Crime, Punishment, and Myopia*

David S. Lee
UC Berkeley and NBER

Justin McCrary
University of Michigan

July 2005

Abstract

Economic theory predicts that increasing the severity of punishments will deter criminal behavior by raising the expected price of committing crime. This implicit price can be substantially raised by making prison sentences longer, but only if offenders' discount rates are relatively low. We use a large sample of felony arrests to measure the deterrence effect of criminal sanctions. We exploit the fact that young offenders are legally treated as adults—and face longer lengths of incarceration—the day they turn 18. Sufficiently patient individuals should therefore significantly lower their offending rates immediately upon turning 18. The small behavioral responses that we estimate suggest that potential offenders are extremely impatient, myopic, or both.

JEL Codes: D9, K4

*Brett Allen provided excellent research assistance. We are grateful to Jeff Kling for detailed suggestions on a previous draft. For useful comments we thank Josh Angrist, David Autor, David Card, Kerwin Charles, Ken Chay, Gordon Dahl, John DiNardo, Michael Greenstone, Vivian Hwa, Christina Lee, Steve Levitt, Robert McMillan, Enrico Moretti, Daniel Parent, Dan Silverman, and seminar participants at the Center for Advanced Study in the Behavioral Sciences, Columbia University, the NBER Summer Institute, the Massachusetts Institute of Technology, the University of Illinois at Urbana-Champaign, the University of Rochester, the University of Toronto, and the University of Wisconsin. Sue Burton of the Florida Department of Law Enforcement was instrumental in obtaining the data we use here. McCrary acknowledges funding from the Center for Local, Urban, and State Policy at the University of Michigan. Any errors are our own.

I. Introduction

Crime continues to be an important social and economic issue in the United States. While crime rates have fallen in the recent past, the cost of controlling crime has not. From 1970 to 1999, criminal justice system expenditures as a share of national income increased 170 percent, and the ratio of criminal justice employees to the population grew 115 percent. Over the same period, the fraction of the population that is incarcerated increased 395 percent, making the U.S. incarceration rate the highest in the world (Maguire and Pastore 2000, Chaddock 2003).

Crime also continues to be an active area of economic research. Recent studies have investigated a wide range of potential factors, including the effect of police and incarceration (Levitt 1997, Di Tella and Schargrodsky 2004, Levitt 1996, Levitt 1998), conditions in prisons (Katz, Levitt and Shustorovich 2003), education (Lochner and Moretti 2004), social interactions and peer effects (Case and Katz 1991, Glaeser, Sacerdote and Scheinkman 1996, Gaviria and Raphael 2001, Kling, Ludwig and Katz 2005, Jacob and Lefgren 2003), and family circumstances and structure (Glaeser and Sacerdote 1999, Donohue and Levitt 2001). Economists have also considered the returns to education among recent prison releasees (Western, Kling and Weiman 2001), the impact of criminal histories on labor market outcomes (Grogger 1995, Kling 2004), the impact of wages and unemployment rates on crime (Grogger 1998, Raphael and Winter-Ebmer 2001), and the effect of incarceration on the supply of crime in the economy (Freeman 1996, 1999).

Much of this research is grounded in Becker's (1968) framework, in which criminal behavior is considered the outcome of a simple cost-benefit calculation. Arguably, the core message of the economic model of criminal behavior is that it can be discouragedby raising its expected "price". As a practical matter, for serious offenses such as murder, robbery, or burglary, increases in expected price may be achieved either through increases in the likelihood of punishment, or increases in the length of prison sentences.

In addition, the ability to raise the price of crime crucially depends on how much potential offenders discount their future welfare. On the one hand, if potential offenders value their future as much as their present welfare, an increase in sentence lengths from 3 to 6 years, for example, would double the expected price. The same effect could be achieved by holding sentence lengths fixed and doubling the probability of punishment. That is, with patient offenders, certainty and severity are substitutes (Becker 1968).

On the other hand, the population at risk of committing crime is generally thought to have very high discount rates (Wilson and Herrnstein 1985, Katz et al. 2003). If potential offenders value their future significantly less than their present welfare, then the same increase in sentence lengths from 3 to 6 years will have a much more modest impact than it would if they were patient. Indeed, in the extreme case of infinite discount rates, increasing sentence lengths would have no effect, and increasing the likelihood of punishment would be the only way to deter crime.

This paper measures the deterrence effect of incarceration length on criminal behavior. As the empirical literature has emphasized, there are two key difficulties in measuring this magnitude. First, criminal sanctions may be endogenously determined. For example, high crime rates may induce state and local governments to make criminal sanctions more punitive (Ehrlich 1973, Levitt 1997, Levitt 2004a). Second, even ignoring any potential endogeneity, the large estimates of the impact of more punitive sanctions generated from cross-regional variation have an ambiguous interpretation (Levitt 1996, Kessler and Levitt 1999). One interpretation is that sanctions are significantly discouraging criminal *behavior* by making it less attractive—a true deterrence or price effect. The alternative interpretation is that more punitive criminal sanctions reduce crime mechanically, by removing highly crime-prone individuals from the community—the incapacitation effect. This latter explanation has nothing to do with behavioral responses to incentives.

To address these problems, this paper uses (1) a large, person-level, longitudinal data set on arrests with information on exact date of birth and exact date of offense, and (2) a quasi-experiment generated by criminal law. Specifically, we take advantage of the following fact: when an individual is charged with a crime that occurs before his 18th birthday, his case is handled by the juvenile courts.[1] But if the offense is committed on or after his 18th birthday, his case must be handled by the adult criminal court, which is known to administer more punitive criminal sanctions.[2] Thus, when a minor turns 18, there is an immediate increase in the expected price of crime.

We argue that even if all other determinants of criminal offending change rapidly by age, they do not change discontinuously at 18. This allows us to attribute any discontinuous drop in offense

[1]While in principle the case may then be transferred to the adult criminal court, this is rare generally (Snyder and Sickmund 1999). We examine juvenile transfer empirically in Section IV.

[2]In Florida, as in most states, the (criminal) age of majority is 18. Several other states have legislated age cutoffs at 16 or 17 (Bozynski and Szymanski 2003).

rates at 18 to a behavioral response to more punitive sanctions. Our approach circumvents the endogeneity of a community's response to crime since it does not use any cross-regional variation in criminal sanctions. Equally important, our high-frequency, longitudinal data permit us to isolate a pure deterrence effect, rather than a combination of deterrence and incapacitation effects.

Our theoretical and empirical results are as follows. First, we use a standard economic model of crime to provide quantitative predictions of the deterrence effect. Assuming exponential discounting with a 0.95 discount factor, the theory predicts at least a 55 percent decline in the odds of arrest in response to a tripling of incarceration lengths when the minor turns 18. The theory also predicts that the proportional response to a marginal increase in incarceration lengths should be larger in the near future (e.g., 3 versus 9 weeks) than in the distant future (e.g., 3 versus 9 years).

Second, using data on the universe of felony arrests in Florida from 1995-2002, we document the following important fact: there is no significant drop in offending at age 18. The point estimates are small in magnitude, with standard errors small relative to predicted effect sizes; even our most negative effects are consistent with annualized discount factors less than 0.1. The estimates are consistently small across types of crime and characteristics of the individual or jurisdiction.

The estimates are significantly smaller than those found by existing studies that have examined the reduced-form impact of more punitive sanctions on crime (for a summary, see Levitt 2004b). A plausible explanation for these differences is that estimates in the existing literature are operating mostly, and perhaps entirely, through an incapacitation effect. In particular, if incapacitation begins immediately—e.g., due to an initial detention upon arrest—then almost all analyses examining crime rates at annual aggregates will estimate a combination of deterrence and incapacitation effects, rather than deterrence alone.[3] We provide evidence consistent with this view, finding large incapacitation effects that are apparent even within a few weeks of an arrest.

We consider three potential explanations for our results. First, it is possible that potential offenders either have systematically incorrect information about punishments, or behave irrationally. Our economic model, calibration exercise, and supporting empirical evidence tell us that individuals with reasonable discount factors *should* reduce their criminal propensities upon turning 18. But, as a matter of fact, they may be making systematic errors in under-estimating the likelihoods and severity of punishments, either at a given point in time, or in the transition from juvenile status to

[3] A notable exception is the approach of Kessler and Levitt (1999).

adulthood.

Second, it is possible that offenders are behaving rationally, but have extremely low discount factors, perhaps much smaller than 0.1. Taken literally, this would imply that criminals perceive the difference between prison sentences of 2 years and 20 years as negligible.

We consider this implication to be somewhat unlikely, and thus tend to favor a third interpretation: offenders have hyperbolic time preferences, with very low short-run discount factors but more reasonable long-run discount factors. Such a theory of present-biased preferences would imply that criminals view the price difference between 2 and 6 weeks of incarceration to be small, but the price difference between 2 and 6 years to be large. This kind of myopia could explain why we observe such small effects: most incarceration lengths are in fact short, so the sub-population of offenders threatened by long prison sentences may be too small to measurably affect the overall estimates of deterrence.

We conclude that, for any of the above interpretations, increases in incarceration lengths—from current levels—are unlikely to significantly deter criminal behavior. The remainder of the paper is organized as follows. Section II outlines our conceptual framework and identification strategy. Section III describes the data. We report our results in Section IV. Section V discusses alternative interpretations, Section VI discusses policy implications, and Section VII concludes.

II. Conceptual Framework

We begin this section by reviewing the standard economic model of crime, embedding it in a stochastic life-cycle framework.[4] We use the model for two main purposes. First, we clarify the identifying assumptions of our empirical strategy. Essentially, we are assuming that all determinants of criminal behavior are evolving "smoothly" as the potential offender turns 18—except for the severity of criminal sanctions, which rises immediately upon the 18th birthday.

Second, after imposing some additional structure, we generate quantitative predictions for the drop in offense rates at age 18. Specifically, we suppose that offenders (1) are rational, (2) have accurate information about punishments, and (3) have exponential time preferences and annual discount factors of 0.95. The theory predicts at least a 55 percent decline in the odds of arrest at 18, in response to a three-fold increase in the length of incarceration, for an elasticity of -0.28.[5] We

[4]We discuss the highlights of the model here and provide details in Appendix A.

[5]This is in the middle of the range of estimates from the empirical literature (Levitt 2004b).

illustrate how these predictions change with different discount factors and time horizons.

After presenting our empirical results, we return in Section V to the model, extending it to allow for hyperbolic discounting and contrasting the predictions from the two approaches.

II.A. The Economic Model of Criminal Behavior

Consider a dynamic version of the model given in Becker (1968). Throughout his (infinite-period) lifetime, the individual chooses whether or not to commit a crime in each period t during which he is free. If he chooses to commit a crime, with probability p_t he will be apprehended and punished, leading to the payoff U_t^s, the lifetime utility properly discounted to period t. With probability $1 - p_t$, the offense is undetected, and he obtains lifetime utility U_t^{cr}. If he chooses not to commit the crime in period t, he receives lifetime utility U_t^a. Thus, at each point in time t, the individual commits crime if and only if

$$p_t U_t^s + (1 - p_t) U_t^{cr} > U_t^a$$

where we assume $U_t^s < U_t^a < U_t^{cr}$. Equivalently, he will commit crime if and only if the probability of apprehension is less than a "reservation probability", given by

$$\overline{p}_t = \frac{U_t^{cr} - U_t^a}{U_t^{cr} - U_t^s} \tag{1}$$

For any given p_t, as the rewards to crime rise (an increase in U_t^{cr}), the payoff to legal activity falls (a decrease in U_t^a), or the criminal sanctions become less punitive (an increase in U_t^s), the individual is more likely to commit an offense, because it is more likely that p_t will fall below the reservation probability $\overline{p}_t$.

To allow for a stochastic element to the model, suppose that in each period t, the individual receives a random draw from a distribution of criminal opportunities. Each opportunity is more or less attractive as p_t is low or high. For example, the least valuable opportunity is one for which apprehension is assured, (i.e., one for which $p_t = 1$). Letting $F_t(\cdot)$ denote the cumulative distribution function of p_t, the probability that the individual commits a crime in period t is thus

$$\Pr(\text{Crime in } t) = F_t(\overline{p}_t)$$

II.B. Identification Strategy

Our identification strategy exploits the fact that in the United States, the severity of criminal sanctions depends discontinuously on the age of the offender at the time of the offense. In all 50 U.S. states, offenders younger than a certain age, typically 18, are subject to punishments determined by the juvenile courts. The day the offender turns 18, however, he is subject to the more punitive adult criminal courts. The criminal courts are known to be more punitive in a number of different ways. For example, as we document below, the expected length of incarceration is significantly longer when the offender is treated as an adult rather than as a juvenile.

