HIV infection risk and condom use among sex workers in Senegal: Evidence from the list experiment method

Aurélia Lépine*[1], Carole Treibich[2], Khady Gueye[3], Cheikh Tidiane Ndour[3], and Peter Vickerman[4]

[1]University College London
[2]University of Grenobles-Alpes
[3]Ministry of Health of Senegal, HIV/STI division
[4]University of Bristol, Bristol Medical School

Abstract

Background: Social desirability bias, which is the tendency to under-report socially undesirable health behaviours, significantly distorts information on sensitive behaviours that is gained from self-reports. As a result, self-reported condom use among high-risk populations is thought to be systematically over-reported, and it is impossible to identify the determinants of condom use.

Objective: The main objective of the paper is to elicit unbiased information on condom use among FSWs using the double list experiment method to analyse the role of HIV infection and exposure to HIV prevention methods in condom use. More specifically, we estimate whether condom use differs between HIV-positive and HIV-negative FSWs. In addition, we estimate the role of FSWs' registration and participation in a pre-exposure prophylaxis demonstration project in condom use.

Method: We designed a list experiment to elicit condom use information from 786 FSWs in Senegal who were surveyed in 2015 and 2017. Using the list experiment method, participants were randomly assigned to one of two groups (treatment or control) and were asked to report the number of statements they agreed with. Respondents assigned to the control group were presented with three non-sensitive items, while those allocated to the treatment group were presented with the same three statements plus the sensitive item (e.g., "I used a condom during my last intercourse with a client"). Comparing the average number of sentences that were agreed with in both groups provides an estimation of the condom use rate in the treatment group, and estimating such prevalence for several sub-groups allows us to identify the role of HIV infection risk in condom use.

Results: We found that the percentage of FSWs using condoms in their last sexual intercourse with a client was 80% in 2015 and 78% in 2017, which was significantly lower

*Corresponding author

than the 97% obtained in the face-to-face surveys in both waves. When estimating condom use among sub-groups with the list experiment method, we found that condom use among HIV-positive FSWs was only 34%, which was 47 percentage points lower than the condom use among HIV-negative FSWs. We also found that registered FSWs are more likely to use condoms than clandestine FSWs are. However, we did not find any difference in condom use between FSWs who were enrolled in the pre-exposure prophylaxis (PrEP) demonstration project and those who were not enrolled.

Conclusion: Health policies should aim to increase condom use among HIV-positive FSWs.

1 Introduction

Condom use is the main preventive tool available for limiting the spread of sexually transmitted infections (STI), including HIV. Given that the consistent use of condoms is known to be the most cost-effective way to prevent HIV transmission (Cohen et al., 2004; Creese et al., 2002; Mitchell et al., 2015), condom use is the cornerstone of any HIV prevention strategy in most countries. Promoting condom use often involves multiple interventions, such as awareness campaigns and the provision of free condoms. However, the evaluation of the effects of such policies is made problematic by the impossibility of directly observing the targeted groups' levels of condom use. While the use of prostate-specific antigen is viewed as a gold standard biomarker for recent unprotected vaginal intercourse, its high financial, ethical and logistical cost is a barrier to its introduction in behavioural surveys. In addition, the use of prostate-specific antigen cannot reveal whether an unprotected act occurred with a client or with a non-commercial sexual partner. For these reasons, most behavioural surveys conducted to estimate the impact of condom promotion interventions are based on self-reported condom use (see Foss et al. (2007) for a review).

A common feature in surveys of female sex workers (FSWs) is a very high level of self-reported condom use (Treibich and Lépine, 2019). However, such safe behaviours are not consistent with the high prevalence of HIV and other STIs measured in FSWs (Dureau et al., 2016). Underestimating condom use may be even more of a concern when considering stigmatised groups that are highly targeted by preventive services, such as FSWs. Senegal is a Muslim-dominated country, and sex outside marriage is forbidden in Islam. As a result, sex workers in Senegal confront a variety of stigmatising discourses (Foley, 2017). Given their experiences with discrimination and stigma, social desirability bias may be a prominent issue when collecting information on their sexual behaviours in an open survey. This observation raises the question of whether the direct elicitation of condom use information, such as through face-to-

face interviews, can provide an accurate estimates of condom use and can therefore be used to measure the impact of condom-based interventions. This concern is supported by evidence that self-reported levels of condom use among FSWs are poorly associated with prostate-specific antigen (Aho et al., 2010; Gallo et al., 2007; Liu et al., 2016; Weir et al., 1999). Thus far, only the polling box method has been used to overcome the biased reporting of condom use among FSWs (Hanck et al., 2008). Our paper provides new evidence on the role of HIV status and different public health interventions in condom use in Senegal, a country where FSWs are up to 9 times more likely to be infected with HIV/AIDS and where there is an HIV/AIDS prevalence of 6.6% in 2015 (APAPS and IRESSEF, 2015).

Senegal is a particularly interesting country for the study of the link between HIV prevention strategies and condom use among FSWs. First, sex work in Senegal is regulated by a public health intervention where FSWs are required to register with a healthcare centre and be regularly screened for STIs, including HIV/AIDS. Our survey waves took place in June 2015 and August 2017 and we collected the HIV status of registered FSWs from these medical records. In addition, a demonstration project recruited 200 FSWs in Dakar between July 2015 and December 2016 to evaluate the feasibility of pre-exposure prophylaxis (PrEP) amongst FSWs in Senegal. PrEP consists of giving low doses of antiretroviral drugs to HIV-negative individuals to reduce their risk of being infected with HIV (Donnell et al., 2010; Grant et al., 2010; Karim et al., 2010). Daily oral PrEP is currently recommended by the World Health Organization (WHO) as an additional prevention choice for all population groups at substantial risk of HIV infection (WHO, 2015). The introduction of this new HIV prevention strategy for high-risk groups could help to eradicate HIV/AIDS epidemics in low-incidence contexts, such as Senegal. Evidence from PrEP trials shows that PrEP has high efficacy if taken consistently. However, there are concerns that it may result in risk compensation, i.e., an increase in risky behaviours resulting from interventions that reduce the perceived risk of infection (Blumenthal and Haubrich, 2014). Indeed, there is some apprehension from civil society organisations in Senegal that PrEP could wipe out decades of condom promotion campaigns targeting FSWs.

Our paper reports condom use estimations based on different designs of an indirect elicitation method in a two-wave survey among FSWs in Senegal: the list experiment method (Wave 1 in 2015) and the double list experiment method (Wave 2 in 2017). In particular, in 2015, participants were randomly assigned to two groups ("treatment" or "control"). The control group was asked to state with how many non-sensitive items they agreed with and the treatment group was asked the same set of questions with the addition of a sensitive statement asking whether they had used a condom during their last sex act with a client. This methodology was extended in the second wave of the survey through the use of a double list experiment design (Droitcour et al., 1991), i.e., the use of two different lists of non-sensitive items where respondents served

sequentially as treatment and control groups (or vice versa) to increase the estimate's precision.

The list experiment method has been extensively used for surveys to elicit information on vote preferences (Gonzalez-Ocantos et al., 2012; Holbrook and Krosnick, 2010), illegal migration (McKenzie and Siegel, 2013), and the use of microfinance loans (Karlan and Zinman, 2012) as well as opinions on topics such as same sex marriage (Lax et al., 2016), racism (Krumpal, 2013), abortion (Bell and Bishai, 2019; Ghofrani et al., 2018; Moseson et al., 2017a,b,c) and female genital cutting (De Cao and Lutz, 2018; Gibson et al., 2018). Previous studies that applied the list experiment method to estimate condom use concluded that condom use was overestimated by 11 points among college students in the United States (LaBrie and Earleywine, 2000) and by 14 points among young men in Uganda, but condom use was neither overestimated among young women (Jamison et al., 2013) nor among teenagers in Colombia (Chong et al., 2013). However, despite the increasing popularity of the list experiment to elicit sensitive behaviours, the method can fail. In addition, a growing number of studies have showed that the list experiment method sometimes produces unreasonable estimates of sensitive behaviours (Bell and Bishai, 2019; Chuang et al., 2019; Haber et al., 2018; Kramon and Weghorst, 2019). Even a larger prevalence of the sensitive behaviour determined by this method compared to direct reports does not prove that the estimated prevalence is the correct one. In addition, the list experiment method works by adding random noise to the data, which can increase standard errors, creating a trade-off between validity and efficiency (Blair et al., 2018), and might lead to the list experiment method being applied to samples that are too small to be useful (Blair et al., 2018).

