Transcript of the December 9, 2021 webinar:
Analyzing Data from the National Crime Victimization Survey

DARYL FOX: Good afternoon, everyone, and welcome to today's webinar, Analyzing Data from the National Crime Victimization Survey, hosted by the Bureau of Justice Statistics. This time, it's my pleasure to introduce Heather Brotsos, Chief, Victimization Statistics Unit within the Bureau of Justice Statistics, for some welcoming remarks and some introduction. Heather?

HEATHER BROTSOS: Good afternoon. My name is Heather Brotsos. I'm chief of Victimization Statistics at the Bureau of Justice Statistics. And I'm thrilled that you've made time in your busy schedules to join us today for our workshop on Analyzing Data from the National Crime Victimization Survey. So before we get started, we'd love to get to know you all a little bit. If you can take a minute to introduce yourself in the chat, how familiar you are with NCVS, how you plan to use the data, and what your preferences are with statistical software programs.

While you're doing that, I'd like to introduce you to our all-star lineup of panelists. Today, we have six BJS statisticians with us: Erika Harrell, Grace Kena, Rachel Morgan, Susannah Tapp, Alexandra (Lexy) Thompson, and Jennifer (Jenna) Truman. So next, we'll take a look at the agenda. And to kick things off, Erika, Jenna, and Grace are going to present some introductory and background information on the NCVS. Next up, we have two special guests, Keith Hullenaar and Tara Martin, who will each present a short presentation on how they're currently using NCVS data. After a 10-minute break, Rachel and Lexy are going to walk you through some demonstrations. You'll be able to follow along using the materials that were sent to you before the—before the event this morning. If you've not received those materials, let us know in the chat and we'll make sure that you get those before the demonstrations so you can follow along. We'll also be monitoring the chat throughout the workshop and we'll leave some time at the end for the Q&A period. We got a lot to cover today, so with that, I'm going to hand it over to Erika to get us started.

ERIKA HARRELL: My name is Erika Harrell. I am a statistician in the Victimization Unit here at BJS. And I'm here to give some background information on our survey. What is the National Crime Victimization Survey, the NCVS? It started in the early '70s as the NCS or the National Crime Survey. It underwent a redesign in the late '80s and early '90s. And in 1992, the survey was renamed the National Crime Victimization Survey or NCVS. The NCVS is one of the nation's two major sources of information on criminal victimization, with the other being the FBI's Uniform Crime Reporting Program. Next slide.

The NCVS, or NCS, was created with a number of goals in mind but the main—one of the main goals was to measure crime that was not reported to the police. Unreported crime was a gap in the UCR Program because the UCR Program only relied on police reports and the NCVS was able to get at those crimes that were not reported to police. The NCVS provides a measure of victim risk and a measure of police reporting overall changes in it by asking crime victims about whether or not they reported a crime—their incident to the police. Next slide.

BJS contracts with the U.S. Census Bureau to collect data for the NCVS. The data are collected year round and the census interviewers attempt to interview everyone in a sample household age12 or older. This is done once every six months to the same household. Seven—this is done seven times over the course of about 3.5 years.

The first interview is done in person while the second through seventh interviews are done either by phone or in person, based on the preference of the respondent. All interviews are self-report interviews. The person just tells the interviewer what has happened in this incident base where we collect characteristics on each crime incident. Each—next slide.

What data are collected? The NCVS, we collect information on nonfatal personal crimes such as rape, sexual assault, robbery, aggravated assault, simple assault, and personal larceny, and also household property crimes such as burglary, trespassing, motor vehicle theft, and other types of household theft. We also collect information on victim and offender characteristics such as age, race, ethnicity, and gender. We also collect information on crime characteristics such as the location of crime and whether or not the offender had a weapon. We also collect information on police reporting and also victim experiences with the criminal justice system such as whether or not the case made it to trial or not. The NCVS is the only source of national data on a variety of topics such as crimes against persons with disabilities and intimate partner violence and injury. Next slide.

For more information on the NCVS Program, please see the very first link. It will take you to a page on the BJS—on the BJS website which describes—giving an overall description of the NCVS. Now, our data and codebooks are housed at the National Archive for Criminal Justice Data which is at the University of Michigan. Those can be accessed at the second link on the slide. The workshop materials that you should have received today that you will probably be using the most will be these three: the Technical Documentation, the Variance User's Guide, and the Criminal Victimization, 2020 annual bulletin, which is basically our annual report that we use to get—to

present our first release of a year's worth of NCVS data. And now I think I should turn it over to Jenna to talk about the survey instruments.

JENNIFER TRUMAN: Thanks, Erika. And thanks, everyone, again for attending today. We're really excited to have you all here. And I'm going to be talking about the NCVS survey instruments and how they're organized for those that are not familiar. So next slide, please.

So the NCVS is a two-stage design survey instrument, and there are several parts. So the first part when we go into households is called the NCVS Control Card. So we are rostering the household and from there, each individual person goes into a crime screener and we ask about victimization they may have experienced. If they were a victim, then we ask at the crime incident report about those individual incidents. If they were not, then they move on to some additional person-level demographic questions. And depending on the time, we also collect different supplement data. And Grace will talk a little bit more about this later. But we collect data on things like identity theft, school crime, and stalking. Next slide.

So a little bit more about the NCVS Control Card. So this is that household roster and this is included in a copy of the workshop materials that you received. So essentially, this is collecting a record from each sample unit. So as Erika said, we sample household. So this collects information at the household level. It's used by the Census Bureau, the field representative, to locate and then also confirm that they've contacted the correct sample household. Generally, we collect information from the household respondents about each person in the household and so we try to get the most knowledgeable adult household member. And then that roster that we're collecting contains demographic information from every person that's living in that household. So we're collecting things like age, race, Hispanic origin, sex, marital status, and education. Next slide, please.

So next, once we have the household information, each person that we interview, everyone 12 or older, is asked a basic screen questionnaire. And so these questions are designed to determine whether or not that person experienced any crimes, whether they were committed against the household or committed against that person. And now in terms of the household, it would just be the household respondent that would be recording those types of crimes. And then we're asking about crimes that they may have experienced in the previous six-month reference period. So the household respondents, as I just said, are asked about property crimes, and then all respondents 12 or older are asked about personal crime. And the questions are written in a short queue format, so the interviewer is reading a "stem" about whether or not the

respondent had experienced a certain type of incident and then gives examples of the types of incidents to prompt the respondent's memory. So if they weren't necessarily thinking about something, that's what they're intended to do, is kind of to prompt their memory and determine whether or not they've experienced that. Next slide.

So one thing to know is that we don't generate any of the crime estimates from the screener questions. So the screener is designed just to trigger what we call incidents, just again using those cues to figure out and remind people and ask if they have experienced anything. And so we're not—but we're not—we're ultimately classifying the types of crime from the crime incident report which we'll talk about here in a second. So if no incidents are recorded, then the interviewer proceeds person by person until the questionnaire is completed for each person 12 or older in that household. Again, we have a copy of this screening questionnaire in your workshop material, so please refer to that and let us know if you have any other questions. Next slide, please.

So if—for those folks that have experienced victimization, they would then turn to the Crime Incident Report. And so the Crime Incident Report, or the NCVS-2, is used to gather detailed information about the incidents that they reported in the screener. We complete one crime incident report, or CIR, for each incident reported. It is all incident-based, so the questions are focused on details regarding that criminal incident. And, again, as I was just kind of saying, ultimately, our crime estimates that we report outcomes from the CIR. So the information that we're capturing there and confirming that a crime was committed then is what ends up in your type of crime coding from here. Next slide.

So within the Crime Incident Report, the types of things that we're asking about, so we're asking about location and presence. We're asking about things—about if they were attacked or if they were threatened, medical—or injury and medical care consequences and things like emotional consequences, whether or not there were any actions taken against the offender, anything about the offender that they might know. So things like the relationship to the offender, you know, age, race, sex of the offender, if they know that information. If things were—you know, if theft—if they experienced a theft whether or not it was an attempted theft or a completed theft. If there was any property damage and also if they reported to police. So one of kind of our main measures that it is important for them to get with whether or not they were reported to police so we can then report out whether or not, you know, we both—are capturing both reported and not reported to police crimes here. Next slide.

So the other thing that—to keep in mind for victimization when we're collecting them is there are—we're asking information about what we call series crimes or series victimizations. So these are victimizations that are very similar and occur so frequently that the victim isn't—is unable to recall each kind of individual event. So as I said, we're asking about each incident when we can. So they've told us—told us in the screener they experienced something, then we asked in the CIR if they—about each of those incidents. But for something like a domestic violence or intimate partner violence that may be occurring frequently, they may not be able to recall each kind of individual incident and details about those. And so in the CIR, we ask them if that was the case. So if it occurred more than six times, it's kind of where we start asking about the series crimes. And if they're not able to tell us details about each individual event we'll classify it as a series victimization and we'll only collect the detailed information on the most recent incident in that series. And so while it is more similar for, you know—or it'd be more, you know, prevalent for something like intimate partner violence or, you know, other attacks that may have happened, it can exist for any type of crime. So they've experienced something frequently enough and they only can tell us kind of details about that, you know—not details about all of them. We'll collect just the first— or the most recent one and collect that detail there. Next slide.

In addition for Census Bureau, our data collector, and so they do the data processing for us. And in addition to the information that we collect from the questions that are in the CIR, we also use what—narrative. So the last—the last question in the CIR, the interviewer writes the summary of an incident, and then those narrative or incident summaries are compared to kind of what was collected in the CIR. And so the interviewers are, you know, trained to kind of pay particular inform—you know, particular attention to certain questions, things that are important for classification. So something like location and presence because if we know—you know, if some—if the person wasn't present, then it wouldn't—it wouldn't be a personal attack to the person. Things like whether or not it was an actual physical completed attack, if it was an attempted attack, if there was theft involved. So something, you know, potentially was missed in the CIR or something, you know, didn't happen, the interviewers have that kind of opportunity to—and have that information in the narrative, provide that summary, and then our Census Bureau colleagues, when they're processing the data, they're actually looking at that and using it for editing and coding, coding ultimately that type of crime, adding details if it was missed and editing the data. And the details that we're collecting from the CIR then go into our classification. And so all of this detail that we've collected, ultimately is how we classify and then record out on the crime. And, again, just as a reminder, there's also a copy of the Crime Incident Report in your workshop materials so you'll have that there to reference. Next slide.

So here's the crime classification. So from all this information that we get from the instruments, we then end up here at this crime classification. And as Erika mentioned, we're classifying both personal and household-level crimes. And for person-level crimes, we have kind of overall violent crime. As I mentioned, they're—we're collecting information about both completed, attempted, and threatened, rape and sexual assault, robberies, assault which—both aggravated and simple assault. And the difference there between aggravated and simple is whether or not there was a weapon or serious injuries. So if there was a weapon or serious injury involved then it is classified as aggravated assault. And then, lastly, we have kind of personal theft, which is like purse snatching, pick-pocketing. We collect that as well. And so this is the person-level crime classification. Next slide.

And then we have the household-level crime classification or property crimes. So from the CIR, we then classify these types of crimes. So we have burglary which we have completed or attempted, motor vehicle theft, again, completed or attempted, and then there's the household-level theft. And we're able to classify, you know, again, completed, attempted, and then the actual amount of theft as well. And next slide. And I think I will be handing this over to Grace. Thank you again for your time. And, Grace, take it away.

GRACE KENA: Thank you, Jenna. Good morning or good afternoon, depending on when you are. Thanks—where you are. Thanks again for joining us. I'll be speaking to you today about how to access NCVS data and statistics. So depending on your needs, there are three ways to do this. Through our statistical reports, through our new online data tool, and through our data files. Next slide.

BJS releases NCVS data through several reports, including our Annual Criminal Victimization Bulletin shown here on screen, and these data are taken from the core NCVS data collection. There are also topical reports that use core data such as hate crime victimization and crimes against persons with disabilities. And then reports such as stalking victimization are based on our stalking supplement, which I'll talk more about in a little bit. Next slide.

NCVS data are also available through our new data tool, the N-DASH. For those who might be familiar with our former NVAT tool, this tool is based on the same types of data generally but just uses a more expanded set of modern presentations. Just another note that the NVAT and the NCVS API are being phased out by early next year in favor of the N-DASH but a new API is in the works for development in the future. So on screen here are a few example charts that you can view and download from the N-DASH. So on the N-DASH site which is listed here on screen for you, you

can access a tool overview and a user's guide that will guide you through how to use the tool. Also available are quick graphics with preset graphics and tables, as well as custom graphics that you can manipulate according to your needs and preferences. If you do get a chance to check the tool out, do let us know what you think. Next slide.

In terms of data files, as Erika mentioned, BJS archives its public-used data at the National Archive of Criminal Justice Data or NACJD. And NCVS data can be found there as well. So data from the core survey, the supplements, and other one-time data files are available there. Next slide.

Each data set has a reference number and includes the date that the data set was released or revised which can be helpful if you need to know what the most up-to-date file is for whatever reason. Each set of files also includes the data files themselves, a codebook, and data documentation. The core NCVS data are available in single-year data files as well as concatenated files which include all data years from 1992 to the most recent data year. For the core collection, the data files include person files, household files, and incident files which Rachel and Lexi will talk about in more detail coming shortly. These core files are released annually along with or not long after the Criminal Victimization Bulletin Report is released. And data files for each of the supplements are also available through the NACJD. Next slide.

Restricted-use data files are available through the Federal Statistical Research Data Centers or FSRDCs or RDCs, depending on your acronym of choice. With these files, you can access more variables than are available through the public-used files. These internal files begin with the 2005 NCVS data for the core, and they include more detailed geographic information and less restrictive top coding of variables than you'll find on the public-used files. In order to get approval, you have to submit an application where you'll need to describe where you want to work in terms of RDC location as well as your research and your specific need for the data you're requesting. More information about this process is available at the link on your slide.

So on screen, here are examples of some NCVS data files that are available through the RDC. And, typically, these files are also released on similar timelines as the accompanying BJS reports and public-used data. So it isn't only—for anyone who isn't familiar, only NCVS data that's available there, all of the other data that the Census Bureau collects is there as well. And so if anyone has research interests that would involve linking files to the NCVS, those options are available. Next slide.

So as mentioned earlier, in addition to the core NCVS, supplemental surveys are also administered periodically to eligible respondents at the end of each NCVS interview.

