

1. Additional Details on the 1980s Prison Boom

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 the following experiment 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 admission rate for each group in the first BJS Prison Survey: 1979.¹ 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 calculate λ^y to satisfy:

$$TotAdmitRate^y = \sum_{r,a,e} \lambda^y \phi_{r,a,e}^{1980} \pi_{r,a,e}^y$$

where $\phi_{r,a,e}^{1980}$ is the 1980 admission rate for each demographic cell (r, a, e) where $r \in \{black, white\}$ is race, $e \in \{employed, nonemployed\}$, $a \in \{18 - 24, 25 - 34, 35 - 54\}$. Then, $\pi_{r,a,e}^y$ 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 cell's share in the total demographics. We then calculate predicted rates for each age group a as:

$$AdmitRate_a^y = \sum_{r,e} \lambda^y \phi_{r,a,e}^{1980} \pi_{r,a,e}^y$$

Figure 1 shows our results. It shows that admissions rates for the youngest group, ages 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, by the late 1990s, the oldest age group, 35–54, has rates substantially higher than predicted.² These patterns are informative about the dynamic role of deterrence and are largely consistent with the theory developed in the paper.

¹ We limit our analysis to males. We use SEER data to measure total population counts for each cell. We adjust this data to be only for those with high school diplomas or less using decennial census data and interpolating linearly for non-decade years.

² The BJS Prison Survey is only conducted every seven years providing limited data points for this analysis.

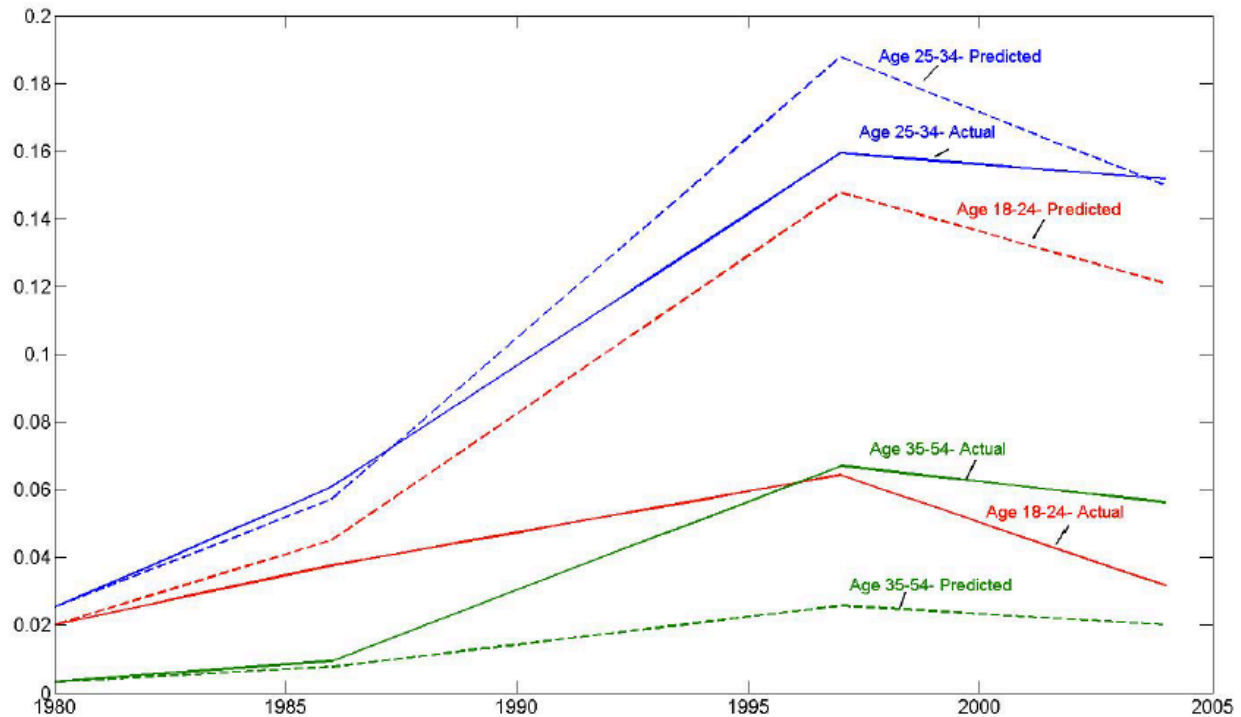


Figure 1. Actual values calculated from Bureau of Justice Statistics prison surveys (US Department of Justice 2010) and census data. Predicted values calculated by holding admission rates fixed to 1979 levels and raising rates by the same proportion for each age group, adjusting for demographics (race and employment).

The left pane of figure 2 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 can be accounted for by pre-existing differences prior to 1980.

Stability in Length of Time Served

Prior to the late 1990s, state-level inmate records were neither uniform nor well kept. There was no reporting requirement to federal authorities. 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 years in 1984 and 2.73 years in 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 percent 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 a great deal of 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 percent of total state and federal admissions at their peak in the 1990s and represent less than 20 percent of admissions in 2010.³ 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.

The importance of all crime categories to the trend motivates the decision to include all crimes in our analysis. We also choose a parsimonious approach in both the model and the data targets. That is, we make no distinction between the four major categories of crime: violent, property, drug, and other crime. This is because criminals often operate in more than one category of crime, and there is great heterogeneity in these patterns.

2. 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.

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

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 $g'^X(\pi) < 0$ (i.e., that punitive policy deters).

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

³ Calculation from Bureau of Justice Statistics data.

prison experience yesterday has on crime today. For what follows, we assume that β may be 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.⁴ 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 2.1, 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 2.1 (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 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 regardless 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 addresses how a change in punitive policy (π) affects aggregate crime and incarceration rates.⁵

⁴ This model will not be able to reconcile the monotonic decline of crime with the non-monotone shape of incarceration. This is 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.

⁵ It is assumed the maximum age is M , but these proofs will also apply to $\lim_{M \rightarrow \infty}$, so long as parameters are appropriately restricted such that crime is finite.

Proposition 2.2 (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 $\sum_{a=0}^M A_a$, and
- 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:

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

By inspection, the right-hand side is increasing in both β and the sequence A_a , and so the inequality is 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.

Corollary 2.3 (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 $\sum_{a=0}^M A_a$, and
- increasing(decreasing) in π if $g''^X(\pi) >> 0$ ($g''^X(\pi) << 0$).

Proof. Omitted.

We now consider the effect of a policy change along the transition. The main result, summarized in proposition 2.4, explicates the existence of cohort effects.

Proposition 2.4 (The cohort born immediately before an increase in π has higher age-specific 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:*

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

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

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

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

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

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

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

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 $X_{t-j+s,s}$. We want to show the following relationship:

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

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.

Corollary 2.5 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 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 2.5 (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:*

- The transition path for crime is non-monotone iff

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

- 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 for the period after the policy change. Then:

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

That $C_0 > C(\pi_0)$, follows directly from proposition 2.4. Simple algebra comparing C_0 and $C(\pi_1)$ provides the necessary and sufficient condition provided in the statement of this corollary.

Corollary 2.6 (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 2.4.

Identification

The manner in which all four effects enter this simple model is crucial for identification. Conceptually, the age effect impacts growth rates while the time and cohort effects impact levels. Finally, we have shown that the age profile can also be changed overtime by changes in the policy π_t . This motivates a two-step procedure to separate these components into growth rate and level effects. We begin by estimating the growth rate effects, generating residuals from predicted values, and then estimating time and cohort effects on these residual levels.

- **Step 1 – Growth effects.** The crime rate of cohort j in periods t and $t - 1$ are:

$$\begin{aligned} c_{j,t-1} &= \Pi_{s=j}^{t-1} [(\phi + \beta \pi_{s-1})] * g^x(\pi_j) \times A_{t-1-j} \\ c_{j,t} &= \Pi_{s=j}^t [(\phi + \beta \pi_{s-1})] * g^x(\pi_j) \times A_{t-j} \end{aligned}$$

- Taking logs of the growth rate $\frac{c_{j,t-1}}{c_{j,t}}$, we have:

$$\begin{aligned} \log(c_{j,t-1}) - \log(c_{j,t}) &= \log(\phi + \beta \pi_{t-1}) + \log\left(\frac{A_{t-1}}{A_t}\right) \\ &= \beta_g \mathbf{I}_t + \beta_a \mathbf{I}_a + \epsilon_{j,t} \end{aligned}$$

- The last line shows the regression strategy, adding $\epsilon_{j,t}$ as the error term. We will regress the growth rate of cohort-level crime upon dummies for time \mathbf{I}_t and age \mathbf{I}_a . Since we only have two of the three (age, cohort, and time) effects, we do not suffer from perfect co-linearity.

- **Step 2 – Use predicted crime growth rates to predict cohort’s life-cycle crime profiles.** Next, generate predicted values $\hat{c}_{j,t-1}$ by multiplying the crime rate at age 18 by the predicted age-cross-time growth rates estimated in this equation. Then use the residuals $\tilde{c} = c_{j,t-1} - \hat{c}_{j,t-1}$ in the following equation to estimate level effects.
- **Step 3 – Level effects.** The cohort specific level, $g^X(\pi_j)$, did not factor into the growth equation but factors into the level. We now estimate that effect as well as contemporaneous time level effects, which affect all individuals equally regardless of their past crime behavior, using the following regression:

$$\tilde{c}_{j,t} = \beta_t \mathbf{I}_t + \beta_c \mathbf{I}_c + \epsilon_{j,t}$$

- Here we again break the co-linear problem in separating age, cohort, and time effects by only featuring two out of the three: time $\beta_t \mathbf{I}_t$ captured by year dummies and cohort $\beta_c \mathbf{I}_c$ captured with cohort dummies.

A summary of the procedure is to first estimate time-varying growth rate effects, then generate residuals from the predicted values, and finally estimate time and cohort effects on these residuals. The first stage (dealing with growth rates) interprets age as the *time-invariant growth and decay* in crime and incarceration for individuals of all cohorts as they get older. The *time-variant change in growth and decay* in crime and incarceration for all individuals of all cohorts and ages from one year to the next is interpreted as the effect of policy changes. The second stage of our regression deals with levels. We estimate the time and cohort effects to best match the life-cycle profile of cohorts. The cohort component is a constant initial level for each cohort from which the life-cycle profile is created using the first stage regression. In other words, it shifts a single cohort’s age profile from the first stage up or down. The further time-level effects fill in gaps for years when individuals of all ages increase crime. The critical difference between the time effects in the first and second stages is that in the first stage, individuals are affected in proportion to their prior behavior whereas in the second stage it is a common level increase for all individuals.

Data and Construction of Outcome Variables

Our two measures of involvement with the criminal justice system are arrest rates and prison admission rates. An important difference between the measures is that arrest rates span all offenses—misdemeanors and felonies—and need not be accompanied by an actual conviction. In comparison, admissions will be defined as a conviction for a new crime (not just a violation of probation or parole) accompanied by a state or federal prison sentence—usually more serious felonies receive a sentence of a year or more.

Arrest Rates

Arrests rates are reported directly by the Federal Bureau of Investigation as part of its Uniform Crime Reporting Program. Rates are reported by coarse age groups, gender, and race. They are also categorized by type: violent crime (murder, forcible rape, robbery, assault, etc.); property crime (burglary, larceny, arson, etc.); and other (fraud, weapons violations, prostitution, drug violations, etc.). At the time of writing, these data were available from 1980 to 2015.

Admission Rates

We calculate admission rates by combining data on admissions from the restricted dataset: National Corrections Reporting Program 1983–2000 (NCRP), accessed through the US Department of Justice

(2010), and the US decennial census and current population survey, accessed through IPUMS (Ruggles 2004).

The NCRP is a restricted access data set maintained by the Bureau of Justice Statistics (BJS). It is a compilation of prison admission, release, parolee, and prison stock data 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 consistency checks from the audit study of Neal and Rick (2014). Next, we include only states that have a consistent time series from 1984 onwards. This leaves us with 12 state prison reports: California, Georgia, Illinois, Michigan, Minnesota, New Jersey, New York, North Dakota, Ohio, South Carolina, Washington, and Wisconsin. Within these states, our sample is limited to males entering prison in the calendar year with a new conviction. This means that we exclude anyone entering prison without a new conviction (e.g., returned escapees or parole violators). We match state-age cells with population data from the census for the same cells. We linearly interpolate population estimates between 1980 and 1990 and between 1990 and 2000. We use CPS annual data following 2000. Finally, we calculate admission rates by crime category according to two metrics: (i) category of offense with the longest sentence; 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 three 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.

2.2 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 3 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-1980s 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 younger ages. Finally, the temporary bumps upward 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:

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

$$\text{st} \quad \beta^Y = 0 \quad \text{if } a < 26$$

$$\beta^Y \geq 1 \quad \text{if } a > 25$$

The independent variables D^T , D^C , and D^A are respectively dummies for time, age, and cohorts.⁶ 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 1 so that it only captures the flattening of the life-cycle profile.

⁶ The data for arrests only provides five-year age bands. Accordingly, we measure cohorts and time in five-year intervals.

We estimate the regression equation using non-linear least squares. The results from the full regression specification and from an alternative specification with $\beta^Y = 0$ for all ages are presented in table 1 and figure 2.2. The interpretation of the latter is to hold the shape of the age profile fixed over time instead of allowing it to flatten.

Table 1: Regression Results: Prison Admission on New Charges

	Base: Time-varying Age Profile	Robustness: Time-invariant Age Profile
β^T_{1990}	0.704 **	0.378 **
β^T_{1995}	0.914 **	0.902 **
β^T_{2000}	0.913 **	0.910 **
β^T_{2005}	0.734 **	0.758 **
β^T_{2010}	0.687 **	0.673 **
β^C_{1945}	0.239 *	0.244 *
β^C_{1950}	0.307 **	0.308 **
β^C_{1955}	0.399 **	0.397 **
β^C_{1960}	0.501 **	0.495 **
β^C_{1965}	0.524 **	0.525 **
β^C_{1970}	0.483 **	0.453 **
β^C_{1975}	0.295 **	0.283 **
β^C_{1980}	0.250 **	0.253 **
β^C_{1985}	0.223 **	0.209 **
β^A_{15-19}	2.836 **	2.823 **
β^A_{20-24}	3.678	3.678
β^A_{25-29}	2.621 **	2.729 **
β^A_{30-34}	2.029 **	2.108 **
β^A_{35-39}	1.546 **	1.605 **
β^A_{40-44}	1.104 **	1.142 **
β^A_{45-49}	0.738 **	0.760 **
β^A_{50-54}	0.449 **	0.460 **
β^A_{55-59}	0.271 **	0.278 **
β^Y_{1990}	0.341 **	
β^Y_{1995}	2.26 *	
β^T_{2000}	0.136	
β^T_{2005}	0.267 *	
β^T_{2010}	0.123	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

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

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

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 2.2. The cohort effects display either no non-monotone effects 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

upward trend in admission rates over the time period. In panel (b), there are no age effects and so the older and younger cohorts that are only seen for part of their life at older or 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 disproportionately 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—is a much better fit.

3. Estimation

3.1 Probability of Incarceration Conditional on a Crime (π)

Choosing a value of π for the initial steady state and calculating the change in π during the 1980s is not completely straightforward. This is because not all crimes are reported to the police. The closest study to ours, Imrohoroglu, 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. This is not an apt strategy for our paper since we would like to include all crimes to provide a complete view of the total criminal justice system and how it has changed overtime. Specifically, we cannot use the clearance rate for drug crimes. The clearance rate for these crimes is not reported since the incidence of these and other victimless crimes are not reported to the police. However, we do have complete information on arrests, convictions, and incarceration on all crimes by category for cases processed in state courts, through the Bureau of Justice Statistics “Felony Sentences in State Courts” series published for the years 1986 to 2009 (1974, 1979, 1986, 1991, 1997). These reports collect data from a subset of US counties. For our analysis we use the years 1986 to 2002—the longest time period for which these reports also include national estimates.

To get a clearance rate for drug trafficking crimes, we impute the total crimes by assuming that the clearance rate for drug trafficking crimes is equal to the average clearance rate for crimes in the categories of murder, robbery, aggregated assault, and burglary. From there we calculate the number of incarcerations per crime in the usual way: by dividing total admissions in these categories by total reported crimes; but we also include total admissions to federal prisons in these years in our count of admissions.⁷ Using this procedure, we arrive at a value of $\pi = 2.4$ for 1986 and a value of $\pi = 7.4$ for 2002. Since the increase in prison admissions started in the late 1970s, we round the initial π down to 2, which is also consistent with the estimations of Pettit (2012). Since the admission rates to prison peaked prior to 2002, we round the later π up to 8.

The rates at each stage are provided in table 2 to provide comparison between the imputed and non-imputed results and to understand at what point in the criminal justice process the π is changing. These figures bolster our claim that there has been a large increase in π . The listed average rates of conviction/arrest and incarceration/conviction do not contain imputed values and have almost doubled for every category over this time period. This conversion also strengthens the claim that policy has changed substantially over the period.

⁷ While federal prisons do hold prisoners convicted of crimes in other categories, such as immigration, the majority of federal prisoners are serving sentences related to drug trafficking and weapons crimes.

Table 2: Source: BJS- Felony Sentences in State Courts Series

1983	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
Drug Trafficking	15.6*	41.2	26.2	10.8*
Total	15.6*	29.6	21.7	2.4*†
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
Drug Trafficking	20.3*	79.9	54.3	11.0*
Total	20.3*	46.9	33.6	7.8*†

¹ * imputed value;

² † also include Federal Prison Admissions;

³ Arrests are adults only and Convictions are felony convictions only.

3.2 Estimation Procedure

The estimation procedure is a mixture of simulated method of moments (SMM) 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 $Y = \{\eta, c, \delta, v, v_p, \eta^{hc}, \xi, \psi_e, \psi_u, \varsigma, \rho, A\}$ as the set of these parameters. Among these parameters, A is a residual parameter. Once the rest of the parameters are determined, A solves the following equation:

$$A = \frac{1}{\eta} \frac{\partial EV_u^{1,0}(h_0, \eta)}{\partial \eta}$$

which basically guarantees that all individuals choose the crime rate η when they start the economy by solving their early life choice problem. This leaves us 11 parameters to be estimated. We estimate these parameters by minimizing equally weighted squares 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, Y solves:

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

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. Targeted and Non-targeted Statistics

4.1 Employment and Wages: National Longitudinal Study of Youth 1979 (NLSY79)

We use data from the July 18, 2013, release of the NLSY79. The NLSY79 includes a nationally representative panel of respondents that were 14 to 22 years old in 1979. Respondents were surveyed annually from 1979 to 1994 and biannually thereafter. The sample is restricted to black or white males who did not graduate high school by age 25.⁸

The NLSY79 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. Incarceration status is observed once per year after 1980 and asked retrospectively in 1980.

In our model, all jobs are found through search, and there is no intensive margin. Accordingly, we define employment in the NLSY79 sample as any non-self-employed job worked a median of 35 to 100 hours per week over the employment relationship. We match each job to its characteristics using the NLSY79 Employer History Roster (EHR). Hourly wage for each job in each week is also taken from the EHR.⁹ We use CPI to calculate wages in 1987 dollars and exclude wages less than \$2 or greater than \$200 per hour as missing.

Wage Regression

Our theory of wage dynamics is a Ben-Porath 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 pre-determined grid. To calibrate the grid points and the transition probabilities, we follow Kitao, Ljungqvist, and Sargent (2017)—a quantitative paper that introduces a quadratic life-cycle wage profile into the Ljungqvist and Sargent (1998) framework. Figure 6 below confirms that our sample of interest also exhibits a quadratic life-cycle wage profile, and so this approach is well suited for our needs.

