Vera INSTITUTE OF JUSTICE

Whitepaper: Counting How Counties Contribute to State Prisons

Oliver Hinds, Olive Lu, and Jacob Kang-Brown
March 20, 2018

Abstract

Understanding how counties in the United States use incarceration is crucial for developing criminal justice policy and for driving reform. Previously, we assembled a dataset containing jail incarceration information at the county level for the years 1970-2015. Here, we aim to compile information about how counties contribute to state prison populations. Toward this goal, we used data made available by the Bureau of Justice Statistics (BJS) through the National Corrections Reporting Program (NCRP), which contains individual state prison admission and release records for the years 1983-1999, and individual prison term records for the years 2000-2015. Since records in the NCRP data include county of commitment, we were able to aggregate records to estimate prison admissions and population at the county level. We supplemented data from the NCRP with data from state Department of Corrections reports when the NCRP data was missing or corrupt. We validated our estimates by summing data across all counties in a state, then comparing these with state-level prison admissions and population data available through the BJS National Prisoner Statistics Program. Through this comparison process and other validation approaches, we identified several challenges that impede accurate estimation of county-level statistics using the NCRP data. The result of our efforts is a novel U.S. county-level dataset of prison population and admissions counts from 1983-2015.

Introduction

Incarceration is a local phenomenon. The overwhelming majority of people in U.S.prisons and jails have been sent there by city, town, or county officials.[1] Despite this localization, most analysis of incarceration data occurs at the state or national level. Understanding, evaluating,

[1] "it is local courts that produce prison-sentenced felons, so explanations of the rapid increase in incarceration numbers should attend to how those court practices were transformed from the late 1970s on." Mona Lynch, "Mass incarceration, legal change, and locale: Understanding and remediating American penal overindulgence," *Criminology & Public Policy 10*, no.3 (2011): 673-698

and changing the way we incarcerate people requires that we analyze incarceration data at the county level.

Recently, we mapped jail incarceration statistics in U.S. counties from 1970-2015.[2] Jails are an important aspect of incarceration, and are often overlooked in national conversations around mass incarceration, especially ones that focus on data and statistics. Compared to jail data, data on prison usage at the state and national levels is widely available, but such data is not currently available at the county level. Although a recent *New York Times* article with analysis by John Pfaff mapped prison admissions at the county level for 2006 and 2014, no data on prison incarceration and admissions over time at the county level has previously been described.[3] Our goal with this work is to produce a dataset of prison incarceration and admissions for as many years and counties as possible.

The primary source for this dataset is the National Corrections Reporting Program (NCRP), which has been produced by the Bureau of Justice Statistics (BJS) since 1983.[4] The NCRP contains person-level incarceration data, including information on charges, demographics, prison admission and release dates, and the county from which the person was committed to prison. We aggregated this data by county of commitment to yield estimates of the number of people in prison and the number of new people admitted to prison from each county with available data from 1983-2015. We report counts of prison population and admissions by race and gender only when counts for these measures are high enough to avoid the possibility that individuals would be personally identifiable.

Working with the NCRP person-level data presents several data quality challenges. Several states did not participate in the NCRP every year (or at all), and there are some clear data errors specific to county-of-commitment reporting. In addition, states have different thresholds for determining who goes to prison and who serves a sentence in jail, based on charge and sentence length. In addition, estimating prison populations prior to 2000 presents challenges because the NCRP records from that period contain admissions and release data that are not explicitly connected to the same individual. This means that the effect of missing or erroneous data cannot be easily assessed or accounted for.

These aggregation and estimation challenges dictated that we perform extensive data validation and cleaning. The prison population and admissions timeline for each county in the U.S. was manually inspected for obvious low-level errors, omissions, and discontinuities. Then, state-level validation was performed by aggregating counts across all counties within a state and

---

[2] Jacob Kang-Brown and Ram Subramanian, *Out of Sight: The Growth of Jails in Rural America* (New York: Vera Institute of Justice, 2017)

[3] Josh Keller and Adam Pearce, “This small Indiana county sends more people to prison than San Francisco and Durham, N.C., combined. Why?” *The New York Times,* September 6, 2016, https://www.nytimes.com/2016/09/02/upshot/new-geography-of-prisons.html

[4] Bureau of Justice Statistics, “Data Collection: National Corrections Reporting Program (NCRP),” https://www.bjs.gov/index.cfm?ty=dcdetail&iid=268

comparing the total with the corresponding data from the National Prisoner Statistics (NPS) program, similar to the validation protocol used in the *New York Times* article referenced above. [5] Once errors in the NCRP data were identified, we supplemented that data with data collected directly from state Department of Corrections (DOC) reports. Usually, these reports do not provide complete prison population and admissions data disaggregated by race, ethnicity, and gender, so we preferred to rely on NCRP data when possible, even if state DOC reports were available.

