Punitive social policy and vital inequality

Elias Nosrati, Lawrence P. King

Geographical inequalities in life and death are amongst the world's most pronounced in the United States. However, the driving forces behind this macroscopic variation in population health outcomes remain surprisingly understudied, both empirically and theoretically. The present article steps into this breach by assessing a number of sociologically informed hypotheses surrounding the underlying causes of such spatial heterogeneity. Above and beyond a range of usual suspects, such as poverty, unemployment, and ethno-racial disparities, we find that a hitherto neglected *explanans* is prison incarceration. In particular, through the use of previously unavailable county-level panel data and a compound instrumentation technique suited to isolating exogenous treatment variation, high imprisonment rates are shown to substantially increase the population-wide risk of premature death. Our findings contribute to the sociology of punishment and population health by relating the rise of the carceral state to the amplification of geographically anchored unequal life chances.

Introduction

The unequal distribution of disability, disease, and death in the United States is amongst the most pronounced in the world and, as documented by a number of recent studies (e.g., Chetty et al. 2016), the health gap between the top and the bottom of the social order is rapidly increasing. A distinctive feature of this inequality is its geographical patterning, as evidenced by large and persistent spatial variation in key outcomes such as life expectancy at birth (Dwyer-Lindgren et al. 2017). There appears to exist a strong correspondence between social and physical space, yet rigorous empirical studies of such geographical heterogeneity remain scarce. The present article probes the macroscopic forces that might explain this phenomenon and identifies an institution of major sociological interest — the prison — as an important yet largely neglected determinant of socially constructed unequal life chances of human organisms, or *vital inequality* (Therborn 2013: 49). Using previously unavailable panel data at the county level and a novel instrumentation technique, we provide causal evidence that penal expansion has deepened geographically anchored disparities in survival and well-being. Our investigation unfolds in three principal

steps. First, we survey previous scholarship seeking to explain the spatial manifestation of vital inequality, and we articulate a unified theory of hyper-incarceration as a vector of ill health. Our argument centres around the dual nature, both material and symbolic, of human capabilities to function and flourish, upon which the punitive regulation of social ills is hypothesised to have a durably corrosive impact. Second, we present a multi-stage empirical analysis of county-level life expectancy and premature mortality risk in which we compare and contrast the explanatory power of various variables of sociological interest and where we also introduce a compound instrumental variable for incarceration suited to isolating exogenous treatment variation within counties over time. We complement our panel models with an investigation of between-county inequalities using a matched regression approach. Finally, we discuss the implications of our analysis for the discipline of sociology and the future of public policy. Our findings provide new empirical evidence surrounding the nexus of punishment and population health, whilst also introducing a compound instrumentation technique that can readily be employed by other scholars of incarceration. In addition, our paper adds to a nascent research domain that extends the analytic purview of sociology to include vital inequality as a central object of enquiry.

What explains spatial variation in population health?

In sociology, vital inequalities are usually explained with reference to the social determinants of health (Berkman, Kawachi, and Glymour 2014), including cultural frames (Hall and Lamont 2009) and institutional arrangements (Beckfield et al. 2015). A distinctively sociological approach is to relate the distributional dynamics of population health to the wider determinants of unequal life chances. These determinants, however, are typically used to account for disparities in individual-level outcomes rather than broader spatial configurations. The most dominant framework for studying key outcomes at a higher level of aggregation is offered by the rich literature on neighbourhood effects wherein community contexts are viewed as powerful predictors of well-being (Sampson 2003), but there is — to our knowledge — no explicit sociological account of what drives macro-level geographical heterogeneity in population health in the United States (or elsewhere). The textbook answer to the question "What accounts for regional disparities in life expectancy?" would most likely be fourfold: material deprivation, race, crime and violence, and access to healthcare. Key explanatory variables would thus include income and poverty, labour market conditions and employment status, ethno-racial and demographic factors, as well as health insurance coverage.

In this paper, we empirically test a series of interrelated hypotheses corresponding

to this sociological approach — that geographical variation in material, demographic, and institutional conditions drives geographical variation in population health outcomes. To the default list of hypothesised explanatory variables, we add an important object of sociological enquiry that forms a distinctive feature of American public policy, namely high rates of incarceration. Not only have sociologists been interested in explaining gargantuan growth in the nation's correctional population — which has ballooned to reach over 2 million individuals behind bars and almost 7 million individuals in total (National Research Council 2014) — but they have also examined how this punitive upsurge has impacted durable patterns of social inequality (Western 2006). In particular, the criminal justice system has been construed as a vector of unequal life chances, in the broadly Weberian sense of probabilistically defined opportunity structures by which groups and individuals access and appropriate social goods. However, the precise nature of the relationship between high rates of incarceration and human welfare remains unclear, as previous scholarship has tended (more often than not) to ignore that form of inequality of which the notion of "life chances" offers an almost literal echo, namely inequalities in life and death. In the following section, we provide an empirical and theoretical rationale for our emphasis on the penal state.

Punishment and population health: theory and evidence

Previous social scientific scholarship has viewed population health as an index of "social success" (Hall and Lamont 2009). According to this literature, a successful society is one that enhances the functional capabilities of its members to live healthy and meaningful lives, both individually and collectively (Sen 1999). A central argument here is that such capabilities are not merely a function of technical biomedical innovations nor of simple behavioural templates, but are dually shaped by historically situated uses of material and symbolic resources. Human health, in other words, is a product of social relations involving institutional practices and cultural repertoires, systems of positions and networks of dispositions. However, whereas the linkages mapping various social phenomena to health and illness have been extensively debated (Link and Phelan 1995) and empirically reviewed (Berkman, Kawachi, and Glymour 2014), the upstream determinants of capability expansion remain opaque (Evans 2009).

To specify the antecedents of functional capabilities on a macroscopic scale, a more recent body of scholarship has viewed population health not only as socially mediated but also as a locus of distributional conflict (Therborn 2013). According to this line of argument, human organisms are embedded within social constellations that are structured by the institutional organisation of social power (Beckfield 2018). Vital inequalities are therefore rooted in the durably institutionalised power of some over

the material life chances of others. We take our cue from this literature by viewing the prison as a potent force in the (re)making of vital inequality insofar as it moulds, from above, the social relations that impinge on human welfare: it is a politically rooted institution that actively stratifies a population, delineates and aligns group boundaries, and thus "realises" (Bourdieu 1993) material and symbolic divisions (Shannon and Uggen 2012).

We consider high rates of incarceration — or, more precisely, hyper-incarceration (Wacquant 2010) — to form part of a broader public policy repertoire that, in turn, is expressive of distributional struggles. This conceptualisation of penality as public policy is historically motivated, as initial the conception of the prison in late sixteenth-century Europe was as "an instrument of social policy with regard to beggars" (Geremek 1994: 207) geared towards the coercive regulation of poverty at the dawn of modern capitalism. A means of warehousing landless vagrants uprooted by the enclosure movement and of curbing the social convulsions wrought by the sudden commodification of labour, the carceral wing of the state was thus, from its origins, an institutional force capable of conferring upon the social determinants of health their causal efficacy. Contemporary evidence has shown that America's punitive upsurge constitutes a similarly oriented way of managing various social ills — from economic deprivation to ethno-racial enmity — via the ideology of "law and order" (National Research Council 2014; Wacquant 2009; see also Muller and Schrage 2019). However, due to the politics of federalism and locally nested differences in organisational cultures that shape the political use of public and private institutions, there is regional variation in the kinds of policy scripts that prevail in each location. Such structural variation can help account for geographically patterned mortality burdens across the United States (Kunitz 2015). This is what motivates our hypothesis of a spatially manifested linkage between incarceration and vital inequality.

Causes and mechanisms

From a rich body of prior research (Massoglia and Pridemore 2015; Wildeman and Wang 2017), we identify three principal mechanisms by which punitive social policies, as epitomised by high rates of incarceration, can impact human functional capabilities. *Relegation* is the mechanism by which the criminal justice system causes groups and individuals to experience downward social mobility. This happens not only through immediate physical seclusion but also, subsequently, through the indelible blemish of a criminal record (Pager 2003), which increases the likelihood of experiencing poverty (Rabuy and Kopf 2015) and cements a sense of social dishonour (Western 2018; Manza and Uggen 2008). Relegation impacts not only those who are incarcerated but also the communities in which they live, where the removal primarily

of working-age men has permanent spillovers in the form of disrupted social ties and territorial stigmatisation. Sociologists have paid particular attention to how parental incarceration impacts "fragile families" (Wildeman and Western 2010) and their neighbourhoods (Clear 2007; Sampson and Loeffler 2010; Leibbrand et al. 2019), precipitating a sequence of intergenerationally transmitted and spatially concentrated adversities that shape functional development over the life course (Wildeman 2009; Wildeman and Muller 2012; Wakefield and Wildeman 2013; Turney 2014). In short, relegation is a mechanism whereby incarceration acts upon and activates the social determinants of health.

Amplification is the mechanism by which already operant social determinants of health are magnified or exacerbated. A key example is the tangled interconnection between poverty and punishment fostered by the punitive treatment of social disadvantage. Evidence suggests that aggressive policing tends to target activities such as public urination, rough sleeping, or begging for food (Thompson 2010: 712). At the same time, whereas rates of homelessness are 21 per 10,000 population for the general public, for those who have been imprisoned once or more than once, the corresponding numbers are 141 and 279, respectively. In other words, for those with multiple encounters with the criminal justice system, homelessness rates are thirteen times that of the overall populace. According to data from the Bureau of Justice Statistics, high rates of recidivism imply that almost half of those released from prison are rearrested within one year, over two-thirds within three years, and over 80% within nine years, thus creating a structural interlock between penal confinement and material hardship (Couloute 2018). Hyper-incarceration, then, can work as an amplifier of the social causes of ill health amongst society's most vulnerable.