We follow a cohort of youth longitudinally and examine whether there is a discontinuous drop in their offense rates when they turn 18. Our data contain exact date of birth and exact date of offense, allowing us to examine the timing at high frequencies. A high frequency approach is important, because it allows us differentiate between secular age effects (Grogger 1998, Levitt and Lochner 2001) and a response to the increase in the price of crime.

Illustrating our approach, Figure I plots criminal propensities at each point in time, under two hypothetical scenarios. In the first scenario, represented by the solid line, the individual is patient and possesses a low discount rate (high discount factor). The probability of offense drops discontinuously at the age of 18, because at that age the expected cost of crime jumps significantly, leading crime to be less attractive.[6]

In the second scenario, the individual is impatient and possesses a high discount rate. Here, even facing the same increase in incarceration lengths as above, the individual perceives the jump in the price of crime to be much more modest. As a result, the drop in criminal propensity is not as great. Our empirical analysis amounts to determining whether the data are relatively more consistent with the first or second scenario.

Our approach does not require that the determinants of criminal behavior be constant throughout the individual's life. Instead, it relies on the arguably plausible assumption that determinants of criminal propensity—other than the severity of punishments—do not change discontinuously at 18. For example, we are arguing that on average, wages do not "jump up" the day after an

[6] Note that we are estimating what is known in the criminology literature as "general deterrence", which is the effect of the threat of punishment on criminal behavior. General deterrence is the phenomenon traditionally addressed in the economics of crime. By contrast, Pintoff (2004) identifies what is known in criminology as "specific deterrence", or the extent to which a prison experience will lower post-release criminal propensity.

individual turns 18, and that the distribution of criminal opportunities does not systematically and dramatically change between the day before and day after the 18th birthday. As another example, we consider it plausible that the ability of law enforcement to apprehend an offender does not discontinuously change between the week before and the week after the agent's 18th birthday.

We emphasize that we only believe that "all other factors" are roughly constant when examining offense rates in relatively short intervals (e.g., one day, or one week). By contrast, we consider it likely that all determinants of criminal behavior are changing significantly from year to year. In the age range of 17, 18, or 19, for example, youth could be graduating from high school, starting a new job, or developing physiologically and psychologically in ways that could affect underlying criminal propensities.

Formally, let t_0 be the week following the individual's 18th birthday, and t be measured in week intervals so that $t_0 - 1$ is the week before his 18th birthday. We are arguing that it is plausible that $U^{cr}_{t_0} = U^{cr}_{t_0-1}$, $U^{a}_{t_0} = U^{a}_{t_0-1}$, and $F_{t_0}(\cdot) = F_{t_0-1}(\cdot)$: the lifetime utilities from the crime, from the legitimate activity, and the distribution of available criminal opportunities are approximately the same for the two periods. The only difference between periods $t_0 - 1$ and t_0, we are arguing, is that lifetime utility will be different if apprehended: $U^{s}_{t_0} \neq U^{s}_{t_0-1}$.

This key smoothness assumption implies that

$$\Pr(\text{Crime at } t_0 - 1) = F_{t_0}(\overline{p}_{t_0-1}), \text{ where } \overline{p}_{t_0-1} = \frac{U^{cr}_{t_0} - U^{a}_{t_0}}{U^{cr}_{t_0} - U^{s}_{t_0-1}}$$

$$\Pr(\text{Crime at } t_0) = F_{t_0}(\overline{p}_{t_0}), \text{ where } \overline{p}_{t_0} = \frac{U^{cr}_{t_0} - U^{a}_{t_0}}{U^{cr}_{t_0} - U^{s}_{t_0}}$$

Thus, we obtain the unambiguous prediction that the discontinuous jump in the punitiveness of sanctions at age 18 will lead to a corresponding discontinuous drop in the probability of offending.

II.C. Estimation

We implement our identification strategy by providing an empirical analogue to the profile depicted in Figure I. We follow a fixed cohort of individuals and estimate the probability of offending for every week, starting at the 17th birthday, and ending on the 19th birthday. Specifically, we estimate the logit

$$\ln\left(\frac{F_t\left(\overline{p}_t\right)}{1 - F_t\left(\overline{p}_t\right)}\right) = X_t'\alpha + D_t\theta \qquad (2)$$

where $X_t \equiv \left(1, t-t_0, (t-t_0)^2, \ldots, (t-t_0)^q\right)'$, q is the order of the polynomial, t indexes age in weeks, t_0 is the week of the 18th birthday, and D_t is 1 if $t \geq t_0$, and is 0 otherwise.[7] The reduced-form parameter of interest is θ, the discontinuous change in the log-odds of committing an offense when the youth turns 18 and immediately becomes subject to the adult criminal courts.

The logit model is a simple way to parametrically describe the time profile of criminal propensities. A completely non-parametric approach would proceed as follows. Let N be the total number of individuals in our sample of 17 year-olds, and let n_1 denote the number of individuals who commit their first post-17 offense in the first week, n_2 the number who commit their first offense in the second week, and so on. We could then simply compute the first three values of $F_t\left(\overline{p}_t\right)$ as $\frac{n_1}{N}$, $\frac{n_2}{N-n_1}$, $\frac{n_3}{N-n_1-n_2}$, for example.[8]

We present these nonparametric estimates graphically in our empirical analysis. In addition, we use the logit in Equation (2) to fit the nonparametric estimates to a parametric form and to provide more efficient estimates of the discontinuity at age 18.

Note that we utilize arrest information as a proxy for criminal offending—as is common in the crime literature—since error-free criminal offending data does not exist. Thus, strictly speaking, we cannot estimate Equation (2), which expresses the odds of *offending* as a function of age.

Nevertheless, this does not pose a problem for our identification strategy, since we have already assumed that the distribution $F_t(\cdot)$ is evolving smoothly and does not discontinuously change at 18. As discussed above, this means that the risk of apprehension for any fixed criminal opportunity does not change abruptly upon an individual's 18th birthday. Thus, we henceforth re-define $F_t(\overline{p}_t)$ as the probability as of time t of being *arrested* for the first time since turning 17, defining $F_t(\overline{p}_t) \equiv F_t^*(\overline{p}_t)E_{\overline{p}_t}$, where $F_t^*(\cdot)$ is the distribution of criminal opportunities, and $E_{\overline{p}_t} \equiv E_t[p_t|p_t < \overline{p}_t]$ is the expected apprehension rate conditional on committing the crime. Since both $F_t^*(\cdot)$ and $E_{\overline{p}_t}$ are monotonically increasing in $\overline{p}_t$, the qualitative predictions are unchanged when using arrests.

II.D. Predicted Magnitudes

In this sub-section, we address the following question: How large should we expect θ to be if individuals are forward-looking and have rational expectations? The answer to this question

[7] The vector X_t can also include interactions of the polynomial in t with the indicator D_t.

[8] This evolving "risk-set" is also used in estimation of the logit model. Thus, the data is an unbalanced panel, where there are N observations for the first period, $N - n_1$ for the second, $N - n_2$ for the third, and so on. Efron (1988) points out that the logit model described estimates the hazard function.

requires quantifying the difference between punishments in the juvenile and adult systems. Our calibration exercise focuses on the differences in expected incarceration length. We discuss other qualitative differences between the punitiveness of the two systems at the end of this sub-section.

Appendix B provides some simple estimates of the expected duration of incarceration, conditional on an arrest. For Florida, the average incarceration length for juveniles arrested for Index crimes range from 2.7 to 6.5 weeks. Adult incarceration lengths are 3.1 to 7.4 times as long.[9] We emphasize that these numbers are conditional on arrest, and will represent an average of long periods (for convicted felons) and very brief periods of incarceration (for those who are released shortly after an arrest). Thus, we broadly define "incarceration" to be any period in which an arrestee is in custody of the criminal justice system (e.g., awaiting trial).

Using these estimates in conjunction with age-specific arrest rates and assumptions on discount factors, we derive quantitative predictions for the size of the reduced-form discontinuity

$$\theta \approx \ln\left(\frac{F_{t_0}(\overline{p}_{t_0}(\delta, S_J, S_A))}{1 - F_{t_0}(\overline{p}_{t_0}(\delta, S_J, S_A))}\right) - \ln\left(\frac{F_{t_0}(\overline{p}_{t_0-1}(\delta, S_J, S_A))}{1 - F_{t_0}(\overline{p}_{t_0-1}(\delta, S_J, S_A))}\right) \tag{3}$$

where the notation emphasizes that $\overline{p}_{t_0}$ and $\overline{p}_{t_0-1}$ are functions of the weekly discount factor δ, incarceration length for juveniles, S_J, measured in weeks, and the incarceration length for adults, S_A. To make these quantitative predictions, we invoke the following assumptions:

A1. **Linear Approximation of** $F^*_{t_0}(\cdot)$. Consider a first-order approximation around $\overline{p}_{t_0} = 0$ (where $F^*_{t_0}(0) = 0$), so that we have $F^*_{t_0}(\overline{p}_{t_0}) \approx f \cdot \overline{p}_{t_0}$, with f a positive constant.[10] With this approximation, it follows that $E_{\overline{p}_{t_0}} \equiv E[p_{t_0}|p_{t_0} \leq \overline{p}_{t_0}] \approx \frac{1}{2}\overline{p}_{t_0}$ and $F_{t_0}\left(\overline{p}_{t_0}\right) \approx \frac{1}{2} f \overline{p}^2_{t_0}$.

A2. **Stationarity.** In each period, the flow utilities from committing a crime undetected, from the alternative legal activity, and from being caught, are u_{cr}, u_a, and u_s, respectively, for all $t \geq t_0 - 1$. Similarly, the distribution of criminal opportunities is the same over time: $F^*_t(\cdot) = F^*(\cdot)$ for all $t \geq t_0 - 1$. Furthermore, we assume that no matter how many incarceration spells the offender has experienced, he faces the same dynamic optimization problem whenever he is released.

A3. **Structure of Punishments.** If apprehended, the individual must remain incarcerated for S_J (if caught as a juvenile) or S_A (if caught as an adult) periods. After this incarceration, the individual is released and faces the original decision problem. For each period during the incarceration, he receives flow utility u_s. This implies that an offender who is

[9] Our estimates are in the the range of the case-control matching approach to sentencing differences found in the criminological literature (for a review, see Redding 2003).

[10] Glaeser and Sacerdote (2000, Appendix 1) use a first-order approximation in a similar context.

apprehended as an adult and is incarcerated for S_A periods evaluates his time in prison as $u_s\left(1+\delta+\delta^2+\cdots+\delta^{S_A-1}\right)=u_s+\frac{\delta-\delta^{S_A}}{1-\delta}u_s$.

As we explain below, with these simplifications, if we choose $S_J=2$ (weeks), $S_A=6$, $\delta=0.999$ (an annualized discount factor of 0.95), the expected probability of apprehension $E_{\overline{p}_{t_0}}$ to be 0.1, and $F_{t_0}\left(\overline{p}_{t_0}\right)$ to be 0.0025, the predicted θ is -1.54, or about an 80 percent decline in the odds of arrest at 18 in response to an increase in the length of incarceration from 2 to 6 weeks.

II.E. Construction of Prediction for θ

We use the numerical example discussed above to illustrate the construction of the predicted θ. In order to compute the predicted θ, it is sufficient to solve for the following unknowns: f, u_{cr}, u_a, and u_s. To solve for f, we first use the relation $E_{\overline{p}_{t_0}}\approx\frac{1}{2}\overline{p}_{t_0}$ and our estimate of $E_{\overline{p}_{t_0}}=0.1$, which implies $\overline{p}_{t_0}=0.2$. The estimate of $E_{\overline{p}_{t_0}}$ is roughly based on FBI statistics on "clearance rates"—the fraction of reported crimes that end in an arrest of a suspect. Appendix B provides further details on this number.[11] Using $\overline{p}_{t_0}=0.2$, the relation $F_{t_0}\left(\overline{p}_{t_0}\right)\approx\frac{1}{2}f\overline{p}_{t_0}^2$, and our own estimate of the probability of arrest at age 18 of $F\left(\overline{p}_{t_0}\right)=0.0025$ yields $f=0.125$.

