**eAppendix 1: Size estimates from PLACE 2014 and PLACE 2016**

PLACE uses a venue-based approach to estimate the number of each key population in the sampled municipality directly. Details about methods and estimator used to obtain direct estimates are available elsewhere (47). Briefly, a comprehensive list of public venues (such as bars and clubs) where key populations meet new sexual partners was developed based on interviews with a broad array of community informants (such as taxi drivers, security guards, and shop owners) in each area. A stratified random sample of venues was visited to ascertain the estimated number of each key population at that site between 11pm and 2am on Saturday night. When necessary, this number was further adjusted for the number of venues visited per person during this time frame. Finally, estimates from sampled venues were up-weighted to represent all venues in the municipality based on the comprehensive list of venues for that municipality.

To ensure municipalities with direct size estimates were representative of all municipalities in the country, the PLACE 2016 study selected three municipalities from each of the nine health regions using stratified random sampling. Of the 27 municipalities selected, there were 17 not covered with the 2014 PLACE study and therefore included in PLACE 2016. Additionally, with the help of local HIV program decision makers, we purposively selected three other municipalities thought to have characteristics not well represented in the original sample for a total of 20 municipalities for data collection.

The PLACE study conducted in 2016 differed from the 2014 PLACEs study in that it collected information only at the venue level from knowledgeable informants rather than from a probability sample of people patronizing the venue, and therefore was more rapid and required fewer resources to conduct. In the PLACE 2016 study, information was collected from both people knowledgeable about the venue, such as an owner or manager, and key population members at the venue. These informants were identified by community mobilizers working with the study team.

**eAppendix 2: Direct estimates from municipalities with data collection**

Table 1. Direct estimates by municipality

| Municipality | Data collection | FSW (%) | MSM (%) | Trans (%) |
|---|---|---|---|---|
| 1 | PLACE 2014 | 4.42 | 1.72 | 0.35 |
| 2 | PLACE 2014 | 1.98 | 0.36 | 0.21 |
| 3 | PLACE 2014 | 0.81 | 0.36 | 0.01 |
| 4 | PLACE 2014 | 7.01 | 1.64 | 0.47 |
| 5 | PLACE 2014 | 2.61 | 1.28 | 0.19 |
| 6 | PLACE 2014 | 0.16 | 0.05 | 0.12 |
| 7 | PLACE 2014 | 0.05 | 0.04 | 0.01 |
| 8 | PLACE 2014 | 2.79 | 1.54 | 0.01 |
| 9 | PLACE 2014 | 4.61 | 0.42 | 0.26 |
| 10 | PLACE 2014 | 4.24 | 1.29 | 0.50 |
| 11 | PLACE 2014 | 1.70 | 0.43 | 0.28 |
| 12 | PLACE 2014 | 17.93 | 3.60 | 1.84 |
| 13 | PLACE 2014 | 2.32 | 0.30 | 0.15 |
| 14 | PLACE 2014 | 2.18 | 1.11 | 0.17 |
| 15 | PLACE 2014 | 4.78 | 0.60 | 0.03 |
| 16 | PLACE 2014 | 2.01 | 0.33 | 0.03 |
| 17 | PLACE 2014 | 3.87 | 2.87 | 0.38 |
| 18 | PLACE 2014 | 1.49 | 0.31 | 0.01 |
| 19 | PLACE 2014 | 0.29 | 0.27 | 0.02 |
| 20 | PLACE 2014 | 4.12 | 2.22 | 0.19 |
| 21 | PLACE 2014 | 1.38 | 0.16 | 0.03 |
| 22 | PLACE 2014 | 1.91 | 0.03 | 0.06 |
| 23 | PLACE 2014 | 1.31 | 0.32 | 0.16 |
| 24 | PLACE 2014 | 3.23 | 0.81 | 0.32 |
| 25 | PLACE 2014 | 2.82 | 1.41 | 0.18 |
| 26 | PLACE 2014 | 2.37 | 0.48 | 0.11 |
| 27 | PLACE 2014 | 1.24 | 0.38 | 0.08 |
| 28 | PLACE 2014 | 1.81 | 0.28 | 0.10 |
| 29 | PLACE 2014 | 0.98 | 0.82 | 0.04 |
| 30 | PLACE 2014 | 1.19 | 0.24 | 0.05 |
| 31 | PLACE 2016 | 6.40 | 1.44 | 0.22 |
| 32 | PLACE 2016 | 1.73 | 0.74 | 0.07 |
| 33 | PLACE 2016 | 2.08 | 0.54 | 0.03 |
| 34 | PLACE 2016 | 5.82 | 1.47 | 0.29 |
| 35 | PLACE 2016 | 3.22 | 1.12 | 0.13 |
| 36 | PLACE 2016 | 5.86 | 0.55 | 0.16 |
| 37 | PLACE 2016 | 1.34 | 2.29 | 0.04 |
| 38 | PLACE 2016 | 3.13 | 0.44 | 0.02 |

