

Combining Data and Crime

Sharon L. Lohr

Arizona State University, Tempe AZ 85287, sharon.lohr@asu.edu

Proceedings of the 2018 Federal Committee on Statistical Methodology (FCSM) Research and Policy Conference

Abstract

One hundred years ago, crime statistics were collected differently by each community—if collected at all. After a brief tour of the history of United States crime statistics, I evaluate the strengths and shortcomings of today's data sources for homicide, violent and property crime, and fraud. Where do we have high quality data, and how might various sources be leveraged to provide a better picture of crime? What are the error properties of the sources and the combined estimates? What are some possible unintended consequences of combining data sources?

1. Introduction

We all know the challenges being faced today by the federal statistical system. Response rates and funding levels continue to decline. At the same time, demand for more granular data—more frequent, and with more geographic detail—is increasing, along with a demand for more privacy. Agencies are being asked to do more and more with fewer resources.

In this paper, I discuss some of those challenges in the context of crime statistics. Section 2 starts with a brief history of the major data sources currently available for estimating the amount of crime, and Section 3 assesses the quality of those data sources. Sections 4 and 5 present the view that breaks in series are not anomalies of a data collection, but reflect errors in the statistics that are not captured in the margin of sampling error. Section 6 argues that using multiple data sources not only provides more information, with possibly lower costs, but can improve our assessments of the accuracy of estimates. The paper concludes with some possible unintended consequences of combining data, and thoughts for the future.

2. Sources of Crime Data: A Brief History

I'd like to begin this history about 100 years ago, with the contributions of the economist, social scientist, and statistician Edith Abbott. Before that time, the Census Bureau compiled statistics on prisoners and some states and cities collected statistics on arrests and convictions, but data were sporadic and of spotty quality.

Abbott was born in Grand Island, Nebraska in 1876 and earned a PhD in Economics from the University of Chicago in 1905. She was at the forefront of many of the social developments of the time, residing at Hull House with her sister Grace Abbott as well as with Jane Addams, Alice Hamilton, Florence Kelley, and Julia Lathrop (University of Chicago, 2018). In 1924, Abbott became the first American woman to be dean of a graduate school in the United States, at the School for Social Services Administration of the University of Chicago.

For the purposes of our story, Abbott served as the statistician for the Chicago City Council Committee on Crime in 1914 and 1915. Her report (Abbott, 1915) went far beyond the other statistical summaries of the day, and had four compelling messages that still resonate today.

1. Accurate statistics are needed to use resources efficiently and to dispel myths about crime.

Abbott's report addressed the popular belief, encouraged by newspaper accounts about organized crime (Lurigio and Binder, 2013) that immigrants were disproportionately responsible for crime in Chicago. She compared the percentages of native- and foreign-born men arrested and convicted of crimes, from Chicago police statistics, with the percentages of native- and foreign-born men in Chicago from the 1910 census. She found that the data available to her indicated that the foreign-born were responsible for less crime than their proportion among adult men in Chicago (although she did not consider age-adjusted arrest rates, as Moehling and Piehl, 2009, argued should have been done).

Abbott was well aware of the deficiencies of the available data. She wrote: "The question as to how far these statistics of 'nativity' are trustworthy must, of course, be considered. In general, the method of having information about country of birth hurriedly entered by a police officer at the time of an arrest or an arraignment would undoubtedly result in many errors" (Abbott, 1915, p. 51).

2. Knowledge about crime is important to human welfare, and crime statistics merit the same degree of attention and resources as statistics about mortality.

In 1915, Chicago was part of the national “registration area” for mortality statistics. Abbott hoped that similar progress could be made in collecting standardized statistics about crime, and endorsed the recommendations from Koren (1911) to establish a uniform scheme for recording data.

3. Crime statistics can be used to improve the quality of policing.

Abbott used the data to form comparison groups and identify outliers. For example, in a follow-up to the 1915 report (Abbott, 1922), she reviewed a comparison of two precincts with similar populations and demographics but different police captains. Further investigation showed that they actually had similar amounts of crime complaints but one precinct had discarded three-quarters of the complaints.

4. No single data source can give a full picture of crime, and data from one source can be used to identify errors and improve the quality of other data sources.

Abbott brought “together from various sources the available statistics, published or unpublished, relating to the subject of crime in Chicago” (Abbott, 1915, p. 19). She used statistics from the police department, the municipal court, the adult probation office, the superintendent of the House of Correction, and the 1910 census to obtain a fuller picture of crime in Chicago. She also used statistics from some sources to investigate errors in other sources. For example, Abbott found the data on criminal complaints to be unreliable. In some years the number of arrests exceeded the number of complaints. The common practice of “desk sergeants writing verbal complaints on slips of paper, placing them on a spindle and tearing them up when an officer reports thereon” (Abbott, 1915, pp. 80–81) did not lead to accurate statistics.