So, for example, the target population of the 2017 fraud supplement was persons aged 18 or older. So though the core survey is administered to persons aged 12 or older, only people within that reference population would get the supplement at the end of their core interview. And supplements are typically in the field for six months at a time, so generally either January through June or June through July. So why do supplements? The core NCVS is already a very lengthy and expansive survey and so these supplements are used to capture emerging issues that can't be easily incorporated into the core instrument and they allow BJS to better capture the changing landscape of crime. Recent supplements include Identity Theft, which is typically administered every two years, the Police-Public Contact Survey which is typically administered every two to three years, the School Crime Supplement which is sponsored by the National Center for Education Statistics and that's generally every other year, and the Fraud and Supplemental Victimizations or Stalking Surveys which have been on less regular schedules of late.

As with the core data, NCVS supplements are available through the archive as a single flat file with the NCVS variables from the core attached. The one exception to that is the PPCS, which contains only a subset of demographic data from the NCVS interview and most of the responses from the PPCS questionnaire itself. And the reason for that is just as a disclosure avoidance precaution in not linking the PPCS data to the full NCVS because of any information that might be available publicly about police contacts that would lead people to be able to identify respondents.

Turning back to the data availability codebooks for the supplements—oh. Go ahead. Thank you. Codebooks for the supplements generally include GVFs or generalized variance functions that can be used to conduct analysis. Alternately, you can use direct estimation through the Complex Sample option in SPSS. And as mentioned earlier, the restricted-used files are also available through the Census RDCs. Next slide.

Now, I'm pleased to introduce our two special guests for today, Dr. Keith Hullenaar and Dr. Tara Martin. Both Keith and Tara have done a lot of research and analysis using the NCVS and both are former attendees of this workshop. So we're very pleased to have them today. They're going to be sharing with us a little bit of work that they've done using the NCVS just to give a flavor of the kinds of questions that can be answered using the data. So we welcome them. We're excited to have them today, and we just generally welcome the work that others do to expand the reach of this survey. So I will turn it over to Keith with that.

KEITH HULLENAAR: Grace, I want to echo, you know, what you said, like since I've—since I've started working with the NCVS, the team has been incredibly supportive of

some of the work I've done and some of the work my research team has done using this data set. And they are—you can go to the next slide right here. So I've been working, you know, with the NCVS for around eight years. I remember I was—the first time I ever saw the dataset, you know, all the rows and columns were—I was actually an undergrad where I examined, I think, the association between reporting to the police and with the sexual violence victimization. And, as Grace said, I took this workshop four years ago and this was sort of a game-changer for me, you know, being exposed to the experts who [INDISTINCT]. So I kind of went from—you know, like that day and the weeks afterward when I was able to message them and ask questions was sort of like when I went from being sort of an amateur sort of user to someone who's a lot more proficient and had a much better understanding of the dataset. So I think that you're in—you're in the right place if you want to use the NCVS for your research and your work.

So a little bit of background about me is that I work at Harborview Injury Prevention and Research Center at University of Washington. And we're sort of an eclectic group of folks who are really interested in violence and injury prevention efforts that advance equity. In terms of education, I received my Ph.D. in Criminology from Pennsylvania State University. And since my grad program, I've been using the NCVS to broadly study violence issues, but my research kind of falls into three different buckets, the first being victim help-seeking, the second being health consequences of violence, and the third, you know, I'm really interested in violence both perpetration and the engagement of adolescents and youth violence. You know, and here, I've been really focused on examining patterns and trends but also how youth violence is policed both at the schools and outside of the schools. And what I want to be talking about today is sort of how I'd use the work and the data from the NCVS to do my research on victim help-seeking and health consequences and some of the stuff I've been doing there. Next slide.

So one of the major projects I've actually been looking at recently with the—with the group at HIPRC and University of Washington is using the NCVS to study police report—police reporting and health-care use after violent injury. So I sort of mentioned in the introduction here, we have an understanding that not all violence is reported to the police. We often refer to this as sort of a dark figure, right? But when it comes to actually violent injuries, we pay less attention to how there's also a dark figure when it comes to the violence that's treated and untreated by health-care providers. Now this is really important because how we survey and how we monitor bodily injury often comes from administratively from police data or police reports and also health-care records. And what we can do, you know, instead of like expending a lot of resources and time trying to match whether a case that's in the police datasets also on the

health-care record, which many studies have done, what we thought we could do is actually use the National Crime Victimization Survey to examine, you know, more simply, you know, if a victim experienced a violent injury, that's serious enough, you know, that they became serious enough to get treated at all by medical care whether it's a Band-Aid or going to the hospital, are they engaging police reporting or are they seeking health care from formal health-care providers? Can—this could kind of give us a sense about, you know, what proportion of violent injuries are reported to the police and also what proportion of violent injuries may be in actual health-care records and be able to pull from the data? And what NCVS does, it gives us a national-level picture of that question. And so what we wanted to do was ask the—this simple but kind of complex question, what are the differences in terms of violent injuries between police reports and health-care records? All right. Next slide.

So what we did is we used the NCVS data, and you're going to learn about this stuff and how to—how to calculate these sorts of rates, is we use the NCVS data and, specifically, cases of victimization that involved injuries that received any type of medical care, again anything from like a Band-Aid to, you know, going to the hospital, and whether these were reported to the police or treated by health-care providers, that is hospitals, doctors offices, clinics. We wanted to know the distribution, you know, both the divergence and the overlaps between police reporting and health-care use. As you could see in the figure in front of you, a lot of these violent injuries are both reported to the police and involved some sort of health-care provider. The second most common sort of option is that the injury involves pulling a police report and then the least common is when the injury involves health care but no police report at all. So in this case, only health care. But, you know, grouping all violent injuries together is, you know, not very informative because we know that victims may seek health care only when they deem the injury is serious enough. So they may avoid health care because it can be very costly for them in some cases. And they may avoid health care for more minor injuries than compared to more serious injuries. So, Daryl, you can go next slide here.

We found that this was the case actually in the data. As you can see, in the minor injuries portion, you see that a lot of the cases for violent—for minor injuries, in this case, it'll be cuts, bruises, scrapes, and scratches. A lot of them involved only a police report rather than healthcare here, whereas for serious injuries, what we found, as we looked at sort of the distribution of the proportion of cases that went to police and health care, only police report, and only healthcare. Now, for serious injuries, most of the cases have only a police report—or most of the cases involved a police report and also health care. And, in fact, when you look at other types of—the distributions for serious injuries, we found that for serious injuries actually in the NCVS data, a serious

injury was just as likely to go to the police as it was to go into healthcare, which suggests that, at least for serious injuries, health-care records and police reports may capture the same proportion of incidents of the dark figure, you know, and have a similar sort of dark figure. But for minor injuries, you know, intuitively find that, you know, health care doesn't do a really good job for the—capturing those minor violent injuries and you may want to rely on police reports in that regard. All right. Next slide.

Now, there's some limitations here, you know, with the NCVS classification of minor and serious injury, it is more than purely theoretical and empirical because you don't have like medical sort of measures of how serious the incident was, like an injury severity score. The health care-used measures in the NCVS don't capture ambulance services. NCS that includes population—excludes populations that are most at risk for violent injury such as people experiencing homelessness. But I think there's a lot of latitude here, what can be used—the intel the NCVS could be used to advance this research agenda. Specifically, I'm thinking about health care use across different populations is an important question that we need to answer and also how victims care for their longer-term problems such as physical, emotional symptoms. Next slide.

And the other project we worked on, this is project number two, is we wanted to use the health consequences of violent victimization measures provided by NCVS to specifically examine how the victim-offender relationship was associated with the harm and symptoms that victims experienced later on after the incident occurs, because we know that criminal justice and health policy is very focused on family and intimate partner violence. We know it affects arrest outcomes and we know that health-care screening often—in interventions, there's a lot of effort and money spent to identify family and intimate partner violence, suggesting that knowing the attacker might be an important risk factor for harm and that it may actually produce greater harm. And some research supports this. But we wanted to use the NCVS data to sort of answer that question. And specifically, we wanted to know whether there might be differences in the victim-offender relation as it relates to harm between those who are injured by violence and those who are uninjured by violence. Next slide.

So this sort of chart right here is a little bit confusing but what you're—what you're looking at in front of you are sort of estimates from a multi-variable model. And each estimate compares a victim-offender relationship category with the reference category being some—the victim was more relationally distant with their offender. So in this top case, you'll see the circle, that's comparing family and intimate partner violence to stranger violence. And, for instance, the triangle that compares family and intimate partner violence to acquaintance violence. And for the diamond, we see acquaintance violence versus stranger violence.

And what we examined is we examined four different outcomes that are in the NCVS data for—in terms of health consequences. And for the red and the blue, we compare the association for those who—for those victims who reported no injury and for those victims who reported injury. And what you can see here is that almost all the points are to the right of the vertical line which in general suggests that as victim—for victims who know their attacker, they tend to report worse health consequences than victims who do not know their attacker or know their attacker less well. And as you can see for some of these, the blue dot is to the right of the red dot suggesting that the victim-offender relationship may have more of an impact on victims who are uninjured than victims who are injured by violence. And this is sort of just a flavor of what you can do with the NCVS data health consequences. And there's many research questions that you can ask here. All right. Next slide.

Just a word of caution though, the NCVS does not include like validated or physical, mental health scales that are often used in medical literature such as the DSM-V diagnostic tool, the SF-12/36. This is about [INDISTINCT] acceptable for some journals but not all. And I think that there's a lot of opportunity here that future research can examine correlates of long-term physical and emotional symptoms. Next slide.

And to conclude, you know, in my opinion, the NCVS has been incredibly valuable to my career, and I think it's, you know, it truly is the leading survey on violent victimization and few datasets compare to its detail and breadth. And with recent redesigns and the continuing of supplements, such as the school crime supplement, the NCVS provides rare insight into violence and health. But as you participate in the workshop today, I think you really need to pay attention to what the NCVS can do really well and what the NCVS cannot do. So be very mindful of the limitations of the NCVS and its measures and its design, which I think you'll get really great insight into today. And that concludes my presentation. I'll hand it off to Tara.

TARA MARTIN: Hi, everyone. My name is Tara Martin, and I am currently an assistant professor at the University of Arkansas at Little Rock. Today, I'm going to be talking about using the NCVS to study domestic violence. And so, unsurprisingly, my research areas include domestic and intimate partner violence. I also study help-seeking behaviors and then methodological approaches to studying victimizations. So you'll see in my presentation that I have done a lot of manipulating of the data files to answer the specific research questions that I have. So I'm going to give you some insight into—how to do that.

Like Keith, I attended this workshop in 2017 I think was the year that I did it. And it has also been invaluable to me. And having access to the statisticians, knowing who to reach out to for help has been very valuable to me being able to work with the NCVS.

So I wanted to start with talking about some strengths of the NCVS for domestic violence research. Obviously, we all are here because we know the NCVS does cover crimes that are not reported to the police. So a major strength for domestic violence research is that unreported crime. Another benefit is that it also includes victim and offender relationship information. The NCVS does ask, "How did the respondent know the offender," and some of the categories are—categories like spouse or ex-spouse, parent or stepparent, child or stepchild, brother and sister, other relatives, but it also does include boyfriends, girlfriends, and exes. And if you're interested in other kinds of people in the home, there are measures for roommates or people like borders that do live in the house as well.

The NCVS also captures all those incident victim and household characteristics, which are good if you want to determine the correlates with domestic violence. There are also help-seeking measures. They of course do ask, "Were the police informed," but they also further ask, "How did the police find out?" They ask questions about whether or not—why it was or was not reported to the police. For example, a person could say that it was a personal or private matter that they took care of it themselves

The other help-seeking measures that are included, I'll talk about some of their limitations as well, are questions about the medical care which Keith was mentioning. So they ask if you were injured to the extent that you received medical care, and then they ask specific types of care that you could receive. There's also a question about seeking other kinds of professional help. That might include responses like receiving counseling, taking medication, anything related to the distress caused by the incident. Then there is another question I think is less well known that asks, "Did you receive any help or advice from any office or agency other than the police that deals with victims of crime?" So I have used these help-seeking measures in a recent study that's currently under review to look at help-seeking behaviors varying by sexual orientation, which is one of the newer questions that the NCVS included in a redesign.

Another strength of the NCVS is that it does have that longitudinal data. It dates the—in—the National Crime Survey dates to 1972, and there are ways to use that data in comparison to the modern NCVS. Because they—you can weight the data to account for differences in how the questions were asked. Some of those behaviorally redesigned questions for some insight into how that could be done, Eric Baumer and Janet Lauritsen in 2010, their reporting to the police journal article talks about how they

can compare that longitudinally. But then people might also be interested in repeat victimization so the NCVS does provide that panel data, which now includes seven interviews over three and a half years. Next slide, please.

There are also some limitations for studying domestic violence research. So I wanted to talk about these limitations so you have some expectations about what you might be able to do with the data. The NCVS does exclude higher-risk populations which would include people that are homeless, it also excludes incarcerated people. And then children under the age of 12 are not included. If you are interested in child abuse, you might be missing a large portion of people you're interested in. It also doesn't include all types of domestic violence or intimate partner violence. If you are theoretically interested in these ideas, you might be interested in something that's not criminally violent like coercive control or financial control. And you're not really able to capture that with the NCVS. Repeat victimization can also be—only measured for three years or 3.5 now that we include that first bounded interview. And this next limitation is a big one for me, and then I see that I skipped—that's help-seeking, so I'll come back up to that. But subsetting the sample often requires pooling the data across multiple years. Once you start getting into very specific types of the sample, for example, if you wanted to look at people with—who experienced intimate partner violence but reported it to the police, had their offender arrested, and they also experienced a major injury, you're going to get really small numbers of people, which means that your estimates become unreliable, your standard errors get really large, and it's hard to detect trends. So keep that in mind that you might need to pool the data across multiple years. But there are limitations to the help-seeking measures going up a couple of bullets. For example, if you wanted to use that, "Did you get any kind of professional help question?" You only get asked that question if you experienced emotional distress for a month or more. So there is the potential that people sought help but didn't experience that distress for a month or more.

And then this last bullet point that I have here is that I think that there might be a possible threshold for what people consider criminal victimization, particularly with domestic or intimate partner violence. And the reason I say that is because the NCVS regularly has lower estimates of intimate partner violence in particular compared to other nationally representative datasets. So maybe people do take that criminal victimization component and don't think of something that applies to their lives. Even though the NCVS questions are designed really well to capture people you might not think of or behaviors you might not think of. Next slide, please.

The BJS reports, the criminal victimization reports regularly discuss prevalence estimates, so I wanted to briefly talk about what the difference is there. A person could

experience multiple victimizations in a year. So I have here this person who experienced three incidents in three victimizations throughout the year. However, that person only counts as one victim. While their victimizations are counted multiple times, they do represent one person. Next slide, please.

