

2021 TUS-CPS Data User Webinar Series

Linkage of the TUS-CPS to ASEC CPS Supplement

Carolyn Reyes-Guzman, PhD, MPH

Program Director & Epidemiologist

NCI, Tobacco Control Research Branch

Disclaimer

The views and opinions expressed are the presenters' and do not necessarily represent the views, official policy, or position of the U.S. Government, U.S. Department of Health and Human Services or any of its affiliated institutions or agencies.

Agenda

Questions?

Please type into the chat box and moderators will ask them at the end of the webinar

- 1. How do TUS and ASEC fit in CPS?*
- 2. Weights in TUS and ASEC*
- 3. Linking TUS and ASEC CPS Supplements*
- 4. Applied example*
 - Smoking and vaping home rules, and smoke-free policy attitudes among residents of multi-unit housing*

How do TUS and ASEC fit in CPS?

What is TUS-CPS?

- TUS is another CPS supplement, every 3-4 yrs, with 3 time points each cycle
- Survey cycle typically fielded for July (Yr 1) + January & May (Yr 2)
- TUS data can be used by researchers to:
 - Monitor tobacco control progress and assess long-term cross-sectional population trends;
 - Track tobacco health disparities;
 - Evaluate tobacco control programs; and
 - Examine national, state and county*-level data

*where available in these smaller geographical unit areas

How ASEC fits into General CPS

- ASEC = Annual Social and Economic **Supplement** (“March”) of CPS
- Universe: civilian noninstitutional U.S. population + members of the Armed Forces living off post or living with their families on post, when ≥ 1 civilian adult lives in the same household
 - (TUS-CPS only includes civilian noninstitutional U.S. population)
- Housing units selected using a probability sample, similar to CPS & TUS
- States, regions and divisions are identified in their entirety; some county level data available for some metropolitan areas, similar to TUS*
- Captures detailed employment information among persons ≥ 15 yrs old

*ASEC: 1,300 out of 3100 U.S. counties included in the sample. <https://www.census.gov/programs-surveys/saipe/guidance/model-input-data/cpsasec.html>

CPS	ASEC
CPS monthly sample size ~60,000 occupied households each month	ASEC annual sample size ~98,000 households; includes oversample of Hispanic households (6,500)
Design: multi-stage stratified sample of housing units; households follow 4-8-4 “panel” design, allowing year-to-year comparisons of same household	Design: multistage stratified sample of housing units: hierarchical file structure, with three record types present: Household, Family, and Person
Constant sample replenishment without excessive respondent burden	Annual supplement, so no 4-8-4 design; allows for annual time series comparisons

https://cps.ipums.org/cps/sample_designs.shtml

<https://www.census.gov/topics/population/foreign-born/guidance/cps-guidance/cps-vs-asec.html>

<https://www.icpsr.umich.edu/web/NADAC/studies/37652>

What is ASEC useful for?

- Provides monthly labor force data, and supplemental data on work experience, income, noncash benefits, housing and migration
- Work experience: employment status, occupation, and industry
 - Weeks + hours per week worked, reason not working full-time, total income and supplemental income components
- Noncash benefits: food stamps, school lunch program, employer-provided group health insurance plan, employer-provided pension plan, personal health insurance, Medicaid, Medicare, or military health care, and energy assistance
- Employment & income questions refer to previous calendar year

Weights in TUS and ASEC

ASEC File Structure + Weights

- As of 2019, ASEC data available in single ASCII (.dat), CSV and SAS files (CSV and SAS files split into 1 file each for HH, PP, FM records)
- ASEC weights account for file structure: a supplement household weight, a supplement family weight, and a supplement person weight
- The final weight is used to produce population estimates for the various items covered in the monthly CPS
- Differences in the questionnaire, sample and data uses for the CPS ASEC Supplement are factored into the ASEC Supplement weight

Dealing With TUS-CPS Weights

- The 2018-19 dataset includes a full-sample weight for self respondents (SmplWgt)
- “Full sample weights are created to compensate for differential selection probabilities, nonresponse, and under-coverage of the target population of U.S. adults”¹
- But there is a second type of weights available for the TUS-CPS too –
 - “Replicate weights, which can accommodate various types of statistical analyses, are created to more accurately estimate standard errors by accounting for the complex survey design”¹