Abbott’s argument about needing statistics to dispel myths about crime was echoed by statisticians, sociologists, and police departments in the early part of the 20th century. The popular conception was that crime was everywhere, and the development of the Uniform Crime Reports (UCR) in the late 1920s was motivated in part by a desire to counteract sensationalist newspaper reports about crime waves (Maltz, 1977). The UCR focused on complaints to the police instead of statistics about arrests or convictions, taking the view that statistics closer to the crime would be more accurate.

A major contribution of the UCR was the development of uniform definitions of crime. For many years, the UCR collected only counts of offenses; they later added more details for homicide through the Supplementary Homicide Reports in 1961. The UCR is currently transitioning to collecting more detailed information about all incidents through the National Incident-Based Reporting System (NIBRS), which started in 1991. In 2016, 37.1 percent of law enforcement agencies submitted their UCR data through NIBRS, with the goal of having all agencies participating in the NIBRS program by 2021 (Federal Bureau of Investigation, 2017b).

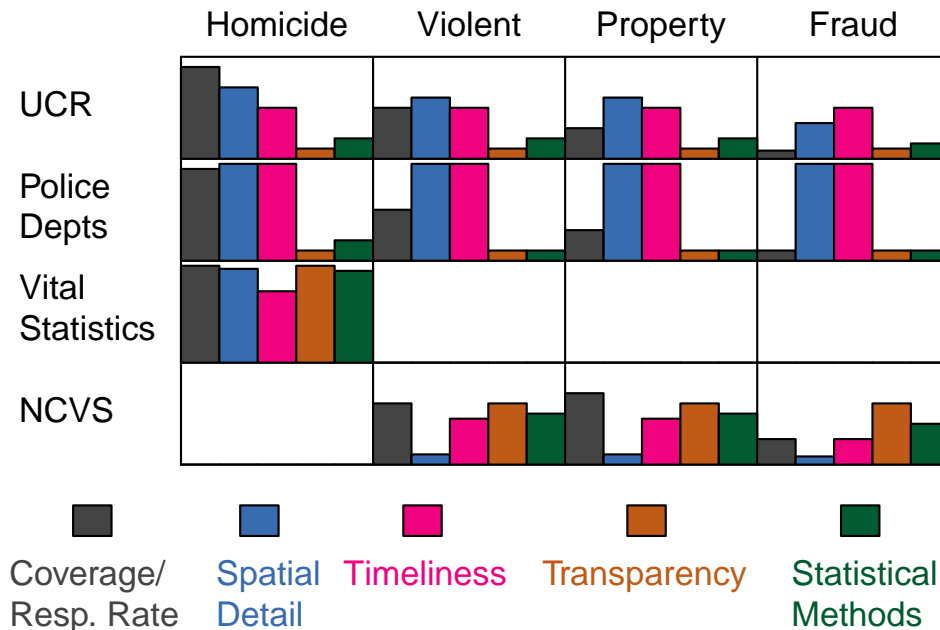
At about the same time that the UCR was launched, in 1933 the National Vital Statistics System (NVSS) achieved participation from all 48 states and the District of Columbia (Hetzel, 1997). This gave a national source of information about crime, but was limited to homicide.

The next major development occurred about 50 years ago, another period of increased concern about crime. Under the Johnson administration, the President’s Commission on Law Enforcement and Administration of Justice (1967) called for multiple sources of data about crime, and was particularly concerned about the lack of information about fraud. The report also took the view that because of unevenness in police procedures, a survey sample would give more uniform crime measures. The survey now known as the National Crime Victimization Survey (NCVS) resulted from this report. From the beginning, the NCVS was viewed as an alternative source to the UCR which would provide information on crimes not known to or recorded by the police.

I date the fourth source of data considered in this paper to 1994, with New York City’s adoption of the CompStat program. This type of computer-based predictive policing system is now used by most large police departments. These involve massive databases of information about past crimes and risk factors for locations and, in some cases, persons.

In Chicago, for example, the Crime Summary map at <http://gis.chicagopolice.org/> displays crime counts in Chicago census tracts or community areas, including incidents up to the day before. Clicking on an incident on the

Figure 1: Crime Data Sources: Subjective Quality Ratings



Crime Incidents map displays the time, block where the incident occurred, type of location (for example, residence, gas station, parking lot, street), type of offense, weapon use, and whether it was a domestic incident. Of course the internal Chicago police department database contains much more information that is not available to the public.

Many of the city databases contain much more detail about incidents than the summary statistics from the UCR. In contrast to the UCR, NVSS, and NCVS, the police department websites make data available almost immediately and with high spatial detail.

3. Quality of Crime Data Sources

Figure 1 presents my assessment of aspects of quality of these four sources for four types of crime: homicide, violent crime, property crime, and fraud. I want to emphasize that these subjective assessments are based on my review of the data sources and do not reflect the views of any organization or agency. The picture is included as an example of how one might compare the strengths and deficiencies of different data sources for meeting information needs.