As shown in Appendix A, assumptions A2 and A3 result in the expressions

$$\overline{p}_{t_0-1}=\frac{u_{cr}-u_a}{u_{cr}-u_s+\frac{\delta-\delta^{S_J}}{1-\delta}((1-\delta)V-u_s)} \tag{4}$$

$$\overline{p}_{t_0}=\frac{u_{cr}-u_a}{u_{cr}-u_s+\frac{\delta-\delta^{S_A}}{1-\delta}((1-\delta)V-u_s)} \tag{5}$$

where

$$(1-\delta)V=\frac{F^*(\overline{p}_{t_0})(1-E_{\overline{p}_{t_0}})u_{cr}+(1-F^*(\overline{p}_{t_0}))u_a+\frac{1-\delta^{S_A}}{1-\delta}F^*(\overline{p}_{t_0})E_{\overline{p}_{t_0}}u_s}{1+\frac{\delta-\delta^{S_A}}{1-\delta}F^*(\overline{p}_{t_0})E_{\overline{p}_{t_0}}} \tag{6}$$

It can be shown that—as with all expected utility problems of this type—optimal choices are invariant to affine transformations of the flow utilities. Thus, it is innocuous to normalize $u_s=1$ and $u_a=2$, for example. After substituting (6) into (4) and (5), we obtain two equations with two unknowns, u_{cr} and $\overline{p}_{t_0-1}$. Using $\delta=0.999$, $S_J=2$, and $S_A=6$, we obtain $u_{cr}\approx 3.52$, and $\overline{p}_{t_0-1}\approx 0.43$. We then use Equation (3) to calculate $\theta\approx -1.54$.

[11] It is worth noting that these data may overstate the probability of arrest, since pressure to "clear" offenses may lead to inflated counts. This may cause our calibration to be too conservative regarding the magnitude of θ.

II.F. Intuition Behind Prediction for θ

A more detailed discussion of the derivation of (4), (5), and (6) is provided in Appendix A. For now, we simply note a number of intuitive features of these expressions. First, when $\delta = 0$, $\overline{p}_{t_0}$ and $\overline{p}_{t_0-1}$ collapse to $(u_{cr} - u_a)/(u_{cr} - u_s)$, the reservation probability from the static economic model of crime (see, for example, Freeman 1999). This is intuitive since, for example, no criminal opportunity will be worth doing when $u_{cr} = u_a$, leading to $F(\overline{p}_{t_0}) = F(\overline{p}_{t_0-1}) = 0$.

Second, V is the expected discounted utility for the rest of an individual's life at any point in which he is free to commit a crime; $(1 - \delta)V$ is its "annuitized" value. In this numerical example, $(1 - \delta)\, V \approx 2.02$; it is strictly higher than $u_a = 2$, because there is an "option value" of being able to engage in crime. The per-period cost of being incarcerated is $(1 - \delta)V - u_s$. The total cost to the apprehended individual is $u_{cr} - u_s + \frac{\delta - \delta^{S_A}}{1-\delta}\left[(1 - \delta)V - u_s\right]$. Note that V is the same at both $t_0 - 1$ and t_0, because in both periods, the individual will face the same (adult) problem in the subsequent period.

Third, equation (6) shows that $(1 - \delta)V$ is a weighted average of the three flow utilities u_{cr}, u_s, and u_a, with weights given by the "steady-state" probabilities of receiving those utilities in a randomly chosen time period.

Finally, the intuition behind the calculation can be summarized as follows. First, estimates of the probability of being apprehended $E_{\overline{p}_{t_0}}$ allow us to approximate the reservation probability $\overline{p}_{t_0}$. Our first order approximation of $F(\cdot)$ and estimate of $F\left(\overline{p}_{t_0}\right)$ give us an approximation of how much the arrest probability changes with a change from $\overline{p}_{t_0-1}$ to $\overline{p}_{t_0}$. The reservation probability $\overline{p}_{t_0}$ yields information on the relative payoff of crime u_{cr}. Intuitively, a large $\overline{p}_{t_0}$ means that $u_{cr} - u_a$ is large relative to $u_a - u_s$; a small $\overline{p}_{t_0}$ implies just the opposite. Finally, with a value for u_{cr} we may calculate $\overline{p}_{t_0-1}$, and this allows us to calculate θ.

II.G. Sensitivity of Predictions

The use of exponential discounting leads to an important implication: increases in incarceration lengths for short periods (e.g., 3 versus 9 weeks) will lead to a *larger* percentage response—that is, a larger θ—than will increases at longer periods (e.g., 3 versus 9 years). We illustrate this in Figure II. In the top panel of Figure II, we maintain $\overline{p}_{t_0} = 0.2$, but vary (S_J, S_A) from $(1, 3)$ to $(100, 300)$

weeks. For each value of S_J, we consider what θ would be if S_A were 3 times as long as S_J.[12]

The solid squares show the predicted θ, when $\delta = 0.999$ (corresponding to an annual discount factor of 0.95). The predicted θ begins close to -1.54 (consistent with an 80 percent decline in the odds of arrest) and declines only slightly in magnitude as S_J and S_A increase. That the predicted magnitude should decline as S_J increases is intuitive. Periods in the near future are given roughly equal weight, so that a tripling of incarceration lengths from 5 to 15 weeks will represent a roughly three-fold increase in the cost. On the other hand, for an increase from 5 to 15 years, utility in the additional 10 years is weighted much less than utility in the initial 5 years.

This principle is illustrated more dramatically by the solid circles, which performs the same calibration, but setting $\delta = 0.987$ (corresponding to an annual discount factor of 0.50). Under this scenario, an increase in incarceration from 100 to 300 weeks will lead to $\theta = -0.5$, or a 40 percent reduction in the odds of an arrest at 18. The calculation for an increase in incarceration from 5 to 15 years (not shown in the figure) yields $\theta = -0.05$, or a 5 percent decline in the odds of arrest at 18. The solid triangles present the analogous computations for $\delta = 0.957$ (corresponding to an annual discount factor of 0.1), which exhibits a similar, but more extreme profile than the other discount rates considered.

In the bottom panel of Figure II, we illustrate the same calculations for $\overline{p}_{t_0} = 0.5$, which is the largest $\overline{p}_{t_0}$ consistent with clearance rates, as described in Appendix B. Note that this value implies that an individual will commit an offense if he perceives the probability of being apprehended to be 0.5 or less. The three curves have a similar shape, but the predictions using $\overline{p}_{t_0} = 0.5$ are roughly half as large in magnitude as those using $\overline{p}_{t_0} = 0.2$.

II.H. Other Factors Affecting θ

The calculation of our predicted θ is based on a very parsimonious model, and clearly ignores a few potentially important factors. First of all, the adult and juvenile justice systems differ in ways other than in the expected incarceration length. There is general agreement that the conditions of incarceration are worse in adult than juvenile facilities (Myers 2003).[13] One inmate quoted in

[12] Note that for each value of S_J, we re-calculated the predicted θ as described above, so there is a unique u_{cr} (and V) associated with each S_J, S_A pair.

[13] Arguably the qualitative aspect of confinement most salient for a youthful offender is sexual assault. A widely-cited study estimates that juvenile offenders held in adult prison were sexually assaulted at a rate five times that of juvenile offenders held in juvenile facilities (Forst, Fagan and Vivona 1989).

Levitt (1988) observed that "[When] you are a boy, you can be put into a detention home. But you can go to jail now [as an adult]. Jail ain't no place to go."[14] Moreover, these qualitative differences may be most important for youthful offenders.[15]

Second, we also do not incorporate the other advantages of being apprehended as a juvenile rather than as an adult. For example, it is generally easier to expunge or seal arrest or conviction records if the offense was committed as a juvenile rather than an adult. There may also be a labor market penalty to having a criminal record (compared to a clean record).[16]

Third, our dynamic model of criminal behavior does not allow for a richer kind of intertemporal decision making. That is, we are assuming that the distribution of opportunities is strictly exogenous: stealing a car today does not affect the opportunities for stealing a car next week. A richer model would allow the individual to explicitly choose the optimal timing of a particular crime.

Although our simple model ignores these factors, we conjecture that most reasonable formalizations of these additional factors would work to produce a θ that is larger in magnitude, for the same S_J, S_A and δ. In addition, as we show in Appendix B, the ratio S_A/S_J could be as high as 7. In sum, we believe that our predictions for the magnitude of θ are probably somewhat conservative.

III. Data and Sample

Our analysis uses an administrative database maintained by the Florida Department of Law Enforcement (FDLE). Essentially, the data consist of all recorded felony arrests in the state of Florida from 1989 to 2002. The database includes exact date of birth, gender, and race for each person. For each arrest incident, there is information on the date of the offense, date of arrest, the county of arrest, the type of offense, whether or not the individual was formally charged for the incident, and whether or not the incident led to a conviction and prison term. Most importantly, the data are longitudinal: each arrest incident is linked to a person-level identifier. A more detailed description of the database and its construction is provided in Appendix C.

Our data has two important advantages over typically available arrest data. First, the data are

[14] The quotation is from interview research conducted by Glassner, Ksander, Berg and Johnson (1983, p. 219, brackets in original).

[15] The youngest inmates in an institution are typically targeted for sexual victimization (Robertson 1999, Saccenti 2000). The age pattern of victimization is relevant since the marginal adult offender, if confined, will be youngest among his peer inmates, whereas the marginal juvenile offender, if confined in a juvenile facility, will be oldest among his peer inmates.

[16] We suspect this effect is minor in light of evidence given in Kling (2004).

high-frequency, allowing us to pinpoint age at offense in weeks rather than years. Second the data are longitudinal. In our context, these aspects are necessary to distinguish between deterrence and incapacitation effects.

To illustrate the value of these aspects of our data, we briefly illustrate the inference problems that arise when using the more readily available, but coarser, cross-sectional data on arrests. Figure III plots the frequency distribution of age (by year), for all arrestees and those arrested for index crimes,[17] as computed from the two most readily available, public-use data sources on arrests: the National Incident-Based Reporting System (NIBRS) and the Uniform Crime Reports (UCR).

The figure shows clear secular age effects in arrest rates, with an apparent trend break around 18. This is consistent with a change in criminal behavior due to more severe punishments for adults at 18. But it is also consistent with criminal behavior declining as youth finish high school and enter the labor market (Grogger 1998, Lochner 2004). A third—and arguably simpler—interpretation of the peak at 18 is that it is entirely due to an incapacitation effect. Starting at 18, arrestees are confined for longer periods, and therefore are temporarily removed from the community, and thus cannot commit crime. Fewer and fewer arrests at older ages will occur as offenders are incapacitated, leading to a mechanical trend break at 18.

By comparison, our approach guarantees that we isolate only a deterrence effect. Specifically, our main estimation sample consists of all persons who have been arrested prior to age 17. Starting at the 17th birthday, we begin tracking this population over time, and for each subsequent day, estimate the probability of being arrested. A discontinuous drop in this arrest rate between day 365 and 366 will be due to deterrence. This is because the population still at risk of experiencing their first-post-17 arrest will be virtually identical for days 365 and 366.

There are a number of advantages of our main estimation sample compared to the remainder of youth who have *not* been arrested as of age 17.[18] First, our sample will include individuals who will more likely be affected by the increase in sanctions. In particular, it seems likely that for this group, there is a positive net benefit to criminal activity. After all, an individual in our sample

[17] Index crimes are murder, robbery, aggravated assault, rape, burglary, larceny, and motor vehicle theft.

[18] The choice of an age 17 threshold in constructing our sample was based on two competing interests. A higher threshold would have the advantage of raising the sample size, but it would give us a shorter period from which to estimate the probability of arrest on the left side of the discontinuity threshold. In the extreme, if we had chosen all those who had been arrested prior to age 17.99, there would be almost no data to estimate the left side of the discontinuity. A lower threshold would yield more periods from which to estimate the left part of the function (as shown in Figure I), but would result in a lower number of individuals present in the first period.

has already been arrested at least once by 17, suggesting that at least one crime was worthwhile to the juvenile. By contrast, those who have not been arrested as of 17 could potentially include many youth who have virtually no chance of committing a serious crime. Indeed, this law-abiding population may well be at a "corner solution", whereby $u_a > u_{cr} > u_s$. We would not expect to see any impact of more punitive sanctions for this sub-population.