The third and final mechanism is that of *corrosion*, which creates and compounds various modalities of "social sundering" (Therborn 2013: 22–28). This involves the enduring decline of social cohesion driven by fractured social networks, neighbourhood violence, and fading collective imaginaries (Patillo, Weiman, and Western 2004). The importance of stable systems of meaning and collective representation, of recognition and reciprocity, and of communal cooperation and solidarity to the fostering of welfare and well-being has been highlighted in previous sociological scholarship on population health (Hall and Lamont 2009; Sampson 2003). We connect this insight to empirical research on incarceration by arguing that punitive interventionism at the bottom of the class structure inflicts long-lasting harm not only by socially relegating groups and individuals or by amplifying pre-existing adversity, but also by warping the moral valence of salutogenic social relations. This is conveyed, on the one hand, by the high risk of death in the immediate period after release from prison, especially from suicide or violence (Binswanger et al. 2007; Zlodre and Fazel 2012), and, on the other hand, by impacted local communities being thrust into collective disarray (Clear 2007; Western

2018). Such fissuring of social ties and the breakdown of moral orders are potent means by which population health is durably affected.

Running through all three mechanisms are two distinct modes of causal efficacy. On the one hand, the impact of criminal justice on population health can occur acutely, such as in the case of police killings and their spillovers (Bor et al. 2018) or the adverse experience of parental incarceration, for parents and children alike (Wildeman, Goldman, and Lee 2019; Turney 2018; Wakefield and Wildeman 2013). On the other hand, the rapid growth of the carceral state over several decades has translated into a set of chronic exposures embedded in the "wear and tear" of everyday life. It is plausible to surmise that the emergence of incarceration as a normalised stage in the life course of young African American men with low levels of formal education (Pettit and Western 2004) forms a vector of cumulative biological burdens whereby the lived experience of social adversity is deposited in the human body in the form of neuroendocrine traits that govern core pathogenic parameters (Seeman et al. 2001; Massey 2004; Wolfe, Evans, and Seeman 2012), including the transcriptional dynamics of the genome (Cole 2009; Slavich and Cole 2013). Both acute and chronic stressors can thus be at work in mapping punitive social policy to vital inequality (Massoglia 2008).

In summary, we posit that hyper-incarceration is a powerful institutional force that mediates and modifies, amplifies and aggravates the dynamics of vital inequality. Our argument weds theory and evidence from prior research, yet identifies a gap that our subsequent empirical analysis seeks to address: although previous studies have shed light on the effects of imprisonment on health at the level of individuals and communities, evidence at the population level remains sparse (Nosrati et al. 2018), especially when it comes to the assessment of premature mortality (Pridemore 2014) and its distinctively spatial patterning (Chetty et al. 2016; Dwyer-Lindgren et al. 2017). In addition, most of the extant literature relies on a limited number of data sets and methods that produce causal identification strategies of varying plausibility (Massoglia and Pridemore 2015), notably in constructing appropriate comparison groups (Johnson and Easterling 2012; Wildeman, Wakefield, and Turney 2013). In what follows, we seek to avoid these pitfalls by using a new county-level data set and a novel instrumentation technique suited to isolating exogenous treatment variation.

Hypotheses, data, and methods

Hypotheses

Against this conceptual and empirical backdrop, our principal hypothesis is that variation in population health can, at least in part, be explained by variation in rates of incarceration, above and beyond other economic and institutional factors. Through whatever pathway — be it relegation, amplification, or corrosion — we find it

plausible that areas experiencing significant expansions of their penal apparatus and rapid swelling of their correctional populations will suffer from poorer human welfare, and that this association is causal. Our secondary hypothesis is that this effect will manifest strongly across the human life course, but especially before one reaches old age. This is not only because the criminal justice system targets younger age groups, but also in light of prior evidence on the intergenerational impacts of parental incarceration on childhood and adolescence.

Data

We use three alternative outcome variables to operationalise vital inequality. The first is life expectancy at birth, one of the most common indices of human welfare. However, this measure is more sensitive to child mortality than to mortality in adult life, and may thus shroud heterogeneity across the life course. For this reason, we also examine two alternative measures of premature mortality risk: the probability of dying between the ages of (a) 25 and 45 and (b) 45 and 65. Given the demographic profile that is disproportionately affected by the penal state, these are the age ranges on which we expect high rates of incarceration to be most impactful. All three outcome variables are drawn from the Institute for Health Metrics and Evaluation (IHME 2017).

To test our hypotheses, we employ two sets of previously unavailable data at the US county level. On the one hand, we use cross-sectional data from 2014 — the most recent year with the best data coverage — that capture key variables of sociological interest. These are county-level median household income (in constant US dollars), unemployment rates, labour force participation rates, poverty rates (as per the federal poverty line), absolute income mobility (defined as the fraction of children earning more than their parents), income inequality (measured by a Gini index within the bottom 99% of the income distribution), residential segregation by race, the population fraction of African Americans, Hispanics, or other ethnic minorities, the population fraction of high school graduates, the percentage of the population without any form of health insurance, violent crime rates, and prison incarceration rates (see below). These data and their sources are described in Appendix Table A1. In addition, to capture (part of) the historical legacy of institutionalised racial domination, we hand-code a binary indicator of whether a state is a former slave state or not.

On the other hand, we use county-level panel data between 1983 and 2014. 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 compiled into annual county-level rates per 100,000 residents aged 15–64 (Hinds et al. 2018). 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 are also excluded from the final analysis. From the above list of variables for which multiple county-year observations are available, we employ a set of baseline controls that are associated with both the treatment and the outcome, namely median household income, annual rates of violent crime, and the fraction of each county population that is African American. These variables are 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 for the panel data are reported in Appendix Table A2.

Methods

We commence our analysis with a series of simple cross-sectional regression models, estimated using ordinary least squares, in which each of the three outcome variables is regressed against a hypothesised explanatory variable, coupled with state-fixed effects. To discern differences in explanatory power and to avoid issues of multicollinearity, we present a separate model for each predictor. To test our hypotheses regarding the driving forces of spatial heterogeneity in population health, we then compare (a) coefficient sizes and (b) relative explanatory power in terms of R^2 across all models.

To examine the causal relation between incarceration and health, we posit the following data-generating process:

$$Y_{it} = T_{i[t-1]}\beta + X_{it}\theta + \mu_i + \varphi_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} denotes one of the three alternative outcome variables as measured in county i at time t; the treatment variable $T_{i[t-1]}$ is the county-level incarceration rate per 100,000 population, lagged by one year to allow for delayed effects; X is a vector of control variables; μ and φ capture unit- and time-fixed effects, respectively; and ε is a stochastic error term. Our principal quantity of interest is β, which is a causal effect parameter to be estimated. However, in an observational study such as this, we do not control the source of variation in the treatment variable, nor can we know for sure if our controls are sufficient to isolate exogenous variation in the treatment. Corresponding parameter estimates may therefore suffer from endogeneity bias. We visualise this identification problem in Figure 1, where the estimated relation between the treatment variable T and the outcome variable Y may be 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 population health, such as locally contingent environmental shocks or welfare-related policy changes.

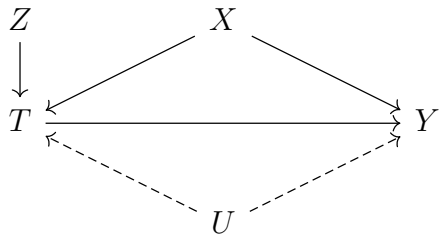

Figure 1: Causal graph depicting the effect of the treatment variable (T = incarceration) on the outcome (Y = life expectancy or premature mortality risk), identified via a compound instrument (Z = unit-specific average incarceration rate $\times$ annual nationwide correctional spending), net of both measured covariates (X = {household income, violent crime, demographics}) and unmeasured confounders (U = {e.g., environmental shocks, local welfare policy}).

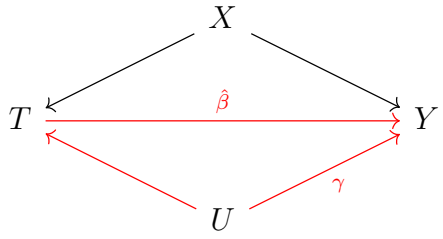

Figure 2: Causal graph depicting the potential sensitivity of the estimated effect $\hat{\beta}$ of the treatment variable (T = incarceration) on the outcome (Y = life expectancy or premature mortality risk) to unmeasured confounding (U = {e.g., environmental shocks, local welfare policy}), net of a set of measured covariates (X = {household income, violent crime, demographics}).

One possible solution to this issue is to construct an instrumental variable Z which is correlated with the treatment but uncorrelated with any other variables in the causal system, thereby isolating exogenous variation in T. We propose a compound instrument derived from the interaction between the unit-specific average exposure to incarceration and annual nationwide correctional expenditure per capita. In other words, $Z_{it} = \overline{T}_i \times C_t$, where $\overline{T}_i$ 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 Appendix Figure A1. We thus obtain a two-stage regression model with the following selection equation:

$$T_{it} = Z_{it}\tau + X_{it}\eta + \alpha_i + \delta_t + \upsilon_{it}. \tag{2}$$

We then re-specify the model in Equation (1) as follows, with $\widehat{T}$ being a vector of fitted values from Equation (2):

$$Y_{it} = \widehat{T}_{i[t-1]}\beta + X_{it}\theta + \mu_i + \varphi_t + \varepsilon_{it}. \tag{3}$$

Any instrumental variable has to meet two principal criteria. First, the relevance criterion requires that the instrument is in fact predictive of variation in the treatment. In other words, there must exist a pathway from Z to T such that aggregate per capita correctional spending correlates with county-level rates 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 τ of Z on T is set to be null, obtaining a χ^2 test statistic of 2,310 on 1 degree of freedom ($p < 0.001$). Hence Z comfortably satisfies the benchmark for identifying a strong instrument. Second, the exclusion criterion requires that the instrument is orthogonal to the outcome over and above its effect on the treatment. The identifying assumption is therefore that the outcome of interest in counties with different propensities to incarcerate will not be affected by changes in aggregate correctional spending other than through the impact of incarceration itself: that is, as per Figure 1, the only pathway linking Z to Y goes via T. We thus argue that the instrument is exogenous because unit-specific shocks in incarceration that deviate from a county's long-run average are generated only by punitive policy shifts that occur independently of any particular county. We control for the endogenous relation between T and Y potentially induced by any time-invariant propensity of counties with a prior health disadvantage to adopt more punitive policy measures by adjusting for county-fixed

effects, whereas year-fixed effects help account for broader changes in the national public health landscape that affect all counties simultaneously.

