ONLINE APPENDIX: Dynamics of Deterrence: A Macroeconomic Perspective on Punitive Justice Policy

1 Datasets and Variable Construction

The empirical counterpart to incarceration in the model is incarceration in a state or federal prison. This excludes incarceration in local jails which are usually for misdemeanors, sentences shorter than one year, or for detainment without a conviction. By contrast, the convictions resulting in state or federal imprisonment are typically serious felonies with a sentence of a year or more. We treat the crime data similarly and focus on serious crimes that would likely be charged as felonies. When measuring admissions we exclude, where possible, admission due to a parole or probation violation; a transfer; a return escapee; or those incarcerated in prisons without a conviction (often immigration violations which have an increasing trend of their own).

1.1 National Corrections Reporting Program (NCRP).

Prison admissions, stocks, and sentence lengths are computed using a panel from the National Corrections Reporting Program (NCRP, accessed through ICPSR United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics (2013)). The NCRP is a restricted access offender-level administrative dataset set maintained by the Bureau of Justice Statistics. Detailed tabulations on prison admission, release, parolee, and prison stock data are reported to the Department of Justice by individual states. We clean the NCRP data by the following criteria. First, we restrict the sample to states meeting internal consistency checks and data completeness requirements from the audit study of Neal and Rick (2014).[1] Next, we include only states that have a consistent time series from 1985-2016; and consistently report the category of offense and whether it was a new court commit.[2] This leaves us with a sample consisting of 17 states: California, Florida, Georgia, Illinois, Kentucky, Michigan, Minnesota, Missouri, New Hampshire, New Jersey, New York, North Dakota, Ohio, South Carolina, Utah, Virginia, and Wisconsin. Trends in this sample of states follow similar broad trends as estimated by the Bureau of Justice Statistics for the nation (see Figure 9) and account for 42-60% of total national admissions over the time period.

[1] We depart from Neal and Rick (2014) in that we do not impute missing data on demographics and offense.

[2] Researchers interested in more recent data will find improved reporting with 38 states providing some kind of data after 2000.

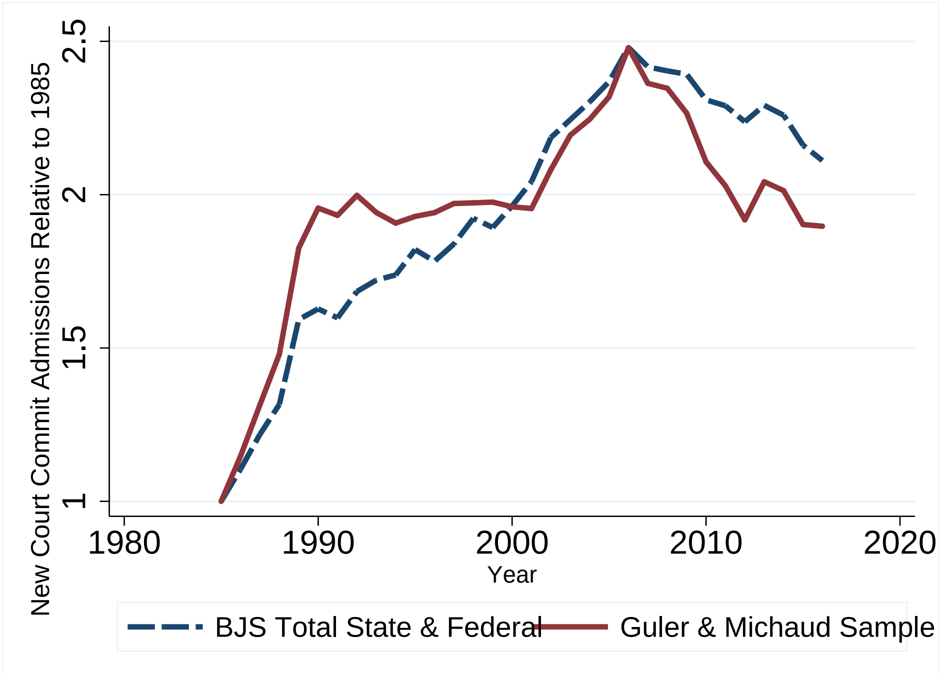


Figure 1: Total admissions count from new court commitments relative to total admissions counts from new court commitments in 1985 from each the National Prisoner Statistics (Bureau of Justice Statistics) and from our 17 state prison sample of NCRP micro data.

Gender and birth date of each offender are fully reported across states with few missing cases. We restrict the sample to adult men only. Some states, however, have internally inconsistent age demographics in stocks and admissions. As a result, we drop Florida, Ohio, and Virginia when comparing outcomes by age or cohort. Education fields are frequently missing and therefore we do not restrict our incarceration sample by education.

The NCRP data are available starting in 1983. We drop 1983 and 1984 because, even within our restricted sample, we find elevated occurrences of missing or incomplete data for our variables of interest in these years. National estimates on male admissions and the male prison population for 1978 onward are published by the Bureau of Justice Statistics constructed from National Prison Statistics data.[3] These estimates do not, however, striate by offense type which is necessary for our analysis. We impute the time-series of admissions and stocks striated by type of crime for the years 1978 through 1984 under the assumption that the composition of offenses are unchanged between 1978 and 1985. In other words, we apply the shares of violent, property, and other crimes that we calculate from our 1985 sample to the national estimates from 1978-1984 to impute the years missing from the NCRP micro data. We remove years of data in the case that a State is missing data, has admissions

[3]These data are easily accessed using the Correction Statistical Analysis Tool (CSAT) at https://csat.bjs.ojp.gov .

under 50, or has a change in stock by over 50%. This amounts to 8 observations. We impute the data for each of these years by linearly interpolating from the two adjoining years.

We append data from US Decennial Census and Current Population Survey (Census, accessed through IPUMS Ruggles (2004)) to calculate admission and prison population rates relative to a reference population. We define the reference population as males without a college degree age 18-35. We choose this reference population to control for the aging of the population driven by the baby-boomers. Crime is concentrated among men under 40. If we do not restrict the sample in this way, incarceration rates in the later years would be under estimated, driven by increases in the population of baby boomer old men who account for a tiny fraction of prisoners.

Offenses are classified into property, violent, and other according to given NCRP code categories. Examples of violent crime include: murder, kidnapping, rape, robbery (armed and unarmed), assault, and blackmail. Examples of property crimes include: burglary, arson, theft, and destruction of property. Crimes categorized as "other" include: drug trafficking, weapons offenses, obstruction, fraud, and tax and revenue violations. We categorize an individual as having one or more of each type of offense according to two metrics: (i) category of offense with the longest sentence only; and (ii) category of any offense with a sentence. The latter implies that a single admission can fall under multiple categories of crime if there are multiple offenses spanning more than one category. Note that the NCRP data only lists up to 3 offenses, prioritizing more serious offenses. This means that offenses such as trespassing or possession of drugs are likely to be omitted even if they add to total sentence length. Our main analysis in the text classifies a prisoner as having a property offense if they are charged with any property offense, and similarly for violent, but our discussion of crime specialization addresses most serious offense as well.

Prison duration is calculated to provide a stationary distribution in the initial calibration that replicates the stocks in the data given the admissions. The model imposes a constant hazard rate of prison exits, or a geometric distribution of prison lengths. The computed duration consistent with all of these features gives a median spell of 30 months in prison for violent crimes and 10 months for property crimes. This duration is similar to Perkins (1992) who calculate a median spell of 27 months in state prisons for violent crimes and 12 months for property crimes in an analysis of BJS administrative survey data. We hold the median time served to be constant throughout our time-series analysis. This choice is consistent with the empirical findings of both Neal and Rick (2014) and Raphael and Stoll (2009).

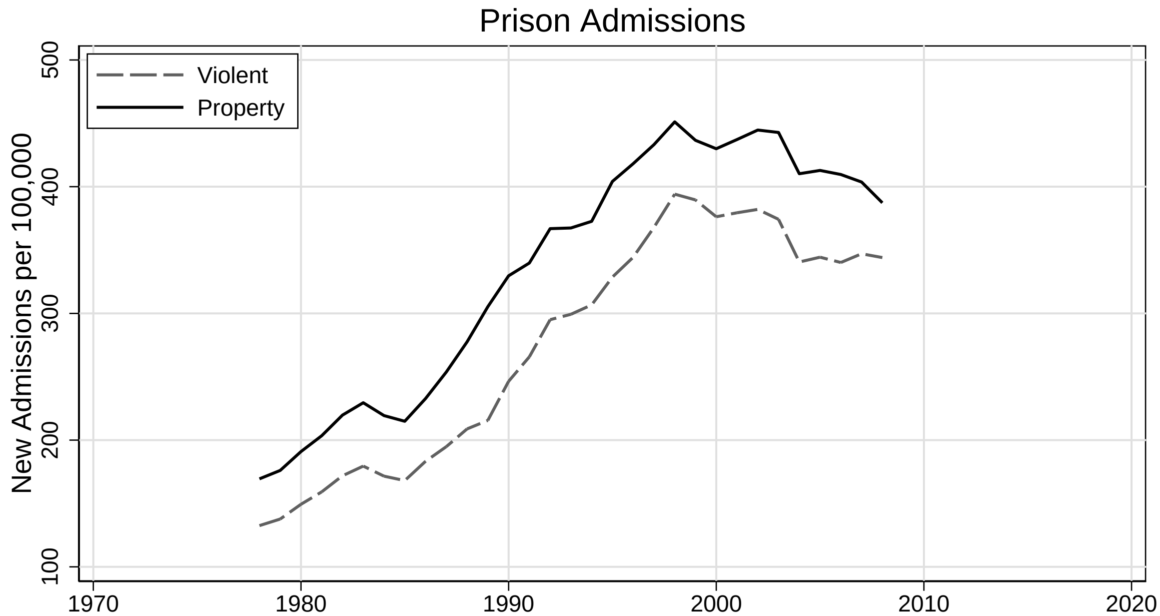


Figure 2: Admission rates for new convictions per 100,000 adults estimated from NCRP and Census data.

1.2 National Crime Victimization Survey (NCVS)

Crime volumes are calculated using the National Crime Victimization Survey 1979-2020.[4] The NCVS is a nationally representative survey of respondents age 12 and older. The survey identifies respondents who have been a victim of crime within the past six months and asks further questions about the nature of that crime. We use incident-level weights to estimate national crime incidence by type of crime: property, violent, and other; taking only offenses that are likely to be charged as felonies. We take this as a measure of total crime of each type committed in the U.S. annually. We combine this with aggregated BJS data on prison admissions from the National Prison Statistics program and define the probability of prison admission conditional on crime by dividing the NCVS crime measure by the BJS prison admission measure.

Caution must be used when comparing our estimates of the probability of prison conditional on crime to other similar concepts in the literature that instead focus on arrest or any type of incarceration. For example, İmrohoroğlu, Merlo, and Rupert (2004), considered property crime alone in their calibration and set their probability of apprehension to equal the clearance rate for these crimes and therefore find higher values than we do because they

[4]The NCVS includes all offenses reported to the survey whereas the Uniform Crime Report includes only crimes known by authorities. Further, we find a high incidence of missing offense codes in the UCR for our time period. For these reasons we use NCVS instead of UCR.

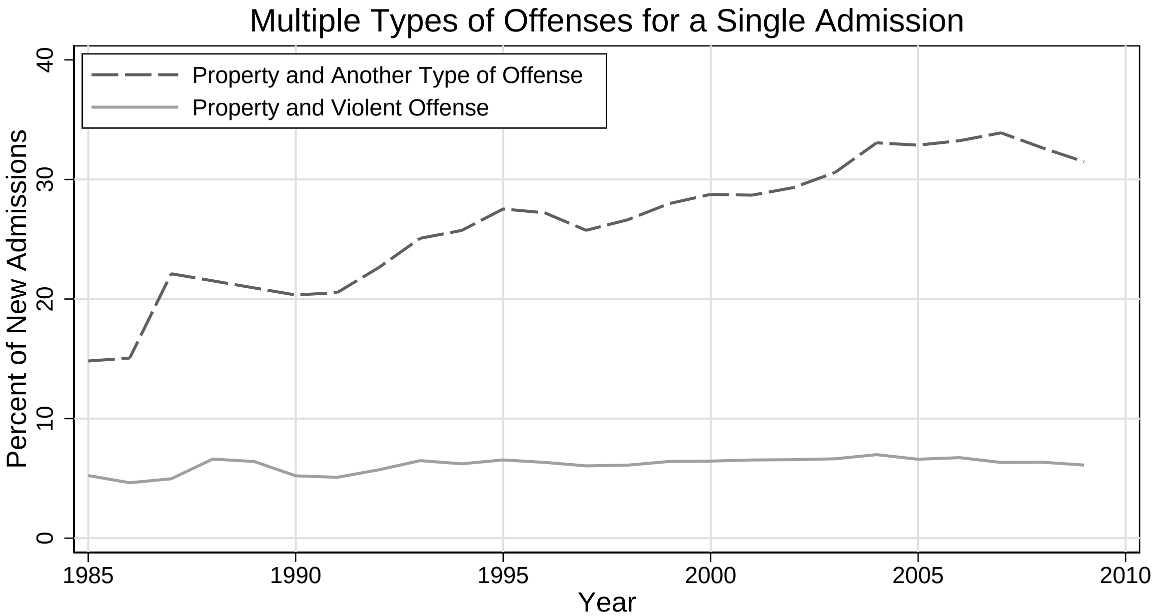


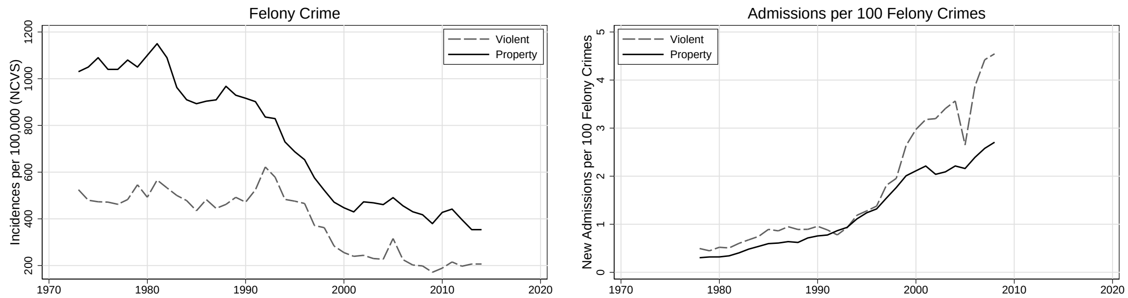
Figure 3: Percent of admissions for a property crime that include crimes in other categories (violent, drug, or other).

do not require the crime to be serious enough to result in imprisonment. They include misdemeanor crimes that would produce jail incarceration only. It is also the case that papers that use the UCR statistics will have lower levels of crime compared to the NCVS and therefore higher apprehension rates.