| | | | | |
|---|---|---|---|---|
| 39 | PLACE 2016 | 5.03 | 1.06 | 0.02 |
| 40 | PLACE 2016 | 2.64 | 1.02 | 0.31 |
| 41 | PLACE 2016 | 2.68 | 0.79 | 0.01 |
| 42 | PLACE 2016 | 10.45 | 1.66 | 0.15 |
| 43 | PLACE 2016 | 5.78 | 2.45 | 0.08 |
| 44 | PLACE 2016 | 1.57 | 0.03 | 0.03 |
| 45 | PLACE 2016 | 4.90 | 2.38 | 0.21 |
| 46 | PLACE 2016 | 0.02 | 0.12 | 0.02 |
| 47 | PLACE 2016 | 2.92 | 0.98 | 0.13 |
| 48 | PLACE 2016 | 13.73 | 3.66 | 0.27 |
| 49 | PLACE 2016 | 8.38 | 1.51 | 0.10 |
| 50 | PLACE 2016 | 4.32 | 1.09 | 0.02 |

Table 2. Summary statistics describing direct estimates over all municipalities

| Population | Mean | Median | Standard deviation | Range |
|---|---|---|---|---|
| FSW | 3.59 | 2.66 | 3.34 | 0.02 – 17.93 |
| MSM | 1.02 | 0.80 | 0.89 | 0.03 – 3.66 |
| Trans | 0.17 | 0.12 | 0.27 | 0.01 – 1.84 |

