November 2020

Mass Incarceration, COVID-19, and Community Spread:

Methodological Appendix

Gregory Hooks
Department of Sociology
McMaster University
ghooks@mcmaster.ca

Companion report for *Mass Incarceration, COVID-19, and Community Spread* and *The early arrival of COVID-19 in counties and regions with large prison and jail populations* (available through Prison Policy Initiative (https://www.prisonpolicy.org/). This report provides details on measurement, methods, and estimation procedures.

I discussed these analyses and this report with Prison Policy Initiative staff (Wendy Sawyer and Peter Wagner) in June 2020. We discussed the likely community transmission of COVID-19 linked to mass incarceration. We agreed that it was an important issue and that I would pursue research on this topic. I collaborated with Wendy Sawyer in preparing *Mass Incarceration, COVID-19, and Community Spread*. However, Gregory Hooks (and not the Prison Policy Initiative) is solely responsible for the underlying statistical analyses[1] (and any errors or misinterpretations). This report complements *Mass Incarceration, COVID-19, and Community Spread* and the companion study *The early arrival of COVID-19 in counties and regions with large prison and jail populations* by providing details on data sources, measurement decisions, analytic approach, and statistical findings. This report is focused on methodological issues; *Mass Incarceration, COVID-19, and Community Spread* and *The early arrival of COVID-19 in counties and regions with large prison and jail populations* provide detail on the implications of the findings and includes policy recommendations.

Prisons and jails were uniquely vulnerable to the global COVID-19 pandemic. The US commitment of mass incarceration has long been criticized (Widra and Hayre 2020). The ongoing commitment to mass incarceration in the United States is an expensive, longstanding, and chronic policy failure. The impacts are widely felt (and the Prison Policy Initiative has helped to document this failure and these impacts, https://www.prisonpolicy.org/). As of 2018, it not only had the world's highest incarceration rate, but with less than 5% of the world's population, the United States accounted for more than 20% of all incarcerated persons in the world (Walmsley 2018). Mass incarceration fails those incarcerated, depriving them of civil and political rights, exposing them to unsafe and unhealthy living conditions, and burdening them with stigma long after they return to the community. Nor does mass incarceration help the larger society. Instead it distorts and compromises political participation, does little to reduce crime, and imposes steep burdens on incarcerated persons, families, and communities. Experts have condemned the United States for violating international norms, placing the health of incarcerated persons and staff at risk, and eschewing community-based options while relying on restrictive and punitive options. Given the longstanding failures of mass incarceration, prisons and jails were uniquely vulnerable to the COVID-19 pandemic.

> ***Part 1: Did mass incarceration bring COVID-19 sooner and more severely? (pp. 3-9)***
> Several logistic regression analyses were undertaken to answer this question. The dependent measures were unwanted COVID-19 events. The results of logistic regression (expressed as odds ratios) confirm the heightened risk posed by mass incarceration in a county and in nearby counties (as the number of incarcerated persons increased, so did the likelihood of these unwanted events). **This analysis is presented in the companion study, *The early arrival of COVID-19 in counties and regions with large prison and jail populations*.**
>
> ***Part 2: Did mass incarceration contribute to higher COVID-19 caseloads? (pp. 9-19)***
> To address this issue, Poisson regression models were used to estimate if and how many additional COVID-19 cases (per 100,000 residents) were linked to incarceration. For these analyses, the number of incarcerated persons per square mile — in the county and in the larger multicounty economic area (see Johnson and Kort [2004]) — were included as independent measures. The choice of Poisson regression, measurement of variables, analyses, and results are discussed below. **This is the focus of the main report, *Mass Incarceration, COVID-19, and Community Spread*.**

[1] Stata (statistical software for data science, https://www.stata.com/) was employed for all data management and statistical analyses discussed in this report.

Due to several advantages they offer, I selected counties as the unit of analysis. Counties cover the entire U.S. territory – therefore understanding dynamics at work at the county level provides insights into national trends. In addition, since counties are nested in states, the role of common factors such as institutional and political issues that affect counties belonging to different states can be assessed (Moller, Alderson and Nielsen 2009). Finally, county boundaries are relatively stable over time and, because data are compiled nationally by centralized agencies, data quality and availability are greatly enhanced (Isserman et al. 2009).

These analyses also consider the relationship between incarceration and COVID-19 spread at the level of multicounty BEA economic areas. Because the people who work in, provide services to, and otherwise interact with prisons and jails live and commute across a number of counties, infections linked to correctional facilities are *not* restricted to the counties in which a given facility is located. With commuting pattern being the primary criteria, more than 3,100 counties were sorted into 179 economic areas (Johnson and Cort 2004). The emphasis that BEA economic areas place on commuting makes these geographies well-suited for this analysis of "community spread" related to prisons and jails across a multi-county area. Those who work in one county but live in another county are exposed to the novel coronavirus in more than one county, and if they become infected, they can infect people in more than one county. I use these economic areas to get more directly at the question of "community spread" from correctional facilities outward to neighboring counties, using the same measures of incarceration and COVID-19 caseloads as in the county-level analyses.

To focus on incarcerated populations within the larger BEA area, I excluded each county's own incarcerated population in the BEA-level analyses. That is, for each BEA economic area, I aggregated the incarcerated populations of every *other* county in the BEA area, but did *not* count those held in the county itself. Thus, the county-level analysis shows how a county's incarcerated population has contributed to its own COVID-19 caseload, while the BEA area-level analysis shows how the incarcerated populations held in *other, nearby, economically-connected counties* may have contributed to the spread of COVID-19.

Part 1: Did mass incarceration bring COVID-19 sooner and more severely?

Data on COVID-19 cases are made available by the *New York Times*. Compiling data provided by state and local health officials, the *New York Times* (2020) has made available: "a series of data files with cumulative counts of coronavirus cases in the United States, at the state and county level, over time. [The *New York* Times collected] this time series data from state and local governments and health departments in an attempt to provide a complete record of the ongoing outbreak." These data are updated daily for each county, making it possible to examine several *unwanted* COVID-19 events:

- Presence: at least one case in the county by **April 1st**
- Presence: at least one case in the county by **May 1st**
- Significant caseload: more than 15 cases by **May 1st**
- Major outbreak: more than 250 cases by **May 1st**

Logistic regression was employed to assess the degree to which mass incarceration contributed to making these events more likely (net of other predictors). (For an overview of logistic regression, see Pampel 2000; Statwing 2020.)

The 2010 Census (U.S. Census Bureau 2020) provides data on the number of incarcerated persons in each county. The 2010 Census also provided information for several control variables (Black population, American Indian/Alaskan Native population, and Hispanic population). The Robert Wood Johnson Foundation (2020) made available several county-level health indicators: average life expectancy, adults *without* health insurance (percentage), and diabetes prevalence (for discussion of this data source and these measures in studies of COVID-19, see Chin et al. 2020). Table 1 provides information on the variables included in these analyses, including: data source, mean, and standard deviation.

Table 1
Information on Data Sources and Descriptive Statistics

Variable	Unit	Mean	Standard deviation	Source
Confirmed cases of COVID-19 (dependent measures)	Two possible values for each of the following events: • Presence: at least one case in the county by **April 1st** • Presence: at least one case in the county by **May 1st** • Significant caseload: more than 15 cases by **May 1st** • Major outbreak: more than 250 cases by **May 1st**	 0.678 0.878 0.515 0.128	 0.468 0.327 0.500 0.335	*New York Times* 2020
Black population (U.S. Census Bureau 2010)	Percent of total population	8.883	15.000	U.S. Census Bureau 2020
Native American / Alaskan Native population (2010)	Percent of total population	2.024	7.747	U.S. Census Bureau 2020
Hispanic population (2010)	Percent of total population	8.284	13.191	U.S. Census Bureau 2020
People detained by ICE in county (2020)	Two possible values: 1 if 2 or more people detained; 0 if 0 or 1 people detained	0.049	0.215	U.S. Immigration and Customs Enforcement 2020
People detained by ICE in multicounty BEA economic area (2020)	Two possible values: 1 if more than 25 people detained; 0 if 25 or fewer people detained	0.416	0.493	U.S. Immigration and Customs Enforcement 2020
Incarcerated persons in county (2010)	Six possible values: 1: 0 people (below 1st percentile) 2: 1-11 people (1-24th percentile) 3: 12-99 people (25th-49th percentile) 4: 100- 624 people (50th – 74th percentile) 5: 625-2,032 people (75th – 89th percentile) 6: more than 2,033 people (90th – 100th percentile)	3.672	1.322	U.S. Census Bureau 2020