The bars from left to right are for different features of quality. The gray bar in the left of each box is my score for each source on how much of the crime it captures (coverage) and response rates. The second, blue, bar rates the level of spatial detail or granularity. The third bar, magenta, is on how fast the data are available—timeliness. The fourth bar in each box, brown, rates the transparency of the source about the details of data collection, and the completeness and usability of the documentation. The green bar on the right of each box rates the quality of the statistical methodology used in design and estimation.

In my assessment, we have several high quality data sources about homicide, but much less information about fraud.

Uniform Crime Reports. The UCR probably captures most homicides. But it has incomplete coverage for other crimes because many are not reported to the police. In 2016, according to the NCVS, about 51 percent of serious violent victimizations and 36 percent of property victimizations were reported to the police (Morgan and Kena, 2017). The 2014 NCVS Identity Theft Supplement estimated that only 8 percent of identity theft victims reported the incident to law enforcement agencies (Harrell, 2016).

A number of factors affect the accuracy of UCR statistics. It is a voluntary program and law enforcement agencies report their own data. There are some documented instances of measurement and classification error, in which an agency may record a robbery as a purse-snatching or an aggravated assault as a simple assault.

The UCR is currently updating its web site and statistical methodology with the transition to NIBRS and the development of new user-friendly data tools, and I expect that the documentation about the methods will be more comprehensive and easier to find after this updating is completed. I think it would be helpful to have a Statistical Methodology link under the Documentation section of ucr.fbi.gov, which would put all of the technical documentation about statistical methods and quality assurance in one place. This section could also give measures of uncertainty for the statistics, which are not exact counts but are affected by missing data and measurement errors. The imputation methods currently used for the summary reporting system count data are largely unchanged from those proposed in 1958 (Federal Bureau of Investigation, 1958) and do not exploit more recent developments in imputation methodology. However, ongoing research (see, for example, LaValle et al., 2014) has identified improvements that can be made to the imputation procedures for NIBRS data, and I anticipate that the imputation procedures will be modernized as the transition to NIBRS progresses.

Police Department Data. The UCR data come from police departments, and I give the police department data similar scores on coverage because they are affected by the same issues. But there are two areas in which the police department information excels: much more detailed information is available, and it is available much faster.

The relatively low ratings of the police department data for transparency and for statistical methodology are not intended as criticisms of these sources, but reflect their primary purpose as administrative data collected to improve policing. Statistical uses are secondary, and statistical methodology would only be involved if these data sets were used to produce official statistics as is done by the UCR. Also, there are practical reasons for not making the methodology and algorithms used by police departments transparent—if all of the methods were posted on the internet, would-be criminals could use them too.

National Vital Statistics System. I think the vital statistics homicide data, which come from death certificates, are probably about the best crime data we have. The NVSS has an ongoing quality improvement program and is also a leader in efforts to combine data sources. Twenty-seven states currently participate in the National Violent Death Reporting System, which links data from police, death certificates, medical examiner reports, and toxicology reports (Crosby et al., 2016). However, the crime information from this rich data source is limited to homicide.

National Crime Victimization Survey. The NCVS provides information on crimes not reported to the police, and thus has higher coverage for some types of violent crime, property crime, and fraud than the UCR and police department sources. The blue bar for spatial detail is low because the NCVS is a survey—it cannot produce the same level of spatial detail as the data sources that attempt to capture all of the crimes known to law enforcement agencies. In 2015, the NCVS data had records for 103 sexual assaults and 219 motor vehicle thefts. By contrast, the 2015 NIBRS data had detailed information on more than 70,000 sexual assaults and 170,000 motor vehicle thefts. Because of the small sample sizes the NCVS cannot, by itself, produce single-year estimates of characteristics of sexual assaults for subpopulations.

While the NCVS captures crimes not found in the other sources, its response rates have been decreasing, as shown in Figure 2. The response rates were obtained from Table 4–2 of National Research Council (2014) for 1993–2011 and from the annual reports of *Criminal Victimization* for 2012–2016. The overall person-level response rate for each year was approximated by multiplying the household-level response rate by the conditional person-level response rate. The household response rate was relatively steady and above 90 percent until 2010, but since then there have been sharp declines. In 2016 the household response rate was about 78 percent, and the overall person-level response rate was about 65 percent.

