoso, J. B. (2010). Child human trafficking victims: Challenges for the child welfare system. *Evaluation and Program Planning*, *33*, 311–316. doi:10.1016/j.evalprogplan.2009.06.018

Gibbs, D. A., Henninger, A. M., Tueller, S. J., & Kluckman, M. N. (2018). Human trafficking and the child welfare population in Florida. *Children and Youth Services Review*, *88*, 1–10. doi:10.1016/j.childyouth.2018.02.045

Goudie, I. B. J., & Goudie, M. (2014). Who captures the marks for the Petersen estimator? *Journal of the Royal Statistical Society, Series A*, *170*, 825–839. doi:10.1111/j.1467-985X.2007.00479.x

Goździak, E. (2015). Data matters: Issues and challenges for research on trafficking. In M. Dragiewicz (Ed.), *Global human trafficking: Critical issues and contexts* (pp. 23–38). London & New York, NY: Routledge.

King, R., Bird, S. M., Overstall, A. M., Hay, G., & Hutchinson, S. J. (2013). Estimating prevalence of injecting drug users and associated heroin-related death rates in England by using regional data and incorporating prior information. *Journal of the Royal Statistical Society, Series A*, *177*, 209–236. doi:10.1111/rssa.12011

Langton, L., Berzofsky, M., Krebs, C., & Smiley-McDonald, H. (2012). Victimizations not reported to the police, 2006-2010. Bureau of Justice Statistics, US Dept. of Justice, (NCJ 238536). Retrieved from https://www.bjs.gov/content/pub/pdf/vnrp0610.pdf 10.1094/PDIS-11-11-0999-PDN

Louisiana Department of Children and Family Services. (2018). *Human trafficking, trafficking of children for sexual purposes, and commercial sexual exploitation: Annual report*. Retrieved from http://www.dcfs.louisiana.gov/assets/docs/searchable/Child%20Welfare/PlansReports/Human%20Trafficking%20Report%202018.pdf

Louisiana Revised Statutes. (2014). *46:2161.1 – Human trafficking victims services plan; adults*. Retrieved from https://www.lawserver.com/law/state/louisiana/la-laws/louisiana_revised_statutes_46-2161-1
Love, H., Hussemann, J., Yu, L., McCoy, E., & Owens, C. (2018). *Justice in their own words: Perceptions and experiences of (in)justice among human trafficking survivors*. Washington, DC: Urban Institute.
Lutnick, A. (2016). *Domestic minor sex trafficking: Beyond victims and villains*. New York, NY: Columbia University Press.
Manrique-Vallier, D., Price, M. E., & Gohdes, A. (2013). Multiple systems estimation techniques for estimating casualties in armed conflicts. In T. Seybolt, B. Fischhoff, & J. Aronson (Eds.), *Counting civilian casualties: An introduction to recording and estimating nonmilitary deaths in conflict* (pp. 77–93). New York, NY: Oxford University Press.
Marshall, B. D. L., Shannon, K., Kerr, T., Zhang, R., & Wood, E. (2010). Survival sex work and increased HIV risk among sexual minority street-involved youth. *Journal of Acquired Immune Deficiency Syndromes*, *53*, 661–664.
Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. Chicago, IL: University of Chicago Press.
Middleton, J. S., Gattis, M. N., Frey, L. M., & Roe-Sepowitz, D. (2018). Youth experiences survey (YES): Exploring the scope and complexity of sex trafficking in a sample of youth experiencing homelessness. *Journal of Social Service Research*, *44*, 141–157. doi:10.1080/01488376.2018.1428924
Morgan, R. E., & Kena, G. (2018). Criminal Victimization, 2016: Revised. Bureau of Justice Statistics, US Dept. of Justice, (NCJ 252121). Retrieved from https://www.bjs.gov/content/pub/pdf/cv16.pdf
Murphy, L. T. (2017). *Labor and sex trafficking among homeless youth: A ten-city study*. New Orleans: Loyola University New Orleans, Modern Slavery Research Project.
Murphy, L. T., Taylor, R., & Bolden, C. L. (2015). *Trafficking and exploitative labor among homeless youth in New Orleans*. New Orleans, LA, USA: Modern Slavery Research Project.
Nawyn, S. J., Kavakli Birdal, N. B., & Glogower, N. (2013). Estimating the extent of sex trafficking: Problems in definition and methodology. *International Journal of Sociology*, *43*(3), 55–71. doi:10.2753/IJS0020-7659430303
Newton, P. J., Mulcahy, T. M., & Martin, S. E. (2008). *Finding victims of human trafficking*. Report on Grant # 2007VTBX0001 for the National Institute of Justice. Bethesda, MD: National Opinion Research Center.
Nicholson, A., Dang, M., & Trodd, Z. (2018). A full freedom: Contemporary survivors' definitions of slavery. *Human Rights Law Review*, *18*, 689–704. doi:10.1093/hrlr/ngy032
O'Brien, E. (2010). Dark numbers: Challenges in measuring human trafficking. *Dialogue E-Journal*, *7*(2). Retrieved from https://eprints.qut.edu.au/48257/
Office on Trafficking in Persons, Administration for Children and Families. (2017). *State data collection efforts on human trafficking*. Retrieved from https://www.acf.hhs.gov/sites/default/files/otip/fs_data_collection_state_data_collection_efforts.pdf
Research Network on the Legal Parameters of Slavery (2012). *Bellagio-Harvard guidelines on the legal parameters of slavery*. Retrieved from http://www.law.qub.ac.uk/schools/SchoolofLaw/FileStore/Filetoupload,651854,en.pdf 10.1094/PDIS-11-11-0999-PDN
Shively, M., Kling, R., Berninger, A., Christopher, L., Rodriguez, B., & Nadel, M. (2019). Advancing human trafficking prevalence estimation: Key findings from developing and field testing hidden population method. Presented at National Academy of Sciences, Engineering, Medicine: A Workshop on Estimating the Prevalence of Human Trafficking in the United States (April 9, 2018). Retrieved from https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_192207.pdf.
Silverman, B. W. (2018). Model fitting in multiple systems analysis for the quantification of modern slavery: Classical and Bayesian approaches. arXiv:1902.06078 [stat.AP]. Retrieved from https://arxiv.org/abs/1902.06078
Tyldum, G. (2010). Limitations in research on human trafficking. *International Migration*, *48*(5), 1–13. doi:10.1111/j.1468-2435.2009.00597.x
Tyldum, G., & Brunovskis, A. (2005). Describing the unobserved: Methodological challenges in empirical studies on human trafficking. *International Migration*, *43*(1–2), 17–34. doi:10.1111/imig.2005.43.issue-1-2
Zhang, S. X., Spiller, M. W., Finch, B. K., & Qin, Y. (2014). Estimating labor trafficking among unauthorized migrant workers in San Diego. *The Annals of the American Academy of Political and Social Science*, *653*, 65–86. doi:10.1177/0002716213519237