Using the list experiment method, we further investigated the determinants of condom use. We focused on HIV status and HIV prevention strategies (i.e., the FSW registration policy, provision of free condoms and PrEP demonstration project). More precisely, we linked information from medical records for registered FSWs collected in both survey waves with the list experiment results to compare condom use for HIV-positive and HIV-negative FSWs. We further investigated the role of two main HIV prevention strategies in condom use: sex work registration and PrEP. We investigated whether past participation in the PrEP demonstration project was associated with lower condom use. To do so, we made use of the fact that the second wave of our survey took place seven months after the end of the PrEP demonstration project targeting FSWs in the Dakar region and that our sample included roughly 60% of all the FSWs who participated in the PrEP demonstration project (115 out of 200). Although we cannot investigate whether current PrEP use is associated with lower condom use, as PrEP was not made available to the participants at the end of the demonstration project, we can investigate whether the use of PrEP over a 15-month period led to a long-term decrease in condom use.

2 Sample and descriptive statistics

The participants were all FSWs working in Dakar, with the sample being stratified by registration status (registered versus non-registered FSWs). Registered FSWs were recruited using medical records from four (of the five) STI centres located in the suburb of Dakar (Rufisque, Pikine, Mbao, and Sebikotane), while non-registered sex workers were recruited through the leaders of sex worker groups and NGO staff. All FSWs were asked to come to the healthcare centre, where they were interviewed in dedicated private rooms. We randomised the allocation of participants to the treatment or control group based on their "arrival" number[1]. Each interview lasted 1.5 hours on average and aimed to collect socio-economic, behavioural and psychological information. The response rate was close to 100% in the population of registered and non-registered FSWs. In addition, HIV status was collected from the medical records of registered FSWs in both survey waves. Ethical clearance was obtained from the London School of Hygiene & Tropical Medicine ethics committee and the national ethics committee in Senegal (reference numbers SEN15/15 and SEN17/24), and written consent was obtained from participants.

Wave 1 took place in June and July 2015. At that time, we collected information from 651 FSWs, and this included their phone number and address as well as consent to be contacted again in future studies. In August 2017, roughly two years after Wave 1, we tried to follow-up with all participants who participated in Wave 1. We first attempted to contact each FSW by phone, and the phone call was made by the leaders of sex worker groups in the case of non-registered sex workers and by midwives in the case of registered sex workers. If there was no answer or if the phone number was no longer valid, our team investigated the cause using information from the sex worker group leaders and their social network. For those who were known to be in Dakar, we then asked group leaders to go to the physical address of the FSW. We were able to reinterview 440 sex workers (67% of the participants of Wave 1) out of which 62 respondents had quit sex work and were thus not asked about their last paid sexual intercourse. We tested whether the participants whom we lost contact with were different from those who remained in the survey. The supplementary file S4 shows average characteristics for two subsets: sex workers who were lost to follow up (Wave 1 only) and those who remained in the second round (Waves 1 & 2). Looking at this table, we can see that sex workers who were lost to follow up did not seem to be different than those who were able to participate in Wave 2 for most of the individual characteristics.

[1]Each enumerator had to interview approximately 50 FSWs per survey wave. The "arrival" number refers to the enumerator's ranking sheet. Odd numbers were first allocated to the control group in list A (in both waves) and then to the treatment group for list B (in Wave 2). Thus, the "arrival" number does not refer to the identification code of the FSW and was not assigned by the enumerator, who did not decide the order of FSWs to be interviewed.

In addition, we recruited another 135 new participants. Hence, the final data set contains information from a total of 786 different FSWs, i.e. 17% of the total population of sex workers in Dakar according to the last sex worker population census (APAPS, 2011-2012). Figure S5 in the supplementary file summarizes the number of respondents who answered the different lists, and Figure S6 in the supplementary file presents the samples for various sub-group analyses.

Descriptive statistics are presented in Table 1. FSWs were, on average, 36 years old in Wave 1. Approximately two-thirds of the participants were divorced. The average income from sex work was approximately 230 USD in both waves. Six and 9% of the sample were HIV positive in Wave 1 and Wave 2, respectively, and 19% participated in the PrEP demonstration project that occurred between the two survey waves.

3 Methods

3.1 Implementation of list experiments

The list experiment or item count technique is an indirect questioning method implemented to limit untruthful answers caused by social desirability bias. The principle of the list experiment method is to allocate respondents randomly to two different groups: a control group and a treatment group. Individuals allocated to the control group are presented with a number of non-sensitive statements. They are not asked to say whether they agree with each of the statements but are asked only how many they agree with. The same statements are presented to the treatment group; the difference is that a sensitive statement is added to the series of non-sensitive statements. Assuming that the two groups have a similar opinion on the non-sensitive statements, one can deduce the share of individuals in the treatment group that agree with the sensitive item by comparing the average number of agreed-to statements in each group (Blair and Imai, 2012; Glynn, 2013; Imai, 2011).

In our survey, the participants in the control (treatment) group were presented with the following question: *"I [the interviewer] will read three (four) statements. I will then ask you how many of these statements you agree with. You should not tell me which specific statement you agree with but rather the number of statements you agree with. I will give you three marbles and you have to hold them in your right hand. Keep both of your hands behind your back. For each of the statements, if you agree with it, please transfer one marble from your right hand to your left hand behind you. If you do not agree with it, please do not transfer any marble. At the end, I would like to know the total number of statements you agreed with. This number should correspond to the number of marbles you have in your left hand. I will now read the statements."*

Condom use was elicited directly and indirectly in the same survey. The list experiment question

was always asked first, and the direct question was asked later on in a section of the questionnaire focusing on sexual acts and clients' characteristics. We carried out the list experiment in 2015 among FSWs to estimate the use of condoms with their last client. In 2017, we again collected data on this population and extended this methodology through the implementation of a double list experiment (Droitcour et al., 1991). Such a design consists of presenting two lists to the respondents. Every FSW responded to both lists, with FSWs in the treatment group for list A becoming the controls for list B and vice versa (Hadji et al., 2016). The use of this double list experiment design was chosen to increase statistical power for sub-group analysis. The statements used in the two list experiments are presented in Figure 1 along with the methodology used to estimate the prevalence of condom use with each list.

3.2 List experiment hypothesis

The effectiveness of the list experiment methodology is based on three assumptions: (i) the randomisation of the treatment, (ii) the absence of any design effect, and (iii) the absence of "liars". First, the individuals allocated to each group must be similar such that, on average, they agree with the same number of non-sensitive statements. Second, the addition of the sensitive item must not change the sum of affirmative answers to the control items. Finally, as pointed out by Kuklinski et al. (1997), the choice of the control items needs to be such that individuals are not urged to provide dishonest responses. Individuals may be pushed to provide untruthful answers if they no longer benefit from privacy because they either agree or disagree with all of the non-sensitive items. We refer to these as the ceiling effects and floor effects, respectively; this assumption is also known as the no liar assumption. Glynn (2013) highlighted that to eliminate this problem, there should be one non-sensitive item that most participants will agree with and another non-sensitive item that most participants will disagree with. Finally, Blair and Imai (2012) advised choosing non-sensitive items that are related to the topic of the behaviour or opinion investigated in the list experiment to prevent suspicion on the part of respondents.