Incarcerated persons in multicounty BEA economic area (2010)	Six possible values: 1: 0-1,419 people (1-9th percentile) 2: 1,420-4,482 people (10th - 24th percentile) 3: 4,483-12,438 people (25th-49th percentile) 4: 12,439 – 24,419 people (50th – 74th percentile) 5: 24,420 – 41,148 people (75th – 89th percentile) 6: more than 41,148 people (90th – 100th percentile)	3.572	1.471	U.S. Census Bureau 2020
Average life expectancy (2017)	Years	77.445	3.001	Robert Wood Johnson Foundation 2019
Adults *without* health insurance (2017)	Percent of adult population	13.288	6.092	Robert Wood Johnson Foundation 2019
Diabetes prevalence (2017)	Percent of population	11.626	2.597	Robert Wood Johnson Foundation 2019
Metro / nonmetro	Two possible values: 1 if metro; 0 if nonmetro	0.371	0.483	U.S. Department of Agriculture 2020

Logistic regression models were estimated for each of the COVID-19 events. The results are summarized in Table 2 (next page). Because the focus is on the degree to which mass incarceration increases the likelihood of unwanted COVID-19 events, Table 2 reports odds ratios for independent variables (for guidance on interpreting odds ratios, see UCLA Statistical Consulting Group 2020).

Table 2
Summary of Logistic Regression Analysis for Variables Predicting COVID-19 Events
(odds ratio, 3,071 counties)

	Presence in county: At least one case (April 1)	**Presence in county: At least one case (May 1)**	**Significant caseload: > 15 cases (May 1)**	**Major outbreak: > 250 cases (May 1)**
Incarcerated persons in county (categorical variable)	1.804***	2.038***	2.014***	2.282***
(ordinal measure, see Table 1)	(0.083)	(0.195)	(0.087)	(0.155)
Incarcerated persons in BEA area (categorical variable)	1.823***	1.298***	1.149***	1.116#
(ordinal measure, see Table 1)	(0.047)	(0.086)	(0.045)	(0.071)
Black	1.038***	1.127***	1.055***	1.047***
(percentage, 2010)	(0.005)	(0.030)	(0.005)	(0.007)
Hispanic	0.993	1.004	0.998	1.025***
(percentage, 2010)	(0.004)	(0.007)	(0.008)	(0.007)
Native Amer. / Alaskan Native	1.005	1.002	0.998	1.057***
(percentage, 2010)	(0.006)	(0.008)	(0.004)	(0.012)
People detained by ICE (county)	2.124*	1.910	1.635*	2.027**
(dummy variable, see Table 1)	(0.689)	(1.417)	(0.402)	(0.480)
People detained by ICE (BEA economic area)	1.137	1.254	1.140**	1.240
(dummy variable, see Table 1)	(0.123)	(0.223)	(0.149)	(0.203)
Average life expectancy	1.036	0.952#	1.053*	1.116**
(years, 2017)	(0.021)	(0.028)	(0.022)	(0.040)
Adults *without* health insurance	0.970**	0.926***	0.967***	0.927***
(percentage, 2017)	(0.009)	(0.014)	(0.009)	(0.016)
Diabetes prevalence	0.947*	0.951	0.941*	0.790***
(percentage, 2017)	(0.024)	(0.037)	(0.024)	(0.035)
Metro / nonmetro	3.903***	6.024***	4.065***	6.960***
(dummy variable, see Table 1)	(0.604)	(2.796)	(0.496)	(1.054)
Constant	0.021*	71.080	0.001***	0.000***
	(0.038)	(188.42)	(0.003)	(0.000)
Pseudo R^2	0.206	0.251	0.292	0.446

*** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; # $P < 0.10$

For each independent measure, Table 2 reports the odds ratio, standard error, and statistical significance. As reported below the table, when the p-value is less than 0.001, there is less than 1 chance in 1,000 that this finding would occur by chance. If the p-value is less than 0.01, this would occur less than 1 time in 100 (and so on). The odds-ratio estimates the impact of the independent measure in making the outcome of interest more likely. For example, consider the "metro/nonmetro" variable's impact on the likelihood of a major outbreak by May 1st. The reported odds-ratio is 6.960 and the p-value is less than 0.001 (indicating that the variable is highly significant). When compared to a nonmetro county, a metro county was 690% more likely to have experienced a major outbreak. The "metro/nonmetro" variable is a dummy variable, with only two possible values. Interpreting continuous measures and categorical variables is more challenging but follows the same basic logic. The odds-ratio reported in Table 2 can be interpreted as the percentage change in the likelihood an event occurs. An odds-ratio above 1.00 (above 100%) suggests that this variable makes the event more likely; an odds ratio below 1.00 (less than 100%) suggests that the variable makes the event less likely. As this research is focused on the impacts of mass incarceration, and measures of mass incarceration (in the county and the BEA area) are categorical variables, additional steps were taken to assist interpretation of the results (see below, Table 3).

In broad terms, the control variables performed as anticipated. Both the percentage Black population and the metro/nonmetro distinction were found to increase the likelihood of each of the four COVID-19 events. The impacts of other controls were mixed. The percentage Hispanic and American Indian / Alaskan Native only had a statistically significant impact on the most severe outbreaks (more than 250 cases as of May 1st). ICE detention facilities in a county increased the likelihood of 3 of these events, but this relationship failed to attain statistical significance for the presence of at least 1 COVID-19 case on May 1st. The BEA-level measure of ICE detention only met the threshold for statistical significance when predicting a significant outbreak (more than 15 cases) by May 1st. Results for the several health-related measures in the county were weak and inconsistent. Diabetes prevalence tended to make COVID-19 events *less* likely (odds ratio below 1.00), while average life expectancy is linked to increased likelihood of more serious COVID-19 events (more than 15 and more than 250 cases by May 1st).

Net of these controls, it is clear that mass incarceration elevated the risk of unwanted COVID-19 events. For both the number of incarcerated persons in the county and in the multicounty BEA economic area, the likelihood of COVID-19 events increases as the number of incarcerated people increases. The odds ratio for the county-level measure ranges from 1.804 (for presence of at least one COVID-19 case as of April 1st) to 2.282 (for a major outbreak of more than 250 cases by May 1st). Stated otherwise, this indicates that mass incarceration in a county roughly doubles the likelihood of COVID-19 events (ranging from 180.4% to 228.2%). When the focus is on the impact of mass incarceration in the BEA area, a similar, but less pronounced trend is in evidence. For this variable, the odds ratio was statistically significant for each event and ranged from 1.823 (or 80.3% more likely to have at least one case by April 1st) to 1.116 (or 11.6% more likely to report more than 250 cases by May 1st).

While Table 2 is helpful in demonstrating that mass incarceration made COVID-19 outbreaks more likely, discussing this in terms of odds ratio is not intuitive (Social Science Computing Cooperative 2014). Results "can often be made more tangible by computing predicted or expected values for hypothetical or prototypical cases…. Such predictions are sometimes referred to as margins" (Williams 2012, pp. 308-09). To facilitate interpretation, the logistic regression estimates were used to calculate the marginal impact of mass incarceration: all other variables in the model were held constant (at their respective means) and the impact of mass incarceration was estimated. The results of these estimations are presented in Tables 3 and 4.

Table 3

Marginal Impact of Mass Incarceration (County-Level) on COVID-19 Events (3,071 counties)
(likelihood of COVID-19 event when all other variables held constant, at their respective means)

	One (1) or more cases (April 1)	**One (1) or more cases** (May 1)	**More than 15 cases** (May 1)	**More than 250 cases** (May 1)
0 persons incarcerated	44.2%	78.8%	23.6%	1.6%
1-11 persons incarcerated (1-24th percentile)	56.3%	87.0%	34.5%	3.4%
12-99 persons incarcerated (25-49th percentile)	68.1%	92.7%	47.3%	6.4%
100-624 persons incarcerated (50-74th percentile)	78.3%	96.1%	61.0%	11.3%
625-2,032 persons incarcerated (75-89th percentile)	86.1%	98.0%	73.8%	18.4%
More than 2,032 persons incarcerated (top 10%)	91.6%	99.0%	83.9%	28.2%

Note: These results are displayed in Figure 1 in *The early arrival of COVID-19 in counties and regions with large prison and jail populations*.

