Supporting Information

Kirk 10.1073/pnas.1501987112

SI Text

This study draws upon prison release and reincarceration data from the DPS&C to estimate the effect of the concentration of parolees in geographic space on subsequent reincarceration rates. Data on the residential location and reincarceration of Louisiana parolees may be requested from the DPS&C Office of Information Services. I also draw upon ZIP-code and parish-level data from the Louisiana Department of Labor, Geolytics, the US Postal Service (USPS), and the Supreme Court of Louisiana. These data are used to control for observed differences in ZIP code and parish conditions across time and space, to isolate the specific effect of parolee concentration on reincarceration rates. Contextual variables at the ZIP-code level include concentrated disadvantage, the proportion of renters, the housing supply (i.e., the ratio of dwellings to population size), the number of providers of prisoner-reentry resources and social services (e.g., counseling, education, employment, and health resources), the average time served in prison by ex-prisoners released to the ZIP code, and the prior recidivism rate (from 2003). Control variables at the parish level include average weekly wage and the average caseload per judge in the parish criminal court.

ZIP Code Variables. Concentrated disadvantage is a scale of economic disadvantage, created via principal components analysis, based on the following time-varying indicators from 2005 and 2006 Geolytics sociodemographic estimates: median income and the percentages of adults (aged 25+) who have a high school education or less and of female-headed families with children.

Proportion of renters is an indicator of the proportion of households in a ZIP code that reside in rental properties. This variable is taken from 2005 and 2006 Geolytics estimates.

Ratio of dwellings to population is a measure of the count of inhabitable residential and commercial addresses in each ZIP code divided by the population count in the ZIP code across two time periods—the fourth calendar quarter of 2005 and the fourth quarter of 2006. Although it is not possible with the available data to distinguish the number of residential addresses from the number of commercial addresses, this measure is nevertheless designed to serve as a proxy for the supply of housing in a ZIP code. This address data are made available, by quarter, from the USPS and is distributed by the US Department of Housing and Urban Development (www.huduser.org/portal/datasets/usps.html).

These USPS data are based on the universe of all addresses with mail delivery. In addition to counting the total number of addresses per ZIP code, the USPS defines vacant properties as addresses that were actively receiving mail in the past but which are no longer occupied. An address is considered vacant if the residence, office, or building has not been occupied in at least 90 d. The USPS categorizes addresses that are blighted or have been demolished as "no stat" addresses. Additionally, buildings or residences still under construction would also receive this designation. No-stat addresses are not actively receiving mail. The measure used in the analysis represents the difference between the total number of addresses and the number of no-stat addresses in a ZIP code, divided by the population size of the ZIP code.

"Resource providers" is an indicator of the extent of social services available in a ZIP code for formerly incarcerated individuals. For the purposes of connecting parolees to social service providers, the Louisiana DPS&C maintains a list of both government and nonprofit service providers. This list contains a total of 2,420 nonduplicate addresses, of which I was able to geocode 2,258 (93.3%). The range of services provided include counseling, education, employment and job training, driving instruction, substance abuse treatment, child care, food, shelter, or housing services, medical services, and transportation. The inclusion of this measure of prisoner-reentry resource providers in statistical models is designed to account for ZIP-code differences in the extent of services available to ex-prisoners, given that the availability of services likely affects recidivism rates.

Average time served in prison is an indicator derived from DPS&C data of the average time served in prison for the most recent incarceration for prisoners released to each ZIP code in a given prison release cohort. Inclusion of this measure in statistical models is designed to account for geographic variation in the risk of recidivism. In other words, some ZIP codes may have greater recidivism rates because those ZIP codes have higher risk offenders (i.e., the nonrandom sorting of different types of parolees). This measure of the average number of years served in prison by those prisoners released to a ZIP code proxies for this variation in risk.

Prior recidivism rate (from 2003) is an indicator derived from DPS&C data of the recidivism rate among parolees released to each ZIP code in 2003. Controlling for the prior recidivism rate helps account for unmeasured factors that predict geographic variation in recidivism.

Parish-Level Variables. Average weekly wage is a measure of the average weekly wage (in 2000-adjusted dollars) in each parish during the two separate periods: September to December 2005 and September to December 2006. Data are from the Louisiana Department of Labor.

Judge caseload is a measure of average caseload across parish judges, and it is derived from information contained within Louisiana Supreme Court annual reports (www.lasc.org/press_room/annual_reports/). To correspond to the two time periods, I use the average number of cases per judge in 2005 and 2006, respectively, in each parish. I include a control for judge caseload given that such caseloads likely influence whether convicted offenders are sentenced to a term of incarceration (i.e., my dependent variable) or to some other sanction, such as probation.