Table S1 in the supplementary file displays the characteristics of FSWs in the control and treatment groups. We note that the randomisation ensured a balance between the two groups with respect to their observable characteristics. The joint significance tests of a large share of the variables presented at the end of Table S1 confirmed the success of the randomisation (hypothesis (i)). In addition, Blair and Imai (2012) presented two theoretical tests to check the "no design effect" assumption (hypothesis (ii)). More precisely, the absence of a design effect implies that:

$$Pr(Y_i \leq y | T_i = 0) \geq Pr(Y_i \leq y | T_i = 1) \text{ for all } y = 0, ..., 3 \quad (1)$$

$$Pr(Y_i \leq y|T_i = 1) \geq Pr(Y_i \leq y - 1|T_i = 0) \text{ for all } y = 1, ..., 4 \quad (2)$$

In other words, the proportion of individuals in the control group who agree with no more than y statements (y =0,1,2,3) should be greater than this proportion for the treated group (see Row 5 in Table S2), and this latter proportion (for y =1,2,3,4) should be greater than the proportion of individuals in the control group who agree with no more than $y - 1$ statements (see Row 6 in Table S2). We tested this assumption by estimating the Bonferroni-corrected p-value (R-package 'list'). For the three lists (List A – 2015, Lists A and B – 2017), we obtained a value of 1. We therefore cannot reject the null hypothesis of no design effect. We also needed to ensure that the addition of the sensitive item did not modify the answers regarding the non-sensitive statements (hypothesis (iii)). In the supplementary material (cf. Table S2), we showed that the proportion of individuals who disagree with all items in the control group was less than 3% (ranging from 2.3 to 2.8%, depending on the list and wave considered), which guards against the possibility that FSWs in the treatment group were forced to agree with the sensitive item. We also mostly avoided the issue of the ceiling effect because the proportion of respondents in the control group who agreed with all non-sensitive items was also low (below 10%, ranging from 5.2 to 9.7%).

3.3 Empirical strategy

We pooled the results from the three list experiments conducted in 2015 and 2017 to investigate the characteristics of FSWs who did not use a condom during their last sexual intercourse.

$$Y_i = \lambda + \beta T_i + \mathbb{1}(List = A) + \mathbb{1}(Wave = 2015) + \varepsilon_i \quad (3)$$

Equation 3 shows that condom use estimated by the list experiment is implemented by regressing the number of statements the respondents agreed with (Y_i) on the allocation to the treatment group (T_i). The average condom use rate using the list experiment is then given by β and corresponds to the average difference in the number of statements between the control and the treatment groups. To account for our survey design that used two different lists at different time periods, we controlled for the variable $List$ (list A or list B), and we accounted for the change in condom use over time by controlling for the variable $Wave$ (2015 versus 2017).

Following Holbrook and Krosnick (2010) and Imai (2011), we investigated the relationship between condom use and FSW characteristics by interacting the allocation to the treatment group (T_i) with potential factors of condom use (S_i):

$$Y_i = \lambda + \beta T_i + \gamma S_i + \alpha T_i \times S_i + \mathbb{1}(List = A) + \mathbb{1}(Wave = 2015) + \varepsilon_i \quad (4)$$

Specifically, β in equation 4 reports the condom use rate among the sub-group for which $S_i = 0$ (i.e., HIV-negative FSWs), while α indicates the difference in condom use rates between the sub-group for which $S_i = 1$ (i.e., HIV-positive FSWs) and the sub-group for which $S_i = 0$. The p-value of α indicates whether the condom use rate is significantly different between the sub-groups.

Given that FSWs surveyed in 2017 answered list experiments A and B and that 377 FSWs[2] who participated in Wave 2 also participated in Wave 1, we clustered standard errors at the sex worker level in regressions 3 and 4.

4 Results

4.1 Measuring misreporting in condom use

Using face-to-face reported information, we found a very high proportion of FSWs who declared using condoms with their last client (97.3% in 2015 and 96.8% in 2017). In 2015, 69 out of 651 FSWs (10.6%) did not respond to the direct condom use question; this number was 18 out of 513 FSWs (3.5%) in 2017.[3] We compared the characteristics of FSWs who did answer the question and those who did not and found that the two groups differed; we present these differences in the supplementary file S3.

Table 2 presents the results of the list experiment exercise. Note that the two lists implemented in 2017 led to a similar estimation of condom use (78.0% with list A versus 78.4% with list B).[4] The double list experiment design allowed a significant increase in precision, reducing the standard error by 39.3% (36.2%) for list A (list B)[5]. This corresponded to an 8 to 10 percentage point reduction in the CI 95% interval.

When combining data collected in 2015 and 2017, the estimated condom use via the list experiment was 79.0%, which was significantly lower than the 97.0% estimated by the direct question ($p < 0.01$). Over-reporting was estimated to be 19.6 percentage points in 2015 and 17.1 percentage points in 2017.

[2]Note that we tested whether this sample is large enough in another paper using the method developed in (Blair et al., 2018) and conclude that the minimum sample size is reached.

[3]The same enumerators were recruited for the two survey waves and were sent to the same health center. This seems to have increased the confidence and experience of enumerators, which could explain why fewer FSWs did not answer the direct question in 2017 compared to 2015. Indeed, 2.1% of the FSWs who were interviewed in 2015 refused to answer the direct question in 2017, while this number was 6.7% among the new respondents (p-value = 0.01).

[4]These two prevalence rates are not statistically different.

[5]These computations are based on the Panel A estimations (cf. Table 2): $\frac{\text{SE(List A)-SE(List A \& B)}}{\text{SE(List A)}} = \frac{0.061 - 0.037}{0.061} = 0.393$. Similar results were obtained with Panel B.

4.2 HIV infection risk and condom use

We investigated the relationship between HIV infection risk and condom use. Table 3 displays the results obtained for this sub-group analysis.
Importantly, HIV-positive registered FSWs (estimated via medical records) were found to report much lower levels of condom use than HIV-negative registered FSWs (33.9% vs. 80.5%, p-value=0.009).
We then investigated the existence of condom use differences according to exposure to various HIV prevention strategies. We found a borderline significant increase in condom use among registered FSWs, with condom use among registered FSWs being 10.8 percentage points higher than amongst non-registered FSWs (84.2% vs. 73.4%, p-value=0.095). However, we neither found that PrEP participation led to decreased condom use (80.7% vs. 77.6%, p-value=0.755) nor found that receiving free condoms led to increased condom use (79.0% vs. 80.9%, p-value = 0.770).

5 Discussion

The list experiment suggests that FSWs in Dakar over-reported condom use by 18.0 points in face-to-face interviews (97.0% vs. 79.0%, p-value < 0.001). Our results provide some evidence about the factors affecting condom use. Sub-group analysis shows that condom use among HIV-positive FSWs was only 29.3%. Despite the difficulty in determining a causal effect of HIV status on condom use, this finding is worrying since it suggests that the riskiest sex acts are unprotected. Second, we found that condom use is significantly higher for registered FSWs, which suggests that this policy may reduce risky behaviours. However, we did not find that FSWs who participated in the PrEP demonstration project had lower condom use.

The fact that the two list experiments conducted in 2017 led to a similar estimation of condom use (78.0% vs 78.4%) confirms that the choice of non-sensitive items related to the topic of the sensitive item did not affect the results as long as the list experiment hypotheses were fulfilled. The high misreporting in our study can likely be explained by the characteristics of the targeted population. Because they are stigmatised, FSWs feared disclosing socially unacceptable behaviours to the enumerators. Hence, by guaranteeing anonymity, the use of indirect elicitation methods is relevant to this population. Nonetheless, we acknowledge that condom use may still be over-estimated. While those methods guarantee privacy in response to survey participants, they cannot help with participants who do not want to reveal their true behaviour.