Table 3 examines "scenarios" focused on mass incarceration. For each scenario, a specified level of mass incarceration is assumed; all other variables (see Table 2) are held constant at their respective means (an "average county"). The results are estimates of the percentage of counties that would experience a COVID-19 event. For example, if there were no people were incarcerated in an "average county," we would expect that 44.2% would have confirmed at least one case of COVID-19 by April 1st. The likelihood goes up as the number of people incarcerated increases. More than three-fourths (78.3%) of otherwise average counties in the 50th-75th percentile for mass incarceration (100-624 people incarcerated), would have at least one case by April 1st. For counties in the top 10% for mass incarceration, the likelihood is 91.6%. The trends are compelling for each COVID-19 event, including the most severe outbreaks (more than 250 cases by May 1st). Only 1.6% of counties with no incarcerated people would be expected to have a serious outbreak by May 1st. However, 11.3% of counties in the 50th-74th percentile for mass incarceration could expect a major outbreak by that date. And, for an otherwise "average" county above the 90th percentile for mass incarceration (more than 2,032 people incarcerated), more than one-fourth (28.2%) would be expected to confront a major outbreak by May 1st.

Table 4
Marginal Impact of Mass Incarceration (at the BEA Area Level) on COVID-19 Events (3,071 counties)
(likelihood of COVID-19 event when all other variables held constant, at their respective means)

	One (1) or more cases (April 1)	**One (1) or more cases** (May 1)	**More than 15 cases** (May 1)	**More than 250 cases** (May 1)
0-1,419 (below 10th percentile)	64.7%	88.4%	48.6%	11.5%
1,420-4,482 (10th-24th percentile)	67.6%	90.5%	50.8%	12.2%
4,483-12,438 (25th-49th percentile)	70.4%	92.2%	53.1%	12.8%
12,439-24,419 (50th-74th percentile)	73.1%	93.7%	55.4%	13.6%
24,420-41,148 (75th-89th percentile)	75.7%	95.0%	57.7%	14.3%
More than 41,148 (90th percentile and above)	78.2%	96.0%	60.0%	15.1%

Note: These results are displayed in Figure 2 in *The early arrival of COVID-19 in counties and regions with large prison and jail populations*.

Table 4 is similar to the preceding table in that it examines "scenarios." However, Table 4 is focused on levels of incarceration in the multicounty BEA area (and not the county itself). Several levels of mass incarceration are considered, while holding all other variables constant. It is important to keep in mind that the number of people incarcerated *in the county* is also assumed to be average (and there are several hundred people incarcerated in the "average" county). With this significant level of incarceration assumed to be in the "average" county, Table 4 suggests that even a county in a BEA area with relatively few incarcerated people is still at considerable risk of these COVID-19 events. Consider a county below the 10th percentile for incarcerated persons in the multicounty BEA area (fewer than 1,420 people in jail or prison in the BEA area). This county would have a 64.7% chance of at least one COVID-19 case by April 1st, and an 88.4% chance of a confirmed case by May 1st. This otherwise average county would have a roughly 50-50 chance (48.6%) of having 15 or more cases by May 1st, and this county would have a 11.5% chance of a major outbreak (more than 250 cases) by May 1st. For each of these events, an increase in incarcerated people in the BEA makes each of these COVID-19 events more likely.

Part 2: Did Mass Incarceration Contribute to Higher COVID-19 Caseloads?

The examination of mass incarceration's contribution to COVID-19 caseloads makes use of the same data sources (for the most part), but employs different estimation procedures. Using county-level data made publicly available by the *New York Times* (2020), I calculated the number COVID-19 cases confirmed between May 1, 2020 and August 1, 2020. As is common in health research, the dependent variable is *not* the absolute number of cases. Instead, it is the number of COVID-19 cases per 100,000 residents. Poisson regression was employed to estimate impacts on COVID-19 caseloads per 100,000 residents. When the dependent measure is a count (as it is in this case): "Poisson distributions represent an efficient method of estimating probabilities of events, particularly where the population size is large and the probability of an event is relatively low. This technique is often used with highly positively skewed distributions" (Osborne 2017, p. 283). In cases in which a dependent variable is not positively skewed and Poisson regression is not required, Poisson regression does *not* generate biased results; results would converge with ordinary least squares regression (ibid.). Because studies of population health often rely on count data (comparable to the dependent variable in these analyses), health researchers frequently rely on Poisson regression (see, for example, Frome and Checkoway 1985; Population Health Methods [Columbia University] 2020).

Many of the same independent variables that were included in logistic regression models that estimated the likelihood of COVID-19 events (see above, Tables 1 and 2) were also used in these estimates. There are, however, several important differences. First, both the county and BEA-level measure of mass incarceration are focused on the density of mass incarceration, i.e., the number of incarcerated persons per square mile. In preliminary models (results available upon request), alternative measures were used, including total number of people incarcerated and the number of incarcerated persons as a share of the total population. These alternative measures did provide evidence that mass incarceration increased COVID-19 caseloads. However, when mass incarceration was measured in terms of incarcerated people per square mile, the findings were more consistent (i.e., statistically significant across variations in modeling) and displayed stronger effects. Based on this preliminary modeling, I concluded that density does have an impact on the role that mass incarceration plays in the spread of COVID-19. Second, preliminary models also provided evidence that the impact of mass incarceration is different in metro and nonmetro counties. In light of this preliminary finding, I created slope-dummy interaction terms (Yobero 2017), yielding the following four measures of mass incarceration:

- Incarcerated persons per square mile (***county***) if ***metro*** (0 if *nonmetro*)
- Incarcerated persons per square mile (***county***) if ***nonmetro*** (0 if *metro*)
- Incarcerated persons per square mile (***multicounty BEA area***) if ***metro*** (0 if *nonmetro*)
- Incarcerated persons per square mile (***multicounty BEA area***) if ***nonmetro*** (0 if *metro*).

Finally, it is likely that the state-level institutional context and public health response influences the spread of COVID-19. For this reason, state fixed effects (a dummy variable for each state is included in the model) are incorporated into the analysis to capture omitted information about state-wide factors, such as public policies and pandemic mitigation efforts. Table 5 provides information on the variables included in these analyses, including: data source, mean, and standard deviation.

Table 5
Information on Data Sources and Descriptive Statistics

Variable	Unit	Mean	Standard deviation	Source
Confirmed cases of COVID-19 per 100,000 residents – May 1- August 1 (dependent measure)*	Cases per 100,000 residents	514.835	767.135	*New York Times* 2020
Confirmed cases of COVID-19 per 100,000 residents – as of May 1st*	Cases per 100,000 residents	21.663	58.856	*New York Times* 2020
Incarcerated persons per square mile (***county***) ***– metro***	Incarcerated persons per square mile	0.890	5.045	U.S. Census Bureau 2020
Incarcerated persons per square mile (***county***) – ***nonmetro***	Incarcerated persons per square mile	0.382	1.317	U.S. Census Bureau 2020
Incarcerated persons per square mile ***(BEA economic area***) ***– metro***	Incarcerated persons per square mile	0.461	0.846	U.S. Census Bureau 2020

Incarcerated persons per square mile ***(BEA economic area*) – *nonmetro***	Incarcerated persons per square mile	0.463	0.638	U.S. Census Bureau 2020
Black population U.S. Census Bureau (2010)	Percent of total population	8.885	14.503	U.S. Census Bureau 2020
Native American / Alaskan Native population (2010)	Percent of total population	1.965	7.382	U.S. Census Bureau 2020
Hispanic population (2010)	Percent of total population	8.291	13.195	U.S. Census Bureau 2020
Asian-American population (2010)	Percent of total population	1.374	2.812	U.S. Census Bureau 2020
Population density	Persons per square mile	259.281	1,724.937	U.S. Census Bureau 2020
People detained by ICE in county (2020)	Count	13.116	119.335	U.S. Immigration and Customs Enforcement 2020
People detained by ICE in multicounty BEA economic area (2020)	Count	245.599	548.525	U.S. Immigration and Customs Enforcement 2020
Less than 9th grade education (2010)	Percent of adult population	5.035	3.612	U.S. Census Bureau 2020
Non-citizens (2010)	Percent of population	55.413	19.053	U.S. Census Bureau 2020
Residents of nursing homes (2010)	Count	478.390	1307.910	U.S. Census Bureau 2020
Residents of group quarters (other than correctional facilities, nursing homes, military bases and university dormitories) (2010)	Count	24.356	104.117	U.S. Census Bureau 2020
Meatpacking plants experiencing severe COVID-19 outbreaks (county)	Dummy	0.024	0.152	Environmental Working Group 2020
Meatpacking plants experiencing severe COVID-19 outbreaks (BEA economic area)	Dummy	0.677	1.136	Environmental Working Group 2020

Note: A dummy variable for each state was included in addition to the variables listed above.

*For several cities (New York City, Kansas City [Missouri], and Joplin [Missouri]), local authorities only provide COVID-19 information for the city. However, these cities span county boundaries. In these cases, I distributed COVID-19 data to counties on the basis of population shares. This is not an ideal solution and, no doubt, introduced measurement imprecision.

As reported in Table 5, the 2010 Census (U.S. Census Bureau 2020) provides data for most sociodemographic measures (including the number of people incarcerated in each county and BEA area). News reports emphasized the frequency and severity of COVID-19 outbreaks at meatpacking plants. For this reason, making use of data compiled and shared by the Environmental Working Group, counties and BEA areas impacted by meatpacking plant outbreaks were identified (dummy variable). Table 5 does not include public health measures compiled and made available by the Robert Wood Johnson Foundation (2020). In preliminary models, a wide range of control variables – including these public health variables and other sociodemographic measures (e.g., poverty rate, median household income, etc.) were included. However, these models displayed high levels of collinearity. The control variables summarized in Table 5 were found to play a role in predicting COVID-19 caseloads (without collinearity).

Table 6
Predicting Confirmed COVID-19 Cases (May 1 – August 1, 2020): Poisson Regression Model
(N= 3,114 counties, coefficient and standard error multiplied by 100 to improve readability)

	Coefficient (Standard Error)
Confirmed cases of COVID-19 per 100,000 residents – as of May 1st	0.014***
	(0.000)
Incarcerated persons per square mile (*county*) – *metro*	0.330
	(0.003)
Incarcerated persons per square mile (*county*) – *nonmetro*	3.967***
	(1.110)
Incarcerated persons per square mile *(BEA econ. area)* – *metro*	10.362***
	(2.184)
Incarcerated persons per square mile *(BEA econ. area)* – *nonmetro*	7.351**
	(2.474)
Black population U.S. Census Bureau (2010)	1.045***
	(0.104)
Native American / Alaskan Native population (2010)	1.212***
	(0.229)
Hispanic population (2010)	1.909***
	(0.175)
Asian-American population (2010)	2.013***
	(0.603)
Population density	-0.003*
	(0.001)
People detained by ICE in county (2020)	-0.004
	(0.005)
People detained by ICE in multicounty BEA economic area (2020)	0.006**
	(0.002)
Less than 9th grade education (2010)	1.560***
	(0.579)
Non-citizens (2010)	0.483**
	(0.070)
Residents of nursing homes (2010)	0.006***
	(0.001)
Residents of group quarters (other than prisons, jails, nursing homes,	-0.032**
military bases and university dormitories) (2010)	(0.011)
Meatpacking plants experiencing severe COVID-19 outbreaks	21.211**
(county)	(6.140)
Meatpacking plants experiencing severe COVID-19 outbreaks	4.467***
(BEA economic area)	(1.183)
Pseudo R-square	0.6212

*** P < 0.001; ** P < 0.01; * P < 0.05; # P < 0.10

In Poisson regression analysis, an influential case exerts an outsized influence on the estimation. In so doing, the estimation is distorted by the influential case, making it less accurate for remaining cases included in the model. Preliminary models were inspected for influential cases (focusing on the Cooks D statistic). A number of the least populated counties exerted an outsized influence on the estimates (for several counties, Cooks D was above 25 – and in one case above 100). Recall that the dependent measure is the number of cases per 100,000. As such, even a modest increase in COVID-19 cases for a county with a small population results in a dramatic spike in the dependent variable. Data on outbreaks at meatpacking plants (Environmental Working Group 2020) were incorporated into the analyses in an effort to address these problems. While doing so proved helpful, problems persisted. Recent reporting indicates that meatpacking plants are not the only factor leading to COVID-19 outbreaks in nonmetro counties. There have been outbreaks across a range of agricultural sectors, including canning and poultry facilities. Detailed county-level data on these outbreaks were *not* available when these analyses were conducted; it is recommended that future research attempt to include more fine-grained analyses of outbreaks in the agriculture and food processing sectors.

To reduce biases in the estimates, twenty-six (26) counties (Cooks D above 3) were dropped from the sample when generating the estimates presented in Table 6 (leaving 3,114 of 3,140 counties in the sample). The total population in counties dropped from the analysis tended to be quite low: half of these counties were home to fewer than 10,000 people; 35 of the counties had fewer than 75,000 people. However, likely reflecting the severe outbreak in New York City, as well as the reporting challenges for the counties comprising the city (see above), two of New York City's boroughs (the Bronx and Manhattan) were also identified as influential cases and dropped from these analyses.

The coefficients for mass incarceration measures (at the county and BEA levels) were significant with and without these influential cases, but the coefficients were slightly larger with the full sample. The results presented here estimate a somewhat smaller impact attributable to mass incarceration, but I believe these more conservative estimates are more reliable.

On the whole, control variables performed as anticipated. That is, the following variables were found to be statistically significant and the impact was in the expected direction:

- Confirmed cases of COVID-19 per 100,000 residents – as of May 1st
- Black population U.S. Census Bureau (2010)
- Native American / Alaskan Native population (2010)
- Hispanic population (2010)
- Asian-American population (2010)
- People detained by ICE in multicounty BEA economic area (2020)
- Less than 9th grade education (2010)
- Non-citizens (2010)
- Residents of group quarters (other than correctional facilities, nursing homes, military bases and university dormitories) (2010)
- Meatpacking plants experiencing severe COVID-19 outbreaks (county)
- Meatpacking plants experiencing severe COVID-19 outbreaks (BEA economic area)

It was not anticipated that population density would be inversely (negative coefficient) and significantly linked to COVID-19 growth. This may reflect that in contrast to the initial wave, COVID-19 spread more rapidly outside of major metropolitan areas in the summer of 2020. Finally, in this estimate, I did not find a statistically significant relationship between "people detained by ICE in county" and COVID-19 caseloads. In the spring of 2020, COVID-19 swept through ICE facilities (much as it did through prisons). This non-finding at the county level (but a positive relationship at the larger BEA area level) might be the result of COVID-19's rapid spread in earlier months.

As noted, because preliminary analyses provided evidence that the effects of mass incarceration were different in metro and nonmetro counties, four measures of mass incarceration (incarcerated persons per square mile) were included in the estimation presented in Table 6 (metro - county; nonmetro - county, metro - BEA economic area, and nonmetro - BEA economic area). As this research is focused on mass incarceration, the results for relevant variables are reproduced and highlighted in Table 7.

Table 7
Highlighting the Impact of Mass Incarceration on COVID-19 Caseloads
(coefficient and standard error of Poisson regression estimation, as reported in Table 6)

	County (incarcerated persons per sq. mi.)	**BEA economic area (incarcerated persons per sq. mi.)**
Metro county (1,166 counties, 225,082 average population)	*Not statistically significant*	10.362*** (2.184)
Nonmetro county (1,974 counties, 23,441 average population)	3.967*** (1.110)	7.351** (2.474)

Three of these four variables achieved statistical significance, providing evidence that mass incarceration contributed to growing COVID-19 caseloads over the summer of 2020. However, for metro counties, the county-level measure of mass incarceration failed to achieve statistical significance. For this reason, the coefficient (as reported in Table 6) is not copied here, and this measure is *not* used when estimating the additional COVID-19 caseloads attributable to mass incarceration. For nonmetro counties, the coefficient at the county level is 3.967, and at the BEA level it is 7.351. For metro counties, the BEA economic area coefficient is 10.362. While these positive and significant coefficients provide strong evidence that mass incarceration accelerated the community transmission of COVID-19 across the United States, interpreting Poisson regression coefficients is not intuitively obvious. To facilitate interpretation, their marginal impacts were calculated (Table 8).

