Modifying State Sample Sizes for the National Crime Victimization Survey

Samantha Spiers, Sandra Peterson, and David Hornick, U.S. Census Bureau, Demographic Statistical Methods Division

Presented at the 2019 Joint Statistical Meeting
Denver, CO
July 27 – August 1, 2019

Background

- The National Crime Victimization Survey (NCVS) is a nationally representative household survey that collects information on criminal victimization, both reported and not reported to police.
- In 2016, the NCVS sample was redesigned and increased to support state estimates for the largest 22 states (called "boost states") using three years of data.
- The NCVS was designed with the following goals:
  - National 1-year violent crime estimate with no more than a 3.68% CV
  - State-level 3-year violent crime estimates with no more than 10% CVs
- The sample was designed with these assumptions:
  - National and state-level violent crime rates are 2%.
  - State-level design effects (DEFFs) are the same as the national DEFFs from 1996-2010 NCVS data.
    - DEFF calculations did not include all instances of repeated crimes (called "series crimes").

Research Objectives

- Recalculate target person interviews using 2016-2018 NCVS.
- Compare sample sizes using only national assumptions to those using state-specific assumptions.
- Examine the effect of series crimes on the DEFFs and sample sizes.

Methods

We calculated sample sizes under two scenarios:

Scenario 1. Only national assumptions (2% national and state-level crime rate, national DEFFs)

Scenario 2. National and state-specific assumptions (2% national crime rate, state-specific rates from 2016-2018 NCVS, national and state-specific DEFFs)

Methods (continued)

Design Effects: $deff = {}^{v(\hat{p})}/_{v_{SRS}(\hat{p})} = v(\hat{p}) / \frac{\hat{p}(1-\hat{p})}{n}$

where $\hat{p}$ is the crime rate, $v(\hat{p})$ is the variance, $v_{SRS}(\hat{p})$ is the variance under SRS, and n is the half-year person interviews.

Table 1: National 1-Year and 3-Year Design Effects

	Including Series Crimes		Not Including Series Crimes	
Estimate	**SR**	**NSR**	**SR**	**NSR**
1-year*	4.98	7.22	2.34	2.99
3-year	4.92	6.86	2.34	4.18

Source: U.S. Census Bureau internal data from the 2016-2018 NCVS

Note: SR = Self-Representing Primary Sampling Unit (PSU), NSR = Non-Self-Representing PSU

* 1-year design effects are an average of the design effects from the 2016, 2017, and 2018 NCVS

Table 2: State-Level 3-Year Design Effects

	Including Series Crimes			Not Including Series Crimes		
PSU	**Mean (SD)**	**Min**	**Max**	**Mean (SD)**	**Min**	**Max**
SR	4.96 (2.88)	1.77	11.71	2.14 (0.38)	1.61	2.96
NSR	7.59 (7.92)	0.11	27.11	3.80 (3.28)	0.14	11.50

Source: U.S. Census Bureau internal data from the 2016-2018 NCVS

Sample Sizes:

1. Calculate national half-year person interviews needed.

$$n = \frac{(1-\hat{p}_{U.S.})}{CV^2 \hat{p}_{U.S.}} [W_{SR} deff_{SR} + W_{NSR} deff_{NSR}]$$

where $\hat{p}_{U.S.}$ = 0.02, CV = 0.0368, W_{SR} = 0.66, and W_{NSR} = 0.34.

2. Allocate n proportionally to all states.
3. Evaluate if enough sample in boost states to meet 10% CV goal.

$$CV_k = \frac{\sqrt{v(\hat{p}_k)}}{\hat{p}_k} \text{ where } v(\hat{p}_k) = (W_{SR,k})^2 \left[deff_{SR,k} \frac{\hat{p}_{SR,k}(1-\hat{p}_{SR,k})}{3n_{SR,k}}\right] + (W_{NSR,k})^2 \left[deff_{NSR,k} \frac{\hat{p}_{NSR,k}(1-\hat{p}_{NSR,k})}{3n_{NSR,k}}\right]$$

where $W_{SR,k}/W_{NSR,k}$ is the proportion of persons in state k in the SR/NSR strata

4. If needed, increase sample in boost states until CV_k = 10%.

Results

Table 3: Sample Sizes not Including Series Crimes

	Current	Scenario 1	Scenario 2
National	144,100	133,800	145,400
Boost States	126,400	114,400	125,900
Balance	17,700	19,500	19,500

Source: U.S. Census Bureau internal data from the 2016-2018 NCVS

Note: Sample sizes rounded to nearest hundred.

Table 4: Sample Sizes Including Series Crimes

	Current	Scenario 1	Scenario 2
National	144,100	271,900	275,500
Boost States	126,400	228,400	232,000
Balance	17,700	43,600	43,600

Source: U.S. Census Bureau internal data from the 2016-2018 NCVS

Note: Sample sizes rounded to nearest hundred.

- Sample sizes are much larger when series crimes are included due to large DEFFs from larger variance in the violent crime rate.
- When series crimes are not included, Scenario 1 sizes are smaller than the current ones, and Scenario 2 sizes are mostly larger than the current.

Conclusions

- Incorporating series crimes into the sample size calculation produces sizes that are too large to be feasible.
- Sizes increase when using state-specific assumptions, indicating that more sample may be needed to meet state-level CV goals.
- State-level assumptions are limited by being from only one three-year timeframe. We will continue with the current sizes and re-examine them after more state-level data is collected.

References

Fay, R. E., & Li, J. (2012). *Rethinking the NCVS: Subnational goals through direct estimation*. Presented at the 2012 Federal Committee on Statistical Methodology Conference, Washington D.C., January 10-12, 2012.

National Crime Victimization Survey (2017). National Crime Victimization Survey, 2016: Technical documentation. https://www.bjs.gov/content/pub/pdf/ncvstd16/pdf.

United States® Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov