



## Incarceration and mortality in the United States

Elias Nosrati<sup>a,\*</sup>, Jacob Kang-Brown<sup>b</sup>, Michael Ash<sup>c</sup>, Martin McKee<sup>d</sup>, Michael Marmot<sup>e</sup>, Lawrence P. King<sup>c</sup>

<sup>a</sup> University of Oxford, Merton College, Merton Street, OX1 4JD, Oxford, UK

<sup>b</sup> Vera Institute of Justice, 233 Broadway, 12th Floor, New York, NY, 10279, USA

<sup>c</sup> University of Massachusetts Amherst, Department of Economics, Crotty Hall, 412 North Pleasant Street, University of Massachusetts, Amherst, MA, 01002, USA

<sup>d</sup> London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT, London, UK

<sup>e</sup> Institute of Health Equity, Department for Epidemiology & Public Health, University College London, 1-19, Torrington Place, WC1E 7HB, London, UK

### ARTICLE INFO

#### Keywords:

Incarceration  
Mortality  
Instrumental variables  
Matching

### ABSTRACT

The ongoing COVID-19 pandemic has spotlighted the role of America's overcrowded prisons as vectors of ill health, but robust analyses of the degree to which high rates of incarceration impact population-level health outcomes remain scarce. In this paper, we use county-level panel data from 2927 counties across 43 states between 1983 and 2014 and a novel instrumental variable technique to study the causal effect of penal expansion on age-standardised cause-specific and all-cause mortality rates. We find that higher rates of incarceration have substantively large effects on deaths from communicable, maternal, neonatal, and nutritional diseases in the short and medium term, whilst deaths from non-communicable disease and from all causes combined are impacted in the short, medium, and long run. These findings are further corroborated by a between-unit analysis using coarsened exact matching and a simulation-based regression approach to predicting geographically anchored mortality differences.

### Introduction

The advent of COVID-19 has spotlighted the role of America's overcrowded prisons as vectors of ill health (Akiyama, Spaulding, & Rich, 2020; Saloner, Parish, Ward, DiLaura, & Dolovich, 2020). As of April 28, 2021, 395,915 prisoners have contracted the disease and 2572 have died as a result. Amongst prison staff, corresponding figures are 110,136 and 201, respectively (Marshall, 2021). However, prior to the ongoing pandemic, causal evidence of the link between high rates of incarceration and infectious disease mortality at the population level has been scarce. More generally, despite the historically unprecedented expansion of the American penal state since the 1970s, imprisonment has rarely been construed as a distal determinant of population health in its own right (Wildeman & Wang, 2017). In this paper, we use spatially disaggregated time-series data and a novel instrumental variable approach to examine how local prison admission rates impact age-standardised death rates at the US county level. By drawing on extant scholarship on the health impacts of penal expansion, we hypothesise a causal association between high imprisonment rates and county-level mortality that is operant above and beyond the role of

factors like income, education, or violent crime. Moreover, we hypothesise that high incarceration rates impact not only those who pass through the criminal justice system but also local populations at large. We provide a comprehensive panel data analysis — to our knowledge, the first of its kind — assessing how incarceration affects geographically anchored patterns of mortality from communicable and non-communicable diseases and all causes combined.

### Background and hypotheses

Since the early 1970s, the American penal state has undergone a historically unprecedented expansion. After 50 years of relative stability, in 1973 the national jail and prison incarceration rate stood at 161 residents per 100,000 population. In 2007, the corresponding figure was 767. In absolute numbers rather than rates, this amounts to a shift from just under 400,000 to over 2.3 million individuals behind bars — a sevenfold increase in less than four decades (National Research Council, 2014). Beyond these aggregate numbers, imprisonment has emerged as a new stage in the life course of men of colour who find themselves at the bottom of the class structure (Pettit & Western, 2004; Western, 2006;

\* Corresponding author. University of Oxford, Merton College, Merton Street, OX1 4JD, Oxford, UK.

E-mail address: [elias.nosrati@merton.ox.ac.uk](mailto:elias.nosrati@merton.ox.ac.uk) (E. Nosrati).

<https://doi.org/10.1016/j.ssmph.2021.100827>

Received 12 March 2021; Received in revised form 1 May 2021; Accepted 21 May 2021

Available online 1 June 2021

2352-8273/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Wacquant, 2009). This is evidenced by how the cumulative risk of experiencing parental incarceration by age 14 amongst African American children born to high-school dropouts exceeds 50% (Wildeman, 2009).

A rich body of evidence has related penal expansion to declining health and deepening health inequality (for recent reviews, see Massoglia & Pridemore, 2015; Wildeman & Wang, 2017). In particular, overcrowded correctional facilities have been linked to infectious disease transmission (Massoglia, 2008; Ndeffo-Mbah, Vigliotti, Skrip, Dolan, & Galvani, 2018) — a linkage that has been further spotlighted by the ongoing COVID-19 pandemic (Akiyama et al., 2020; Hooks & Sawyer, 2020; Reinhart & Chen, 2020; Saloner et al., 2020). Upon release from prison, former inmates experience mortality rates close to thirteenfold that of the comparable populace and are especially vulnerable during the first two weeks post-release, notably via acute stress-related psychosocial mechanisms (Binswanger et al., 2007; Zlodre & Fazel, 2012).

Moreover, previous scholarship has documented the ways in which high rates of incarceration act in cascading ways upon other social determinants, or ‘fundamental causes’ (Link & Phelan, 1995), of health. Chief amongst such upstream determinants are the social and economic decay of neighbourhoods (Sampson & Loeffler, 2010), the disruption of social and family ties (Wildeman & Muller, 2012; Wildeman, Schnittker, & Turney, 2012), adverse childhood experiences related to parental incarceration (Wildeman, 2009; Turney, 2014), and enduring material deprivation and hardship in areas of concentrated disadvantage (Western, 2018). For instance, the incarceration of a family member has been shown to impair the well-being of non-incarcerated partners and children, notably due to declining household income, reduced parental investment, and unstable social relationships (Wildeman & Muller, 2012; Turney, 2014). At the community level, the criminal justice system plays a pivotal role in shaping the trajectories of neighbourhoods by removing prime-age men from their local communities, fragmenting family relationships, and eroding social ties (Western, 2006; Sampson & Loeffler, 2010; National Research Council, 2014; Western, 2018).

Against this backdrop, we hypothesise that high rates of incarceration have a causal impact on a range of mortality outcomes, not only at the level of the individual but on a population level. Drawing on the extant literature, our argument is that the experience of incarceration may prove deeply consequential not only for those who are incarcerated but also for their families, friends, and broader social connections (Gilmore, 2007; Wildeman & Muller, 2012; Wildeman & Wang, 2017). Thus the causal pathways from imprisonment to mortality most likely involve diverse modalities of ‘social sundering’ (Therborn, 2013, pp. 22–28) whereby the material and symbolic fabric of social life is eroded, for individuals and collectives alike (see also Wilkinson & Pickett, 2010; 2018). Given this plurality of pathways and mechanisms, we therefore expect the hypothesised causal effect to manifest in the form of mortality from both communicable and non-communicable diseases, within and between units of analysis over time.

Previous studies have shed light on the effects of incarceration on health at the level of individuals and communities. However, robust evidence at the population level remains scarce, notably when it comes to the geographical patterning of different types of mortality rates, although we note two recent associational studies by Kajeepeta, Ruth-erford, Keyes, El-Sayed, and Prins (2020, 2021). Our macroscopic approach allows us not only to generate a broad overview of how incarceration shapes population health, but also eschews the methodological challenges intrinsic to the use of individual-level survey data (however, see Daza, Palloni, & Jones, 2020), such as producing plausible causal identification strategies and constructing appropriate comparison groups (see Johnson & Easterling, 2012; Wildeman, Wakefield, & Turney, 2013). In what follows, we test our hypotheses by using county-level panel data and a novel instrumentation technique suited to isolating exogenous treatment variation.

## Data and methods

### Outcome and control variables

The first of our three outcome variables is the county-level age-standardised mortality rate from communicable, maternal, neonatal, and nutritional diseases per 100,000 county population between 1983 and 2014. Taken together, these highly correlated causes of death form a ‘level 1’ category within the framework of the Global Burden of Diseases, Injuries, and Risk Factors Study, the methodology of which is employed to generate comparable age-standardised mortality metrics across county-years. To avoid terminological clutter, however, we henceforth use ‘communicable’ as a shorthand for ‘communicable, maternal, neonatal, and nutritional’. The second outcome variable is mortality from all non-communicable diseases and the third is all-cause mortality. All three of these variables are publicly available from the Institute for Health Metrics and Evaluation (IHME, 2017). Our treatment variable is the county-level annual prison admissions rate, generated by the Vera Institute of Justice using state corrections sources and the National Corrections Reporting Program by the Bureau of Justice Statistics which are converted into annual county-level rates per 100,000 residents aged 15–64 (Hinds, Lu, Wallace-Lee, & Kang-Brown, 2020). Six states — Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont — are excluded from the analysis due to lack of consistently collected prison admissions data. Due to certain discrepancies between our data sources in measuring county boundaries and accounting for changes to counties over time, the state of Virginia and a handful of counties from other states (77 counties in total out of an initial sample of 3004) are also excluded from the final analysis. We employ a set of baseline control variables that are associated with both the treatment and the outcome, namely annual rates of violent crime, median household income, high school graduation rates, and the fraction of each county population that is African American, Hispanic, or any other non-White ethnic minority. These variables are all available from the US Census Bureau, except for the measure of violent crime which is extracted from the Federal Bureau of Investigation’s Uniform Crime Reporting Program. Descriptive statistics are reported in Table 1.

**Table 1**  
Descriptive statistics.

Statistic	N	Mean	St. Dev.	Min	Max
Mortality from communicable disease	65,237	50	13	15	263
Mortality from non-communicable disease	65,237	841	115	247	1499
All-cause mortality	65,237	972	140	323	1832
Incarceration rate per 100,000 population	65,237	268	205	0.0	2583
Violent crime rate per 100,000 population	65,237	278	267	0.0	5972
Median household income (\$)	65,237	47,105	11,709	17,583	125,705
High school graduation rate	65,237	0.8	0.1	0.3	1.0
Fraction African Americans	65,237	0.1	0.1	0.0	0.9
Fraction Hispanics	65,237	0.1	0.1	0.0	1.0
Fraction other ethnic minority	65,237	0.02	0.1	0.0	0.9

Notes: All variables, listed in the first column, are measured at the county level. The second column lists the number of observed county-years. The three outcome variables — communicable, non-communicable, and all-cause mortality rates per 100,000 population — are taken from the Institute for Health Metrics and Evaluation US Health Map database (IHME, 2017). The incarceration rate is per 100,000 population aged 16–64 and is constructed by the Vera Institute of Justice (Hinds et al., 2020). The measure of violent crime is extracted from the Federal Bureau of Investigation’s Uniform Crime Reporting Program. All remaining variables are taken from the US Census Bureau.

### Instrumental variable models

To empirically assess how incarceration affects county-level mortality rates, we posit the following data-generating process:

$$Y_{it} = T_{i,t-k}\beta + X_{it}\theta + \mu_i + \varphi_t + \varepsilon_{it}, \quad (1)$$

where  $Y_{it}$  denotes one of the three alternative outcome variables as measured in county  $i$  at time  $t$ ; the treatment variable  $T_{i[t-k]}$  is the county-level incarceration rate per 100,000 population, lagged by  $k$  years to allow for delayed effects;  $X$  is a vector of control variables;  $\mu$  and  $\varphi$  capture unit- and time-fixed effects, respectively; and  $\varepsilon$  is a stochastic error term. Our principal quantity of interest is  $\beta$ , which is a causal effect parameter to be estimated. By standardising our predictor variables, this parameter is interpreted as the excess number of deaths per 100,000 county population caused by a standard deviation increase in incarceration rates, net of all controls. However, as visualised in SI Figure A1, we face a potential identification problem wherein the estimated relation between the treatment variable  $T$  and the outcome variable  $Y$  is biased by some unmeasured confounder  $U$ , even after controlling for observed covariates  $X$ . In our case,  $U$  might denote unobserved variables that simultaneously affect incarceration and mortality, such as locally contingent healthcare or welfare-related policy shocks.