[2018-2019 Tobacco Use Supplement to the Current Population Survey \(TUS-CPS\) User Guide for Conducting Weighted Analyses](#)

Linking TUS-CPS File to Replicate Weights

- Reminder of TUS-CPS 2018-19 data: 160 replicate weights
- To run analyses, first need to merge TUS respondents with their corresponding replicate weights by ID
- “ID” variable in survey data file is unique within survey year and survey month (see slides 18-19)
- Same ID variable was created for replicate weight files
- Files merged by SurYear, SurMonth and ID

Do You Also Need to Use ASEC Weights?

- Depends on your research question
 - Users should examine the overlap between TUS and ASEC with respect to the core (general) CPS variables they will examine (e.g., demographic vars)
 - If the % of matching in the TUS-ASEC samples is questionable, users should recalculate a new weight using both TUS-CPS and ASEC weights
 - There are also replicate weights in ASEC to consider if users need to recalculate weights

Analyses Using TUS-CPS Replicate Weights

- When conducting variance estimation for the TUS-CPS with replicate weights, Fay's method is applied (a variation of balanced repeated replication, or BRR)
- When linking TUS to other CPS supplements, use **Fay=0.75**, SAS; **ADJFay=16**, SUDAAN

Linking TUS and ASEC CPS Supplements

Assumptions

- Since we were mostly interested in tobacco use questions within the TUS, with respect to housing questions in ASEC, we first examined the proportion of our overlap sample on key core sociodemographic variables in both ASEC and TUS to determine if a weight recalculation was warranted

Steps in Linking TUS + ASEC

- Merge 2019 January and May TUS-CPS survey files (not interested in July 2018 TUS because of longer time window with March 2019 ASEC)
- Merge 2019 January and May TUS-CPS replicate weight files
- Combine TUS survey + TUS replicate weight data
- Merge ASEC household + person files using household ID
- Merge ASEC and TUS files
- Create match variables for sociodemographic variables in both TUS and ASEC to examine overlap of samples
- Output matched dataset and create data steps in final merged file

Linking Select TUS-CPS Months to Replicate Weights*

1

```
options ls=255 ps=55 nocenter nofmterr validvarname=upcase mprint; ods html close; ods html;
```

```
Proc Format;
```

```
Value PrtAgeF  
  18-34 = "18-34"  
  35-44 = "35-44"  
  45-54 = "45-54"  
  55-HIGH = "55+";
```

```
Libname TUS "data\";
```

```
filename ASEC2019 "L:\...\asec2019_pubuse.dat" lrecl=1092;
```

```
/* read in two time points from TUS 2019 that have overlap with ASEC 2019 */
```

```
Data TUSCPS19;
```

```
Set TUS.CPSJAN19 TUS.CPSMAY19;
```

```
If PRPerTyp=2 & PrtAge>18 & Intrview=1 & PRS64=1; *adult civilian records, self respondents;
```

```
If (HRMonth=1 & HRMIS In (1,2,5,6)) | (HRMonth=5 & HRMIS In (3,4,7,8)); *January is HRMonth =1, May is  
HRMonth=5;
```

```
If HRMonth=1 Then H_MIS=HRMIS+2; *matching with March data if Jan = 1 then add 2 months for March overlap;
```

```
Else If HRMonth=5 Then H_MIS=HRMIS-2; *matching with March data by subtracting 2 months from next panel;
```

```
PERIDNUM=HRHHID||HRHHID2||Put(PULineNo,Z2.); *person id based on 2 household ids;
```

```
→ ID=catx("_", HRYEAR4, HRMONTH, QSTNUM, OCCURNUM);
```

```
Run;
```

```
Proc Sort Data=TUSCPS19;
```

```
→ By ID; Run;
```

*Before this step, download and read in Jan and May 2019 TUS-CPS data + replicate weights with SAS programs: <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data>