While a number of methods can be used as alternatives to self-reported face-to-face interviews, our study highlights the high potential of the double list experiment as a tool for eliciting less-biased estimates of condom use in behavioural surveys conducted in low-income countries. Our

results show that the double list experiment method has the advantage of allowing sub-group analyses, and by increasing statistical power, the double list experiment method allows the investigation of the role of characteristics that are not frequent in the population (e.g., HIV infection) in sensitive behaviour. Our conclusion challenges results presented in other recent studies. For instance, Bell and Bishai (2019) found that the results of the list experiment method showed a smaller proportion of the sensitive behaviour than the results of the direct question method did. However, the authors showed that the reason for such findings lies in issues in the implementation of the list experiment, assuming that participants have mentally enumerated the treatment list items differently from the control list items. Another paper by Chuang et al. (2019) concluded that the list experiment has weak internal consistency. These authors implemented several double list experiments to measure the prevalence of sensitive sexual behaviours in African countries. They found that the prevalence estimated from the two lists differed strongly for at least half of the behaviours estimated. Looking at the design of those lists, one can note that the discrepancies in the results may have been caused by the violation of several key assumptions of the list experiment method (e.g., design effect, ceiling and floor effects). Violations of those assumptions have led to a lack of confidentiality of the answers in some lists, while such confidentiality was guaranteed in other lists. Finally, Haber et al. (2018) found that the list experiment method had poor external validity for eliciting HIV status and comparing the prevalence obtained with the list experiment method to objective measures (biological markers). However, the use of non-sensitive items unrelated to HIV status surely explains why the authors found no difference between the elicited and self-reported serostatus. Indeed, the mix of sentences such as *"I prefer bananas over grapes"* or *"I played football yesterday"* along with the sensitive item may make the sensitive item stand out too much, especially considering the stigma attached to the sensitive item under study (HIV infection). List-experiment implementation guidelines stress the need to use non-sensitive items related to the sensitive item that is of interest (Blair and Imai, 2012). While the abovementioned studies differed in design, the failure of the list experiment method in these studies was always due to the violation of key assumptions of the methodology. As a result, the list experiment method has the potential to improve the data quality of sexual and health surveys, but special attention needs to be given to the survey design.

Our study had several limitations. First, HIV status was only available for registered FSWs, which limits the generalisability of our results to the whole population of FSWs. Second, we were not able to assess the effect of being on PrEP on condom use, given that the PrEP project ended six months before the second wave of data collection. The study showed that PrEP does not lead to risk compensation in the long term, but we will investigate the effect of PrEP on condom use after its introduction for FSWs in Senegal in early 2020. Third, the study recruited all active registered sex workers in the health centres located in the suburb of Dakar,

and the sample of registered FSWs is likely to be representative of this population. However, non-registered FSWs were recruited using snowball sampling, and this sample is likely to over represent non-registered FSWs who are connected to FSW groups and non-governmental organizations. As a result, the use of our selected sample may lead to the overestimation of the bias in condom use obtained with direct questioning. In addition, we were not able to include FSWs under the age of 18 for ethical reasons. Finally, we were not able to determine a causal effect of FSW characteristics on condom use. For instance, it is possible that there are confounders that affect both the decision to register as a FSW or participate in the PrEP demonstration project and the decision to use condoms. In fact, registered and non-registered sex workers differ in their characteristics (Ito et al., 2018); hence, it is not possible to conclude that there is a causal effect of registration on condom use. Additional research using the list experiment in the context of quasi-experimental designs is required to establish a causal effect of registration and/or PrEP use on condom use.

Future research on the use of the list experiment method to elicit information on sexual behaviours could be conducted along three axes. First, future research on condom use measurement should aim to test the validity of the results obtained with the list experiment. This could be done by comparing the results obtained with the list experiment to the detection of prostate-specific antigen. Second, although the list experiment method has been successfully used in low-literacy settings (De Cao and Lutz, 2018; Ghofrani et al., 2018; Gibson et al., 2018; Moseson et al., 2015), additional research should assess the validity of using the list experiment method in low literacy settings without the use of marbles, which would better support its use in national surveys. Finally, statistical methods should be developed to use the list randomised variable as a left-hand side variable.

6 Conclusion

We used list experiments to investigate the role of HIV infection and HIV prevention in condom use. Our results confirmed the existence of a high social desirability bias among this high-risk group in Senegal. When analysing the determinants of condom use, we provided alarming evidence that HIV-positive FSWs have very low rates of condom use. The results suggest that the list experiment provides a promising technique for improving the reporting of sensitive behaviours among a low-literacy population in a resource-poor setting as well as a method for identifying barriers to condom use in these settings.

References

Josephine Aho, Anita Koushik, Soumaila Laye Diakité, Kovana Marcel Loua, Vinh-Kim Nguyen, and Selim Rashed. Biological validation of self-reported condom use among sex workers in guinea. *AIDS and Behavior*, 14(6):1287–1293, 2010.

APAPS. Etude pour l'estimation du nombre de ts et de hsh dans la région de dakar. Technical report, commanditée par la DLSI, le CNLS et FHI, 2011-2012.

APAPS and IRESSEF. Enquête nationale de surveillance combinée des ist et du vih/sida (ensc). Technical report, APAPS, 2015.

Suzanne O. Bell and D. Bishai. Can a list experiment improve validity of abortion measurement. *Studies in Family Planning*, 50(1):43–61, 2019.

Graeme Blair and Kosuke Imai. Statistical analysis of list experiments. *Political Analysis*, 20 (1):47–77, 2012.

Graeme Blair, Alexander Coppock, and Margaret Imai Moor. When to worry about sensitivity bias: Evidence from 30 years of list experiments. 2018.

J. Blumenthal and R. Haubrich. Risk compensation in prep: An old debate emerges yet again. *The virtual mentor : VM*, 16(11):909–915, 2014.

Alberto Chong, Marco Gonzalez-Navarro, Dean Karlan, and Martin Valvidia. Do information technologies improve teenagers' sexual education? evidence from a randomized evaluation in colombia. *NBER Working Papers 18776*, 2013.

Erica Chuang, Pascaline Dupas, Elise Huillery, and Juliette Seban. Sex, lies, and measurement: Do indirect response survey methods work? 2019.

D.A. Cohen, S.-Y. Wu, and T.A. Farley. Comparing the cost-effectiveness of hiv prevention interventions. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 37(3):1404–1414, 2004.

A. Creese, K. Floyd, A. Alban, and L. Guinness. Cost-effectiveness of hiv/aids interventions in africa: a systematic review of the evidence. *The Lancet*, 359(9318):1635–1642, 2002.

Elisabetta De Cao and CLemens Lutz. Sensitive survey questions: Measuring attitudes regarding female genital cutting through a list experiment. *Oxford Bulletin of Economics and Statistics*, 80(5):871–892, 2018.

D. Donnell, J. M. Baeten, J. Kiarie, K. K. Thomas, W. Stevens, C. R. Cohen, . . ., and C. Celum. Heterosexual hiv-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. *The Lancet*, 375(9731):2092–2098, 2010.

Judith Droitcour, Rachel A. Caspar, Michael L. Hubbard, Teresa L. Parsley, Wendy Visscher, and Trena M. Ezzati. The item-count technique as a method of indirect questioning: A review of its development and a case study application. *Measurement Errors in Surveys*, pages 185–210, 1991.

J. Dureau, K. Kalogeropoulos, P. Vickerman, M. Pickles, and M. C. Boily. A bayesian approach to estimate changes in condom use from limited human immunodeficiency virus prevalence data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 65(2):237–257, 2016.

Ellen E. Foley. Regulating sex work: subjectivity and stigma in senegal. *Culture, health & sexuality*, 19(1):50–63, 2017.

AM. M. Foss, M. Hossain, P.T. Vickerman, and C.H. Watts. A systematic review of published evidence on intervention impact on condom use in sub-saharan africa and asia. *Sexually Transmitted Infections*, 83(7):510–16, 2007.

M.F. Gallo, F.M. Behets, M.J. Steiner, S.C. Thomsen, W. Ombidi, S. Luchters, C. Toroitich-Ruto, and M.M. Hobbs. Validity of self-reported 'safe sex' among female sex workers in mombasa, kenya - psa analysis. *International Journal of STD & AIDS*, 18(1):33–38, 2007.

Marziyeh Ghofrani, Fariba Asghari, Maryam Kashanian, Hojat Zeraati, and Akbar Fotouhi. Prevalence of induced abortion in iran: A comparison of two indirect estimation techniques. *International Perspectives on Sexual and Reproductive Health*, 44(2):73–79, 2018.

Mhairi A. Gibson, Eshetu Gurmu, Beatriz Cobo, Maria M. Rueda, and Isabel M. Scott. Indirect questioning method reveals hidden support for female genital cutting in south central ethiopia. *Plos One*, 13(5):e0193985, 2018.

Adam N. Glynn. What can we learn with statistical truth serum? design and analysis of the list experiment. *Public Opinion Quarterly*, 77(S1):159–72, 2013.

Ezequiel Gonzalez-Ocantos, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. Vote buying and social desirability bias: Experimental evidence from nicaragua. *American Journal of Political Science*, 56(1):202–17, 2012.

R. M. Grant, J. R. Lama, P. L. Anderson, V. McMahan, A. Y. Liu, L. Vargas, . . ., and M. E. Ramirez-Cardich. Preexposure chemoprophylaxis for hiv prevention in men who have sex with men. *New England Journal of Medicine*, 363(27):2587–2599, 2010.

Noah Haber, Guy Harling, Jessica Cohen, Tinofa Mutevedzi, Franck Tanser, Dickman Gareta, Kobus Herbst, Deenan Pillay, Bärnighausen, and Günther Fink. List randomization for eliciting hiv status and sexual behaviors in rural kwazulunatal, south africa: a randomized experiment using known true values for validation. *BMC Medical Research Methodology*, 18: 46, 2018.

M. Hadji, F. Asghari, M. Yunesian, P. Kabiri, and A. Fotouhi. Assessing the prevalence of publication misconduct among iranian authors using double list experiment. *Iranian Journal of Public Health*, 45(7):897–904, 2016.

Sarah E. Hanck, Kim M. Blankenship, Kevin S. Irwin, Brooke S. West, and Trace Kershaw. Assessment of self-reported sexual behavior and condom use among female sex workers in india using polling box approach: A preliminary report. *Sexually Transmitted Diseases*, 35 (5):489–494, 2008.

Allyson L. Holbrook and Jon A. Krosnick. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 74(1):37–67, 2010.

Kosuke Imai. Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association*, 106(494):407–416, 2011.

Seiro Ito, Aurélia Lépine, and Carole Treibich. The effect of sex work regulation on health and wellbeing of sex workers: Evidence from senegal. *Health Economics*, 27(11):1627–1652, 2018.

Julian C. Jamison, Dean Karlan, and Pia Raffler. Mixed-method evaluation of a passive mhealth sexual information texting service in uganda. *Information Technologies & International Development*, 9(3):1–28, 2013.

Q. A. Karim, S. S. A. Karim, J. A. Frohlich, A. C. Grobler, C. Baxter, L. E. Mansoor, . . ., and Z. Omar. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of hiv infection in women. *Science*, 329(5996):1168–1174, 2010.

Dean S. Karlan and Jonathan Zinman. List randomization for sensitive behavior: An application for measuring use of loan proceeds. *Journal of Development Economics*, 98(1):71–75, 2012.

Eric Kramon and Keith Weghorst. (mis)measuring sensitive attitudes with the list experiment: Solutions to list experiment breakdown in kenya. *Public Opinion Quarterly*, 83(S1):236–263, 2019.

Ivar Krumpal. Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity*, 47(4):2025–47, 2013.

J.H. Kuklinski, M.D. Cobb, and M. Gilens. Racial attitudes and the 'new south'. *The Journal of Politics*, 59(2):323–49, 1997.

J. W. LaBrie and M. Earleywine. Sexual risk behaviors and alcohol: Higher base rates revealed using the unmatched-count technique. *Journal of Sex Research*, 37(4):321–26, 2000.

Jeffrey R. Lax, Justin H. Phillips, and Alissa F. Stollwerk. Are survey respondents lying about their support for same sex marriage? lessons from a recent list experiment. *Public Opinion Quarterly*, 80(2):510–33, 2016.

H. Liu, D. E. Morisky, X. Lin, E. Ma, B. Jiang, and Y. Yin. Bias in self-reported condom use: Association between over-reported condom use and syphilis in a three-site study in china. *AIDS Behavior*, 20(6):1343–1352, 2016.

David McKenzie and Melissa Siegel. Eliciting illegal migration rates through list randomization. *Migration Studies*, 1(3):253–57, 2013.

K.M. Mitchell, A. Lépine, F. Terris-Prestholt, K. Torpey, H. Khamofu, M.O. Folayan, J. Musa, J. Anenih, A.S. Sagay, E. Alhassan, J. Idoko, and P. Vickerman. Modelling the impact and cost-effectiveness of combination prevention amongst hiv serodiscordant couples in nigeria. *AIDS*, 29(15):2035–44, 2015.

Heidi Moseson, Moses Massaquoi, Christine Dehlendorf, Luke Bawo, Bernice Dahn, Yah Zolia, Eric Vittinghoff, Robert A. Hiatt, and Caitlin Gerdts. Reducing under-reporting of stigmatized health events using the list experiment: results from a randomized, population-based study of abortion in liberia. *International Journal of Epidemiology*, 44(6):1951–1958, 2015.

Heidi Moseson, Caitlin Gerdts, Christine Dehlendorf, Robert A. Hiatt, and Eric Vittinghoff. Multivariable regression analysis of list experiment data on abortion: results from a large, randomly-selected population based study in liberia. *Population Health Metricsvolume*, 15: 40, 2017a.

Heidi Moseson, Emily Treleaven, Caitlin Gerdts, and Nadia Diamond-Smith. The list experiment for measuring abortion: What we know and what we need. *Family Planning*, 48(4): 397–405, 2017b.

H.S. Moseson, C. Gerdts, S. Baum, K. White, K. Hopkins, J.E. Potter, and D. Grossman. Measuring texas women's experiences with abortion self-induction using a list experiment. *Contraception*, 96(4):272, 2017c.

Carole Treibich and Aurélia Lépine. Estimating misreporting in condom use and its determinants among sex workers: Evidence from the list randomisation method. *Health Economics*, 28(1):144–160, 2019.

S.S. Weir, R.E. Roddy, L. Zekeng, and K.A. Ryan. Association between condom use and hiv infection: a randomised study of self reported condom use measures. *Journal of Epidemiology and Community Health*, 53(7):417, 1999.

WHO. Who expands recommendation on oral pre-exposure prophylaxis of hiv infection. Technical report, Geneva: World Health Organization, 2015.

Figure 1: (Double) list experiment design

Respondents assigned to Group 1	***Respondents assigned to Group 2***	***Estimated prevalence of condom use***
List A	**List A**	
1. It is safer to bring a client home than going to the hotel	1. It is safer to bring a client home than going to the hotel	*I used a condom during my last intercourse with a client*
2. *I used a condom during my last intercourse with a client*		
3. I prefer that the client pays me before the sexual intercourse	2. I prefer that the client pays me before the sexual intercourse	
4. Monday is the day I have the greatest number of clients	3. Monday is the day I have the greatest number of clients	
Number of agreed statements : $G1_A$	Number of agreed statements : $G2_A$	P_A = average ($G1_A$) - average ($G2_A$)
List B	**List B**	
1. The majority of my clients are Senegalese	1. The majority of my clients are Senegalese	
	2. *I used a condom during my last intercourse with a client*	
2. I usually spend the whole night with my client	3. I usually spend the whole night with my client	
3. I usually solicit clients by phone	4. I usually solicit clients by phone	
Number of agreed statements : $G1_B$	Number of agreed statements : $G2_B$	P_B = average ($G2_B$) - average ($G1_B$)

Notes: List A was implemented in 2015 and 2017 while list B was implemented only in 2017.
Respondents assigned to group 1 serve as treated units for list A and as controls for list B while respondents assigned to group 2 serve as controls for list A and as treated for list B.