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, the regression also includes individual fixed (γ_i) effects 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

⁸ GED-holders are included in the sample. This is especially relevant since many incarcerated individuals earn a GED in prison or are mandated to do so as part of their parole release.

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

¹⁰ The data are censored with a maximum age of 50 on account of the single-cohort panel structure of the NLSY79.

spells aggregated; (b) non-employment and prison spells separated; and (c) non-participation, unemployment, and prison spells separated.

Table 3: Wage Regressions table

	(1) Wagem_ b/se	(2) Wagem_ b/se	(3) Wagem_ b/se	(4) Wagem_ b/se	(5) Wagem_ b/se
Age (yrs)	0.02377*** (0.00033)			0.02375*** (0.00033)	0.02219*** (0.00034)
Age ²	-0.00045*** (0.00001)			-0.00045*** (0.00001)	-0.00041*** (0.00001)
Experience (mo)		0.00160*** (0.00003)			
Experience ²		-0.00000*** (0.00000)			
Age 25-34			0.12937*** (0.00178)		
Age 35-50			0.21097*** (0.00206)		
Non-Employed (mo)	-0.00436*** (0.00011)	-0.00433*** (0.00011)	-0.00468*** (0.00011)	-0.00437*** (0.00011)	
Non-Employed ²	-0.00001* (0.00000)	-0.00001 (0.00000)	-0.00001 (0.00000)	-0.00001* (0.00000)	
Jail Last Yr				-0.39459*** (0.05093)	-0.40735*** (0.05091)
Non-Participant (mo)					-0.00318*** (0.00033)
Non-Participant ²					0.00005** (0.00001)
Unemployed (mo)					-0.00642*** (0.00034)
Unemployed ²					0.00013*** (0.00001)
Constant	1.58557*** (0.00264)	1.59744*** (0.00267)	1.67146*** (0.00198)	1.58605*** (0.00264)	1.59659*** (0.00275)
Observations	182071	182071	182071	182071	182071

Standard errors in parentheses

Experience is total months over lifetime.

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

We also use the weekly data from the NLSY79 to calculate labor market flows to compare with the model's calibration. The states and flows are identified as follows. Employment is defined in the same way as described above for the Mincer regression specification with experience. Non-employment is categorized as either "unemployed" or "non-participant," depending on respondents' answers to a question about their labor market status. If a respondent has a job that 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 subsequent four weeks. If we see the respondent becomes "employed" in the subsequent four weeks and the tenure is greater than one month, then we count the individual as having been continuously employed. In other words, if the respondent regains employment at tenure greater than one month, we conclude the transition is spurious and drop it as a true transition.

Table 4: Weekly Employment Transition Rates

Status at $t - 1$ Status at t	Employed			Non-Employed	Unemployed	Non-Participant
	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.

5. Characteristics at Arrest: Survey of Inmates of State Correctional Facilities

The Survey of Inmates of State Correctional Facilities from the Bureau of Justice Statistics is a 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 specific sample selection is black or white males, ages 18 to 64 at survey date, with less than a high school diploma.¹¹ Further, the sample is restricted to inmates entering the prison in the 1970s.¹² The percentage of the demographic group in prison uses total population in the demographic group estimated from the 1980 decennial census.

Targeted & Non-Targeted Statistics from the Prison Survey 1979

Population in State Prison	
Age 18-24	0.026
25-34	0.043
35-64	0.005
Estimated Pop. State & Federal Prison	
Age 18-24	0.028
25-34	0.047
35-64	0.005
Employed Month of Crime	
Age 18-64	0.699
Unemployed Month of Crime	
Age 18-64	0.153
NiLF Month of Crime	
Age 18-64	0.144

¹¹ Specifically, this is coded as not having completely attended the 12th grade (4th year of high school).

¹² All observations are weighted with frequency weights provided in the survey construction. Employed includes both part- and full-time.

Total prison populations, including state and federal facilities, were estimated by scaling populations to adjust for the share of males with sentences greater than one year in state prisons in 1979, out of federal and state facilities combined (92.7 percent). We accessed National Prison Statistics using the Corrections Statistical Analysis Tool available at bjs.gov.

6. Recidivism of Prisoners Released Series, 1983, 1994, and 2005

The Bureau of Justice Statistics organized the compilation of demographic and criminal history data for prisoners released in 1983, 1994, and 2005. The data cover a representative sample of 16,000 released prisoners in 1983 and 38,624 released prisoners in 2005 from California, Florida, Illinois, Michigan, Minnesota, New Jersey, New York, North Carolina, Ohio, Oregon, and Texas.¹³ 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. The 2005 survey significantly increased the number of states involved to 30 and followed individuals over five years instead of three.

We obtained access to the restricted microdata for the first two surveys but not for the 2005 iteration. Our statistics for recidivism in the 2000s instead come from restricted microdata we 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 by the Florida Department of Corrections between 1996 and 2004. Outcomes for each released individual are available from state criminal records for three years following their release. We restrict our analysis to individuals released from prisons after 2000 for comparability.

There are obvious hazards in comparing the Florida survey to the 1983 and 1994 surveys. Clearly, Florida on its own is less representative than the 11 states used in those surveys, but it is consistently a top-3 state in number of state prison inmates accounting for 7 to 10 percent of the prison population. The greater concern is that the survey covers only recidivism activities taking place within Florida. For this reason we would expect to under-estimate recidivism activities relative to the 1983 and 1994 surveys. However, a comparison of reimprisonment rates over a three-year time horizon with those reported for the 2005 iteration of the Recidivism of Prisoners Released Survey provides some confidence in the comparability of the Florida survey. We calculated a total three-year reimprisonment rate of 36 percent from our subsample of the Florida data, which is remarkably close to the same statistic of 36.1 percent in the BJS report from the 2005 Recidivism of Prisoners Released Survey.¹⁴

We are interested in a single dimension of recidivism most consistent with our model and measurements in other datasets: reimprisonment for a new felony charge. The table below presents trends in this statistic by the age partition used in our model.¹⁵

¹³ Arizona, Delaware, and Virginia were added in the 1994 survey, but we exclude them for consistent comparison across surveys.

¹⁴ 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.

¹⁵ Caution must be used when comparing these data with the BJS summary papers on the surveys. Our analysis of the microdata exactly replicates these reports when using the “received” records from prisons and jails to identify reincarceration. Using this

3-yr Re-imprisonment Rate on a Felony Charge			
Age	1983 (<i>a</i>)	1994 (<i>a</i>)	2000-2003 (<i>b</i>)
18-24	22.1	29.8	42.6
25-34	18.3	26.3	36.4
35-64	9.6	20.1	30.4
Total (18-64)	17.8	25.0	35.0

Re-imprisonment rate for inmates released from state prison during the stated year.

a authors' calculations from the Recidivism of Prisoners Released Series (BJS) micro-data

b authors' calculations from Criminal Recidivism in a Large Cohort of Offenders Released from Prison in Florida 2004-2008 (FDOC) micro-data

7. Recidivism of Felons on Probation, 1986–1989

Recidivism of Felons on Probation, 1986–1989, is a data release from the Bureau of Justice statistics. These data include information from sentencing records, probation files, and criminal histories collected in 1989, pertaining to individuals under felony probation in 1986. The sample includes 32 urban and suburban areas. There are 12,369 observations from a representative sample of the 81,927 total individuals on probation in 1986. Although the time period is slightly later than our initial calibration, we use these statistics to calibrate parameters of the model related to skill shocks by employment and jail status, which we assume are time and policy invariant.

The specific sample selection is black or white males, ages 18 to 64 at survey date, with less than a high school diploma.¹⁶ Observations are weighted using the provided “probation case weight” variable, which considers demographics, offense, location, and sentence.¹⁷

measure, we match their 40 percent three-year recidivism rate for 1983, which breaks down to 51 percent, 38 percent, and 29 percent for young, middle, and old age groups, respectively. However, this measure includes jails, which we are not considering in other datasets, and it includes reconfinement 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.

¹⁶ Here we cannot distinguish between GED and non-GED high school graduates, so GEDs are excluded.

¹⁷ Statistics by wage for those with more than 60 percent of weeks employed are pooled due to small sample by race. The pool includes 446 individuals.

Targeted Statistics from Recidivism of Felons 1986-1989		
Statistic	Proportion	3 Year Recidivism
Total Pop	1	17.1
Wks Employed		
>60%	31.4 (29.0-33.6)	9.7 (7.4-12.0)
40-60%	18.5 (16.75-20.25)	12.9 (9.1-16.7)
<40%	50.1 (49.0-52.2)	23.3 (20.8-25.8)
Hourly Wage		
< \$5	22.7 (17.7-27.6)	12.8 (4.5-21.1)
\$5-\$10	47.7 (41.9-53.6)	14.1 (7.8-20.5)
>\$10	29.6 (24.4-34.8)	10.1 (3.7-39.0)
Sample: Emp >60%	Proportion	3 Yr Recidivism
< \$5	21.9 (16.5-27.2)	6.3 (1.2-11.4)
\$5-\$10	51.7 (45.1-58.2)	10.5 (4.8-16.3)
>\$10	26.5 (20.9-32.0)	6.9 (0.1-13.6)

Data note: Weeks employed are for the first year after release. Workers earning less than the minimum wage (\$3.10) are treated as unemployed. Wages are truncated above at \$10 with 24 percent of employed individuals earning more than \$10.

8. Decennial Census and Current Population Survey

We use data from the decennial census and current population surveys to calculate labor market statistics for our focus population: white and black males without high school diplomas.

The specific sample selection is males, ages 18 to 64 at survey date, with less than a high school diploma,¹⁸ civilian, non-institutionalized. We considered a subject employed if he was employed at the time of the survey and also worked 50 to 52 weeks in the previous year. Similarly, we considered a subject

¹⁸ Specifically, less than a high school diploma is coded as (*educd* < 50 & *educd* > 1), and we also consider adding (*educd* == 63 || *educd* == 62). We are trying to include GEDs, but we omit those completing 12th grade with a diploma.

unemployed if he was unemployed at the time of the survey and was also unemployed for more than three weeks in previous last year.

Employment Rate and Employment to Population Ratio		
Statistic	Employment Rate	Employment to Pop
1970	95.4	81.9
Young (18-24)	90.3	66.9
Middle (25-34)	95.2	88.6
Old (35-64)	96.3	83.5
1980	90.1	71.1
Young (18-24)	82.0	61.5
Middle (25-34)	87.9	77.9
Old (35-64)	93.3	72.7
1990	87.7	64.4
Young (18-24)	79.0	56.4
Middle (25-34)	85.8	72.5
Old (35-64)	91.5	64.3
2000	88.5	59.1
Young (18-24)	82.0	55.7
Middle (25-34)	88.8	67.6
Old (35-64)	91.3	57.8
2010	79.9	54.5
Young (18-24)	66.4	38.1
Middle (25-34)	79.9	65.8
Old (35-64)	84.2	56.8

For the estimation of the model, we target statistics averaged across the 1970 to 1980 current population survey for our sample. First the employment rate is 70 percent. As a robustness check on wages, we compare growth rates implied by our NLSY79 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 and multiplied by average hours per week.¹⁹

¹⁹ We set hourly wages below \$2 or above \$200 (1980 dollars) equal to missing.

Median Hourly Wage by Age Group (1980-dollars)			
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

9. 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 and have identical human capital and criminal capital (low), and we assume that prison has no explicit effect on job-finding probability.

Let V_p , V_u , V_e represent the value functions for incarcerated, unemployed, and employed individuals, respectively. We can formulate these value functions as:

Incarcerated Individual:

$$rV_p = \tau(V_u - V_p) \quad (9.1)$$

Unemployed Individual:

$$rV_u = b + \lambda_w \max\{V_e - V_u, 0\} + \eta \int \max\{\pi(V_p - V_u) + \kappa, 0\} dH(\kappa) \quad (9.2)$$

Employed Individual:

$$rV_e = wh + \delta(V_u - V_e) + \eta \int \max\{\pi(V_p - V_e) + \kappa, 0\} dH(\kappa) \quad (9.3)$$

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

Lemma 9.1. *There exists $\kappa_u^*(\kappa_e^*)$ such that unemployed (employed) individual commits every crime if the reward is higher than or equal to $\kappa_u^*(\kappa_e^*)$, and κ_u^* is given by:*

$$\kappa_u^* = \pi(V_u - V_p) \quad (9.4)$$

$$\kappa_e^* = \pi(V_e - V_p) \quad (9.5)$$

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

Corollary 9.2. *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 9.3. *The values for incarceration, V_p , unemployment, V_u , and employment, V_e , can be expressed as:*

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

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 (i.e., 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 fewer crimes. However, the net effect on incarceration probability is ambiguous. Notice that incarceration probability in the model is $\pi\eta(1 - H(\kappa_i^*))$ 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.

9.1 Firm's Problem:

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

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

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:

$$\lambda_f = \frac{\kappa(r + \delta + \eta(1 - H(\kappa_e^*))\pi)}{(1 - w)h} \quad (9.6)$$

Lemma 1. *An increase in the probability of incarceration, $\pi\eta(1 - H(\kappa_e^*))$, 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 (9.6). There is an inverse relationship between the expected duration of a new match and the probability of a worker ending the match by going to prison. 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 that 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.²⁰

9.2 Steady-State Flows:

The equations characterizing the steady state are as follows:

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

where p and u are the measures 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:

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

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:

$$\frac{u}{1-p} = \frac{\delta + \eta\pi}{\lambda_w + \delta + \eta\pi} \quad (9.7)$$

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

10. Proofs

From equation (9.1) we have:

$$V_u = \frac{r+\tau}{\tau} V_p \quad (10.1)$$

We can express the difference between the value of employment and incarceration by combining equations (9.1) and (9.3):

$$(r + \eta(1 - H(\kappa_e^*))\pi + \delta)(V_e - V_p) = wh + (\delta - \tau)(V_u - V_p) + \eta\bar{\kappa}_e \quad (10.2)$$

²⁰ 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 it may or may not hold otherwise (proposition 2).

where $\bar{\kappa}_e = \int_{\kappa_e^*} \kappa dH(\kappa)$. Substituting (9.4) and (9.5) into the above equation, we get:

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

Similarly, using equations (9.1) and (9.2), 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 (9.4) and (9.5) into the above equation, we get:

$$(r + \lambda_w + \tau) \frac{\kappa_u^*}{\pi} + \eta(1 - H(\kappa_u^*))\kappa_u^* = b + \lambda_w \frac{\kappa_e^*}{\pi} + \eta\bar{\kappa}_u \quad (10.3)$$

Then, equations (10.2) and (10.3) 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 (10.2) and (10.3), 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 (10.3):

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

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

Using integration by parts, we can express equation (10.2) 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:

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

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 (10.3) and the implicit function theorem, we have:

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

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 (9.7) 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.

11. Value Functions

11.1 Individuals

Given the five-dimensional heterogeneity of the individuals, the value function for the individuals depends on five state variables: age, criminal type, addiction type, labor market status, and skill. We denote $V_i^{x,m,k}(h)$ as the value of an individual with labor market status $i \in \{e, u, p\}$, addiction type $x \in \{a, na\}$, criminal type $k \in \{0, 1\}$, skill level h , and age m . Employed individuals receive wage, wh , which potentially depends on the skill level and market wage rate.

For notational convenience we denote $I_f^{x,m,k}(h)$ as the indicator function for the continuation of the match between the firm and the individual of type x, k , skill level h , and age m . $I_f^{x,m,k}(h) = 1$ if the match continues, and $I_f^{x,m,k}(h) = 0$ if the match dissolves:

$$I_f^{x,m,k}(h) = \begin{cases} 1 & \text{if } V_e^{x,m,k}(h) \geq V_u^{x,m,k}(h) \\ 0 & \text{o. w.} \end{cases} \quad (11.1)$$

For notational convenience we define the following value functions:

$$\begin{aligned} V_{eu}^{x,m,k}(h) &= I_f^{x,m,k}(h)V_e^{x,m,k}(h) + (1 - I_f^{x,m,k}(h))V_u^{x,m,k}(h) \\ V_{px}^{x,m,k}(h) &= v_x V_p^{-x,m,k}(h) + (1 - v_x)V_p^{x,m,k}(h) \\ V_{ijx}(h) &= v_x V_{ij}^{-x,m,k}(h) + (1 - v_x)V_{ij}^{x,m,k}(h) \end{aligned}$$

The value of an employed worker with skill level h , age m , and type x, k becomes the following equation:

$$\begin{aligned} wh + \eta_m \int \max\{V_{ec}^{x,m,k}(h, \kappa) - V_e^{x,m,k}(h), 0\} dF(\kappa) + \eta_m^x (V_{eca}^{x,m,k}(h) - V_e^{x,m,k}(h)) + \\ rV_e^{x,m,k}(h) = \psi_e (V_{eu}^{x,m,k}(f_e(h)) - V_e^{x,m,k}(h)) + \delta (V_u^{x,m,k}(h) - V_e^{x,m,k}(h)) + \\ \vartheta_m (V_{eu}^{x,m+1,k}(h) - V_e^{x,m,k}(h)) + \xi^x (V_{eu}^{-x,m,k}(h) - V_e^{x,m,k}(h)) \end{aligned} \quad (11.2)$$

s.to

$$\begin{aligned} V_{eca}^{x,m,k}(h) &= \pi V_{px}^{x,m,-k}(h) + (1 - \pi)V_{eux}^{x,m,k}(h) \\ V_{ec}^{x,m,k}(h, \kappa) &= \pi V_{px}^{x,m,-k}(h) + (1 - \pi)V_{eux}^{x,m,k}(h) + \kappa \end{aligned}$$

Unemployed individuals receive unemployment benefits. Denoting b as the replacement ratio, the unemployment benefits for an individual with human capital level h becomes bh . The value of an unemployed individual with skill level h , age m , and type x, k can be written as follows:

$$\begin{aligned}
 rV_u^{x,m,k}(h) = & bh + \eta_m \int \max\{V_{uc}^{x,m,k}(h, \kappa) - V_u^{x,m,k}(h), 0\} dF(\kappa) + \eta_m^x (V_{uca}^{x,m,k}(h) - V_u^{x,m,k}(h)) + \\
 & \lambda_w^{k,m} (V_{eu}^{x,m,k}(h) - V_u^{x,m,k}(h)) + \xi^x (V_u^{-x,m,k}(h) - V_u^{x,m,k}(h)) + \\
 & \vartheta_m (V_{eu}^{x,m+1,k}(h)) + \psi_u (V_u^{x,m,k}(f_u(h)) - V_u^{x,m,k}(h))
 \end{aligned} \tag{11.3}$$

s.to.

$$\begin{aligned}
 V_{uca}^{x,m,k}(h) &= \pi V_{px}^{x,m,-k}(h) + (1 - \pi) V_{ux}^{x,m,k}(h) \\
 V_{uc}^{x,m,k}(h, \kappa) &= \pi V_{px}^{x,m,-k}(h) + (1 - \pi) V_{ux}^{x,m,k}(h) + \kappa
 \end{aligned}$$

Lastly, incarcerated individuals receive no benefits. The value of an incarcerated individual becomes the following:

$$rV_p^{x,m,k}(h) = \tau (V_u^{x,m,k}(h) - V_p^{x,m,k}(h)) + \vartheta_m (V_p^{x,m+1,k}(h) - V_p^{x,m,k}(h)) + \tag{11.4}$$

$$\xi^x (V_p^{-x,m,k}(h) - V_p^{x,m,k}(h)) + \psi_p (V_p^{x,m,k}(f_p(h)) - V_p^{x,m,k}(h)) \tag{11.5}$$