To address this concern, we construct an instrumental variable,  $Z$ , that is suited to isolating exogenous variation in  $T$ . To do this, we adopt a compound instrument derived from the interaction between the county-specific average exposure to incarceration over the sample period and annual nationwide per capita correctional spending. This instrument meets the relevance criterion insofar as increasing correctional expenditure is predictive of higher rates of incarceration. It also meets the exclusion criterion insofar as annual aggregate correctional spending is independent of any given county, to the effect that unit-specific shocks that deviate from a county's long-run average exposure to imprisonment result from a treatment assignment mechanism that is orthogonal to that county's potential outcomes. In other words, the outcome of interest in units with varying propensities to incarcerate will not be affected by changes in aggregate correctional spending other than through the impact of incarceration.

We believe that this proposed instrumentation method constitutes an advance in the study of the incarceration-health nexus. A recent study by Weidner and Schultz (2019) uses a cross-sectional design in which correctional spending alone is used as an instrumental variable. We argue that the methodological framework of the present paper provides a more stringent framework for causal inference by virtue of the timeseries dimension of our data. Not only are year- and unit-specific attributes netted out by de-meaning through entities, but lagged effects are also incorporated into our model design. The two-way fixed-effects model thus constitutes a rigorous approach that eliminates any confounders that either remain stable over time — such as county- or state-level institutional factors — or form part of any aggregate time trends, whilst also allowing for dynamic relationships. This combination of factors leads us to believe that we are better positioned to isolate exogenous shocks that operate above and beyond individual units' default exposure to incarceration.

We thus obtain an instrument  $Z_{it} = \bar{T}_i \times C_t$ , where  $\bar{T}_i$  is county  $i$ 's average incarceration rate over the sample period and  $C_t$  is the aggregate per capita expenditure on the construction and maintenance of correctional facilities across all states in year  $t$ . The latter variable is obtained from the Bureau of Justice Statistics' Justice Expenditure and Employment Series and is measured every few years. A spline function is then applied to impute missing values through interpolation between observed years, the result of which is visualised in SI Figure A2. Our two-stage regression model now has the following selection equation:

$$T_{it} = Z_{it}\tau + X_{it}\eta + \alpha_i + \delta_t + v_{it} \quad (2)$$

We then re-specify the model in equation (1) as follows, with  $\hat{T}$  being

a vector of fitted values from equation (2):

$$Y_{it} = \hat{T}_{i,t-k}\beta + X_{it}\theta + \mu_i + \varphi_t + \varepsilon_{it}. \quad (3)$$

We set  $k \in \{1,5,10\}$  to assess the short-run, medium-run, and long-run effects of incarceration. To empirically assess the strength of the chosen instrument, we compare the model in equation (2) to a restricted first-stage regression in which the effect  $\tau$  of  $Z$  on  $T$  is set to be null, obtaining a  $\chi^2$  test statistic of 8057 on 1 degree of freedom ( $p < 0.001$ ). Hence  $Z$  comfortably satisfies the benchmark for identifying a strong instrument. Our model accounts for (a) any time-invariant confounders, even if these are unobserved, by isolating variation within counties over time and (b) any aggregate trends that affect all counties simultaneously.

### Matching and between-county models

We complement our analysis of within-county variation over time with a model of long-run mortality differences between counties by averaging across units over our sample period. Despite the fact that fixed-effects models are typically preferred when seeking to infer causation and that between-unit variation rarely yields plausible estimates of a causal relationships, we are nevertheless interested in the between-county variation because a sole focus on within-county variation over time prevents us from examining a key quantity of interest, namely the magnitude of disparity in mortality burdens between counties. However, in order to render the corresponding parameter estimates more plausible, we employ matching as a non-parametric form of pre-processing the data (Ho, Imai, King, & Stuart, 2007; Iacus, King, & Porro, 2019).

The goal of matching is to reduce inefficiency, bias, and model dependence. It is a non-model-based approach to preparing the data for parametric analysis with a view to mimicking experimental research designs. In non-technical terms, matching seeks to select units of analysis (counties) that are similar if not identical to one another in all respects except for one: whether or not they are exposed to a key variable of interest. In the present case, the quantity of interest is the effect of high rates of incarceration on mortality rates, above and beyond the endogenous associations between incarceration and factors like income, education, or crime. Applying a matching algorithm will help 'match' counties that share key characteristics, except that some have high incarceration rates and others have low incarceration rates. This will facilitate a more precise account of the link between penal expansion and the local mortality burden. In more technical terms, let  $Y_i$  designate the outcome variable of interest (mortality), let  $T_i \in \{0,1\}$  designate a dichotomous treatment variable (low versus high incarceration rates), and let  $X_i$  designate a series of pre-treatment covariates (income, education, crime, demographic composition, etc.). Then the treatment effect  $\beta$  on a treated unit  $i$  is  $\beta = Y_i(T_i = 1) - Y_i(T_i = 0)$ . However, the last term of this equation,  $Y_i(T_i = 0)$ , is an unobserved counterfactual. One can estimate this quantity with  $Y_j$  from control units (indexed by  $j$ ) that are matched on relevant covariates (i.e.,  $X_i \approx X_j$ ) such that the estimated counterfactual quantity,  $\hat{Y}_i(T_i = 0)$ , is equal to  $Y_j(T_j = 0)$ . Unmatched units are pruned from the data set to improve empirical covariate balance between treatment and control groups in the sample, and the parametric model is applied to the pruned rather than to the raw data. As a result, the functional form of the parametric specification is subject to less arbitrary model dependence.

In the analysis below, we employ what is known as coarsened exact matching. This form of matching proceeds as follows. For lack of being able to match on exact values of continuous covariates, this algorithm temporarily 'coarsens' the covariates  $X$  into subcategories (defined via a non-parametric histogram estimator). It then applies exact matching on the coarsened  $X$ ,  $c(X)$ , before sorting observations into strata, each with unique values of  $c(X)$ . Any stratum with zero treated or control units is pruned from the data set. The algorithm then passes the original (uncoarsened) units — except for the pruned ones — on to the matched

data set that is used in the parametric analysis (for further details, see Iacus et al., [2012]).

After obtaining a matched data set, we regress  $Y$  on  $T$  alone using simple ordinary least squares regression, since covariate balance is obtained through matching. We then adopt a simulation-based approach to presenting key quantities of interest (King, Tomz, & Wittenberg, 2000). We collect our model estimates in the stacked column vector  $\hat{\psi} = \{\hat{\beta}, \hat{\sigma}^2\}$ , which forms the mean of a multivariate Normal distribution with variance equal to the model covariance matrix,  $\hat{V}(\hat{\psi})$ . We reduce model dependence by drawing tens of thousands of numbers from this distribution and averaging uncertainty across the simulated parameter estimates. This further allows us to simulate counterfactuals by comparing expected values of each of the outcome variables across treatment states, with  $T = 0$  for counties below mean exposure and  $T = 1$  for counties with above mean exposure to incarceration. The simulated expected values are used to visualise uncertainty surrounding model parameters and to directly compare the distributions of  $E(Y | T = 0)$  and  $E(Y | T = 1)$ .

### Sensitivity analyses

Given the lack of instrumentation in this cross-sectional setting, we refrain making causal claims. However, we conduct a simple non-parametric sensitivity analysis that allows us to precisely quantify the amount of unmeasured confounding that would in theory be required to eliminate our estimated treatment effect  $\hat{\beta}^\wedge$ . For (theoretically dichotomised) treatment and control units, let  $U$  denote an unmeasured confounder. Then the bias factor,  $B$ , is defined as the difference between  $\hat{\beta}^\wedge$  and what  $\hat{\beta}^\wedge$  would have been had we controlled for  $U$  as well, net of our other control variables. We assume that  $U$  is binary and that the effect of  $U$  on  $Y$  is the same across both treatment states (i.e., no  $U$ -by- $T$  interaction). We then define

$$\gamma = E(Y | U = 1, T, X) - E(Y | U = 0, T, X)$$

As the effect of the unmeasured confounder on the outcome, net of the treatment and control variables. We also define

$$\delta = P(U = 1 | T = 1, X) - P(U = 1 | T = 0, X)$$

As the difference in the prevalence of the unmeasured confounder between the treatment and control groups. Then it can be shown that  $B = \gamma \times \delta$  (VanderWeele & Arah, 2011; VanderWeele, 2015, pp. 68–69). In assessing the sensitivity of our model coefficients to unmeasured confounding, we ask how large  $\gamma$  would have to be in order to reduce our estimated effect size  $\hat{\beta}^\wedge$  to zero. We address this question by visualising how  $B$  changes as the two sensitivity parameters (co-)vary across a range of possible values.

As a final robustness check, we run a series of cross-sectional regressions with data from 2014 alone (the year with the best data coverage), without matching. For this particular year, we have access to additional control variables that help inform the sensitivity analysis, including residential segregation by race, unemployment and poverty rates, and intergenerational income mobility. These additional data and their sources are described in SI Table A1. Due to issues of multicollinearity, we present a series of regression models in which the control variables are added and removed one at a time. We then assess how the coefficient for incarceration changes in response to each new covariate. All analyses are conducted in R version 4.0.2 (R Core Team, 2020).

## Findings

### Panel data regressions

Table 2 shows results for a set of two-way fixed-effects regressions wherein the incarceration variable is instrumented as described above. The outcome and treatment variables are residualised with respect to the

**Table 2**  
Instrumented two-way fixed effects regression models.

	Communicable	Non-Communicable	All-Cause
Incarceration rate ( $t - 1$ )	2.9 (2.1, 3.7)	26.0 (22.0, 30.0)	26.0 (22.0, 30.0)
Incarceration rate ( $t - 5$ )	0.8 (0.2, 1.4)	20.0 (16.0, 24.0)	20.0 (16.0, 24.0)
Incarceration rate ( $t - 10$ )	-0.6 (-1.1, 0.01)	13.0 (9.0, 17.0)	15 (11.0, 19.0)

Notes: The outcome variables are age-standardised mortality rates from communicable diseases in the second column, from non-communicable diseases in the third column, and from all causes in the fourth column. The incarceration variable, lagged by one, five, and ten years, is instrumented as described in the Data and methods section. The corresponding parameter estimates are interpreted as the excess number of deaths per 100,000 county population caused by a standard deviation increase in incarceration rates, after adjusting for violent crime, median household income, high school graduation rates, fraction African Americans, fraction Hispanics, and fraction other ethnic minority (not displayed). 95% confidence intervals derived from robust standard errors are shown in parentheses below each parameter estimate.