Linking Select TUS-CPS Months to Replicate Weights

2

```
FileName Rep1901 "L:\...\jan19srrep.dat" lrecl=1617;  
FileName Rep1905 "L:\...\may19srrep.dat" lrecl=1617;
```

```
%Macro Rep19(File);  
  Data &File;  
    Infile &File;  
    Input @001 QstNum 5.  
          @006 OccurNum 2.  
          @008 SmplWgt 10.  
          @018 (repwgt1-repwgt160) (10.);  
  If SmplWgt=0 & (repwgt1=0 | repwgt1=.) Then Delete;
```

```
  Data &File;  
    Set &File;  
    Array Wgts(161) SmplWgt repwgt1-repwgt160;  
    Do I = 1 to 161;  
      Wgts(I)=Wgts(I)/10000;  
    End;  
    Drop I;
```

```
%Mend Rep19;
```

```
%Rep19(Rep1901); Run;  
%Rep19(Rep1905); Run;
```

```
  Data Repls19;  
    Set Rep1901(In=In1)  
        Rep1905(In=In2);  
    HRYear4=2019;  
    If In1 Then HRMonth=1;  
    Else If In2 Then HRMonth=5;  
    ID=catx("_", HRYEAR4, HRMONTH, QSTNUM, OCCURNUM);  
  Run;  
  
  Proc Sort Data=Repls19;  
    By ID;  
  Run;
```

Linking Select TUS-CPS to ASEC Files

3

```
Data TUSTotal (drop = /* include variables to drop */);
Merge TUSCPS19(In=In1)
      Reps19(In=In2);
By ID;
If In1;
Reps=In2;
Run;

Proc Freq Data=TUSTotal;
Table Reps/Missing;
Run;

/* Read in 2019 ASEC data */
Data HH_2019(Keep=H_SEQ HUnits H_HHType H_MIS H_HHNum
             H_TENURE HPUBLIC HLORENT HUNDER18 H_IDNUM)
      Per_2019(Keep=PH_SEQ A_LineNo A_Age A_Sex GREG
              GTMETSTA HEFAMINC A_HGA A_MARITL PEMLR PRDRaceA
              PERIDNUM );

Infile ASEC2019; *household and person data all in 1
ASCII public use file;
Input @001 RecType 1. @;
If RecType=1 Then Do;
```

```
Input @008 H_HHNUM 1.
      @009 H_IDNUM $Char20.
      @029 H_SEQ 5.
      @061 H_HHType 1.
      @064 H_MIS 1.
      @076 HUNDER18 2.
      @078 HUnits 1.
      @089 H_TENURE 1.
      @326 HLORENT 1.
      @327 HPUBLIC 1.;
Output HH_2019;
End;
```

```
Else If RecType=3 Then Do;
Input @002 A_LineNo 2.
      @012 PERIDNUM $Char22.
      @036 PH_SEQ 5.
      @043 GREG 1.
      @060 GTMETSTA 1.
      @065 HEFAMINC 2.
      @079 A_Age 2.
      @087 A_HGA 2. /* education */
      @090 A_MARITL 1.
      @092 A_Sex 1.
      @153 PRDRaceA 2.
      @271 PEMLR 1.; /* labor code */
Output Per_2019;
End;
Run;
```

Linking Select TUS-CPS to ASEC Files

4

```
Data HH_2019;  
  Length HHID $ 20;  
  Set HH_2019;  
  HHID=H_IDNUM;  
Run;
```

```
Data Per_2019;  
  Length HHID $ 20;  
  Set Per_2019;  
  HHID=SubStr(PERIDNUM,1,20);  
Run;
```

```
Proc Sort Data=HH_2019;  
  By HHID;  
Run;
```

```
Proc Sort Data=Per_2019;  
  By HHID A_LineNo;  
Run;
```

```
Data ASEC2019;  
  Merge HH_2019(In=In1)  
        Per_2019(In=In2);  
  By HHID; *household id variable ;  
  If In1 & In2;  
Run;
```

```
Proc Sort Data=ASEC2019;  
  By H_MIS PERIDNUM;  
Run;
```

```
Proc Sort Data=TUStotal;  
  By H_MIS PERIDNUM;  
Run;
```