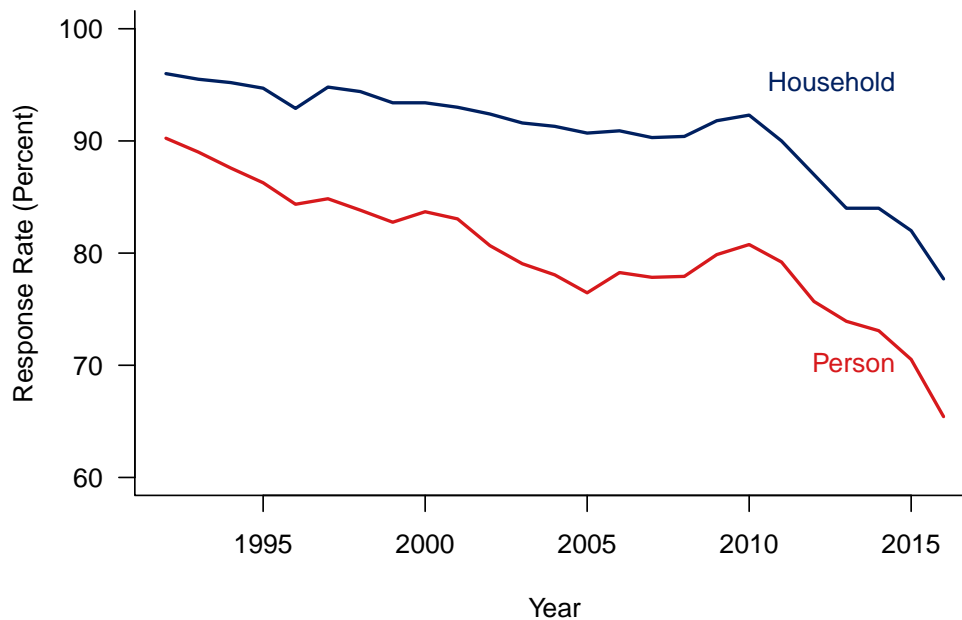
The overall response rates do not give the full picture of NCVS representativeness. The NCVS began publishing detailed nonresponse information for demographic groups in 2016, the year in which the household response rate dipped below 80 percent, and the published response rate reports for other years do not contain the same level of detail. However, we can look at possible effects of nonresponse on the representativeness of the sample through the lens of the final weights, which include nonresponse adjustments.

The lines in Figure 3 display the ratios

$$\frac{\text{median final weight of respondents in age group}}{\text{median final weight of respondents age 60 and over}}$$

for each year and for the age groups 12–17, 18–24, 25–34, 35–59, and 60+. Each point on the line for ages 18–24, for example, shows the ratio of the median weight for 18–24-year-olds in that year to the median weight for the persons

Figure 2: NCVS Household and Person-Level Response Rates, 1993–2016



age 60 and over in the same year. This gives a measure of weight disparity that is independent of the sample size. If the survey had a constant sampling fraction and full response, we would expect all of these lines to be constant at one. The further away from one the ratio is, the larger the weight disparity. Figure 3 shows a steady increase in the weights of the younger age groups relative to the older age groups through 2015.

The sharp increase in 2016 may be partly attributable to different demographic characteristics in the new primary sampling units, since the NCVS is now oversampling certain states to produce state-level estimates. But much of it is due to the low response rates in the younger age groups—in 2016 the conditional person-level response rate (not including household nonresponse) was 64.6 percent for persons age 12–17 and 73.5 percent for persons age 18–24, compared with 93.6 percent for persons age 65 and older (U.S. Department of Justice, Bureau of Justice Statistics, 2017, p. 17). The survey thus has the highest nonresponse and the highest relative weights for the age groups that experience the highest rates of violent crime.

4. Standard Errors from Design-Based Inference

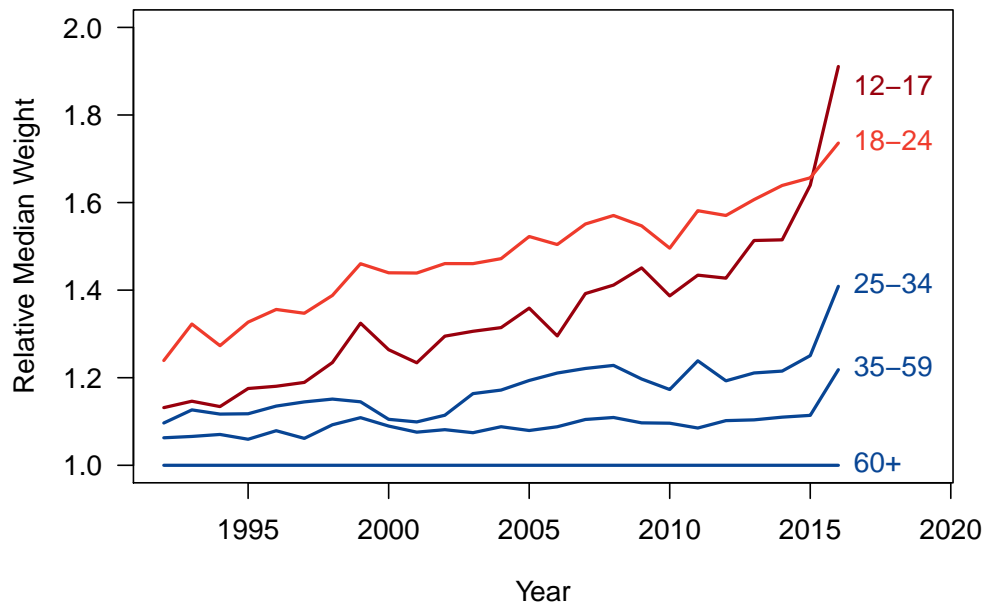
Response rates for the NCVS and other household surveys have been declining. Yet the standard errors are computed the same way as they were when response rates were higher. At what point does the response rate become too low for design-based standard errors to provide proper inferences?