control variables listed above, but to avoid redundant clutter, we only display our key quantities of interest, namely the lagged treatment effects. All model specifications in Table 2 yield Wald test statistics for joint significance of more than 500 on 7 degrees of freedom. We find, as shown in the second column, that a standard deviation increase in rates of imprisonment in one year causes 2.9 excess deaths from communicable diseases per 100,000 county population in the following year (95% CI: 2.1, 3.7;  $p < 0.001$ ). Expressed as a percentage change, this amounts to a 7.2% increase in the local communicable death rate — a substantively large effect size. Five years later, as shown in the same column of the second row, the corresponding number is 0.8 (95% CI: 0.2, 1.4;  $p < 0.01$ ), or a 5.6% increase. However, we find no robust effect a decade later, as shown in the same column of the last row. The third column shows that higher incarceration leads to 26 excess deaths from non-communicable diseases (95% CI: 22.0, 30.0;  $p < 0.001$ ) in the short run, 20 excess deaths (95% CI: 16.0, 24.0;  $p < 0.001$ ) in the medium run, and 13 excess deaths (95% CI: 9.0, 17.0;  $p < 0.001$ ) in the long run per 100,000 population. In percentage terms, these effects are equivalent to a rise in non-communicable deaths by 3.4%, 3.2%, and 2.8%, respectively. Finally, as shown in the last column, the treatment effects for mortality from all causes is 26 excess deaths (95% CI: 22.0, 30.0;  $p < 0.001$ ) in the short run, 20 excess deaths (95% CI: 16.0, 24.0;  $p < 0.001$ ) in the medium run, and 15 excess deaths (95% CI: 11.0, 19.0;  $p < 0.001$ ) in the long run per 100,000 population, corresponding to increases of 3.5%, 3.1%, and 3.0%, respectively.

These findings provide strong evidence in favour of our principal hypothesis, namely that high rates of incarceration impact population-level mortality outcomes in short, medium, and long run. The fact that our estimated coefficients for communicable and non-communicable disease deaths — insofar as they partition the outcome space — do not add up mechanically to the coefficient for all-cause mortality is most likely due to differences in age-specific mortality rates by cause of death, to the effect that the two categories do not sum (exactly) to unity when age-standardised on a separate basis.

For the sake of comparison, we also run non-instrumented versions of the two-way fixed-effects models. As shown in SI Table A2, these consistently produce smaller parameter estimates, but remain robust. We surmise that the discrepancy in effect sizes derives from attenuation bias in the non-instrumented panel regression — possibly due to measurement error or omitted variable bias — or from differences between the *local* average treatment effect estimated by the instrumented models and the *population* average treatment effect estimated by the non-instrumented models (see Card, 2001).



### Between-county matched regressions

We proceed to the between-county analysis by applying a matching algorithm to time-averaged versions of all our covariates after splitting counties into those with above versus below mean exposure to incarceration. The matching procedure results in a pruned data set composed of 1694 counties. As reported in SI Table A3, the diagnostics reveal a high degree of balance improvement since the empirical covariate distributions in both the treatment and control groups are now similar, meaning the smaller sample size strengthens rather than undermines the subsequent statistical inference. We then regress our three outcome variables on the treatment variable using simple ordinary least squares, the results of which are displayed in Table 3. A standard deviation increase in incarceration rates is associated with 4.3 excess communicable deaths (95% CI: 3.7, 4.9;  $p < 0.001$ ), 44.2 excess non-communicable deaths (95% CI: 39.7, 48.7;  $p < 0.001$ ), and 56.1 excess all-cause deaths (95% CI: 50.8, 61.4;  $p < 0.001$ ). Expressed in terms of semi-elasticities, this amounts to a 8.9%, 5.3%, and 5.9% increase in mortality, respectively. We note that the larger between-unit effect sizes are expected insofar as they reflect greater inter-county (as opposed to intra-county) variation in mortality outcomes.

To get a more intuitive sense of what these numbers mean in substantive terms, we predict and plot the conditional expectation of each outcome variable by treatment status using a simulation-based approach, as described in the Data and methods section. Fig. 1 shows that in counties with low rates of incarceration, mortality from communicable diseases is expected to be 44.4 deaths per 100,000 population (95% CI: 43.8–45.0;  $p < 0.001$ ). In counties with high rates of incarceration, the corresponding number is 52.0 (95% CI: 51.4–52.7;  $p < 0.001$ ). For non-communicable diseases, a shift from control to treatment increases the expected mortality rate from around 800 (95% CI: 795–806;  $p < 0.001$ ) to around 879 deaths per 100,000 population (95% CI: 873–884;  $p < 0.001$ ). For all-cause mortality, the corresponding numbers are 917 (95% CI: 910–923;  $p < 0.001$ ) and 1016 (95% CI: 1009–1022;  $p < 0.001$ ), respectively.

Fig. 2 visualises the sensitivity analysis using the parameter estimates from Table 3, with  $\delta$  denoting the degree of selection on the unmeasured confounder across the two treatment states (ranging from 0 to 1, with higher values indicating a higher prevalence of the confounder in the treatment group — i.e., in counties with high rates of incarceration), and  $\gamma$  denoting the magnitude of the effect of  $U$  on the outcome, above and beyond that of the treatment and other controls, that would be required to completely eliminate the effect of incarceration on the three outcome variables. The reader will note that even for unusually high levels of selection on the unmeasured confounder, the effect of  $U$  on the outcome would have to be large in order to nullify that of incarceration, especially for non-communicable disease deaths and all-cause mortality. For instance, even when the difference in the prevalence of the

**Table 3**  
Between-county matched regression models.

	Communicable	Non-Communicable	All-Cause
Incarceration rate	4.3 (3.7, 4.9)	44.2 (39.7, 48.7)	56.1 (50.8, 61.4)
Multiple $R^2$	14.8%	18.4%	20.2%
Observations	1694	1694	1694

Notes: The outcome variables are age-standardised mortality rates from communicable diseases in the second column, from non-communicable diseases in the third column, and from all causes in the fourth column. The association between treatment and outcome is estimated by applying a simple linear regression model to a pruned data set that is pre-processed using coarsened exact matching. Counties are matched on the variables listed in the Data and methods section (see also SI Table A3). All variables are time-averaged over the sample period. Parameter estimates are interpreted as the number of excess deaths associated with a standard deviation increase in incarceration rates. 95% confidence intervals are shown in parentheses below each parameter estimate.

confounder between the treatment and control groups is as high as 0.8 — an unlikely scenario —  $U$  would have to cause around half a dozen excess communicable deaths, over 50 non-communicable deaths, and almost 75 all-cause deaths per 100,000 population — above and beyond the effects of  $T$  and  $X$  — to eliminate our model estimates. We believe it is plausible that most relevant confounders are already included in our matrix of covariates. As such, a more plausible value of  $\delta$  would be at the lower end of the X-axis in Fig. 2. At, say,  $\delta = 0.2$ , the effect of  $U$  on  $Y$  would have to be nearly 25 excess communicable deaths, around 225 non-communicable deaths, and around 275 all-cause deaths, which seems highly improbable. In short, an unusually large amount of unmeasured confounding would be needed to explain away the estimated impact of incarceration on between-county mortality patterns.