How we checked % TUS-ASEC overlap sample

5

- Linked TUS-CPS to ASEC files and created key “match” variables

```
Data Together;
Merge TUStotal (In=In1)
      ASEC2019 (In=In2);
By H_MIS PERIDNUM;
TUS_Rec=In1;
ASEC_Rec=In2;
If In1 & In2 Then Do;
  If PEXsex=A_Sex Then Match_Sex=1;
Else Match_Sex=0;
  If PrtAge=A_Age Then Match_Age=1;
Else If (PrtAge=(A_Age+1)) | (PrtAge=(A_Age-1)) Then Match_Age=2;
Else Match_Age=0;
  If PTDTRACE=PRDTRACEA Then Match_Race=1;
Else Match_Race=0;
End;
Run;
```

Output Final TUS-ASEC dataset

```
/* output final dataset applying match flags for analyses */  
  
data final;  
set together;  
if TUS_Rec=1 & ASEC_Rec=1 & Match_Sex=1 & Match_Age In (1,2) & Match_Race=1;  
  
/*recode demographic variables of interest */  
run;  
  
Proc Tabulate Data=Final;  
  Title2 "TUS-ASEC merged dataset";  
  Class HRYear4 PrtAge;  
  Var PWSRWgt;  
  Table PrtAge="TUS Age" All="Total",  
        HRYear4="Year"* (N="Sample Size"*F=Comma8.  
          PWSRWgt=" "* (Sum="Population"*F=Comma12. PctSum<PrtAge  
All>="Percent"*F=7.2) );  
  Format PrtAge PrtAgeF.;  
Run;
```

How we checked % TUS-ASEC overlap sample


- Check key “match” variables (note: code in SAS-callable SUDAAN)

```
proc crosstab data=together filetype=sas design=BRR;
Weight SmplWgt; * this is the TUS self-response weight;
RepWgt repwgt1-repwgt160/ADJFay=16;
Class Suryear Match_Sex Match_AGE Match_RACE;
Tables Suryear*(Match_Sex Match_AGE Match_RACE);
Print   NSum = "Sample Size"
        WSum = "Population Size"
        RowPer = "Percent"
        SERow = "Standard Error"
        LowRow = "Lower 95% CI"
        UpRow = "Upper 95% CI"
        /NSumFmt=F8.0 WSumFmt=F10.0;


Run;
```


How we checked % TUS-ASEC overlap sample


Age match in TUS and ASEC

MATCH_AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	142755	77.58	142755	77.58
 0	249	0.14	143004	77.72
1	33473	18.19	176477	95.91
2	7525	4.09	184002	100.00

Race match in TUS and ASEC

MATCH_RACE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	142755	77.58	142755	77.58
 0	69	0.04	142824	77.62
1	41178	22.38	184002	100.00

Sex match in TUS and ASEC

MATCH_SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
.	142755	77.58	142755	77.58
 0	76	0.04	142831	77.62
1	41171	22.38	184002	100.00

How we checked % TUS-ASEC overlap sample

- Next, we checked distributions for these key “match” variables in both merged TUS-ASEC and TUS-only datasets (e.g., age group)

```
proc crosstab data=final filetype=sas design=BRR;
Weight SmplWgt;
RepWgt repwgt1-repwgt160/ADJFay=16;
Class Suryear PEXsex NewAGE NewRACE;
Tables Suryear*(PEXsex NewAGE NewRACE);
Print   NSum = "Sample Size"
        WSum = "Population Size"
        RowPer = "Percent"
        SERow = "Standard Error"
        LowRow = "Lower 95% CI"
        UpRow = "Upper 95% CI"
        /NSumFmt=F8.0 WSumFmt=F10.0;
Run;
```

