**METHODOLOGICAL APPROACH TO STUDY THE HIDDEN POPULATION: A CASE OF TRAFFICKED WOMEN IN MEXICO**

**DR. ARUN KUMAR ACHARYA**
**PROFESOR-INVESTIGADOR**
**INSTITUTO DE INVESTIGACIONES SOCIALES**
**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN**
**BIBLIOTECA UNIVERSITARIA "RAÚL RANCEL FRÍAS"**
**AVENIDA ALFONSO REYES, 4000 NTE, C.P. 64440, MONTERREY, N.L**
**PHONE/FAX: (0052) 81 8329 4237**
**Email: acharya_77@yahoo.com**

# METHODOLOGICAL APPROACH
# TO STUDY HIDDEN POPULATIONS:
# THE CASE OF TRAFFICKED WOMEN IN MEXICO

**Arun Kumar Acharya**
Instituto de Investigaciones Sociales
Universidad Autónoma de Nuevo León, México

**ABSTRACT**

For a variety of reasons, there is always an inconvenient to get data from the key subpopulations, particularly those whose behaviors are illegal or illicit. Non-probability sampling approaches such as network sampling are the simplest to carry out, but are subject to a high risk of sampling bias, whereas most of the probability sampling methods considered are limited because they are only adequate under certain circumstances and for some groups.

Considering the difficulties to obtain quality of data and information, in this paper an attempt has been made to stand out an appropriate method and technique to study hidden populations. It explains the method and technique used to study trafficked women in Mexico City. Using the time-space sampling and network techniques, the research identified seventy-three trafficked women, sixty were interviewed through a semi-structured questionnaire and thirteen through an in-depth interview.

**Key words:** Research methodology, hidden populations, trafficked women, Mexico.

**INTRODUCTION**

The problem of collecting accurate information about the behavior and composition of social groups arises in many areas of research. In most cases, standard sampling and estimation techniques developed over the past 70 years provide a means for collecting such information (Salganik and Heckathorn, 2004). However, there are a number of important groups for which these techniques are not applicable.

For example, the challenge arises when groups are hidden —that is, no sampling frame exists for them— and, as the behaviors in which members of this groups engage are either illegal or illicit, they generally prefer not to participate in surveillance data-collection activities (Heckathorn, 1997). Thus, meaningful surveillance requires that sampling strategies are both feasible and capable of producing unbiased estimates —or more real estimates with minimal levels of bias—, devised for population subgroups that are not efficiently captured using conventional surveillance data-collection strategies (Magnani, *et al.,* 2005).

The limitation of current statistical methodology has received attention not just from public health researchers, but also from statisticians and sociologists. Standard sampling and estimation techniques require the researcher to select sample members with a known probability of selection. In most cases, this requirement means that researchers must have a sampling frame, almost of all members in the population (Salganik and Heckathorn, 2004). However, for many populations of interest, such list does not exist. In vulnerable environments or settings these

groups will generally consist of trafficked women, drug users, HIV/AIDS patients, men who have sex with men (MSM), among others (Semaan, Lauby, and Liebman, 2002).

There is an important rationale for conducting surveillance among the 'high-risk' group of trafficked women, especially when they are trafficked into prostitution. In this kind of subgroups, a researcher wishing to study a population without a sampling frame could attempt to construct such a frame. For a number of populations, this frame construction is impractical or impossible due, in the first place, to small size of the target population and, secondly, to the difficulty of locating its members. This difficulty could be due to the behaviors' sensitive nature (for example, injection drug users) or simply because it is difficult to distinguish members of the target population from members of the general population (for example, jazz musicians). These special populations that cannot be studied using standard sampling and estimation techniques are called hidden populations (Mills, *et al.,* 1998).