To our knowledge, no other studies of incarceration have used this instrumentation method. Other instrumental-variable approaches in the extant literature have relied either on broader measures of aggregate-level policy environments (e.g., DeFina and Hannon 2013) or on individual-level variation derived from the random assignment of judges with different sentencing propensities (e.g., Kling 2006). 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 believe that the methodological setup of our paper provides a more stringent framework for causal inference by virtue of the time-series dimension of the 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.

A central threat to our identification strategy is the presence of non-parallel trends across counties with different treatment exposure levels. This is because our approach is akin to a difference-in-difference design wherein the effect of T on Y is compared across counties assigned to (continuous) treatment states as a function of nationwide correctional spending over time. The key assumption of such a design is that the control units provide an appropriate counterfactual of the trend that the treated units would have followed had they not been treated. To test this assumption, we visually inspect aggregate trends for all three outcomes of interest, stratified by whether units experience above or below mean exposure to incarceration. As reported in Appendix Figures A2–A4, we find no indication of the parallel-trends assumption being violated.

Another threat to causal inference is the persistence of residual confounding. Given that we cannot empirically verify that our instrument is strictly exogenous, the probability of such unmeasured confounding is non-zero. To address this concern, 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 $\hat{\beta}$. Let

$$\hat{\beta} = \mathbb{E}(Y \mid T = 1,\ X) - \mathbb{E}(Y \mid T = 0,\ X)$$

denote the expected difference in the outcome variable Y for (theoretically dichotomised) treatment and control units, respectively, net of our matrix of controls X, and let U denote an unmeasured confounder. Then the bias factor, $\mathcal{B}$, is defined as

the difference between $\hat{\beta}$ and what $\hat{\beta}$ would have been had we controlled for U as well. We make the simplifying assumptions 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). For the most part, these assumptions merely serve to simplify the presentation without substantively impacting the outcome of the sensitivity analysis. We now define

$$\gamma = \mathbb{E}(Y \mid U = 1,\, T,\, X) - \mathbb{E}(Y \mid U = 0,\, T,\, X)$$

as the effect of the unmeasured confounder on the outcome, net of the treatment and control variables, as visualised in FIGURE 2. We also define

$$\delta = \mathbb{P}(U = 1 \mid T = 1,\, X) - \mathbb{P}(U = 1 \mid T = 0,\, X)$$

as the difference in the prevalence of the unmeasured confounder between the treatment and control groups. The bias factor is then readily obtained as the product of these two sensitivity parameters: $\mathcal{B} = \gamma \times \delta$ (VanderWeele and Arah 2011; VanderWeele 2015: 68–69). In assessing the sensitivity of our model coefficients to unmeasured confounding, we ask how large γ would have to be in order to reduce our estimated effect size $\hat{\beta}$ to zero. We address this question by visualising how $\mathcal{B}$ changes as the two sensitivity parameters (co-)vary across a range of possible values.

Assessing between-county variation

Although fixed-effects regressions are nearly always preferred in analyses of panel data, we complement the investigation of within-county variation over time with an analysis of between-county variation. This is for two principal reasons. On the one hand, despite its many virtues, the "within" estimator eliminates most of the variation — and hence information — captured by the data and does not always lend itself to readily interpretable predicted values of substantive interest. On the other hand, by isolating inequalities across units, the "between" estimator is geared towards our principal quantity of interest, namely geographical disparities. This allows us to simulate sociologically relevant counterfactuals and compare differences in outcomes in an intuitive way.

However, we acknowledge the methodological flaws of the "between" estimator — which are compounded by the fact that we cannot instrument for incarceration in a cross-sectional setting. In order to render the corresponding parameter estimates more plausible, we employ coarsened exact matching as a non-parametric form of pre-processing the data (Ho et al. 2007; Iacus et al. 2018). The goal of matching is to reduce inefficiency, bias, and model dependence by selecting units of analysis that are similar to one another in all respects except for their treatment status. Thus, in our case, counties are "matched" with other counties that share key characteristics, with

the exception 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 vital inequality. In formal terms, we let Y denote the outcome variable of interest, $T_i \in \{0, 1\}$ is a dichotomised treatment indicator encoding below versus above mean exposure to incarceration, and X is the set of pre-treatment covariates (violent crime, median household income, high school graduation rates, and the county ethno-racial composition). The treatment effect β on a treated unit i is $\beta_i = Y_i(T_i = 1) - Y_i(T_i = 0)$. 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, $\widehat{Y}_i(T_i = 0)$, is equal to $Y_j(T_j = 0)$. Coarsened exact matching temporarily "coarsens" the covariates X into sub-categories using 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]). All analyses are conducted in R version 4.0.2 (R Core Team 2020).

Findings

Preliminary models

A set of preliminary results are displayed in Table 1. These models include key hypothesised predictors of the three outcome variables. To avoid issues of multicollinearity, only a subset of available variables are selected on the basis of overall model fit. In particular, the variable measuring the population percentage without health insurance turns out to be highly correlated with median household income and is therefore omitted from these specifications. Including the variable does not alter any other parameter estimates, but reverses the sign of the variable itself whenever it is used together with median household income. We do, however, assess the robustness of our principal findings to controlling for health insurance coverage in our sensitivity analysis (see below). Moreover, all models are adjusted for state-fixed effects. Each outcome variable is log-transformed, whilst all continuous predictors are standardised by subtracting the mean and dividing by the standard deviation. Parameter estimates are thus interpreted as the (approximate) percentage change in the outcome variable associated with each standard deviation increase in the predictor. We find, as suspected, that local economic conditions, in the form of median household income and unemployment rates, are strongly associated with cross-county variation in life expectancy and premature mortality risk. The same goes

for regional variation in the demographic composition of counties, as well as variation in violent crime rates. Overall, these findings suggest that the sociological inclination to look for economic and historical-institutional explanations of geographical heterogeneity in population health is well-founded. Moreover, we find that our emphasis on the penal state is justified: higher rates of incarceration are significantly associated with all three outcomes above and beyond the other covariates, and they contribute meaningfully to the overall model fit. However, these simple correlational findings do not allow us to draw any strong conclusions, especially when it comes to questions of causality. To assess the causal nature of the observed association between incarceration and population health, we proceed to our instrumented panel data analysis.

Panel regressions

Table 2 displays results from three separate two-way fixed-effects instrumental variable regressions in which, once again, the outcome variable is log-transformed and the treatment variable is standardised by subtracting the mean and dividing by its standard deviation. This allows us to interpret the treatment effect as the percentage change in the outcome caused by a standard deviation increase in rates of incarceration. The variables are also residualised with respect to median household income, violent crime, and fraction African Americans or other ethnic minority, such that coefficients are interpreted as net effects. All variance estimators are consistent with serial autocorrelation, heteroskedasticity, and unit clustering. We observe that a standard deviation increase in the treatment variable in one year is estimated to cause a 0.7% drop in life expectancy at birth in the following year (95% CI: 0.6, 0.8; $p < 0.001$). The risk of dying between the ages of 25 and 45 is increased by 7.8% (95% CI: 6.6, 9.0; $p < 0.001$), whereas the risk of death for the 45–65 age group is increased by 8.1% (95% CI: 6.9, 9.3; $p < 0.001$). These are substantively large effect sizes that offer strong support in favour of our principal hypothesis, namely that punitive social policy, operationalised as high rates of incarceration, has a detrimental impact on human welfare. Moreover, the expectation that health in midlife is most heavily affected appears to be vindicated.

We now assess the robustness of our estimated treatment effect to the presence of unmeasured confounding. To save space, we limit the sensitivity analysis to the model of life expectancy, but similar (though even more robust) results apply to the other two outcome measures. Figure 3 visualises variation in the bias factor $\mathcal{B}$, as defined earlier, across a range of possible values of the two sensitivity parameters δ and γ. The X-axis denotes the degree of selection on the unmeasured confounder across the two treatment states (ranging from 0 to 1, with higher values indicating a higher

Table 1: Preliminary cross-sectional regression models.

	Life expectancy	$\mathbb{P}$(death 25–45)	$\mathbb{P}$(death 45–65)
Median household income	1.0%	−12.3%	−9.6%
	(0.1)	(0.6)	(0.4)
Unemployment rate	−0.6%	5.0%	2.9%
	(0.1)	(0.9)	(0.7)
Fraction African Americans	−0.2%	1.9%	1.7%
	(0.1)	(0.5)	(0.5)
Fraction other ethnic minority	−0.3%	2.9%	1.9%
	(0.1)	(0.9)	(0.6)
Violent crime rate	−0.2%	2.2%	2.1%
	(0.05)	(0.4)	(0.4)
Prison incarceration rate	−0.5%	3.9%	4.0%
	(0.1)	(0.4)	(0.4)
R^2	76%	81%	79%
Observations	1,990	1,990	1,990

Notes: The log-transformed outcome variable is life expectancy at birth in the year 2014. Each row is a separate regression of the outcome on the predictor variable listed in the first column. All models are adjusted for state-fixed effects, with the exception of the model with the former slave state indicator. All continuous regressors are standardised by subtracting the variable mean and dividing by its standard deviation. Parameter estimates are interpreted as the percentage change in the outcome variable associated with a standard deviation increase in each predictor. Robust standard errors are shown in parentheses. All parameter estimates are statistically significant at $p < 0.001$.