We add validation to our chosen values of the probability of prison conditional on crime by comparing to complete information on arrests, convictions, and incarceration on all crimes by category for cases processed in state courts provided by the Bureau of Justice Statistics "Felony Sentences in State Courts" series published for several years in the time period 1986-2009. These reports collect data from a subset of US counties. For comparison to our analysis we use the years 1986 and 2002; the longest time period for which these reports also include national estimates.

Table 1 reports the BJS estimates. While these figures do not provide a complete analysis of all property crimes or all violent crimes, nor do they provide weighted averages, our estimates are reasonably within their range. These statistics also suggest that policy, not apprehension technology, has driven the increase in incarceration per crime since the arrest rate has not increased much, if at all, for these crimes.

Another point of comparison is Pettit (2012) who estimates the probability of prison for all crimes to be approximately 2 percent in the 1970s. A crime-weighted average of our estimations for property and violent crimes is close to this figure.



(a) Crime (b) Prison Admissions per Crime

Figure 4: Crime per 100,000 adults and prison admission rates per crime calculated from NCRP, Census, and NCVS data.

1986	Arrest per Reported Crime	Conviction per Arrest	Incarceration per Conviction	Incarceration per Reported Crime
Murder	84.7	56.4	53.7	45.5
Robbery	20.7	37.7	32.8	6.8
Aggregated Assault	36.7	12.5	8.9	3.3
Burglary	8.9	35.6	26.2	2.3
2002	Arrest per Reported Crime	Conviction per Arrest	Incarceration per Conviction	Incarceration per Reported Crime
Murder	79.0	70.2	66.7	52.7
Robbery	19.3	47.2	40.6	7.9
Aggregated Assault	45.9	23.3	16.5	7.6
Burglary	9.4	49.9	35.9	3.4

Table 1: Source: BJS- Felony Sentences in State Courts Series

1.3 Survey of Inmates of State Correctional Facilities (Prison Survey)- 1974, 1979, 1986, 1997, 2004

The Survey of Inmates of State Correctional Facilities is a restricted-access, representative survey of inmates in adult correctional facilities. We use the 1979 survey consisting of approximately 12,000 inmates in 300 institutions for the initial calibration of the model. The calibration sample is restricted to male inmates entering the prison in the 1970's. All observations are weighted with frequency weights provided by the survey to construct a nationally representative sample. These weights account for non-response.

These data include prisoner's responses to their labor market characteristics *at the time they committed the crime for which they are currently incarcerated for*. Table 2 provides mean the distribution of prisoners across labor market statuses striated by age and type of

Targeted Statistics from the Prison Survey 1979				
		Violent	Property	All, Incl other
Prevalence of prison by age 35 (%)		2.4%	1.7%	
Employed Month of Crime (%)				
	Age 18-64	71.5	71.8	71.0
	Age 18-35	71.2	71.2	70.5
Unemployed Month of Crime (%)				
	Age 18-64	14.1	14.6	13.9
	Age 18-35	15.3	15.7	15.1
NiLF Month of Crime (%)				
	Age 18-64	14.1	13.2	14.7
	Age 18-35	13.4	12.8	14.1
Mean income if Empl				
	Age 18-64	27,029	26,413	27,264
	Age 18-35	25,786	24,980	25,934

Table 2: Source: BJS- Survey of Inmates of State Correctional Facilities 1979 series

crime. Employed includes both part and full time. The general patterns are that there are little differences in the employment to population ratio across violent and property criminals but property criminals are more likely to be in labor force and have lower mean earnings conditional on being employed.

Table 2 also provides estimates of the extensive margin of crime: the percentage incarcerated for a crime by age 35. These estimates are constructed using the estimates of lifetime incidence of imprisonment in a Federal or State prison calculated in Bonczar (2003). The estimate of incidence of imprisonment by age 35 for each violent and property crimes is calculated by multiplying the lifetime incidence for males in Bonczar (2003) by the share of inmates whose most serious crime is violent or property, respectively. This calculation is consistent with our calibration strategy which assumes that first time admissions for individuals over 35 are zero. This choice is motivated by the estimates in Bonczar (2003) that the share first admissions after age 35 to be less than 5% of all admissions in 1979.

Table 3 displays statistics related to the degree of criminal specialization over the life course. These statistics are calculated from the 1997 wave of the Prison Survey because this wave includes questions asking current prisoners about past offenses resulting in incarceration. There is some specialization. Prisoners serving time for a property offense are less likely to have a prior violent offense but specialization is far from complete. 27% of prisoners serving time for a property offense have served time in the past for a violent offense and 38.7% of prisoners serving time for a violent offense have served time in the past for a

Current and Prior Admissions from the Prison Survey 1997			
Current Offense	First Timer	Prior Non-Violent	Prior Violent
Property	15.5	57.21	27.2
Violent	30.6	38.7	35.8
Drug	26.5	52.1	21.5

Table 3: Charges in multiple types of crime for different prison admissions. Note the question is asked slightly differently for violent crime, allowing a single individual to report both violent and non-violent priors, resulting in shares that add up to more than 100%.

non-violent offense.

1.4 Recidivism of Prisoners Released Series, 1983 and 1994

The Bureau of Justice Statistics organized the compilation of data tracking three years of post-released outcomes for prisoners released in 1983 and 1994. These restricted data cover a representative sample of 16,000/272,111 released prisoners in 1983/1994 from California, Florida, Illinois, Michigan, Minnesota, New Jersey, New York, North Carolina, Ohio, Oregon, and Texas.[5] Prisoners in these states comprise approximately two-thirds of the prison population. The files have two layers of data. The first layer includes socio-demographic data and corrections records data at the time of inmate release. The second layer contains information on subsequent events over the three years after release including arrest, imprisonment, and non-criminal data.

Our statistics for recidivism in the 2000's come from restricted micro-data we have obtained from the study: "Criminal Recidivism in a Large Cohort of Offenders Released from Prison in Florida, 2004-2008". The study provides similar variables to the Recidivism of Prisoners Released Series for over 156,000 offenders released from the Florida Department of Corrections between 1996-2004. Outcomes for each released individual are available from state criminal records for 3 years following their release. We restrict our analysis to individuals admitted to prison after 2000 for comparability.

There are a couple of notes regarding the comparison of the Florida survey to the 1983 and 1994 surveys. While Florida on its own is less representative than the 11 states used in those survey, it is consistently a top-3 state in number of state prison inmates accounting for 7-10% of the total national prison population. The greater concern is that the survey covers only recidivism activities taking place within Florida. For this reason we would expect

[5] Arizona, Delaware, and Virginia were added in the 1994 survey, but we exclude them for consistent comparison across surveys.

to under-estimate recidivism activities relative to the 1983 and 1994 surveys. However, a comparison of re-imprisonment rates over a three year time horizon with those reported for the 2005 iteration of the Recidivism of Prisoners Released Survey provides confidence in the comparability of the Florida Survey. We calculated a total 3-year reimprison rate of 36% from our sub-sample of the Florida data which is remarkably close to the same statistic of 36.1% in the BJS report from the third report of the Recidivism of Prisoners Released Survey series from 2005[6] We cannot use the 2005 Recidivism of Prisoners Released directly because we were unable to secure access to the micro data but it seems our Florida sample is a reasonable substitute.

We are interested in a single dimension of recidivism most consistent with our model and measurements in other data sets: re-imprisonment for a new felony charge. The table below presents trends in this statistic by the age partition used in our model.[7]

Age	1983		1994		2000-2003*	
	Violent	Property	Violent	Property	Violent	Property
6 months	7.9	11.9				
1 year	13.5	19.9				
2 year	19.3	27.1				
3 year	22.2	30.7				
4 year	23.7	32.5				
5 year	24.7	33.8				
	Total 3-year Recidivism					
18-24	41.2	64.0	37.8	41.0	40.7	48.8
25-34	26.2	32.6	31.9	40.3	33.4	49.6
35-64	13.9	27.0	25.1	35.6	21.8	44.3
Total (18-64)	22.2	30.7	34.0	39.3	32.6	47.7

Table 4: 3-year Re-imprisonment Rate on a New Felony Charge. *2000-2003: Florida only.

[6]That report cautions against comparisons across years because of the stark demographic changes of the prisoner population towards older individuals. However, this is exactly what our theory predicts! It is not a problem for us to compare recidivism rates across these surveys with recidivism rates from our model data because the (endogenous!) age demographics in our model are changing in a similar way as the data.

[7]Caution must be used when comparing these data with the BJS summary papers on the surveys. Our analysis of the micro-data exactly replicates these reports when using the "Received" records from prisons/jails to identify re-incarceration. Using this measure we match their 40% 3-year recidivism rate for 1983 which breaks down to 51%/38%/29% for young, middle, and old respectively. However, this measure includes both jails which we are not considering in other datasets and includes re-confinement for violation of conditions of release, probation, or parole which we also do not model and do not include in the admission data from the NCRP data.

1.5 Decennial Census and Current Population Survey

We use data from the Decennial Census and Current Population Surveys to calculate population counts and labor market statistics for our focus population. The sample includes males who are ages 18-35 at survey date with a high school degree or less.

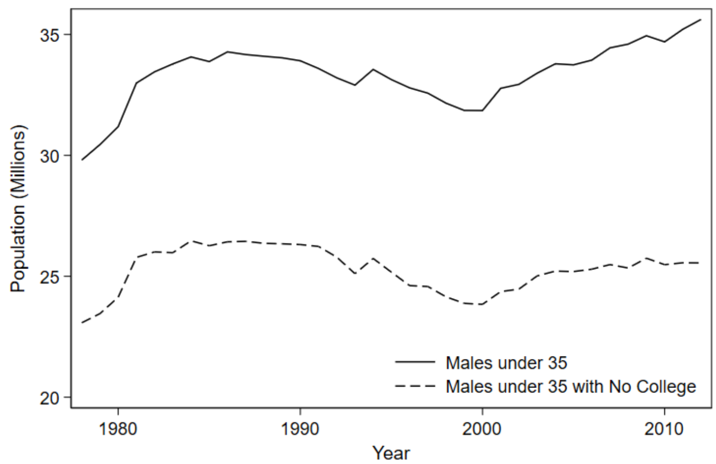


Figure 5: Population, CPS.

Figure 5 shows the total estimated population for our sample over our time of study. Note that there is little population growth for this group over time. For this reason our incarceration rates may look different from what is shown in other studies that measure the incarceration rate as total prisoners divided by the total adult population. Dividing instead by the population of males under 35, as we do, gives a better measure of incarceration hazards because prison admissions are overwhelmingly young males.

Figure 6 shows employment statistics for our sample. Our calibration target is full year employment defined as working 50-52 weeks in the reference year. We focus on the employment to population ratio because we do not model the participation decision.

For the estimation of the model, we target statistics averaged across the 1970-1980 Current Population Survey for our sample. First the employment rate is 76.2%. As a robustness check on wages, we compare growth rates implied by our NLSY regression to life-cycle wage growth (cross-section) in the 1980 Census. Wages are constructed as total labor income divided by the product of weeks worked in the year times average hours per week.[8]

[8] We set hourly wages below \$2 or above \$200 (1980) dollars equal to missing

Figure 6: Employment to Population, Males with a High School Degree or less, CPS.

Statistic		Value	Growth	Standard Dev.
Young (18-24)				
	All	5.41		6.38
Middle (25-34)			Growth Young to Middle	
	All	6.95	22.1%	6.69
Old (35-64)			Growth Middle to Old	
	All	8.45	21.5%	7.67
Total (18-64)			Growth Young to Old	
	All	7.58	56.2%	7.37

Table 5: Median Hourly Wage by Age Group (1980-dollars)

1.6 National Longitudinal Study of Youth 1979 (NLSY79)

We use data from the July 18, 2013 release of the NLSY79. The 1979 cohort of the NLSY consists of representative panel of 12,686 young men and women age 14-22 during their first interview in 1979. The timing of the study makes it appropriate to calibrate the initial steady state of the model to since we are targeting the late 1970's. Respondents were surveyed annually from 1979 to 1994 and biannually thereafter. The sample is restricted to Black or White males that do not have a college degree. The NLSY asks questions about crime and incarceration outcomes in both contemporaneous and retrospective questions. Within this sample, 19% report incarceration in jail or prison at some point in their lives.

The NLSY data includes variables on both labor market outcomes and incarceration. Labor market variables, including labor force participation, employment and unemployment status, hourly wages, and job characteristics are available on a weekly frequency. In our model, all jobs are found through search and there is no intensive margin. Accordingly, we define employment in the NLSY sample as any non-self employed job worked a median of 35-100 hrs per week over the employment relationship. We match each job to its characteristics using the Employer History Roster. Hourly wage for each job in each week is also taken from the Employer History Roster.[9] We use CPI to calculate wages in 1987 dollars and exclude wages less than \$2.00 or greater than \$200.00 per hour as missing.

Our theory of wage dynamics is a "learning by doing" type of progression following Ljungqvist and Sargent (1998). Wages increase probabilistically following a period of employment and decrease probabilistically following a period of non-employment along a predetermined grid. To calibrate the grid points and the transition probabilities, we follow Kitao, Ljungqvist, and Sargent (2017). This is a quantitative paper that introduces a quadratic life-cycle wage profile into the Ljungqvist and Sargent (1998) framework. Figure 7 below confirms our sample of interest also exhibits a quadratic life-cycle wage profile and so this approach is well-suited for our needs.