11.2 Firms

$$rV_f^{k,m} = -c_{k,m} + \lambda_f^{k,m} \int (J^{x,m,k}(h) - V_f) d\mu_u(h, x|m, k) \tag{11.6}$$

$$\begin{aligned}
 rJ^{x,m,k}(h) = & \left\{ (1 - w)h + \delta (V_f^{k,m} - J^{x,m,k}(h)) + \psi_e (J^{x,m,k}(f_e(h)) - J^{x,m,k}(h)) + \right. \\
 & \left(\eta_m (1 - F(\kappa_e^{x,k,m}(h))) + \eta_m^x \right) \pi (V_f^{k,m} - J^{x,m,k}(h)) + \vartheta_m (J^{x,m+1,k}(h) - J^{x,m,k}(h)) \Bigg\} \\
 & \left(\left(\eta_m (1 - F(\kappa_e^{x,k,m}(h))) + \eta_m^x \right) (1 - \pi) v^x + \xi^x \right) (J^{-x,m,k}(h) - J^{x,m,k}(h))
 \end{aligned} \tag{11.7}$$

where $\kappa_e^{x,k,m}(h)$ denotes the criminal reward threshold for the employed individual with characteristics (x, k, m, h) .

11.3 Equilibrium

A stationary equilibrium consists of value functions for individuals V, J decision rules I_f , policy functions κ_u and κ_e for the unemployed and employed, market tightness $\theta^{k,m}$, job offer arrival rates $\lambda_w^{k,m}$, worker arrival rates $\lambda_f^{k,m}$, and stationary measure of individuals μ , such that:

1. Given market tightness and job and worker arrival rates, policy functions s_u and s_e and value functions V, V_f , and J solve (11.2) through (11.6), and decision rule I_f solves (11.1).

2. Stationary measure μ is consistent with the decision rules.
3. Perfect competition among firms should result $V_f^{k,m} = 0$.
4. Job arrival rates and worker arrival rates satisfy (9.6).

12. Additional Results

12.1 The Intensive Margin: Recidivism Disaggregated by Age over the Transition

A key non-targeted prediction of the calibrated model is that the intensive margin of crime should increase while the extensive margin of crime should decrease when policy becomes more punitive. In other words, although crime falls, recidivism should increase. Furthermore, recidivism should increase more for older ages. Figures 7(a) and 7(b) below show two measures of recidivism calculated from the Recidivism of Prisoners Released data series. The sample is restricted to our demographics of interest: white and black males without college experience. It is also restricted to those serving sentences of five years or less.

Figure 7(a) shows the percent of individuals rearrested for a new felony crime within three years of release. This is the preferred measure because it focuses more on the behavioral incidence of crime by removing one additional layer of time-varying discretion regarding whether an arrest translates to a prison sentence. The results are striking and in line with the theory. Recidivism increases for almost all age groups over time, but it increases the most for older individuals. By 2005, the recidivism curve is almost age-independent. This is striking given that all individuals over age 40 are pooled.

Figure 7(b) shows a similar but less stark result for recidivism measured as the percent re-imprisoned on a new offense. This excludes reimprisonment due to technical violations of probation or parole but does not factor in the changing role of prior offenses in sentencing over time. The general pattern is still there—the age profile increases and becomes flatter—but the interpretation is a bit more convoluted.

Figure 12.1 in the main text plots the change in the recidivism rate implied by the model across different age groups. It shows the ratio of recidivism rates in the final steady state as a ratio of the recidivism rates in the initial steady state. As in the data, the most significant increase in the recidivism rate happens for the old age group.

12.2 Comparative Statistics for Probability of Getting Caught (π)

As we explained in the main text, the model has the potential to generate non-monotonic response to the change in probability of getting caught, π . As π increases, we expect the incarceration to increase since, conditional on committing crime, individuals are more likely to get caught and sent to prison. However, in the model, individuals can respond in two dimensions to the change in incarceration probability. First, they can increase the crime threshold resulting in fewer crimes upon receiving opportunity. Moreover, they can reduce their early life choice of crime propensity, which again results in lower likelihood of crime. Lastly, the change in the policy results in changes in the distribution of individuals across human capital and labor market status. Such a change has an impact on the crime propensity, as explained in the main text.

In this section, we provide how overall incarceration changes in response to a change in π . Figures [9\(a\)](#) through [9\(d\)](#) below provide the comparison across steady states. As the probability of getting caught increases, the incarceration rate increases initially (figure 9[a]). However, when π reaches around 10 percent, the incarceration rate reaches the maximum and starts declining. The decline comes from the offsetting effects of a decline in the conditional crime probability (figure 9[b]), a decline in the fraction of individuals with high criminal capital (figure 9[c]), and a decline in the early life choice of crime arrival rate (figure 9[d]).

12.3 Results for Robustness

The parameter we do not have data to identify is the standard deviation of crime reward distribution. In this section, we provide some robustness checks with respect to this parameter. We re-estimate the model by changing the standard deviation of the crime reward distribution and generate the statistics corresponding to the re-estimated parameters. We run two robustness checks. In the first experiment, we decrease the standard deviation of the crime reward distribution by 50 percent, and in the second one we increase it by 50 percent, compared to its benchmark value.

In each of these counterfactuals, we adjust the log-mean of the distribution to keep the mean of the distribution constant. Figures [10\(a\)](#) through [10\(d\)](#) below show the results of these robustness checks. Overall qualitative results do not change. However, quantitative results are sensitive to the choice of the variance. Although the decrease in the variance does not generate significant changes in the statistics compared to the benchmark economy, the increase in the variance amplifies the results.

These results imply that the quantitative effects provided in the paper can be interpreted as conservative numbers. It is possible that the actual effects might be larger than we presented in the paper. Pinning down the actual quantitative effects requires a discipline to identify the variance parameter. One way to do this is to take a stand on the observed benefits of the crime. Several papers do this by focusing only on property crimes. However, since our definition of crime is more general, we have avoided this path.

12.4 Alternative Calibration

Table [5](#) below provides the estimated numbers corresponding to counterfactual models we consider in the paper. Column 1 is for the benchmark model. Column 2 (no prison flag) is for the model when we remove the prison flag assumption. Column 3 (no criminal capital) is for the model when we assume there is no heterogeneity across criminal capital and everyone has low criminal capital. Column 4 (low variance) is for the model when we set the standard deviation of crime reward distribution to half of its value in the benchmark model. Column 5 (high variance) is for the model when we set the standard deviation of crime reward distribution to twice its value in the benchmark model.

In each case, we obtain the calibrated parameters by minimizing the sum of the square of the percentage distance between same data moments and model implied moments as in the benchmark calibration, except for the model with no criminal capital. Since the model with no criminal capital removes the parameters about the criminal capital process (η , ξ , ν , nu_p), we only target incarceration rates for young and middle-age, employment rates for young and middle-age, average unemployment duration in the whole population, the regression coefficients, and the change in incarceration for young individuals over transition.

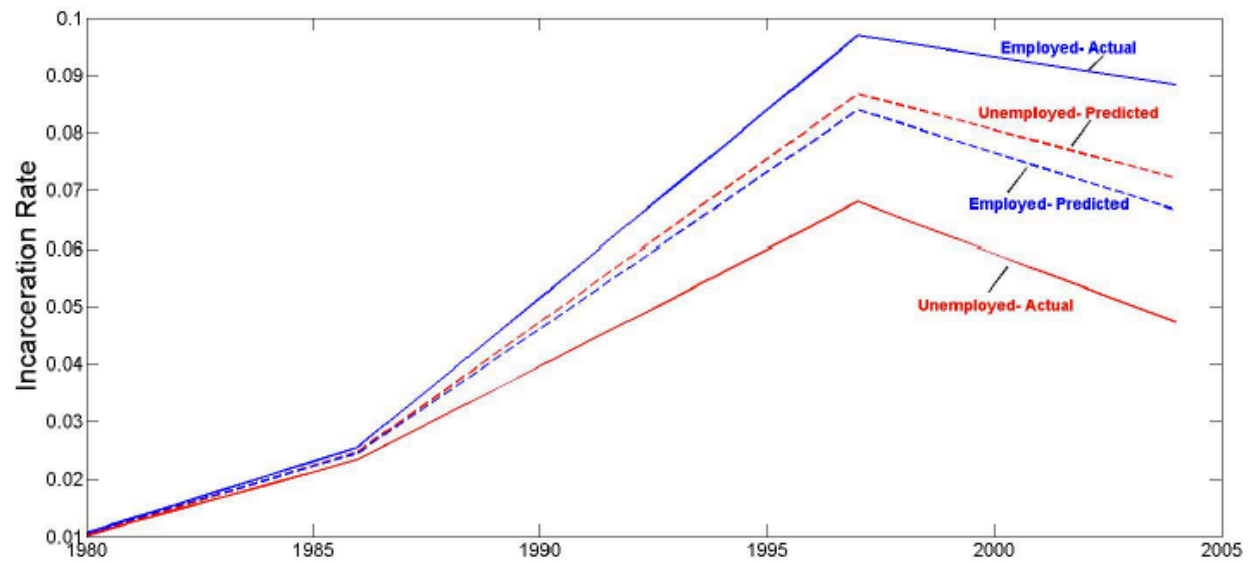
Table 5: Calibrated Parameters

Parameter	Benchmark	No Prison Flag	No Criminal Capital	Low Variance	High Variance
η	0.038	0.036	0.075	0.012	0.075
c	133.5	132.0	143.4	114.4	116.1
δ	0.015	0.015	0.015	0.014	0.014
ς	0.37	0.38	0.375	1.44	1.66
ψ_e	0.011	0.010	0.018	0.013	0.015
ψ_p	0.014	0.014	0.032	0.014	0.016
ψ_u	0.014	0.014	0.032	0.014	0.016
ρ	1.91	1.88	1.35	4.95	1.75
η^{hc}	0.077	0.075	-	0.088	0.068
ξ	0.028	0.024	-	0.100	0.010
ν	0.031	0.032	-	0.025	0.223
ν_p	0.36	0.40	-	0.23	0.037
A	1.27	1.13	0.61	282395	2.59