Table 2: Two-way fixed effects instrumental variable regression models

	Life expectancy	$\mathbb{P}$(death 25–45)	$\mathbb{P}$(death 45–65)
Incarceration rate $(t-1)$	-0.7%	7.8%	8.1%
	(0.06)	(0.6)	(0.6)
Wald test	413	453	461
Observations	57,084	57,084	57,084

Notes: The log-transformed outcome variables are life expectancy at birth in the first column, the probability of death between the ages of 25 and 45 in the second column, and the probability of death between the ages of 45 and 65 in the third column. The incarceration variable, lagged by one year, is instrumented as described in the Data and methods section. The corresponding parameter estimates are interpreted as the percentage change in the outcome variable caused by a standard deviation increase in incarceration rates — net of median household income, violent crime rates, and fraction African Americans. Standard errors consistent with serial autocorrelation, heteroskedasticity, and unit clustering are shown in parentheses below each parameter estimate. All parameter estimates are statistically significant at $p < 0.001$.

prevalence of the confounder in the treatment group, i.e., in counties with higher rates of incarceration), whereas the Y-axis denotes the magnitude of the effect of U on the outcome, above and beyond that of the treatment and the control variables, that would be required to completely eliminate the effect of incarceration on life expectancy at birth. In light of the argument concerning the exogeneity of our chosen instrument, we believe it is plausible that the amount of residual confounding — if there is any — remains moderate. As such, the most likely values of δ would be at the lower end of the X-axis in FIGURE 3. At, say, $\delta = 0.1$, U would have to cause an excess within-county fall in life expectancy such that $\gamma \approx -6.5\%$ to nullify the effect of incarceration. Given the relative magnitudes at stake, this seems highly unlikely. For the sake of argument, assume that the bias factor is 10% of $\hat{\beta}$ — i.e. $\mathcal{B} = 0.07\%$. Then the bias-adjusted effect of high incarceration rates would still be -0.63% (95% CI: -0.53, -0.73; $p < 0.001$), which is equivalent to around half a year of life expectancy. Overall, the sensitivity analysis suggests that an unusual amount of unmeasured confounding would be needed to cast doubt upon our causal estimates.

CROSS-SECTIONAL REGRESSIONS

We proceed to pre-processing the data using coarsened exact matching. We first assign units to treatment and control groups, defined as above versus below mean exposure to incarceration, respectively. We then apply the matching algorithm which results in a pruned data set composed of $N = 1,679$ counties, with 1,064 counties in the control group and 615 counties in the treatment group. Counties are matched on time-averaged versions of the following control variables: violent crime, median household income, high school graduation rates, and the county fraction of African Americans, Hispanics, or other non-White ethnic minorities. 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. Complete matching results are reported in APPENDIX TABLE A3. Using simple ordinary least squares, we then estimate a "between" model,

$$\overline{Y}_i = \overline{T}_i\beta + \xi_i$$

where the (time-averaged) outcome variable $\overline{Y}$ is regressed on $\overline{T}$ alone since covariate balance is obtained through matching. We log-transform $\overline{Y}$ and rescale $\overline{T}$ as done above. TABLE 3 displays the results of three separate regressions, one for each outcome variable. We see that a standard deviation increase in incarceration is associated with a 1.1% decline in life expectancy at birth (95% CI: 1.0, 1.2; $p < 0.001$) and a rise in the two measures of premature mortality risk of 10.6% (95% CI: 9.6, 11.6; $p <$

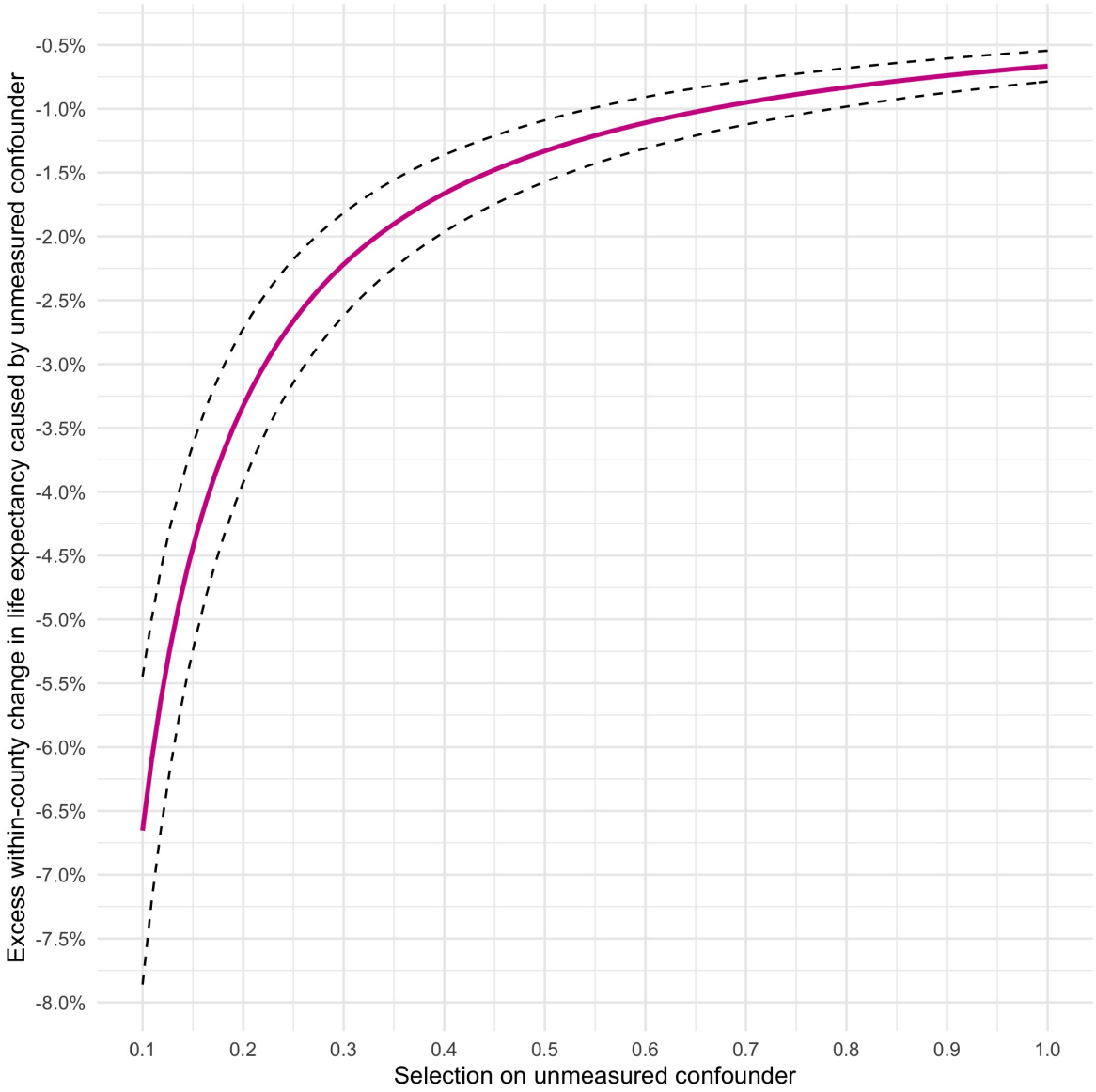

FIGURE 3: Sensitivity analysis plot to assess residual confounding of the estimated effect $\hat{\beta}$ of incarceration on life expectancy at birth as per TABLE 2. Values of δ (X-axis) and γ (Y-axis) that lie on the solid violet line would completely eliminate the estimated effect. The dotted black lines denote the lower and upper bounds of the corresponding 95% confidence interval. Values below the plotted curve would reverse the sign of the estimated effect.

0.001) and 7.9% (95% CI: 7.2, 8.8; $p < 0.001$), respectively. We note, moreover, that incarceration appears to account for no less than one-fifth of all variation in all three outcome variables. We conduct the same sensitivity analysis as before, again for life expectancy at birth, as shown in FIGURE 4. Given the lack of instrumentation, it is hard to surmise the amount of unmeasured confounding that is reasonable to expect in the case of this model. Nonetheless, we see that even at unusually high levels of selection on the unmeasured confounder — e.g., $\delta = 0.8$ — γ would still need to exceed the estimated effect of $\overline{T}$ in order to cancel out $\hat{\beta}$. At more moderate levels of δ, say $\delta = 0.2$, the net impact of U (or $\overline{U}$) on $\overline{Y}$ would have to be nearly -5.5% to nullify $\hat{\beta}$. Overall, this indicates that a non-negligible portion of the reported association is likely to be causal and thus that incarceration is a determinant not only of within-county changes in population health but also of between-county inequalities.

To get a better intuitive sense of what these numbers mean in substantive terms, we predict the conditional expectation of each outcome variable given different levels of $\overline{T}$. We first (re)define control units as those with incarceration rates at one standard deviation below the mean and treatment units as those with incarceration rates at one standard deviation above the mean. We then calculate the corresponding conditional expectations of $\overline{Y}$ across 100,000 simulated values of the stacked column vector $\hat{\psi} = \{\hat{\beta}, \hat{\sigma}^2\}$, derived from the outputs of a model similar to that in TABLE 3, except the outcome variable is no longer log-transformed and the treatment is not rescaled. This vector forms the mean of a multivariate Normal distribution with variance equal to the model covariance matrix $\hat{\mathbb{V}}(\hat{\psi})$. We may thus obtain simulated parameter values

$$\tilde{\psi} \sim \mathcal{N}(\hat{\psi}, \hat{\mathbb{V}}(\hat{\psi}))$$

which are used to predict and plot life expectancy and premature mortality risk by treatment status, whilst simultaneously accounting for estimation uncertainty (see King, Tomz, and Wittenberg [2000] for methodological details). The results of this procedure are shown in FIGURE 5. We see that a shift from low to high incarceration rates is associated with a drop in life expectancy from over 77 years in the control group to just over 75.5 years in the treatment group. When it comes to the probability of dying between the ages of 25 and 45, we see that high rates of incarceration shift this probability from around 2.9% in the control group to around 3.5% in the treatment group. Finally, for the 45–65 age bracket, a shift from low to high incarceration rates is associated with a rise in the probability of death from under 13.5% to nearly 15.5%. Overall, these results suggest that differences in incarceration rates can account for a substantial portion of disparities in health outcomes across US counties.

Table 3: Between-county matched regression models

	Life expectancy	$\mathbb{P}$(death 25–45)	$\mathbb{P}$(death 45–65)
Incarceration rate	−1.1%	10.6%	7.9%
	(0.1)	(0.5)	(0.4)
R^2	18%	21%	20%
Observations	1,679	1,679	1,679

Notes: The log-transformed outcome variables are life expectancy at birth in the second column, the probability of death between the ages of 25 and 45 in the third column, and the probability of death between the ages of 45 and 65 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 Appendix Table A3). All variables are time-averaged over the sample period. Parameter estimates are interpreted as the percentage change in the outcome variable associated with a standard deviation increase in incarceration rates. Standard errors are shown in parentheses below each parameter estimate. All parameter estimates are statistically significant at $p < 0.001$.