The inability to collect information about hidden populations has complicated existing studies and forced researchers to focus on other problems (Heckathorn, 1997). Imagine a researcher who whishes to begin a study of, for example, a huge problem, and tries to get a sample from the hidden population from which he/she could generalize. A first thought might be to attempt to construct a sampling frame, and then, once the frame is complete, to select people with known probability of selection from the frame. This approach brings the problem back under the purview

of standard sampling and estimation (Salganik and Heckathorn, 2004). However, this *frame construction* is extremely expensive and probably impossible—imagine researcher trying to make a list of all sex workers in a large city.

Another approach would be to reach a large number of people via *random-dialing* and then screen them for membership in the hidden population. Again, this approach is extremely costly and potentially very inaccurate (Salganik and Heckathorn, 2004). If a small portion of a population in a city belongs to the hidden population, it would take much more screening interviews to yield a small sample size. Further, it is unlikely that the resulting sample would be a random sample from the hidden population, because many members are not reachable by phone, or would not reveal their behavior to an unknown interviewer over the phone. The nature and magnitude of these biases would be unknown, and thus there would be no way to generalize from this type of sample to the entire hidden population.

Another way to collect information about the behaviors and composition of hidden populations is to take a sample of the target population in an institutional setting, for example, injection drug users in a drug rehabilitation program. This can provide researchers with valuable information, but since the members of the population who enter institutional settings are a nonrandom sample from the hidden population, it is impossible to use samples from institutional settings to make accurate estimates about the entire hidden population.

The other method to obtain information from hidden populations is the *network perspective.* Using the extra information available in the social network we can design a sampling and estimation scheme that, in many cases, is both cheaper and more accurate than the methods commonly used. This new sampling and estimation method is called *respondent driven sampling*. The basic idea behind this method is that respondents are selected not from a sampling frame, but from the friendship network of existing members of the sample (Salganik and Heckathorn, 2004).

The sampling process begins when researchers select a small number of *seeds,* which are the first people to participate in the study. These seeds then recruit others to participate in the study. This process of existing sample members recruiting future sample members continues, until the desired sample size is reached.

In this project an attempt has been made to develop a research methodology to study hidden populations. The present paper explains the method and technique used to study trafficked women in Mexico City during the years 2002-2005. Women trafficked for prostitution and sexual exploitation are kept in closed doors. It is hard to find information about them, because they are not easily visible, and are mixed up with women who are in prostitution on their will. Beside this, trafficked women are part of an organized criminal network. In order to study a hidden population, such as trafficked women, a special methodology is needed, which I explain as follows.

**EXISTING METHODS FOR STUDYING HIDDEN POPULATIONS**

This section explains existing literature on research methodology to study the hidden population. Most studies about the characteristics of hidden populations use either targeted sampling or time-space sampling. These approaches differ in their strengths and weaknesses, but one problem is that, for many hidden populations, there is no systematic or principled way to use the information collected to make inferences about the population from which samples are drawn.

For example, in targeted sampling Watters and Biernacki (1989) stated that researchers use a number of outreach techniques to attract a sample of people of the hidden population. This technique is also sometimes called "*street outreach*" because it generally involves sending fieldworkers into the streets to find and recruit members of the hidden population. Thus, targeted sampling does succeed in giving researchers access to a large sample of non-institutional members of the population.

However, this sampling technique is clearly not a random sample; where all people have the same probability of selection. For example, in studies of injection drug users, for safety concerns often require that the interview should conduct only during the day, when most drug scenes are less active. Also, injection drug users who do not congregate in public are almost certain to be missed in the sampling process. There is no way to know the magnitude of these selection biases, so it is

not possible to generalize from the sample to the target population (Salganik and Heckathorn, 2004).

On the other hand, in their study, Muhir, *et al.* (2001) indicates that an alternative to targeted sampling is the time-space sampling. According to them, under this method, first an ethnographic fieldwork is done to construct a sampling frame identifying members of the target population when they gather at a specific location i.e. parks, bars etc. Once the place, day and time has fixed, it taken into consideration as the primary sampling units. Then these units are randomly selected, in some cases with probabilities based on the expected sample yield at the location, and members of the target population entering the place are intercepted and interviewed. As the location, day and time are sampled with a known probability, it is possible to use the sample to make statistical inferences about the population that attends the identified places, but unfortunately, all places are not accessible (Stueve, *et al.,* 2001). For example, some studies do not consider sample from locations with low expected sample yields because it is prohibitively costly. Additionally Kanouse, *et al.,* (1999) stated that private places are generally not accessible to researchers and in some studies due to safety concerns it further limit the choice of places.