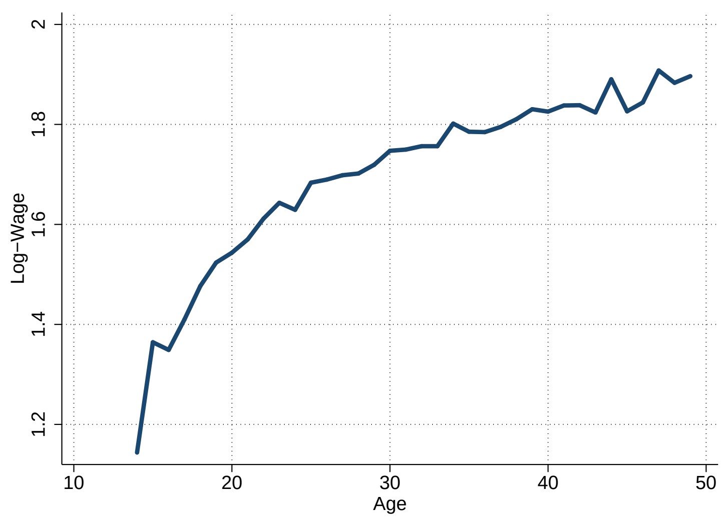


Figure 7: Mean Log-Wage by Age, NLSY 79

We construct a regression where the dependent variable is the natural-log of the hourly wage ($ln(w_{it})$). The regression includes a quadratic term to capture the typical life-cycle wage profile (A_{it}) which will be used to set transition probabilities and the grid shape for the employed. The regression also includes a quadratic transformation of the length of total non-employment over the past two years (N_{it}). This is motivated by the wage scarring literature showing persistent wage effects from periods of non-employment (Michaud (2018)). Finally,

[9]If a worker is employed in two jobs in the same week, we consider the longest held job.

individual fixed (γ_i) effects are included to control for level differences across individuals as we are concerned with growth rates, not levels.

$$ln(w_{it}) = \alpha + \beta^A A_{it} + \beta^{A2}(A_{it})^2 + \beta^N N_{it} + \beta^{N2}(N_{it})^2 + \gamma_i + \epsilon_{it}$$

We consider two variations on measures for the life-cycle: (1) age and (2) measured experience (months of employment).[10] We also provide robustness as to the type of non-employment: (a) all non-employment spells aggregated; (b) non-employment and prison spells separated; (c) non-participation, unemployment, and prison spells separated. Table 6 shows the estimates and standard errors for each specification along with the persistence of the residual in the AR(1) corresponding to each specification.

Labor market flows are constructed at the weekly frequency with NLSY data. The states and flows are identified as follows. Employment is defined in the same way as described above for the Mincer regression. Non-employment is categorized into "unemployed" or "non-participant" according to a question asking the respondent's job search status. If the respondent has a job, but that job does not meet our requirement to be classified as "employed", we categorize the individual as "unemployed". We use the tenure variable to clean for spurious flows including transitory changes in hours that would move a respondent across states. We do this as follows. If we see a switch from "employed" to any of our non-employment categories at time "t", we then check the tenure variable reported for the next 4 weeks. If we see the respondent becomes "employed" in the next four weeks and the tenure is greater than one month, then we count the individual as having had been continuously employed. In other words, if they regain employment at tenure greater than one month, we conclude the transition is spurious and drop it as a true transition.

2 Additional Details on The Prison Boom 1980s-

Contribution of Demographics- Age, Race, and Employment Status. To be sure the expansion in incarceration and the cohort results are not driven by changes in the racial or employment status composition of these groups, we perform a sort of shift-share analysis on prison admissions data. We divide the population into 12 cells covering the intersections of each of two race groups: black and white, two employment groups: employed and non-employed, and three age groups: 18-24, 25-34, and 35-54. We then calculate the prison

[10] The data are censored with a maximum age of 50 on account of the single-cohort panel structure of the NLSY.

admission rate for each group in the first BJS Prison Survey: 1979.[11] Employment in this survey is a self-report of status at the time of arrest. Next, we predict the admission rate for each age group as follows. We first calc λ^y to satisfy:

$$TotAdmitRate^y = \sum_{r,a,e} \lambda^y \phi^{1980}_{r,a,e} \pi^y_{r,a,e}$$

where $\phi^{1980}_{r,a,e}$ is the 1980 admission rate for each demographic cell (r, a, e) where $r \in \{black, white\}$ is race, $e \in \{employed, nonemployed\}$, $a \in \{18-24, 25-34, 35-54\}$. $\pi^y_{r,a,e}$ is the share of each demographic cell in year y. Therefore, λ^y is the percent increase in admissions rate necessary to match the total admission rate in year y, holding year 1980 relative behavior of each demographic cell fixed, but adjusting for changes in each cells share in the total demographics. We then calculate predicted rates for each age group a as:

$$Adm\hat{i}tRate^y_a = \sum_{r,e} \lambda^y \phi^{1980}_{r,a,e} \pi^y_{r,a,e}$$

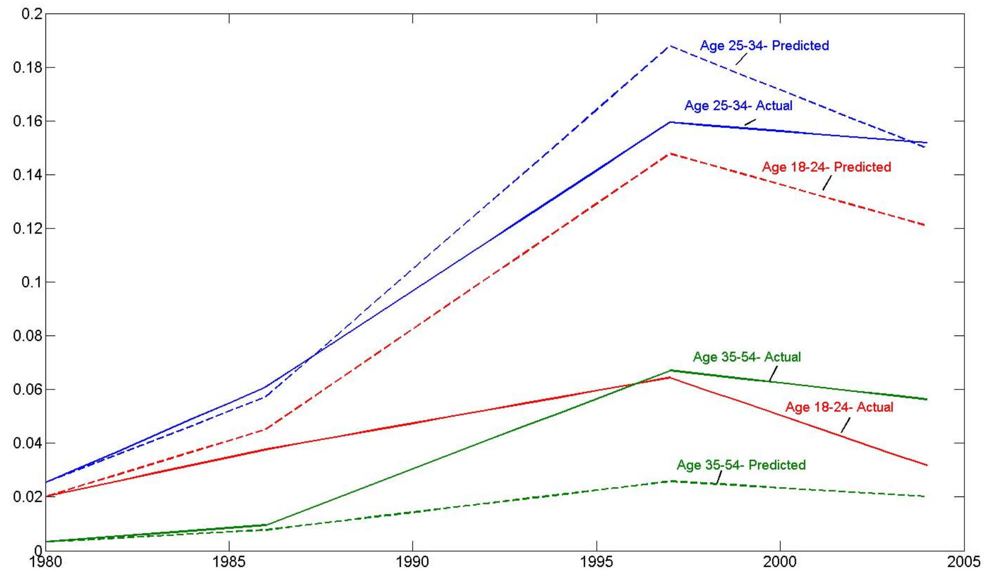
Figure 8 shows our results. It shows that admissions rates for the youngest group, age 18-24 have been consistently lower than predicted by 1979 behavior. The middle age group, 25-34, is sometimes lower and sometimes higher than predicted. Finally, the oldest age group 35-54, has rates substantially higher than predicted by the late 1990s.[12] These patterns are informative about the dynamic role of deterrence and are largely consistent with the theory developed in the paper.

The left pane of Figure 9 displays predicted and actual incarceration rates by employment status, controlling for changes in race and age composition across the two over time. Observe that the increase in incarceration rates occurred for both employed and non-employed individuals. Therefore aggregate rates cannot be accounted for by changes in employment status. Finally, the right pane shows that whites account for only slightly less of the increase than would be predicted from pre-1980 outcomes. In other words, the high incarceration rates of blacks relative to whites were not driven by policy changes but accounted for by pre-existing differences prior to 1980.

Stability in Length of Time Served. Prior to the late 1990's, state-level inmate records were not uniform nor well-kept. There was no reporting requirement to federal authorities.

[11]We limit our analysis to males. We use SEER data to measure total population counts each cell. We adjust this data to be only for those with High School degrees or less using Decennial Census data and interpolating linearly for non-decade years.

[12]The BJS Prison Survey is only conducted every seven years providing limited data points for this analysis.



(a) Predicted and Actual Rates by Age

Figure 8: Actual values calculated from BJS Prison Surveys and Census data. Predicted values calculated by holding admission rates fixed to 1979 levels, and raising rates by the same proportion for each age group, adjusting for demographics (race and employment).

Thus, there is no historical time series of the sentence length served in state and federal prisons. Neal and Rick (2014) and Pfaff (2011) infer sentence lengths from admission, stock, and release data on successive age cohorts of inmates in a sample of state prisons. Both papers concur that the median length of time served in prison for new admits did not change much over the last few decades. This agrees with other papers such as Raphael and Stoll (2009). Therefore, we keep the parameter governing the duration of time served (probability of release) constant through our transition experiment. We set this parameter to provide an expected duration of 2.7 years following Raphael and Stoll (2009) who find an average duration served of 2.64/2.73 in 1984/1998. During this time period, the BJS did report that the time served for Federal prisoners did increase from 15 months to 29 months. However, federal prisoners represent just a small portion of total prison inmates, a consistent 7% during this time.

A Note on the Role of Drug Crime and Enforcement One hypothesis is that criminality associated with drug markets, particularly associated with cocaine, is important to understanding the aforementioned incarceration trends. This hypothesis is met with skepticism in the Criminology and Economics literatures. A look at the data reveals why. It is true that prison admissions involving a drug charge have been the category with the largest expansion over the past 30 years. However, the rise in admissions based on drug felonies can only account for 33% of total state and federal admissions at their peak in the 1990's and represent less than 20% of admissions in 2010.[13] Sentences for drug felonies are relatively short and so prisoners currently serving for a drug offense comprise an even smaller share of the stock relative to the flow.

3 Simplified Model of Dynamics.

In this section we establish a simplified model of crime and incarceration guided by the main mechanisms in our full quantitative model. We use it to derive an empirical strategy to estimate age, time, and cohort effects from semi-aggregated panel data given a set of assumptions that are consistent with our theory. Let $C_{j,t}$ and $I_{j,t}$ be the crime and incarceration rates, respectively, of cohort j at time t. These are our outcomes in the data for which we are interested in measuring cohort effects. The relationship between these variables over time is provided by the following equations.

$$\begin{array}{rl} \text{Incarceration Rate} & I_{j,t} = \pi_t C_{j,t} \\ \text{Initial Crime Choice} & X_{j,0} = g^X(\pi) \\ \text{Evolution of Crime Rate} & C_{j,t} = X_{j,t} A_a + T_t \\ & X_{j,t} = (\phi + \beta \pi_{t-1}) X_{j,t-1} \end{array}$$

The interpretation of this model in relation to our research is as follows. The policy variable is π_t: the probability of incarceration conditional on committing a crime. It is exogenous and can change over time. The first line presents the result that, assuming a large population, the incarceration rate for cohort j at time t is equal to that cohort's crime rate $C_{j,t}$ multiplied by the incarceration probability π_t.

The remaining equations explicate an extreme version of the cohort effects found in the full structural model. In the full model, choices made under the policy prevalent during youth persistently affect outcomes even as the policy changes later in life. Here, we model that cohort effect as a permanent component $X_{j,0}$ interpreted as an initial crime choice. The initial crime choice is given by a function $g^X(\pi) \in [0,1]$. We assume this function is twice continuously differentiable in $(0,1)$ and that $g'^X(\pi) < 0$, ie: that punitive policy deters.

[13]Calculation from Bureau of Justice Statistics Data.

The final two lines show the evolution of a cohort's crime rate given the initial crime choice and the evolution of the policy. First, the cohort's last period crime rate $X_{j,t-1}$ has a persistent effect on today's crime rate $X_{j,t}$. The coefficient term $(\phi + \beta\pi_{t-1})$ has the following interpretation. The term $\phi < 1$ captures the direct effect crime today has on crime tomorrow. The term $\beta\pi_{t-1}$ captures the effect that a prison experience yesterday has on crime today. For what follows, we assume that β is larger than zero in which case a prison experience increases future crime or at least slows its decay. Both ϕ and β can be interpreted as some persistent criminal capital. The age effect is A_a. In the data, crime peaks before age 20 and declines over the life-cycle. Incarceration is hump shaped, peaking between 25-35 before declining.[14] Therefore A_a will be larger or smaller than one capturing the growth and decay of crime related to the life cycle not otherwise captured by the criminal capital process. Time shows up in two ways. The first is a level effect T_t. The second is through changes in the policy π_t over time.

The first two propositions present steady state comparative statics with respect to π. For these, we suppress the time and cohort subscripts. The first result from this model, summarized in Proposition 1, is that the age profile of crime looks different in steady states with different incarceration probabilities π. In particular, as π increases crime is more persistent over the life-cycle resulting in higher incarceration rates for old individuals relative to young.

Proposition 1 (A steady state with a higher punitive policy exhibits higher crime and incarceration at older ages relative to young). *Let two policies $\hat{\pi} > \pi$ be given and $\hat{X}_a$ and X_a be the persistent component of crime at age a in the steady-state for each policy, respectively. Then:*

$$\frac{\hat{X}_a}{\hat{X}_{a-s}} > \frac{X_a}{X_{a-s}} \quad \forall \quad s \in (1, a)$$

Proof. We begin by showing $\frac{\hat{X}_a}{\hat{X}_{a-1}} > \frac{X_a}{X_{a-1}}$, then the remainder cases for $s \in (1, a)$ can be

[14] This model will not be able to reconcile the monotonic decline of crime with the non-monotone shape of incarceration. This is both because crimes in the data include less serious offenses and because incarceration sentences depend on past criminal records. Instead of including these features, we instead estimate the model twice in the data for each series to capture the two different concepts of crime.

completed by induction. Expanding, the result is immediate:

$$\begin{aligned}\frac{\hat{X}_a}{\hat{X}_{a-1}} &= \frac{(\phi+\beta\hat{\pi})\hat{X}_{a-1}}{\hat{X}_{a-1}}\\ &= (\phi+\beta\hat{\pi})\\ &> (\phi+\beta\pi)\\ &= \frac{(\phi+\beta\pi)X_{a-1}}{X_{a-1}} = \frac{X_a}{X_{a-1}}\end{aligned}$$

The inequality holds since it is given that $\hat{\pi} > \pi$ □

The change in the life-cycle profile at the steady state when the policy increases occurs irregardless of the elasticity of the initial crime choice. Crime becomes more persistent over the life cycle through the prison experience so long as $\beta > 0$. Since incarceration is hump-shaped over the life-cycle, this implies that the peak of life-cycle incarceration will move to older ages for β sufficiently large. This result is particularly important for how we think about time and age effects in the data. It is consistent with shifts towards incarceration at older ages that are salient in the data and suggests the permanent component of this shift can be interpreted as the effect of changes in policy.