And I wanted to discuss the difference there because one person could potentially be accounting for a large number of victimizations. And there could be different policy implications for repeat victims versus a large spread of one-time victims. And so this is data coming from my dissertation, looking at trends in intimate partner violence. And the top line indicates the incidence rate which is similar to victimization rate. And you can see that it's higher than the bottom line which is the prevalence rate, which does indicate that there are repeat victims in the sample, which is unsurprising. But because we might react or respond differently policy-wise, it's important to consider that there are different ways to manipulate the NCVS to answer those prevalence versus incident questions. And so, manipulating the datasets has been one of my biggest challenges, my biggest foes in working with the NCVS. Because if you want to look at prevalence, you have to aggregate the incidents for one person in the incident file, then you merge that file to the person file, which would also be required if you wanted to look at non-victims to see if there are different correlates of victimization. And then you have—you might—you want yearly estimates, so you need to account for the fact that people are interviewed twice in a year. So you create yearly indicators before running your estimates. So it is a bit of a process. And I wanted to talk about another process. Next slide, please.

Where you might be interested in looking at repeat victimization. This also requires manipulating the NCVS and the files that are available to download. If you wanted to look at repeat victimization, the easiest way I would say to do that, is again you need to merge the incident file to the person file. And then the files that you download do come in the long format, which means that there are multiple rows per person. So you also need to restructure the dataset into a wide format, which you can do by grouping the variables by the person identifiers. And then you get a nice dataset that has all the indicators for one person on a single row. However, when looking at repeat victimization, do not forget that the series victimizations exist. There are indicators on a single row of data, so you need to consider that that would also be repeat victimization you're interested in, and don't forget to note that when you're calculating repeat victimization. Next slide, please.

So I wanted to give you a couple of practical tips for using the NCVS. One of them is to understand your measures and the skip patterns to make sure that you are not seeing a bunch of out of—out-of-universe people in your sample. Then you also want to check

your subsample sizes. I recommend doing this early on to see if your research question is feasible as is or if you might need to pull your dataset over multiple years. And then one of the most important things is knowing where to find help. The technical documentation, the statisticians are going to give you all this great information but in my experience, technical documentation is good for understanding changes in the designs. The variance estimation guides are good for knowing when to calculate your rows if you're pooling data. And then don't forget that there are methodology sections in the prior BJS reports which give good insight into how the estimates are actually calculated so you can replicate those. And then, that is all I have. So if you have any questions about anything that I've said, feel free to reach out to me.

HEATHER BROTSOS: Thank you so much, Tara and Keith, very interesting presentations. We are right up against the break, but before we do that, we've got a lot of questions queuing up. So we're going to try to knock off a couple of those before we all leave for our break. And if you have questions for Keith or Tara, please get those into the Q&A and we'll try to get those addressed before the break. So the first question we have, Keith, this is for you. Ingrid asks, "Are there any chance these analyses, police report versus health-care-seeking, were done just for intimate partner violence?"

DR. KEITH HULLENAAR: So we don't have, like, the—it's done separately for intimate partner violence but another part of the paper is that we address—is that we try to use predictors and a model. And these are just very simple like bivariate associations to determine whether intimate partner violence is more likely to be in health care as opposed to, for instance, police reports. And I believe one of the things that we found, I'm pulling out the paper right now for—with regards to that. Just give me a second. Is that, we found that in only health-care data, that intimate partner violence as compared to stranger violence is actually underrepresenting in health care, and only healthcare data. So in this case, if you were to use only health-care records and administrative data, you would put in according to what the NCVS assessments are. You potentially underestimate intimate partner violence. And this matches some of the files I had in a previous paper, you know, that, you know, if you look at the victimizations, actually intimate partner violence is somewhat less likely to actually end up in health care than stranger victims at least—versus stranger violence, I should say.

HEATHER BROTSOS: Great. Thank you. All right. We've got another question. From someone asking about how to drill into the data to view race, age of victims, other demographic information. I swear this one wasn't a plan, but we just released our new N-DASH—data visualization dashboard, if you can use to play around with all of that data and view it in a very beautiful visualization, so we'll put the link into the chat or the

Q&A so that you have that link at your fingertips and you can play around with that data. Another question about, "I noticed the date of the census data was 2016 but we've got—when will that be more—when will that be updated using more recent census data?" So, Rachel, I think what you are referring to is our sample redesign. So every 10 years, we take data from the decennial, and we update the sample design that in a way to account for, you know, shifts in population across all of our—the—all of our coverage. And so, our next sample redesign update is going to be in 2026. We do that every 10 years. And we'll be using data from the 2020 decennial as an input into that. So stay tuned for more information on what we've got planned for that. Scrolling through some of these other questions, a way to submit suggestions for future NCVS supplements, you can always reach out to our Ask BJS email address. We'll put that into the chat too to provide suggestions. There is a lot of work that goes into new supplements. So that's not a, kind of, flipping the switch on or off kind of decision. But we're always interested in, you know, more ideas for what we can be doing and what we can be collecting with the NCVS. Another question, "Where to find the crime incident report data?" We will paste those links back into the chat and then when Lexi and Rachel get into the demos, they will probably cover a little bit of that too. Going through—from Charlie, "Can you talk about the upcoming NCVS redesign?" So, Charlie if you're interested in hearing about the work on that we've been doing on the redesign, I would suggest that you join us next week for our second webinar on updates on NCVS. We will be talking a little bit about some of those activities at that webinar and would love for you to come and hear us. And then Jennifer asks about looking at data by state. We are also working very hard to examine subnational estimates and are getting close to being able to put out some information on the work that we've been doing in the subnational program. We're not quite there yet but hope that we will be able to share more about the work we're doing with subnational estimation very soon. So with that, I think we will save the rest of the questions for the end.

ALEXANDRA THOMPSON: Hello, everyone. Welcome back from your break. I hope it was a good one. Now that you've learned about, you know, the NCVS—more about the NCVS, how the data are collected, and even heard from some practitioners that use the data more frequently, we're now going to go through some demonstrations on how you can run your own analysis and replicate some of the estimates from our annual reports. Next slide.

In case you haven't conducted analysis with the NCVS before, there are a number of resources available to you to kind of get you started and provide some helpful tips on analyzing the data. On NACJD's website, which you heard about earlier where we put up all of our data from the core NCVS and its supplements, there is a resource guide

and a codebook available for each dataset. Both of those go into more detail of the variables that are available, other types of information, along with the resource guide has some information about linking data and some helpful variables for that process.

We heard Dr. Martin talk about the technical documentation, which is also in your workshop materials that you should have received for this webinar. And that contains even more in-depth information on all the topics we've talked about today, like weighting, NCVS sample, along with more information on how to calculate estimates. There's also some users' guides on how to calculate standard errors for estimates in the NCVS like—or if you wanted to use generalized variance function parameters for your standard errors or if you wanted to use direct variance estimation. For example, like using BRR standard errors or Taylor series to calculate your standard errors. We also have included an example spreadsheet for calculating standard errors using GVF, which you'll see Rachel use later in the demonstration as well. There is also a learning guide on NACJD's website, which goes through a more step-by-step process of how you can calculate estimates using the NCVS. We've also some—we've also provided some of our more popular recodes, such as our type of crime recode, and our recodes for victim-offender relationship in your workshop materials as well. Next slide.

As I mentioned, we'll be replicating some estimates from the Criminal Victimization 2020 Bulletin. The estimates we'll be calculating today are highlighted. We'll be looking at the total number and the rate of violent victimization and the total number and rate of overall property victimizations. Next slide.