Statistical Methods. As described in the main text, a DID estimation strategy was used to estimate the effect of the concentration of prisoner reentry on reincarceration rates. A key assumption of the DID approach is that the change in the reincarceration rate would be the same across treatment and control neighborhoods if both experienced the same change over time in the concentration of parolees. In the absence of any kind of change in the concentration of parolees, the temporal change in reincarceration would be the same for treatment and control groups. Satisfying this parallel-trends assumption becomes problematic when some factor besides the treatment affects the treatment group but not the control group. In the main text I restricted the analysis to ZIP codes located outside of New Orleans to make a more plausible case that I have satisfied the parallel-trends assumption that is core to the DID framework. It is possible that Hurricane Katrina affected the New Orleans area in unmeasured ways, such that there were additional factors affecting treatment neighborhoods in New Orleans that did not affect control neighborhoods elsewhere in the state.

Though there are sound methodological reasons for excluding New Orleans ZIP codes from the main analysis, for the sake of

thoroughness, I also estimated models including these ZIP codes. These results are found in Table S1. Consistent with the results presented in Table 1, the significant positive interaction between parolee concentration and the time period ($\delta = 0.116$) indicates that ZIP code reincarceration rates are a positive function of the extent of the concentration of parolees. Ultimately, the inferences are not sensitive to whether I include or exclude ZIP codes from New Orleans in my analysis.

Testing for Spatial Autocorrelation. Given the possibility that recidivism rates in surrounding ZIP codes—or similarly, that predictors of recidivism in surrounding ZIP codes such as the concentration of parolees—influence the rate of recidivism in focal neighborhoods, I estimated a supplementary analysis incorporating the spatial lag of recidivism rates into my models. I created a spatially lagged recidivism measure using a queen-based contiguity spatial weight matrix. The queen criterion designates neighborhoods as contiguous with a focal neighborhood if they share a common border or vertex. In contrast to the main analysis, which used a negative binomial regression, for this analysis with a spatial lag I used an ordinary least-squares (OLS) regression. I used an OLS model given the challenges of estimating a count model with endogenous spatial lags. For this model, I measured the dependent variable as the proportion of new parolees who were reincarcerated within 1 y. Model results are found in Table S2.

Consistent with the results presented in Table 1, I find that the concentration of parolees in a neighborhood is positively predictive of reincarceration rates. I do not find evidence that recidivism rates in focal neighborhoods are influenced by recidivism rates in surrounding neighborhoods, whether in the reduced (model 1) or full (model 2) models. This finding does not necessarily mean that there is no spatial dependence with recidivism. Measurement of spatial association depends upon the scale of the geographic unit of analysis. Hence, whereas the rate of recidivism in a ZIP code does not appear to depend upon the rate of recidivism in contiguous ZIP codes, it might be the case that there is spatial dependency at a smaller unit of analysis such as the street block.

Table S1. Difference-in-differences estimates of Louisiana Reincarceration, including New Orleans

Variables	Coefficient	Robust SE
Intercept	3.373	(4.670)
Concentration of parolees	−0.032	(0.045)
Year 2006 (vs. 2005)	0.227	(0.059)***
Concentration of parolees × year 2006	0.116	(0.044)**
Concentrated disadvantage	0.042	(0.022)*
Proportion renters	0.291	(0.360)
Average weekly wage	−0.117	(0.034)***
Ratio dwellings to population	−0.003	(0.157)
Nearby service providers	−0.013	(0.134)
Judge caseloads	−0.002	(0.006)
Average time served	−0.082	(0.042)*
Prior recidivism rate (2003)	0.237	(0.187)

The dependent variable is the 1-y reincarceration rate. The coefficients and SEs for average weekly wage, nearby service providers, and judge caseloads are multiplied by 100.
*$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$ (two-tailed test).

Table S2. OLS estimates with a spatially lagged dependent variable

	Model 1		Model 2	
Variables	Coefficient	Robust SE	Coefficient	Robust SE
Intercept	0.175	(0.015)***	−0.168	(1.467)
Spatial lag of recidivism	0.018	(0.035)	−0.002	(0.031)
Concentration of parolees	−0.019	(0.020)	−0.028	(0.018)
Year 2006 (vs. 2005)	0.041	(0.021)$^{+}$	0.040	(0.021)$^{+}$
Concentration of parolees × year 2006	0.049	(0.028)$^{+}$	0.047	(0.028)$^{+}$
Concentrated disadvantage			−0.001	(0.011)
Proportion renters			−0.039	(0.111)
Average weekly wage			−0.026	(0.013)$^{+}$
Ratio dwellings to population			−0.068	(0.084)
Nearby service providers			0.117	(0.058)*
Judge caseloads			−0.001	(0.003)
Average time served			−0.015	(0.010)
Prior recidivism rate (2003)			−0.095	(0.052)$^{+}$

The dependent variable is the 1-y reincarceration rate, constructed by dividing the number of recidivists from a release cohort in a given ZIP code by the number of parolees in the release cohort. The coefficients and SEs for average weekly wage, nearby service providers, and judge caseloads are multiplied by 100.
$^{+}P \leq 0.10$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$ (two-tailed test).