The second set of results address how a change in punitive policy (π) affects aggregate crime and incarceration rates.[15]

Proposition 2 (Conditions for increased punitive policy to decrease crime.)**.** *Let $\mathbb{C}(\pi)$ be the aggregate crime rate. For a given π, the likelihood of $\frac{\partial\mathbb{C}(\pi)}{\partial\pi} < 0$ is:*

- *decreasing in β.*
- *decreasing in $\sum_{a=0}^{M} A_a$.*
- *increasing(decreasing) in π if $g''^X(\pi) >> 0$ ($g''^X(\pi) << 0$).*

Proof. The aggregate crime rate given π is the sum of crime across all age groups:

$$\mathbb{C}(\pi) = g^X(\pi)\sum_{a=0}^{M}(\phi+\beta\pi)^a A_a$$

Then, since we assume $\frac{\partial g^X(\pi)}{\partial\pi} < 0$, it is true that $\frac{\partial\mathbb{C}(\pi)}{\partial\pi} < 0$ iff:

$$-\frac{\partial g^X(\pi)}{\partial\pi} > \frac{\partial\sum_{a=0}^{M}(\phi+\beta\pi)^a A_a}{\partial\pi}$$

[15] It is assumed the maximum age is M, but these proofs will also apply to $\lim_{M\to\infty}$, so long as parameters are appropriately restricted such that crime is finite.

By inspection, the right-hand side is increasing in both β and the sequence A_a, and so the inequality less likely to hold for larger values of these parameters. Also, the left-hand side is increasing in π if $g''^X(\pi) >> 0$, and so the inequality is more likely to hold for larger values of π if g^X is strictly convex. $\square$

Corollary 1 (Response of incarceration to increased punitiveness.)**.** *Let $\mathbb{I}(\pi)$ be the aggregate incarceration rate. For a given π, is the likelihood of $\frac{\partial \mathbb{I}(\pi)}{\partial \pi} < 0$ is:*

- *decreasing in β.*
- *decreasing in $\sum_{a=0}^{M} A_a$.*
- *increasing(decreasing) in π if $g''^X(\pi) >> 0$ ($g''^X(\pi) << 0$).*

Proof. Omitted. $\square$

We now consider the effect of a policy change along the transition. The main result, summarized in Proposition 3 explicates the existence of cohort effects.

Proposition 3 (The cohort born immediately before an increase in π has higher age-specific crime and incarceration rates at all ages than all cohorts it precedes and follows.)**.** *Let an initial π_0 be given. Denote with hat notation the variables related to the cohort born at $\bar{t}-1$ where $\bar{t}$ is when the policy is changed to $\pi > \pi_0$. Then:*

$$\begin{aligned} C_{\hat{j},t} &> C_{t-(\hat{j}+s),s} \quad \forall \quad t > \bar{t}+1 \quad \textit{and} \quad s \neq \bar{t}+1 \\ I_{\hat{j},t} &> I_{t-(\hat{j}+s),s} \quad \forall \quad t > \bar{t} \quad \textit{and} \quad s \neq \bar{t} \end{aligned}$$

Proof. Expanding $C_{\hat{j},\bar{t}}$, we have:

$$\begin{aligned} C_{\hat{j},\bar{t}} &= X_{\hat{j},\bar{t}} A_{\bar{t}-j} + T_{\bar{t}} \\ &= (\phi + \beta\pi_0) X_{\hat{j},\bar{t}-1} A_{\bar{t}-j} + T_{\bar{t}} \\ &= \Pi_{\tau=\hat{j}}^{\bar{t}} [(\phi + \beta\pi_0) * X_{j,0}] A_{\bar{t}-j} + T_{\bar{t}} \\ &= \Pi_{\tau=\hat{j}}^{\bar{t}} [(\phi + \beta\pi_0) * (g^X(\pi_0))] A_{\bar{t}-j} + T_{\bar{t}} \\ &= [(\phi + \beta\pi_0)^{\bar{t}-\hat{j}-1}] * (g^X(\pi_0)) A_{\bar{t}-j} + T_{\bar{t}} \end{aligned}$$

Since time and age effects are invariant to the policy change, we can ignore them. It suffices to show that $X_{\hat{j},t} > X_{t-\hat{j}+s,s}$ for all $t > \bar{t}$ and all $s \neq \bar{t}$. The evolution of $X_{\hat{j},t}$ for $t > \bar{t}$ is:

$$X_{\hat{j},t} = [(\phi + \beta\pi)^{t-\bar{t}}] * [(\phi + \beta\pi_0)^{\bar{t}-\hat{j}-1}] * (g^X(\pi_0))$$

First, let us consider prior (older) cohorts at the same age in the past: $s < \bar{t} + 1$. Their persistent component is $X_{t-\hat{j}+s,s}$. We want to show the following relationship:

$$\begin{aligned} X_{t-(\hat{j}+s),s} &= [(\phi + \beta\pi_0)^{t-\hat{j}-1}] * (g^X(\pi_0)) \\ &< [(\phi + \beta\pi)^{t-\bar{t}}] * [(\phi + \beta\pi_0)^{\bar{t}-\hat{j}-1}] * (g^X(\pi_0)) \\ &= X_{\hat{j},t} \qquad \forall \quad t > \bar{t} \quad \text{and} \quad s < t \end{aligned}$$

This inequality holds because $\pi > \pi_0$ and $\beta \in (0, \infty)$. Now, for later (younger) cohorts at the same age in the future: $s > \bar{t} + 1$. Their persistent component is $X_{t-(\hat{j}+s),s}$. We want to show the following relationship:

$$\begin{aligned} X_{t-(\hat{j}+s),s} &= [(\phi + \beta\pi)^{t-\hat{j}+s-1}] * (g^X(\pi)) \\ &< [(\phi + \beta\pi)^{t-\bar{t}}] * [(\phi + \beta\pi_0)^{\bar{t}-\hat{j}-1}] * (g^X(\pi_0)) \\ &= X_{\hat{j},t} \qquad \forall \quad t > \bar{t} \quad \text{and} \quad s > \bar{t} \end{aligned}$$

It suffices to show, for any n:

$$[(\phi + \beta\pi)^n] * (g^X(\pi)) < [(\phi + \beta\pi_0)^n] * (g^X(\pi_0))$$

Which holds for $\pi > \pi_0$ and $g'^X(\pi) < 0$, both as assumed. $\square$

Corollary 2 establishes an additional restriction required for the cohort effect to translate to a non-monotone transition in crime and incarceration. With respect to crime: it is essentially required that the increase in the incidence and impact of prison on future crime to not be too large relative to the initial crime choice. It does not require that crime fall in the new steady state, but that would suffice. The condition is more stringent with respect to incarceration since the change in incarceration rate is the change in the crime rate times the change in the policy π. Still, it is not required that crime fall in the new steady state in order for the transition to be non-monotone, but again this would be sufficient.

Corollary 2 (The transition path of crime and incarceration after an increase in punitiveness are non-monotone if the elasticity of the initial choice is sufficiently large relative to the effect of prison on criminal persistence.)**.** *Let an initial* π_0 *be given and consider the economy at a steady state for that* π_0*. Assume at time-zero the policy switches permanently and unexpectedly to* $\pi_1 > \pi_0$*. Then:*

- *a) The transition path for crime is non-monotone iff*

$$\frac{g^x(\pi_0)}{g^x(\pi_1)} > \frac{\sum_{a=0}^{M-1}(\phi+\beta\pi_1)^a+1}{\sum_{a=0}^{M-1}(\phi+\beta\pi_0)^a+1}$$

- *b) The transition path for crime is non-monotone iff*

$$\frac{\pi_0 g^x(\pi_0)}{\pi_1 g^x(\pi_1)} > \frac{\sum_{a=0}^{M-1}(\phi+\beta\pi_1)^a+1}{\sum_{a=0}^{M-1}(\phi+\beta\pi_0)^a+1}$$

Proof. Let $C(\pi_0)$ and $C(\pi_1)$ be the steady state aggregate crime rate at the two policies, π_0 and π_1, respectively. Let C_0 be the aggregate crime rate the period after the policy change. Then:

$$\begin{aligned}
C(\pi_0) &= g^x(\pi_0)\sum_{a=0}^{M}(\phi+\beta\pi_0)^a A_a \\
C_0 &= g^x(\pi_0)(\phi+\beta\pi_1)[\sum_{a=0}^{M-1}(\phi+\beta\pi_0)^a A_a + 1] \\
C(\pi_1) &= g^x(\pi_1)\sum_{a=0}^{M}(\phi+\beta\pi_1)^a A_a
\end{aligned}$$

That $C_0 > C(\pi_0)$, follows directly from Proposition 3. Simple algebra comparing C_0 and $C(\pi_1)$ provides the necessary and sufficient condition provided in the statement of this corollary. □

Corollary 3 (If crime falls in the new steady state with increased punitiveness, then the transition path of crime and incarceration to that steady state are non-monotone.)**.** *Let an initial π_0 be given and consider the economy at a steady state for that π_0. Assume at time-zero the policy switches permanently and unexpectedly to $\pi_1 > \pi_0$. Let $C(\pi)$ and $I(\pi)$ be the aggregate crime and incarceration rates at the steady state of policy π. Then, if $C(\pi_0) > C(\pi_1)$, the transition between steady states is non-monotone for both crime and incarceration.*

Proof. Follows straightforwardly from Proposition 3. □

3.1 Results.

First we check to see that the data are consistent with the assumptions and predictions of the simple model with respect to the age-profile. We assumed that the age profile peaks early on and then decays at a constant rate. We then derived that a change in the policy π should change the shape of the age-profile by increasing mass at older ages. Figure 10 shows that the data are consistent with these features. In particular, the shape of the age profile in the earliest year available, close to the pre-1980's "steady state" has a shape that exhibits a strikingly constant decay over the life-cycle. The same is true for the last year available only for arrests: 2010. This is the closest we get to the model's predicted new steady state if we consider the major policy changes to have happened in the early 1980s. It is also true that the latest year available for each series puts a relatively higher mass on older ages compared to young. Finally, the temporary bumps upwards in the age profile over time departing from a constant decay are hints at cohort effects along the transition. We now estimate these effects more formally following our regression strategy.

The full regression specification is:

$$\begin{aligned} I_{a,c,t} &= (\beta^T \mathbf{D^T} + \beta^C \mathbf{D^C}) * (\beta^A \mathbf{D^A} * \beta^Y \mathbf{D^T}) \\ \text{st} & \quad \beta^Y = 0 \quad \text{if} \quad a < 26 \\ & \quad \beta^Y \geq 1 \quad \text{if} \quad a > 25 \end{aligned}$$

The independent variables D^T, D^C, and D^A are respectively dummies for time, age, and cohorts.[16] Although time enters in two ways, we refer to β^T as the "time effect". The cohort effect is β^C. Age effects are multiplicative to time and cohort: β^A. Finally, we allow the age effect to change over time (β^Y) only after the peak of the life-cycle incarceration curve. We also impose that this coefficient be greater than or equal to one so that it only captures the flattening of the life-cycle profile.

We estimate the regression equation using non-linear least squares. The results from the full regression specification and from an alternative specification without time varying age profiles ($\beta^Y = 0$ for all ages) are presented in Figure 11. In the first specification, the coefficients on β^Y are significant at the 95% level suggesting that the age profile flattens over time. None of cohort effects, however, are statistically different at the 90% confidence level if we include the β^Y's or not.

An alternative common in the criminology literature considers a linear model of age, time, and cohort effects according to the following specification:

$$I_{a,c,t} = (\beta^T \mathbf{D^T} + \beta^C \mathbf{D^C}) + \beta^A \mathbf{D^A}$$

[16] The data for arrests only provides 5 year age bands. Accordingly, we measure cohorts and time in five year intervals.

Since the regressors are co-linear, typical empirical implementations estimate cohort effects with two regressions. One considers cohort and time effects only and the other considers cohort and age effects only. The cohort estimates from these regressions are presented in Figure 12. The cohort effects display no or much reduced non-monotone effects. In Panel (a), cohort effects are monotone increasing because there are no year effects and so the cohort effects must capture the strong upwards trend in admission rates over the time period. In Panel (b), there are no age effects and so the older/younger cohorts that are only seen for part of their life at older/younger ages are capturing some of the age effects. This is a problematic specification because the increase in the admission rates over time is affecting each age group equally. It implies X more bodies per 1,000 entering prison, but in reality the increases in prison admissions come disproportionably from ages with high admission rates in 1980s. The proportional model we use in which the time effects scale the age profile of admissions proportionally are a much better fit.

4 Computational Appendix

4.1 Computational Algorithm for the Stationary Model

The state variables of the value and policy functions are human capital, crime capital (low and high), labor market status (incarcerated, unemployed and employed), prison flag and age (young, middle and old). We discretize the human capital into 21 grid points and approximate the age and labor market status dependent human capital process using Tauchen's method. The equilibrium objects are the market tightness for each age and prison flag, θ_{km}, individual value function, V_p, V_u, V_e, firm value functions, J_f, V_f and stationary distribution, Γ.

- Make a guess on the market tightness, , θ^0_{km}.
- Given the market tightness, compute the job arrival rates, λ^{km}_w, and worker arrival rates, λ^{km}_f using equations 4 and 5
- Iterate on the value functions V_p, V_u, V_e, and J_f, until convergence using equations 1-7. This step will also yield the decision rules and policy functions
- Using the individual decision rules and policy functions, iterate the distribution until convergence
- Given the stationary distribution and firm's value function, compute θ^1_{km} using equation 7

- If $\max_{k,m} |\theta^1_{km} - \theta^0_{km}| < \epsilon$ stop. Otherwise, update $\phi\theta^0_{km} = \theta^1_{km} + (1-\phi)\theta^0_{km}$ where $\phi = 0.8$, and return to step 2.