This slide is kind of an overview of what we'll go through in the demonstration. I'll start off today using SPSS and showing you the three different files that are available to you on NACJD's website and were mentioned by Grace earlier in the presentation. I'll start with the incident file, which is used to generate the number of violent and property crimes. And the incident file in terms of its structure, it contains an entry for every incident that we recorded through the survey. And for each incident, it has information such as what was the type of crime the victim experienced, what was the location of the incident, what the victim-offender relationship was, or whether the crime was reported to the police, among lots of other variables that you could look at related to the incident. All this goes through the person file, which is used to generate the number of persons aged 12 or older in the US. And the person data file has an entry for each person that we surveyed, regardless of whether they were a victim of a crime or not. And also included in this dataset is information about that person like their race or ethnicity, their sex, their—and other demographic variables related to that person. We also have the household data file which we'll conclude with to generate the total number of households in the U.S. And similar to the person file that has entries for

both households that had a person who reported a victimization and households where no one reported the victimization. And there's also variables in the household file like, is that, residents in an urban, suburban, or rural area, is that in the eastern part of the United States? So other information about the residents. And then once we have the estimates calculated in SPSS, Rachel will go over how to calculate the rates and standard errors in Excel.

And I will switch over now to SPSS to show you an example syntax file. Just give me a moment to share my screen. So you should be able to see—to see the syntax file in SPSS. And this is also included in your workshop materials. If you'd like to go through this on your own time in more detail, you'll notice that along with the code, there are also some notes in gray. And these are just notes that we like to put in throughout our process to document what we're doing, not only for ourselves but also in case someone wants to verify the data and helps them keep track of, you know, what we're calculating and what's going on and what we're trying to accomplish. I will go over this in more detail to show you kind of the step-by-step process. I'm actually going to do that in Excel rather than in SPSS because that way you'll be able to see both the code that we use and what type of output it generates more easily at the same time. So typically, you would run this code in SPSS, you know, go through all these steps. Again, I'll go through it in more detail. And then you can output—or you can export what output you receive into Excel. And then you'll see how we use that when Rachel shows us how to calculate the rate.

But I am going to switch over to Excel again just for this presentation to—more easy for you to more easily see kind of what code we use and what it generates. All right. So you should be able to see this Excel spreadsheet showing the output. And if I can't, maybe someone other—one of the other presenters can let me know. So this is kind of the—what you get when you do that final step of exporting the data and your output, but just going it through this way for the presentation.

So the first step is you need to actually pull the file that you want to work with. And like I said, we'll be starting with the incident file here. It's got that kind of three label on—when you download it from NACJD's website. And before we get to calculating the actual estimates, we like to prepare the data a little bit to make analysis easier. For example, we incorporate certain recodes depending on what we're trying to analyze. And so here you'll see some examples of our type of crime recode, again, which is included in your workshop materials. This variable here, V4529, is the type of crime code variable that is already prepared in the dataset when you download it from NACJD. And it contains all these different types of crimes like #7 here is robbery without an injury. And when we're preparing a report, we usually report on kind of

broader categories of crime, which are listed here like rape and sexual assault, robbery, aggravated assault, simple assault, et cetera. And so this step is just recoding those smaller types of crime into the broader categories.

We also have a different type of crime recode. More specifically, this one is to look at violent crime excluding simple assault, which is another broader category we report on in addition to violent crime. Do you see—as this goes into the variable of new crime too. And then the final type of crime recode, which is more of our focus today, and we'll use throughout this syntax is this new crime variable which categorizes crime into violent, personal larceny, and property victimizations.

This text is mostly just for kind of cleaning up these new variables that we created, adding, you know, a variable label or name, and then just adding some more documentation to the different categories. Another important step to take before analyzing NCVS data is to make sure that you exclude victimizations that occurred outside of the United States. We do this for all of our reports. The variable here, V4022, refers to the location of the incident. And so we're excluding any victimization that occurred outside the US or had a value equal to one. And then this line, this frequencies line, just generates a table so we can double-check if that category was removed. So that frequency line generates this table here, you can see there's no one. So we know that that category was properly excluded.

Now we'll be able to calculate the actual estimates. The NCVS estimates are weighted either for the incident or for victimizations. We're looking at victimization today. So we'll weight by the variable series weight, which is the victimization weight adjusted for serious crimes. And then to get the victimization by violent and property crimes is this section of code here. We'll select for only the year 2020 which is all that we're interested in. And then we'll look at it by those different crime variables. So that section of code generates this table right here. And so you can see there's, you know, 2020, this is the year that we want. And then we have violence, personal larceny, and property crime. So here are the estimates that we are interested in for our purposes today. We also like to look at the un-weighted count or the number of sample cases, just to make sure that none of the categories fall below 10. Currently, for our reports, we flag estimates if they are based on 10 or fewer sample cases. So we just want to double-check that before we proceed. So we just take—we remove the weight so that we're looking at the un-weighted count, we prepare a very similar table, and then double-check these numbers here. Again, we got 2020, violent, personal larceny, and property, none of these numbers fall below 10. So we're good to go. None of these estimates need to be flagged. And then moving on, once we have the victimization estimates, we now want to look at the number of persons age 12 or older in the US.

For this, we have to switch over to the person file, which is this command right here. We'll then weight the data using the person weight that is included in the data when you download it. We're again selecting for only 2020 and just getting that total estimate of persons in the U.S. age 12 or older. So that code generates this table right here and here is the estimate. And then we'll follow a very similar process for calculating the number of households in the US. We'll first grab the household file which is this line of code right here. And then weight by the household weight, select for only 2020, just to get that total estimate of the number of households. So this table gets generated right here and again, we have the estimate sitting in this cell right here. Once all of that is completed, we can just export the data, which is this—which is caused by this section of code right here. Again, normally all this would be done in SPSS, but I just did it in Excel to show—just to show both the code and the output more easily associated with each other. I will then turn it over to Rachel, who will show you how to calculate the rates and standard errors in Excel.

RACHEL MORGAN: Thanks, Lexi. All right, I am going to—if you can pass the screen sharing over to me, please. All right, and I put it in the chat. If you were following along with Lexi, then you already opened up the NCVS workshop replication example Excel file. If you haven't opened that yet, please do that now as we're going to walk through that document. All right. And if one of my colleagues would interrupt me if they cannot see this, please.

So there are four tabs in this document and one of them is the output that Lexi just exported and just walked through, so that is the second tab. So it has all of the SPSS code and then output. And sometimes we like to highlight some of the major codes that we use or commands that we use just to make it easier to see what that command is doing. So that's just a preference on our end. And I did that here so we could more easily see the distinctions between the sections that we're going to talk about. The first tab is the victimization estimate tab. So we're going to go to this one in just a moment. But this is where we're going to calculate the rate and link the numbers and the unweighted numbers to this tab from the output. And then the last two tabs are how we're going to generate the standard errors. And for these standard errors today, we're going to use GVF so with the NCVS data, you can calculate standard errors in a number of different ways. We use—we have in the past used GVF a lot and they are calculated by the Census Bureau and basically fitted the estimates—the crime estimates are fitted to a regression model. And there is more detail in the technical documentation, a lot more detail on the GVS, there are four sets of GVS that we'll go through in a few minutes. And then the other way we calculate standard errors, which we're not going to do today is through direct variance estimation, which my colleagues talked about a little bit, either using SAS or SPSS. And in your workshop materials,

there's a lot of documents on details, depending on what software program you use, and how to generate standard errors using direct variance estimation, so either using a Taylor series or using BRR methods. And if you do go that route and have questions, feel free to reach out to us at another time, because we're not going to get into details about direct variance estimation today. All right.