Second, it is plausible that those who have already been arrested by age 17 are more likely to understand that there is a difference between the juvenile and adult criminal justice systems; they may even have been warned about this fact upon their initial arrest. Glassner et al. (1983) provide anecdotal evidence to support this.[19]

Third, our sample is not likely to be affected by expungement or sealing of criminal records. In Florida, as in most states, it is possible to have criminal records sealed or expunged. If juvenile records were systematically missing relative to adult records, then we would be biased against finding a deterrence effect. As we demonstrate empirically in Section IV, by selecting on having at least one juvenile record, we focus on a subpopulation whose entire juvenile arrest record is observed.

Overall, we believe our sample is reasonably representative of the population that is likely to be affected by the increase in sanctions. Table I reports some summary statistics for our main estimation sample. We begin with 64,073 individuals who have been arrested at least once by their 17th birthday. As is common in criminal justice data sets, 80 to 90 percent of these arrestees are male, and roughly 50 percent are non-white. The first three columns present information on the first, second, and third arrest for these individuals. Age at first arrest is about 15, and the most common category of offense is property crime, followed by violent crime. At second and third arrests, offenders are over 16 and 17 on average, respectively, and the offense mix is similar, with drug crimes becoming somewhat more prevalent. At all three arrest snapshots, individuals are distributed evenly among small, medium, and large counties.[20]

Column (4) of Table I presents means for the first arrest since 17, which is observed for slightly

[19] For example, responding to a question regarding how he knew that sanctions were more punitive after the age of majority, one twelve-year-old interviewed by the authors who was earlier arrested for stealing from cars responded that the police had told him so: "Police come in our school and a lot of stuff, and I get caught and they tell [sic] me that" (p. 220).

[20] We classified counties according to total arrests in the FDLE data. Medium counties are Franklin, Palm Beach, Duval, Pinellas, Polk, Escambia, and Volusia. Large counties are Miami-Dade, Broward, and Orange. Remaining counties are classified as small.

less than half of our main estimation sample. The sex and gender composition of those observed at their first arrest since 17 is similar to those observed at their first, second, and third arrests, as is the county-size distribution. Offenses are distributed somewhat more evenly among the four crime types described. The final column reports the same statistics for all arrests where the individual was 17 or 18 years old at the time of arrest. From the means, it is apparent that our sample is broadly representative of this larger arrest population.

IV. Results

Our main result is summarized by Figure IV, an empirical analogue to Figure I. Each open circle in the figure represents the fraction of the population that is arrested for each week subsequent to the 17th birthday. For example, the first circle shows that about 0.005 of the sample are arrested within a week of their 17th birthday. In the week of their 18th birthday almost 0.0025 are arrested. These numbers are calculated as a proportion of the sample that has not yet been arrested; that is, the top panel of Figure IV is an empirical hazard function for being arrested, starting at age 17. The solid line gives predicted probabilities of arrest, based on maximum likelihood estimates of the logit in equation (2).[21]

The figure shows little indication of a systematic drop in arrest rates at the age of 18. The arrest probability literally does fall between the week before and after the 18th birthday, but that drop does not appear to be unusual, as compared to week-to-week differences ranging from age 17 to 19.

For comparison, the bottom panel of Figure IV plots the analogous weekly arrest probabilities for those who were arrested at least once by age 19. We track the arrest records for individuals in this "falsification" sample for two years, from 19 to 21. At each point over this time span, the offenders face the same punishment schedule (i.e., that of the adult criminal justice system) leading to arrest probabilities which are smooth in age, as expected. Empirically, the top and bottom panels of the figure are remarkably similar.

Table II reports estimated discontinuities in arrest probabilities at 18, based on the logit model of Equation (2). These estimates support the inference suggested by Figure IV: the drop in arrests

[21]Here, X_t is a cubic polynomial in t. These predictions correspond to estimation of the model in column (1) of Table II.

at 18 is small in magnitude and statistically insignificant. The estimated discontinuity is roughly -0.018, with a standard error of about 0.047. These estimates imply that we can statistically rule out values of θ more negative than -0.111.[22] These estimates are much smaller than the predictions of our theoretical model—even with an annual discount factor of 0.1, our model predicts that θ should be in the range of -0.8 to -1.54.[23]

The estimated discontinuity is robust to changes in specification. Moving from left to right in Table II, we control for an increasing number of factors. Column (1) gives our most parsimonious model, controlling only for a juvenile/adult dummy and a cubic polynomial in age at current arrest; column (8) gives our most complex model, adding controls for race, size of county in which the baseline arrest[24] occurred, offense type of baseline arrest, and a quintic polynomial in age at baseline arrest. In each column, the added controls are good predictors of the probability of arrest, but in no case does including additional controls affect the estimated discontinuity importantly.

Appendix Table III explores the sensitivity of the estimates to functional form. It reports the estimated θ for different orders of the polynomial, ranging from a linear to a quintic polynomial in time and allowing for interactions of the polynomial with the juvenile/adult dummy. The models are also tested against an unrestricted specification, where the polynomial and the dummy are replaced with a full set of week-dummies. Overall, the linear and quadratic specifications are apparently too restrictive, and can be statistically rejected by a test against the unrestricted model. For richer specifications, including ones that include a linear term interacted with juvenile/adult status, the point estimates range from -0.065 to 0.029, with none of the estimates being statistically significant. A similar pattern is found when baseline covariates are included.

We interpret our results as showing a small behavioral response to more punitive criminal sanctions. Given our standard errors, we cannot rule out small deterrence effects, but our estimates are precise enough to easily rule out the benchmark magnitudes predicted by annualized discount factors of 0.95, 0.5, and even 0.1.

We now consider some important potential threats to our analysis and interpretation.

[22] Use of a one-sided test implies rejection of any θ more negative than -0.095.

[23] Below, we inflate our estimates to account for juvenile transfer.

[24] For those in our estimation sample, we refer to the first pre-17 arrest as the “baseline arrest”.

IV.A. Transfers of Juveniles to the Adult Criminal Court

The first threat to the validity of our research design is the possibility of a lack of a discontinuity in the "treatment". That is, while all adults are handled by the criminal courts, and most minors are handled by the juvenile courts, all states allow a juvenile offender to be transferred to the criminal courts to be tried as an adult. This is particularly of concern in Florida, where prosecutors have discretion on whether to try a juvenile arrestee as an adult.

In principle, prosecutors could be more likely to request that the case be transferred to the criminal justice system when the arrestee is almost 18. In the extreme case, *all* arrestees aged 17.8 or 17.9 could be transferred to the adult court, which would result in no discontinuous jump in the punitiveness of criminal sanctions and hence no "treatment".

Our data allow us to empirically rule out this possibility. The top panel of Figure V plots the probability of being formally charged as an adult as a function of the age at the first post-17 arrest. Each open circle represents the number of individuals against whom a formal prosecution was filed, expressed as a fraction of those who were arrested in that particular week.[25] There is a striking upward discontinuity at the age of 18; apparently, those who are arrested just before their 18th birthday have about a 0.2 probability of being formally prosecuted as an adult, while those arrested just after their 18th brithday have a 0.6 to 0.7 probability. The latter probability is not 1, because not all arrestees will have formal charges filed against them.

The bottom panel of Figure V provides further evidence of a discontinuity in the treatment, using a different measure of punishment. It instead plots the probability that the arrestee is eventually convicted and sentenced to either state prison or a county jail. Again, the figure shows a flat relationship between this measure of punitiveness and the age at arrest. There is a noticeable jump at age 18, from about 0.03 to 0.17.

We quantify these discontinuities in Table III, which reports coefficient estimates from different OLS regressions. Columns (1) through (5) simply regress the dummy variable for whether the individual was prosecuted on the juvenile/adult status dummy, a cubic polynomial in age at arrest, the same cubic polynomial interacted with the juvenile/adult dummy, and other covariates. Columns (6) through (8) further include the age at the baseline arrest as additional controls. Across spec-

[25]Therefore, the sample for this figure is the same as that underlying the top panel of Figure III.

ifications, the discontinuity estimate of about 0.40 is relatively stable. Appendix Table IV is an analogous table for the probability of being convicted as an adult and sentenced to prison or jail.

IV.B. Effects by Sub-groups and Crime Types

Another potential threat to the research design is the possibility that offenses committed by juveniles and adults have different likelihoods of being recorded in our data. For example, it is possible that law enforcement may exercise discretion in formally arresting an individual, based on age. In particular, if the probability of arresting an individual, conditional on the same offense, is substantially higher for an 18.1 year old than a 17.9 year old, then it is theoretically possible that the small effects we observe are a combination of a negative deterrence effect and a positive and offsetting jump in the arrest probability, due to law enforcement discretion.

There are a number of reasons why believe this is probably not occurring in our data. First, our analysis focuses on very serious crimes, where it seems unlikely that an officer would be willing to release a suspect without an arrest, purely on the basis of the individual's age. Index crimes include murder, robbery, rape, assault, burglary, larceny and motor vehicle theft, and involve a victim. We suspect that the pressure to capture a suspect is too great for officers to be willing to release an individual suspected of committing an Index crime. By contrast, for relatively less serious crimes such as misdemeanors, or victimless crimes such as drug possession, the selective release of offenders based on age is more plausible.

Second, our main estimation sample consists of individual who *already* have a recorded formal arrest as of age 17, when we begin following their arrest experiences. Thus, it seems unlikely that the law enforcement agency will exercise leniency in recording an arrest: as of age 17, it is too late to do anything to keep the youth's felony arrest record clean.

Third, our analysis focuses on arrests since 1994, the year of Florida's Juvenile Justice Reform Act (JJRA), which required that felonies committed by juveniles be forwarded to the state for inclusion in the criminal history records maintained by the FDLE.[26] The impact of this law on

[26] The implications of this Florida law was summarized by an attorney general opinion in 1995: "Under Florida law, crime and police records regarding crime have been a matter of public record. With limited exceptions, however, the identity of a juvenile who committed a crime has been protected. With the enactment of Chapter 94-209, Laws of Florida, an omnibus juvenile justice reform measure, the Legislature has amended the confidentiality provisions relating to juvenile offenders to allow for greater public dissemination of information. The clear goal of the Legislature was to establish the public's right to obtain information about persons who commit serious offenses, regardless of age" (Butterworth 1995, p. 274).

the prevalence of juvenile records is shown in Appendix Figure I. This figure show the ratio of arrests of 17-year-olds to 18-year-olds, by year and month, from 1989 to 2002. There is a marked discontinuity in the ratio after October 1994, the effective date of the JJRA.

Finally, if juvenile and adult arrests had different likelihoods of being recorded in our data, we would expect to observe significant heterogeneity in the estimated θ, by different groups of individuals, and different crime types, since it is likely that any off-setting measurement problems will vary by characteristics of the individual, as well as by crime type. Figure VI provides evidence contrary to this prediction. The top panel of the figure disaggregates the arrest probabilities from the top panel of Figure IV into two components: property and violent crime. The figure shows that the estimated discontinuity is essentially the same for the two categories of crime.

We also estimate θ separately by sub-groups defined by key correlates of arrest propensities, and find no evidence of significant negative effects for some groups being masked by positive effects of other sub-groups. Table IV reports estimates from interacting the juvenile/adult dummy with race, size of county of the baseline arrest, and offense of the baseline arrest. The estimates for these different sub-groups range from -0.07 to 0.09. These estimates are generally of small magnitude; moreover, none of the 20 are statistically significant. Finally, we fail to reject the null hypothesis that the interaction effects are equal to the main effect presented in Table II, and this holds for all specifications considered.

Although our analysis focuses on Index crimes, we show for completeness the results for all remaining offenses in the bottom panel of Figure VI. We consider the potential for arrest discretion to be the most serious for these non-Index crimes, which include "victimless" offenses such as drug possession, or resisting arrest felonies, for example. Here, the cubic polynomial predictions do show a small perverse discontinuity, although the simple means, represented by the squares and circles, do not reveal an obviously compelling jump at age eighteen. Thus, the relative smoothness for these arguably "worst-case" crime categories, as well as their similarity between the drug and non-drug stratification are also consistent with minimal arrest discretion at age eighteen.

IV.C. Expungement of Records

Here, we describe in greater detail how the longitudinal nature of our data allows us to circumvent problems caused by potential expungement and sealing of juvenile records. In general, the ability to expunge or seal one's records could potentially generate a positive discontinuity in arrest probabilities. That is, if we observe a much greater number of arrestees of age 18.1 compared to 17.9, this could be entirely due to some fraction of those of age 17.9 erasing their juvenile records, which would therefore not be included in our database. More generally, this expungement effect could offset any true negative deterrence effect.