Though, the time-space sampling methods have limitations, but it produces probability samples of the population that attend place, day and time, which are accessible to researchers. In some situations this place-attending population differs in unknown ways from the true population of interest. For example, Salganik and

Heckathorn (2004) indicate that drug injectors who appear in public places accessible to researchers are probably not the representative of all drug injectors. This problem introduces unknown bias into the estimates and a bias that could be substantial in some situations and minimal in others.

The above authors put their argument that target sampling as well as the time-space sampling techniques could be best possible sampling techniques to locate the hidden population. But there are other studies on hidden population which describes that techniques like hierarchically ascending classification, cross classification and network sampling (respondent-driven sampling) are the possible research methodologies which could be use to obtain a good result to study the hidden population, which I have elaborate in following.

One of the useful method which can be use to study the hidden population is hierarchically ascending classification analysis, which is often called as cluster analysis. In this method researcher need to constructs successively more general classes or types according to the similarities between the descriptive variables being analyzed to study hidden population (Van Meter, 1990). Similar to the hierarchically ascending method, another method to study the hard to reach population is cross-classification analysis. This method not only analyzes the similarity between variables for all individuals, but also similarity between individuals for all variables. These similarities are then combined or crossed to form coherent blocks of individuals' variables, which are called "polythetic classes" (Van Meter, 1990).

Van Meter (1990) in his study indicates that in the case of drug users, "street ethnographers" or simply "'big city savvy" furnish the necessary knowledge to find these persons, who will act as initial contact nominees, each of which is interviewed using an established questionnaire, and then asked to nominate other drug users to be interviewed. Using the network sampling, Van Meter interviewed thirty-five drug users. This means that a respondent-driven technique can indeed make the sample grow rapidly. This methodology also assures that no individual is interviewed twice and that specific quotas for sex, age, or other social characteristics are met by the nomination procedure.

Similar to Van Meter, in a study of Cocaine users in three different cities (i.e. Rotterdam, Munich, and Rome) of the European Community, Avico, *et al.* (1988) stated that network sampling can quickly and efficiently survey a relatively large sample of current cocaine users. In two months of part-time fieldwork in these three cities using a common instrument and sampling design they interviewed 153 active cocaine users. This is one of the largest samples of cocaine users ever collected in either Europe or America.

In another research on hidden population, Van Meter conducted a comparative research on abortion in Protestant, Catholic, and Muslim cultures. He used the network sampling to identify the women who have performed abortion. He argued that network sampling would appear to be one of the very few valid methods to provide sufficient systematized data to permit a formal analysis and comparison of

this social phenomenon in different cultures (Van Meter, 1986). On the other hand, in a study on trafficking of children in Thailand —conducted by the Southeast Asian Ministers of Education Organization (SEAMEO)— two hundred trafficked children were interviewed though the network sampling method.

Moreover, Van Meter (1990) argues that combination of network sampling and cross-classification analysis is well adapted to the study of hidden populations. This methodology does have problems with the calculations of general population estimates and the explanation of variance. Through the complementary combination of these two methodologies are extremely useful tool in comparative sociological research (Van Meter, 1990). The above studies indicate that in order to study a hard-to-reach population, suitable and appropriate methods are needed. Because of methodological problems, no such in-depth study has ever been done to explore trafficked women, for which I believe this research will help to fill this gap.

**A NEW APPROACH TO STUDY THE TRAFFICKED WOMEN: AN EXAMPLE OF MEXICO CITY**

*Methodology*

In this section I develop a new approach for studying the hidden population of trafficked women. This study was conducted in Mexico City during 2002 to 2005. Sixty trafficked women were interviewed using a semi-structured questionnaire, and thirteen through an in-depth interview.