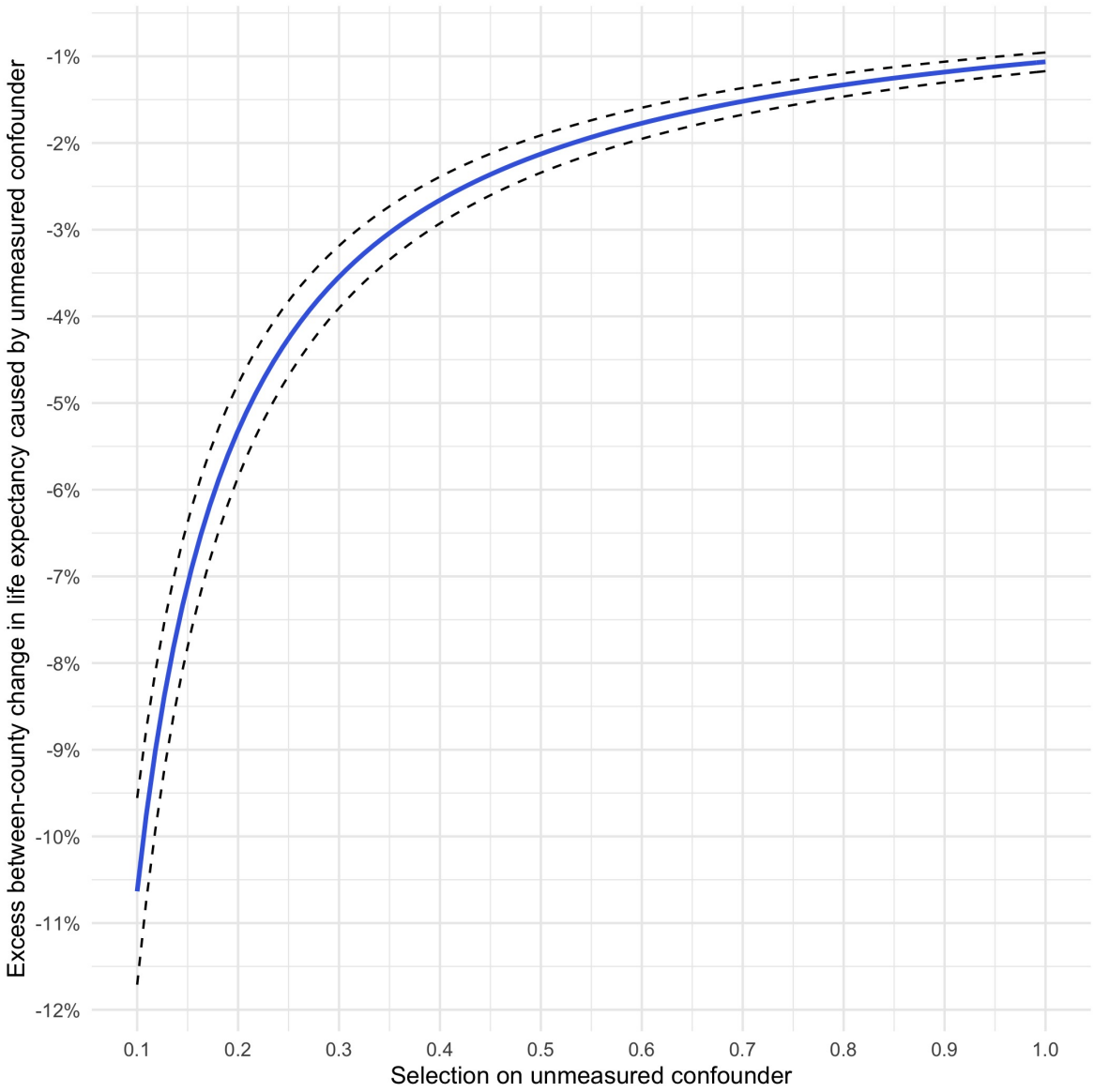

Figure 4: Sensitivity analysis plot to assess unmeasured confounding of the estimated effect $\hat{\beta}$ of incarceration on life expectancy at birth as per Table 3. Values of δ (X-axis) and γ (Y-axis) that lie on the solid blue line would completely eliminate the estimated effect. The dotted black lines denote the lower and upper bounds of the corresponding 95% confidence interval. Values above the plotted curve would reverse the sign of the estimated effect.

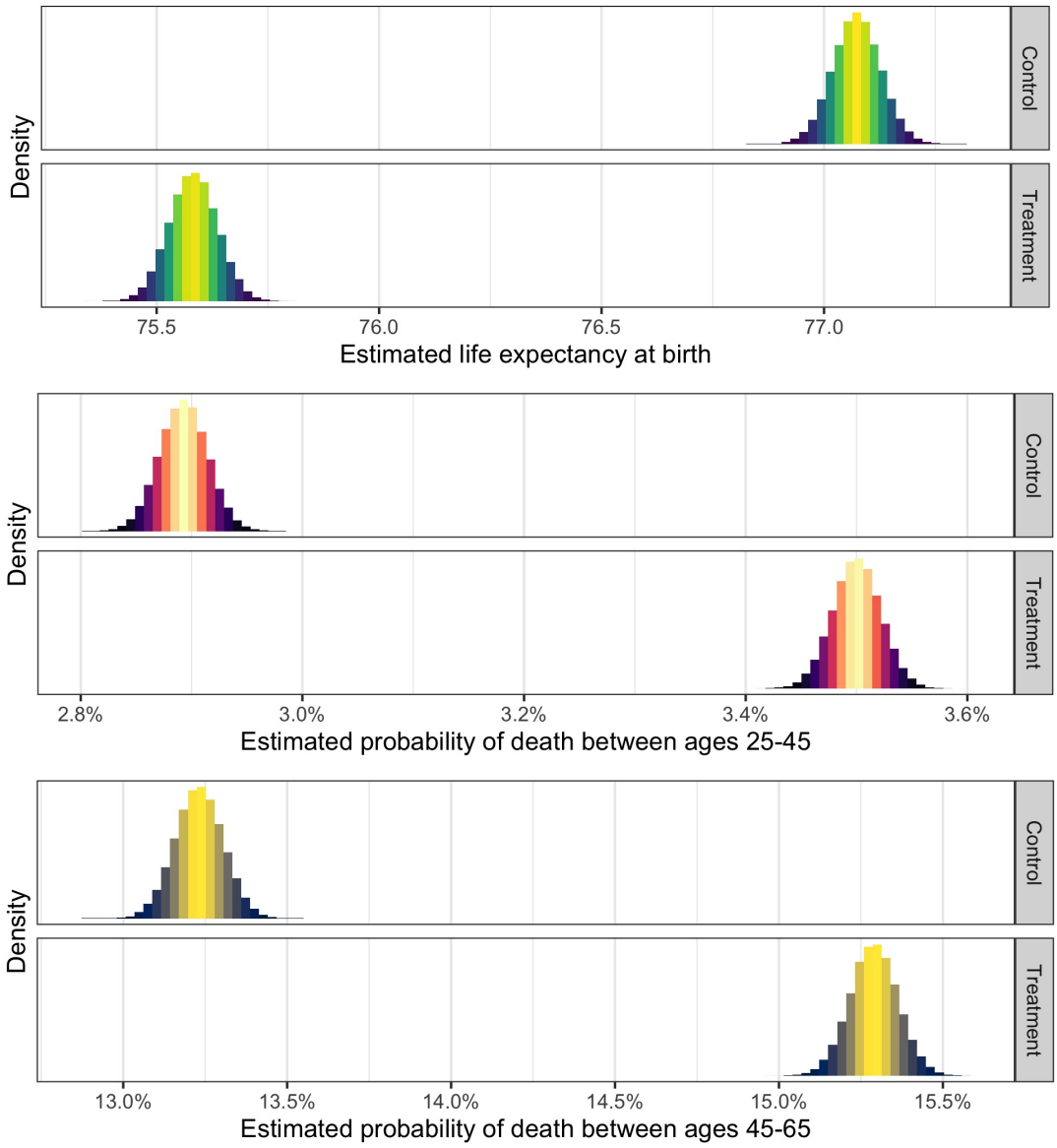

Figure 5: Density plots of expected outcome values conditional on treatment state. In the top panel, the outcome variable is life expectancy at birth, in the middle panel the outcome variable is the probability of death between ages 25–45, and in the bottom panel the outcome variable is the probability of death between ages 45–65. 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 Appendix Table A3). All variables are time-averaged over the sample period. $N = 1,679$.

Alternative specifications

For the sake of comparison and completeness, we also run a series of cross-sectional models in which we assess the robustness of the association between incarceration and health to additional control variables from 2014, as described in Appendix Table A1. To avoid over-specification, we add and remove one control variable at a time. However, as above, we adjust for state-fixed effects in all models. Results are displayed in Appendix Tables A4–A6 for each of our three outcome variables. We note that the estimated coefficient of incarceration remains remarkable stable across all specifications, which further confirms the robustness of the hypothesised relation between rates of imprisonment and population health.

Concluding discussion

Our analysis provides a unified sociological account of substantial geographical variation in life expectancy at birth and premature mortality risk across the United States. We draw on previous scholarship on the distal determinants of vital inequality and generate a framework for the study of population-level (as compared to individual-level) health outcomes. We relate the distributional dynamics of mortality and life expectancy to a range of sociological factors, and we spotlight one of these — punitive social policy — that is rarely invoked in the existing population health literature, despite being a major object of sociological enquiry.

We see our paper as a twofold contribution to the sociological study of incarceration and to the science of vital inequality. We provide an empirically grounded yet theoretically informed account of the causal linkages between high imprisonment rates and three alternative health outcomes, and our estimated effects prove to be substantially large and highly significant. We offer an analytically unified theory of hyper-incarceration as a vector of ill health that spotlights downward social mobility, cumulative disadvantage, and the breakdown of collective efficacy as key components, and we generate causal evidence in support of our principal hypotheses using a stringent two-way fixed effects instrumental variable regression, complemented by various investigations of between-county associations. Our findings provide new empirical insights surrounding the nexus of punishment and population health which may inform future policy-making geared towards criminal justice reform. Insofar as the prison embodies a distinctly American nexus of class and race, our analysis also speaks to ongoing political debates surrounding issues of social inequality, racial justice, and human welfare. In the process, we have introduced a novel instrumentation technique that can readily be employed by other scholars of incarceration, whilst also adding to a nascent research domain that extends the analytic purview of sociology to include vital inequality as a central object of enquiry,

especially by bringing a distinctly macroscopic lens to bear on the analysis of population health.