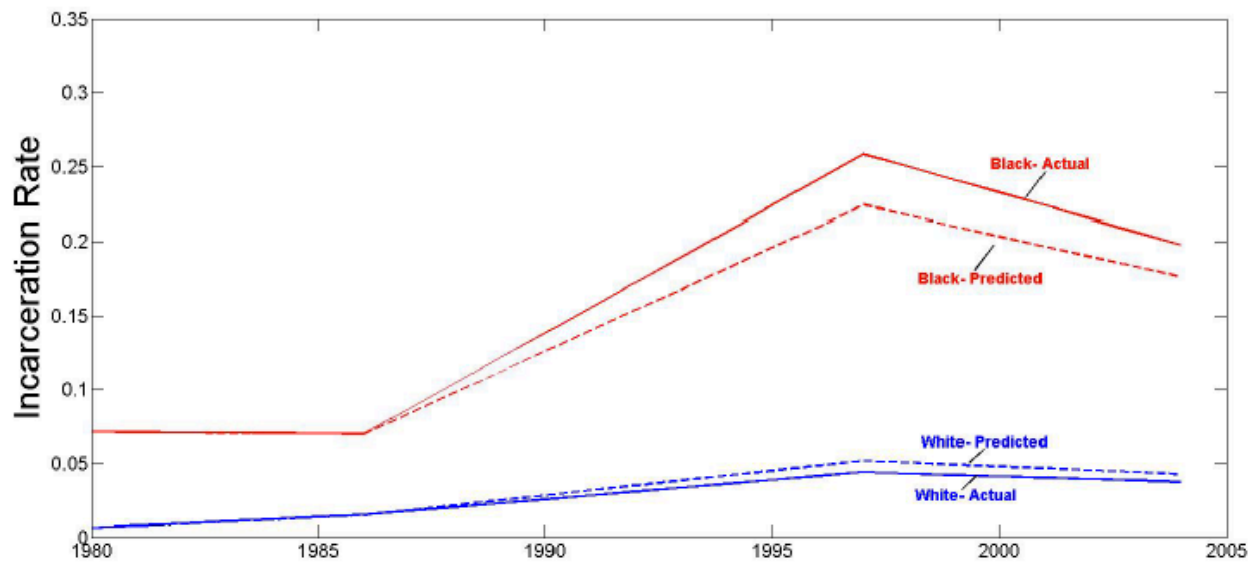
Notes: The Table shows the internally calibrated parameters of the model with different assumptions. The first column is for the benchmark economy. The second column corresponds to the economy without prison flag. The third column is for the economy without criminal capital. The fourth column is the economy with the variance of reward distribution set to half of the benchmark economy. The fifth column is the one with the variance of the reward distribution set to twice the one in the benchmark economy.

References

- Bhati, Avinash Singh. 2010. "Criminal Recidivism in a Large Cohort of Offenders Released from Prison in Florida, 2004–2008." Inter-university Consortium for Political and Social Research [distributor].
- Bureau of Justice Statistics and F. B. of Prisons (1974, 1979, 1986, 1991, 1997). Survey of inmates in state and federal correctional facilities. Inter-university consortium for political and social research.
- Carson, E. A. and J. Mulako-Wangota (2017). Bureau of justice statistics tables. data retrieved from Corrections Statistical Analysis Tool (CSAT)- Prisoners, www.bjs.gov.
- Cunniff, M. N. A. o. C. J. P. (1994). Recidivism of felons on probation, 1986–1989: [united states].
- Durose, Matthew R., Alexia D. Cooper, and Howard N. Snyder. Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010. Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, 2014.
- İmrohoroglu, Ayse, Antonio Merlo, and Peter Rupert. 2004. "What Accounts for the Decline in Crime?." *International Economic Review* 45 (3): 707–29.
- Kitao, Sagiri, Lars Ljungqvist, and Thomas J. Sargent. 2017. "A Life-Cycle Model of Trans-Atlantic Employment Experiences." *Review of Economic Dynamics* 25 (April): 320–49.
- Ljungqvist, Lars, and Thomas J. Sargent. 1998. "The European Unemployment Dilemma." *Journal of Political Economy* 106 (3): 514–50.
- Michaud, Amanda M. 2018. "A Quantitative Theory of Information, Worker Flows, and Wage Dispersion." *American Economic Journal: Macroeconomics* 10 (2): 154–83.
- Neal, Derek, and Armin Rick. 2014. "The Prison Boom and the Lack of Black Progress after Smith and Welch." NBER Working Paper 20283. National Bureau of Economic Research, Cambridge, MA.
- Pettit, Becky. 2012. *Invisible Men: Mass Incarceration and the Myth of Black Progress*. New York: Russell Sage Foundation.
- Pfaff, John F. 2011. "The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices." *American Law and Economics Review* 13 (2): 491–531.
- Raphael, Steven, and Michael A. Stoll. 2009. "Why Are So Many Americans in Prison?" In *Do Prisons Make Us Safer?: The Benefits and Costs of the Prison Boom*. New York: Russell Sage Foundation, 27.
- Ruggles, Steven, et al. 2004. "IPUMS USA: Version 3.0 [Dataset]." Minneapolis, MN: Integrated Public Use Microdata Series.
- US Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. 2010. *National Corrections Reporting Program Series, 1983–2000*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. <https://www.icpsr.umich.edu/icpsrweb/NACJD/series/38>.

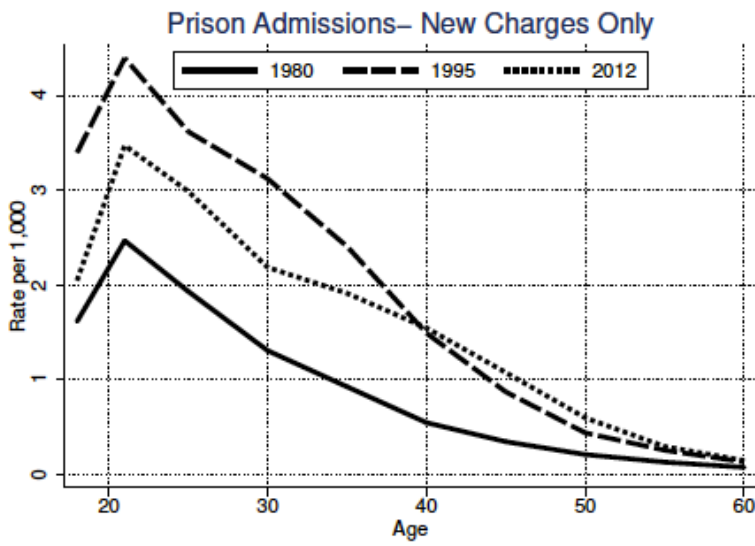


(a) Predicted and Actual Rates by Employment Status

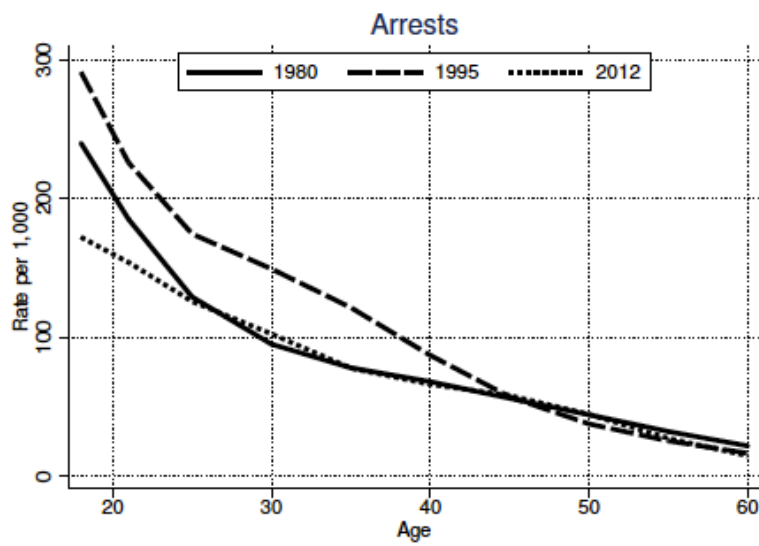


(b) Predicted and Actual Rates by Race

Figure 2. 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, employment status, and age).