Here is what three pioneers of probability sampling said:

In the Census Bureau it is usually assumed that if the required information is obtained from more than 95 per cent of the designated households one is entitled to feel fairly secure in assuming that the sample was taken in conformance with sampling theory, even though assumptions may be necessary for the remaining 5 per cent. It has been found that for some purposes trouble arises even when making assumptions for only 5 per cent. (Hansen and Hurwitz, 1949, p. 365)

A sample that is 95 or 98 percent a probability-sample and the other 5 or 2 percent a judgment-selection or judgment-adjustment for refusals, for people not at home, etc., may still be an excellent sample, although it is important to investigate the remaining 5 or even 2 percent as soon as possible. (Deming, 1950, p. 13)

Figure 3: Ratios of Median Weight for Age Groups to Median Weight for Age 60+, 1993–2016



Under some circumstances, said Hansen and Hurwitz and Deming, it might be acceptable to have a response rate as low as 95 percent. Below 95 percent, however, the design-based standard error may be suspect. And they had doubts about the inference even with 95 percent response—a level much higher than most response rates today.

Figure 4 displays a graph of the estimated number of serious violent victimizations (rape, robbery, aggravated assault) from the NCVS along with 95 percent confidence intervals (the data are from Bureau of Justice Statistics, 2018). The confidence intervals, computed using the design properties, are quite narrow because of the large sample sizes.

The graph shows a spike in number of serious violent victimizations in 2006. In that year, new primary sampling units were chosen based on the 2000 census. In addition, the NCVS replaced paper-and-pencil interviews with computer-assisted personal interviews and made other methodological changes (Rand and Catalano, 2007). Similarly, the 2016 NCVS had new primary sampling units chosen using information from the 2010 census.

The spike occurring with the design change in 2006 indicates that the cross-sectional NCVS standard errors that are reported every year may be too small, and the standard errors for the estimates of change for 2005–2006 and 2015–2016 are also too small. If the spike does not represent a real change in the victimization rate, the standard error should be wider to reflect the uncertainty about the estimate. The standard errors are based entirely on sampling variability, which is only part of the error. They do not capture errors from measurement or missing data. Poststratification reduces the standard errors further, making the assumption that the weighting adjustments have removed all nonresponse bias. Poststratification is justified theoretically when there is full response, but may be shrinking the standard errors too much when response rates are low.

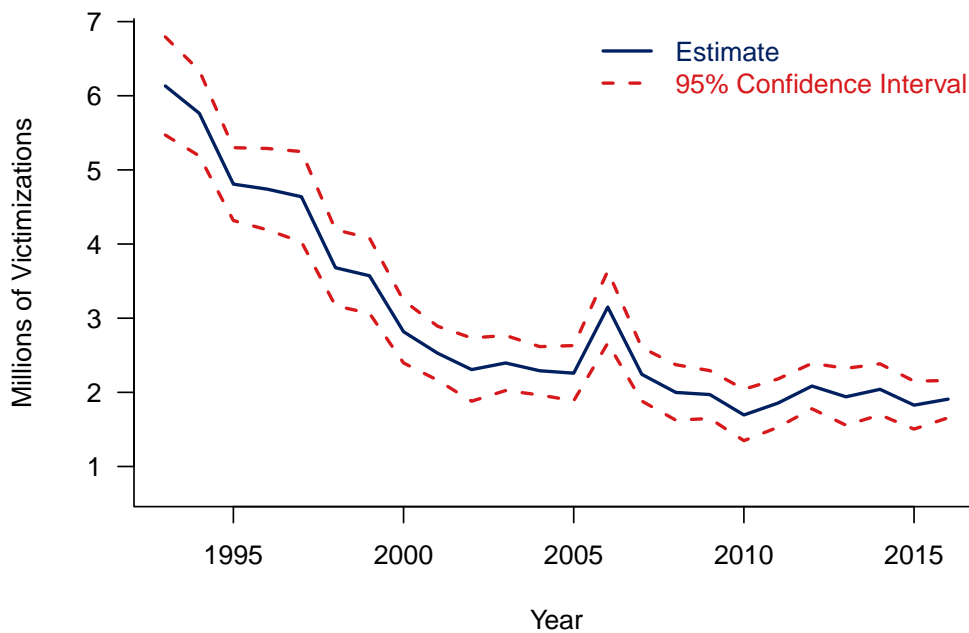
No standard errors are reported for the UCR and police department statistics. Almost all of the errors in these sources are from missing data or measurement errors.

5. Common Causes and Special Causes of Variation

Years such as 2006 and 2016 are treated as anomalies in the NCVS series. Rand and Catalano (2007) and Morgan and Kena (2017) stated that because of changes in methodology and the introduction of new primary sampling units, national estimates from 2006 and 2016 are not comparable with NCVS data from previous years.