Table 8
Marginal Impact of Mass Incarceration on Confirmed COVID-19 Cases (May 1 – August 1)
(additional cases when all other variables held constant, at their respective means)

		ADDITIONAL CASES PER 100,000 RESIDENTS			ADDITIONAL CASES PER 100,000 RESIDENTS	
	Incarcerated persons per sq. mi. - County	Metro - County	Nonmetro - County	Incarcerated persons per sq. mi. - BEA area	Metro - BEA area	Nonmetro - BEA area
0 percentile	0 incarcerated persons per sq. mi.	0	0	0 incarcerated persons per sq. mi.	0	0
25th percentile	0.004 incarcerated persons per sq. mi.	0	0.12	0.27 incarcerated persons per sq. mi.	21.46	15.35
50th percentile	0.059 incarcerated persons per sq. mi.	0	1.87	0.74 incarcerated persons per sq. mi.	61.45	43.64
75th percentile	0.31 incarcerated persons per sq. mi.	0	9.74	1.40 incarcerated persons per sq. mi.	120.38	84.61
90th percentile	1.80 incarcerated persons per sq. mi.	0	116.26	1.94 incarcerated persons per sq. mi.	171.95	119.82
95th percentile	3.44 incarcerated persons per sq. mi.	0	280.88	2.29 incarcerated persons per sq. mi.	205.83	142.66

Notes:
- Because the coefficient for county-level measures of mass incarceration failed to attain statistical significance for metro counties, it is assumed that increased incarceration in metro counties does not result in additional cases of COVID-19.
- These results are displayed in Figures 1 and 2 in *Mass Incarceration, COVID-19, and Community Spread*.

Table 8 presents several "scenarios" focused on mass incarceration. For each scenario, a specified level of mass incarceration is assumed; all other variables are held constant at their respective means (an "average county"). The results (using coefficients from Poisson regression estimates) are in terms of the number of additional confirmed cases of COVID-19 that would be anticipated due to mass incarceration. An "average" county could expect roughly 780 cases of COVID-19 between May 1st and August 1st (2020). Table 8 reports *additional* cases on top of this baseline.

For example, if there were no people incarcerated in an "average" *nonmetro* county, we would expect no additional cases. Furthermore, mass incarceration has a negligible impact at the 50th percentile (0.059 incarcerated people per square mile) among nonmetro counties: we would expect 1.87 additional cases per 100,000 residents. But the additional caseload rises sharply as mass incarceration rises: we would expect 9.74 additional cases per 100,000 residents at the 75th percentile, 116.24 at the 90th percentile, and 280.88 at the 95th percentile. In broad terms, the same basic pattern is found at the BEA area level – for both metro and nonmetro counties. However, when the focus is on the BEA area level, additional COVID-19 cases rise significantly for "average" counties at the 25th percentile (21.46 additional cases per 100,000 residents for metro counties and 15.35 for nonmetro counties). The spike in cases rises as the density of incarcerated people increases. For an "average" metro county in a BEA area at the 95th percentile for mass incarceration (2.29 incarcerated people per square mile), we would anticipate an additional 205.83 cases per 100,000 residents – on top of the baseline caseload. For the "average" nonmetro county at the 95th percentile, the estimated additional caseload would be 142.66 cases per 100,000 residents.

Table 9 (next page) continues to focus on the interpretation of the Poisson regression estimation summarized in Table 6. Whereas the preceding discussion (Table 8) presented hypothetical "average" counties, the attention in Table 9 shifts to tracking the experiences of several counties. In preparing *Mass Incarceration, COVID-19, and Community Spread*, I used Marion Correctional Institution (MCI), Marion County, and the Columbus-Marion-Chillicothe (Ohio) economic area as examples (see *COVID-19 spread faster in counties with large prisons – and to nearby counties: Marion Correctional Institution (Ohio) is a disturbing example)*:

- Marion Correctional Institution – and the rapid spread of COVID-19 in April 2020 – was discussed to highlight the poor management of the pandemic in prisons and jails.
- The sharp rise in COVID-19 cases in Marion County (home to Marion Correctional Institution) was reviewed to document the risks that mass incarceration posed to the counties in which prisons are located.
- The Columbus-Marion-Chillicothe (Ohio) economic area was used to highlight the economic ties and commuting patterns that link mass incarceration in one county to the community spread of COVID-19 across the broader BEA economic area.

Once again, I return to counties in the Columbus-Marion-Chillicothe (Ohio) economic area to highlight trends. The experiences of four counties in this BEA economic area are highlighted in Table 9. Two are metropolitan counties: Franklin County (home to Columbus) and Delaware County (immediately north of Franklin County); two are non-metropolitan counties: Marion County (home to Marion Correctional Institution) and Ross County (home to several large prisons).

.

Table 9
Impact of Mass Incarceration on Confirmed COVID-19 Cases (May 1 – August 1)
Selected Counties, Columbus (Ohio) BEA Economic Area

		INCARCERATION IN COUNTY		INCARCERATION IN BEA AREA		MASS INCARCERATION CONTRIBUTION TO COVID-19 CASELOAD (CASES PER 100,000 RESIDENTS)				
County	**Population 2018**	**Incarcerated persons**	**incarcerated persons per sq. mi.**	**Incarcerated persons**	**Incarcerated persons per sq. mi.**	**Incarcerated in county (a)**	**Incarcerated in BEA area (b)**	**All other causes (c)**	**Total cases per 100,000 a + b + c**	**Additional cases: mass incarceration[c]**
Delaware (metro)	197,008	195[a]	0.42[a]	32,184	2.27	0[a] (0%)	95.54 (21%)	360.04 (79%)	455.58 (100%)	188.14 (21%)
Franklin (metro)	1,275,333	2,353[a]	4.32[a]	30,026	2.12	0[a] (0%)	141.89 (20%)	577.88 (80%)	719.77 (100%)	1,809.57 (20%)
Marion (nonmetro)	77,051	4,749	11.75	27,630	1.95	402.16 (37%)	144.13 (13%)	533.07 (49%)	1079.36[b] (100%)	421.26 (51%)
Ross (nonmetro)	65,344	5,647	8.15	26,732	1.89	149.09 (28%)	69.88 (12%)	319.92 (59%)	539.70[b] (100%)	143.04 (41%)

[a] Recall (see Table 6) that Poisson regression estimation did *not* find a statistically significant relationship between incarcerated persons in the county and COVID-19 caseloads for metro counties (Delaware County and Franklin County). For this reason, incarcerated persons in metro counties were *ignored* when calculating the contribution of mass incarceration to COVID-19 caseloads. Because I found a significant relationship in nonmetro counties, county-level incarceration (incarcerated persons per square mile) was included for these counties (Marion County and Ross County in Table 9).
[b] "Incarcerated in county (a)," "Incarcerated in BEA area (b)," "All other causes (c)" does *not* sum to 100% due to rounding.
[c] In the Poisson regression estimation, the dependent measure is COVID-19 cases per 100,000 residents. To calculate the number of additional cases due to mass incarceration, the population of the county must be considered: additional cases per 100,000 residents * (population / 100,000):

- Delaware: 95.54 * (197,008 / 100,000) = 188.14
- Franklin: 141.89 * (1,275,333 /100,000) = 1,809.57
- Marion: (402.16 * (77,051 /100,000)) + (144.13 * (77,051 /100,000)) = 421.26
- Ross: (149.09 * (63,344 /100,000)) + (69.88 * (63,344 /100,000)) = 143.04

As of August 1st, the United States had confirmed 4,456,389 cases of COVID-19 (World Health Organization 2020). Based on the Poisson regression model reported here, I estimate that 566,804 COVID-19 cases in the United States (12.7%) are linked to mass incarceration. However, mass incarceration's contribution to community spread is uneven across counties. Consider the 4 counties that are the focus of Table 9. Because the Columbus economic area is home to more than 30,000 incarcerated people, mass incarceration played a disproportionate role in each of the four counties. Roughly 20% of the COVID-19 caseload in both of the metro counties (Delaware and Franklin) is linked to the high levels of mass incarceration in the larger BEA area (compared to the 12.7% national average). Although Delaware County has relatively few incarcerated people and Columbus reported more than 2,000 incarcerated people, because these are metro counties, this county-level incarceration was ignored in these calculations. For the two nonmetro counties, large prisons in the county played a major role. For Marion County, it is estimated that 37% of the COVID-19 caseload is linked to Marion Correctional Institution – and when combined with the effect of mass incarceration in the multicounty BEA area – over half of all cases in the county (51%) are tied to mass incarceration. A comparable trend is in evidence for Ross County (28% of cases linked to county-level incarceration and 12% tied to BEA-level incarceration). For Ross County, county and BEA-level incarceration combine to account for 41% of COVID-19 cases.