How we checked % TUS-ASEC overlap sample

- TUS-only dataset

SURYEAR		AGE_TUS				
		Total	1	2	3	4
Total	Sample Size	45148	9548	7364	6821	21415
	Population Size	250517589	74524958	40810638	40645064	94536929
	Percent	100.00	29.75	16.29	16.22	37.74
	Standard Error	0.00	0.13	0.03	0.03	0.07
	Lower 95% CI	.	29.49	16.23	16.16	37.60
	Upper 95% CI	.	30.01	16.35	16.29	37.88
	2019	Sample Size	45148	9548	7364	6821
	Population Size	250517589	74524958	40810638	40645064	94536929
	Percent	100.00	29.75	16.29	16.22	37.74
	Standard Error	0.00	0.13	0.03	0.03	0.07
	Lower 95% CI	.	29.49	16.23	16.16	37.60
	Upper 95% CI	.	30.01	16.35	16.29	37.88

How we checked % TUS-ASEC overlap sample

- TUS-ASEC merged dataset

SURYEAR		Age Group				
		Total	18-34	35-44	45-54	55+
Total	Sample Size	40909	8199	6599	6203	19908
	Population Size	224852125	63812953	36485875	36854992	87698305
	Percent	100.00	28.38	16.23	16.39	39.00
	Standard Error	0.00	0.25	0.15	0.14	0.22
	Lower 95% CI	.	27.89	15.93	16.12	38.56
	Upper 95% CI	.	28.88	16.53	16.67	39.45
	2019	Sample Size	40909	8199	6599	6203
	Population Size	224852125	63812953	36485875	36854992	87698305
	Percent	100.00	28.38	16.23	16.39	39.00
	Standard Error	0.00	0.25	0.15	0.14	0.22
	Lower 95% CI	.	27.89	15.93	16.12	38.56
	Upper 95% CI	.	28.88	16.53	16.67	39.45

Applied Example

Smoking and vaping home rules, and smoke-free policy attitudes among residents of multi-unit housing

- 2019 (March) ASEC Public Use ASCII Data¹ matched to 2019 TUS-CPS² (January + May surveys)

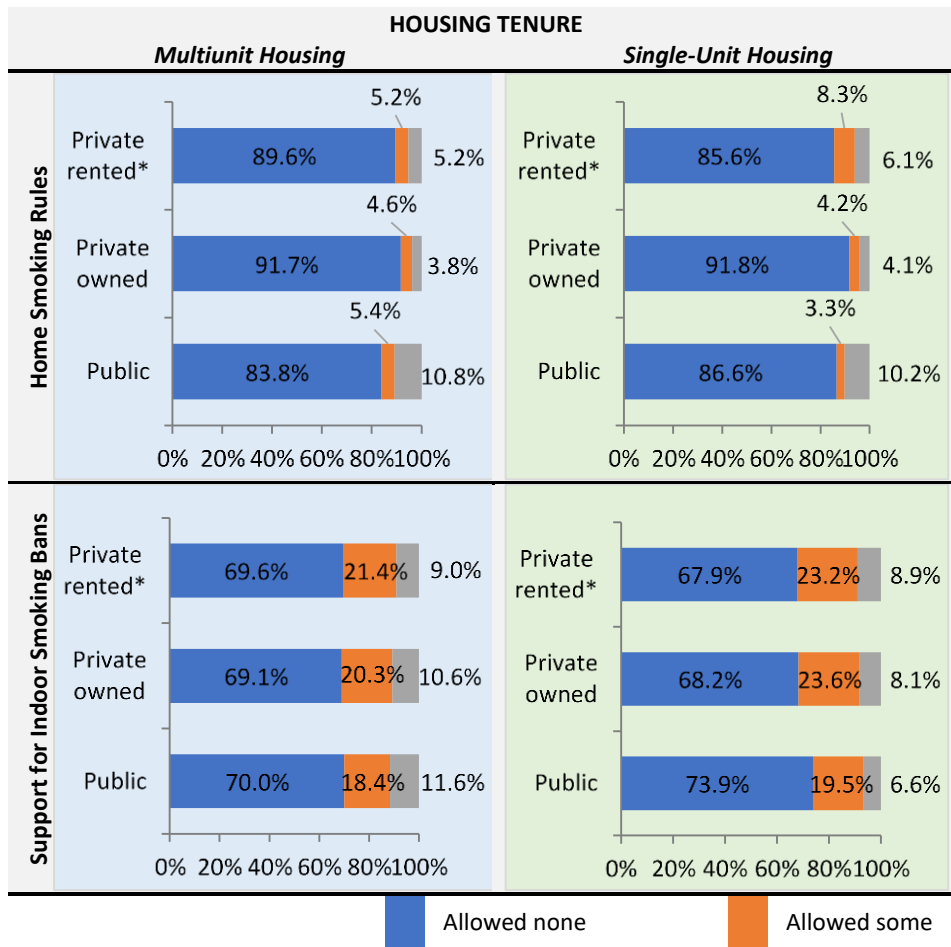
	MULTI-UNIT HOUSING					SINGLE-UNIT HOUSING				
Variable	N=59,910,550	n=9,662	Current Smoker n=1,151 (10.8%)	Former smoker n=1,850 (15.6%)	Never smoker n=6,661 (73.6%)	N=162,349,276	n=30,634	Current Smoker n=3,484 (10.7)	Former smoker n=6,797 (19.3)	Never smoker n=20,353 (70.0)
HOUSING TENURE										
Public	5,182,886	905	188 (15.5)	194 (9.5)	523 (7.5)	1,082,507	198	44 (1.2)	42 (0.6)	112 (0.6)
Private owned	15,278,762	2,659	203 (17.6)	638 (33.2)	1,818 (25.0)	131,991,591	25,684	2,522 (69.2)	5,926 (85.5)	17,236 (82.0)
Private rented*	39,448,902	6,098	760 (66.9)	1,018 (57.4)	4,320 (70.7)	29,275,178	4,752	918 (29.6)	829 (13.9)	3,005 (17.4)

¹<https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.2019.html>

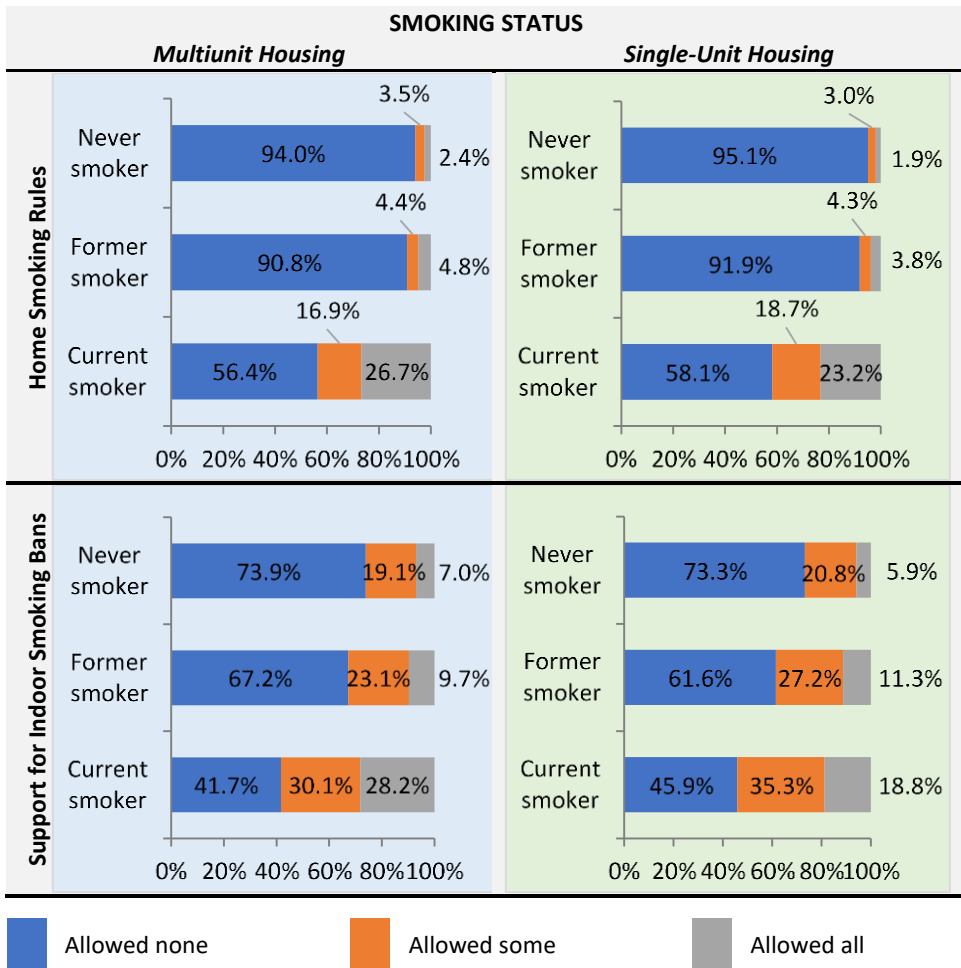
²<https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data>