So if everyone would go to the output tab, we're going to scroll down to—one thing I did want to note when Lexi was talking about crimes that occurred outside of the United States, just to give you some more context. So, you know, we're interviewing folks that their households are located in the US, but they may have experienced the crime when they were on vacation or something, you know, in the Caribbean or in another country and we don't want to include those in the count of crimes that occurred in the United States for the year because they didn't occur in the United States. So that was just one extra point I wanted to make on that—on that note.

All right, so around line 123, we have the estimate for table 1, which is the violent crime table in the bulletin. So we have—we're going to pull the violence number and the property number. And if you're not familiar with linking in Excel, this is how we do it because someone is—someone else in the unit is verifying all of the numbers that we produce, all the reports we put out. So we link things so it's easier for that person to find exactly where that number came from. If they just had, you know, a number in one of these spreadsheets, and they had no idea where it came from, and they're trying to match it to something, it could take them forever to do that. So we link things to make it easier. So if you will—so push the equals button on your keyboard in this numbers column and the violent crime row, and then move over to the output tab, and select the cell that has the violence number, so it's 4.5 million, and then hit enter, it will link it. And so you can see up here at the top, that it tells you exactly where that number came from. So that's the key for us so that we know where that estimate came from. And then if we go down here to the property crime row, press the equals key again, and then move over and select the property crime estimate, which is about 12 million property crimes, press enter, and then we have our number of property crimes.

We're also going to pull in the unweighted estimates and we do this just so it's all in the same—in the same spreadsheet for colleagues to see but you can just like, you know, take a quick like view at these estimates and see that none are less than 10. But for the purposes of this, we'll do—we'll do an equals again, and then we'll link these estimates just so we have them all in the same space.

And I know there was a question in the chat earlier about looking at unweighted estimates versus weighted estimates. We always look at the unweighted numbers just

to check the sample sizes and make sure that they look okay and the estimates are going to be reliable to put out. But we always publish the weighted estimates because we are—we're putting out, you know, national estimates on victimization that occurred in the United States. So we always want to weight the estimates to the population 12 or older, or the population of males or females, et cetera. So we checked the unweighted ones, but we're always publishing on the weighted estimates. And there are a number of weights on the—on the files—on the incident file and if you scroll down to the end of the workshop slides, there's some more details that we didn't have time to discuss today about the different weights. And there's also a lot of information in the codebooks through NACJD with extra details on the weights. We use the series_weight and that's the one that adjusts for series victimizations because we don't want those series victimizations to be outliers in the data and be kind of messing with what the estimates may look like. But, you know, as long as—I think as long as in your analysis, you have a reason to use a non-series weight or you note that when you're writing your journal article or doing your research, I think it's up to the researcher but for us, we use this series victimization weight. So I hope I answered that question, if not, feel free to follow up.

All right, and then we're going to scroll down a little bit more to get the person population. And this is the population of persons in the United States, age 12 or older in 2020 and we are going to—again, let's go—we can put them right here, so if we press the equal sign and then we're going to link to 278 million, and then we'll fix the format of that in just a second. But let's also link the household, so keep scrolling, and we have 127 million, and then I'm going to right-click here, format cell, just to make them whole numbers. All right. So now we're going to calculate the rates. So our rates are—it is going to be the numerator, which is—so we're going to do equals, we're going to do the number of violent victimizations as the numerator divided by the number of persons as the denominator and then multiplied times a thousand. So we're looking at a violent victimization rate per 1,000 persons age 12 or older. And then we're looking at a property crime victimization rate, we're going to do the numerator, which is the number of property crimes divided by the number of households, times a thousand. So that is going to be and hopefully that matches, you know, we can always check against the em dash tool and check against, you know, the criminal victimization bulletins, to make sure that your estimates match up with ours.

And now, we're going to move over to the significance testing sheets. Now, these sheets—these sheets are very special, I would say, Census Bureau created these for us. And there are a lot of built-in formulas and things. So we highlighted the cells that you really need to worry about and fill in when you're trying to generate standard errors. And you don't need to worry about any other cells. There are clean copies of

the sig sheets saved, or in the materials that Daryl sent to you. So there's one sig sheet we use if you are trying to get the standard error for a number so, you know, the 4.5 million or the 12 million, and then there's another sig sheet for rates and percent. So in this case, we're going to look at rates. But if you see right up here in row two, enter one if you want to compare rates, enter two if you want to compare percent. So say you wanted to look at the percentage of violent victimizations that were reported to police, you would change this number to a two, it would change column E to percent and then you could go about filling it out the way that I'm about to show you. But in this case, we're going to look at rates per thousand so we're going to change that to a number one.

This other question here about rhos, so I believe Tara mentioned this briefly, but the rhos are correlation coefficients. So because the NCVS is interviewing the same household over a three-and-a-half year-period of time, we need to make sure that the correlation coefficients are accounted for. So the other document I would like you guys to open that was in your materials, if you haven't already, is the GVF parameters in rhos, which is another Excel document. And this document, there may be special formatting on here where you can't change things or you can't delete or something. Just because that's how we deal with it in-house because it is such a very—a well-used and protected document that we have because it has all the GVF from, what, 1992 to 2020 so we want to make sure it stays exactly as it is. So we're going to go down to 2020 here. And this date right here is when Erika added the 2020 GVF in the documents, so don't worry about that. But there are four sets of parameters or GVF for each year. And there is more detail in the text doc on this. But there are two for person crimes, so violent crimes, and then two for property crimes. So depending on what estimates—or what standard errors you're trying to generate. So, and then the other— one of the other tabs, you don't use these last two tabs, but the second tab is the rhos tab and your document may not have the second tab. I think I opened a different one. But the first two tabs are the same as ours. So don't worry about it if yours doesn't have these other two. But these are the one and two-year correlation coefficients are all the main types of crime. So, for this, we're not going to use it for this demonstration because we're just looking at 2020 estimates, but if we're going to compare 2019 rate of violent crime to the 2020 rate of violent crime, we would select the row that is in row nine for crimes of violence and select the 2019, 2020 row and put it in over here on the right in this keyed Rho part of the document. But we're not going to do that today, but just so you know. And if you have questions followed up on the rhos since we're not using those today, please just email us and reach out.

So what we're going to do, let's start with the sig number sheets. This is the total violent crimes. So what we're going to use is we're going to use the overall person

crime GVF estimates. We're not going to use the domain estimates, those are for if you're disaggregating violent crime types or looking at violent crime reporting to police or victim/offender characteristics, et cetera, then you would use these domain estimates. But since we're looking at the overall violent crime estimate, we're going to use the overall GVF.

