

Methodology and Data Notes

DATA SOURCES

[Long Sentences by the Numbers](#) uses data from the [National Corrections Reporting Program](#) (NCRP) Public-Use File distributed by the National Archive of Criminal Justice Data to generate trends analyses. The NCRP compiles individual-level data on admissions and releases from state prisons and post-confinement community supervision. Details regarding the NCRP and this specific file can be found [here](#).

METHODOLOGY

The analysis begins in 2005 because significant levels of prison population data were incomplete across multiple jurisdictions for earlier years. Only states that reported admissions and releases in the NCRP Term Records File for all years (2005 – 2020) are included in this analysis. The Term Records File was used for all analyses except for those referencing populations serving a long sentence. Data for those analyses were drawn from the Year-End Population File.

A total of 23 states were included in the analyses, representing 68% of the total United States population. These include: California, Colorado, Florida, Georgia, Illinois, Indiana, Kentucky, Minnesota, Mississippi, Missouri, Nebraska, New York, North Carolina, North Dakota, Oklahoma, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Utah, Washington, and Wisconsin.

Data for admissions and releases include people admitted to prison on a new criminal charge, and exclude people admitted to prison for parole revocation. Unless otherwise noted, release data exclude individuals whose releases are classified as “other” (which includes commutation/pardon, death, transfer, AWOL, escape, and other categories not listed).

Expected length of stay was estimated using the reciprocal exit rate described by [Patterson and Preston](#). While several estimation methods exist to generate expected length of stay, given the specific focus on sentences of 10 years or more, the reciprocal exit rate was selected as it performs the same or better than other methods. In Patterson and Preston’s (2007) analysis of individuals who were sentenced for murder, the reciprocal exit rate underestimated expected length of stay by 0.2 years at a 5% growth rate. In contrast, growth adjusted models overestimated expected length of stay by 1.7 years at a 5% growth rate. At a 5% reduction rate, the reciprocal exit rate underestimated expected length of stay by five years, while the growth models overestimated this value time by three to six years. As estimation methods are sensitive to large changes in prison populations, the 2020 expected length of stay estimate should be viewed with caution, regardless of which method is used.