### Cross-sectional regressions

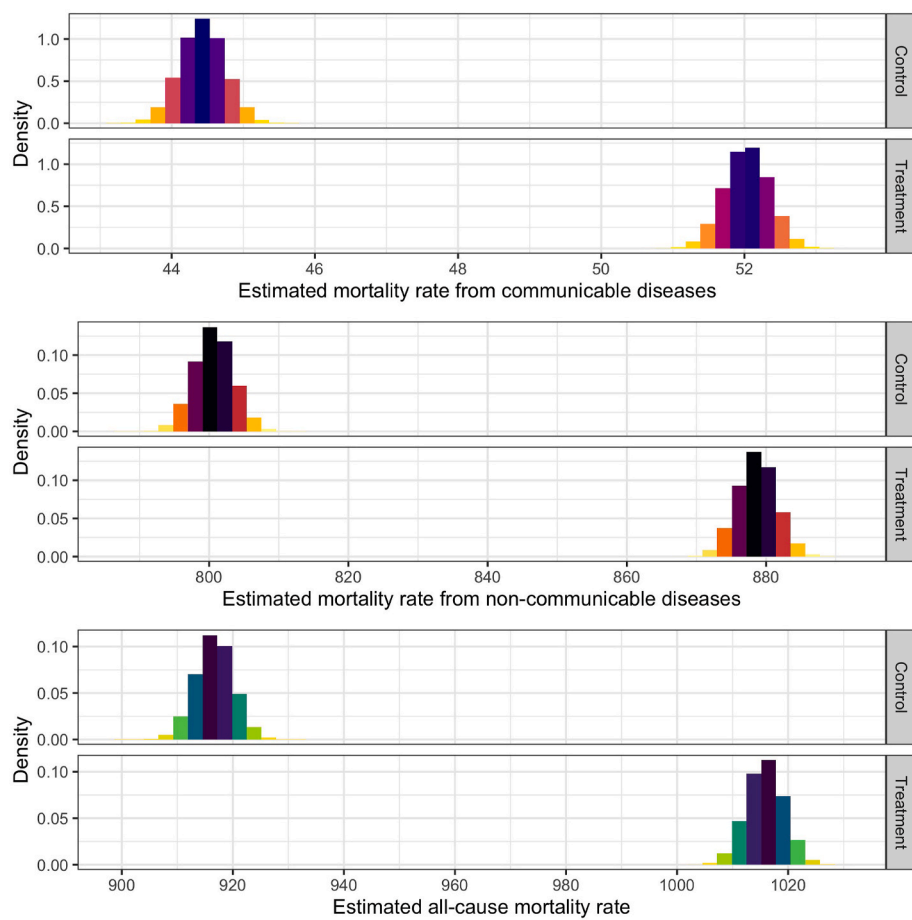
Finally, we assess the robustness of the hypothesised relation between incarceration and mortality by running a series of cross-sectional regressions with additional data from 2014, without pre-processing the data via matching. Additional control variables include the local unemployment and poverty rates, a measure of absolute income mobility at the county level, income inequality as measured by a local Gini index, residential segregation by race, and the percentage of the county population with no health insurance. To avoid over-specification, we add and remove one control variable at a time. However, we adjust for state-fixed effects in all models. As reported in SI Tables A4–A6, the estimated association between incarceration and each of the outcome variables remains stable across all specifications, which further confirms the robustness of our principal findings.

### Concluding discussion

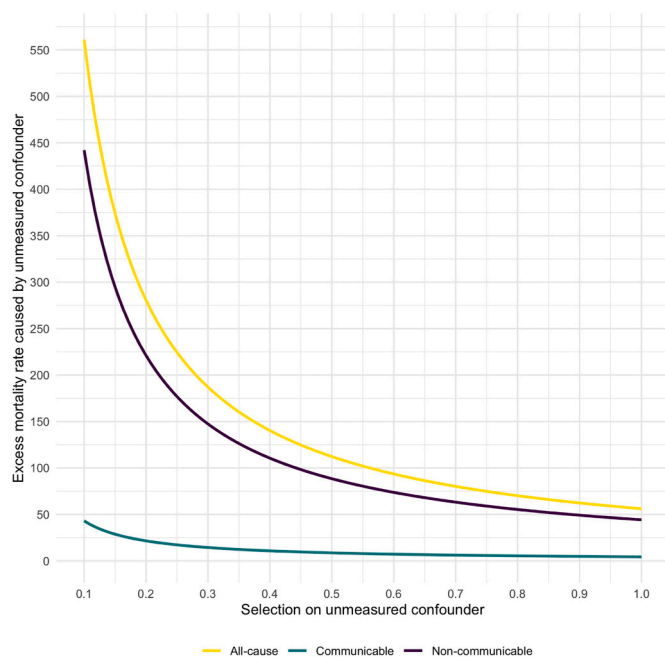
Our findings confirm the hypothesis that high rates of incarceration exert a substantively large impact on county-level mortality rates. Our joint usage of variation within and between units demonstrates that penal expansion can be construed as a distal determinant of declining health and deepening health inequality across the United States. We view our paper as a contribution to a growing literature on the health impacts of incarceration. The novelty of our approach is the combination of a new methodological design with previously unavailable county-level data to study dynamic changes in cause-specific mortality rates. To our knowledge, this is the first analysis to adopt this threefold approach to generate novel empirical evidence at the population level, at a high level of geographical resolution.