Florida law allows individuals who successfully complete a juvenile diversion program to apply to have all juvenile records expunged (Fla. Stat. 943.0582). Apart from this provision, Florida law also mandates that juvenile arrest histories be expunged when the individual turns 24.[27]

Our sample circumvents these two expungement provisions in the following way. First, our estimation sample is restricted to those committing baseline crimes before age 17 but subsequent to January 1, 1995. Therefore, the individuals are no older than 24 by the end of our sample frame, and thus will not be subject to the time-activated expungement. Second, a requirement for inclusion in our sample is an arrest record prior to age 17. These individuals therefore did not have their complete juvenile arrest history expunged.

To illustrate how our sample avoids the expungement problem, we examine the time profile of arrests for the individuals who are *not* included in our sample. For this remainder sample of individuals who were *not* observed as arrested prior to turning 17. Some of these individual's first real arrest will occur before eighteen, and for some it will occur after eighteen. If there is an opportunity for the former group to later expunge their juvenile records, then a positive discontinuity in the number of arrests should occur at age eighteen.

This is the pattern found in Figure VII. This figure is a "stacked" histogram, where the combined total represents the total number of people who are arrested for the first time since turning 17. The histogram is comprised of two populations, those arrests corresponding to our estimation sample (the dark bars) and the remaining, unused observations (the light bars). For the total, there is a striking positive discontinuity at age 18. But this discontinuity is entirely concentrated in the unused sample (the upper part of the stacked graph).[28]

[27]The exception to this is when the individual has committed a serious offense as an adult. "Habitual offenders'" juvenile records are retained by the FDLE until the offender is 26, (Fla. Stat. 943.0585, 943.059).

[28]The dark bars represent the values used in the upper panel of Figure III, except that Figure III normalizes each value with the "at-risk" population at each point in time to provide a probability value.

IV.D. Evidence on Incapacitation

We have argued that expected incarceration length becomes discontinuously more punitive at age 18. If this is correct, then subsequent to the arrest of two individuals—one just older than 18 and the other just younger than 18—we should expect to see the marginal adult offender committing fewer *subsequent* offenses due to this differential incapacitation. The data are strongly consistent with this prediction.

This is shown in Figure VIII. The top panel of the figure plots the average number of arrests that occur *after* the initial arrest as a function of the age at the initial arrest. Specifically, the leftmost open circle indicates that among those who are arrested the week after their 17th birthday, the average number of times they are arrested within the subsequent 30 days is a little less than 0.20. Among those arrested just before their 18th birthday, they are arrested approximately 0.20 times within the subsequent 30 days.

There is a discontinuity at age 18, with 18.02 year olds being arrested half as many times within the 30 day window. We believe the most plausible explanation for this difference is that being handled by the adult courts leads to a longer period of custody than does being processed as a juvenile. The 17.98 year old is released earlier and hence has a greater opportunity to commit more crimes, compared to the 18.02 year old. There seems to be no other differences between the 17.98 and 18.02 year-old arrestees. For example, soon after being initially arrested, the 17.98 year-old will turn 18, and from that point on both groups of individuals will be subject to the same adult regime.

The solid circles and open triangles plot the same kind of graph, except that we examine the number of arrests occurring within 120 and 365 days after the initial arrest. Since the number of arrests is cumulative, the averages are necessarily higher. The discontinuity estimate also grows in magnitude, but seems to level off at longer periods.

Note that the length of the follow-up period—30 days, 120 days, or 365 days—is arbitrary. The bottom panel of Figure VII plots the profile of discontinuity estimates using follow-up lengths ranging from 1 to 365 days. For example, the estimates at 30, 120, and 365 days in the bottom panel—emphasized with large solid triangles—correspond to the discontinuity estimates from the top panel of the figure. Overall, the bottom panel shows that even after 20 days, there is a large divergence in the cumulative number of arrests between those who are arrested as a 17.98 year-old and those arrested at age 18.02, for example. This divergence continues to grow, slowing down at around 100 days after the initial arrest.

We interpret this as evidence that incapacitation effects occur fairly rapidly after an arrest occurs. This finding is important for two reasons. First, it provides a potential explanation for why our magnitudes are so much smaller than those of, for example, Levitt (1996).[29] It is possible that the use of arrest data at relatively coarse intervals (age in years) could lead to estimates that combine both deterrence and incapacitation effects.

Second, Figure VIII constitutes further supporting evidence for our research design. Marginal juveniles and marginal adults, while similar in age, do appear to be given very different punishments, with differences in incarceration lengths occurring in the very short run.

V. Alternative Interpretations

As shown in Section IV.A, 20 percent of the juvenile arrestees in our sample are prosecuted as adults. This "noncompliance" means that our discontinuity estimates are "intent-to-treat" parameters, and must be inflated in order to be comparable to our theoretical predictions. After accounting for both noncompliance and sampling variation, we can rule out predicted values for θ more negative than -0.14.[30,31]

This magnitude is significantly smaller than is predicted by the theoretical model outlined in Section II. For example, consider the θ of smallest magnitude that is consistent with our estimates for S_J, S_A, and $E_{\overline{p}}$, and our choice of δ. As is apparent in Figure II, the magnitude of θ is smaller when we choose S_J and $E_{\overline{p}}$ to be as large as possible. Based on Appendix Tables I and II, this results in $S_J = 7$, $S_A = 21$, and $E_{\overline{p}_{t_0}} = 0.5$. Setting $\delta = 0.1$ leads to a predicted θ of -0.66. Using the estimate from column (1) of Table II, we reject at the 1 percent level the hypothesis that θ is equal to this value.

[29] Levitt (2004b) notes that the range of estimates of the elasticity of crime with respect to expected punishment is roughly -0.10 to -0.40. Our theoretical predictions are comfortably within this range. A 200 percent increase in the expected incarceration is predicted to lower arrest rates by 55-79 percent, leading to elasticity estimates of -0.275 to -0.39.

[30] Juvenile transfer leads us to scale up our estimated discontinuity (standard error) of -0.018 (0.0474) by 1/0.8=1.25, resulting in a point estimate (standard error) of -0.0225 (0.05925).

[31] Using a one-sided confidence region, we can rule effect sizes larger in magnitude than 12 percent.

The unusually low discount factors implied by our calibration exercise suggest that it is worthwhile to consider a number of reasonable alternative interpretations of the estimated magnitudes.

V.A. Myopia and Hyperbolic Discounting

One possibility is that our assumption of exponential discounting is too restrictive. With exponential discounting, an increase in incarceration from 2 to 6 days represents a larger increase in the *relative* "price" of crime, compared to an increase from 2 to 6 weeks, which in turn, represents a larger relative price increase than a change from 2 to 6 years.

A natural alternative to consider is the case of hyperbolic discounting, where short-run discount factors are smaller than long-run discount factors. Under hyperbolic discounting, a tripling of incarceration lengths can potentially have a smaller proportional response in the short-run than in the long run.

To illustrate this, we adopt quasi-hyperbolic discounting, where the discount sequence is given by $1, \beta\delta, \beta\delta^2, \beta\delta^3, \ldots$ with $\beta < 1$. This sequence captures the qualitative aspect of different discount factors for the short-run and the long-run. In the psychology and economics literature, this form of discounting is often used to model tastes for immediate gratification and self-control problems.

Here, we consider the case of the so-called *sophisticated* hyperbolic discounter, who has correct expectations about his own future preferences and actions. In Appendix A, we show that under this scenario, (4) and (5) are modified only slightly to become

$$\begin{aligned} \overline{p}_{t_0-1} &= \frac{u_{cr}-u_a}{u_{cr}-u_s+\beta\frac{\delta-\delta^{S_J}}{1-\delta}\left((1-\delta)V-u_s\right)} \\ \overline{p}_{t_0} &= \frac{u_{cr}-u_a}{u_{cr}-u_s+\beta\frac{\delta-\delta^{S_A}}{1-\delta}\left((1-\delta)V-u_s\right)} \end{aligned} \tag{7}$$

and the expression for $(1-\delta)V$ remains the same as in (6). Using these modified equations, we can compute the predicted θ in the exactly same manner as described in Section II, except that we must choose a value for β.

Figure IX illustrates the predicted θ that would result in response to a tripling of incarceration lengths, for different values of S_J, S_A, and δ, while letting $\beta = 0.05$. For comparison, the same curves from the bottom panel of Figure II (representing the exponential case) are provided in the figure. The open squares present predicted effects under hyperbolic discounting, with $\beta = 0.05$

and $\delta = 0.999$. These predictions show that for very short horizons—when S_J is 1 or 2 weeks—a tripling of the incarceration length at age 18 leads to a θ that is smaller in magnitude than 0.20. This is much smaller than the prediction under exponential discounting, which is about 0.80 in magnitude.

For the hyperbolic agent, the predicted θ *grows* in absolute value as we examine longer time horizons, as intuition would suggest: in the very short run, when discount factors are very low, additional periods of incarceration seem less costly, but in the longer run, when discount factors are higher, those additional periods of punishment become more important. Note that the predicted θ asymptotes to zero in the long run.

The open circles and triangles illustrate the predicted θ for the case of $\beta = 0.05, \delta = 0.987$ (annual discount factor of 0.5) and for the case $\beta = 0.05, \delta = 0.957$ (annual discount factor of 0.1). These two curves show a similar pattern. For very short time horizons, θ is relatively small in magnitude. It becomes more negative for longer horizons, and asymptotes toward the curves for the exponential case.

In summary, the average incarceration length for juveniles ranges between 2.6 and 6.4 weeks, and it is precisely within this time horizon that the exponential model has a difficult time explaining small discontinuity estimates. On the other hand, a certain degree of myopia, which allows short-run discount factors to be smaller than long-run discount factors, can explain small effects in short time horizons.

V.B. Inframarginal Criminals

Another potential explanation for our results is that our sample of offenders are "inframarginal" in the sense that large changes in the price of crime do not affect criminal behavior because the benefits of crime far exceed the costs. That is, in the discussion above, we conjecture that criminal offending does not respond significantly to increased incarceration lengths because the effective price of crime does not change significantly—due to the degree and form of the discounting. Alternatively, the price may be changing significantly, but so few crimes are prevented that the decline is imperceptible.

Figure X illustrates the point. It plots a hypothetical density for p. In the figure, most of the mass of the density is on the left side of the support, implying that most criminal opportunities carry a small risk of apprehension. Thus far, we have emphasized that impatience and myopia could lead the cutoff $\overline{p}$, to be highly inelastic, in which case an an increase in incarceration lengths could result in the small shift of $\overline{p}$ from line A to line B, for example.

But an alternative possibility is that offenders are reasonably patient and not myopic, and that an increase in incarceration lengths lead the cutoff $\overline{p}$ to fall significantly—but from line C to line D. Since the density is low in that region, offense rates would not change significantly under such a scenario. In other words, only a small fraction of criminal opportunities are affected by the change in the cutoff $\overline{p}$.

It should be noted, however, that a natural consequence of this alternative hypothesis is that the elasticity of criminal behavior with respect to law enforcement must be small. Intuitively, if most opportunities' benefits far exceed their costs, increases in the probability of punishment will only slightly affect offending rates. In the graph, increasing police will tend to shift the density of p to the right. This means that the "inframarginal criminal" hypothesis also implies very small marginal effects of police.

Thus, we cannot completely rule out the "inframarginal criminal" hypothesis, but we view it as unsatisfactory. The empirical literature on policing is well-developed, and most studies find large elasticities of crime with respect to policing levels (Levitt 2004a). For example, Di Tella and Schargrodsky (2004) estimate an elasticity of crime with respect to police of roughly -0.3. In sum, if the population that we study is sensitive to the likelihood of punishment, then it must also be true that there is a nontrivial number of criminal opportunities "on the margin" of being worthwhile.