So what I'm going to do is I'm going to select control and copy all three of these numbers, however you wish. And we are going to paste them in this A, B, and C column. And as you can see, in this document, it aligns with the ABC. And I'm just going to say to my colleagues if there's any questions that you feel like I need to answer immediately that folks are asking, please jump in and stop me from moving on. All right. So we are going back to the replication example. So the estimate is just going to be the victimization estimate, so we're going to do an equal sign again, so we can link it. We're going to link to this 4.5 million and since we don't need a row, this is the standard error right here, is in the L. So that is the standard error for this victimization estimate. Now we're going to do property crime, so we're going to go back to the GVF document. Again, we're going to do the overall property crime estimate since we're looking at the total property crime victimization. We're going to copy those three numbers, the ABC. Go back to our sig sheets and copy those in. And then again, we're just going to link our 12 million number, so that we have those. So there, the 263, that is your standard error for property crime victimization that occurred in 2020. All right. And then the sig rate sheet is pretty similar. The only—there's two big differences. The first is that you just do the B and C parameters. You're not going to copy in the A parameter. So we'll go ahead, and actually I'm just—for the sake of time, I'm just going to copy both of these B and C parameters from the other sig sheet. Put it right here. The base is going to be the denominator. So the base is the denominator of the rate that you've created. So for violent crime, it's going to be the number of persons 12 or older. For property crime, it's going to be the number of households, so we're going to just link those, and then the actual rate. So—and in the workshop slides, there are step-by-step instructions for this. Just so if you want to go back and redo it on your own, become more familiar, feel free to do that of course. And then we have the standard errors over here on column M for the rate of violent victimization of 2020 and the rate of property crime in 2020.

The other thing I will point out is, if you scroll down here, there are all these different steps. And this—so for example, this one to two sets corresponds to the set over here on the left, the one to two rows. So let's say if you wanted to do significance setting between the 2019 victimization rate and the 2020 violent victimization rate. What you would do is I'm—we're going to—I'm going to do this quickly, but I want to show this example. We're going to copy these, and so we're going to say this is the 2019 violent

victimization rate. Oh, nope. Sorry, that is 2020; 2019—what happens when I do this on the fly. And you could of course run the 2019 number, off the top of my head, I think it's, what, five—we're going to say 5.2 million is what I'm remembering right now. But of course, you want to get that exact number. So we're just going to put that in here for now as an example. And we're going to pull—because this is 2019, we're going to pull these overall person parameters for 2019. We're going to copy them right here, paste. All right. And so this is set five to six, so the set on the left and we're comparing rows five to six. What we're going to do is we're going to scroll down to five to six, which is right here. And it's going to show you—basically, there's a bunch of other formulas built in here, looking at P values and confidence intervals, and it's going to tell you whether those two numbers are statistically different or the same. And right now, we're showing that this significance is at 10 percent. So then you would be able to report that those estimates were different. So quick example on that just so you know how the steps work at the bottom. Let me—I think those were the main things. Let me check my notes and make sure that I'm not missing anything major to tell you. As I said, there are step-by-step instructions in the slide. And I think—I think that's it. So if we can go back to the slides, I will pass the ball back to Daryl and we'll follow up with the slide on the resources.

All right, so updates from BJS. We talked about some of these in the chat. But subscribe to JustStats which is the BJS listserv, and you'll get up-to-date information on any data releases, report releases, and data tool updates. And of course, you'll get, you know, information on like the webinars. That's how those came through, so we have two webinars next week. Please sign up for those. The first is the updates from BJS on the NCVS. And then, we have another one on Measuring Hate Crimes in the U.S., so both of those will be really interesting. And then of course, please follow BJS on Twitter and Facebook for updates.

Next slide, I think that's the end. So thank you. There's all of our contact information. Please feel free to email us if you have any questions, you know, pick one of us to email and that person can move it to, you know, someone else if needed based on expertise or whatnot. So thank you all very much. And we will open it up for questions and answers now.

HEATHER BROTSOS: Oh, I'm back. First of all, thank you to all of our presenters today. All of our wonderful BJS statisticians as well as our special guests, Keith and Tara, fantastic presentation. We've got just about 12 minutes left for some questions and answers. So if you haven't put your questions in, please do that now. We'll try to get through as many as we can. If we don't get to your question today, feel free to email us at the contact provided on this slide or you can also email the AskBJS—the

AskBJS email box. That is a monitored email box, and we will get routed those questions directly to be able to answer them. So a couple of things to start off with, Rachel, maybe you can take this first one. You went over this a little bit in the demonstration that I think it bears repeating. If you can talk a little bit about the difference between the different types of files that we—that we have within, you know, any given year for the NCVS, what each of those files is.

RACHEL MORGAN: Yes. I was just looking real quick. Daryl, if you will advance to a couple of slides down, there's one on the public-use file structure. Yes, perfect, that one. So the way NACJD releases data files, we release a concatenated file, which is 1992 to the present year. And there are three data files that are released with that concatenated file. There is an incident file, a household file, and then a person-level file. And if you look here, the way the variables are named, indicates what level basically that they were—what part of the survey they were asked. So 2000-level variables are house-hold variables, 3000-level variables are person-level, and then 4000 are incident-level variables. And so you're going to want to download all three of those files from NACJD with the annual files so say just the 2020 download from NACJD, there's going to be five files associated with that. As long as you have the household, the person, and then there is an incident extract file, then you're going to be good to go there. That answer?

HEATHER BROTSOS: Yes, thanks, Rachel. And then related to that, when you download those files, you can download them in different formats. So if SPSS is not your preferred package, you can download them in data or the delimited file or SAS, there's a few different options there when you go to do that download. Our next question is asking about the information collected on—where did it go—on incidents, was there any info on the dataset on when crimes occurred more specific than the year? Lexi, do you want to answer that one?

ALEXANDRA THOMPSON: Yes, we do ask questions about when the incident occurred. If you look in the incident data file, there are two variables related to one, the timing of the incident. One variable is the month the incident occurred and the other variable is the year the incident occurred. Those are variable V4014 for the month, and variable V4015 for the year the incident occurred.

HEATHER BROTSOS: Great. We've also got a couple of questions about the supplement, which we did not cover in today's presentations. We really focused on the core NCVS, but Jenna, would you be able to just give a quick answer about the reference period that's used in the supplement on how that's different from the core and how—or how we collect if we ask those questions differently?

JENNIFER TRUMAN: Yup. So the supplement generally use a 12-month reference period. And so because there are generally prevalence measures and so we're asking if they had experienced things like stalking or identity theft or fraud in the last 12 months, they are, for the most part, structured somewhat similarly in a sense that we connect green asking about their experiences. And then as they indicate that they may have experienced stalking or may have experienced identity theft, then we ask them detailed level crime, like an incident report similar to how we do with the core NCVS. So it's similarly structured there and I also—I know I saw this in the chat so, I'll just address this quickly. There was a question about whether or not there are GVF different for the supplement, and there are. So if you are interested in using the supplement data, they have separate GVF and those are released with the codebooks for that data collection.

HEATHER BROTSOS: Awesome, thank you, Jenna. Scrolling through all of these great questions, I see some that are—that are somewhat specific to a particular type of analysis or a specific method that you're trying to employ. I don't think we'll get into the detail of all of those today, but if you do have those types of questions, those are things you can definitely shoot us an email and we're happy to provide a response to those through email and help you work through any of those issues that are specific to your analysis. Just doing one more scroll through the other and then just one more plug that you can also email AskBJS, and email any of us. And we will—we will get back to you promptly. We are excited for you to use the data and we would love to see what you're doing and how you're using the data in the work that you are doing. So with that, I think we have gotten to the end of our program. I want to thank everybody again for attending and taking time out of your busy schedules to hear about the NCVS. Thank you again to all of our presenters. Thank you to Daryl and Tammy for making this a flawless program technologically and logistically. And we really enjoyed spending the afternoon with you all. So thank you so much and stay tuned for those webinars that we've got next week if you're able to make time. Thanks so much.