But most of the errors underlying those anomalies have been there all along. They only become visible as breaks in the time series when the sample design changes and the new primary units expose some of the nonsampling errors.

Figure 4: Serious Violent Victimizations from NCVS



When the primary sampling units and procedures are carried over to successive years, the errors carry over too and are hidden.

Deming (1986) distinguished between special causes and common causes of variation. Special causes are due to a particular event. For example, in July 2016 the NCVS added questions on citizenship and sexual orientation to the screener questionnaire (U.S. Department of Justice, Bureau of Justice Statistics, 2017, p. 7). Changing the questionnaire is an example of a special cause of variation that might cause response rates and victimization estimates to change, and its effects can be estimated if an appropriate experimental design is used.

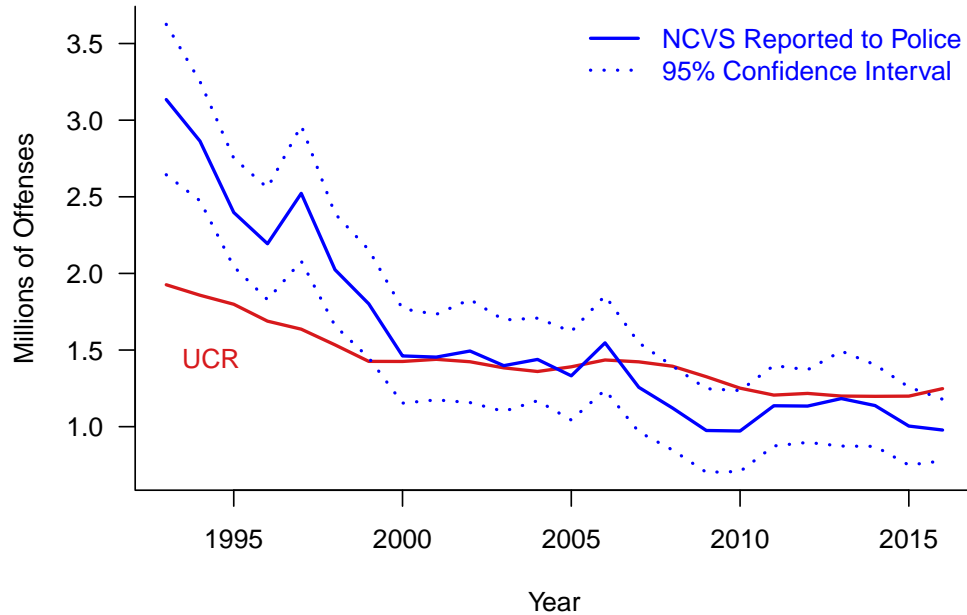
But I argue that the effects of new interviewers and new primary sampling units are common causes of variation. These are systems features, and should be dealt with as part of the system. The spike for 2006 in Figure 4 indicates that the standard errors for the NCVS fail to capture some of the sources of common-cause variation.

How can a survey report a margin of error that includes common-cause sources of variation from measurement, interviewing, and nonresponse error as well as sampling error? Since the nonresponse adjustments use models, models are also needed to inflate the estimated variability to include nonsampling error. Much more research is needed on this problem, but here I would like to mention a few promising approaches.

Spiegelhalter and Riesch (2011) discussed the problem of estimating uncertainty in the arena of estimating risks. They distinguished between aleatory (due to randomness) and epistemic (due to lack of knowledge) uncertainty, and presented a five-level structure for assessing and communicating uncertainty. Lohr and Brick (2017) proposed incorporating the effects of choosing different weighting models through a Bayesian model averaging approach. Measurement error models can be used to include variability from interviewers who may be concentrated in new primary sampling units—variability that was treated as special causes in the 2006 and 2016 NCVS. Another potential approach employs a hierarchical model to capture between-source variability from multiple data sources (Lohr and Raghunathan, 2017; National Academies of Sciences, Engineering, and Medicine, 2017). In all of these, the model-based measure of uncertainty will be larger than from pure design-based inference.

The hierarchical models mentioned in the preceding paragraph combine data from multiple sources to construct a point estimate or posterior distribution that relies on all of the sources. In the next section, I want to highlight a different use of multiple sources of data—to get a more realistic view of the uncertainty about the estimate from a single source.

Figure 5: Serious Violent Crime from UCR and Reported to Police from NCVS, 1993–2016



6. Using Multiple Data Sources to Study Error Properties

Figure 5 plots the estimates of the number of serious violent crimes reported to the police from NCVS from 1993–2016, along with 95% confidence intervals (Bureau of Justice Statistics, 2018). These are compared with the number of rapes, robberies, and aggravated assaults from UCR for the same years, calculated by subtracting the number of murders and nonnegligent manslaughters from the number of violent crimes for each year. The UCR statistics for 1993 to 2014 were downloaded from www.ucrdatatool.gov on January 30, 2018, and the statistics for 2015 and 2016 were obtained from Federal Bureau of Investigation (2017a).