However, we acknowledge the limitations of this study. First of all, as we are unable to empirically verify that our models capture exogenous treatment variation, we recognise that our parameter estimates may suffer from residual confounding or other sources of bias. Nevertheless, in the panel data analysis, the use of a novel instrumentation technique coupled with unit- and time-fixed effects provides a stringent framework for causal inference that minimises the likelihood of obtaining spurious associations. In the cross-sectional analysis, we have sought to adjust for the most important and likely confounders of the relevant relationships and we have used matching as a means of mimicking an experimental research design. Both analyses yield robust parameter estimates and are not subject to high levels of model dependence. Our sensitivity analysis suggests that an inordinate amount of unmeasured confounding would be required to explain away the estimated effect of incarceration on mortality outcomes.

Second, although our findings provide meaningful quantitative estimates of the hypothesised causal associations, we do not have the data to flesh out the relevant pathways or to detail the precise mechanisms by which high rates of incarceration exert the kinds of effects that we propose. We are also not able to capture the broader correctional population, such as those on probation or parole, nor potentially relevant patterns of migration or other demographic changes. Moreover, our data



**Fig. 1.** Density plots of expected outcome values conditional on treatment state. In the top panel, the outcome variable is mortality from communicable disease, in the middle panel the outcome variable is mortality from non-communicable diseases, and in the bottom panel the outcome variable is all-cause mortality per 100,000 population. Each model compares counties with incarceration rates at one standard deviation below the mean ('Control') to those with incarceration rates at one standard deviation above the mean ('Treatment'). The association between treatment and outcome is estimated by applying a simple linear regression model to a pruned data set that is pre-processed using coarsened exact matching. Counties are matched on the variables listed in the Data and methods section (see also SI Table A3). All variables are time-averaged over the sample period.  $N = 1694$ .



**Fig. 2.** Sensitivity analysis plot to assess unmeasured confounding of the estimated effect  $\beta^A$  of incarceration on each of the outcomes in Table 3. Values of  $\delta$  (X-axis) and  $\gamma$  (Y-axis) that lie on the lines would completely eliminate the corresponding effect estimates. Values above the plotted curve would reverse the sign of the estimated effect.

do not allow us to estimate how much of the excess mortality is due to high death rates amongst former prisoners and how much is due to spillover effects on local areas. With the data at hand, we cannot explain in detail why deaths from communicable, maternal, neonatal, and nutritional diseases are affected in the short and medium run, whilst non-communicable disease deaths are also impacted in the long run. We note that the lag between exposure and outcome for communicable diseases is, in almost all cases, short, whereas non-communicable diseases involve distributed lags that can extend over a prolonged period. We suspect that our results reflect the ways in which incarceration acts upon and activates the broader determinants of health by corroding social ties and the collective efficacy of neighbourhoods and communities, a process which in turn becomes durably embodied by local populations in ways that manifest as chronic ill health in the longer run, for example by influencing behaviours, both health promoting (for example, we know that strong social ties are associated with improved blood pressure control) and harming (unhealthy behaviours such as smoking). Of course it would require longitudinal data at the individual and community level to tease this out but we believe that our findings are consistent with the known causal pathways (see Nosrati & King, 2021 for further theoretical discussion). We note, furthermore, that the size of our parameter estimates — although consistent in sign and overall robustness — vary somewhat across model specifications. This most likely reflects differences in variation within and between units over time. To better address all these points, future research should seek to integrate multilevel data that account for the complex interconnections between individuals, neighbourhoods, local communities, and the criminal justice system across time and space.

Third, we estimate an average treatment effect, yet existing research on incarceration shows that the penal state disproportionately targets a

specific fraction of the American population, namely African Americans at the bottom of the social structure (Wacquant, 2009). With our current data and ecological approach, we are unable to assess any potential treatment effect heterogeneity — that is, whether incarceration is more harmful to some than to others. We are also unable to offer a more refined investigation of social and economic factors such as income, education, or ethno-racial division, all of which are imperfectly measured at an aggregate level in our data set.

These limitations do not prevent us from concluding that high rates of incarceration shape unequal life chances in the United States and can harm population health. We provide evidence for a robust and substantively large net causal linkage between incarceration and communicable, non-communicable, and all-cause mortality rates. This implies that protective rather than punitive criminal justice policies may help to shield vulnerable populations from premature mortality and to reduce regional inequalities in health and well-being.

### Declaration of competing interest

None of the authors have any conflicts of interest or financial disclosures to declare.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2021.100827>.

### Ethical statement

This paper does not involve the usage of primary data sources and there are no ethical concerns to report.

### Author statement

Elias Nosrati: Conceptualisation, Data curation, Methodology, Formal analysis and Investigation, Writing – original draft, Writing – review & editing; Jacob-Kang Brown: Data curation, Writing – review & editing; Michael Ash: Writing – review & editing; Martin McKee: Writing – review & editing; Michael Marmot: Writing – review & editing; Lawrence P. King: Conceptualisation, Writing – review & editing.

### References

- Akiyama, M. J., Spaulding, A. C., & Rich, J. D. (2020). Flattening the curve for incarcerated populations — covid-19 in jails and prisons. *New England Journal of Medicine*, 382(22), 2075–2077.
- Binswanger, I. A., Stern, M. F., Deyo, R. A., Heagerty, P. J., Cheadle, A., Elmore, J. G., et al. (2007). Release from prison — a high risk of death for former inmates. *New England Journal of Medicine*, 356(2), 157–165.
- Card, D. (2001). Estimating the return to schooling. *Econometrica*, 1127–1152.
- Daza, S., Palloni, A., & Jones, J. (2020). The consequences of incarceration for mortality in the United States. *Demography*, 57, 577–598.
- Gilmore, R. W. (2007). *Golden gulag: Prisons, surplus, crisis, and opposition in globalizing California*. Berkeley: University of California Press.
- Hinds, O., Lu, O., Wallace-Lee, J., & Kang-Brown, J. (2020). *Reconstructing how counties contribute to state prisons*. New York: Vera Institute of Justice.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15, 199–236.
- Hooks, G., & Sawyer, W. (2020). *Mass incarceration, COVID-19, and community spread*. Northampton, MA: Prison Policy Initiative.
- Iacus, S. M., King, G., & Porro, G. (2019). A theory of statistical inference for matching methods in causal research. *Political Analysis*, 27(1), 46–68.
- Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: coarsened exact matching. *Political Analysis*, 20(1), 1–24.
- IHME. (2017). *United States life expectancy and age-specific mortality risk by county 1980–2014*. Seattle, United States: Institute for Health Metrics and Evaluation.

