

**Sampling and Surveying Hard-to-Reach Populations:  
A Study of Female Labor Migrants in Moscow, Russia**

Victor Agadjanian

Center for Population Dynamics  
Arizona State University (USA)

Natalya Zotova

Institute of Ethnology and Anthropology  
Russian Academy of Sciences (Russia)

Abstract

Because household-based survey designs are notoriously ineffective in studying hard-to-reach groups such as irregular migrants, these groups, however numerically large they may be, are rarely represented in demographic analyses. In this paper, we report on the application of a workplace-based stratified probability sampling design, response rate, and item-specific refusals in a study of irregular female migrants from Tajikistan, Kyrgyzstan, and Uzbekistan working in bazaars, eateries, and small retail outlets in Moscow, Russia. We argue that workplace-based survey, while not flawless, provides a uniquely feasible and cost-effective tool for studying irregular migrants and other marginalized groups.

## **1. Introduction**

Systematic sampling and data collection among hard-to-reach and “hidden” populations have always been a major challenge for social scientists. Marginalized groups such as homeless people, drug users, sex workers, or illegal immigrants are notoriously difficult to access in a systematic way (Faugier and Sargeant 1997; Marpsat and Firdion 1999; Valenzuela 2002). In most demographic studies, where representative household survey is the golden standard for data collection, such marginalized population segments are lost by definition or at best are grossly underrepresented. An additional set of problems concerns the willingness of individuals from marginalized hard-to-reach groups to provide information to survey interviewers, especially about the type of sensitive matters that define their marginality: it can be argued that the same social characteristics and constraints that hinder access to these individuals may also impair their willingness or ability to answer survey questions. This article reports on how these two types of challenges have been addressed in a survey of irregular international female labor migrants conducted in Moscow, Russia, in the autumn of 2010. We define these migrants as “irregular” as most of them are in violation of at least some of the convoluted and arcane laws and regulations that govern foreigners’ entry into, stay, and work in the Russian Federation. The main aims of the survey were to assess the feasibility of a stratified probability sampling approach tailored to the occupational distribution of this hard-to-reach population and to test a battery of sensitive questions such as those dealing with sexual partnerships and HIV/STD risks, as well as those concerning legal status, income, and living arrangements.

Large-scale labor migration from Central Asia to Russia is a historically recent phenomenon. In the early years after the collapse of the U.S.S.R., the nations that emerged from its ruins saw massive emigration of ethnic Russians and other ethnic groups of European origin primarily to

Russia. At the same time, political instability, especially in Tajikistan spurred sizeable flows of refugees. However, since the second half of the 1990s, migration from the region has been increasingly dominated by labor migrants from native ethnic groups, which in the Soviet era were characterized by very low mobility (Ergeshbayev 2006). The Russian Federation has always been the destination country of choice for the overwhelming majority of Central Asian labor migrants; among main Russia-bound migration streams, Tajik migration is the oldest, but economic hardships and endemic unemployment in Kyrgyzstan and Uzbekistan have fueled increased migration from those nations as well (Zayonchkovskaya 2009).

Russia's Central Federal Region, and especially the city of Moscow, Europe's largest metropolis with a population of more than 10 million, and surrounding it Moscow Oblast (region), is a major magnet for labor migrants from Central Asia (Dolotkeldieva 2010; Zayonchkovskaya 2009). According to the 2007 Federal Migration Service data, 35.3%, 27.6%, and 23.4% of legally registered migrants from Kyrgyzstan, Tajikistan, and Uzbekistan, respectively, worked in Moscow and Moscow Oblast (Zayonchkovskaya 2008: 34); the number of only Tajik migrants in Moscow was estimated at between 70,000 to 100,000 (Ivanova 2009)<sup>1</sup>. Although labor migration from Central Asia has been dominated by men, typically employed in the construction and home maintenance and improvement sectors, the numbers of female labor migrants in recent years has been on the rise (e.g., Khusenova 2010; Tajikistan 2010; Tyurukanova 2005; Zotova 2007). Yet there are no reliable statistics on the size and distribution of female Central Asian migrants in Moscow.

---

<sup>1</sup> According to Golunov (2008: 83), the share of Tajiks working in Moscow among all Tajik migrants in Russia is 49%.