Mexico City is the second largest city of the world with a population of 25 million (one fourth of the country's population lives in Mexico City). This city is an important administrative, economic, and cultural capital. Due to its importance, thousands of people migrate every day from peripheral areas, as well as from abroad. This city has the largest sex industry in the country.

It is difficult to say the exact number of women that are in prostitution in Mexico City, but according to Teresa Ulloa, president of the *Regional Coalition against Trafficking of Women and Children in Latin America and the Caribbean,* 400 women enter every day into prostitution[1] in Mexico City, out of which 80% do it against their will (González, 2003). For its high demand of commercial sex work, the trafficking of women to Mexico City in recent years is booming (Acharya and Stevanato, 2005).

Once a woman is brought to a brothel, she looses control upon herself and becomes a slave. Violence, exploitation, and sexual torture are common events in the day-to-day life of these victims. In addition, these women are very vulnerable to different kind of diseases, like RTI/STI and HIV/AIDS (Acharya, 2006). As Mexico is observing a growing pandemic of HIV/AIDS, it is very urgent to carry out studies on trafficking, as there is a close relation between this subgroup of population and these deadly diseases. Even though some organizations —like the *Instituto Nacional de las Mujeres* (National Institute of Women)— have attempted to study this subgroup of population, inadequate techniques and methodologies have made

---

[1] This figure includes trafficked and non-trafficked women.

them abandon such efforts (Acharya, 2006). In the present study I have tried to describe the specific methodology and techniques used to study the hidden population of trafficked women.

For this research project I first carried out an ethnographic field work through my constant visit to the brothel area La Merced of Mexico City. By these visits, using the time-space technique I identified a client. Later, I developed a friendly relation by inviting him to my house, to the cinema, and for lunch and dinner. Once I gained his confidence, I informed him of my research project and asked him to find a brothel's owner (pimp) who could help me in my research. At the beginning he declined to help, as this would be very dangerous, but after one month of regular interaction, he agreed and arranged an appointment with a brothel's owner. After discussing my research with her, I told her that I needed to identify some trafficked women for interviews. I explained the meaning of trafficked women and how it differs from other women who are in prostitution. But she declined to give any information, because it would be very risky for her, as well as for me. However, after several conversations, she told me she could help, but asked about what she would gain. I felt by her words she was asking for some money, so I asked her directly what she wanted. Immediately she responded she would charge me for each piece of information; at last, I agreed to pay her 50 pesos (US$ 5) each.

Since that meeting on, she started to look for trafficked women and whenever she found one, she would call me for the interview. So, in this study, the respondents were selected not from a sampling frame, but from the social network of existing

population members. In this case, rather than attempting to directly estimate from the sample, as in traditional sampling and estimation, first I approached the population through my social network, and from there I drew my sample, as shown in the figure below.
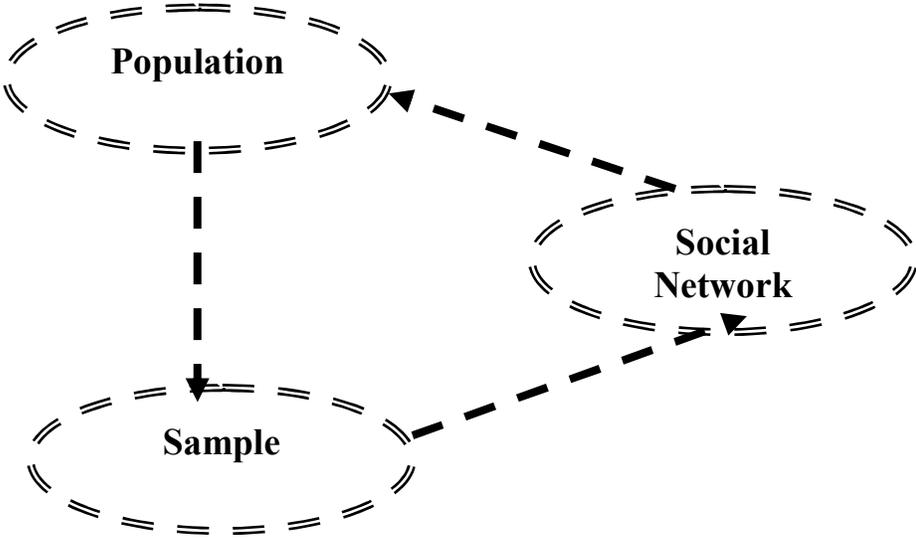
Schematic of Respondent-Driven Sampling