Mass Incarceration, COVID-19, and Community Spread closes by reporting the total number of COVID-19 cases in states, BEA economic areas, and the nation. These estimates were based on the calculations reported here. That is, the additional cases of the four counties (Delaware, Franklin, Marion, and Ross) discussed in the preceding paragraphs were included in the 9,582.8 additional cases for the state of Ohio (as reported in Table 1, *Mass Incarceration, COVID-19, and Community Spread*). Appendix Table 1 (also below) provides a listing of all 50 states and the District of Columbia. Across the entire Columbus-Marion-Chillicothe (Ohio) economic area, I calculated that mass incarceration added 3,458.29 cases to the COVID-19 caseload. Among BEA economic areas, this ranks 27th in the nation. Because Table 2 (*Mass Incarceration, COVID-19, and Community Spread*) only includes the top 25 BEA areas, the Columbus BEA is not included. Appendix Table 2 (also below) provides a listing of all 179 BEA economic areas. Of course, these four counties were included in the national tabulation — 566,804 additional cases linked to mass incarceration for the entire United States (as reported in Table 3, *Mass Incarceration, COVID-19, and Community Spread*).

Conclusion

The conclusion of the report *Mass Incarceration, COVID-19, and Community Spread* comments on the social and political implications of these findings, and this discussion will not be repeated here. In concluding this methodological report, the focus will stay close to estimation decisions and their implications.

There is reason to believe that the estimations presented here should be seen as lower-bound estimates. That is, efforts were made to avoid overstating the impact of mass incarceration. First (as noted earlier), preliminary Poisson regression models provided evidence that influential cases may have distorted the estimation. After dropping suspect cases (i.e., dropping 26 cases with Cooks D above 3.0), the model was re-estimated. When compared to preliminary models, the coefficient for the several mass incarceration measures were *smaller* in the final model (reported in Table 6). While addressing the potential biases posed by influential cases is appropriate, the solution (dropping 26 cases) may have resulted in an underestimation of mass incarceration's impact. Second, for metro counties, the Poisson regression estimates did *not* find a statistically significant relationship between mass incarceration in a county (incarcerated persons per square mile) and COVID-19 cases. For this reason, when calculating the overall impact of mass incarceration on COVID-19 caseloads, it was assumed that prisons and jails in metropolitan counties had no impact. This finding is surprising – and in important respects counterintuitive. However, given the findings reported in Table 6, I believe that this is the appropriate decision. Future research can and should take steps to address influential cases and the surprising finding that the number of incarcerated persons in metro counties does not influence the spread of COVID-19. Doing so might provide evidence that this report has understated the impact of mass incarceration on the spread of COVID-19.

REFERENCES CITED

Alexander, Michelle. 2010. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness.* New York: The New Press.

American Civil Liberties Union. 2012. "Abuse of the Human Rights of Prisoners in the United States: Solitary Confinement." Available online at: https://www.aclu.org/other/abuse-human-rights-prisoners-united-states-solitary-confinement, accessed October 4, 2020.

Chin, Taylor, Rebecca Kahn, Ruoran Li, Jarvis T. Chen, Nancy Krieger, Caroline O. Buckee, Satchit Balsari, Mathew V. Kiang. 2020. "U.S. County-Level Characteristics to Inform Equitable COVID-19 Response." *medRxiv* 2020.04.08.20058248; doi: https://doi.org/10.1101/2020.04.08.20058248.

Environmental Working Group. 2020. "Meatpacking Plant Counties Report Twice the National Average Rate of COVID-19 Infections." Website: https://www.ewg.org/interactive-maps/2020-covid19-meat-plant-outbreaks/map/, accessed October 1, 2020.

Frome, Edward, and Harvey Checkoway. 1985. "Use of Poisson Regression Models in Estimating Incidence Rates and Ratios." *American Journal of Epidemiology* 121:309–323, https://doi.org/10.1093/oxfordjournals.aje.a114001.

Human Right Watch. 2020. "Criminal Justice." Available online at: https://www.hrw.org/united-states/criminal-justice, accessed October 4, 2020.

Isserman, A.M., E. Feser, and D.E. Warren. 2009. "Why Some Rural Places Prosper and Others Do Not." *International Regional Science Review* 32:300-342.

Johnson, Kenneth, and John Kort. 2004. "2004 Redefinition of the BEA Economic Areas," *Survey of Current Business* (November):68-75.

Moller, Stephanie, Arthur Alderson and Francois Nielsen. 2009. "Changing Patterns of Income Inequality in U.S. Counties, 1970–2000." *American Journal of Sociology* 114:1037–1101.

New York Times. 2020. " Coronavirus (Covid-19) Data in the United State." Website: https://github.com/nytimes/covid-19-data, accessed October 15, 2020.

Osborne, Jason. 2017. "Poisson Models: Low-Frequency Count Data as Dependent Variables', pp. 283-304 in *Regression & Linear Modeling: Best Practices and Modern Methods*. Thousand Oaks, CA: Sage [Accessed 5 November 2020], doi: 10.4135/9781071802724.

Pampel, Fred C. 2000. *Logistic regression*. Thousand Oaks, California: Sage [Accessed 4 November 2020], doi: 10.4135/9781412984805.

Population Health Methods (Columbia University). 2020. "Confounders in Time-Series Regression." Website: https://www.publichealth.columbia.edu/research/population-health-methods/confounders-time-series-regression, accessed October 4, 2020

Robert Wood Johnson Foundation 2020. *County Health Rankings & Roadmaps*. Website: https://www.countyhealthrankings.org/, accessed June 23, 2020.

Social Science Computing Cooperative (University of Wisconsin). 2014. "Exploring Regression Results using Margins." Website: https://www.ssc.wisc.edu/sscc/pubs/stata_margins.htm, accessed November 1, 2020.

Southern Center for Human Rights. 2020. "Category: Mass Incarceration." Website: https://www.schr.org/category/mass-incarceration/, accessed October 4, 2020.

Statwing. 2020. "A User-Friendly Guide to Logistic Regression." http://docs.statwing.com/a-user-friendly-guide-to-logistic-regression/, accessed June 26, 2020.

UCLA: Statistical Consulting Group. 2020. "FAQ: How Do I Interpret Odds Ratios in Logistic Regression?" Website: https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-how-do-i-interpret-odds-ratios-in-logistic-regression/, accessed November 1, 2020.

U.S. Census Bureau. 2020. *Your Gateway to Census 2000.* Website: https://www.census.gov/main/www/cen2000.html, accessed March 15, 2020.

U.S. Department of Agriculture. 2020. "2013 Rural-Urban Continuum Codes." Available at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx, accessed August 16, 2020.

U.S. Immigration and Customs Enforcement. 2020.

Walmsley, Ron. 2018. *World Prison Population List* (12th edition). *International Centre for Prison Studies*, accessed October 4, 2020.

Widra, Emily, and Dylan Hayre. 2020. *Failing Grades: States' Responses to COVID-19 in Jails & Prisons.* Prison Policy Initiative, available online at: https://www.prisonpolicy.org/reports/failing_grades.html, accessed October 4, 2020.

Williams, Richard. 2012. "Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects." *The Stata Journal* 12:308–331.

World Health Organization. 2020. *Coronavirus Disease (COVID-19). Situation Report – 194.* Data as received by WHO from national authorities by 10:00 CEST, 1 August 2020. Additional information available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

Yobero, Czar. 2017. "Linear Regression Using Dummy Variables." *RPubs,* available online at: https://rpubs.com/cyobero/dummy-variables, accessed November 5, 2020.

Appendix Table 1
Impact of Mass Incarceration on States (50 states and District of Columbia)
Ranked by Net Additional Cases of COVID-19 Confirmed (May 1- August 1)

Rank	State	Net additional cases	Total population	Net additional cases per 100,000 residents
1	California	113,968.7	39,148,760	291.1
2	Florida	92,981.0	20,598,140	451.4
3	Texas	55,016.6	27,885,196	197.3
4	Illinois	47,298.3	12,821,497	368.9
5	New York	38,915.3	19,618,452	198.4
6	New Jersey	25,139.2	8,881,845	283.0
7	Georgia	24,951.1	10,297,484	242.3
8	Pennsylvania	20,166.3	12,791,181	157.7
9	North Carolina	13,277.2	10,155,624	130.7
10	Maryland	12,513.2	6,003,435	208.4
11	Virginia	12,040.3	8,413,774	143.1
12	Louisiana	10,684.6	4,663,616	229.1
13	Ohio	9,582.8	11,641,879	82.3
14	Arizona	8,754.8	6,946,685	126.0
15	Tennessee	7,866.5	6,651,089	118.3
16	South Carolina	7,716.4	4,955,925	155.7
17	Indiana	6,879.4	6,637,426	103.6
18	Alabama	6,721.5	4,864,680	138.6
19	Massachusetts	6,681.8	6,830,193	97.8
20	Connecticut	6,583.0	3,581,504	183.8
21	Michigan	4,787.3	9,957,488	48.0
22	Mississippi	4,545.3	2,988,762	152.8
23	Washington, DC	3,607.1	684,498	526.9
24	Wisconsin	3,454.1	5,778,394	59.8
25	Missouri	3,381.2	6,090,062	55.5
26	Delaware	3,180.2	949,495	334.9
27	Washington	2,633.7	7,294,336	36.1
28	Kentucky	2,455.2	4,440,204	55.3
29	Arkansas	1,590.6	2,990,671	53.2
30	Oklahoma	1,589.2	3,918,137	40.6
31	Minnesota	1,310.3	5,527,358	23.7
32	Colorado	1,078.2	5,531,141	19.5
33	Rhode Island	932.0	1,056,611	88.2
34	Oregon	930.9	4,081,943	22.8
35	Iowa	875.4	3,132,499	27.9
36	Kansas	645.8	2,908,776	22.2