Several stages of data cleaning were performed to improve correctness and continuity, followed by interpolation at the variable level (total, race, and gender counts individually) to account for missing or poor data. The result is a novel county-level dataset of prison population and admissions counts for the years 1983-2015. This dataset will be made available to the public.

Methods

Sources

National Corrections Reporting Program

Our primary source used to compile the county-level prison population and admissions counts is the NCRP dataset, which has been compiled each year since 2011 by Abt Associates and released by BJS. Prior to 2011, the NCRP was compiled by the U.S. Census Bureau. The NCRP seeks to provide records on all individuals incarcerated in state prison, and has been collected and reported since 1983. Crucial to our goal, records include prison admission and release dates, race, ethnicity, gender, and county of commitment. Every year, each state is asked to complete a questionnaire and to provide a report including information on each person incarcerated in the state. In practice, many incarcerated people are not included in the NCRP, due to incomplete participation in the program.

Conventions for the NCRP release have shifted since its inception.[6] Below we describe the two different conventions.

2000-2015 Recently, the NCRP has been released as a single dataset covering the period since 2000. All individuals that would have been included in individual year datasets are represented in the combined dataset by 'term' records, which contain information on each unique stay in

[5] Bureau of Justice Statistics, "Data Collection: National Prisoner Statistics (NPS)," https://www.bjs.gov/index.cfm?ty=dcdetail&iid=269

[6] W. Rhodes, G. Gaes, T. Rich, Y. Almozlino, M. Astion, R. Kling, and M. Shively, "Observations on the NCRP," *NCRP White Paper,* no. 1 (2012)

prison from admission to release. Computing prison populations and admissions for any year (or any date or date range) is straightforward using this release convention.

However, in assembling term records, Abt Associates uses only complete and reliable data. While this makes sense for compiling complete term records, we are ultimately interested in prison population and admissions, for which term records are not always necessary. Luckily, BJS also releases the 'extra' data: admission, release, and stock records deemed unworthy of inclusion in term records. We have incorporated this extra data into our estimates when possible, as described below.

1983-1999 The NCRP was originally released as independent datasets containing all reported prison admissions and releases for a single year. Counting the number of yearly prison admissions is straightforward given this representation, but computing prison populations without connected term records is more complicated. We describe our method for estimating prison populations below.

State Department of Corrections Reports

To supplement missing or incomplete NCRP data, we included county-level prison data gathered from annual reports released by the Department of Corrections of individual states where available.

State sources :

- Florida: Prison admissions (1999) by county and gender.[7]
- Kansas: Prison admissions (2011-2015) by county. [8]
- Michigan: Prison admissions (2014-2015) by county.
- Mississippi: Prison admissions (2011) by county.[9]
- Ohio: Prison admissions (2002-2006 and 2015) by county and gender.[10]
- Pennsylvania: Prison population (2015) and admissions (1983 and 2015) by county and gender.[11]

---

[7] Florida Department of Corrections, "Florida Department of Corrections Agency Annual Reports," http://www.dc.state.fl.us/pub/annual/

[8] Kansas Department of Corrections, "ARCHIVED: Annual Reports (Corrections Briefing Reports)" https://www.doc.ks.gov/publications/Reports/Archived

[9] *Admissions: General Characteristics,* (Mississippi Department of Corrections, 2011) http://www.mdoc.ms.gov/Admin-Finance/Documents/Annual-Reports/AnnualReport2011/14%20-%20Admissions%20-%20General%20Characteristics.pdf

[10] Ohio Department of Rehabilitation and Correction, "Annual Reports," http://www.drc.ohio.gov/annual-reports

[11] *1983 Annual Statistical Report* (Camp Hill, PA:Pennsylvania Bureau of Correction, 1983) http://www.cor.pa.gov/About%20Us/Statistics/Documents/Old%20Statistical%20Reports/1983AnnualReport.pdf; Michael Krausse, Runglux Kuba, and Dean Lategan, *2015 Annual Statistical Report* (Pennsylvania Department of Corrections, 2015) http://www.cor.pa.gov/About%20Us/Statistics/Documents/Reports/2015%20ASR%20Report.pdf

- South Carolina: Prison population and admissions by county, gender, and race (only Black, White, and Other).[12]
- Washington: Prison admissions (2015) by county.[13]
- Wisconsin: Prison admissions (2015) by county.[14]

National Prisoner Statistics Program

In validating the quality of our county-level prison population and admissions counts, we compared sums across all counties in a state with the state-level measures available through the NPS. Since 1926, the NPS has produced state-level prison population and admissions counts based on state DOC responses to a questionnaire administered by BJS. Most states participate in most years, and when states do not participate BJS attempts to impute missing data or acquire it from alternate sources. For the purposes of creating county-level counts from the NCRP and state DOC reports, we treated the NPS data as the gold-standard, except for recent prison admissions data from Washington, for reasons described in the **Anomalies** section.

Processing Steps

Pre-processing

Before aggregating individual NCRP records at the county level, we took steps to eliminate duplicate records representing the same prison term to prevent counting the same person more than once. By construction of Abt Associates, duplicate records only occur in the 'extra' data (only present from 2000-2015) not used to generate term records, which cannot have duplicates.

Duplicate records can occur in the extra data when a state reports data, but the quality of this data is too low to convert into term records. In this case, there may be multiple records for the same person. For example, a state may report a prison admission event, a number of stock records, and a release event for the same prison term. Although the state-reported record type (admission, release, stock) is provided, and this information could ideally be used to determine

[12] "Distribution of Committing County Inmates Admitted 2015" (South Carolina Department of Corrections, 2015) (https://dc.statelibrary.sc.gov/bitstream/handle/10827/20035/DC_Distribution_of_Committing_County_for_Inmates_Admitted_2015.pdf?sequence=1&isAllowed=y; "SCDC FAQs January 2016" (South Carolina Department of Corrections, 2016) http://www.doc.sc.gov/research/SystemOverview/SCDC_FAQs_January_2016.pdf

[13] "Prison Releases and Admissions by County - Fiscal Years 2006-2017" (Washington State Department of Corrections, August 2017) http://www.doc.wa.gov/docs/publications/reports/200-RE001.pdf

[14] *Prison Admissions: 2000-2016* (Wisconsin Department of Corrections, August 2017), 12, https://doc.wi.gov/DataResearch/InteractiveDashboards/DAIAdmissions2000to2016.pdf

unique cases, one or more of these records will often be missing. Combining all duplicate records allows us to create virtual term records with enough information to create prison population and admissions counts.

To identify duplicates, we search for records with identical county of commitment, race, sex, birth date, and prison admission date. We combine the duplicate records into a single record to maintain data that might be present in some records but missing in others, resulting in 'combined extra' records that serve as another set of term records for our purposes.

Aggregate prison admissions at the county level

The term and 'combined extra' records are processed to yield county-level prison admissions counts. We did not include prison admissions records flagged as returns from court or transfers from another jurisdiction, or people with a total sentence of less than one year. The result is a count of prison admissions by race, ethnicity, and gender for each county/year pair. Note that the NCRP-reported categories for race changed between 1998 and 1999 to separate the previous "Asian or Pacific Islander" category into two new categories, "Asian" and "Native Hawaiian or Pacific Islander".

Aggregate prison population at the county level

Estimating prison population at the county level is straightforward for years since 2000. Similar to the process for counting admissions, we simply count the number of term records with an admission date and release date bounding December 31st of the corresponding year. However, estimating prison populations for previous years is more involved than estimating admissions because there is no dedicated census or stock data released with the NCRP prior to 2000. Our approach to counting people in prison for the years 1983-1999 is to examine all release records in the NCRP and accumulate prison population counts for each year that the person was in prison on December 31st, based on the admission and release dates. To these counts we add counts of people who were released from prison after 1999 based on the term and stock records in the most recent, post-2000 NCRP dataset.

Post-processing

Verification, validation, and correction

Each year BJS reports prison admissions and population from each state. As a sanity check that our county-level estimates were reasonable, we summed all county-level estimates within each state and compared the total to the corresponding state measure from the NPS. Figure 1

shows an example of this process, which identified measures and data from Utah in some years as unreliable.

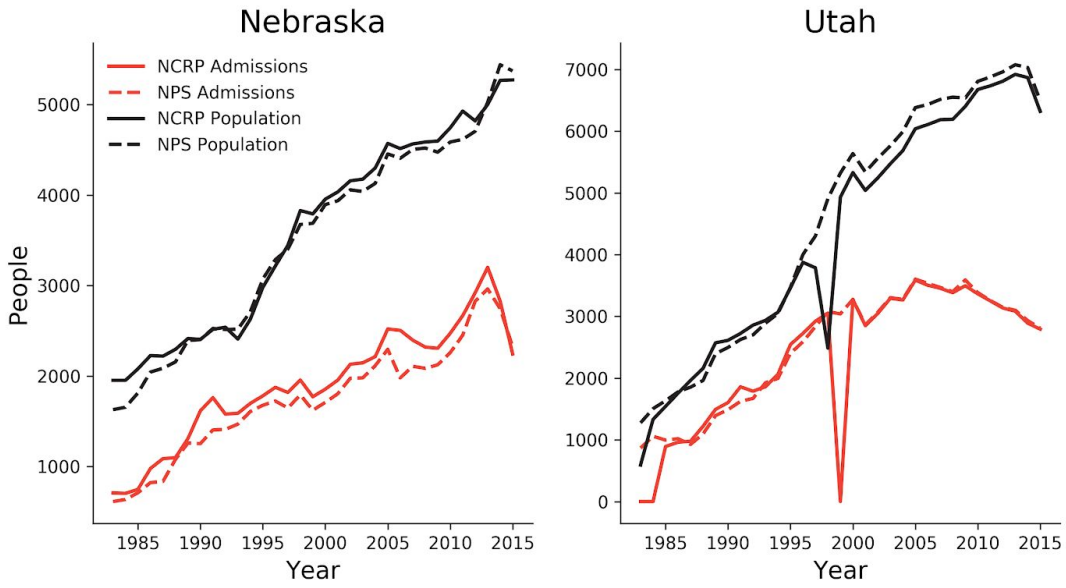

Figure 1: The county-level prison measures were partially verified by comparing the sum within state to the state-level measures reported as part of the NPS. The left panel shows that there is close agreement between the state-aggregated NCRP and NPS prison admissions and population measures over the entire time period for Nebraska. The right panel shows that NCRP data in Utah is unreliable for admissions in 1999 and for population in 1997 and 1998.

Even though comparing with the NPS data revealed many instances where prison admissions or population estimates from the NCRP required correction or interpolation, such comparisons cannot catch data issues at the county level. Figure 2 shows an example where, for the state of New York, comparing with the NPS catches the problem of missing prison admissions in 1983, but misses the problem of incorrect county of commitment coding in 1987, which makes the New York NCRP data unreliable for that year.

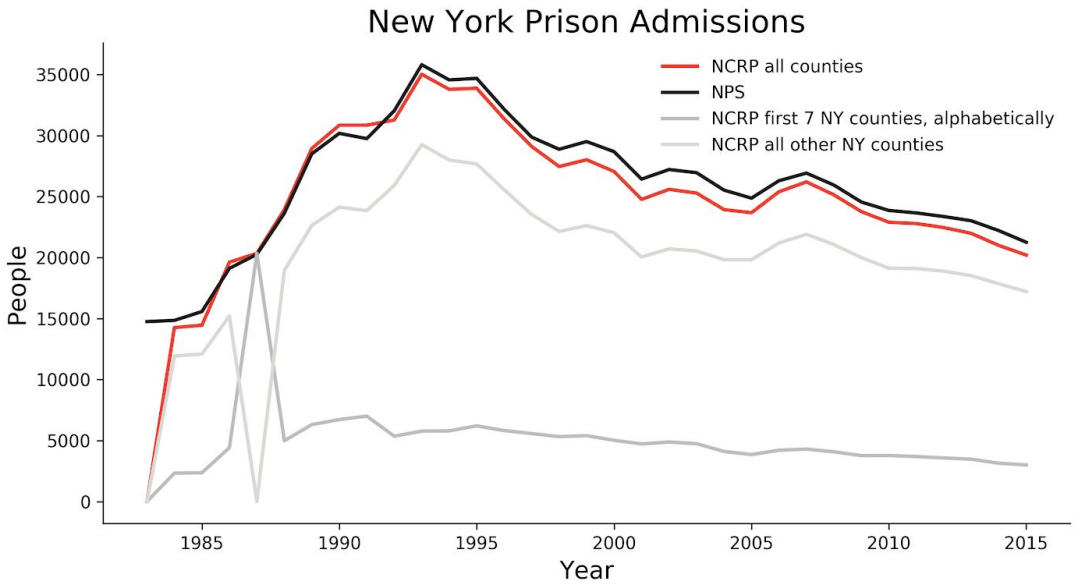

Figure 2: An error in the county of commitment coding for New York in 1987 means that all prison admissions are attributed to the first seven alphabetical counties in the state. Importantly, the sum of admissions across the entire state still matches the NPS-reported number almost exactly.

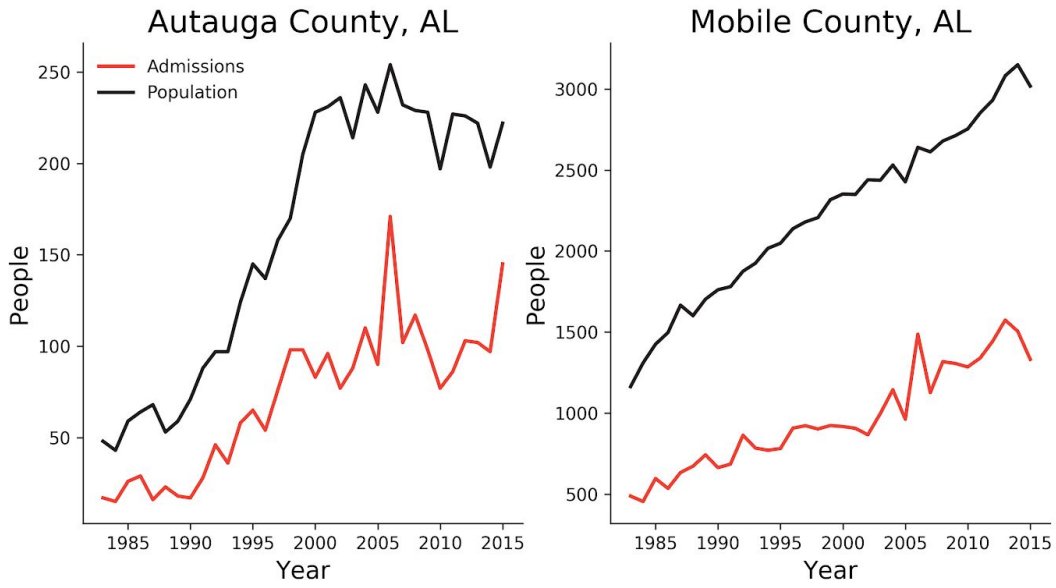

Figure 3: Examples of aggregated county prison admissions and population estimates that appear to be of good quality.

Motivated by the possibility of such problems, we manually inspected all prison admissions and population timeseries for missing data and obvious data entry errors. Figure 3 shows example timeseries for two counties in Alabama that seem complete and consistent, and were

considered reliable. Figure 4 shows timeseries for two counties in Virginia that had implausible fluctuations in both prison admissions and population in 2011, despite the sum of counties matching nicely with the NPS-reported total. Data for these county/year pairs was marked as unreliable and corrected as described in the following section.

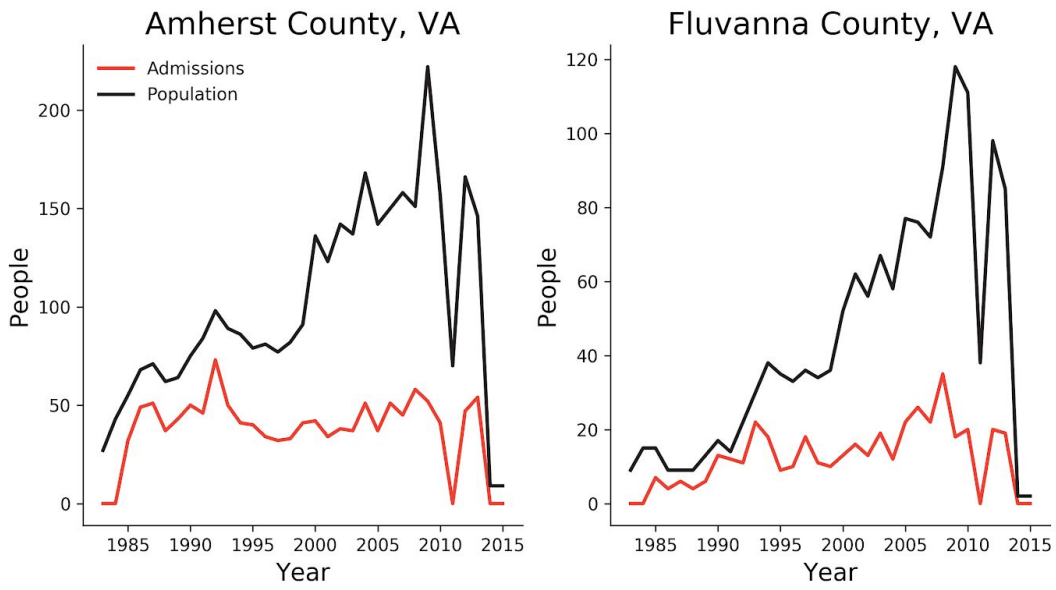

Figure 4: Examples of aggregated county prison admissions and population estimates that have data quality issues for 2011, despite the state total matching the NPS-reported data.

Manually inspecting each county's timeseries revealed several county coding issues in the NCRP. Issues include mis-labeled counties, missing county labels, and county labels that are consistently swapped. A complete accounting of these issues is presented in the Results section.

Data cleanup

Prison admissions and population by county were automatically corrected for unreliable, suspicious, or missing data. Processing steps taken during cleanup are as follows:

1. Data from county codes that were identified as swapped were corrected by unswapping counts.
2. Data collected manually from state DOCs was added.
3. Unreliable, suspicious, or missing data was interpolated.
4. Data that could not be interpolated because it was out of the bounds of reliable data was marked as missing.

Discard small samples

Finally, we discarded individual samples small enough to allow the possibility of individual identification. We used a threshold of 3, and discarded anything equal to or less than the threshold, but greater than 0. When the total count for a sample (the total admissions or population for a county in a particular year) met the discard requirements, all race and gender counts were also discarded. If the total did not meet the discard requirements, but either gender did, both gender samples were discarded, while the total was retained. If two or more race samples met discard requirements, only these samples were discarded. However, if only one race sample met the discard requirements, the race sample with next lowest count was also discarded so that the original small sample could not be recovered by subtraction.

Anomalies

Missing county of commitment in 2015

Due to a reported data processing error, the county of commitment is missing from the term records of Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Virginia, Washington, West Virginia, Wisconsin, and Wyoming in the 2015 vintage of the NCRP.

New York county of commitment in 1987

As shown in **Figure 2**, the county of commitment for counties in New York is incorrectly reported for 1987, with all counties lumped into the first seven alphabetical counties.

North Carolina county of commitment in 1995-1999

Reported county of commitment for North Carolina is shifted by one FIPS code for the years 1995-1999. For example, the data for Perquimans County, with FIPS 37143 was reported as the data from Person County, with FIPS 37145.

Individually swapped county of commitment

County of commitment for records from Clark County, Nevada (32003) was mislabeled as ‘Unknown County’ for the years 1990 and 2000-2006. Similarly, the county of commitment is exchanged on records from Monroe and Montgomery counties, Pennsylvania (42089 and 42091, respectively) for the year 2000 and after.

Washington NPS admissions

While validating our estimates against the NPS, we noticed a large difference between the NPS-reported and NCRP-estimated total number of prison admissions for Washington State for years after 2003. After some research into this discrepancy, we believe that the NCRP-estimated admissions count is more representative of a traditional “prison admission” than what is reported in the NPS (see **Figure 5**).