Despite the large number of Central Asian migrants of both sexes in Moscow and their ubiquitous presence in the city's everyday life, drawing a systematic sample from this population involves considerable challenges. The residence-based sampling methods traditionally used for household surveys are not feasible for a number of reasons. First, many if not most foreign-origin migrants reside in overcrowded apartments (sometimes with twenty or even more tenants per two- or three-room unit) that are often rented or sub-rented illegally; outsiders' access to these apartments is difficult as most people would not even open the door to a stranger.<sup>2</sup> The majority of migrants hardly spend any time in their overcrowded dwellings during the day, even when they are not working. Second, the apartments that house migrants are scattered throughout the city; although migrants tend to settle in lower-cost areas in the city's periphery and outskirts, there is no clear pattern of migrants' residential concentration, especially by origin. Finally, a sizeable portion of migrants may reside (spend nights) at their work places and therefore could not be captured even through most sophisticated residence-based sampling designs.

## **2. Approaches to studying hard-to-reach populations**

The problems of unavailable sampling frame and unknown residential distribution are not unique to migrants in the Russian capital. These problems are pervasive in research on irregular migrants and other hard-to-reach populations. Common solutions for these problems have been convenience and "snowball" sampling (e.g., Chavez [1992] 1998; Ehrlich et al., 2007; Faugier and Sargeant 1997; Organista and Kubo 2005; van Meeteren 2010). Non-probability sampling approaches have been employed in research on migrants in Russia (Poletaev 2005) and China (Li

---

<sup>2</sup> The following conversation with a Kyrgyz migrant in the preparatory stage of our project is telling: Question: "The [Russian Population] Census will start soon [October 2010]. There are a lot of people living in your apartment. What will you do when the enumerators ring your door bell?" Answer: "We simply will not open the door, that's it. No one will open [the door] to the enumerators; we don't open to strangers."

et al., 2004). While cost-effective and generally suited for in-depth explorations in which statistical representativeness is not of import, these approaches almost invariably entails considerable selection bias, which precludes robust statistical inferences.

Respondent-driven sampling (RDS), a refined variety of chain-referral sampling first introduced by Heckathorn (1997; 2002), has shown considerable promise for systematic access to hard-to-reach and “hidden” populations. RDS has been employed to sample injecting drug users (Heckathorn 1997; Heckathorn et al. 2002; Wang et al. 2005), men having sex with men (Yeka et al. 2006); Magnani et al. 2005). However, RDS relies on stringent data assumptions and preconditions (Heckathorn 1997; Salganik and Heckathorn 2004) that limit its practical application. Initial participants (“seeds”) may vary in size of their personal networks and levels of networks homophily; setting up and tracing peer recruitment process may be costly and time-consuming, and weighting can be established only on the basis of respondents’ self-report (Marpsat and Razafindratsima 2010). In fact, recent simulations cast doubts on the accuracy of RDS, especially for epidemiological analyses (Goel and Salganik 2010). Moreover, in RDS there is an implicit assumption of solidarity among members of the target group based on the quality or qualities that distinguish them from the rest of the population (e.g., use of a particular drug), which is necessary to facilitate contacts and recruitment in referral chains. However, as our preliminary fieldwork in Moscow showed, many irregular female migrants, like many representatives of other marginalized groups who find themselves in similarly alienating and hostile environments, were unwilling or able to provide referrals to any other member of their group even for additional remuneration.

An alternative to RDS is the time-space or time-location probability sampling approach whereby study participants are recruited systematically at certain times of day and in certain

venues that they frequent regularly for their group-unique needs or characteristics. For example, time-space sampling has been used to sample homeless people accessing food and other services at homeless shelters and food/clothing distribution points (Ardilly and Le Blanc 2001; Koegel et al. 1996; Marspat and Firdion 2000). Valenzuela (2002) used time-space probability sampling in a study of day laborers in metropolitan Los Angeles. The researchers first compiled a list of all known day laborer hiring sites in the city. Then a random sample of day laborers waiting for work offers at those sites was carried out in a certain part of the day (usually in the early morning hours when hiring at those sites typically occur). Also in Los Angeles County, Kanouse et al. (1999) used over a hundred street areas that accounted for most of street prostitution to sample street sex workers. Their research teams cruised the selected areas at certain time shifts and in certain directions to enumerate all the street workers present; they then approached the women, one by one, for screening and possible interview. The procedure yielded almost a thousand interviews, most of which were conducted at about two-thirds of the selected sites.