*Rent includes no cash rent or rent without payment

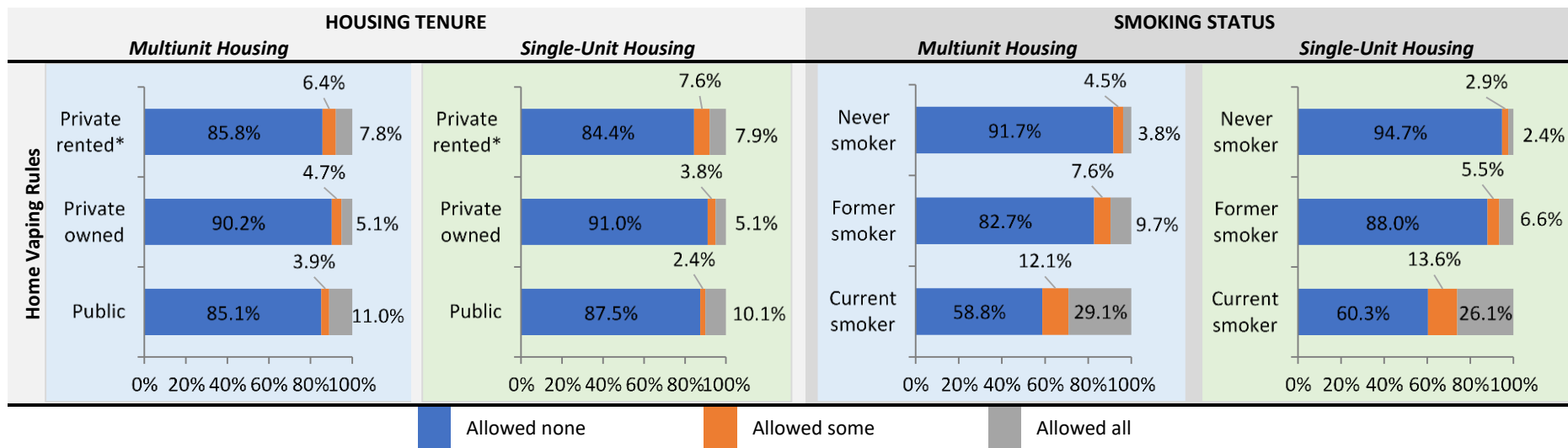
Smoking and vaping home rules, and smoke-free policy attitudes by housing tenure and smoking status¹



Smoking and vaping home rules, and smoke-free policy attitudes by housing tenure and smoking status¹



Smoking and vaping home rules, and smoke-free policy attitudes by housing tenure and smoking status¹



Acknowledgements



Tobacco Use Supplement
Current Population Survey

- Beth Slotman & Harold Bobbitt (Westat/NCI Contract)
- Todd Gibson (IMS/NCI Contract)
- Nalini Corcy (BLH/NCI Contract)
- Dana Chomenko (BLH/NCI Contract)
- Minal Patel (Truth Initiative)
- Maggie Mayer (NCI)
- Kristen Mangold (NCI)
- Benmei Liu (NCI)

TUS-CPS Website - <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps>



**NATIONAL
CANCER
INSTITUTE**

www.cancer.gov

www.cancer.gov/espanol