We acknowledge the limitations of our approach. First of all, despite our systematic conceptualisation of the causal pathways leading from incarceration to vital inequality, we are unable to explicitly test the workings of such mechanisms. Instead, we draw on prior literature to motivate the broader scope of our analysis. Future work should seek to probe the mechanisms of relegation, amplification, and corrosion through the use of large-scale multilevel data sources that follow individuals, neighbourhoods, and broader collectives over extended periods of time. Moreover, our data, despite being of high quality and collected at a relatively fine-grained level of geographical resolution, prevent us from further disaggregating the estimated effects and examining their likely heterogeneity. We also acknowledge that our identification strategy may suffer from unmeasured confounding — especially in our account of between-county inequalities — given that we cannot be certain to have captured purely exogenous treatment variation. However, we provide a simple yet informative sensitivity analysis suggesting that an inordinate amount of confounding must be present to nullify our main results. Our models are highly robust to alternative specifications and provide substantively meaningful estimates of the hypothesised relationships. Our overall conclusion thus remains unambiguous: punitive social policy kills.

Bibliography

Beckfield, Jason. 2018. *Political Sociology and The People's Health.* New York: Oxford University Press.

Beckfield Jason, Clare Bambra, Terje Eikemo, Tim Huijts, Courtney McNamara, Claus Wendt. 2015. "An institutional theory of welfare state effects on the distribution of population health." *Social Theory and Health* 13(3): 227–244.

Berkman, Lisa, Ichiro Kawachi, and Maria M. Glymour, eds. 2014. *Social Epidemiology.* 2nd edition. New York: Oxford University Press.

Binswanger, Ingrid A., Marc F. Stern, Richard A. Deyo, Patrick J. Heagerty, Allen Cheadle, Joann G. Elmore, and Thomas D. Koepsell. 2007. "Release from Prison – A High Risk of Death for Former Inmates." *New England Journal of Medicine* 356(2): 157–165.

Bor, Jacob, Atheendar S. Venkataramani, David R. Williams, and Alexander C. Tsai.

2018. "Police killings and their spillover effects on the mental health of black Americans: a population-based, quasi-experimental study." *The Lancet* 392(10144): 302–310.

Bourdieu, Pierre. 1993. "À propos de la famille comme catégorie réalisée." *Actes de la recherche en sciences sociales* 100: 32–36.

Geremek, Bronislaw. 1994. *Poverty: A History*. Oxford: Wiley-Blackwell.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The Association Between Income and Life Expectancy in the United States, 2001–2014." *Journal of the American Medical Association* 315(16): 1750–1766.

Clear, Todd R. 2007. *Imprisoning Communities: How Mass Incarceration Makes Disadvantaged Neighborhoods Worse*. New York: Oxford University Press.

Cole, Steven W. 2009. "Social Regulation of Human Gene Expression." *Current Directions in Psychological Science* 18(3): 132–137.

Couloute, Lucius. 2018. *Nowhere to Go: Homelessness among formerly incarcerated people*. Prison Policy Initiative. URL: https://www.prisonpolicy.org/reports/housing.html. Accessed 9 September 2020.

DeFina, Robert, and Lance Hannon. 2013. "The Impact of Mass Incarceration on Poverty." *Crime & Delinquency* 59(4): 562–586.

Dwyer-Lindgren, Laura, Amelia Bertozzi-Villa, Rebecca W. Stubbs, Chloe Morozoff, Johan P. Mackenbach, Frank J. van Lenthe, Ali H. Mokdad, and Christopher JL Murray. 2017. "Inequalities in Life Expectancy Among US Counties, 1980 to 2014: Temporal Trends and Key Drivers." *JAMA Internal Medicine* 177(7): 1003–1011.

Evans, Peter. 2009. "Population Health and Development: An Institutional-Cultural Approach to Capability Expansion." Pp. 104–127 in *Successful Societies. How Institutions and Culture Affect Health*. Edited by Peter Hall and Michèle Lamont. New York: Cambridge University Press.

Hall, Peter and Michèle Lamont, eds. 2009. *Successful Societies. How Institutions and*

Culture Affect Health. New York: Cambridge University Press.

Hinds, Oliver, Olivia Lu, and Jacob Kang-Brown. 2018. *Reconstructing How Counties Contribute to State Prisons.* New York: Vera Institute of Justice.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3): 199–236.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2019. "A Theory of Statistical Inference for Matching Methods in Causal Research." *Political Analysis* 27(1): 46–68.
Iacus, Stefano M., Gary King, and Giuseppe Porro. 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, Elizabeth I. and Beth Easterling. 2012. "Understanding Unique Effects of Parental Incarceration on Children: Challenges, Progress, and Recommendations." *Journal of Marriage and Family* 74(2): 342–356.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44: 341–355.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *American Economic Review* 96(3): 863–876.

Kunitz, Stephen J. 2015. *Regional Cultures and Mortality in America.* New York: Cambridge University Press.

Leibbrand, Christine, Erin Carll, Angela Bruns, and Hedwig Lee. 2019. "Barring Progress: The Influence of Paternal Incarceration on Families' Neighborhood Attainment." *Social Science Research* 84: 102321.

Link, Bruce G. and Jo Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior*: 80-94.

Manza, Jeff, and Christopher Uggen. 2008. *Locked Out: Felon Disenfranchisement*

and American Democracy. New York: Oxford University Press.

Massey, Douglas. 2004. "Segregation and Stratification: A Biosocial Perspective." *Du Bois Review: Social Science Research on Race* 1(1): 7–25.

Massoglia, Michael. 2008. "Incarceration as exposure: the prison, infectious disease, and other stress-related illnesses." *Journal of Health and Social Behavior* 49(1): 56–71.

Muller, Christopher and Daniel Schrage. 2019. "The Political Economy of Incarceration in the U.S. South, 1910–1925: Evidence From a Shock to Tenancy and Sharecropping." IRLE Working Paper No. 105-19.

National Research Council. 2014. *The Growth of Incarceration in the United States: Exploring Causes and Consequences*. Edited by Jeremy Travis, Bruce Western, and Steve Redburn. Washington, DC: National Academies Press.

Nosrati, Elias, Michael Ash, Michael Marmot, Martin McKee, and Lawrence P. King. 2018. "The association between income and life expectancy revisited: deindustrialization, incarceration, and the widening health gap." *International Journal of Epidemiology* 47(3): 720–730.

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

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108(5): 937–975.

Patillo, Mary, David Weiman, and Bruce Western, eds. 2004. *Imprisoning America: The Social Effects of Mass Incarceration*. New York: Russell Sage Foundation.

Pettit, Becky and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review* 69(2): 151–169.

Pridemore, William A. 2014. "The mortality penalty of incarceration: evidence from a population-based case-control study." *Journal of Health and Social Behavior* 55(2): 215–33.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rabuy, Bernadette and Daniel Kopf. 2015. *Prisons of Poverty: Uncovering the pre-incarceration incomes of the imprisoned*. Prison Policy Initiative. URL: https://www.prisonpolicy.org/reports/income.html. Accessed 9 September 2020.

Sampson, Robert J. 2003. "The Neighborhood Context of Well-Being." *Perspectives in Biology and Medicine* 46(3): S53–S64.

Sampson, Robert J. and Charles Loeffler. 2010. "Punishment's place: the local concentration of mass incarceration." *Daedalus* 139(3): 20–31.

Seeman, Teresa E., Bruce S. McEwen, John W. Rowe, and Burton H. Singer. 2001. "Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging." *Proceedings of the National Academy of Sciences* 98(8): 4770–4775.

Sen, Amartya. 1999. *Development as Freedom*. Oxford: Oxford University Press.

Shannon, Sarah and Christopher Uggen. 2012. "Incarceration as a Political Institution". Pp. 214–225 in *The Wiley-Blackwell Companion to Political Sociology*. Edited by Edwin Amenta, Kate Nash, and Alan Scott. Oxford: Wiley-Blackwell.

Slavich, George M., and Steven W. Cole. 2013. "The Emerging Field of Human Social Genomics." *Clinical Psychological Science* 1(3): 331–348.

Therborn, Göran. 2013. *The Killing Fields of Inequality*. Cambridge: Polity Press.

Thompson, Heather A. 2010. "Why Mass Incarceration Matters: Rethinking Crisis, Decline, and Transformation in Postwar American History." *Journal of American History* 97(3): 703–734.

Turney, Kristin. 2018. "Adverse Childhood Experiences among Children of Incarcerated Parents." *Children & Youth Services Review* 89: 218–225.

Turney, Kristin. 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, Tyler J. 2015. *Explanation in Causal Inference. Methods for Mediation and Interaction*. New York: Oxford University Press.

Vanderweele, Tyler J. and Onyebuchi A. Arah. 2011. "Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders." *Epidemiology* 22(1): 42–52.

Wacquant, Loïc. 2010. "Class, Race & Hyperincarceration In Revanchist America." *Daedalus* 139(3): 74–90.

Wacquant, Loïc. 2009. *Punishing the Poor: The Neoliberal Government of Social Insecurity*. Durham: Duke University Press.

Wakefield, Sara and Christopher Wildeman. 2013. *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality*. New York: Oxford University Press.

Weidner, Robert R. and Jennifer Schultz. 2019. "Examining the relationship between US incarceration rates and population health at the county level." *SSM Population Health* 9: 100466.

Western, Bruce. 2018. *Homeward: Life in the Year After Prison*. New York: Russell Sage Foundation.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.

Wildeman, Christopher. 2009. "Parental imprisonment, the prison boom, and the concentration of childhood disadvantage." *Demography* 46(2): 265–280.

Wildeman, Christopher and Christopher Muller. 2012. "Mass imprisonment and inequality in health and family life." *Annual Review of Law and Social Science* 8: 11–30.