V.C. Irrational or Uninformed Criminals

Finally, the theory that we utilize presumes that potential offenders are rational, and have correct expectations about the consequences of their behavior. This may simply be untrue. As an example, our empirical analysis shows quite clearly that the probability of being formally prosecuted as an adult rises discontinuously at age 18 (see Figure V). But it may well be that young offenders are unaware of this fact. They may be unaware that the average incarceration length for adults is at least 3 times the length for juveniles, and they may be unaware that being caught as an adult rather than as a juvenile implies an immediate difference in the length of incapacitation (as suggested by Figure VIII). Young offenders may not only be unaware of the actual probabilities of arrest, but also unaware that juveniles and adults are treated differently by the criminal justice system. These are all examples of how young offenders could be making systematic errors in estimation. Each of these possibilities could independently explain why the estimates θ could be so small.[32]

VI. Policy Implications

We view our results as inconsistent with the hypothesis that transferring juveniles to the adult court or lowering the criminal age of majority will reduce crime through deterrence. The evidence does suggest, however, that such policies could lead to a reduction in the incidence of crime through incapacitation. Prison expansion may be desirable, particularly if the crimes prevented by incarceration have a high social cost. On the other hand, the general equilibrium effects described by Freeman (1999) may offset these benefits: even as locked-up criminals are prevented from recidivating, in a long-run equilibrium, previously "crowded-out" criminals may simply take their place.

We also conclude—albeit more tentatively—that increasing the frequency of jail or prison sentences or otherwise lengthening periods of incarceration has limited value as a deterrent. Strictly speaking, our discontinuity estimate measures an evolutionary response to an anticipated change in the price of crime at a given point in time, while uniform increases in incarceration lengths represent an increase in price of crime along all points in the life cycle. It should be noted, however, that the two effects are connected. Specifically, in our model, the former is strictly greater in magnitude than the latter. This is because an anticipated change in punitiveness will raise the total punishment exclusively through the length of incarceration, while a uniform increase throughout the lifecycle will additionally make crime generally less attractive and hence reduce the per-period opportunity cost of being incarcerated.[33]

A caveat to this conclusion involves the case of hyperbolic discounting. If potential offenders have very high short-run discount rates but low long-run discount rates, then in principle, there

[32] On the other hand, as Levitt (1998) suggests, there is reason to believe that young offenders are, at a minimum, aware that it is better to be arrested as a juvenile than as an adult for the same crime.

[33] To see this, compare our discontinuity parameter, $\frac{\overline{p}_{t_0-1}}{\overline{p}_{t_0}} = \frac{u_{cr}-u_s+\frac{\delta-\delta^{S_A}}{1-\delta}((1-\delta)V-u_s)}{u_{cr}-u_s+\frac{\delta-\delta^{S_J}}{1-\delta}((1-\delta)V-u_s)}$, to the deterrence parameter associated with lowering S_A for all adult ages to S_J, given by, $\frac{\overline{p}^*_{t_0}}{\overline{p}_{t_0}} = \frac{u_{cr}-u_s+\frac{\delta-\delta^{S_A}}{1-\delta}((1-\delta)V-u_s)}{u_{cr}-u_s+\frac{\delta-\delta^{S_J}}{1-\delta}((1-\delta)V^*-u_s)}$. The latter is smaller in magnitude than the former because V^*, the new discounted expected utility of being free, is larger than V (since incarceration lengths have fallen).

can be large marginal effects for the sub-population who are to face long prison sentences. On the other hand, our estimates indicate that this sub-population may be relatively small—at least too small to affect the overall estimates for our main sample.

Finally, we note that our findings may indicate that the marginal criminal justice dollar is more effectively spent on raising the *probability* of imprisonment rather than the incarceration length. Even with highly impatient or myopic criminals, doubling the odds of punishment will double the effective price of crime.

VII. Conclusions

A longstanding notion in the economics of crime is that illegal behavior can be deterred by raising either the certainty or severity of punishment. In practice, this tradeoff is highly dependent on criminals' discount rates. This is because the only practical way to increase severity is to lengthen periods of incarceration, which will only raise the price of crime if discount rates are relatively low. If discount rates are very high, then incarceration can do little to deter illegal behavior. By contrast, elasticities with respect to the *likelihood* of incarceration are invariant to discount rates.

Learning about discount rates requires a setting that generates exogenous variation in the timing of future payoffs. We believe that our research design generates such exogenous variation. Upon turning 18, an individual faces an immediate increase in the length of incarceration. In addition, all other determinants of criminal behavior are arguably stable from one week before to one week after the 18th birthday.

Using this approach, we find that criminals are largely unresponsive to this sharp change in future penalties. We attempt to rationalize the magnitudes with a simple expected utility model with exponential discounting. We find this model to be unappealing for two reasons. First, to accommodate the data, the model requires extremely small discount factors, e.g., lower than 0.1. Second, the exponential model predicts the largest elasticities for the incarcerations at the shortest horizons—an increase in incarceration length from 1 to 3 days is predicted to have much larger proportional effects than an increase in incarceration length from 1 to 3 years.

Thus, we favor a less restrictive form of time preferences: that of hyperbolic discounting. A model of myopia with reasonable long-run discount rates predicts an inverted U-shape for elasticity sizes. Elasticities are small in the very short-term (e.g., days), larger in the medium-term (e.g., weeks/months), and small again in the very long-term (e.g., years). Such a model also seems intuitive, as criminal behavior—at least for the kinds of crimes that we focus on—could be thought of as the consequence of a self-control problem and a taste for immediate gratification.

A. Theory Appendix

In this Appendix, we describe in detail how we obtain Equations (4), (5), (6), and (7), which we use for providing our predicted magnitudes of θ. We first describe the case of exponential discounting, and then the case of "sophisticated" hyperbolic agents.

1. Exponential Discounting

The model is similar to a search model of unemployment, in which each period the individual obtains a draw from a distribution of wage offers. In our context, at the beginning of any period t, the individual draws from a distribution of criminal opportunities, indexed by p_t. Let $F^*(\cdot)$ be the cumulative distribution function of p_t. The individual then chooses whether or not to commit the crime. With probability p_t, he is arrested and is incarcerated for a fixed number of periods: if caught as a juvenile $t < t_0$, S_J periods, if caught as an adult $t \geq t_0$, he faces $S_A\,(> S_J)$ periods of incarceration. During this period of incarceration, he cannot commit crime, and upon release, he is free again to commit crime.

In each period, he receives flow utility u_{cr} if he commits the crime undetected, u_s if he is caught in the act or is serving an period of incarceration, and u_a if he abstains from crime. In each period, he maximizes the discounted expected utility

$$E_t\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}u_\tau\right]$$

where δ is the constant discount factor, E_t is an expectation given information at the beginning of period t, and u_τ is a random variable that equals either u_{cr}, u_s, or u_a, depending on the random draw of p_t, the agent's plan, and whether the agent is apprehended. The individual maximizes over all feasible plans, so his choice consists of a sequence of decisions to commit or abstain, conditional on draws of p_t.

We solve recursively, first beginning with the adult's ($t \geq t_0$) problem. Upon observing the draw p_t, the individual has two choices:

1. Commit the crime: obtain the utility $(1-p_t)\left(u_{cr}+\delta V\right)+p_t\left(u_s+\delta u_s+\cdots+\delta^{S_A-1}u_s+\delta^{S_A}V\right)$, or equivalently, $(1-p_t)\left(u_{cr}+\delta V\right)+p_t\left(u_s+\frac{\delta-\delta^{S_A}}{1-\delta}u_s+\delta^{S_A}V\right)$.
2. Abstain from crime: obtain the utility $u_a+\delta V$

where V is the continuation payoff, $E_t\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}u_\tau\right]$, evaluated at the optimal plan. The agent commits the crime if and only if $(1-p_t)(u_{cr}+\delta V)+p_t(u_s+\frac{\delta-\delta^{S_A}}{1-\delta}u_s+\delta^{S_A}V) > u_a+\delta V$. Equivalently, he will commit the crime if and only if $p_t < \overline{p}_t$, where

$$\overline{p}_t = \frac{u_{cr}-u_a}{u_{cr}-u_s+\frac{\delta-\delta^{S_A}}{1-\delta}\left(\left(1-\delta\right)V-u_s\right)}$$

which is the expression given in (5).

Since the problem is stationary for the adult, $\overline{p}_t=\overline{p}$. We may solve for V recursively as follows:

$$\begin{aligned}V &= \int_0^{\overline{p}}\left[\pi\left\{u_s+\frac{\delta-\delta^{S_A}}{1-\delta}u_s+\delta^{S_A}V\right\}+(1-\pi)\left\{u_{cr}+\delta V\right\}\right]f(\pi)d\pi+\int_{\overline{p}}^{1}\left[u_a+\delta V\right]f(\pi)d\pi \\ &= F^*(\overline{p})E_{\overline{p}}\left\{u_s+\tfrac{\delta-\delta^{S_A}}{1-\delta}u_s+\delta^{S_A}V\right\}+F^*(\overline{p})\left(1-E_{\overline{p}}\right)\left\{u_{cr}+\delta V\right\}+\left(1-F^*(\overline{p})\right)\left\{u_a+\delta V\right\}\end{aligned}$$

where $E_{\overline{p}} \equiv E[p_t | p_t \leq \overline{p}]$. Re-arranging, we obtain

$$(1-\delta) V = \frac{\frac{1-\delta^{S_A}}{1-\delta} F^* (\overline{p}) E_{\overline{p}} u_s + F^* (\overline{p}) (1 - E_{\overline{p}}) u_{cr} + (1 - F^* (\overline{p})) u_a}{1 + \frac{\delta - \delta^{S_A}}{1-\delta} F^* (\overline{p}) E_{\overline{p}}}$$

which is the expression given by (6). Note that $(1-\delta) V$ can be viewed as the "annuitized" value of the problem (whenever the individual is free to commit a crime), and is a weighted average of the three flow utilities, where the weights depend on $F^* (\overline{p})$, $E_{\overline{p}}$, δ, and S_A.

Now consider the individual at time $t = t_0 - 1$, the period immediately preceding his 18th birthday. His two choices are

1. Commit crime: obtain utility $(1 - p_t) (u_{cr} + \delta V) + p_t \left(u_s + \delta u_s + \cdots + \delta^{S_J - 1} u_s + \delta^{S_J} V\right)$, or equivalently, $(1 - p_t) (u_{cr} + \delta V) + p_t \left(u_s + \frac{\delta - \delta^{S_J}}{1-\delta} u_s + \delta^{S_J} V\right)$.

2. Abstain from crime: obtain utility $u_a + \delta V$

The only difference here is that S_A has been replaced by the shorter incarceration length S_J. V is exactly the same as before because at time $t_0 - 1$, the individual becomes an adult in the next period and faces the adult (stationary) dynamic problem. This leads to

$$\overline{p}_{t_0 - 1} = \frac{u_{cr} - u_a}{u_{cr} - u_s + \frac{\delta - \delta^{S_J}}{1-\delta} ((1-\delta) V - u_s)}$$

which is the expression in (4).

2. Hyperbolic Discounting

Now consider the problem with the quasi-hyperbolic discount sequence given by $1, \beta\delta, \beta\delta^2, \beta\delta^3 \ldots$ The solution has a similar form to that of the exponential case, with two differences: (1) the solution includes the short run discount factor β, and (2) V has a slightly different interpretation.

At each point in time t, the individual now maximizes

$$E_t \left[u_t + \beta \sum_{\tau = t+1}^{\infty} \delta^{\tau - t} u_\tau \right]$$

Since preferences are now time-inconsistent, we need to make an additional assumption about the agent's beliefs about his preferences in the future (O'Donoghue and Rabin 1999). We consider the case of sophistication, where the individual is aware of his present-biased preferences, and of the future choices he will make due to those preferences. Consider the adult's (stationary) problem first. His two choices are

1. Commit crime: obtain utility $(1 - p_t) (u_{cr} + \beta\delta V) + p_t \left(u_s + \beta\delta u_s + \cdots + \beta\delta^{S_A - 1} u_s + \beta\delta^{S_A} V\right)$, or equivalently, $(1 - p_t) (u_{cr} + \beta\delta V) + p_t \left(u_s + \beta \frac{\delta - \delta^{S_A}}{1-\delta} u_s + \beta\delta^{S_A} V\right)$.

2. Abstain from crime: obtain utility $u_a + \beta\delta V$

which is similar to before, except that V has a different definition.

Here, $V \equiv E_t \left[\sum_{\tau = t+1}^{\infty} \delta^{\tau - t - 1} u_\tau\right]$, evaluated at the optimal contingent plan. Note the change in where the index τ starts. With this definition, $\beta\delta V$ gives the sophisticated agent's evaluation—from the perspective of period t—of all periods starting at $t + 1$.