Appendix 1 – Impact of Mass Incarceration on States

Rank	State	Net additional cases	Total population	Net additional cases per 100,000 residents
37	West Virginia	469.9	1,829,054	25.7
38	New Hampshire	332.7	1,343,622	24.8
39	New Mexico	311.0	2,092,434	14.9
40	Utah	292.1	3,045,350	9.6
41	Nebraska	192.5	1,904,760	10.1
42	Idaho	165.9	1,687,809	9.8
43	Nevada	165.7	2,922,849	5.7
44	South Dakota	37.3	849,954	4.4
45	Maine	22.6	1,332,813	1.7
46	Wyoming	21.9	581,836	3.8
47	Vermont	20.0	624,977	3.2
48	Hawaii	18.9	1,422,029	1.3
49	North Dakota	9.9	752,201	1.3
50	Montana	8.1	1,041,732	0.8
51	Alaska	1.5	730,318	0.2

Appendix 2
Impact of Mass Incarceration on BEA Economic Areas (179 BEA Economic Areas)
Ranked by Net Additional Cases of COVID-19 Confirmed (May 1- August 1)

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
1	Los Angeles-Long Beach-Riverside, CA	94,221.8	20,678,296	455.7
2	New York-Newark-Bridgeport, NY-NJ-CT-PA	64,000.1	23,602,788	271.2
3	Chicago-Naperville-Michigan City, IL-IN-WI	47,690.6	10,457,692	456.0
4	Miami-Fort Lauderdale-Miami Beach, FL	40,823.2	6,855,487	595.5
5	Houston-Baytown-Huntsville, TX	26,005.6	7,809,735	333.0
6	Orlando-The Villages, FL	25,108.0	5,072,299	495.0
7	Philadelphia-Camden-Vineland, PA-NJ-DE-MD	23,700.9	7,145,289	331.7
8	Washington-Baltimore-Northern Virginia, DC-MD-VA- WV	21,580.5	10,040,033	214.9
9	Atlanta-Sandy Springs-Gainesville, GA-AL	19,339.6	8,014,119	241.3
10	Dallas-Fort Worth, TX	14,526.0	8,892,231	163.4
11	San Jose-San Francisco-Oakland, CA	14,279.6	10,432,077	136.9
12	Tampa-St. Petersburg-Clearwater, FL	12,452.6	3,030,047	411.0
13	Raleigh-Durham-Cary, NC	6,902.8	3,512,772	196.5
14	Phoenix-Mesa-Scottsdale, AZ	6,895.1	5,260,048	131.1
15	Boston-Worcester-Manchester, MA-NH	6,372.4	8,594,883	74.1
16	McAllen-Edinburg-Pharr, TX	5,715.4	1,356,787	421.2
17	Fresno-Madera, CA	5,298.3	1,761,235	300.8
18	Charlotte-Gastonia-Salisbury, NC-SC	4,845.3	3,179,708	152.4
19	Baton Rouge-Pierre Part, LA	4,549.0	861,346	528.1
20	Cleveland-Akron-Elyria, OH	4,448.1	4,533,215	98.1
21	Memphis, TN-MS-AR	4,356.6	2,047,494	212.8
22	Indianapolis-Anderson-Columbus, IN	4,323.4	3,556,695	121.6
23	Sarasota-Bradenton-Venice, FL	4,080.4	2,081,951	196.0
24	Detroit-Warren-Flint, MI	3,716.8	6,837,098	54.4
25	Jacksonville, FL	3,665.0	1,884,231	194.5
26	Richmond, VA	3,497.8	1,745,675	200.4
27	Columbus-Marion-Chillicothe, OH	3,458.3	2,850,691	121.3
28	Jackson-Yazoo City, MS	3,326.8	1,661,397	200.2
29	Hartford-West Hartford-Willimantic, CT	3,134.2	2,295,996	136.5

Appendix 2 – Impact of Mass Incarceration on BEA Economic Areas

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
30	San Antonio, TX	2,980.9	2,736,961	108.9
31	St. Louis-St. Charles-Farmington, MO-IL	2,808.1	3,388,001	82.9
32	Gainesville, FL	2,668.2	501,152	532.4
33	Milwaukee-Racine-Waukesha, WI	2,531.2	2,362,430	107.1
34	Birmingham-Hoover-Cullman, AL	2,507.5	1,772,590	141.5
35	Macon-Warner Robins-Fort Valley, GA	2,327.2	672,993	345.8
36	Nashville-Davidson-Murfreesboro-Columbia, TN	2,264.3	3,095,166	73.2
37	Columbia-Newberry, SC	2,203.5	1,116,404	197.4
38	Seattle-Tacoma-Olympia, WA	2,135.0	5,168,694	41.3
39	Virginia Beach-Norfolk-Newport News, VA-NC	2,073.0	1,921,391	107.9
40	New Orleans-Metairie-Bogalusa, LA	2,052.2	1,706,546	120.3
41	Lafayette-Acadiana, LA	2,003.6	867,513	231.0
42	Albany, GA	1,958.9	607,225	322.6
43	Greenville-Spartanburg-Anderson, SC	1,942.1	1,472,084	131.9
44	Savannah-Hinesville-Fort Stewart, GA	1,823.2	902,118	202.1
45	Panama City-Lynn Haven, FL	1,776.0	305,449	581.5
46	Pensacola-Ferry Pass-Brent, FL	1,759.2	748,559	235.0
47	Harrisburg-Carlisle-Lebanon, PA	1,723.0	2,231,844	77.2
48	Austin-Round Rock, TX	1,683.2	2,181,797	77.1
49	Sacramento-Arden-Arcade-Truckee, CA-NV	1,485.2	2,925,434	50.8
50	Greensboro-Winston-Salem-High Point, NC	1,434.5	2,019,723	71.0
51	Louisville-Elizabethtown-Scottsburg, KY-IN	1,357.9	1,656,018	82.0
52	Minneapolis-St. Paul-St. Cloud, MN-WI	1,351.1	5,533,996	24.4
53	Tallahassee, FL	1,349.5	543,410	248.3
54	Huntsville-Decatur, AL	1,343.1	1,136,616	118.2
55	Killeen-Temple Fort Hood, TX	1,324.9	747,217	177.3
56	Cincinnati-Middletown-Wilmington, OH-KY-IN	1,289.8	2,392,211	53.9
57	Corpus Christi-Kingsville, TX	1,281.7	888,458	144.3
58	Kansas City-Overland Park-Kansas City, MO-KS	1,211.5	2,741,889	44.2
59	Augusta-Richmond County, GA-SC	1,066.7	649,218	164.3

