Blending Administrative Data with a Probability Sample of Nonparticipants to Produce National Estimates: The NCS-X NIBRS Estimation Project

#### 2020 Federal Committee on Statistical Methodology

Marcus Berzofsky



www.rti.org RTI International is a trade name of Research Triangle Institute. RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.

#### Background

- The Uniform Crime Reporting (UCR) Program aggregated by the FBI has been the repository of crimes reported to the police for the last 100 years
- Traditionally, police agencies have reported crime in "summary" form aggregate monthly counts
- The FBI is sunsetting "summary reporting system" (SRS) UCR transitioning agencies to the National Incident Based Reporting System (NIBRS)
- NIBRS provides incident-level level information for each reported crime
- But, law enforcement agencies need to upgrade their record management systems in order to submit NIBRS compliant data to the FBI

#### Understanding the Problem: Transitioning to NIBRS



By 2018, **approximately 8,000 out** of 18,000 Law Enforcement Agencies (LEAs) have transitioned to NIBRS.

Reporting agencies are not randomly distributed and skewed towards less populated states

# NCS NATIONAL CRIME STATISTICS EXCHANGE Powering the Transition to NIBRS

Recruiting a probability sample of 400 law enforcement agencies (LEAs)
 who were nonreporting agencies in 2011 to supplement the existing NIBRS
 data

All nonreporting agencies with 750+ police officers were included in these
 400 LEAs

•Combine data from these 400 agencies with data from the 8,000+ existing reporting agencies to produce national estimates



#### NCS-X Sample Agencies (n=400)



Population Served: • 0 • 2,000,000 O 4,000,000 O 6,000,000 O 8,000,000

Agency Type: • 750+ Officer LEAs • Other LEAs

#### **NIBRS Estimation Project: Objectives**

- Determine the optimal method to weight and blend the NCS-X sample with "early adopters" and "late joiners" to produce representative national and state-level estimates
- Validate the estimation methodology



#### Estimation Approach #1: Top-Down Approach

#### **TOP-DOWN APPROACH**

- Develops national weights first
- Uses national weights to develop subnational estimates



### Estimation Approach #1: Top-Down Approach (cont.)

- Advantages
  - The intended approach for the NCS-X sample
  - Most efficient (in terms of precision) approach for national estimates
- Disadvantages
  - Agency weights may represent agencies in different states making state estimates difficult to produce
  - Long term, does not lend itself to subnational estimation



#### Top-Down Approach Weighting/Estimation Strategy



\*72 out of the 400 NCS-X sample agencies are large agencies (750+ officers).

#### **Estimation Options: Intermediate Approach**

#### INTERMEDIATE APPROACH

- Begins with a mix of weights designed for estimation at different levels of geography
- State-level estimates created where able; regional estimates created when needed
- Together state and regional weights can produce national estimates



### Estimation Options: Intermediate Approach (cont.)

- Advantages
  - Provides for state-level estimates sooner than top-down approach
  - Allows most flexibility with the type of subnational areas available for estimation
- Disadvantages
  - Weights are not as efficient as top-down approach, but more efficient than bottom-up approach
  - Subnational estimates will be available on a flow basis



#### Intermediate Approach Weighting/Estimation Strategy

#### Weighting by State



- Weighting by state if coverage ratio exceeds 80%
- For remaining states, produce top-down weights
- Over time, will achieve a 50state design which will allow for simultaneous estimates at the national and state levels

### Weighting Strategy: Naïve Design-Based



- Advantages
  - Closest strategy to the original sample design
  - Theoretically should have the smallest
    amount of bias
- Disadvantages
  - Unclear if original sample design still applicable given level of late joiners since 2011 NCS-X sample drawn
  - Will result in estimates with larger variances than other strategies

### Weighting Strategy: Blending New Joiners



- Advantages
  - Helps smooth the weights compared to the naïve design-based method
  - Since late joiners come from the same pool of agencies the NCS-X sample was drawn they are more likely to be similar to the NCS-X sample agencies than the early adopters
- Disadvantages
  - Deviates from the original sample design
  - Allowing late joiners to represent nonjoiners may introduce bias in the estimates

### Weighting Strategy: Blending All Reporters



• Advantages

- Minimizes the variance in estimates
- Disadvantages
  - Deviates from the original sample design
  - Strategy that is most likely to introduce bias into the estimates

## **Crime Statistics Validation**

- For each blending, weighting, and estimation strategy, we tabulated estimated crime totals and compared against external benchmarks
- Validation Sources:
  - SRS for crime count estimates
  - UCR arrest data for arrest count estimates

Offense Type	Weighted Total among NIBRS Reporters			SRS Total among all LEAs	Relative Bias(%)
	Estimate	SE	RSE	Estimate	
All Crime					
All Violent Crime					
Murder					
Manslaughter					
Rape					
Robbery					
Assault					
Aggravated Assault					
Simple Assault					
All Property Crime					
Burglary					
Larceny					
Vehicle Theft					