The NCVS and UCR statistics in Figure 5 are not measuring exactly the same thing. The populations differ—the NCVS is restricted to U.S. residents ages 12 and older, while the UCR includes crimes against children and nonresidents that are reported to the police—and the offenses are defined differently in the two data collections. Nevertheless, we would expect the two series to be roughly parallel and in fact there is a rough concordance in the numbers of victimizations from 2000 onwards.

However, the comparison in Figure 5 provides another indication that the NCVS standard errors may be too low, particularly in 2016. At present, the UCR summary count data do not have enough detail to identify potential causes of the divergences between the NCVS and UCR statistics. However, the NIBRS system provides more information about victims, offenders, and circumstances of the crime. As the NIBRS data achieve greater coverage and become more representative, we will be able to match populations and subpopulations better between the UCR and NCVS and use divergences to obtain more information about nonsampling errors in both sources.

Figure 6 shows a cross-classification of the types of crimes measured (or not) by UCR and NCVS. The graph in Figure 5 displayed serious violent crimes reported to the police. These are captured by both the NCVS and the UCR and are thus in the box in the upper left corner. Of course, comparing multiple sources will not provide information about the crimes that are missed by all sources, shown in the lower right corner. Nor will multiple sources repair biases that are in the same direction for all sources.

The NCVS was initiated to obtain a measure of the so-called “dark figure of crime”—the crimes not reported to the police in the bottom half of Figure 6. And it captures some of them, shown by the light gray box in the lower left. But with the increasing nonresponse and perhaps nonreporting to the NCVS, the box in the lower right—we might think of this as the “really dark figure of crime”—may be getting bigger. In particular, a substantial amount of fraud may be located in the lower right box, uncaptured by both sources.

Figure 6: Crimes Captured by NCVS and by UCR

	In NCVS	Not in NCVS
In UCR	Reported to police Except 	Homicide Reported to police against: Children, persons in institutions, commercial establishments
Not In UCR	Not reported to police Except 	Not reported to police against: Children, persons in institutions, commercial establishments Not represented in NCVS: Nonresponse Measurement error

Lohr and Raghunathan (2017) and the National Academies of Sciences, Engineering, and Medicine (2017) discussed methods for combining data sources to obtain one set of statistics that relies on all of the sources. But we do not necessarily need to reduce the data to one set of statistics. Instead, we can use one or more sources to learn more about error properties of the other sources, and to produce alternative views of the big picture. The diversity of the data sources is a strength. In particular, we need to cultivate alternative sources for the lower right quadrant in Figure 6.

What are some possible unintended effects of an increased reliance on combining data and multiple sources? First, although there is substantial record linkage activity for police statistics and for the National Violent Death Reporting System, I do not recommend linking individual NCVS records to records in other sources. I strongly advocate requiring affirmative consent for any record linkage of survey data, and this could affect response rates and quality of NCVS data, in addition to increasing privacy concerns.

We also want to be cautious about replacing survey data with model predictions. Everyone is under pressure to use inexpensive and convenient data to obtain estimates cheaper and faster. But just because a model holds today, that does not mean it will hold tomorrow. In addition, inexpensive data sources can dry up or change the information collected. The NCVS continues to be a crucial source of crime data, giving information that is not available from any other national source. At the same time, it is also crucial to continue research and experimentation to improve the NCVS, and to use other data sources to assess its quality and error properties.

What do we need to watch out for if we strive for more comprehensive standard errors? Agencies and survey takers are often judged on their standard errors, which provides a disincentive for obtaining measures of uncertainty that include more than just sampling error. It is easy to get a small standard error using model-based methods—just omit sources that disagree, and do not try models that might give different results. We need industry standards for what alternative weighting models, and what statistical methods, to use—similar to the industry standards that exist for how to calculate and report response rates.

7. Lessons from Edith Abbott's Work

The four points from Edith Abbott's groundbreaking 1915 report all apply today. Some of the statistics about crime found on the internet have dubious properties, and we as a statistical community can combat myths about crime through producing accurate statistics and being transparent about their collection and properties. The vital statistics program continues to be a model, I think, particularly in its ongoing evaluation and quality improvement program. It is also a model for facilitating data sharing and promoting standards for data collection. And we can make use of the multiple data sources to obtain a better picture about crime—and the uncertainty about our estimates of crime— than we would have from one individual source.

So where are we? There are big jobs ahead. Some say that surveys are dinosaurs, and there are much cheaper ways to collect data. But the cheaper data sources have errors, too—they just do not disclose how big those errors are. As there is more and more impetus to rely on alternative data sources, we as statisticians can further the debate by insisting on honest assessments of the accuracy of those sources. We can lead the way by changing the practice for surveys—by having better assessment of the error structures (perhaps using alternative sources to help), and by presenting standard errors that better reflect the accuracy of the statistics.

We are at the same sort of frontier in statistics that the early probability samplers faced in the 1930s. Our frontier, like theirs, is how to make inferences from incomplete data. But, as Edith Abbott often said: “We of the West are not afraid of crossing the frontiers.”

