Leveraging Incident-Level Police Data with Other Sources

The potential of synthetic data and other approaches

Getting Free of the Summary System

- Index crime classification was a product of the time and the limits of a paper and pencil technology.

- We have been free of these constraints for many years, but nothing has come to replace the crime index crime classification.

- We have demonstrated the enlightenment that comes with the ability to aggregate and disaggregate index crime classes.

- Much greater advances are possible when we begin to blend NIBRS with other data sources in a manner only permissible with incident data.

- This presentation focuses on the merging of data from victimization surveys and NIBRS.
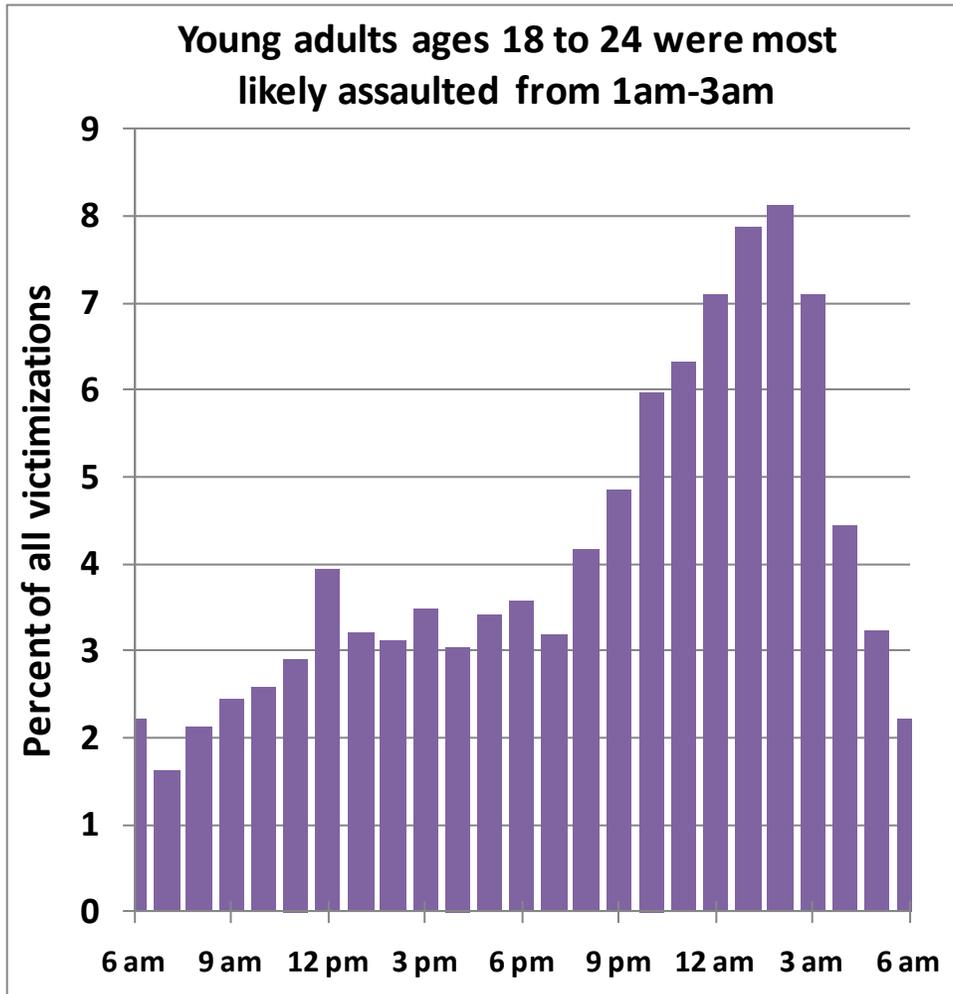
The Limits of Disaggregation

- Demonstrating the enlightenment potential of NIBRS has focused on disaggregation and recombination of attributes of crime events.
  - Other papers in this panel
  - Howard Snyder's rape and sexual assault illustration.
- More enlightenment comes when this ability is combined with data not included in the MIS of police organizations.
  - Police and other consumers are already aware of dis-aggregations in their own data.
  - CIWG emphasized the need to bring other data to bear on
  - Lima illustration—using different denominators.

Uniform Crime Reports Data on Rape 2010

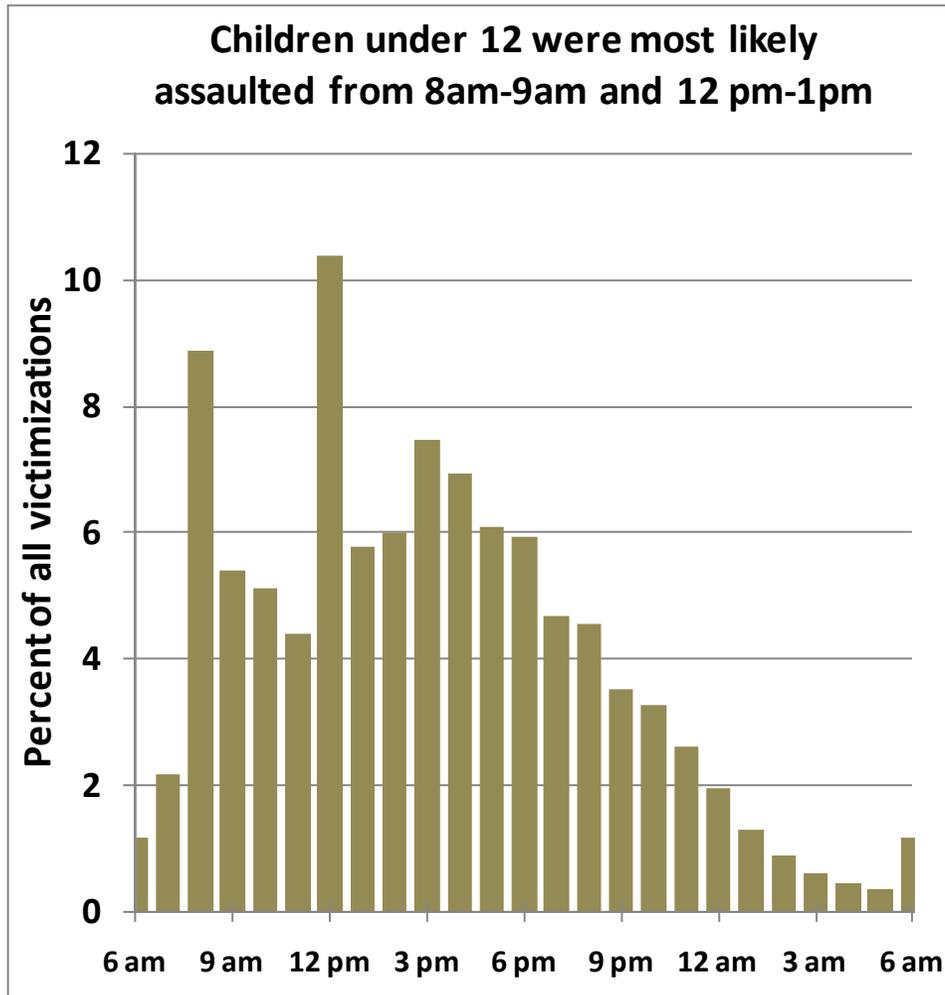| Counts	84,767
Rate Per 100,000 population	27.5

Violent sexual assaults involving young adults ages 18 to 24 years

Young adults ages 18 to 24 were most likely assaulted from 1am-3am

Victims ages 18 to 24 assaulted in a residence					
Offender Age	Intimate	Family	Acquaintance	Stranger	Total
7 to 11	0	0	0	0	1
12 to 17	1	2	16	2	21
18-24	46	7	**184**	21	258
25-34	30	12	**131**	18	191
35-44	6	21	**59**	8	94
45-54	2	18	31	4	54
55-64	0	7	9	1	17
65+	0	4	5	0	10
Total	**86**	**71**	**436**	**54**	**647**
Victims ages 18 to 24 assaulted in a non-residence					
Offender Age	Intimate	Family	Acquaintance	Stranger	Total
7 to 11	0	0	0	0	0
12 to 17	1	0	12	2	15
18-24	9	1	**81**	27	119
25-34	5	1	**64**	35	104
35-44	2	2	39	20	63
45-54	0	2	24	11	37
55-64	0	1	6	2	9
65+	0	0	4	2	6
Total	**17**	**7**	**230**	**99**	**353**
Overall Total	**103**	**78**	**666**	**153**	**1000**

Source: BJS analysis of : Federal Bureau of Investigation, National Incident Based Reporting System, Full reporting states 2006 to 2008.
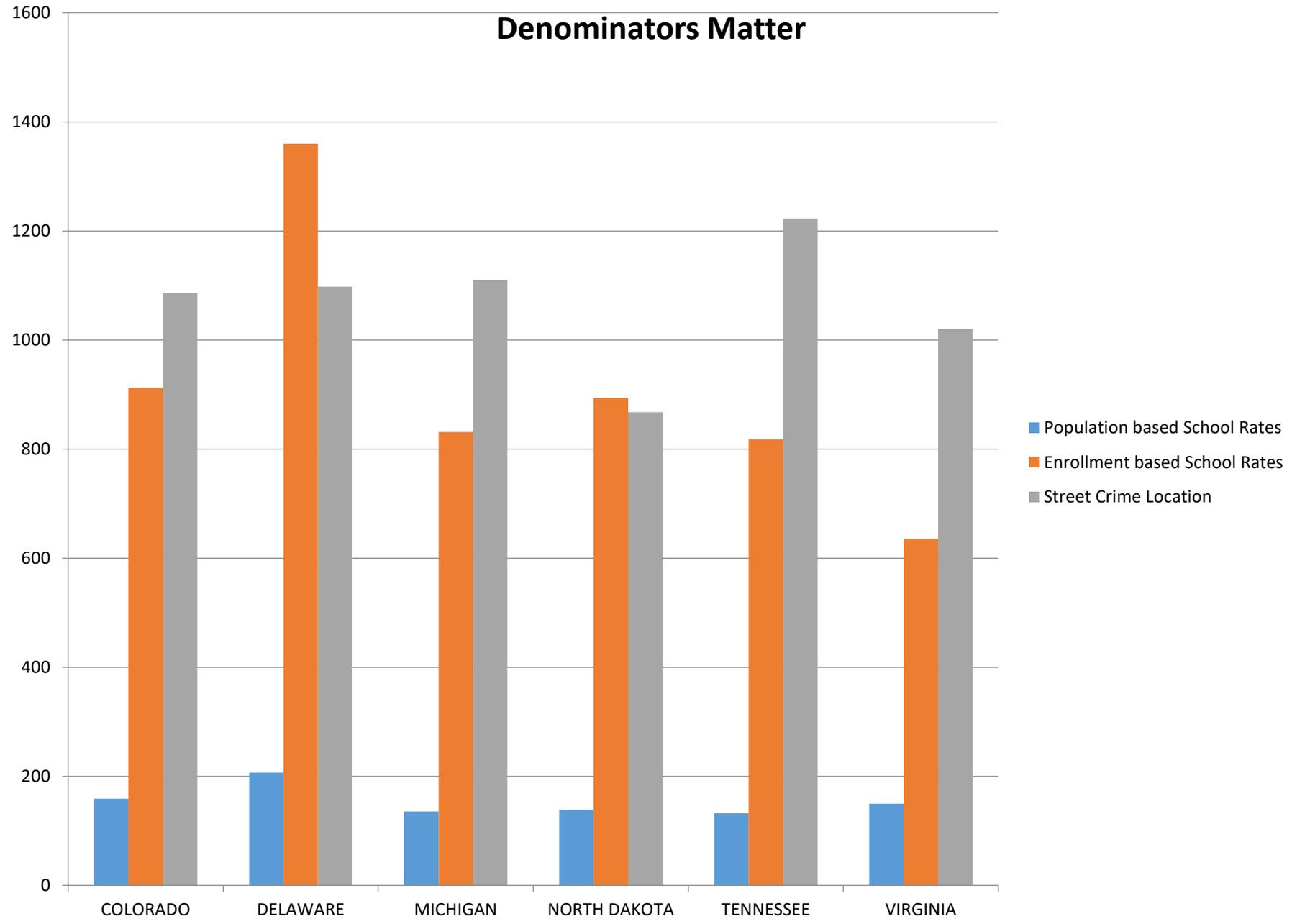
Violent sexual assaults involving children under 12

Children under 12 were most likely assaulted from 8am-9am and 12 pm-1pm

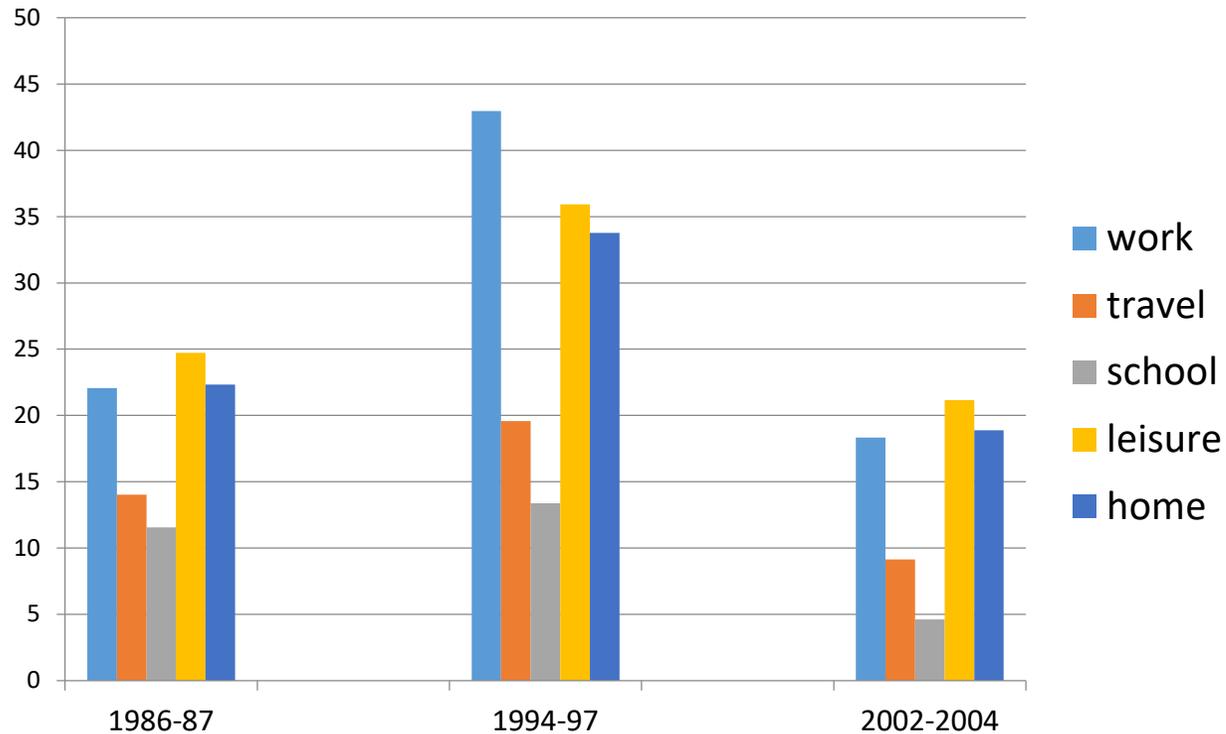| Victims under age 12 assaulted in a residence					
Offender Age	Intimate	Family	Acquaintance	Stranger	Total
7 to 11	0	46	44	1	90
12 to 17	0	**157**	**102**	2	261
18-24	0	49	44	2	95
25-34	0	**108**	**60**	2	171
35-44	0	**78**	44	1	124
45-54	0	49	31	1	81
55-64	0	25	12	0	37
65+	0	13	7	0	20
Total	**2**	**525**	**343**	**10**	**880**
Victims under age 12 assaulted in a non-residence					
Offender Age	Intimate	Family	Acquaintance	Stranger	Total
7 to 11	0	3	22	1	25
12 to 17	0	9	22	2	34
18-24	0	4	7	2	12
25-34	0	8	7	2	17
35-44	0	6	7	2	14
45-54	0	4	5	1	10
55-64	0	2	2	1	5
65+	0	1	2	0	3
Total	**0**	**36**	**74**	**10**	**120**
Overall Total	**2**	**561**	**417**	**20**	**1000**

Source: BJS analysis of data from the FBI's National Incident Based Reporting System, 2006 to 2008.

Denominators Matter

Population based School Rates
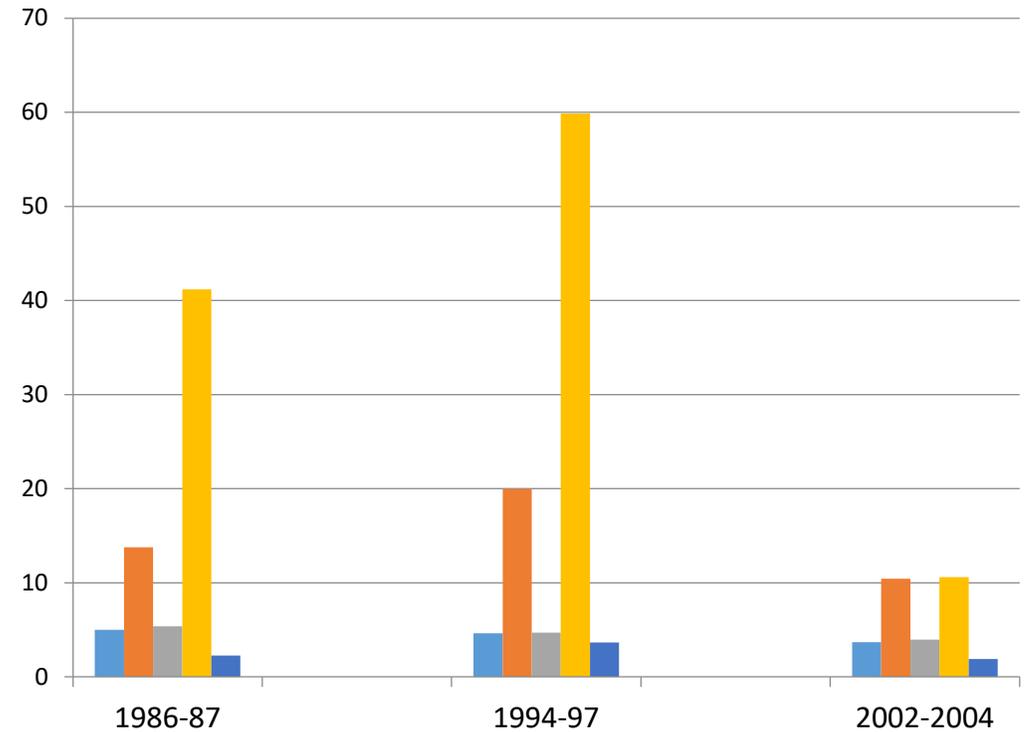Enrollment based School Rates
Street Crime Location

Comparison of Population and Time Adjusted Rates of Personal Victimization

Population Based Victimization Rates by Activity Domain and Time

Personal Victimization Rates by Activity Adjusted forTime

Legend: work, travel, school, leisure, home

More Aggressive Strategies for Integrating Data Sources

- Summary system only allowed for linking of police and other data by geography.
- Incident-level data allows for much more minute matching at the incident or person level.
- Associating of police records with external data was estimate specific
- Incident level data permits the creation of combined data sets that can produce a multiple estimates.
- We will suggest some ways that this could be done and the advantages and obstacles to doing so.

Victim Surveys as an External Data Source

- There are many sources of external data that can be integrated with police administrative record data for enlightenment purposes.
- Victim Surveys were chosen because:
  - Citizen surveys were a source of external data mentioned prominently by the CIWG. Citizens know a lot that the police want to know.
  - Surveys data are collected and available at the incident and person level.
  - Both surveys and police record data speak to a similar population, so there will be substantial overlap which facilitates integration
- Broadening availability of survey and incident level police data for the same jurisdictons.

Strategies for Integrating Surveys and NIBRS data

- Blended estimates of crime rates for jurisdictions.
- Case Control Designs for exploring crime and victimization risk.
- Creating "synthetic" data to fill in the blanks

Blended Estimates of Crime for Jurisdictions

- Victimization surveys cannot easily provide estimates for small areas, e.g. cities or neighborhoods, but there is a great demand for such data.

- Small area estimates of victimization have been done using jurisdiction level police data.

- Blending data from police administrative records and surveys in the same jurisdiction could provide much more accurate small area estimates.
  - Mean Square Error of the estimate is reduced when both data source used
  - Improvements in accuracy would allow routine estimates of victimization at the jurisdiction level.

The Blending Process

- Both the surveys and police records attempt to estimate victimization and crime rates.

- Similar to dual frame estimation where estimates from both frames are combined to provide a more precise estimate of victimization.

- Contribution of each frame can be differentially weighted and sources of bias across the two frames can be modelled.

- Dual frame estimation has been done for higher and lower cost victimization surveys, but not for household and police administrative record frames.

Case-Control Designs to Explore Risk

- Case-control designs provide variance on the dependent variable by sampling on that variable. Valuable for studying rare events.
- Police administration not widely used for studying risk of crime because they lack non-victims with similar information as victims.
  - Are hot spots simply hot people?
  - Crime rates must be tied to areas and area population
- Combining non-victims from surveys with crime victims from police administrative record data would allow analysis of risk.
  - Non-victims are abundant in victim surveys, so most reasonably sized cities will have enough controls.
  - Will need to adjust for selectivity in the case and control lists.

Creating Synthetic Data on Crime

- Both surveys and NIBRS contain data that is unique and data they share in common.
- Can the common data be used to "impute" unique data in the other source?
  - Survey would have extensive data on repercussions of victimization events that police records would not.
  - These data can be used to estimate the impact or costs of offenses handled by the police.
- Cold deck imputation could be used to match cases eligible for imputation across data sources.

Is there enough places with both Surveys and NIBRS?

- NIBRS
  - 12 states are 100 percent NIBRS
  - NCS-X recruiting large cities
- Victim Surveys
  - NCVS has expanded sample in 22 states
  - Many cities have their own survey that would be good for some of these purposes.
- States would seem the most likely places to explore joint use of the NCVS and NIBRS

How much information in common?

- Illustration using instrumentation for 12 city survey conducted in 1999 by the Census Bureau for BJS.

- Of the 12 cities, only one, Springfield, MA had also implemented NIBRS at the time.

Information Available for Linkage

- Attributes of Victims and Offenders
  - Age
  - Race
  - Gender
  - Ethnicity
  - Residence Status
- Attributes of incidents
  - Date and time
  - Type of Crime
  - Location
  - Injury, Loss and Recovery
- Gross categories overlap, but specific evaluation is necessary.

Candidates for Synthetic Estimation

- What does one source have that the other does not?
- Estimating the Impact of Victimization
  - Medical costs and Medical Care--
  - Property Loss and Recovery
  - Injury
  - Time Loss from Work
- Describing Responses to Victimization
  - Police response on the scene
  - Use of Victim Services
  - Arrest and Notification of Arrest

Conclusion

- We must get more aggressive in integrating incident level police data with external sources of information to achieve the full potential of NIBRS.

- We have presented some potential ways of being more aggressive in integrating NIBRS with victim survey data.

- These methods show promise but must be tested and disseminated.

- There are data available that would support this testing.

- I encourage you who are so inclined to walk through the door.