Systematically sampling irregular migrants of different origins is only one of the challenges that a study focusing on this population segment in a city like Moscow faces. Response rate, and especially cooperation rate, is already a major problem in general surveys in urban Russia (Yausheva 2007). Response rate is particularly low in large cities: thus the response rate for the 2004 Generations and Gender Survey in Moscow and St. Petersburg was less than 15% (Houle and Shkolnikov 2005). These rates could be even lower among migrants, because of concerns and fears that this population might have due to its precarious legal and work status, incessant harassment by law enforcement authorities, and pervasive public xenophobia. Members of migrant ethnic groups that are targets of especially harsh treatment and of particular vicious stereotypes, such as darker-skinned Central Asians, may feel disproportionately reluctant to

engage in conversations with strangers. Finally, even if they agree to participate in survey interviews, Central Asian migrant women might be unwilling to answer sensitive questions dealing with their legal status, employment, and marital and sexual relationships and experiences due to a combination of legal and economic insecurities, gender ideologies, and ethno-religious norms.

### **3. Method**

The sampling and data collection procedures examined here were part of a study on sexual risks of female labor migrants from three Central Asian countries in Moscow (the analysis of sexual behavior and risks lies beyond this scope of this paper). Based on ethnographic exploration of employment niches in which female Central Asian migrants typically concentrate, the following three sectors of the city economy were selected: bazaars (*рынки*), eateries (restaurants and cafés), and small retail shops. This approach left out a potentially sizeable portion of migrants who work in other sectors, such as those in home-based personal care for the elderly and in private childcare.<sup>3</sup> We acknowledge this omission as a limitation of the study design. However, it is important to note that the study was focused largely on sexual risks of relatively young women (aged 18-40), while child and personal care occupations typically attract older women. Also, the study design excluded women who work in the sex and related industries; while HIV/STD risks of such women are relatively well known, these risks stem directly from the nature of their trade

---

<sup>3</sup> Anecdotal evidence suggests that even though Central Asian women expect lower wages as nannies than their Russian counterparts, some middle-class Russian families may be reluctant to hire them out of concern for Central Asians' poor knowledge of or accent in Russian that can affect the children's learning and command of the language.

and therefore are subject to different constraints and mechanisms than the risks of women whose work does not explicitly involve provision of sexual services.

The multi-stage stratified sample was designed to reach a total target number of 225 respondents, including equal numbers of respondents from the three sectors and three ethnic groups which account for the overwhelming majority of migrants from the Central Asian region—Kyrgyz, Tajiks, and Uzbeks. It should be mentioned that while each of these ethnic groups is the titular ethnicity in its respective eponymous nation, members of these groups reside throughout Central Asia and often make up substantial non-titular minorities (e.g., Tajiks in Uzbekistan, Uzbeks in Kyrgyzstan). The sampling approach based on ethnicity rather than provenance was justified from the substantive standpoint as one of the interests of the study was to examine ethnocultural variations in the sexual and other outcomes. The target sample therefore was to include nine ethnicity\*sector cells—25 women in each.

In devising the multi-stage sampling strategy, we assumed that the distribution of migrant workers' of the target ethnic groups in all three sampling domains—bazaars, eateries, and retail shops—as well as their spatial distribution within each of these sectors should generally be as close to random as their residential distribution in the Moscow area. The sampling approach differed between bazaars, on the one hand, and cafés/restaurants and shops, on the other, due to the nature and organization of these sectors of urban economy. For the bazaar segment of the sample, at the first stage, we compiled a list of forty Moscow bazaars (excluding very small bazaars and bazaars that specialize primarily in construction equipment and materials, where women rarely work).<sup>4</sup> To ensure that the three ethnic teams of interviewers do not overlap, we

---

<sup>4</sup> Although foodstuffs sections or rows (*продуктовые отделы* or *ряды*) are primary features of almost any bazaar, many if not most of the bazaars also have sizeable sections selling clothing, shoes, and other manufactured goods



randomly selected five bazaars from this list —two large- and three medium- or smaller-size— for each ethnic group. In each bazaar, a standard random walk algorithm (every fifth stall) was used to identify five women of the given ethnicity. For Kyrgyz and Tajiks, the five original bazaars yielded the target number of respondents; for Uzbek women, one more bazaar had to be added to complete the target 25 interviews.

The above two-stage approach in identifying and selecting spatial clusters and selecting individuals within them could not be applied to eateries and shops: canvassing all such businesses in a city of more than ten million was, of course, out of the question. Enumerating them fully on the basis of telephone directories or available registries proved impracticable as these directories and registries are incomplete and outdated. Selecting territorial clusters within the city (e.g., administrative subdivisions or electoral precincts) was also ruled out as such businesses are located very unevenly throughout these clusters. After considering and rejecting the above options, we decided to use the Moscow subway (metro) stations (see the metro map in Figure 1) as the basis for sampling. Much of the city’s economic and social life is pivoted on its subway stations (182 in total, as of the end of 2010), and the types of small businesses that we wanted to sample tend to gravitate to them. For each of the three ethnic groups, we selected two non-neighboring radial metro lines (by matching the lines’ map colors with the colors of the three Central Asian nations’ flags): on these lines, we randomly selected four “peripheral” stations (outside the area delimited by the circular line) and three “central” stations (within the circle).

---

(*вещевые ряды*). The food and “things” stalls are often intermixed, and our sampling procedure did not differentiate between food and non-food sections.

Figure 1 around here

The areas surrounding the selected stations constituted the primary sampling units. All eateries and all small retail shops (excluding large supermarkets or department stores) within the radius of 500 meters of each peripheral station and 300 meters of each central station were canvassed and recorded into two randomly ordered separate lists (restaurants/cafés and retail shops). These businesses made up the secondary sampling units. The interviewers received the maps of the areas on which each establishment was identified with a number from the corresponding list; they visited these establishments during normal business hours in the order indicated on the map looking for respondents with qualifying characteristics (ethnicity and age). If an establishment had several eligible women, one of them was selected randomly (the alphabetical order of the third letter of the first name in the Russian alphabet).

Female interviewers of matched ethnicity asked the selected qualifying woman whether she would be interested in participating in a survey. If the woman agreed, the interviewer would make an arrangement with her about the place and time of the interview. Before starting the interview with a woman who expressed interest, the interviewer would read to the respondent the standard consent letter describing the study, potential risks and benefits, protection of confidentiality, and the recompense (200 rubles, equivalent to roughly 7 USD at that time). If the selected woman refused to participate in the study at the initial contact or did not grant her consent, the interviewer would thank the woman and moved on to the next establishment on her list. The interview would proceed only after consent. The 25 pages-long survey instrument included seven modules and took on average 54 minutes to complete. The study design was approved by the Institutional Review Boards (ethics councils) of the co-authors' respective

institutions. The sampling and interviews were monitored throughout the fieldwork; upon completion of the fieldwork, the interviewers were debriefed on their experience of data collection and specifically on difficulties encountered in the process.

In the remainder of the paper we discuss the response and cooperation rates, choice of time and place for interview by study participants, refusal to answer specific questions, and participants' willingness to provide personal contacts for possible follow-up interviews. We conclude with an assessment of the study methodology with an emphasis on its potential usefulness for demographic research.

## **4. Results**

### **4.1 Response rate**

Although the nature of the chosen sampling design precludes calculation of conventional response rate (number of completed interviews divided by the total number of sampled units), the effectiveness of this design can be assessed in two ways: number of workplace units needed to reach the target number of eligible individuals and agreement of sampled individuals to participate in the survey (cooperation rate).

Based on prior field observations, we estimated that seven stations would be enough to fill the target number of restaurant/café workers of each ethnic group (25). This estimate proved correct for Tajiks and Kyrgyz; for Uzbeks, two additional stations were added. Sampling in retail outlets proved more problematic as those do not gravitate to metro stations to the same extent and do not have as many migrant workers as do eateries. As a result, four additional stations had to be added to reach the target number of Kyrgyz, 10 for Uzbeks, and 15 for Tajiks. In sum, it took interviewers who were looking for Kyrgyz migrants to visit 100 restaurants/café and 89

retail outlets to conduct the target numbers of interviews; in comparison, 137 restaurants/cafés and 118 retail outlets had to be visited to find the target numbers of Uzbek respondents and 120 restaurants/cafés and 153 retail outlets were needed to reach the target sample of Tajiks. These statistics are summarized in Table 1.

Table 1 about here

*Cooperation* rate (number of those who agreed to participate divided by the number of eligible individuals who were contacted) is a particular important measure of success in studies of hard-to-reach and “hidden” populations. Table 2 presents cooperation rates obtained for each sector and ethnic group. As can be seen, the bazaars registered the largest number of refusals. We attribute this to three main factors. First, according to a Russian government resolution initially adopted in 2006 and extended in the following years, only citizens of the Russian Federation are authorized to sell in bazaars (Russian Federation 2006). Bazaar vendors without Russian citizenship therefore might feel especially apprehensive of any unusual interaction with strangers.<sup>5</sup> Second, the setting of a bazaar is expressively public: the vendors are—or feel to be—under constant watch by fellow vendors from neighboring stalls as well as by customers; this continuous informal surveillance may have served as an additional deterrent. And third, the fieldwork started in the bazaars and only after the bazaar sub-sample target had been reached proceeded to eateries and retail shops; the interviewers were simply not as experienced in engaging potential respondents at that early stage of the project (as they admitted in the

---

<sup>5</sup> Despite the law restricting vending in bazaars to citizens of the Russian Federation, the majority of sellers in Moscow bazaars lack Russian citizenship. In our survey sample, only one-fifth of the bazaar subsample said to be Russian citizens.

debriefing interviews). Both restaurants/cafés and shops registered much higher cooperation rates. In fact, in shops, most of which were covered in the late phase of the survey, the overall cooperation rate was nearly 100%.

Table 2 about here

Table 2 also shows that cooperation rate varied noticeably across ethnic groups. Because the three ethnic teams of interviewers worked in parallel, the differences cannot be attributed to cumulative experiences of interviewers (although differences in skills and effort across interviewers cannot be ruled out). Most remarkably, Kyrgyz women showed a much higher cooperation rate in bazaars (77% compared to 55% among both Tajiks and Uzbeks). The ethnic pattern was less obvious in the other two sectors; in fact, both Kyrgyz and Uzbek women working in eateries showed lower cooperation rates than did Tajik women. While a definitive explanation of these patterns is difficult, we suggest that they are reflective of the groups' ethnocultural backgrounds and migration experiences. In the more public environment of bazaars, the greater level of cultural flexibility and openness of Kyrgyz women, who are generally less religious, more Russified, and come from a less rigid gender-ideological background, may have played a role. In contrast, in less publicly scrutinized settings, the traditional cultural background may matter less as more educated (i.e., less traditional) women are selected into those sectors' occupations. For example, 50% of Tajiks and 17% of Uzbeks working in shops had complete higher education compared to 30% and 7%, respectively, among bazaar workers. Interestingly, no similar selection on education was present in the Kyrgyz subgroup, the newest segment of Central Asian female migration: in both the shops and the

bazaars, 26% of Kyrgyz respondents were university educated. Although the educational distribution of the women who agreed to participate in the survey cannot be readily extrapolated to those who did not, it seems reasonable to assume that educational variations by ethnicity among the latter were comparable.

#### **4.2 Choice of place and time of interview**

While the workplace provides the best, if not the only, systematic access to a hard-to-reach working population like the one targeted in this study, it may not offer an ideal venue for conducting the interview because individuals may be both busy at work and feel under close watch of their supervisors, co-workers or fellow vendors. Yet, nine out of ten respondents chose to be interviewed at their workplaces and at the time of first contact. In fact, during the pretest we tried to insist on scheduling the interview on a different day, but this generated a large number of subsequent refusals (women who had agreed to be interviewed at the first contact would refuse at a later phone contact, usually saying that they did not have time).

#### **4.3 Refusal to answer specific questions**

The analysis of responses to specific items showed that the questions that we had expected to be most sensitive—on legal status, details of sexual behavior (number and characteristics of partners; frequency of vaginal, anal, and oral intercourse), and income—did not generate the anticipated refusals, even though in the absence of any comparable data it is impossible to judge to what extent respondents were truthful in answering those questions. Interestingly, according to the interviewers' debriefing, the questions that made some respondents uncomfortable were those dealing with people who shared their residence (the survey asked about relation to the

respondent, sex, age, place of birth/provenance, education, work status of every person living in the respondent's residence and whether that person spend the previous night at that residence) . As migrants typically share living quarters with a lot of strangers, they felt particularly reluctant to talk about what they saw as other people's lives (if they knew much about those lives at all). The burden of supplying information about a large number of people typically sharing living quarters with respondents was probably also a factor: in a few cases, respondents refused to continue after providing details on only several of their co-residents. In another case, the perceived burden resulted in a refusal to answer questions on own migration history. In all, however, the number of refusals to answer specific questions was negligible.

#### **4.4. Providing personal contact information for a follow-up interview**

At the end of the interview, all the respondents were asked to provide a contact (usually a mobile phone number) in case researchers needed to conduct a follow-up interview for an additional monetary reward. Just over one-third of them, 35 percent, agreed to supply their contact information. Table 3 presents the results of a binomial logistic regression model that predicts the likelihood of having agreed to provide personal contact information from respondent's ethnicity, occupational sector, age, partnership status, education, income, and legal status. As the results indicate, ethnicity and sector of urban economy were the only two characteristics that had statistically significant effects on the outcome. Thus, women working in restaurants and cafés and in retail outlets were significantly less likely to provide personal contacts for subsequent interviews than were bazaars women. We believe that this difference is due, at least to some extent, to the selective cooperation rate described above: those bazaar women who had agreed to

participate in the survey felt rather comfortable with the idea of being interviewed and therefore did not mind to do it again, especially if it involved additional recompense.

Table 3 about here

Tajik women, all other things being equal, were more willing to share their personal contact information than were Uzbek and especially Kyrgyz women (the difference between Uzbeks and Kyrgyz was not statistically significant). This ethnic pattern requires further investigation. Most instructively, however, legal status did not have any effect on the likelihood of agreeing to provide personal contact information: the sign of the regression coefficient suggests that women with a more secure legal status (those who held the citizenship of the Russian Federation or had permanent or temporary residence permits) were less likely to supply contact information, but the coefficient is not significantly different from zero. Income, education, and partnership status were also irrelevant to the respondents' willingness to provide personal contact details.

## **5. Conclusion**

Hard-to-reach, marginal groups often fall through the cracks of large demographic studies, which almost invariably employ residence-based surveys for data collection. Our study shows that workplace-based systematic sampling approaches can offer a viable, cost-effective alternative to traditional household survey sampling designs when a sampling frame is not available, especially in large metropolitan areas such as Moscow where response rates even in general household-based surveys are below 15% and where, as the 2004 GGS experience shows, it takes on average seven visits to complete a survey interview (Haule and Shkolnikov 2005). Like RDS and time-



location sampling approaches, the approach used in our study does not yield a fully representative sample of the target population (Wang et al., 2005). Yet it does provide a powerful and relatively inexpensive tool to systematically access, sample, and survey irregular labor migrants and probably other similarly marginal groups that is far superior to approaches of the convenience sampling family. This approach is also more efficient than RDS as it does not involve a lengthy multi-stage referral process, which can also be quite costly because subjects are paid both to participate in the study and to recruit others. Importantly, unlike relatively small hard-to-reach groups such as the homeless, drug addicts, or high sexual risk groups which are typically of interest to epidemiologists or substance abuse researchers (e.g., Heckathorn et al. 2002; Kanouse et al. 1999; MacKellar et al. 2007; Magnani et al. 2005; Malekinejad et al. 2008; Pollack et al. 2005; Wang et al. 2005), the target group of our study represents a fairly large and growing segment of the city population, which is well familiar to almost every city dweller but remains “invisible” to demographers and other social scientists relying on household survey data.

Our experience also suggests that the marginality of irregular migrants and similar groups does not translate to a particularly low response rate. While the overall response rate between residence-based and work-based surveys cannot be directly compared, the cooperation rate achieved in our study is comparable to those typically recorded in household surveys of general population in urban settings (e.g., Heeringa 1997) or other types of time-location sampling (e.g., Kanouse et al. 1999). True, the cooperation rate in our study varied across the three types of work sites and across the three ethnic groups; for most of these variations we believe that we have found plausible, even if tentative, explanations. However, it is also important to recognize that the relatively high share of refusals in bazaars was compensated by a large concentration of and easy access to eligible respondents there; in contrast, while finding eligible respondents in

eateries and shops took more time and effort, the cooperation rate in those two settings was generally much higher than in bazaars. Although we cannot directly assess the influence of legal status on agreement to participate in the survey, it is noteworthy that legality did not significantly affect respondents' willingness to provide contact information for possible subsequent interviews. Finally, and perhaps most reassuringly, we observed no refusals to answer such sensitive questions as those dealing with income, details of sex life, and legal status.

This project was but a first step toward designing and fielding a large-scale, multi-site study of female migrants' vulnerabilities and risks. While the sampling design of the larger study may be adjusted to balance specific research questions with available resources, workplace-based sampling will be the most feasible and cost-effective sampling approach.

## References

- Ardilly, P., and Le Blanc, D. (2001). Sampling and weighting a survey of homeless persons: A French example. *Survey Methodology* 27(1): 109-118.
- Chavez, L.R., F.A. Hubbell, S.I. Mishra, and R.B. Valdez. 1997. Undocumented Latina immigrants in Orange County, California: A Comparative analysis. *International Migration Review* 31(1): 88-107.
- Dolotkeldieva, A. 2010. Les migrants kirghizes à Moscou : politiques publiques, stratégies migratoires et réseaux associatifs. In : Laruelle, M. (ed.) *Dynamiques migratoires et changements sociétaux en Asie centrale*. Editions PETRA: Paris.
- Ehrlich, S. F., Kurt C. Organista, Doug Oman. 2007. Migrant Latino day laborers and intentions to test for HIV. *AIDS and Behavior* 11(5): 743-752.
- Ergeshbayev, U. 2006. Trends and development of migration processes in Kyrgyzstan. In *Migration Perspectives Eastern Europe and Central Asia. Planning and Managing Labour Migration*. Ed. by R.R. Rios. Vienna: International Organization for Migration, 45-54.
- Faugier, J., and M. Sargeant. 1997. Sampling hard to reach population. *Journal of Advanced Nursing* 26(4): 790–797.
- Goel, S. and Salganik, M.J. 2010. Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences* 107(15): 6743–6747.
- Golunov, S.V. 2008. *Regional Dimensions of Transborder Migration to Russia*. Moscow: Aspect Press. (in Russian).
- Heckathorn, D. D. 1997. Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems* 44(2): 174-199.

- Heckathorn, D. D. 2002. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems* 49(1): 11-34
- Heckathorn, D. D., Semaan, S., Broadhead, R.S., and Hughes, J.J., 2002. Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18–25. *AIDS & Behavior* 6: 55–67.
- Heeringa, S. G. *Russia Longitudinal Monitoring Survey (RLMS): Sample Attrition, Replenishment, and Weighting in Rounds V-VII*. <http://www.cpc.unc.edu/projects/rlms-hse/project/samprep/samprep.pdf> (accessed in November 2011).
- Houle, R. and V. Shkolnikov. (2005). *Low response rates in the cities of Moscow and Sankt-Petersburg and GGS-Census comparisons of basic distributions*. Max Planck Institute for Demographic Research: Rostock, Germany.
- Ivanova, T.D. 2009 “Tajiks in Moscow society” In Zayonchkovskaya, Zh.A.(ed.) *Immigrants in Moscow*. Moscow: Tri Kvadrata (in Russian).
- Kanouse, D. E., S.H. Berry, and N. Duan. 1999. Drawing a probability sample of female prostitutes in Los Angeles County. *Journal of Sex Research* 36(1): 45-51.
- Khusenova, N. 2010. La féminisation des migrations de travail tadjikes en Russie. In : Laruelle, M. (ed.) *Dynamiques migratoires et changements sociétaux en Asie centrale*. Editions PETRA: Paris.
- Koegel, P., Burnam, A., and Morton, S. 1996. Enumerating homeless people: Alternative strategies and their consequences. *Evaluation Review* 20: 378-404.
- Li X., Stanton B., Fang X., Lin D., Mao R., Wang J., Cottrell L., and Harris C. 2004. HIV/STD risk behaviors and perceptions among rural-to-urban migrants in China. *AIDS Education and Prevention* 16(6): 538-556.

- MacKellar, D.A., Gallagher K.M., Finlayson, T. Sanchez, T., Lansky, A. and Sullivan P.S. 2007. Surveillance of HIV risk and prevention behaviors of men who have sex with men – A national application of venue-based, time-space sampling. *Public Health Reports* 122 (Suppl. 1): 39-47.
- Magnani, R., K. Sabin, T. Saidel, and D. D. Heckathorn. 2005. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 19 (Suppl. 2): S67-S72
- Malekinejad, M., Johnston, L.G., Kendall, C., Franco, L.R., Kerr, S., Rifkin, M.R., and Rutherford, G.W. 2008. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: A systematic review. *AIDS and Behavior* 12: S105–S130.
- Marpsat, M., and Firdion, J-M. (1999). The homeless in Paris: a representative sample survey of users of services for the homeless. In D. Avramov (editor) *Coping with Homelessness: Issues to be Tackled and Best Practices in Europe*. Aldershot, Brookfield, USA, Singapore, and Sydney: Ashgate Publishing.
- Marpsat, M., and N. Razafindratsima. 2010. Survey methods for hard-to-reach populations: Introduction to the special issue. *Methodological Innovations Online* 5(2): 3-16.
- Organista, K. C., and Kubo, A. 2005. Pilot survey of HIV risk and contextual problems and issues In Mexican/Latino migrant day laborers. *Journal of Immigrant Health* 7(4): 269-281
- Poletaev, D. 2005. Women-migrants from foreign countries in Russia. *Vestnik Evrazii* 1: 18-31 (in Russian).
- Pollack L.M., Osmond D.H., Paul J.P. and Catania, J.A. (2005). Evaluation of the Center for Disease Control and Prevention’s HIV behavioral surveillance of men who have sex with men: Sampling issues. *Sexually Transmitted Diseases* 32(9): 581-589.

- Russian Federation. 2006. On setting an acceptable share of foreign workers in the area of retail commerce on the territory of the Russian Federation for 2007” Resolution of the Government of the Russian Federation no.683, 15 November, <http://www.rg.ru/2006/11/16/kvota1-doc.html> (accessed in November 2011).
- Salganik, M. J. and D. D. Heckathorn. 2004. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology* 34:193-239.
- Tajikistan. 2010. *Tajikistan Foundation: Analytical Report*. Moscow: Tajikistan Foundation (in Russian).
- Tyurkanova, E. V. 2005. Gender aspects of labor migration from CIS countries to Russia.” *Diaspori* 1: 48-64 (in Russian).
- Valenzuela, A. Jr. 2002. Working on the margins in metropolitan Los Angeles: Immigrants in day-labor work. *Migraciones Internacionales* 1 (2): 6-28.
- van Meeteren, M. 2010. *Life without Papers: Aspirations, Incorporation and Transnational Activities of Irregular Migrants in the Low Countries*. Rotterdam: Erasmus Universiteit.
- Yausheva, A.A. 2007. Building a sample: In search of new possibilities. *Sotsiologicheskie Issledovaniia* 8 (August): 90-96 (in Russian).
- Wang, J., R. G. Carlson, R. S. Falck, H. A. Siegal, A. Rahman, and L. Li. 2005. Respondent-driven sampling to recruit MDMA users: A methodological assessment” *Drug and Alcohol Dependence* 78: 147-157.
- Zayonchkovskaya, Zh.A. 2008. Population migration from Central Asia. Pp. 24-34 in E.V. Abdullayev (ed.) *Labor Migration in the Republic of Uzbekistan: Social, Legal, and Gender Aspects*. Tashkent: UNDP and Gender Program of Swiss Embassy in Uzbekistan (in Russian).

Zayonchkovskaya, Zh.A.(ed.). 2009. *Immigrants in Moscow*. Moscow: Tri Kvadrata (in Russian).

Zotova, Natalya A. 2007. Women labor migrants from Tajikistan (older age group)” *Vestnik Evrazii* 2 (36): 72-88 (in Russian).

Yeka, W., Maibani–Michie, G., Prybylski, D., and Colby, D. 2006. Application of respondent driven sampling to collect baseline data on FSWs and MSM for HIV risk reduction interventions in two urban centres in Papua New Guinea. *Journal of Urban Health* 83(6 Suppl.): i60-72.

Table 1. Survey sample statistics for eateries and retail shops

Sector and Ethnicity	Number of metro stations initially selected	Number of metro stations added	Number of establishments identified	Number of establishments visited
<i>Eateries</i>				
Kyrgyz	7	0	140	100
Uzbek	7	2	165	137
Tajik	7	0	137	120
Overall	21	2	442	357
<i>Shops</i>				
Kyrgyz	7	4	107	89
Uzbek	7	10	118	118
Tajik	7	15	153	153
Overall	21	29	378	360



Table 2. Cooperation rate by sector and ethnicity

Sector and Ethnicity	Number of those contacted	Number of those who agreed to participate	Cooperation rate (number agreed/number contacted*100)
<i>Bazaars</i>			
Kyrgyz	31	24	77%
Uzbek	44	24	55%
Tajik	47	26	55%
Overall	122	74	61%
<i>Eateries</i>			
Kyrgyz	33	25	76%
Uzbek	32	25	78%
Tajik	27	25	92%
Overall	92	75	81%
<i>Shops</i>			
Kyrgyz	25	25	100%
Uzbek	27	25	93%
Tajik	25	25	100%
Overall	77	75	97%
<i>Total sample</i>	<i>291</i>	<i>224</i>	<i>77%</i>

Table 3. Logistic regression predicting provision of personal contact information upon survey completion

Covariate	Coefficient (standard error)
Retail	-1.144 (0.422) **
Eatery	-1.003 (0.421) *
Bazaar (reference)	0
Kyrgyz	-2.376 (0.467) **
Uzbek	-1.659 (0.402) **
Tajik (reference)	0
Russian citizen or legal resident	-0.642 (0.444)
Neither citizen no legal resident (reference)	0
Age	-0.003 (0.029)
Lives with a married or unmarried partner	0.129 (0.331)
Does not live with a partner (reference)	0
Higher education	-0.147 (0.392)
Less than higher education (reference)	0
Number of years lived in Moscow	0.107 (0.070)
Monthly income, logged	-0.733 (0.515)
Number of observations	224

*Note:* \* significant at  $p < 0.5$ ; \*\* significant at  $p < 0.01$ .

Figure 1. Moscow metro map