References

- Abbott, E. (1915). Statistics relating to crime in Chicago. In *Report of the City Council Committee on Crime of the City of Chicago*, pp. 17–88. Chicago, IL: City of Chicago.
- Abbott, E. (1922). Recent statistics relating to crime in Chicago. *Journal of the American Institute of Criminal Law & Criminology* 13, 329–358.
- Bureau of Justice Statistics (2018). Number of serious violent victimizations by reporting to the police, 1993–2016. Generated using the NCVS Victimization Analysis Tool at www.bjs.gov on February 3, 2018.
- Crosby, A. E., J. A. Mercy, and D. Houry (2016). The National Violent Death Reporting System. *American Journal of Preventive Medicine* 51(5), S169–S172.
- Deming, W. E. (1950). *Some Theory of Sampling*. New York, NY: Dover.
- Deming, W. E. (1986). *Out of the Crisis*. Cambridge, MA: Massachusetts Institute of Technology Center for Advanced Engineering Study.
- Federal Bureau of Investigation (1958). *Uniform Crime Reports for the United States, Special Issue*. Washington, DC: United States Government Printing Office.
- Federal Bureau of Investigation (2017a). *Crime in the United States, 2016*. Washington, DC: FBI. <https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016>. Last accessed November 15, 2017.
- Federal Bureau of Investigation (2017b). FBI releases 2016 NIBRS crime statistics in report and CDE, promotes transition of agencies. https://ucr.fbi.gov/nibrs/2016/resource-pages/nibrs-2016_summary.pdf. Last accessed May 8, 2018.
- Hansen, M. and W. Hurwitz (1949). Dependable samples for market surveys. *Journal of Marketing* 14, 363–372.
- Harrell, E. (2016). Victims of identity theft, 2014. Technical Report NCJ 248991, Bureau of Justice Statistics, Washington, DC. <https://www.bjs.gov/content/pub/pdf/vit14.pdf>. Last accessed February 25, 2018.
- Hetzel, A. M. (1997). *U.S. Vital Statistics System: Major Activities and Developments, 1950–95*. Hyattsville, MD: National Center for Health Statistics.
- Koren, J. (1911). Criminal statistics in the United States. *Journal of Criminal Law and Criminology* 2(4), 568–572.
- LaValle, C. R., S. M. Haas, and J. J. Nolan (2014, June). *Testing the Validity of Demonstrated Imputation Methods on Longitudinal NIBRS Data*. Charleston, WV: West Virginia Office of Research and Strategic Planning. http://djcs.wv.gov/ORSP/SAC/Documents/WV_Impute2ReportJan2014_Final.pdf. Last accessed December 19, 2017.
- Lohr, S. L. and J. M. Brick (2017). Roosevelt predicted to win: Revisiting the 1936 *Literary Digest* poll. *Statistics, Politics and Policy* 8(1), 65–84.
- Lohr, S. L. and T. E. Raghunathan (2017). Combining survey data with other data sources. *Statistical Science* 32(2), 293–312.

- Lurigio, A. J. and J. J. Binder (2013). The Chicago outfit: Challenging the myths about organized crime. *Journal of Contemporary Criminal Justice* 29(2), 198–218.
- Maltz, M. D. (1977). Crime statistics: A historical perspective. *Crime & Delinquency* 23(1), 32–40.
- Moehling, C. and A. M. Piehl (2009). Immigration, crime, and incarceration in early twentieth-century America. *Demography* 46(4), 739–763.
- Morgan, R. E. and G. Kena (2017). Criminal victimization, 2016. Technical Report NCJ 251150, Bureau of Justice Statistics, Washington, DC. <https://www.bjs.gov/content/pub/pdf/cv16.pdf>. Last accessed December 30, 2017.
- National Academies of Sciences, Engineering, and Medicine (2017). *Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps*. Washington, DC: The National Academies Press.
- National Research Council (2014). *Estimating the Incidence of Rape and Sexual Assault*. Washington, DC: The National Academies Press.
- President’s Commission on Law Enforcement and Administration of Justice (1967). *The Challenge of Crime in a Free Society*. Washington, DC: U.S. Government Printing Office. <https://www.ncjrs.gov/pdffiles1/nij/42.pdf>. Last accessed February 15, 2018.
- Rand, M. and S. Catalano (2007). Criminal victimization, 2006. Technical Report NCJ 219413, Bureau of Justice Statistics, Washington, DC.
- Spiegelhalter, D. J. and H. Riesch (2011). Don’t know, can’t know: embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A* 369(1956), 4730–4750.
- University of Chicago (2018). Edith Abbott. School of Social Service Administration. <http://www.ssa.uchicago.edu/edith-abbott>. Last accessed February 1, 2018.
- U.S. Department of Justice, Bureau of Justice Statistics (2017). *National Crime Victimization Survey 2016 Codebook, ICPSR 36828*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor].