

Data Needs and Methodological Challenges for Understanding the Long-Term Consequences of Criminal Justice Contact (in Longitudinal Survey Data)

Christopher Wildeman¹
Duke University & ROCKWOOL Foundation Research Unit

March 28, 2024

¹ A big thanks to Robert Apel, Martin Eiermann, Alexandra Gibbons, and Hedy Lee for helpful comments.

Overview


- ▶ Thanks!
- ▶ Why surveys deserve our attention
- ▶ Data needs and methodological challenges

Thanks!

Why surveys deserve our attention


They include core health outcomes

- ▶ Such as self-rated health and severe functional limitations
- ▶ Which are highly predictive of loads of other health outcomes
- ▶ And don't require contact with the medical system to do so.
- ▶ Some, such as the NLSY79, also include information about cognition and receiving an Alzheimer's or dementia diagnosis²

²The Cognition Module asks roughly 240 questions intended to test cognitive functioning including, among many other things, word recall questions, backwards counting questions, subtraction tests, and animal naming. 


They include core health outcomes

- ▶ Such as self-rated health and severe functional limitations
- ▶ Which are highly predictive of loads of other health outcomes
- ▶ And don't require contact with the medical system to do so.
- ▶ Some, such as the NLSY79, also include information about cognition and receiving an Alzheimer's or dementia diagnosis²

²The Cognition Module asks roughly 240 questions intended to test cognitive functioning including, among many other things, word recall questions, backwards counting questions, subtraction tests, and animal naming. 

They include core health outcomes

- ▶ Such as self-rated health and severe functional limitations
- ▶ Which are highly predictive of loads of other health outcomes
- ▶ And don't require contact with the medical system to do so.
- ▶ Some, such as the NLSY79, also include information about cognition and receiving an Alzheimer's or dementia diagnosis²

²The Cognition Module asks roughly 240 questions intended to test cognitive functioning including, among many other things, word recall questions, backwards counting questions, subtraction tests, and animal naming. 

They include core health outcomes

- ▶ Such as self-rated health and severe functional limitations
- ▶ Which are highly predictive of loads of other health outcomes
- ▶ And don't require contact with the medical system to do so.
- ▶ Some, such as the NLSY79, also include information about cognition and receiving an Alzheimer's or dementia diagnosis²

²The Cognition Module asks roughly 240 questions intended to test cognitive functioning including, among many other things, word recall questions, backwards counting questions, subtraction tests, and animal naming.

And they also include

- ▶ Some fun bonus outcomes like self-reported earnings!
- ▶ As Kornfeld and Bloom (1999:180) showed 25 years ago, mismeasurement of earnings in SSI/UI data is especially severe for the justice-involved, making this benefit of surveys vital

And a rich set of confounders

- ▶ Social marginalization
- ▶ Myriad health behaviors
- ▶ Crime and delinquency

And they are (mostly) paid for!

- ▶ Leverages existing data infrastructure
- ▶ Though, to skip ahead a bit, possible enhancements to existing survey data (such as differentiating whether someone was living in a prison or a jail or merging in rich administrative data like Mueller-Smith might discuss) would go a long way

Data needs and methodological challenges

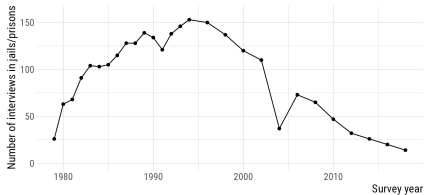
Data limitations

- ▶ Short jail stays likely far undercounted
- ▶ Attrition in all surveys (though some do quite well)
- ▶ Some only have self-reports of incarceration, and anything on arrests/charges/convictions is based solely on self-reports
- ▶ Representative structure can make causal modeling difficult

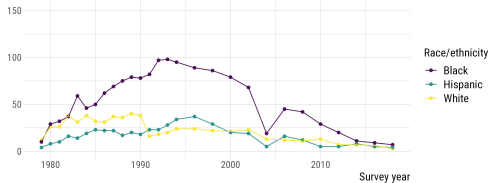
Data needs

- ▶ Enhancements would help!
- ▶ Can differentiate prison and jail in the NLSY series
- ▶ Could move beyond self-reports (via residence or linking)
- ▶ This would help build up an associational research base

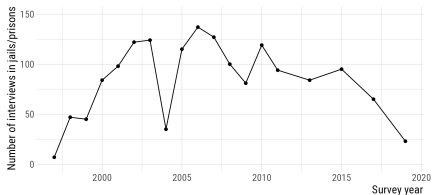
Number of NLSY79 interviews in jails/prisons by year



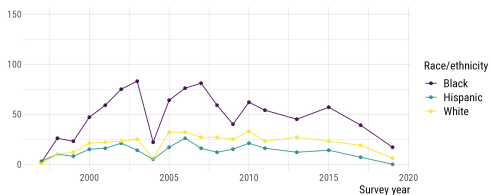
Number of NLSY79 interviews in jails/prisons by race/ethnicity and year



Number of NLSY97 interviews in jails/prisons by year



Number of NLSY97 interviews in jails/prisons by race/ethnicity and year



In general, survey years refer to the primary year of data collection; i.e. NLSY97 round 1 (1997-98) is listed as 1997.

Methodological obstacles

- ▶ Relatively limited methods for estimating long-term effects that can be used with most survey data we currently have (although can use geo-identified data to isolate policy shocks)
 - ▶ Lack of exogenous variation
 - ▶ Incarceration is often a repeated event, making common econometric methods (like judge IV studies) less helpful
- ▶ In short, it is much easier to identify incarceration as a **risk factor** for poor health than as a driver of poor health

Thanks again!