The expressions for the sophisticated agent's thresholds at time t_0 and $t_0 - 1$ are

$$\overline{p}^{sp} = \overline{p}_t^{sp} = \frac{u_{cr} - u_a}{u_{cr} - u_s + \beta \frac{\delta - \delta^{S_A}}{1-\delta} \left((1-\delta) V - u_s \right)}, \text{ for } t \geq t_0$$

$$\overline{p}_{t_0 - 1}^{sp} = \frac{u_{cr} - u_a}{u_{cr} - u_s + \beta \frac{\delta - \delta^{S_J}}{1-\delta} \left((1-\delta) V - u_s \right)}$$

which are given in (7). The superscript *sp* denotes the solution for a sophisticated hyperbolic agent.

Even though V here has a slightly different definition, the expression for it is identical to (6). This is because $V \equiv E_t \left[\sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1} u_\tau \right] = E_{t+1} \left[\sum_{\tau=t+1}^{\infty} \delta^{\tau-t-1} u_\tau \right] = E_t \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} u_\tau \right]$. Thus, we have

$$(1-\delta) V = \frac{\frac{1-\delta^{S_A}}{1-\delta} F^* \left(\overline{p}^{sp} \right) E_{\overline{p}^{sp}} u_s + F^* \left(\overline{p}^{sp} \right) \left(1 - E_{\overline{p}^{sp}} \right) u_{cr} + \left(1 - F^* \left(\overline{p}^{sp} \right) \right) u_a}{1 + \frac{\delta - \delta^{S_A}}{1-\delta} F^* \left(\overline{p}^{sp} \right) E_{\overline{p}^{sp}}}$$

Although V has the same form as in the exponential case, its actual value will be different, because it is evaluated at the steady-state threshold $\overline{p}^{sp}$ rather than at $\overline{p}$.

In general, holding all else equal—including keeping δ constant—a lower β leads to higher levels of crime, for two reasons. First, the discounted cost is lower, due to $\beta < 1$. Second, the expected opportunity cost of incarceration, given by $(1-\delta) V - u_s$, will be smaller for the hyperbolic agent. This is because the sophisticated hyperbolic correctly knows that he has a self-control problem and will commit "too much" crime in the future. More formally, we know that $\overline{p}$—the threshold for the exponential case—maximizes V. Therefore, any other threshold $\overline{p}^{sp}$ must make V smaller.

B. Details on Selected Calculations

In this Appendix, we describe how we obtain rough estimates of (1) the average length of incarceration conditional on an arrest for both juveniles (S_J) and adults (S_A), and (2) the expected probability of arrest $E_{\overline{p}} = E \left[p | p < \overline{p} \right]$. These quantities are used in our computation of predicted magnitudes of θ in Section II.

1. Average Incarceration Lengths: S_J and S_A

To estimate the average number of weeks that an arrestee can expect to spend incarcerated, we obtain the cumulative number of person-weeks that are spent incarcerated for a given year, and divide by the total number of arrests that occur within the year. As long as jail/prison populations and arrest numbers are reasonably stable, this should provide the average number of weeks spent incarcerated per arrest.

In the juvenile courts, a serious criminal offender will be placed in secure "detention" (the rough equivalent of jail), where they await an adjudication by the juvenile court judge. If they are found to be guilty, they are committed to a residential placement facility (the rough equivalent of a prison). In the table, column (2) (labeled "jail") includes juveniles in secure detention awaiting an adjudication as well as adults who are unconvicted, but awaiting court proceedings. Note that in the United States, jails not only incarcerate those awaiting hearings and trials, but also those who have been convicted to short prison terms. Therefore, the prison population includes adults in a state correctional facility as well as those incarcerated in jail who are serving a sentence. For juveniles, the prison population (column (3)) includes those in juvenile residential placement.

Our estimates of the jail and prison populations are complied from the 1999 Census of Jails, the 2000 Census of State and Federal Correctional Facilities, and the 1999 Census of Juveniles in

Residential Placement. Our estimates of the number of arrests come from the 1999 FBI Uniform Crime Reports.

The first two numbers in column (1) of Appendix Table I provide the number of arrests for juveniles (younger than 18 years old) and adults (18 and older). Column (2) provides the stock of juveniles and adults incarcerated in jail awaiting court proceedings. Provided that these population numbers are reasonable estimates of the average daily population throughout the year, this number multiplied by 52 gives the number of person-weeks spent incarcerated in "jail" throughout the year.

The next column takes the ratio of the first two columns to produce an average duration of incarceration conditional on an arrest. For juveniles, it is about 0.59 weeks, and for adults it is about 2.21 weeks. The ratio of these is given in the same column (3.77).

Column (3) provides prison populations for both juveniles and adults, and the subsequent column divides by the number of arrests to give the average length of incarceration in juvenile or adult prison—2.06 and 6.09—conditional on an arrest. It is important to note that this average will include many zeroes, for those who are not convicted/committed, or for whom formal charges are dropped. The final column adds the two averages. Overall, the expected length of incarceration conditional on an arrest is about 2.65 weeks for a juvenile, and 8.30 for adults.

Our analysis focuses on Index crimes. More minor crimes are likely to lead to very short periods of custody, and many offenders—particularly for misdemeanors—may be released almost immediately after a formal arrest. Thus, the numbers in the first set of rows are probably a lower bound on the incarceration lengths, conditional on an Index arrest.

To obtain an upper bound, the second set of rows re-computes the average durations using the number of Index Crime arrests as the denominator. This is an upper bound since surely some non-Index Arrests lead to a positive incarceration length. Overall, these show a much larger average duration for adults, with the estimated ratio S_A/S_J as large as 8.86 for the average duration in jail. Among the six ratios reported in the upper part of the table, it appears that a ratio of $S_A/S_J = 3$ is probably a conservative estimate of the relative punitiveness of adult criminal sanctions, in terms of incarceration lengths. For comparison, we perform the same computation for the United States as a whole. The bottom part of the table shows that the corresponding estimates of S_A/S_J are of a similar magnitude.

2. Expected Apprehension Rates (Conditional on Crime): $E\left[p|p<\overline{p}\right]$

$E\left[p|p<\overline{p}\right]$ is the expected probability of being apprehended, conditional on committing a crime. We provide a rough estimate of this quantity using so-called "clearance rates" from the FBI Uniform Crime Reports. A reported crime is "cleared by arrest", when an incident is followed up by law enforcement, and results in arresting an alleged offender.

Column (2) of Appendix Table II reports clearance rates for all FBI Index Crimes, and the various sub-categories. The overall rate is 0.20. There are two reasons why this almost certainly an upwardly biased estimate of $E[p|p<\overline{p}]$. First, not all criminal incidents are reported to the police, and therefore the denominator of Column (2) is probably too small. To address this, we obtained estimates of the rate of reporting victimizations to the police from the National Criminal Victimization Survey. By multiplying these rates (Column 3) by Column 2, we obtain an arguably more accurate estimate of the probability of arrest conditional on committing an Index crime. The average for all Index crimes is about 0.08, with the lowest for larceny (0.06), and the highest for assault (0.26) and murder (0.49).

The second reason why both Column (2) and the final column should be considered an upper bound for $E[p|p<\overline{p}]$ is that it may well be true that some of the arrestees are mistakenly arrested; that is, the numerator of both Column (2) and the final column may be too large.

For this reason, we believe that the estimates in final column are probably still too large. Nevertheless—since a larger estimate of $E[p|p < \overline{p}]$ will lead to larger absolute magnitudes of θ, we use the values 0.10 and 0.25 to generate conservative predictions for θ.

C. Data Appendix

Our data set is constructed using four electronic files maintained by the Florida Department of Law Enforcement (FDLE). The FDLE refers to these files as the arrest, date of birth, judicial, and identifier files. They constitute the key elements of Florida's Computerized Criminal History (CCH) system, which is maintained by the Criminal Justice Information System (CJIS) division of the FDLE. We obtained from the FDLE records on all felony arrests for the period 1989 to 2002.

1. Construction of Data Set

We construct our data set as follows. First, we begin with the arrests file, which contains a person identifier [34], the offense date, the arrest date, the charge code, and the arresting agency. Each record of the arrest file pertains to a separate offense. [35] The total number of records in the arrest file we received is 4,498,139. Because a single arrest event may result in multiple records (due to multiple offenses), we collapse the data down to the level of the (1) person identifier and (2) arrest date, coding the offense as the most serious offense with which the individual was charged on that date. [36] There are 3,314,851 unique arrest-person observations.[37] Our collapsing procedure is essentially the "Hierarchy Rule" instituted by the Federal Bureau of Investigation (FBI) in the Uniform Crime Reporting (UCR) system for crime counting purposes (Federal Bureau of Investigation 2004, p. 10).

We similarly collapse the judicial file down to the person-arrest-date level. The judicial file represents all arrests that result in a formal prosecution. For each collapsed observation, if any of the potentially multiple arrests led to a conviction and prison or jail sentence, a prison or jail sentence was associated with the person-arrest-date. The collapsed judicial and arrest files were then merged on the unique person-identifier-arrest-date pair. Then, using the person identifier, which is present in all four files, race from the identifier file was merged on, and birthday was merged on from the birth date file.

2. Date Variables

The key variables we utilize from the arrest file are the person identifier, the arrest date, and the offense code. The offense date is missing for many observations, so we use the arrest date to proxy for the date of crime commission. This is due primarily to a reporting problem—officers do not always submit information on the offense date. On the other hand, there are no missing values for the arrest date. Among the 1,948,096 records with information on offense date, every one of those records has an associated arrest date. 90% of those arrest dates are equal to the offense date,

[34]This person identifier is constant across the various FDLE files.

[35]Roughly speaking, a record of the arrest file corresponds to the triple of (1) person identifier, (2) arrest date, and (3) charge code. Conceptually, the named triple will not uniquely identify a record due to the possibility of multiple arrest events for the same crime on the same day. Practically, there also appear to be some minor errors with double-counting in the file (e.g., two such triples, one with a missing offense date and another with an offense date equal, as usual, to the arrest date). However, the number of unique triples in the data is 94.3% of the total record count. We conclude that neither the conceptual distinction nor the double-counting issue is important empirically.

[36]We defined the seriousness of an offense according to...

[37]Thus, the average arrest event is associated with 1.36 charges.

and over 93% of those arrest dates occur within the first week subsequent to the stated offense date.

To further assess the validity of date of arrest as a proxy for date of offense, we obtained data from the Miami Police Department, which recorded arrest and offense dates for all charges pertaining to arrests made between July 1999 and December 2002. For the 272,494 arrests we obtained, 257,263 have a valid offense date, and 91.3% of those have offense and arrest dates that are identical, and with 95.8% of arrests occurring within the first week after the offense date. Focusing only on felony arrests, we find that of the 33,698 felony arrests, 32,033 have valid offense dates, and of these 78.9% have identical arrest and offense dates, and 90.6% have associated arrest dates that fall within a week of the offense date.

References

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, March/April 1968, *76* (2), 169–217.

Bozynski, Melanie and Linda Szymanski, "National Overviews," *State Juvenile Justice Profiles*, 2003, Pittsburgh: National Center for Juvenile Justice,.

Butterworth, Bob, *Annual Report of the Attorney General of the State of Florida*, Tallahassee: Attorney General's Office, 1995.

Case, Anne C. and Lawrence F. Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," *NBER Working Paper #3705*, 1991.

Chaddock, Gail Russell, "U.S. Notches World's Highest Incarceration Rate," *Christian Science Monitor*, August 18, 2003.

Di Tella, Rafael and Ernesto Schargrodsky, "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack," *American Economic Review*, March 2004, *94* (1), 115–133.

Donohue, John J., III and Steven J. Levitt, "The Impact of Legalized Abortion on Crime," *Quarterly Journal of Economics*, May 2001, *116* (2), 379–420.

Efron, Bradley, "Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve," *Journal of the American Statistical Association*, June 1988, *83* (402), 414–425.

Ehrlich, Isaac, "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, May/June 1973, *81* (3), 521–565.

Federal Bureau of Investigation, *Uniform Crime Reporting Handbook, 2004*, Washington, D.C.: Federal Bureau of Investigation, 2004.

Forst, Martin, Jeffrey Fagan, and T. Scott Vivona, "Youths in Prisons and Training Schools: Perceptions and Consequences of the Treatment-Custody Dichotomy," *Juvenile and Family Court Journal*, 1989, *40* (1), 1–14.

Freeman, Richard B., "The Labor Market," in James Q. Wilson and Joan Petersilia, eds., *Crime*, Oakland: Institute for Contemporary Studies, 1995, pp. 171–191.