4.2 Computational Algorithm for the Transitional Model

We set the transition period $T = 15600$ corresponding to 300 years. parameters that change along the transition are the incarceration probability, π, mean of the crime reward distribution, μ^k, and the individual productivity, z. These transitions are introduced as linear changes over 30 years, and each gradual change is introduced as surprise and permanent.

Given these assumptions, the transition of the model is solved as follows:

- Solve the initial steady-state of the model as in 4.1. Store the initial distribution as μ^0
- Solve the final steady-state of the model as in 4.1. Store the final steady-state value functions.
- Make a guess on the market tightness along the transition $\theta^{t,0}_{km}$ for all $t \in [1, T]$.
- Given the market tightness, compute the job arrival rates and worker arrival rates along the transition.
- Starting from the last period, compute the value functions iterating on the value functions as in 4.1 until the first period.[17]
- Given the value functions and decision rules, iterate on the distribution starting from μ^0.
- Given the distribution along the transition and the firm's value function, compute the updated market tightness, $\theta^{t,1}_{km}$
- If $\max_{k,m,t} |\theta^{t,1}_{km} - \theta^{t,0}_{km}| < \epsilon$ stop. Otherwise, update $\phi\theta^{t,0}_{km} = \theta^{t,1}_{km} + (1-\phi)\theta^{t,0}_{km}$, and return to step 4.

4.3 Estimation

The estimation procedure is a mixture of Simulated Method of Moments and Indirect Inference. There are 12 parameters to be estimated in the model. The details of these parameters are explained in the Calibration section of the main text. We denote $\Upsilon = \{\eta^1, c, \delta, \mu^{e,2}, \mu^{e,3}\mu^{u,1}, \zeta^3, \rho_h, \sigma_h, \nu, \eta^{1,hc}_a, \mu^k\}$ as the set of these parameters. We estimate these

[17]Here we do not use the next period value functions since we assume all the changes along the transition are introduced unexpectedly and as permanent.

parameters by minimizing equally weighted square of percentage distance between model simulated moments and data moments. Denoting Ω_M as the model generated moments and Ω_D as the data moments, Υ solves:

$$max_{\Upsilon} \left(\frac{\Omega_M - \Omega_D}{\Omega_D} \right) W \left(\frac{\Omega_M - \Omega_D}{\Omega_D} \right)^T$$

where W is the identity matrix. The construction of the moments are explained in the Calibration section of the main text. Some of these moments are generated by running the same regression both in the real-life data and model simulated data.

4.4 Sensitivity of Moments to Selected Parameters

There is no analytical mapping of the parameters to the model moments. In Section 4.3, we discuss how the parameters are identified through the selected moments. In this section, we graphically show how certain parameters are linked to certain moments following that discussion.

Although each parameter affects all the moments calibrated, certain parameters have stronger effects on certain moments. Figure 13 plots the sensitivity of crime related moments to the crime related parameters that affect the moment strongest. In each plot, the dot shows the calibrated value of the parameter, and the line shows the change in the corresponding moment to the change in the selected parameter. The figure highlights the monotonic and strong relation between each of the parameter and the corresponding moment.

Figure 14 plots the same sensitivity analysis for the labor market related moments and parameters. Again, each of the moment is strongly and monotonically related to the corresponding parameter.

4.5 Calibration of the Model with Violent Crime

We use the same externally set parameters as in the benchmark model as none of these parameters is specific to the type of crime. The only externally set parameter changes is the arrest probability. This value is as the ratio of new prison admits for property crime estimated from NCRP's NPS restricted micro-data to the reported crimes estimated from the National Crime Victimization Survey (NCVS) for 1979-1980. For property crime, it is calculated as 0.3% whereas for violent crimes it is 0.5%. The rest of the internally calibrated parameters are listed in Table 8. Table 9 shows the performance of the model in matching the moments targeted.

4.6 Steady-State Results with Violent Crime

Figures 15 and 16 plot the comparison of both steady-states. Table 12 compares both steady-states along several moments.

4.7 Transitional Dynamics with Violent Crime

Figures 17-24 reproduce the same set of plots we present in the main text for the property crime. See the main text for explanations of each figure.

4.8 Calibration of the Alternative Models

To highlight the significance of including criminal capital into our benchmark model, we recalibrate the model by removing criminal capital and using alternative assumptions to generate the persistence in criminal activity. In each alternative model, we target the same set of moments and use the same estimation procedure. Table 14 presents the calibrated parameter values for each alternative model. The column labelled as "M0" corresponds to the benchmark model. Below, we briefly describe the assumptions for alternative models:

- **M1: No Criminal Capital** This is the model without criminal capital and no other ingredients added to the model. The absence of criminal capital removes three parameters: probability of gaining high criminal capital upon committing a crime, ν, additional crime arrival rate for high criminal capital individuals, η_a^{hc}, and the probability of losing high criminal capital in old age, ζ^3. We also add another parameter, crime arrival rate for old individuals ($\eta^3 = \zeta^3$), to be able to match the incarceration rate for old individuals.

- **M2: Higher Human Capital Depreciation** In this model, we keep high criminal capital assumption, but remove the presence of irrational crimes for the high criminal capital. Instead, we assume that high criminal capital individuals' human capital depreciates at a faster rate, and reduce the mean of the human capital process for all individuals, including employed, unemployed and incarcerated, by η_a^{hc}. This model has the same set of parameters as the benchmark model. The only difference is the interpretation of η_a^{hc}.

- **M3: More Crime Opportunities** This model is the same as the benchmark model. The only difference is that additional crimes high criminal capital individuals receive need not necessarily be committed. They are the same types of crimes low criminal capital individuals receive but at a higher rate. η_a^{hc} captures the additional crime arrival rate for the high criminal capital individuals.

- **M4: Better Crime Opportunities** This is very similar to M3. The only difference is that high criminal capital individuals do not receive crime opportunities at a higher rate, but from a distribution with a higher mean. η_a^{hc} captures the increase in the mean of crime reward for the high criminal capital individuals.

- **M5: Ex-ante Heterogeneity in Criminal Capital** In this model, we assume there is ex-ante heterogeneity in the criminal capital, and ν represents the fraction of high criminal capital individuals. However, in this model, we assume committing a crime does not allow individuals to gain high criminal capital, i.e. there is no possibility for moving from low criminal capital to high criminal capital. However, we still keep the assumption of the possibility of losing high criminal capital when old to be able to match old incarceration rate.

- **M6: Higher Arrest Probability for the Incarcerated** In this model, we again remove the arrival of irrational crimes for the high criminal capital individuals. Instead, we assume that high criminal capital individuals face a higher probability of arrest upon committing a crime.

5 A Simple Theoretical Model

Here we provide a simpler version of the model presented in the paper. We assume that the only source of ex-post heterogeneity across individuals is the employment status. That is, we assume all individuals are infinitely-lived, have identical human capital and criminal capital (low), and prison has no explicit effect on job finding probability.

Let V_p, V_u, V_e represent the value functions for an incarcerated, unemployed and employed individual, respectively, we can formulate these value functions as:

Incarcerated Individual:

$$rV_p = \tau \left(V_u - V_p\right) \tag{1}$$

Unemployed Individual:

$$rV_u = b + \lambda_w \max\left\{V_e - V_u, 0\right\} + \eta \int \max\left\{\pi\left(V_p - V_u\right) + \kappa, 0\right\} dH\left(\kappa\right) \tag{2}$$

Employed Individual:

$$rV_e = wh + \delta\left(V_u - V_e\right) + \eta \int \max\left\{\pi\left(V_p - V_e\right) + \kappa, 0\right\} dH\left(\kappa\right) \tag{3}$$

As long as $wh > b$, we have a cut-off rule for the crime decision.

Lemma 1. *There exists $\kappa_u^*\left(\kappa_e^*\right)$ such that unemployed (employed) individual commits every crime if the reward is higher than or equal to $\kappa_u^*\left(\kappa_e^*\right)$, and κ_u^* is given by*

$$\kappa_u^* = \pi\left(V_u - V_p\right) \tag{4}$$

$$\kappa_e^* = \pi\left(V_e - V_p\right) \tag{5}$$

We can also prove that $V_e - V_u > 0$. This gives us the following corollary:

Corollary 4. *If $wh > b$, then the threshold of crime reward for employed is higher than for unemployed: $\kappa_e^* > \kappa_u^*$.*

The above corollary implies each employed individual commits fewer crimes than an unemployed individual. Thus, employment is an individual characteristic that "deters" crime.

We can further characterize the values for employment, unemployment and incarceration as functions of the cut-off values κ_e^* and κ_u^*.

Lemma 2. *The values for incarceration, V_p, unemployment, V_u, and employment, V_e, can be expressed as:*

$$\begin{aligned} V_p &= \frac{\tau\kappa_u^*}{r\pi} \\ V_u &= \frac{\left(r+\tau\right)\kappa_u^*}{r\pi} \\ V_e &= \frac{r\kappa_e^* + \tau\kappa_u^*}{r\pi} \end{aligned}$$

Using these values, we can show the following two important propositions:

Proposition 1. *If $\delta > \tau$, an increase in the probability of getting caught, π , increases the crime threshold for both unemployed and employed, i.e. decreases the crime rate.*

Proposition 2. *If $wh > b$, an increase in the job offer arrival rate increases the crime threshold for the unemployed (decreases the crime rate). Furthermore if $\delta > \tau$ ($\delta < \tau$) an increase in the job offer arrival rate increases (decreases) the crime threshold for the employed and decreases (increases) the crime rate.*

These propositions show that as the criminal policy becomes more punitive (an increase in π), unemployed and employed individuals respond by committing less crimes. However, the net effect on incarceration probability is ambiguous. Notice that incarceration probability in the model is $\pi\eta\left(1-H\left(\kappa_i^*\right)\right)$ for $i \in \{u, e\}$. An increase in π directly increases this probability. However, as individuals respond, κ_i^* increases, which decreases this probability. The net effect depends on the magnitude of these two forces. The second important message of the above propositions is the effect of a change in job arrival rate on crime propensities. The second proposition shows that as labor market opportunities for individuals improve (an increase in job arrival rate), individuals decrease their crime propensities.

5.1 Firm's Problem:

Let J be the value of a match between a firm and individual and V_f be the value of a vacancy. We have the following flow equations for the firm:

$$\begin{aligned} rV_f &= -k + \lambda_f\left(J - V\right) \\ rJ &= (1-w)h + \delta\left(V_f - J\right) + \eta\left(1 - H\left(\kappa_e^*\right)\right)\pi\left(V_f - J\right) \end{aligned}$$

Lastly, the free-entry condition pins down the market tightness:

$$V_f = 0$$

Combining the free-entry condition with the above value functions, we get

$$\lambda_f = \frac{k\left(r + \delta + \eta\left(1 - H\left(\kappa_e^*\right)\right)\pi\right)}{(1-w)h} \tag{6}$$

Lemma 1. *An increase in the probability of incarceration, $\pi\eta\left(1-H\left(\kappa_e^*\right)\right)$, increases λ_f, the worker arrival rate for the firms, which means job offer arrival rate for the workers, λ_w, decreases.*

This is an immediate consequence of equation (6). The expected duration of a new match decreases as the probability a worker ends the match by going to prison increases. Shorter match duration implies lower profits for the firm. To maintain the equilibrium zero expected profits condition, fewer firms post vacancies and the job arrival rate for workers decreases. Referring back to a worker's problem, observe a decrease in the job arrival rate results in lower crime thresholds. Thus, if the deterrence effect of an increase in incarceration policy (the probability of prison) is small enough such that total flows into prison increase, then the equilibrium response of firms to post fewer vacancies can further reduce the deterrence effect of the policy. [18]

5.2 Steady-State Flows:

The equations characterizing the steady-state are as follows:

$$\begin{aligned} p\tau &= u\eta\pi f_u + (1-u-p)\,\eta\pi f_e \\ u\left(\lambda_w + \eta\pi f_u\right) &= p\tau + (1-u-p)\,\delta \\ (1-u-p)\left(\delta + \eta\pi f_e\right) &= u\lambda_w \end{aligned}$$

where p and u are the measure of incarcerated and unemployed individuals, respectively, and f_u and f_e are the probability of committing crime conditional on receiving an opportunity. That is, $f_u = (1 - H(\kappa_u^*))$ and $f_e = (1 - H(\kappa_e^*))$. Then, we have

$$\begin{aligned} u &= \frac{\tau\left(\delta + \eta\pi f_e\right)}{\left(\tau + \eta\pi f_e\right)\left(\lambda_w + \eta\pi f_u + \delta\right) + \eta\pi\left(\delta - \tau\right)\left(f_u - f_e\right)} \\ p &= \frac{\eta\pi\left(f_e + u\left(f_u - f_e\right)\right)}{\tau + \eta\pi f_e} \end{aligned}$$

If we assume that $f_u = f_e = 1$, which is the case when the crime reward distribution is degenerate and the reward is sufficiently large, then we can show that the unemployment rate, which is defined as $\frac{u}{1-p}$ becomes

$$\frac{u}{1-p} = \frac{\delta + \eta\pi}{\lambda_w + \delta + \eta\pi} \tag{7}$$

Proposition 3. *An increase in the probability of getting caught, π, increases the unemployment rate.*

6 Proofs

Proof. From equation (1) we have

$$V_u = \frac{r+\tau}{\tau} V_p \tag{8}$$

[18] This requires that crime increases when the job arrival rate falls. By Proposition 2, this requirement is true with certainty if the job duration is shorter than the prison duration, but may or may not hold otherwise (Proposition 2).

We can express the difference between the value of employment and incarceration by combining equations (1) and (3) :

$$(r + \eta (1 - H(\kappa_e^*)) \pi + \delta) (V_e - V_p) = wh + (\delta - \tau)(V_u - V_p) + \eta \bar{\kappa}_e$$

where $\bar{\kappa}_e = \int_{\kappa_e^*} \kappa dH(\kappa)$. Substituting (4) and (5) into the above equation we get

$$(r + \delta) \frac{\kappa_e^*}{\pi} + \eta (1 - H(\kappa_e^*)) \kappa_e^* = wh + (\delta - \tau) \frac{\kappa_u^*}{\pi} + \eta \bar{\kappa}_e \tag{9}$$

Similarly, using equations (1) and (2), we can express the difference between the value of unemployment and incarceration as:

$$(r + \eta (1 - H(\kappa_u^*)) \pi + \lambda_w + \tau) (V_u - V_p) = b + \lambda_w (V_e - V_p) + \eta \bar{\kappa}_u$$

Again, substituting equations (4) and (5) into the above equation we get

$$(r + \lambda_w + \tau) \frac{\kappa_u^*}{\pi} + \eta (1 - H(\kappa_u^*)) \kappa_u^* = b + \lambda_w \frac{\kappa_e^*}{\pi} + \eta \bar{\kappa}_u \tag{10}$$

Then, equations (9) and (10) give us κ_u^* and κ_e^*. Given these thresholds, we can express the value functions as

$$\begin{aligned} V_p &= \frac{\tau \kappa_u^*}{r\pi} \\ V_u &= \frac{(r + \tau) \kappa_u^*}{r\pi} \\ V_e &= \frac{r\kappa_e^* + \tau \kappa_u^*}{r\pi} \end{aligned}$$

□

Proof. Using implicit function theorem on equation (9) and (10) we have

$$\frac{d\kappa_e^*}{d\pi} = \frac{(\delta - \tau) \frac{d\kappa_u^*}{d\pi} + wh + \eta \int_{\kappa_e^*} (1 - H(\kappa)) d\kappa}{r + \delta + \eta\pi (1 - H(\kappa_e^*))}$$

If $\delta > \tau$, it is immediate to see that $\frac{d\kappa_e^*}{d\pi} > 0$. Using implicit function theorem on equation (10)

$$\frac{d\kappa_u^*}{d\pi} = \frac{\lambda_w \frac{d\kappa_e^*}{d\pi} + b + \eta \int_{\kappa_u^*} (1 - H(\kappa)) d\kappa}{r + \lambda_w + \tau + \eta\pi (1 - H(\kappa_u^*))}$$

Since $\frac{d\kappa_e^*}{d\pi} > 0$, then we also have $\frac{d\kappa_u^*}{d\pi} > 0$. □

Proof. Using integration by parts, we can express equation (9) as

$$(r+\delta)\frac{\kappa_e^*}{\pi} = wh + (\delta-\tau)\frac{\kappa_u^*}{\pi} + \eta\int_{\kappa_e^*}(1-H(\kappa))\,d\kappa$$

Using the implicit function theorem , we have

$$\frac{d\kappa_e^*}{d\kappa_u^*} = \frac{\delta-\tau}{r+\delta+\eta\pi\left(1-H\left(\kappa_e^*\right)\right)}$$

Then, if $\delta < \tau$, we have $\frac{d\kappa_e^*}{d\kappa_u^*} < 0$ and if $\delta > \tau$, we have $1 > \frac{d\kappa_e^*}{d\kappa_u^*} > 0$. Similarly, using equation (10) and the implicit function theorem, we have

$$\frac{d\kappa_u^*}{d\lambda_w} = \frac{\kappa_e^* - \kappa_u^*}{r+\lambda_w+\tau+\eta\pi\left(1-H\left(\kappa_u^*\right)\right)-\lambda_w\frac{d\kappa_e^*}{d\kappa_u^*}}$$

We know that $wh > b$ implies $\kappa_e^* - \kappa_u^* > 0$. Since we also know that $\frac{d\kappa_e^*}{d\kappa_u^*} < 1$, then we have $\frac{d\kappa_u^*}{d\lambda_w} > 0$. This result is independent of the relation between δ and τ. But, depending on the relation between δ and τ we have two opposite results. If $\delta < \tau$, since we have $\frac{d\kappa_e^*}{d\kappa_u^*} < 0$, then we get $\frac{d\kappa_e^*}{d\lambda_w} < 0$. Otherwise, if $\delta > \tau$, we have $\frac{d\kappa_e^*}{d\lambda_w} > 0$. □

Proof. Equation (7) shows the relation between the unemployment rate and π. As π increases unemployment increases. Moreover, an increase in π decreases the offer arrival rate, which further increases the unemployment rate. □

References

Bhati, A. S. (2010-07-29): "Criminal Recidivism in a Large Cohort of Offenders Released from Prison in Florida, 2004-2008," .

Bonczar, T. P. (2003): "Prevalence of Imprisonment in the US Population, 1974-2001," Special Report NCJ 197976, Bureau of Justics Statistics.

İmrohoroğlu, A., A. Merlo, and P. Rupert (2004): "What Accounts for the Decline in Crime," *International Economic Review*, 45(3), 707–729.

Kitao, S., L. Ljungqvist, and T. J. Sargent (2017): "A life-cycle model of trans-Atlantic employment experiences," *Review of Economic Dynamics*, 25 (April), 320–349.

Ljungqvist, L., and T. J. Sargent (1998): "The European unemployment dilemma," *Journal of political Economy*, 106(3), 514–550.

Michaud, A. M. (2018): "A quantitative theory of information, worker flows, and wage dispersion," *American Economic Journal: Macroeconomics*, 10(2), 154–83.

Neal, D., and A. Rick (2014): "The prison boom and the lack of black progress after Smith and Welch," Discussion paper, National Bureau of Economic Research.

of Justice. Office of Justice Programs., U. S. D. (2014-12-05): "Recidivism of Prisoners Released in 1994.," .

Perkins, C. A. (1992): "National Corrections Reporting Program: 1986," National Corrections Reporting Program Series NCJ 132291, Bureau of Justics Statistics.

Pettit, B. (2012): *Invisible Men: Mass Incarceration and the Myth of Black Progress*. Russell Sage Foundation.

Pfaff, J. F. (2011): "The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices," *American law and economics review*, p. ahr010.

Raphael, S., and M. A. Stoll (2009): "Why are so many Americans in prison?," *Do prisons make us safer*, pp. 27–72.

Ruggles, S., e. a. (2004): "Integrated public use microdata series: Version 3.0," .

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics (2013): "National Corrections Reporting Program Series, 1983-2012: [United States]," .

	(1) Wagem_ b/se	(2) resid b/se	(3) Wagem_ b/se	(4) resid b/se	(5) Wagem_ b/se	(6) resid b/se
Age 25-34	0.1301		0.1290		0.1211	
	(0.00)		(0.00)		(0.00)	
Age 35-64	0.2140		0.2107		0.1982	
	(0.00)		(0.00)		(0.00)	
Non-Employed (mo)	-0.0044		-0.0047			
	(0.00)		(0.00)			
Non-Employed (mo)^2	-0.0000		-0.0000		0.0001	
	(0.00)		(0.00)		(0.00)	
Lagged Resid ln(wage)		0.9844		0.9844		0.9843
		(0.00)		(0.00)		(0.00)
Jail Last Yr			-0.0951		-0.0310	
			(0.01)		(0.01)	
Non-Participant (mo)					-0.0036	
					(0.00)	
Non-Participant^2					0.00005	
					(0.00001)	
Unemployed (mo)					-0.0075	
					(0.00)	
Unemployed (mo)^2					0.0002	
					(0.00)	
Constant	1.6647	0.0005	1.6721	0.0006	1.6795	0.0006
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Individuals	1,262	1,262	1,262	1,262	1,262	1,262
Monthly Observations	182071	188364	182071	188364	182071	188364

Standard errors in parentheses

Non-employment, Non-participation, and Unemployment are total months in past two years.

Table 6: Wage Regressions table

Status at $t-1$	Employed			Non-Employed	Unemployed	Non-Participant
Status at t	NE	U	N	E	E	E
By Age						
18-24	1.91	1.01	0.89	3.20	4.12	2.43
25-34	1.04	0.51	0.53	2.56	3.87	1.83
35-50	0.53	0.23	0.31	1.09	1.89	0.75
Total (18-50)						
Never Incarcerated†	1.05	0.54	0.51	2.81	3.82	1.94
Incarcerated w/in last year‡	3.06	1.18	1.88	1.21	2.33	0.97
Total	1.19	0.59	0.60	2.44	3.54	1.67

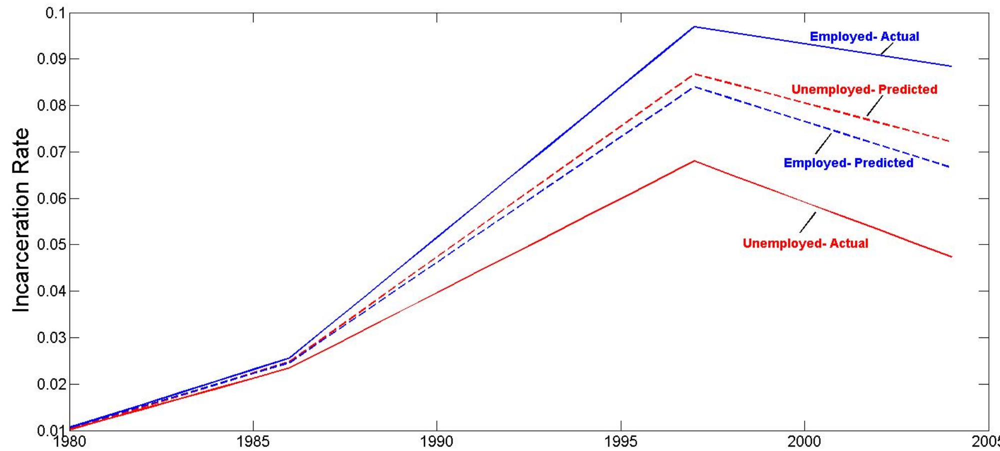
† Never observed as incarcerated in entire sample: age 14-19 to age 50.

‡ We have 36,002 observations of employment status for 257 individuals incarcerated within a last year.

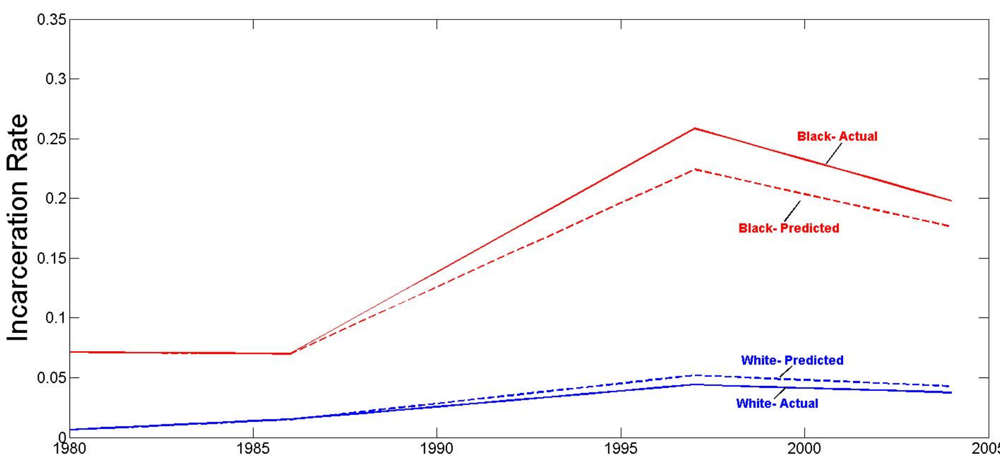
Table 7: Weekly Employment Transition Rates

Parameter	Explanation	Value
η^1	crime arrival rate	0.05%
c	vacancy cost	68.84
δ	separation shock	1.36%
$\mu^{e,2}$	human capital mean-middle employed	0.07
$\mu^{e,3}$	human capital mean-old employed	0.06
$\mu^{u,1}$	human capital mean-nonemployed	0.39
ζ^3	rehabilitation shock	0.35%
ρ_h	human capital persistency	0.94
σ_h	human capital shock std	0.25
ν	prob of being high criminal	0.15
$\eta_a^{1,hc}$	high criminal crime arrival rate	0.65
μ^k	mean crime reward	1.44

Table 8: Calibrated Parameters - Violent: The Table shows the internally calibrated parameters of model with only violent crime.

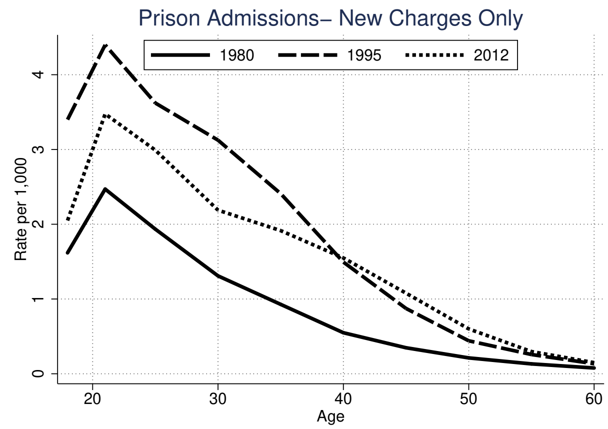


(a) Predicted and Actual Rates by Employment Status

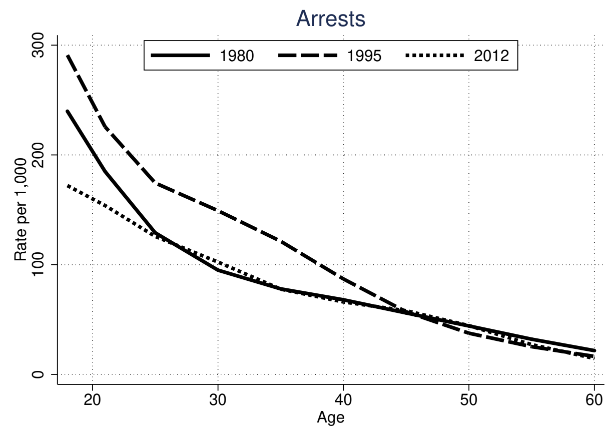


(b) Predicted and Actual Rates by Race

Figure 9: Actual values calculated from BJS Prison Surveys and Census data. Predicted values calculated by holding admission rates fixed to 1979 levels, and raising rates by the same proportion for each group, adjusting for demographics (race/or employment status and age).