## DATA COLLECTION

Every time I arrived at the brothel to collect information from a trafficked woman, I would go to the indicated woman, identify myself, and tell her about the interview. Most of these women recruit their clients at the roadside in front of the hotel where they worked; however, it was not possible to make the interview there, since they were under constant surveillance by the pimp. So, at the time the woman had no clients, I would pay the solicited amount and would take her to the hotel room as a "client." I adopted this method to keep my identity hidden from the pimp and traffickers, and once getting into the hotel room, I would make the interview.

At the beginning of the interview, I would offer the women to come to my house, to another hotel, or to a restaurant, but they would always say no. They said: "…This is our area; it is not possible for us to go out from this place, because *madrina* (madam) does not permit us to work in other places…". They would often say: "...All kinds of services are available in this hotel, so why do you want to go elsewhere?" Since these women were not allowed to leave the hotel, I had to act as a client. In the hotel room, I would make the interview, which lasted around 30 minutes. After completing the talk, I would invite them to help me find other trafficked women, so I decided to use the network technique to interview more trafficked women.

After a few days I found full cooperation from the interviewed trafficked women. A few months later I had constructed a net of trafficked women from whom I was getting the information for my next interviews. In this way, initially using the time-space and later network sampling methods, sixty interviews through a semi-structured questionnaire and thirteen through in-depth interviews were conducted during 2002 and 2005 among women who were trafficked from the countryside, as well as from different urban centers of Mexico, to Mexico City. All these interviews were conducted in Spanish.

Studying a hidden population is always more costly in economic terms than studying other populations. But the important point is that by using this technique and methodology, a thick body of information and description on trafficked women

in Mexico was produced. Thus, multiple scientific interests can be associated with this single methodology, since the network sampling (respondent-driven sampling) technique is cheaper, quicker, and easier to implement than other methods commonly used to study hidden populations (Semaan, Lauby, and Liebman, 2002). This is a significant advantage, because it means that for a given amount of resources, respondent-driven sampling allows researchers to have more study sites or larger sample sizes than other methods.

**DISCUSSION**

For many years researchers have known that the network sampling technique is an extremely useful way to collect samples from hidden populations, such as trafficked women. It also considered that the network sampling method as a mere convenience sample, so hopelessly full of bias that can only be useful for exploratory purposes. However, in this paper I maintain that if network sampling is properly used, it can produce unbiased information.

The positive property of respondent-driven sampling is that it gives us information not just about the individuals in the population, but also about the network connecting them. Currently, researchers are beginning to explore ways of extracting useful social network information from the sample (Heckathorn and Jeffri, 2001). Due to the role that networks play in the transmission of diseases, this information could be extremely useful in public health studies.

Another desirable property of respondent-driven sampling is that sample data can be combined with institutional data to estimate the size of a hidden population (Heckathorn and Jeffri, 2001), which cannot be done with others research methods which I have discussed in the beginning of this paper.

From my point of view, given the critical importance to identify trafficked women, the network sampling (respondent-driven sampling) technique is important. Appropriate sampling approaches are at the core of any high-quality surveillance system, especially when the system is tracking hidden populations and when surveillance data are not interpretable or are inconsistent. I hope this work opens up new possibilities for researchers interested in carrying out substantive research on hidden populations.

**REFERENCES**

ACHARYA, A. (2006). *La esclavitud humana: el tráfico de mujeres en la India y México*, doctoral thesis, Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México (UNAM), Mexico.