Table 1: Descriptive statistics

Variable	*Wave 1 = 2015* Obs	Mean	SD	*Wave 2 = 2017* Obs	Mean	SD
Socio-economic characteristics						
Age (in years)	651	35.88	9.23	513	38.34	9.40
Income from sex work (CFAF)	649	133,387	123,428	507	127,550	111,280
Divorced (%)	651	69.28	46.17	513	68.81	46.37
HIV status						
HIV positive (%, medical record)	219	5.94	23.68	173	8.09	27.35
HIV prevention strategies						
Registered with authorities (%)	650	50.00	50.04	512	49.80	50.05
Received free condoms (%)	641	67.08	47.03	510	60.59	48.91
Participated in the PrEP demonstration project (%)	-	-	-	513	18.91	39.20

Notes: The sample is composed of 651 and 513 FSWs in 2015 and 2017 respectively. Differences in the number of observations for a given year are due to missing information.
1 USD = 588 CFAF in June 2015 and 1 USD = 555 CFAF in August 2017.

Table 2: Estimated condom use and over-reporting

Condom use	Obs	Number of statements: Treatment	Number of statements: Control	Estimated condom use	Clustered SE	95% CI	Self-reported condom use	Over-reporting
Panel A - all observations								
2015 - List A	651	2.50	1.70	0.797	0.057	[0.685; 0.909]	-	-
2017 - List A	513	2.42	1.64	0.780	0.061	[0.660; 0.900]	-	-
2017 - List B	513	2.67	1.89	0.784	0.058	[0.671; 0.897]	-	-
2017 - Lists A & B	1,026	2.55	1.76	0.782	0.037	[0.709; 0.856]	-	-
2015 & 2017 - Lists A & B	1,677	2.53	1.74	0.788	0.032	[0.725; 0.851]	-	-
Panel B - observations for which we have the self-declared condom use								
2015 - List A	582	2.46	1.68	0.777	0.061	[0.657; 0.898]	0.973	0.196
2017 - List A	495	2.43	1.63	0.802	0.062	[0.681; 0.924]	0.968	0.166
2017 - List B	495	2.68	1.89	0.793	0.058	[0.678; 0.907]	0.968	0.175
2017 - Lists A & B	992	2.56	1.76	0.797	0.038	[0.723; 0.872]	0.968	0.171
2015 & 2017 - Lists A & B	1,574	2.52	1.73	0.790	0.033	[0.724; 0.856]	0.970	0.180

Notes: Estimated condom use corresponds to the $\hat{\beta}$ in equation $Y_i = \lambda + \beta T_i + \mathbb{1}(List = A) + \mathbb{1}(Wave = 2015) + \varepsilon_i$, with SE clustered at the FSW's level - equation (3). Differences between the number of observations are due to missing information to the self-reported question. Over-reporting is computed by comparing the self-reported condom use rate with the one estimated with the list experiment method.

Table 3: Condom use by sub-groups

Subgroups (S_i)	Obs	Estimated condom use S_i =No (1)	Estimated condom use S_i =Yes (2)	Difference p-value ±
HIV status				
HIV positive (medical record ⋄) †	565	0.805	0.339	0.009
HIV prevention strategies				
Registered with authorities †	1,674	0.734	0.842	0.095
Received free condoms †	1,629	0.809	0.790	0.770
Participated in PrEP demonstration project ‡	1,026	0.776	0.807	0.755

Notes: † Data from the 2015 and 2017 surveys are considered. Observations from the three lists are used leading to 1,677 observations (651 + 513 × 2). Differences in the number of observations is due to missing information.

‡ This information is available only in the 2017 survey. The lists A and B are used leading to 1,026 observations (513 × 2). Column (1) corresponds to $\hat{\beta}$ and Column (2) to $(\hat{\beta} + \hat{\alpha})$ in equation 4 with SE clustered at the FSW's level:
$Y_i = \lambda + \beta T_i + \gamma S_i + \alpha T_i \times S_i + \mathbb{1}(List = A) + \mathbb{1}(Wave = 2015) + \varepsilon_i$.

± The p-value refers to the significance level of $\hat{\alpha}$.

⋄ Information based on medical records (available only for registered individuals).

Reading indications: For the HIV positive variable (first line of the table), 80.5% of the HIV negative registered FSWs used a condom with heir last client while 33.9% of HIV positive registered FSWs did so.

Supplementary file

Table S1: Tests of randomisation (2015 and 2017 surveys)

Variables	Wave 1 - 2015				Wave 2 - 2017			
	Obs	**Control**	**Treated**	**p-value**	Obs	**Control**	**Treated**	**p-value**
	651	*323*	*328*		*513*	*252*	*261*	
Socio-demographic characteristics								
Age (in years)*	*651*	35.58	36.16	0.421	*513*	38.23	38.46	0.782
Is divorced (%)*	*651*	67.80	70.73	0.419	*513*	65.87	71.65	0.159
Never married (%)*	*651*	25.70	23.17	0.454	*513*	22.22	18.01	0.234
Use condoms as contraceptive method (%)	*495*	52.57	49.17	0.451	*513*	24.60	24.14	0.903
Has a regular partner (%)*	*651*	46.13	41.16	0.202	*513*	48.41	50.57	0.625
Household (HH) size*	*651*	6.26	6.24	0.957	*513*	6.93	7.16	0.634
HH monthly expenditures (CFAF)*	*651*	358,017	349,909	0.757	*513*	365,815	357,365	0.745
Monthly sex revenues (CFAF)*	*649*	134,498	132,299	0.821	*507*	123,872	131,101	0.465
HH received transfers in the past year (%)*	*649*	27.73	25.00	0.431	*510*	25.20	24.23	0.800
HH sent transfers in the past year (%)*	*647*	38.87	38.11	0.843	*512*	23.51	28.35	0.212
Risk aversion in sex (1 to 10)*	*651*	7.76	7.64	0.567	*513*	7.52	7.69	0.467
Preference for future (1 to 10)*	*651*	6.69	6.88	0.457	*513*	7.22	7.74	0.079
HIV knowledge (score 0-8)*	*651*	6.32	6.45	0.186	*513*	6.23	6.24	0.847
Fear of discrimination due to HIV	*614*	67.43	71.61	0.261	*458*	62.22	66.95	0.291
Sex work activity								
Number of clients within a week*	*648*	6.49	6.56	0.893	*513*	8.30	8.41	0.889
Has only occasional clients (%)*	*645*	11.32	14.98	0.170	*513*	4.37	4.60	0.899
Has only regular clients (%)*	*645*	33.02	32.42	0.871	*513*	35.32	36.02	0.869
Last client was an occasional client (%)*	*645*	40.37	47.68	0.062	*513*	25.79	28.35	0.515
Declared use of condom with last client (%)$^\diamond$	*582*	97.60	96.90	0.603	*496*	98.38	95.18	0.044
Work mostly in bars or brothels (%)*	*651*	23.84	26.83	0.381	*513*	36.51	33.33	0.452
Link with the authorities and the health system								
Registered FSW (%)*	*650*	47.68	52.29	0.240	*512*	49.21	50.38	0.790
Received free condoms (%)$^\diamond$	*641*	65.41	68.73	0.372	*511*	64.14	56.92	0.096
Is affiliated to a STD centre*	*648*	72.36	74.01	0.637	*512*	56.75	60.38	0.404
Visited a STI centre in the last month (%)*	*651*	56.97	56.10	0.824	*513*	34.92	36.40	0.728
Did a HIV test in the last 12 months (%)*	*651*	81.11	80.18	0.764	*513*	86.51	81.61	0.131
HIV seropositive (medical record data) (%)	*219*	4.90	6.84	0.548	*173*	11.70	3.80	0.058
Had any STI symptom in the last month (%)*	*646*	20.67	23.55	0.383	513	11.51	15.33	0.206
Participated in the PrEP demonstration	-	-	-	-	*513*	21.83	16.09	0.098
Test of joint significance								
(considering the variables indicated by *):	F(22,606) = 0.70, p-value = 0.843				F(22,479) = 0.87, p-value = 0.642			
(considering the variables indicated by * and $^\diamond$):	F(24,542) = 0.89, p-value = 0.608				F(24,458) = 1.20, p-value = 0.238			

Remark: 651 and 513 FSWs answered to the list experiment questions in 2015 and 2017 respectively. Differences in the number of observations for a given year are due to missing information.