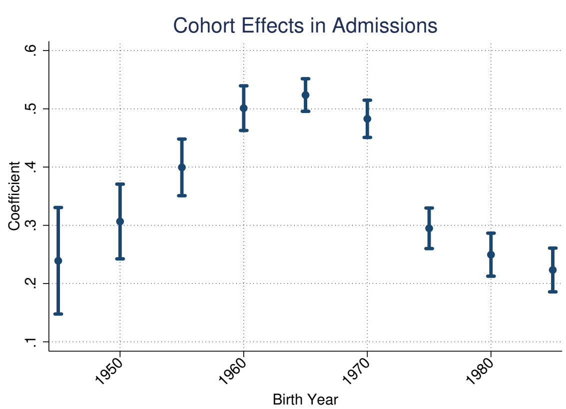


(a) New Prison Admission Rate by Age

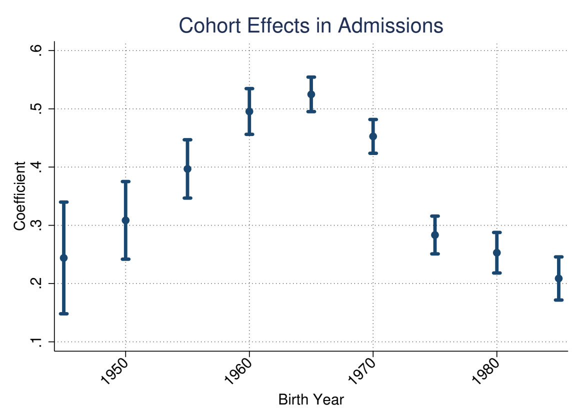


(b) Arrest Rate by Age

Figure 10: Age Profiles of Crime and Incarceration

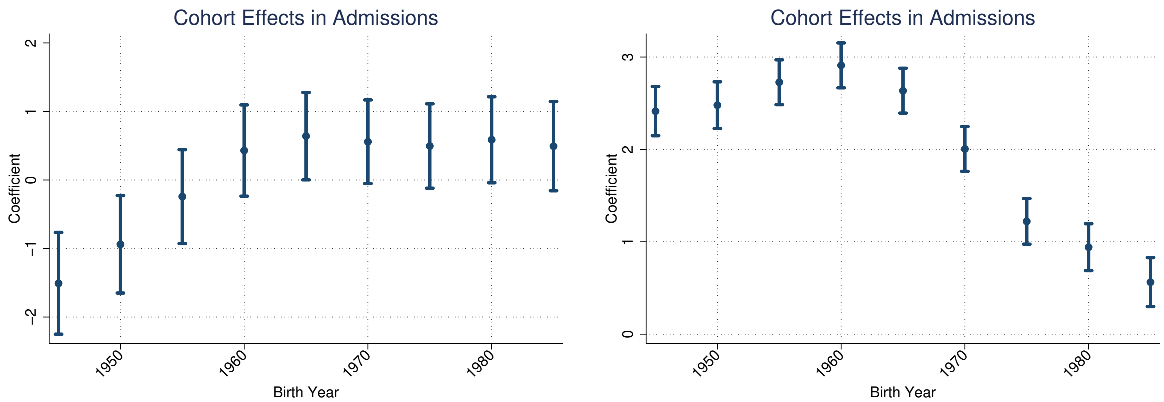


(a) Base: Time-varying Age Profile



(b) Time-invariant Age Profile $\beta^Y = 0$

Figure 11: **Estimated Cohort Effects:** Prison admissions from National Corrections Reporting Program Data and restricted to admissions on new charges only.



(a) Linear- Cohort and Age Effects Only

(b) Linear- Cohort and Time Effects Only

Figure 12: **Estimated Cohort Effects:** Prison admissions from National Corrections Reporting Program Data and restricted to admissions on new charges only.

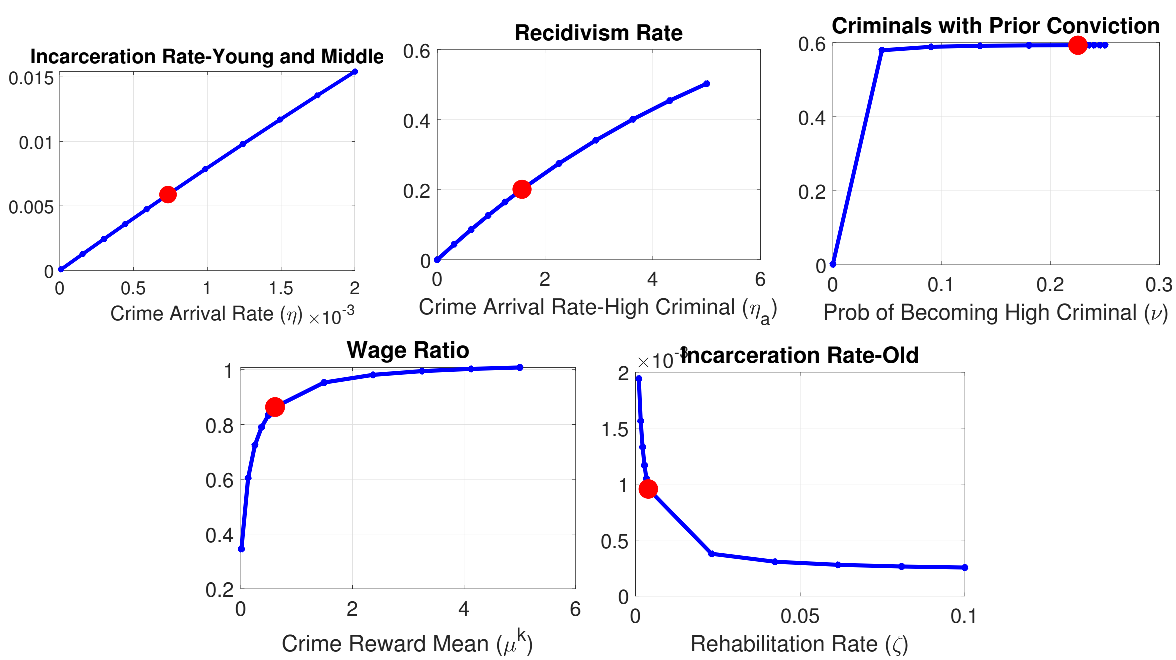


Figure 13: **Crime Related Moments:** The dot shows the calibrated value of the parameter.

Figure 14: **Labor Market Related Moments:** The dot shows the calibrated value of the parameter.

Moment	Data	Model
Incarceration - young and middle	0.44%	0.44%
Incarceration - old	0.09%	0.09%
Unemployment duration	20 weeks	20 weeks
Employment rate - young and middle	76.2%	76.9%
Recidivism rate (1 year)	13.5%	13.4%
Wage Ratio (criminals vs non)	87.7%	87.7%
criminal with prior	53.7%	54.7%
Regression coefficient-β^M	0.13	0.13
Regression coefficient-β^O	0.21	0.21
Regression coefficient-β^N	-0.005	-0.005
income persistency	0.96	0.96
income std	0.20	0.20

Notes: The Table shows a comparison of empirical and simulated moments with only violent crime.

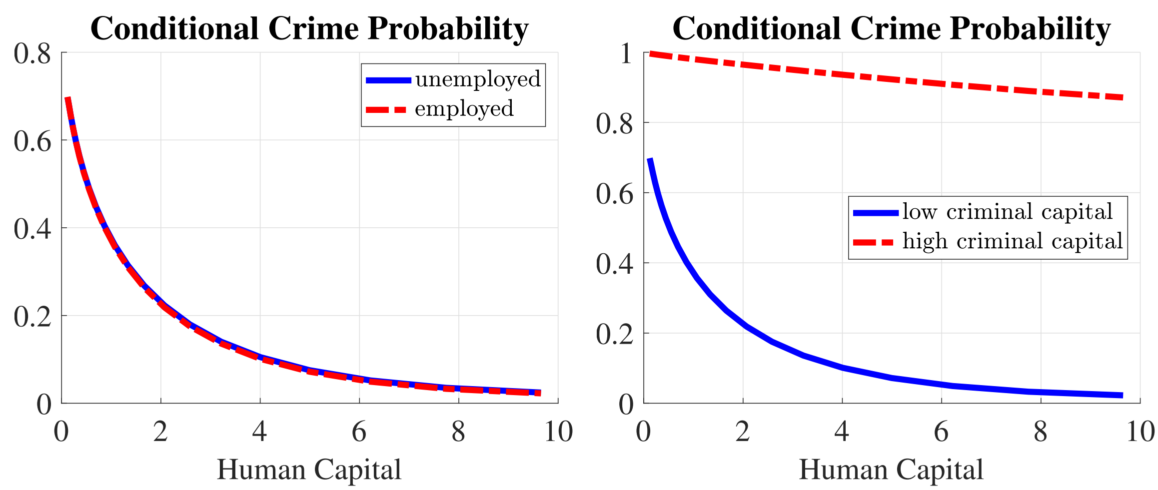
Table 9: Model Match

	Criminals	Overall
Employment rate	75.6%	76.9%
Human capital	0.97	1.18
Prison Flag	54.7%	1.5%
Young and middle population	74.1%	34.0%

Table 10: Characteristics of Criminals

	Estimate	SE	tStat	pValue
Age 25-34	0.07	0.01	14.27	0.0
Age 35-50	0.23	0.01	56.19	0.0
Prison Flag	0.08	0.01	6.32	0.0
Employed	-0.003	0.003	-0.83	0.41
ln(wage)	-0.12	0.0	-65.48	0.0
Constant	0.25	0.01	55.1	0.0

Table 11: Crime Elasticities



(a) Labor Status

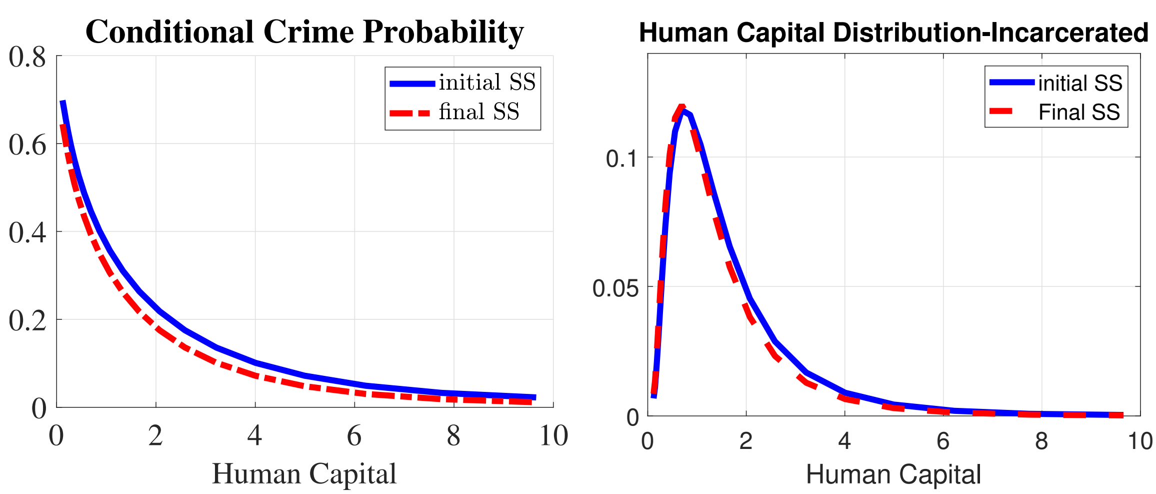
(b) Criminal Capital

Figure 15: **Determinants of Crime:** The figure shows model generated crime probability conditional on receiving an opportunity as a function of human capital, labor market status and criminal capital for a middle-age agent.

Steady-State Variables	SS1 $\pi = 0.5\%$	SS2 $\pi = 2.9\%$
Incarceration	0.44%	1.45%
Crime Rate	0.3%	0.1%
Employment rate	76.9%	74.7%
Recidivism rate-1 year	13.4%	66.8%
Criminals with prison flag	54.6%	80.2%
Frac w/ high criminal capital	0.8%	0.7%
With prison flag	1.5%	2.2%
Share committing 95% of crimes	0.6%	0.05%
Wage ratio	87.7%	79.1%

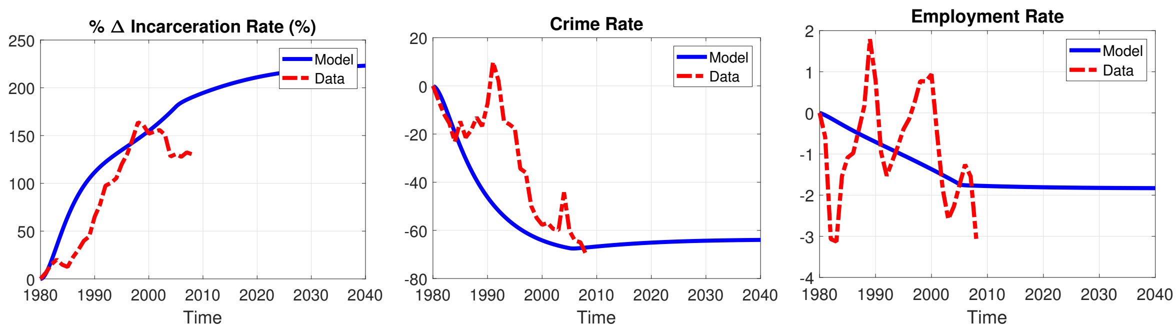
Notes: The Table shows a comparison of two steady states, one with $\pi = 0.5\%$ and one with $\pi = 4.5\%$, productivity 20% lower and crime reward 150% higher.