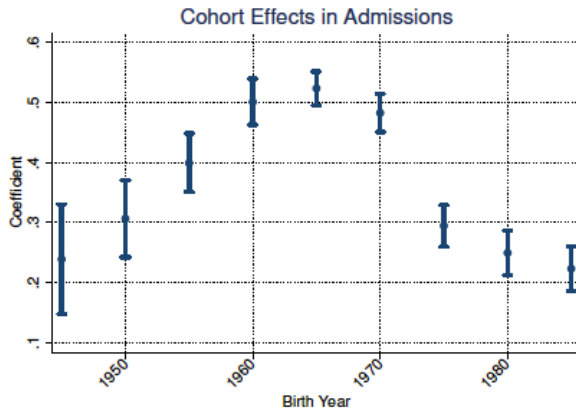


(a) New Prison Admission Rate by Age

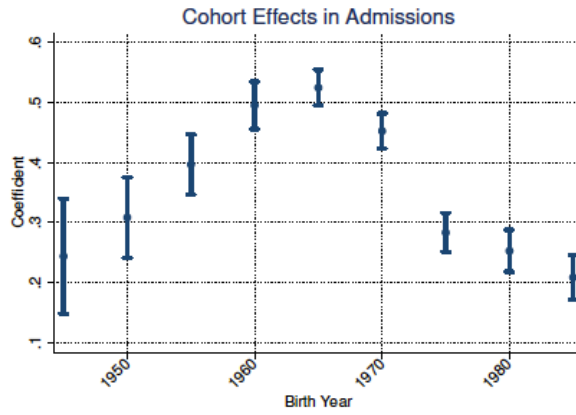


(b) Arrest Rate by Age

Figure 3. Age Profiles of Crime and Incarceration

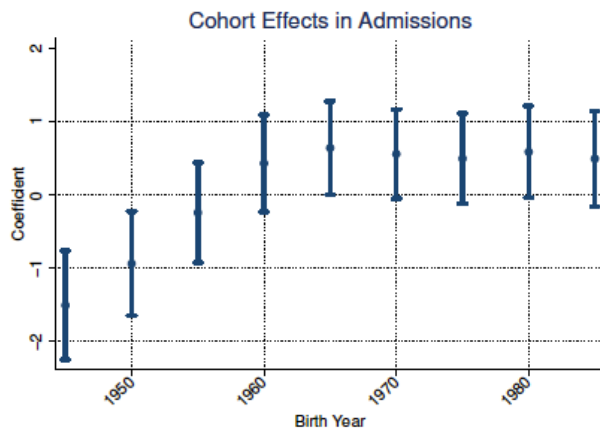


(a) Base: Time-varying Age Profile

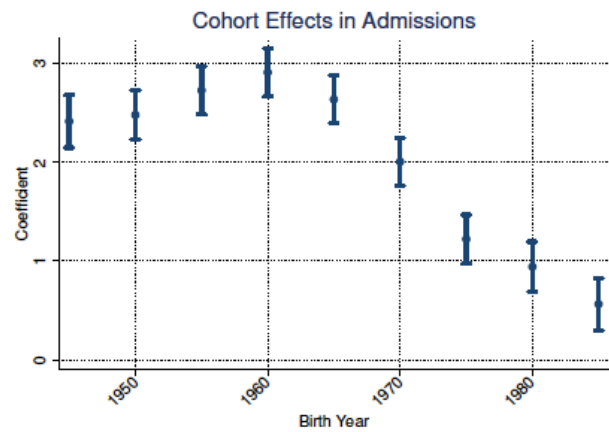


(b) Time-invariant Age Profile

Figure 4. Estimated Cohort Effects: Prison admissions data is from National Corrections Reporting Program data (US Department of Justice 2010) and is restricted to admissions on new charges only.



(a) Linear- Cohort and Age Effects Only



(b) Linear- Cohort and Time Effects Only

Figure 5. Estimated Cohort Effects: Prison admissions is from National Corrections Reporting Program data (US Department of Justice 2010) and is restricted to admissions on new charges only.

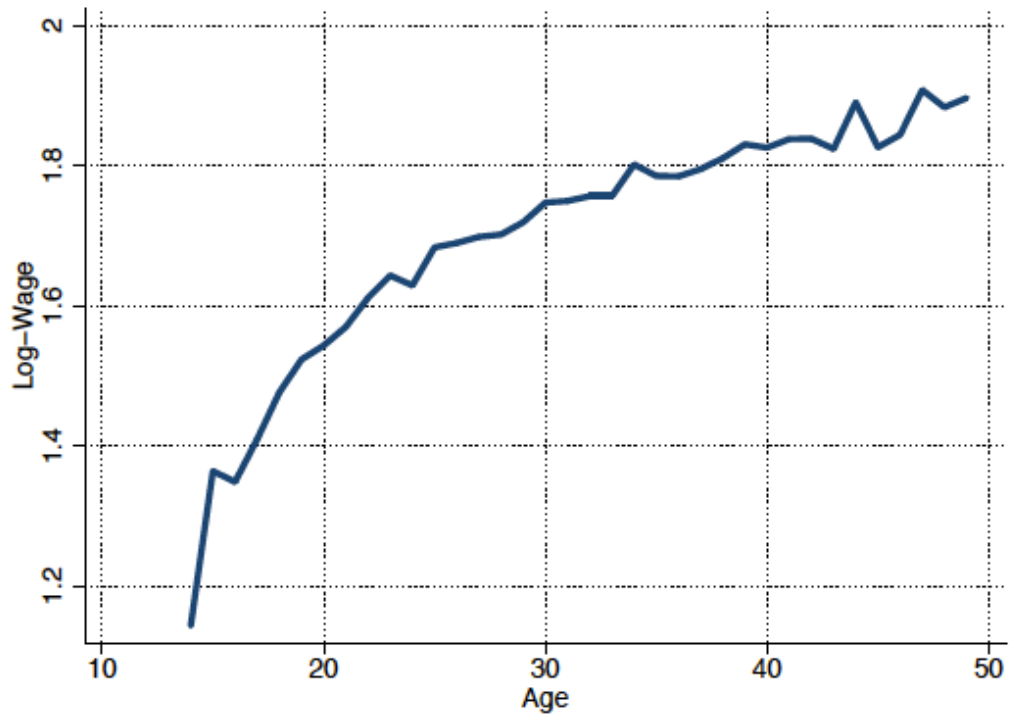
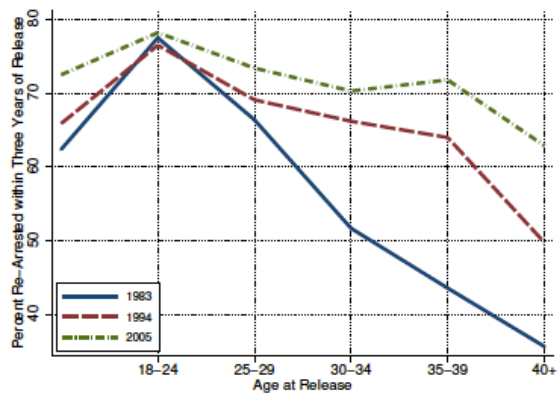
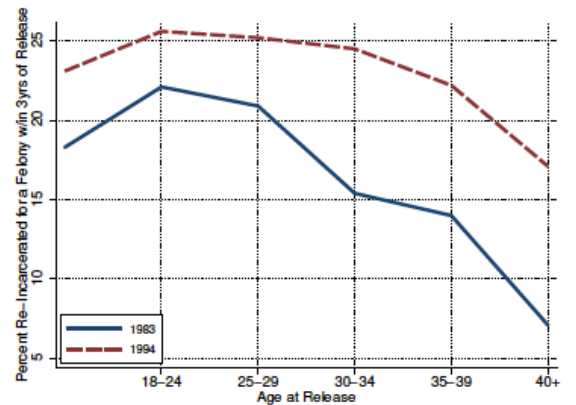


Figure 6. Mean Log-Wage by Age, NLSY79



(a) Felony Arrest



(b) Incarceration for New Offense

Figure 7. Recidivism by age calculated from the Recidivism of Prisoners Released Series (BJS).

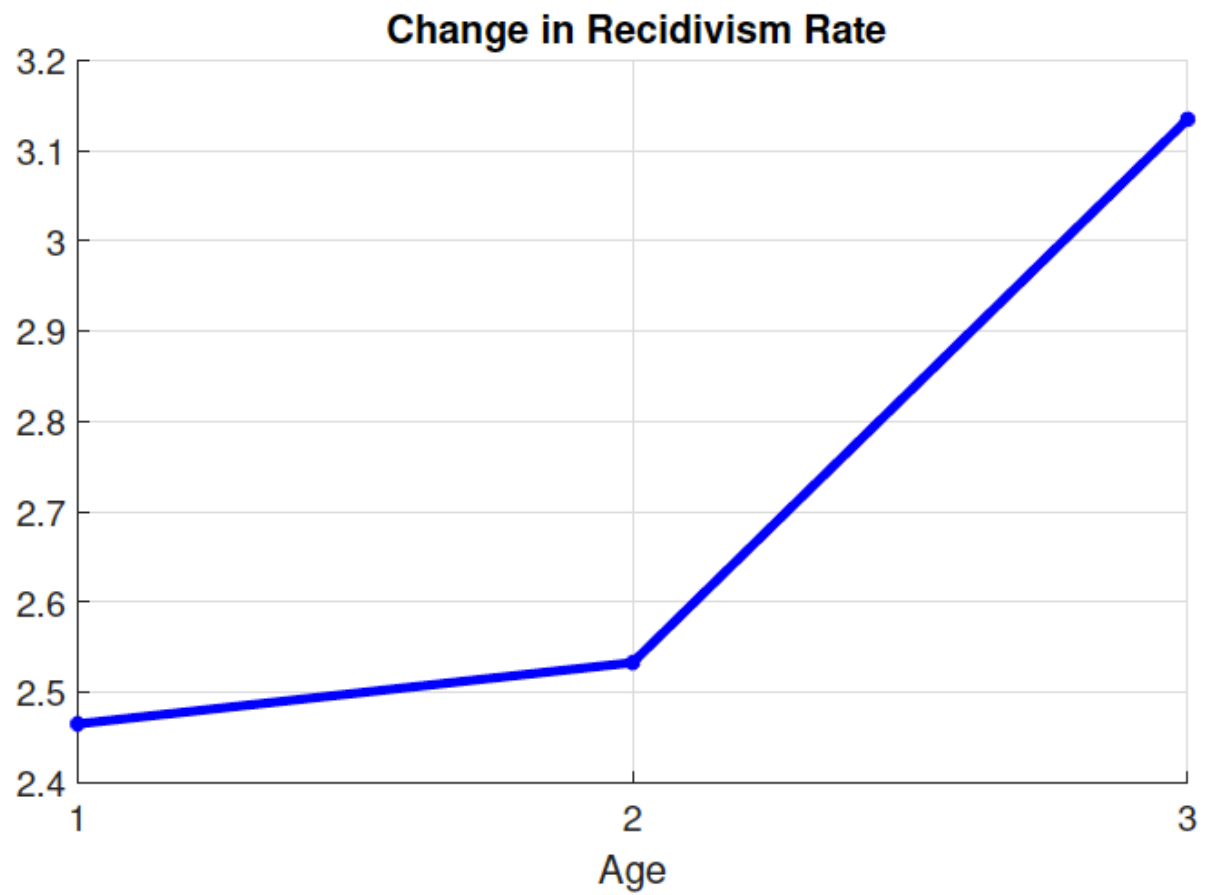


Figure 8. Change in Recidivism: The figure plots the change in the recidivism rate for different age groups across the steady states in the benchmark model as the ratio of the recidivism rate in the final steady state to the recidivism rate in the initial steady state.