Appendix 2 – Impact of Mass Incarceration on BEA Economic Areas

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
60	Myrtle Beach-Conway-Georgetown, SC	1,026.4	1,148,965	89.3
61	Portland-Vancouver-Beaverton, OR-WA	1,009.5	3,341,379	30.2
62	Pittsburgh-New Castle, PA	1,007.1	2,850,837	35.3
63	Montgomery-Alexander City, AL	1,006.4	552,240	182.2
64	Shreveport-Bossier City-Minden, LA	924.1	557,323	165.8
65	Grand Rapids-Muskegon-Holland, MI	902.9	1,992,050	45.3
66	Knoxville-Sevierville-La Follette, TN	893.8	1,250,325	71.5
67	Oklahoma City-Shawnee, OK	878.4	2,192,918	40.1
68	Little Rock-North Little Rock-Pine Bluff, AR	872.4	1,572,035	55.5
69	Columbus-Auburn-Opelika, GA-AL	838.2	494,720	169.4
70	Johnson City-Kingsport-Bristol (Tri-Cities), TN-VA	769.2	860,786	89.4
71	Denver-Aurora-Boulder, CO	753.3	4,558,349	16.5
72	Charleston-North Charleston, SC	743.8	796,815	93.3
73	Lexington-Fayette-Frankfort-Richmond, KY	740.7	1,555,777	47.6
74	Tulsa-Bartlesville, OK	657.0	1,402,716	46.8
75	Monroe-Bastrop, LA	640.4	337,021	190.0
76	Mobile-Daphne-Fairhope, AL	591.7	765,497	77.3
77	Tucson, AZ	544.5	1,192,585	45.7
78	Dayton-Springfield-Greenville, OH	497.0	1,370,927	36.3
79	Greenville, NC	489.4	726,411	67.4
80	Beaumont-Port Arthur, TX	489.1	466,693	104.8
81	Rochester-Batavia-Seneca Falls, NY	487.7	1,517,322	32.1
82	Lake Charles-Jennings, LA	478.9	351,954	136.1
83	Syracuse-Auburn, NY	475.2	1,994,448	23.8
84	Dover, DE	459.1	618,534	74.2
85	Dothan-Enterprise-Ozark, AL	456.8	313,919	145.5
86	Fort Wayne-Huntington-Auburn, IN	434.1	803,095	54.1
87	Asheville-Brevard, NC	389.8	723,270	53.9
88	Peoria-Canton, IL	329.7	868,156	38.0
89	Madison-Baraboo, WI	319.4	1,223,576	26.1
90	State College, PA	318.4	794,772	40.1
91	Evansville, IN-KY	316.0	770,228	41.0
92	Kennewick-Richland-Pasco, WA	314.9	598,390	52.6
93	Colorado Springs, CO	299.2	779,130	38.4
94	Lubbock-Levelland, TX	297.9	466,262	63.9

Appendix 2 – Impact of Mass Incarceration on BEA Economic Areas

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
95	Columbia, MO	293.6	519,968	56.5
96	Springfield, IL	292.3	618,179	47.3
97	Scranton-Wilkes-Barre, PA	291.0	649,802	44.8
98	Roanoke, VA	282.9	826,022	34.2
99	Appleton-Oshkosh-Neenah, WI	277.6	870,776	31.9
100	Midland-Odessa, TX	276.8	659,159	42.0
101	El Paso, TX	276.0	1,208,018	22.8
102	Albany-Schenectady-Amsterdam, NY	272.0	1,386,317	19.6
103	Gulfport-Biloxi-Pascagoula, MS	269.9	433,378	62.3
104	Des Moines-Newton-Pella, IA	267.4	1,316,289	20.3
105	Salt Lake City-Ogden-Clearfield, UT	262.6	2,836,584	9.3
106	South Bend-Mishawaka, IN-MI	239.0	960,164	24.9
107	Charleston, WV	237.0	1,163,888	20.4
108	Cedar Rapids, IA	222.0	536,484	41.4
109	Buffalo-Niagara-Cattaraugus, NY	210.6	1,444,680	14.6
110	Omaha-Council Bluffs-Fremont, NE-IA	170.1	1,115,584	15.2
111	Las Vegas-Paradise-Pahrump, NV	163.3	2,614,169	6.2
112	Davenport-Moline-Rock Island, IA-IL	153.3	486,334	31.5
113	Toledo-Fremont, OH	149.8	979,303	15.3
114	Albuquerque, NM	147.4	953,990	15.5
115	Tupelo, MS	144.0	549,141	26.2
116	Abilene, TX	139.9	229,642	60.9
117	Wichita-Winfield, KS	135.7	1,085,754	12.5
118	Texarkana, TX-Texarkana, AR	130.6	314,091	41.6
119	Boise City-Nampa, ID	124.9	798,302	15.7
120	Champaign-Urbana, IL	113.1	547,732	20.6
121	Amarillo, TX	109.7	503,671	21.8
122	Springfield, MO	105.3	1,040,357	10.1
123	Erie, PA	95.4	500,725	19.0
124	Harrisonburg, VA	84.8	327,344	25.9
125	Clarksburg, WV + Morgantown, WV	80.7	346,893	23.3
126	Fayetteville-Springdale-Rogers, AR-MO	80.4	590,637	13.6
127	Reno-Sparks, NV	79.3	736,401	10.8
128	Jonesboro, AR	69.9	314,001	22.3
129	Sioux City-Vermillion, IA-NE-SD	66.6	377,219	17.6
130	Paducah, KY-IL	62.4	243,389	25.6
131	Fort Smith, AR-OK	60.1	341,675	17.6
132	Lincoln, NE	58.9	433,545	13.6

Appendix 2 – Impact of Mass Incarceration on BEA Economic Areas

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
133	Cape Girardeau-Jackson, MO-IL	53.4	297,625	17.9
134	Traverse City, MI	44.7	284,840	15.7
135	Pueblo, CO	42.0	244,628	17.2
136	Spokane, WA	34.6	890,628	3.9
137	La Crosse, WI-MN	34.1	262,475	13.0
138	Wichita Falls, TX	32.0	187,998	17.0
139	Marinette, WI-MI	29.4	327,294	9.0
140	Eugene-Springfield, OR	28.7	840,261	3.4
141	Pendleton-Hermiston, OR	26.9	147,565	18.2
142	Sioux Falls, SD	26.6	522,564	5.1
143	Wausau-Merrill, WI	24.3	518,243	4.7
144	Topeka, KS	23.7	479,039	4.9
145	Portland-Lewiston-South Portland, ME	21.8	1,009,674	2.2
146	San Angelo, TX	19.6	150,631	13.0
147	Honolulu, HI	18.9	1,422,029	1.3
148	Joplin, MO	15.7	370,331	4.2
149	Duluth, MN-WI	14.2	352,749	4.0
150	Santa Fe-Espanola, NM	12.1	276,447	4.4
151	Idaho Falls-Blackfoot, ID	12.1	358,863	3.4
152	Twin Falls, ID	10.7	194,697	5.5
153	Wenatchee, WA	10.5	273,078	3.8
154	Waterloo-Cedar Falls, IA	9.5	225,921	4.2
155	Salina, KS	8.5	189,056	4.5
156	Kearney, NE	7.0	328,178	2.1
157	Bend-Prineville, OR	6.8	241,191	2.8
158	Forgo-Wahpeton, ND-MN	6.4	327,423	2.0
159	Redding, CA	5.6	361,246	1.6
160	Helena, MT	5.0	281,436	1.8
161	Mason City, IA	4.7	155,260	3.0
162	Burlington-South Burlington, VT	4.7	396,062	1.2
163	Casper, WY	4.2	370,320	1.1
164	Farmington, NM	3.5	224,636	1.6
165	Lewiston, ID-WA	3.3	93,538	3.5
166	Alpena, MI	3.0	228,950	1.3
167	Scotts Bluff, NE	2.9	90,092	3.2
168	Grand Forks, ND-MN	2.7	208,430	1.3
169	Bismarck, ND	2.2	203,971	1.1
170	Rapid City, SD	1.7	238,686	0.7
171	Minot, ND	1.6	166,435	1.0

Appendix 2 – Impact of Mass Incarceration on BEA Economic Areas

Rank	BEA Economic Area	Net additional cases	Total population	Net additional cases per 100,000 residents
172	Anchorage, AK	1.5	730,318	0.2
173	Billings, MT	1.3	369,157	0.4
174	Missoula, MT	1.2	320,831	0.4
175	Great Falls, MT	1.1	147,904	0.7
176	Bangor, ME	0.8	323,139	0.3
177	Aberdeen, SD	0.3	82,805	0.4
178	Flagstaff, AZ	0.3	147,567	0.2
179	San Diego-Carlsbad-San Marcos, CA*	0.0	3,302,833	0.0

*****San Diego is an anomaly. San Diego County is the sole county in the San Diego-Carlsbad-San Marcos, CA BEA economic area. As this is a metro county, incarceration in the county is *ignored* when estimating additions to the COVID-19 caseload. The measure of mass incarceration at the BEA level excludes incarceration in the county (and thus results in San Diego County having zero (0) people incarcerated in the BEA area. This quirk in measurement understates the impact of mass incarceration in this county. The “average” metro county saw an additional 109 cases per 100,000 residents. If San Diego – with a population of 3.3 million – was average in this respect, roughly 3,600 cases of COVID-19 would have been linked to mass incarceration.