Table 12: Steady-State Comparison



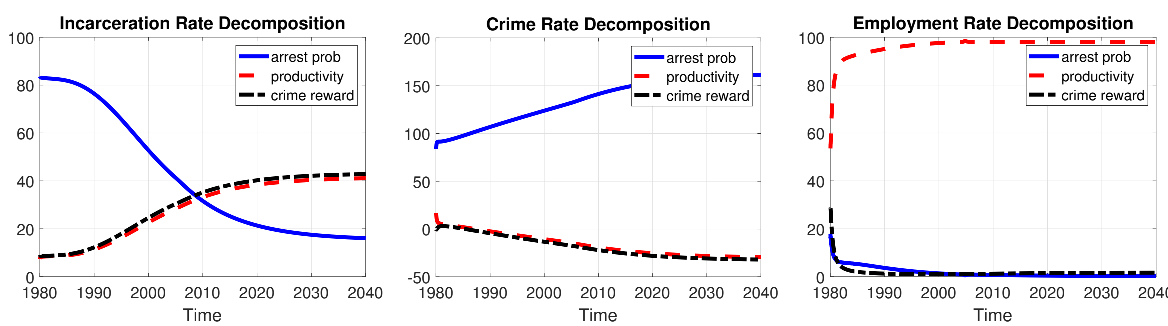
(a) Policy (b) Distribution

Figure 16: **Steady-State Comparison:** The left panel shows model generated crime probabilities conditional on receiving an opportunity as a function of human capital for a middle-age employed individual with low criminal capital and no prison flag across the initial and the final steady-states. The right panel plots the distribution of human capital among the incarcerated across the initial and the final steady-states.



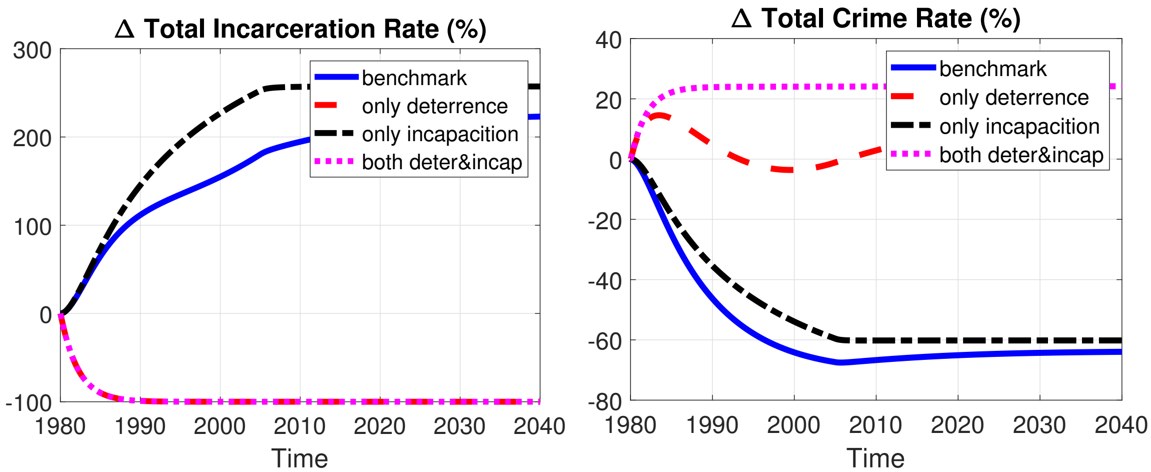
(a) Incarceration (b) Crime (c) Employment

Figure 17: **Transitional Dynamics - Model vs Data:** The figure shows the evolution of incarceration rate, crime rate and employment rate along the transition. The left panel plots the total incarceration rate. The middle one plots the total crime rate and the right panel plots the employment rate relative to their initial steady-state levels. The solid lines correspond to their model counterparts whereas dashed lines correspond to the data.



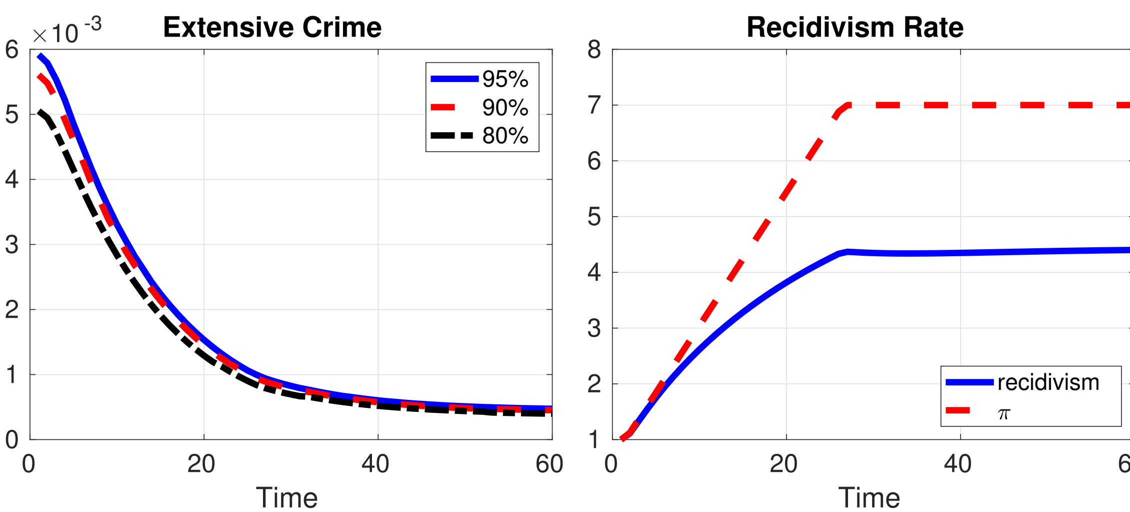
(a) Incarceration (b) Crime (c) Employment

Figure 18: **Transitional Dynamics - Shapley-Owen Decomposition:** Solid lines show the contribution of the change in incarceration probability, dashed line shows the contribution of the change in the productivity, and finally the long-dashed line shows the contribution of the change in the crime reward. The left panel is for the incarceration rate, the middle panel is for the crime rate and the right panel is for the employment rate.



(a) Incarceration (b) Crime Rate

Figure 19: **Incapacitation vs Deterrence:** The figures compare the evolution of incarceration and crime rate along the transition without incapacitation or deterrence effects. The solid line is the benchmark economy. The long dashed line is the economy when incapacitation is eliminated. The dashed line is the economy when all decision rules of the individuals and firms kept at the initial steady-state levels.

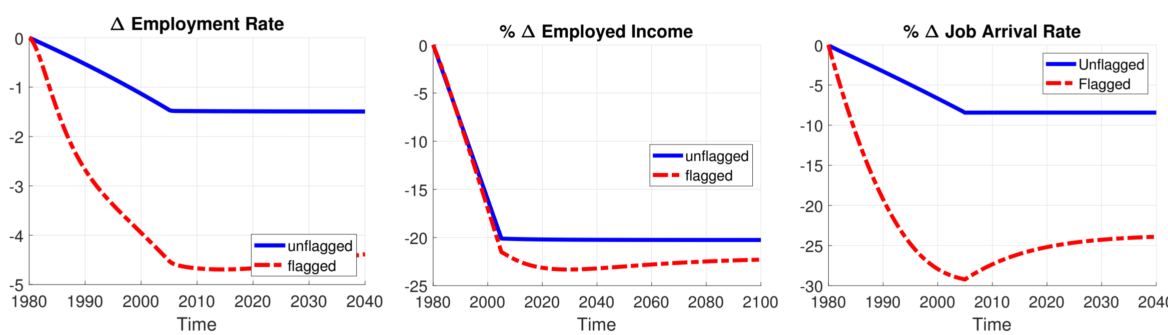


(a) Extensive Crime (b) Recidivism

Figure 20: **Extensive Crime and Recidivism:** The left plots the measure of individuals committing certain shares of aggregate crime along the transition. The solid line is for 95% of crimes, the dashed line is for 90% of crimes and the long-dashed line is for 80% of crimes. The right panel plots the one year recidivism rate together with the arrest probability along the transition. Both recidivism rate and arrest probability are normalized to their initial steady-state level.

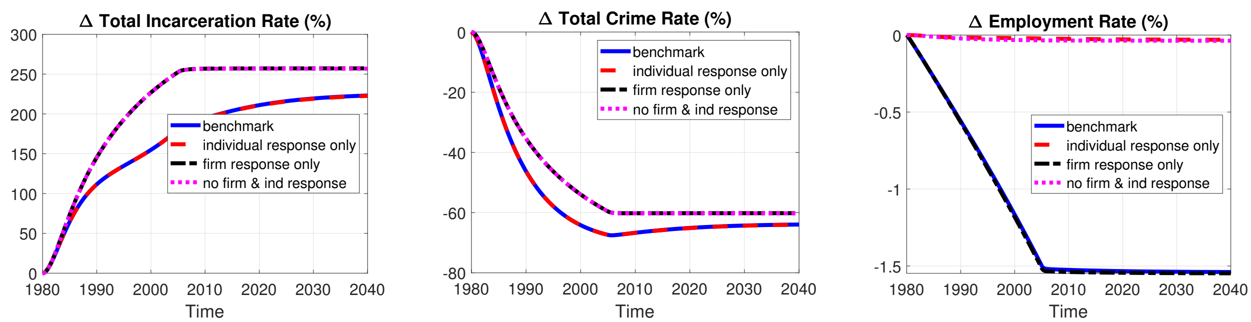
Total 3-year Re-imprisonment			
Age	1983	1994	2000-2003*
18-24	64.0	41.0	48.8
25-34	32.6	40.3	49.6
35-64	27.0	35.6	44.3
Total (18-64)	30.7	39.3	47.7
Expected % of Population Incarcerated by age 35			
	Year of Birth		
	1974-1979	1994	2000-2003*
	1.7	4.0	4.7

Table 13: Upper panel: 3-year Re-imprisonment Rate on a New Felony Charge, 1983 & 1994 Recidivism of Prisoners Released Series (of Justice. Office of Justice Programs. (2014-12-05)); *2000-2003: Florida only, (Bhati (2010-07-29)). Lower panel: estimated from Bonczar (2003) and authors' calculations in NCRP.



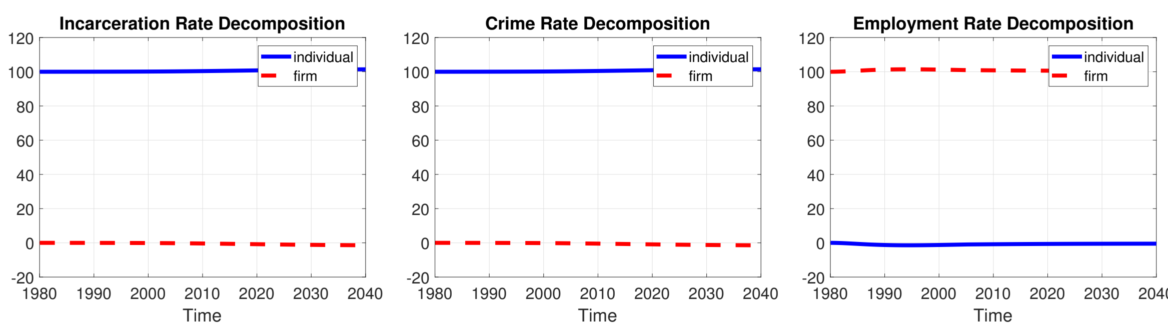
(a) Employment (b) Income (c) Job Arrival Rate

Figure 21: **Employment, Income and Job Arrival Rates across Different Groups:** The figures show the evolution of employment rate and income for individuals with (flag) and without (unflagged) prior incarceration record. The left panel is for employment, the middle panel is for income dynamics, and the right panel is for the job arrival rate of the middle-age individuals. All are changes in percentage points relative to the initial steady-state level.



(a) Incarceration (b) Crime Rate (c) Employment

Figure 22: **Transitional Dynamics: Policy Decomposition:** The figures show the decomposition of the incarceration, crime rate and employment along the transition. The solid line is the benchmark economy. The dashed line is the economy when firms keep the same job creation level. The long dashed line is the economy when individuals keep their criminal policy as in the first steady-state. Lastly, the dotted line is the economy when firm keep the same job creation level, individuals keep their crime choices as in the first steady-state.



(a) Incarceration (b) Crime Rate (c) Employment

Figure 23: **Transitional Dynamics: Policy Decomposition:** The figures show the Shapley-Owen decomposition of individual and firm policy functions. The solid line is the contribution of individual criminal policy and the dashed line is the contribution of firm vacancy policy. The left panel is for the incarceration rate, the middle panel is for the crime rate and the right panel is for the employment rate.

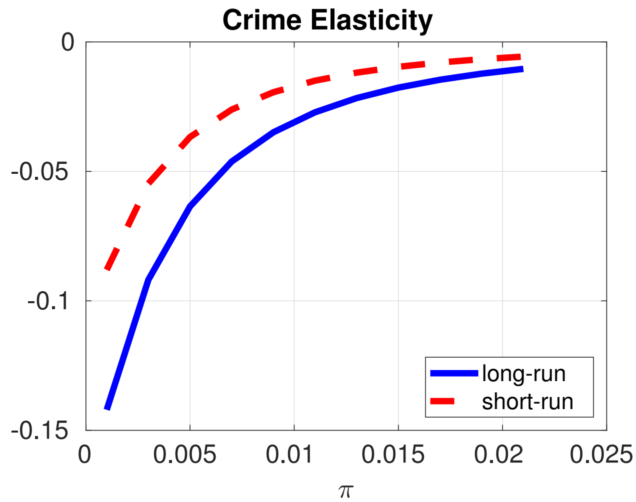


Figure 24: **Crime Elasticities:** The figure plots the aggregate crime elasticity both in the short-run and the long-run.

	M0	M1	M2	M3	M4	M5	M6
η^1	0.08%	2.95%	4.79%	0.05%	3.13%	0.2%	0.75%
c	77.05	82.47	45.78	85.09	73.80	74.0	71.51
δ	1.41%	1.27%	1.37%	1.45%	1.30%	1.40%	1.48%
$\mu^{e,2}$	0.11	0.09	0.03	0.13	0.07	0.1	0.1
$\mu^{e,3}$	0.14	0.13	0.004	0.21	0.08	0.12	0.13
$\mu^{u,1}$	0.17	0.23	0.78	0	0.34	0.23	0.20
ζ^3	0.39%	0.4%	0.54%	0.39%	0.1%	0.35%	0.18%
μ^k	0.44	0.03	0.39	3.50	0.1	0.1	2.88
$\eta_a^{1,hc}$	0.94	0	1.4	0.94	0	0.97	1.0
ν	0.16	0	0	0.10	0	0.04	0.15

Table 14: Calibrated Parameters - Alternative Models: The Table shows the internally calibrated parameters of the alternative models.