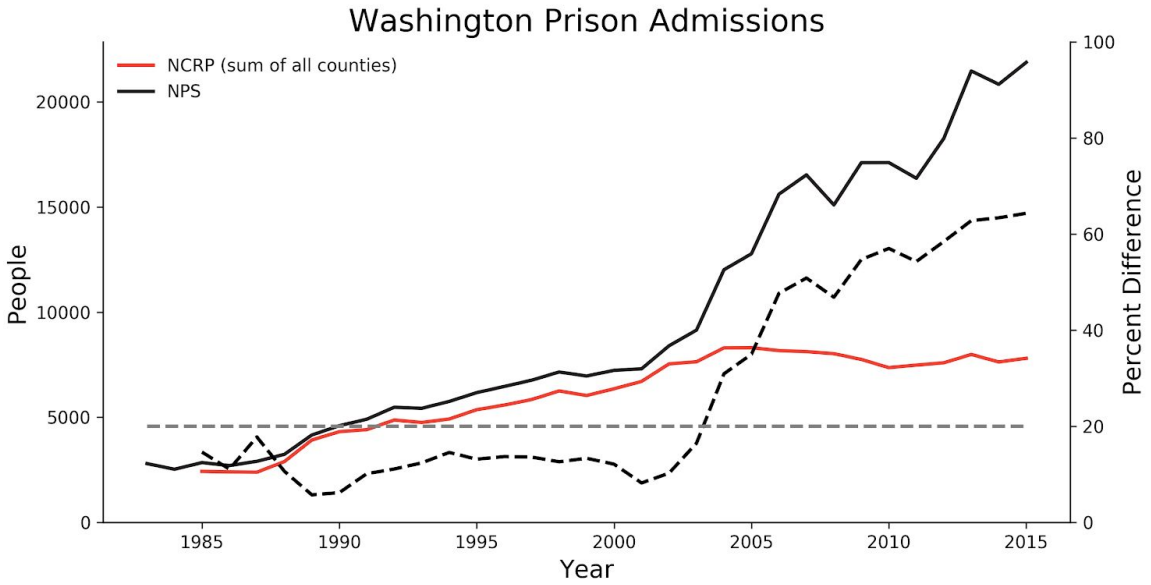

Figure 5: Comparison of the number of prison admissions from Washington reported as part of the NPS and NCRP. From 1985-2003, the numbers agree. After 2003, the percent difference grows rapidly to more than 60 percent in 2015.

From 2004 to 2008, the NPS Methods Notes for Washington stated that, “A recently revised law allows increasing numbers of certain inmates with sentences of less than 1 year to be housed in prison”, but did not refer to admissions specifically.[15] No relevant notes for Washington were included in NPS reports between 2009-2013. Finally, in 2014, an admissions-specific methods note appeared: “Admissions and releases increased due to the implementation of swift and certain sanctions for violation behavior, where an offender is arrested on the spot for violations

[15] Paige M. Harrison and Allen J. Beck, *Prisoners in 2004* (Washington, DC: Bureau of Justice Statistics, October 2005, NCJ 210677), 14; Paige M. Harrison and Allen J. Beck, *Prisoners in 2005* (Washington, DC: Bureau of Justice Statistics, November 2006, NCJ 215092), 13; William J. Sabol, Heather Couture, and Paige M. Harrison, *Prisoners in 2006* (Washington, DC: Bureau of Justice Statistics, December 2007, NCJ 219416), 13; Heather C. West and William J. Sabol, *Prisoners in 2007* (Washington, DC: Bureau of Justice Statistics, December 2008, NCJ 224280), 11; and William J. Sabol, Heather C. West, and Matthew Cooper, *Prisoners in 2008* (Washington, DC: Bureau of Justice Statistics, December 2009, NCJ 228417), 16

and is sanctioned to 1 to 3 days of confinement."[16] In 2016, a similar note is included, "Admissions and release counts of conditional release violators included offenders who received probation sentences and were sent to county jails for a term of fewer than 30 days for violating their probation conditions."[17] The 2015 NPS report contains a similar statement.[18]

People admitted to and held in county jails for less than 1 month, and people who serve sentences of 1 to 3 days do not represent traditional "prison admissions". We chose to accept our NCRP-estimated prison admissions counts, despite large disagreement with the NPS admissions counts post-2003.

[16] E. Ann Carson, *Prisoners in 2014* (Washington, DC: Bureau of Justice Statistics, September 2015, NCJ 248955), 24
[17] E. Ann Carson, *Prisoners in 2016* (Washington, DC: Bureau of Justice Statistics, January 2018, NCJ 251149), 35
[18] E. Ann Carson and Elizabeth Anderson, *Prisoners in 2015* (Washington, DC: Bureau of Justice Statistics, December 2016, NCJ 250229), 22