Table S2: Checking floor, ceiling and design effects for the different lists and waves

Estimated proportions	Source	Number of reported items						Sum
		Obs	0	1	2	3	4	
2015 - List A								
Row 1	Treatment list	*328*	0.006	0.079	0.409	0.424	0.082	1.000
Row 2	Proportion at least		1	0.994	0.915	0.506	0.082	-
Row 3	Control list	*323*	0.028	0.334	0.548	0.090	0	1.000
Row 4	Proportion at least		1	0.972	0.638	0.090	0	-
Row 5	Equation 1		0	0.022	0.277	0.416	0.082	0.796
	SE		(0.010)	(0.031)	(0.032)	(0.015)	-	
Row 6	Equation 2			0.006	0.058	0.013	0.008	
	SE			(0.004)	(0.018)	(0.038)	(0.022)	
2017 - List A								
Row 1	Treatment list	*255*	0	0.078	0.474	0.396	0.051	1.000
Row 2	Proportion at least		1	1	0.922	0.448	0.052	-
Row 3	Control list	*258*	0.023	0.403	0.484	0.090	0	1.000
Row 4	Proportion at least		1	0.977	0.574	0.090	0	-
Row 5	Equation 1		0	0.023	0.348	0.358	0.051	0.781
	SE		(0.009)	(0.035)	(0.036)	(0.014)	-	
Row 6	Equation 2			0.000	0.055	0.127	0.038	
	SE			(0.000)	(0.019)	(0.044)	(0.023)	
2017 - List B								
Row 1	Treatment list	*258*	0.004	0.031	0.349	0.519	0.097	1.000
Row 2	Proportion at least		1	0.996	0.965	0.616	0.097	-
Row 3	Control list	*255*	0.024	0.165	0.710	0.102	0	1.000
Row 4	Proportion at least		1	0.976	0.811	0.102	0	
Row 5	Equation 1		0	0.020	0.154	0.514	0.097	0.785
	SE		(0.010)	(0.027)	(0.036)	(0.018)	-	
Row 6	Equation 2			0.004	0.011	0.196	0.005	
	SE			(0.004)	(0.015)	(0.039)	(0.026)	

Remark: Rows 5 and 6 test the absence of design effect.
Row 5: a positive value indicates that the proportion of individuals in the control group who agree with no more than y statements is greater than this proportion for the treated group.
Row 6: a positive value indicates that the proportion of individuals in the treated group who agree with no more than y statements is greater than the proportion of individuals in the control group who agree with no more than $y - 1$ statements.
The sum of the difference between Row 2 and Row 4 gives the difference-in-means estimator.

Table S3: Comparison between women who did or did not self-report condom use (2015 and 2017 surveys)

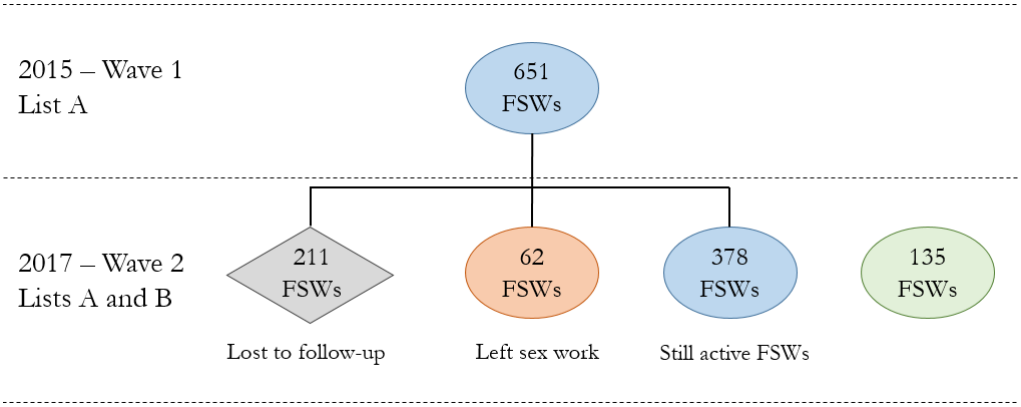
	Wave 1 - 2015				Wave 2 - 2017			
		Self-reported condom use				Self-reported condom use		
Variables	Obs	**No**	**Yes**	**p-value**	Obs	**No**	**Yes**	**p-value**
	651	*69*	*582*		*513*	*18*	*495*	
Socio-demographic characteristics								
Age (in years)	*651*	39.16	35.49	0.002	*513*	38.00	38.356	0.875
Is divorced (%)	*651*	88.41	67.01	<0.001	*513*	61.11	69.09	0.474
Never married (%)	*651*	10.14	26.12	0.004	*513*	27.78	19.78	0.407
Use condoms as contraceptive method (%)	*495*	65.91	49.45	0.037	*513*	27.78	24.24	0.732
Has a regular partner (%)	*651*	40.58	43.99	0.590	*513*	27.78	50.30	0.061
Household (HH) size	*651*	6.78	6.19	0.376	*513*	8.17	7.01	0.373
HH monthly expenditures (CFAF)	*651*	306,967	359,501	0.217	*513*	283,372	364,333	0.252
Monthly sex revenues (CFAF)	*649*	127,328	134,084	0.672	*507*	98,056	128,636	0.253
HH received transfers in the past year (%)	*649*	24.64	26.55	0.733	*510*	27.78	24.59	0.759
HH sent transfers in the past year (%)	*647*	34.33	38.97	0.461	*512*	0.00	26.92	0.010
Risk aversion in sex (1 to 10)	*651*	1.45	2.40	0.006	*513*	1.89	2.41	0.419
Preference for future (1 to 10)	*651*	6.38	6.84	0.286	*513*	7.94	7.47	0.547
HIV knowledge (score 0-8)	*651*	6.49	6.37	0.481	*513*	6.06	6.24	0.389
Fear of discrimination due to HIV (%)	*614*	87.88	67.34	0.001	*458*	73.33	64.33	0.475
Sex work activity								
Number of clients within a week	*648*	6.36	6.55	0.819	*513*	5.67	8.45	0.188
Has only occasional clients (%)	*651*	4.35	14.09	0.023	*512*	5.88	4.44	0.779
Has only regular clients (%)	*650*	35.29	32.13	0.599	*512*	41.18	35.56	0.635
Last client was an occasional client (%)	*645*	28.57	45.70	0.009	*513*	22.22	27.27	0.637
Work mostly in bars or brothels (%)	*651*	50.72	45.88	0.446	*511*	23.53	40.89	0.152
Link with the authorities and the health system								
Registered FSW (%)	*650*	79.41	46.56	<0.001	*513*	27.78	30.51	0.978
Received free condoms (%)	*641*	86.76	64.75	<0.001	*494*	56.25	59.21	0.814
Is affiliated to a STD centre (%)	*648*	85.29	71.72	0.017	*512*	44.44	59.11	0.216
Visited a STI centre in the last month (%)	*651*	79.71	53.78	<0.001	*513*	16.67	36.36	0.087
Did a HIV test in the last 12 months (%)	*651*	88.41	79.73	0.085	*513*	72.22	84.44	0.165
HIV seropositive (medical record data) (%)	219	8.82	5.41	0.441	*173*	0.00	8.14	.
Had any STI symptom in the last month (%)	*646*	27.94	21.45	0.224	*510*	33.33	12.80	0.012
Participated in the PrEP demonstration (%)	-	-		-	*513*	11.11	19.19	0.391

Remark: Differences in the number of observations are due to missing information.