Wildeman, Christopher and Bruce Western. 2010. "Incarceration in Fragile Families." *The Future of Children* 20(2): 157–177.

Wildeman, Christopher, Alyssa Goldman, and Hedwig Lee. 2019. "Health

Consequences of Family Member Incarceration for Adults in the Household." *Public Health Reports* 134(1): 15S–21S.

Wildeman, Christopher, Sara Wakefield, and Kristin Turney. 2013. "Misidentifying the effects of parental incarceration? A comment on Johnson and Easterling (2012)." *Journal of Marriage and Family* 75(1): 252–258.

Wolfe, Barbara, William Evans, and Teresa E. Seeman, eds. 2012. *The Biological Consequences of Socioeconomic Inequalities*. New York: Russell Sage Foundation.

Zlodre, Jakov and Seena Fazel. 2012. "All-Cause and External Mortality in Released Prisoners: Systematic Review and Meta-Analysis." *American Journal of Public Health* 102(12): e67–e75.

Appendix

Table A5: Cross-sectional regression model of premature mortality risk ages 25–45.

Table A6: Cross-sectional regression model of premature mortality risk ages 45–65.

Table A1: Descriptive statistics: cross-sectional data from 2014.

Statistic	N	Mean	St. Dev.	Min	Max
Life expectancy at birth	2,830	77.7	2.4	67.6	86.8
Probability of death ages 25–45 (%)	2,830	3.1	0.9	1.2	8.3
Probability of death ages 45–65 (%)	2,830	13.3	3.1	4.8	32.6
Incarceration rate per 100,000 population	2,107	313	165	21.9	732
Violent crime rate per 100,000 population	2,572	237	194	0.0	2,569
Median household income ($)	2,830	46,515	11,427	22,640	108,477
Unemployment rate	2,830	6.3	2.2	1.2	24.0
Labour force participation rate	2,830	0.6	0.1	0.3	0.9
Poverty rate	2,830	0.1	0.1	0.02	0.5
Absolute income mobility	2,673	47.5	6.1	25.1	68.3
Income inequality	2,830	0.4	0.1	0.2	0.8
High school graduation rate	2,830	0.9	0.1	0.5	1.0
Fraction African Americans	2,830	0.1	0.1	0.002	0.9
Fraction Hispanics	2,830	0.1	0.1	0.004	1.0
Fraction other ethnic minority	2,830	0.04	0.1	0.001	0.9
Racial segregation	2,830	0.1	0.1	0.0	0.7
Percentage uninsured	2,830	18.4	5.5	3.6	41.4

Notes: All variables, listed in the first column, are measured at the county level in the year 2014. The second column lists the number of observations. The unemployment rate variable is taken from the US Bureau of Labor Statistics. All other variables are from the Opportunity Insights database (n.d.). Absolute income mobility is measured as the county-level percentage of children who earn more than their parents in 2014. Income inequality is measured as the county-level Gini index within the bottom 99% of the income distribution. Racial segregation is measured as residential segregation by race. The final variable measures the percentage of the county population without health insurance. The sources of all remaining variables are listed in A2 below.

Figure A1: Observed and imputed values of annual aggregate per capita correctional expenditure.

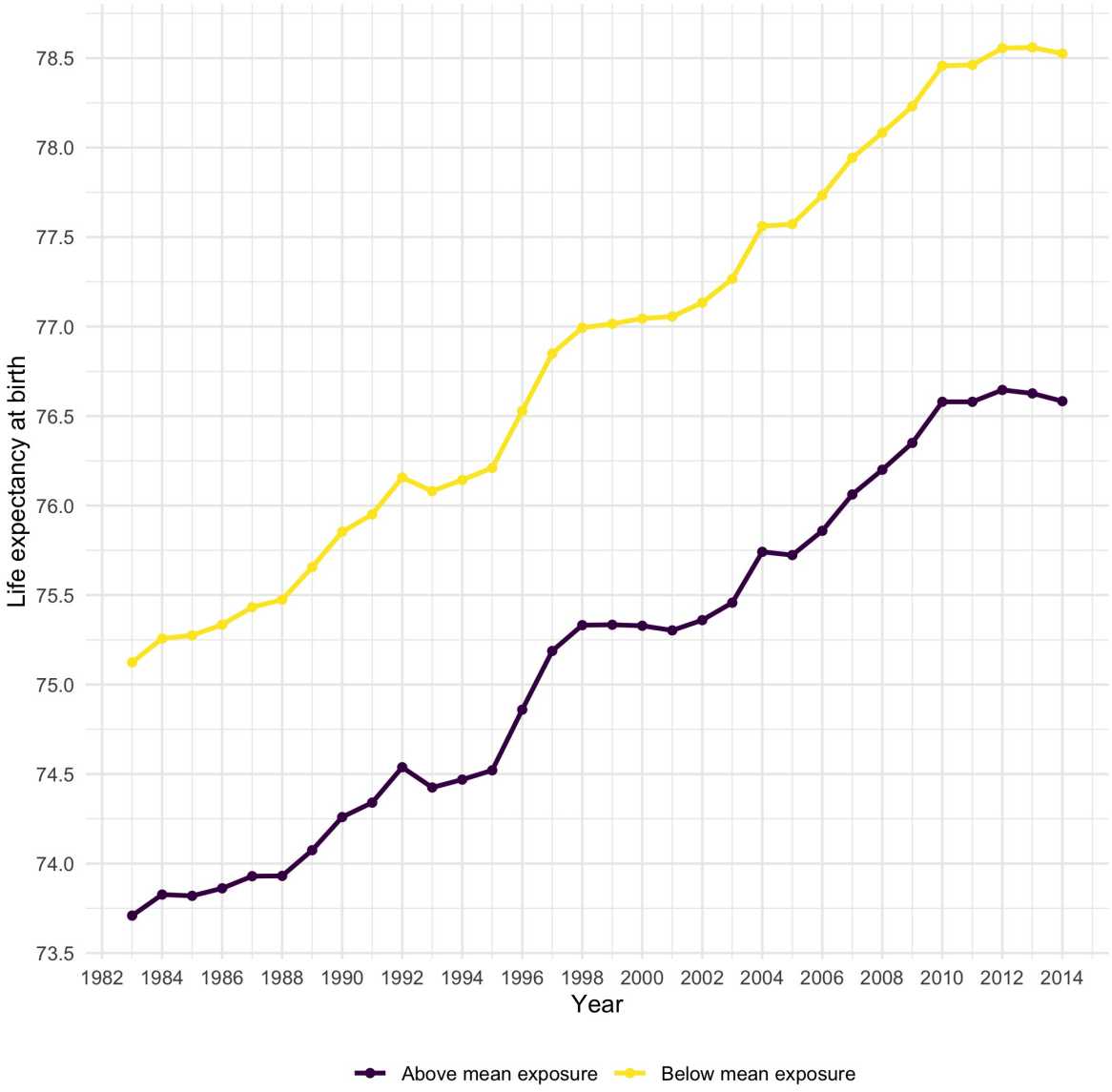

Figure A2: Visual inspection of parallel trends assumption for counties with above versus below average exposure to incarceration. Outcome: life expectancy at birth.

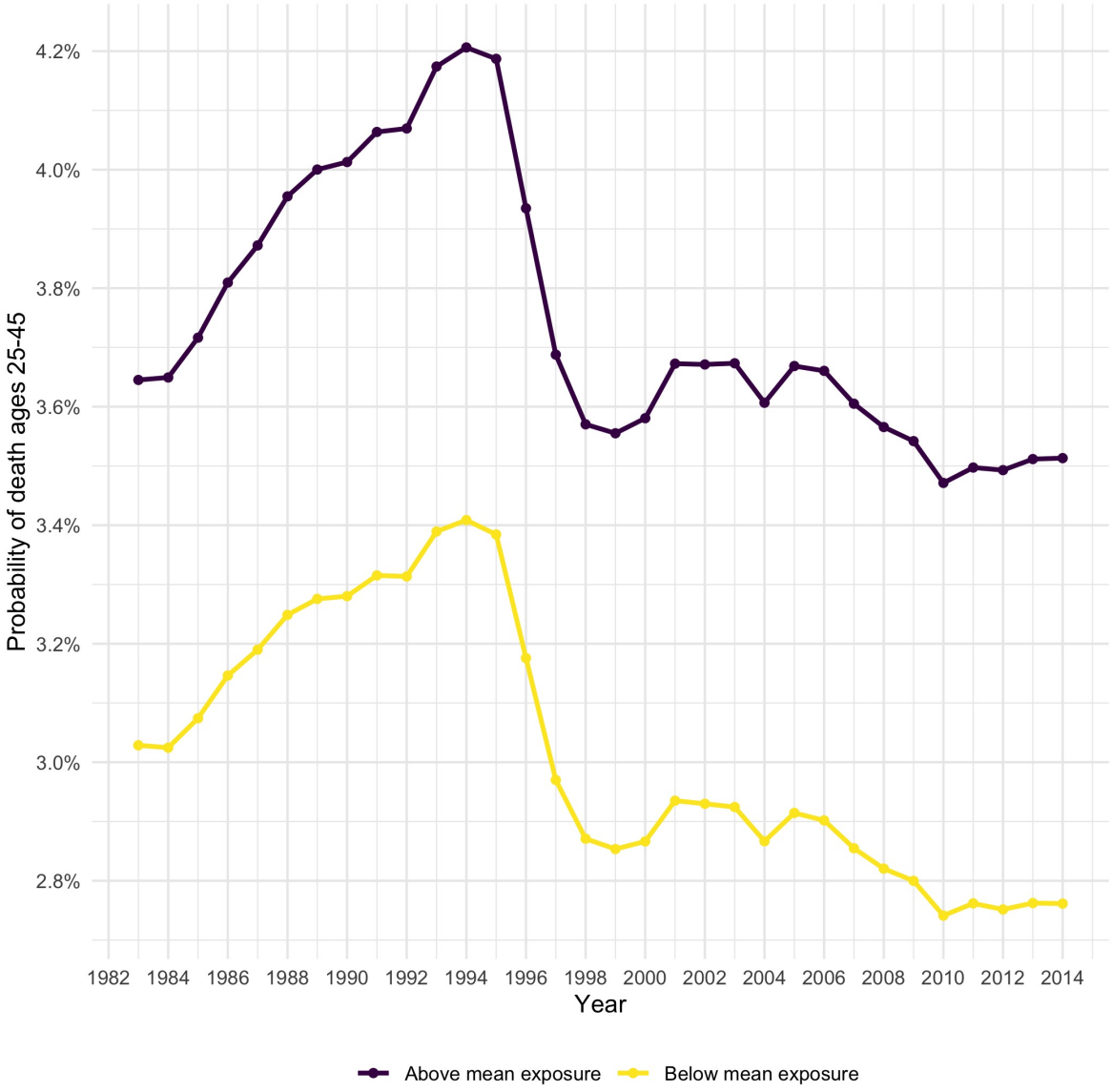

Figure A3: Visual inspection of parallel trends assumption for counties with above versus below average exposure to incarceration. Outcome: probability of death between ages 25–45.