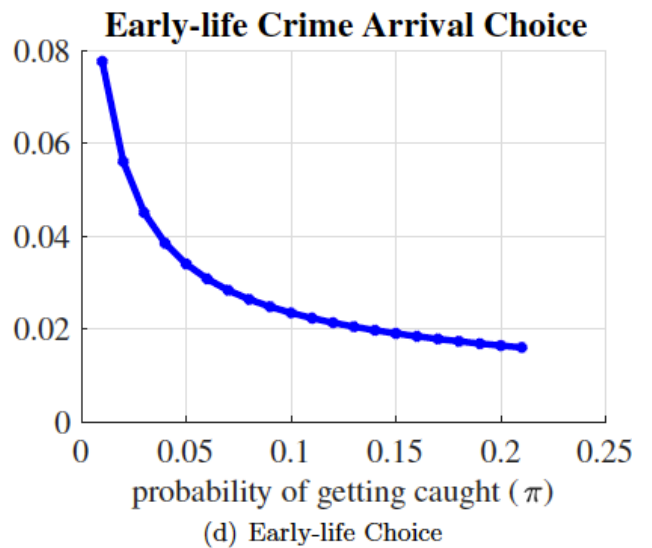
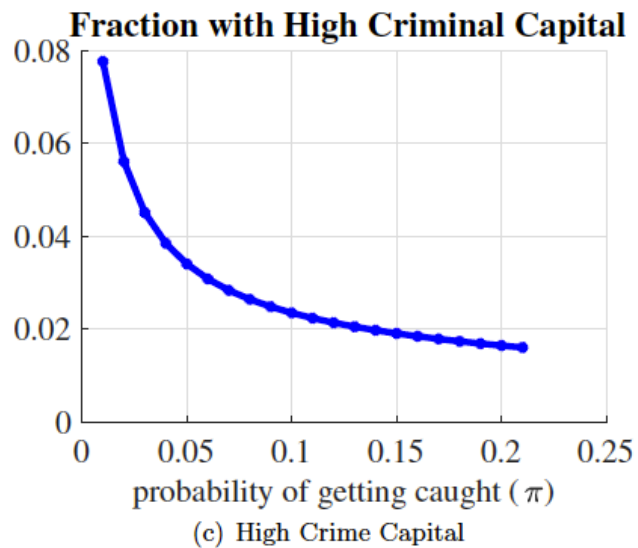
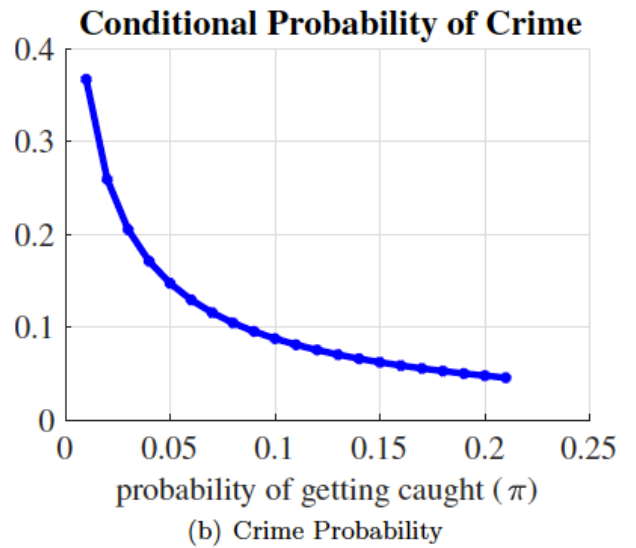
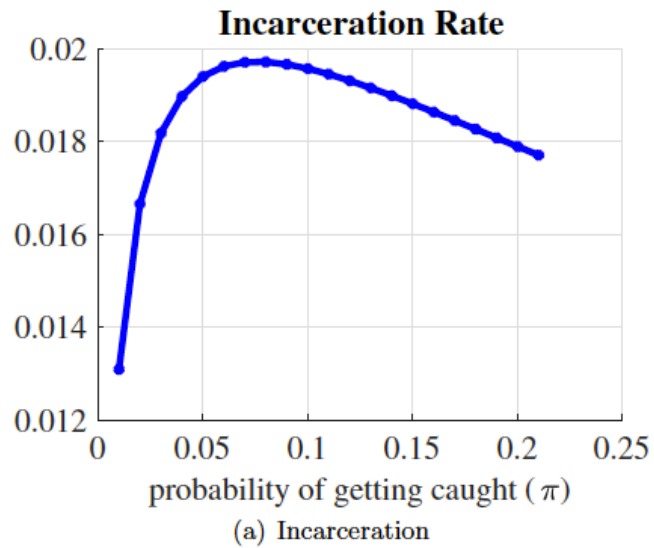


Figure 9. Comparative Statistics: The figures provide the responses of overall incarceration rate, crime probability upon receiving an opportunity (crime reward threshold), fraction of individuals with high criminal capital and choice of crime arrival rate in early life to the changes in probability of getting caught, π , at the steady state.

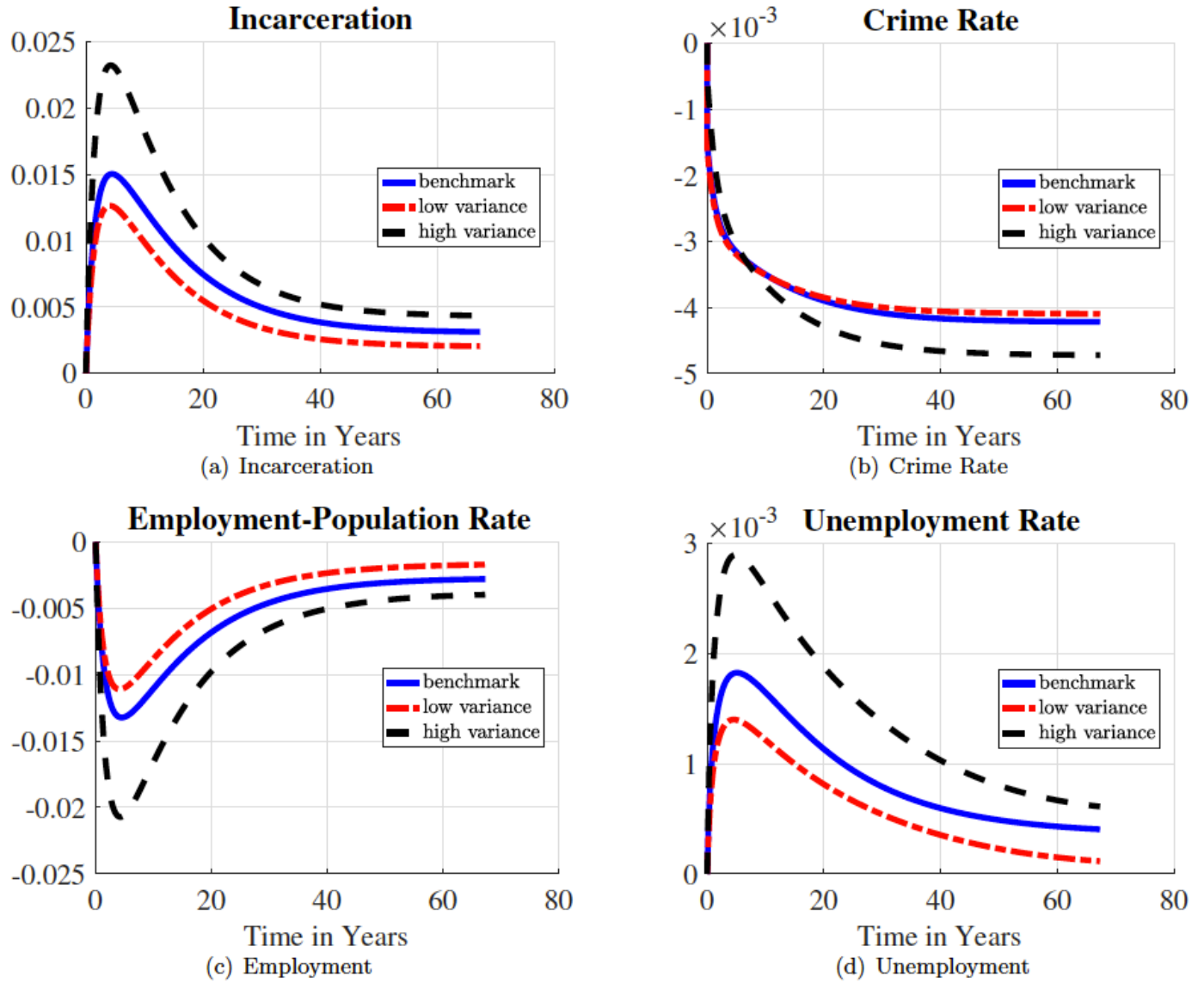


Figure 10. Robustness: The figures compare the evolution of various statistics with respect to a mean-preserving change in the crime reward distribution. They all show the percentage point change compared to the initial steady state. The solid line is the benchmark economy. The long-dashed line corresponds to a 50 percent increase in the standard deviation. In each case, the log-mean of the distribution is adjusted to keep the mean of the distribution constant.