- Johnson, E. I., & Easterling, B. (2012). Understanding unique effects of parental incarceration on children: Challenges, progress, and recommendations. *Journal of Marriage and Family*, 74(2), 342–356.
- Kajeepeeta, S., Mauro, P. M., Keyes, K. M., El-Sayed, A. M., Rutherford, C. G., & Prins, S. J. (2021). Association between county jail incarceration and cause-specific county mortality in the USA, 1987–2017: A retrospective, longitudinal study. *The Lancet Public Health*, 6(4), e240–248.
- Kajeepeeta, S., Rutherford, C. G., Keyes, K. M., El-Sayed, A. M., & Prins, S. J. (2020). County jail incarceration rates and county mortality rates in the United States, 1987–2016. *American Journal of Public Health*, 110(S1), S109–S115.
- King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: improving interpretation and presentation. *American Journal of Political Science*, 44, 341–355.
- Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior. Special issue*, 80–94.
- Marshall, P. (2021). *A state-by-state look at coronavirus in prisons*. <https://tinyurl.com/y8hv3kc5>. (Accessed 28 April 2021).
- Massoglia, M. (2008). Incarceration as exposure: The prison, infectious disease, and other stress-related illnesses. *Journal of Health and Social Behavior*, 49(1), 56–71.
- Massoglia, M., & Pridemore, W. A. (2015). Incarceration and health. *Annual Review of Sociology*, 41, 291–310.
- National Research Council. (2014). In J. Travis, B. Western, & S. Redburn (Eds.), *The growth of incarceration in the United States: Exploring causes and consequences*. Washington, DC: National Academies Press.
- Ndeffo-Mbah, M. L., Vigliotti, V. S., Skrip, L. A., Dolan, K., & Galvani, A. P. (2018). Dynamic models of infectious disease transmission in prisons and the general population. *Epidemiologic Reviews*, 40(1), 40–57.
- Nosrati, E., & King, L. P. (2021). Punitive social policy and vital inequality. *International Journal of Health Services*. In press.
- Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: Race and class inequality in U.S. Incarceration. *American sociological review*, 69(2), 151–169.
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reinhart, E., & Chen, E. L. (2020). Incarceration and its disseminations: COVID-19 pandemic lessons from Chicago's cook county jail. *Health Affairs*, 39(8), 1412–1418.
- Saloner, B., Parish, K., Ward, J. A., DiLaura, G., & Dolovich, S. (2020). COVID-19 cases and deaths in federal and state prisons. *The Journal of the American Medical Association*, 324(6), 602–603.
- Sampson, R. J., & Loeffler, C. (2010). Punishment's place: The local concentration of mass incarceration. *Daedalus*, 139(3), 20–31.
- Therborn, G. (2013). *The killing fields of inequality*. Cambridge: Polity Press.
- Turney, K. (2014). Stress proliferation across generations? Examining the relationship between parental incarceration and childhood health. *Journal of Health and Social Behavior*, 55(3), 302–319.
- VanderWeele, T. J. (2015). *Explanation in causal inference. Methods for mediation and interaction*. New York: Oxford University Press.
- VanderWeele, T. J., & Arah, O. A. (2011). Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*, 22(1), 42–52.
- Wacquant, L. (2009). *Punishing the poor: The neoliberal government of social insecurity*. Durham: Duke University Press.
- Western, B. (2018). *Homeward: Life in the year after prison*. New York, NY: Russell Sage Foundation.
- Wildeman, C. (2009). Parental imprisonment, the prison boom, and the concentration of childhood disadvantage. *Demography*, 46(2), 265–280.
- Wildeman, C., & Wang, E. A. (2017). Mass incarceration, public health, and widening inequality in the USA. *The Lancet*, 389(10077), 1464–1474.
- Wildeman, C., & Muller, C. (2012). Mass imprisonment and inequality in health and family life. *Annual Review of Law and Social Science*, 8, 11–30.
- Wildeman, C., Wakefield, S., & Turney, J. (2013). Misidentifying the effects of parental incarceration? A comment on Johnson and Easterling (2012). *Journal of Marriage and Family*, 75(1), 252–258.
- Wildeman, C., Schnittker, J., & Turney, K. (2012). Despair by Association? The mental health of mothers with children by recently incarcerated fathers. *American Sociological Review*, 77, 216–243.
- Wilkinson, R., & Pickett, K. (2018). *The inner level: How more equal societies reduce stress, restore sanity and improve everyone's well-being*. London: Penguin.
- Wilkinson, R., & Pickett, K. (2010). *The spirit level: Why equality is better for everyone*. London: Penguin.
- Weidner, R. R., & Schultz, J. (2019). Examining the relationship between US incarceration rates and population health at the county level. *SSM Population Health*, 9, 100466.
- Western, B. (2006). *Punishment and inequality in America*. New York, NY: Russell Sage Foundation.
- Zlodre, J., & Fazel, S. (2012). All-cause and external mortality in released prisoners: Systematic review and meta-analysis. *American Journal of Public Health*, 102(12), e67–e75.

### Further reading

Opportunity Insights, n.d. <https://opportunityinsights.org>. Accessed 12 September 2020.