**eAppendix 3: Code for analysis of single simulated dataset**

```
%let i=300; *say 300 total municipalities;

*generate data;
data a;
      *set parameters;
      b1=-1.5; b2=1.5; b3=0.02; b1y=2.3; b2y=3; b3y=1.2;
      *set seed;
      call streaminit(323);*(4432);*(343);
            do i=1 to &i;
                  z1=rand("bernoulli", 0.3);
                  z2=rand("bernoulli", 0.75);
                  z3=rand("bernoulli", 0.2);
                  ps=1/(1+exp(-(-2-b1*0.3-b2*0.75+b1*z1+b2*z2)));
                  s=rand("bernoulli", ps);
                  pmu=1/(1+exp(-(-3.85-b1y*0.3-b2y*0.75-
                  b3y*0.2+b1y*z1+b2y*z2+b3y*z3)));
                  n=floor((rand("normal")**2)*2000+800);
                  y=rand("binomial",pmu, n);
                  logn=log(n);
                  output;
            end;
run;
data a; set a; drop b1-b3; run;

*full data;
ods select none;
proc genmod data=a;
      model y=/link=log dist=poisson offset=logn;
      ods output parameterestimates=full;
run;
data full; set full; if parameter="Intercept";
      muhat=exp(estimate);
      lcl=exp(LowerWaldCl);
      ucl=exp(UpperWaldCl);
      keep muhat lcl ucl;
run;

*complete case;
proc genmod data=a;
      where s=1;
      model y=/link=log dist=poisson offset=logn;
      ods output parameterestimates=cc;
run;
data cc; set cc; if parameter="Intercept";
      muhat=exp(estimate);
      lcl=exp(LowerWaldCl);
      ucl=exp(UpperWaldCl);
      keep muhat lcl ucl;
run;

*IPSW;
proc logistic data=a desc noprint; model s=z1 z2; output out=d p=d; run;
proc logistic data=a desc noprint; model s=; output out=n p=num; run;
```

```
data b; merge a d n; by i; if d<0.0001 then d=0.0001; w=1/d; sw=num/d; run;
proc genmod data=b;
      weight sw;
      where s=1;
      class i(param=ref ref=first);
      model y=/link=log dist=poisson offset=logn;
      repeated subject=i;
      ods output GEEEmpPEst=weighted;
run;
data weighted; set weighted; if parm="Intercept";
      muhat=exp(estimate);
      lcl=exp(LowerCl);
      ucl=exp(UpperCl);
      keep muhat lcl ucl;
run;

*MI - Adapted from Cole SR, Chu H, Greenland S. Multiple-imputation for
measurement-error correction. Int J Epidemiol. 2006;35(4):1074-1081;

%let m=20; *20 imputations;
data m; set _NULL_; run;
proc genmod data=a;
      where s=1;
      model y=z1 z2 /link=log dist=poisson covb offset=logn;
      ods output parameterestimates=est covb=covall;
run;
data bb; set est; keep parameter estimate; run;
proc transpose data=bb out=bb2 prefix=b; run;
data bb2; set bb2; drop b4 _NAME_; run;
data cov; set covall; keep prm1-prm3;run;
*use proc IML to draw from multivariate normal;
proc iml;
      use cov; read all into cov;
      use bb2; read all into mu; mu=mu`;
      v=nrow(cov);
      n=&m;
      seed=222;
      l=t(root(cov));
      z=normal(j(v,n,seed));
      x=l*z;
      x=repeat(mu,1,n)+x;
      tx=t(x);
      create m from tx;
      append from tx;
quit;

data m ;
      set m;
      array col {3} col1-col3; array b {3} b1-b3;
      retain _imputation_ 0;
      _imputation_=_imputation_+1;
      do u=1 to 3;
            b[u]=col[u];
      end;
      drop col1-col3 u;
run;
```

```
data m; set m;  keep _imputation_ b1-b3; run;

data aa; set a; do _imputation_=1 to &m; output; end; run;
proc sort data=m; by _imputation_; run;
proc sort data=aa; by _imputation_; run;
data mi; merge aa m; by _imputation_; run;

data mi; set mi;
      call streaminit(320);
      *impute yhat for nonsampled municipalities only (otherwise yhat=y);
      if s=1 then yhat=y; else yhat=exp(logn+b1+b2*z1+b3*z2);
run;

proc genmod data=mi ;
      by _imputation_;
      model yhat=/link=log  dist=poisson offset=logn;
      ods output parameterestimates=imp;
run;
data imp; set imp; if parameter="Intercept"; run;
proc mianalyze data=imp;
      modeleffects Estimate;
      stderr StdErr;
      ods output parameterestimates=miest;
run;
data mi_est;
      set miest;
      muhat=exp(estimate);
      lcl=exp(lclmean);
      ucl=exp(uclmean);
      keep muhat lcl ucl;
run;

*Augmented IPSW;
%let rep=1000; *for bootstrapped variance;
ods select none;
proc surveyselect data=b out=bootc outhits seed=123 method=urs samprate=1
rep=&rep noprint; run;
data bootcc; set b(in=b0) bootc; if b0=1 then b=0; else b=replicate; drop
replicate numberhits; run;

data b2; set bootcc; if s=0 then y=.; run;
proc genmod data=b2;
      by b;
      weight w;
      model y=z1 z2/link=log dist=poisson offset=logn ;
      output out=drw p=yhat;
run;
proc genmod data=drw ;
      by b;
      model yhat= /link=log dist=poisson offset=logn ;
      ods output parameterestimates=regw;
run;
data regw0; set regw;
      where parameter="Intercept";
      if b=0 then do; muhat=exp(estimate); output; end;
      keep muhat;
run;
```

```
proc means data=regw; where b>0 and parameter="Intercept"; var estimate;
output out=cis std=se; run;
data regwci;
      set regw0;
      if _N_=1 then set cis;
      lcl=exp(log(muhat)-1.96*se); ucl=exp(log(muhat)+1.96*se);
      keep muhat lcl ucl;
run;
data full; set full; sc="Full"; run;
data cc; set cc; sc="Complete case"; run;
data weighted;set weighted; sc="IPSW"; run;
data mi_est; set mi_est; sc="MI"; run;
data regwci; set regwci; sc="Augmented IPSW";  run;

data ests;
      length sc $ 15;
      set full cc weighted mi_est regwci;
      keep sc muhat lcl ucl;
run;

*output results;
ods select all;
proc print data=ests; run;
```