ACHARYA, A. and A. SALAS STEVANATO, A. (2005). "Violencia y tráfico de mujeres en México: Una perspectiva de género", *Estudos Feministas*, Vol. 13, N. 3, Brazil.

AVICO, U., C.D. KAPLAN, D. KORCZAK AND K.M. VAN METER (1988). *Cocaine epidemiology in three European Community cities: A pilot study using a snowball sampling methodology*, Research report, Cocaine Steering Group, Health Directorate, Comisión of the European Communities, Brussels.

GONZÁLEZ, R. (2003). *Violencia contra las mujeres deja un millón de víctimas anuales en México*, CIMAC, Mexico.

HECKATHORN, D.D. (1997). "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations", *Social Problem*, Vol. 44, N. 2.

HECKATHORN, D.D. and JEFFRI, J. (2001). "Finding the Beat: Using Respondent-Driven Sampling to Study Jazz Musicians", *Poetics*, Vol. 28.

KANOUSE, D.E.; BERRY, S.H.; DUAN, N.; LEVER, J.; CARSON, S.; PERLAMN, J.F.; and LEVITAN, B. (1999). "Drawing a Probability Sample of Female Street Prostitutes in Los Angeles County", *Journal of Sex Research,* Vol. 36(1), pp. 45-51.

KOZEL, N.J. and ADAMS, E.H. (1986). "Epidemiology of Drug Abuse: An Overview", *Science*, 21, November.

MAGNANI, R.; SABIN, K.; SAIDEL, T. and HECKATHRON, D. (2005). "Review of Sampling Hard-to-Reach and Hidden Populations for HIV Surveillance", *AIDS*, 19 (Suppl. 2), ISSN 0269-9370.

MILLS, S.; SAIDEL, T.; BENNETT, A.; REHEL, T.; HOGLE, J.; and BROWN, T. (1998). "HIV Risk Behavior Surveillance: A Methodology for Monitoring Behavioral Trends", *AIDS Review*, Vol. 12 (Suppl. 2).

MUHIR, F.B.; LIN, L.S.; STUEVE, A.; MILLER, R.L.; FORD, W.L.; JOHNSON, W.D.; and SMITH, P.J. (2001). "A Venue-Based Method for Sampling Hard to Reach Populations", *Public Health Reports*, Vol. 116 (Suppl. 1).

PCOC (PRESIDENT'S COMMISSION ON ORGANIZED CRIME) (1986). *America's Habit: Drug Abuse, Drug Trafficking, and Organized Crime,* Washington, DC, US Government Printing Office.

SALGANIK, M.J. and HECKATHORN, D.D. (2004). Sampling and Estimation in Hidden Populations using Respondent-Driven Sampling, *Sociological Methodology*, Vol. 34, Issue 1.

SEMAAN, S.; LAUBY, J.; and LIEBMAN, J. (2002). "Street and Network Sampling Evaluation Studies of HIV Risk-Reduction Interventions", *AIDS Reviews,* Vol. 4, pp. 213-223.

STUEVE, A.; O'DONNELL, L.N.; DURAN, R.; DOVAL, A.S.; and BLOME, J. (2001). "Time-Space Sampling in Minority Communities: Results with Young

Latino Men Who Have Sex with Men", *American Journal of Public Health*, Vol. 91(6), pp. 922-926.

VAN METER, K.M. (1986). "Basic Typology and Multi-Method Analysis in the Social Sciences, New Delhi", *ISA World Congress*, India.

—— (1990). "Sampling and Cross-Classification Analysis in International Research, in Comparative Methodology: Theory and Practice in International Social Research", in Else Oyen (ed.), *Sage Studies in International Sociology,* ISBN 0-8039-8326-3.

WATTERS, J.K. and BIERNACKI, P. (1989). "Targeted Sampling: Options for the Study of Hidden Populations", *Social Problems*, Vol. 36(4), pp. 416-430.