Table S4: Attrition

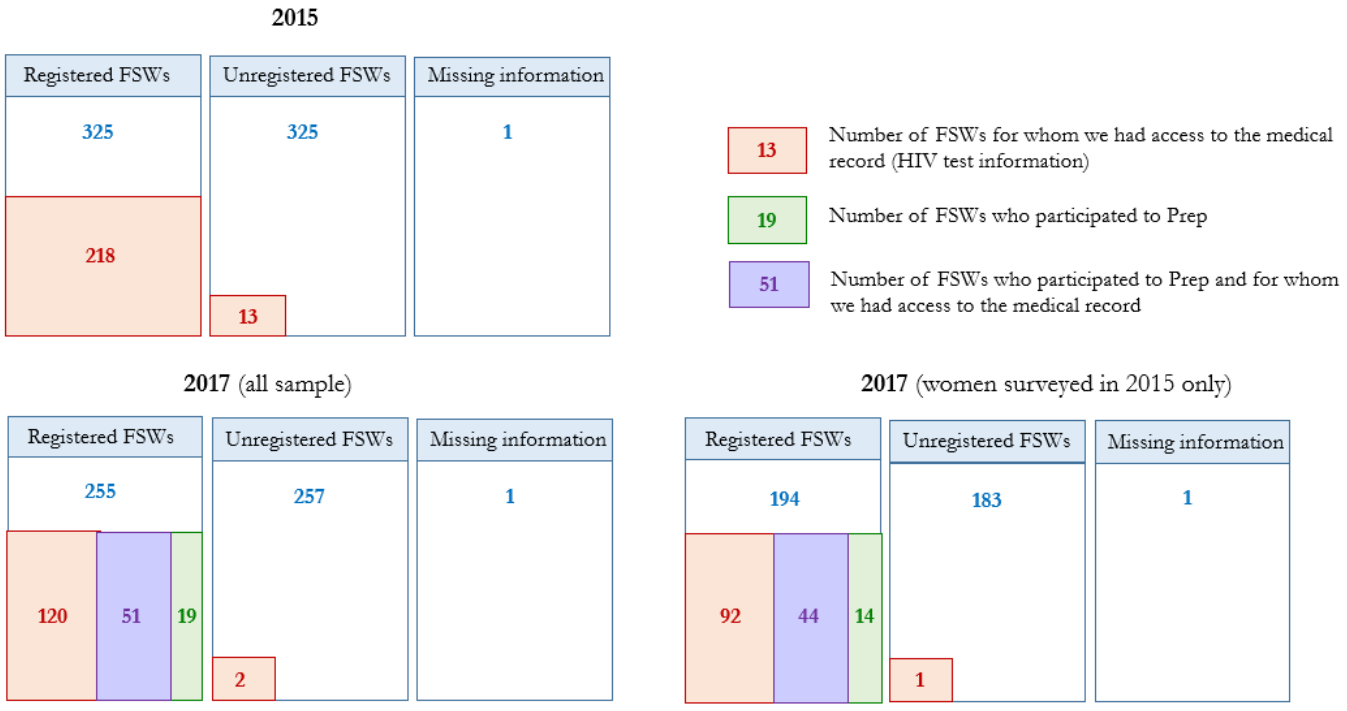
		Mean		
	Obs	**Wave 1 only**	**Waves 1 & 2**	p-value
	651	*211*	*440*	
Socio-demographic characteristics				
Age (in years)	*651*	33.61	36.96	0.000
Is divorced (%)	*651*	66.82	70.45	0.348
Never married (%)	*651*	28.43	22.50	0.099
Use condoms as contraceptive method (%)	*495*	52.41	50.15	0.636
Has a regular partner (%)	*651*	43.13	43.86	0.860
Household (HH) size	*651*	5.52	6.60	0.014
HH monthly expenditures (CFAF)	*651*	336,765	362,165	0.364
Monthly sex revenues (CFAF)	*649*	140,014	130,239	0.346
HH received transfers in the past year (%)	*649*	21.80	28.54	0.068
HH sent transfers in the past year (%)	*647*	36.67	39.36	0.511
Risk aversion in sex (1 to 10)	*651*	2.24	2.33	0.706
Preference for future (1 to 10)	*651*	6.75	6.80	0.863
HIV knowledge (score 0-8)	*651*	6.43	6.36	0.546
Fear of discrimination due to HIV (%)	*614*	67.16	70.73	0.365
Sex work activity				
Number of clients within a week	*648*	6.61	6.49	0.832
Has only occasional clients (%)	*651*	16.59	11.36	0.064
Has only regular clients (%)	*650*	27.96	34.62	0.090
Last client was an occasional client (%)	*645*	46.63	42.79	0.359
Declared use of condom with last client (%)	*582*	95.21	98.22	0.038
Work mostly in bars or brothels (%)	*651*	48.34	45.45	0.490
Link with the authorities and the health system				
Registered FSW (%)	*650*	50.71	49.66	0.802
Received free condoms (%)	*641*	62.32	69.35	0.077
Is affiliated to a STD centre (%)	*648*	71.77	73.80	0.586
Visited a STI centre in the last month (%)	*651*	55.92	56.82	0.830
Did a HIV test in the last 12 months (%)	*651*	79.62	81.14	0.648
HIV seropositive (medical record data) (%)	*219*	4.23	6.76	0.460
Had any STI symptom in the last month (%)	*646*	21.43	22.48	0.764

Remark: Differences in the number of observations are due to missing information.



Notes : 651 female sex workers (FSWs) were interviewed in 2015 and thus answered to list A question.
In 2017, 440 FSWs were re-interviewed but only the still active FSWs answered to list A and list B questions.
135 additional FSWs were interviewed in 2017 and answered to both lists.
Put differently, 651 and 513 (378 + 135) FSWs answered to the list questions in 2015 and 2017 respectively.
In short, 786 (651 + 135) different FSWs answered at least to one list question.

Figure S5: Sample of list experiment respondents in 2015 and 2017



Note: In 2017, there were 70 FSWs involved in PrEP, and among them 51 had medical record.
We can note that 58 of FSWs who had PrEP in 2017 were surveyed in 2015.

Figure S6: Number of FSWs in sample

Table S5: Gain in precision with the double list experiment - Sub-group analysis

		Obs	$\hat{\alpha}$	SE($\hat{\alpha}$)	SE reduction
HIV status					
HIV positive (medical record) ⋄					
	Double list	*346*	-0.345	0.275	
	List A	*173*	0.005	0.367	-0.251
	List B	*173*	-0.650	0.357	-0.230
HIV prevention strategies					
Registered with authorities					
	Double list	*1,024*	0.114	0.075	
	List A	*512*	0.073	0.122	-0.385
	List B	*512*	0.154	0.114	-0.342
Received free condoms					
	Double list	*1,020*	-0.011	0.076	
	List A	*510*	-0.024	0.127	-0.402
	List B	*510*	0.000	0.117	-0.350
Participated to PrEP demonstration project					
	Double list	*1,026*	0.030	0.096	
	List A	*513*	-0.081	0.144	-0.333
	List B	*513*	0.144	0.138	-0.304

Notes: Data from the 2017 survey only is considered.

Equation 4 is adapted in the following way: $Y_i = \lambda + \beta T_i + \gamma S_i + \alpha T_i \times S_i + \mathbb{1}(List = A) + \varepsilon_i$, with SE clustered at the FSW's level. The significance level of α indicate that there exists a difference in condom use between the two sub-groups. $\hat{\alpha}$ and SE($\hat{\alpha}$) refer to the estimated coefficient and related standard error.

⋄ Information based on medical records (available only for registered individuals).

SE reduction is computed in the following way: $\frac{\text{SE(Double list) - SE(List A)}}{\text{SE(List A)}}$