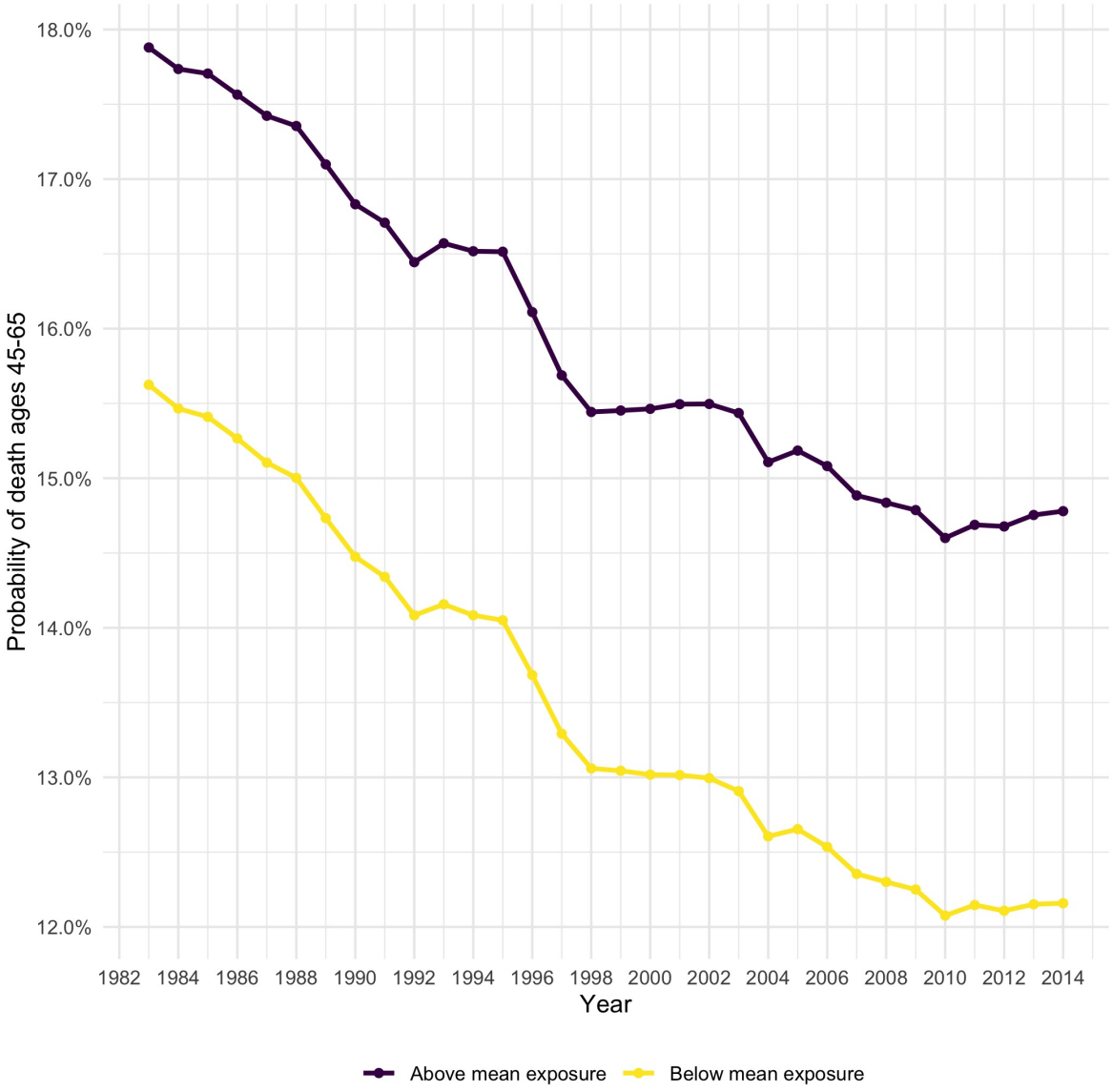

Figure A4: Visual inspection of parallel trends assumption for counties with above versus below average exposure to incarceration. Outcome: probability of death between ages 45–65.

Table A2: Descriptive statistics: panel data from 1983–2014.

Statistic	N	Mean	St. Dev.	Min	Max
Life expectancy at birth	59,153	76.2	2.3	66.2	86.8
Probability of death ages 25–45 (%)	59,153	3.3	0.9	1.2	10.2
Probability of death ages 45–65 (%)	59,153	14.5	3.0	4.8	32.6
Incarceration rate per 100,000 population	59,153	250.1	158.6	6.0	732.1
Violent crime rate per 100,000 population	59,153	281	260	0.0	3,894
Median household income ($)	59,153	47,533	11,853	17,583	125,705
Fraction African Americans	59,153	0.1	0.1	0.0	0.9
Fraction other ethnic minority	59,153	0.02	0.05	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 — life expectancy at birth and the two measures of premature mortality risk — are taken from the Institute for Health Metrics and Evaluation US Health Map database (2017). The incarceration rate is per 100,000 population aged 16–64 and is constructed by the Vera Institute of Justice (Hinds et al. 2018). The measure of violent crime is extracted from the Federal Bureau of Investigation's Uniform Crime Reporting Program. The measure of the county proportion of African Americans is taken from the US Census Bureau.

Table A3: Mean covariate balance obtained from coarsened exact matching.

	Treated	Control	Difference	Improvement
Distance	0.47	0.46	0.01	96%
Violent crime rate	223	212	11	93%
Median household income ($)	$4,4095$	$4,4171$	-76	99%
High school graduation rate	0.73	0.73	-0.0006	99%
Fraction African Americans	0.08	0.08	0.001	99%
Fraction Hispanics	0.04	0.03	0.003	93%
Fraction other ethnicity	0.014	0.014	0.0005	40%

Notes: Results from applying a coarsened exact matching algorithm to time-averaged versions of the variables presented in the Data and methods section. The algorithm produces a matched data set using a dichotomous treatment indicator of above versus below mean exposure to incarceration. The table displays matched covariate levels in the control and treatment groups, as well as the percentage balance improvement as a result of matching.

Table A4: Cross-sectional regression model of life expectancy at birth.

Control variable	Control coefficient	Incarceration coefficient
Unemployment rate	−1.30%***	−0.74%***
	(0.09)	(0.05)
Poverty rate	−1.29%***	−0.74%***
	(0.08)	(0.05)
Absolute income mobility	0.66%***	−0.82%***
	(0.08)	(0.06)
Income inequality	0.09%	−0.89%***
	(0.06)	(0.06)
Racial segregation	−0.02%	−0.88%***
	(0.06)	(0.06)
Percentage uninsured	−0.51%***	−0.84%***
	(0.10)	(0.06)

Notes: The log-transformed outcome variable is life expectancy at birth in the year 2014. Each row is a separate regression wherein the association between incarceration and life expectancy is adjusted for the control variable listed in the first column. All models are also adjusted for state-fixed effects. All regressors are standardised by subtracting the mean and dividing by the standard deviation. Parameter estimates are interpreted as the percentage change in the outcome variable associated with a standard deviation increase in each predictor. Robust standard errors are shown in parentheses below each parameter estimate. Statistical significance levels: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table A5: Cross-sectional regression model of premature mortality risk ages 25–45.

Control variable	Control coefficient	Incarceration coefficient
Unemployment rate	13.8%***	6.6%***
	(0.9)	(0.5)
Poverty rate	13.9%***	6.7%***
	(0.7)	(0.5)
Absolute income mobility	−6.9%***	7.4%***
	(0.8)	(0.6)
Income inequality	1.4%*	7.9%***
	(0.6)	(0.6)
Racial segregation	1.2%*	8.0%***
	(0.6)	(0.6)
Percentage uninsured	10.9%***	7.4%***
	(1.0)	(0.5)

Notes: The log-transformed outcome variable is the probability of death between the ages of 25 and 45 in the year 2014. Each row is a separate regression wherein the association between incarceration and premature mortality risk is adjusted for the control variable listed in the first column. All models are also adjusted for state-fixed effects. All regressors are standardised by subtracting the mean and dividing by the standard deviation. Parameter estimates are interpreted as the percentage change in the outcome variable associated with a standard deviation increase in each predictor. Robust standard errors are shown in parentheses below each parameter estimate. Statistical significance levels: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table A6: Cross-sectional regression model of premature mortality risk ages 45–65.

Control variable	Control coefficient	Incarceration coefficient
Unemployment rate	9.7%***	6.3%***
	(0.7)	(0.4)
Poverty rate	9.9%***	6.3%***
	(0.5)	(0.4)
Absolute income mobility	−5.7%***	6.8%***
	(0.6)	(0.5)
Income inequality	−0.07%	7.4%***
	(0.5)	(0.5)
Racial segregation	0.6%	7.4%***
	(0.4)	(0.5)
Percentage uninsured	5.5%***	6.9%***
	(0.8)	(0.4)

Notes: The log-transformed outcome variable is the probability of death between the ages of 45 and 65 in the year 2014. Each row is a separate regression wherein the association between incarceration and premature mortality risk is adjusted for the control variable listed in the first column. All models are also adjusted for state-fixed effects. All regressors are standardised by subtracting the mean and dividing by the standard deviation. Parameter estimates are interpreted as the percentage change in the outcome variable associated with a standard deviation increase in each predictor. Robust standard errors are shown in parentheses below each parameter estimate. Statistical significance levels: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.