——, "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?," *Journal of Economic Perspectives*, Winter 1996, *10* (1), 24–42.

——, "The Economics of Crime," in Orley Ashenfelter and David E. Card, eds., *Handbook of Labor Economics, Volume 3C*, Vol. 3C, New York: Elsevier-North Holland, 1999, pp. 3529–3571.

Gaviria, Alejandro and Steven Raphael, "School-Based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics*, May 2001, *83* (2), 257–268.

Glaeser, Edward L. and Bruce Sacerdote, "Why is There More Crime in Cities?," *Journal of Political Economy*, December 1999, *107* (2, Part 2), S225–S258.

—— and ——, "The Determinants of Punishment: Deterrence, Incapacitation and Vengeance," *NBER Working Paper #7676*, April 2000.

——— , ——— , and Jose A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics*, May 1996, *111* (2), 507–548.

Glassner, Barry, Margret Ksander, Bruce Berg, and Bruce Johnson, "A Note on the Deterrent Effect of Juvenile vs. Adult Jurisdiction," *Social Problems*, December 1983, *31* (2), 219–221.

Grogger, Jeff, "The Effect of Arrests on the Employment and Earnings of Young Men," *Quarterly Journal of Economics*, February 1995, *110* (1), 51–72.

——— , "Market Wages and Youth Crime," *Journal of Labor Economics*, October 1998, *16* (4), 756–791.

Jacob, Brian A. and Lars Lefgren, "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime," *American Economic Review*, December 2003, *93* (5), 1560–1577.

Katz, Lawence F., Steven D. Levitt, and Ellen Shustorovich, "Prison Conditions, Capital Punishment and Deterrence," *American Law and Economics Review*, Fall 2003, *5* (2), 318–343.

Kessler, Daniel and Steven D. Levitt, "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation," *Journal of Law and Economics*, April 1999, *42* (1, Part 2), 343–363.

Kling, Jeffrey R., "Incarceration Length, Employment, and Earnings," *IR Section Working Paper # 494*, Princeton University, 2004.

——— , Jens Ludwig, and Lawrence F. Katz, "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics*, February 2005, *120* (1), 87–130.

Levitt, Steven D., "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," *Quarterly Journal of Economics*, May 1996, *111* (2), 319–351.

——— , "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997, *87* (3), 270–290.

——— , "Juvenile Crime and Punishment," *Journal of Political Economy*, December 1998, *106* (6), 1156–1185.

——— , "Deterrence," in James Q. Wilson and Joan Petersilia, eds., *Crime: Public Policies for Crime Control*, Oakland: Institute for Contemporary Studies, 2004, pp. 435–450.

——— , "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not," *Journal of Economic Perspectives*, Winter 2004, *18* (1), 163–190.

——— and Lance Lochner, "The Determinants of Juvenile Crime," in Jonathan Gruber, ed., *Risky Behavior Among Youths: An Economic Analysis*, Chicago: University of Chicago Press, 2001, pp. 327–374.

Lochner, Lance, "Education, Work, and Crime," *NBER Working Paper #10478*, May 2004.

——— and Enrico Moretti, "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review*, March 2004, *94* (1), 155–189.

Maguire, Kathleen and Ann L. Pastore, *Sourcebook of Criminal Justice Statistics*, Washington, D.C.: U.S. GPO, 2000.

Myers, David L., "Punishing Violent Youth in the Adult Criminal Justice System," *Youth Violence and Juvenile Justice*, April 2003, *1* (2), 173–197.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, March 1999, *89* (1), 103–124.

Pintoff, Randi, "The Impact of Incarceration on Juvenile Crime: A Regression Discontinuity Approach," *Mimeo, Yale University*, 2004.

Raphael, Steven and Rudolf Winter-Ebmer, "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics*, April 2001, *44* (1), 259–284.

Redding, Richard E., "The Effects of Adjudicating and Sentencing Juveniles as Adults: Research and Policy Implications," *Youth Violence and Juvenile Justice*, April 2003, *1* (2), 128–155.

Robertson, James E., "Cruel and Unusual Punishment in United States Prisons: Sexual Harassment Among Male Inmates," *American Criminal Law Review*, Winter 1999, *36* (1), 1–51.

Saccenti, Brian, "Preventing Summary Judgement Against Inmates Who Have Been Sexually Assaulted By Showing That the Risk Was Obvious," *Maryland Law Review*, 2000, *59*, 642–668.

Snyder, Howard N. and Melissa Sickmund, *Juvenile Offenders and Victims: 1999 National Report*, Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention, 1999.

Western, Bruce, Jeffrey R. Kling, and David F. Weiman, "The Labor Market Consequences of Incarceration," *IR Section Working Paper*, Princeton University, 2001, *450.*

Wilson, James Q. and Richard J. Herrnstein, *Crime and Human Nature*, New York: Simon and Schuster, 1985.

Figure I. Criminal Propensity by Age: Theoretical Predictions

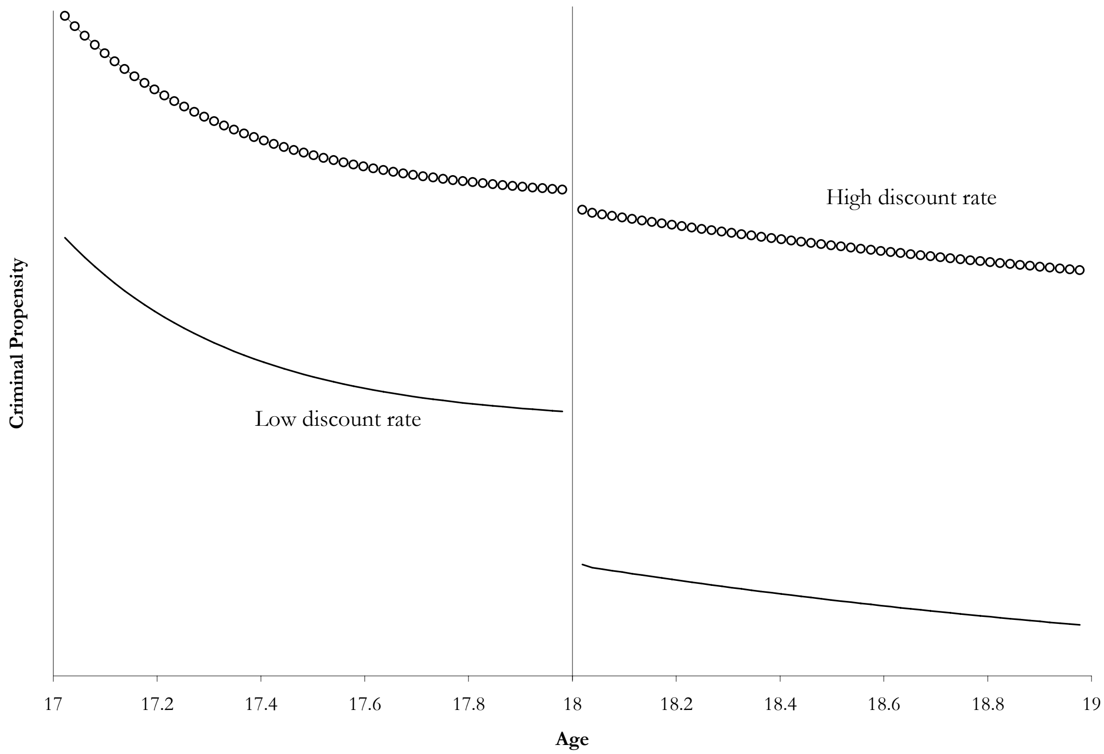
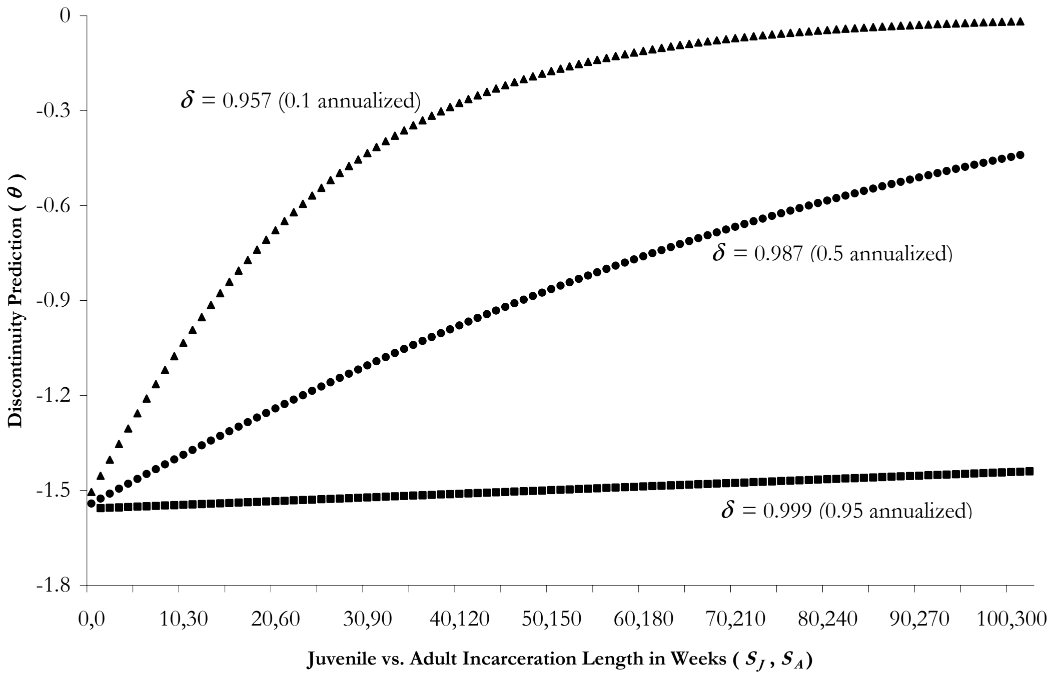


Figure II. Predicted Response to Tripling of Incarceration Length

A. Reservation Probability ($\bar{p}_t$) of 0.2



B. Reservation Probability ($\bar{p}_t$) of 0.5

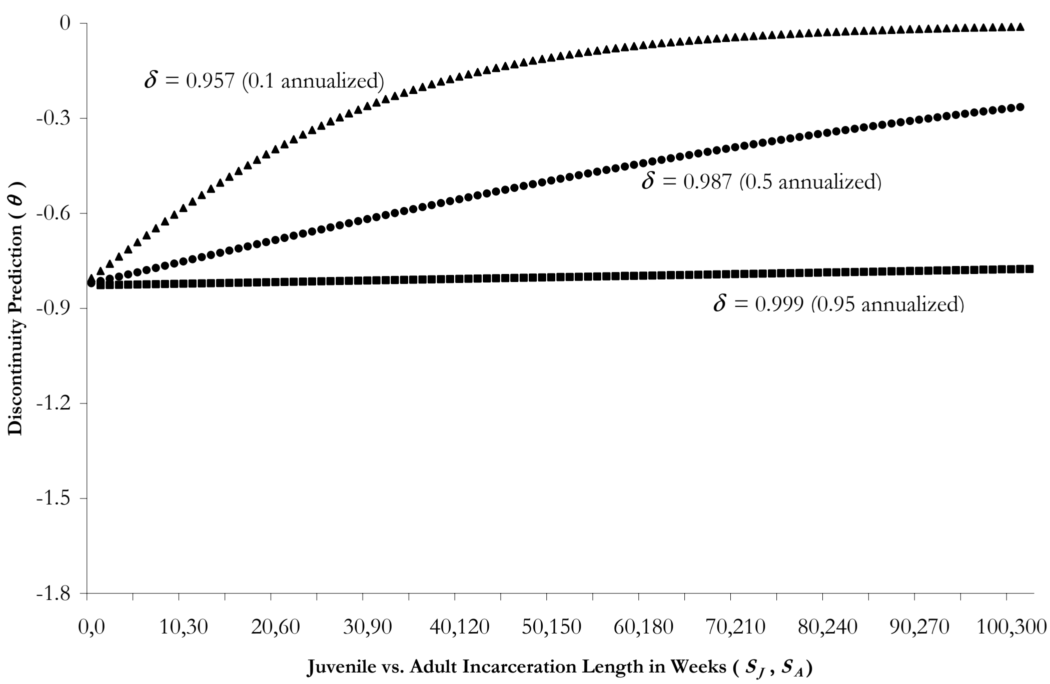
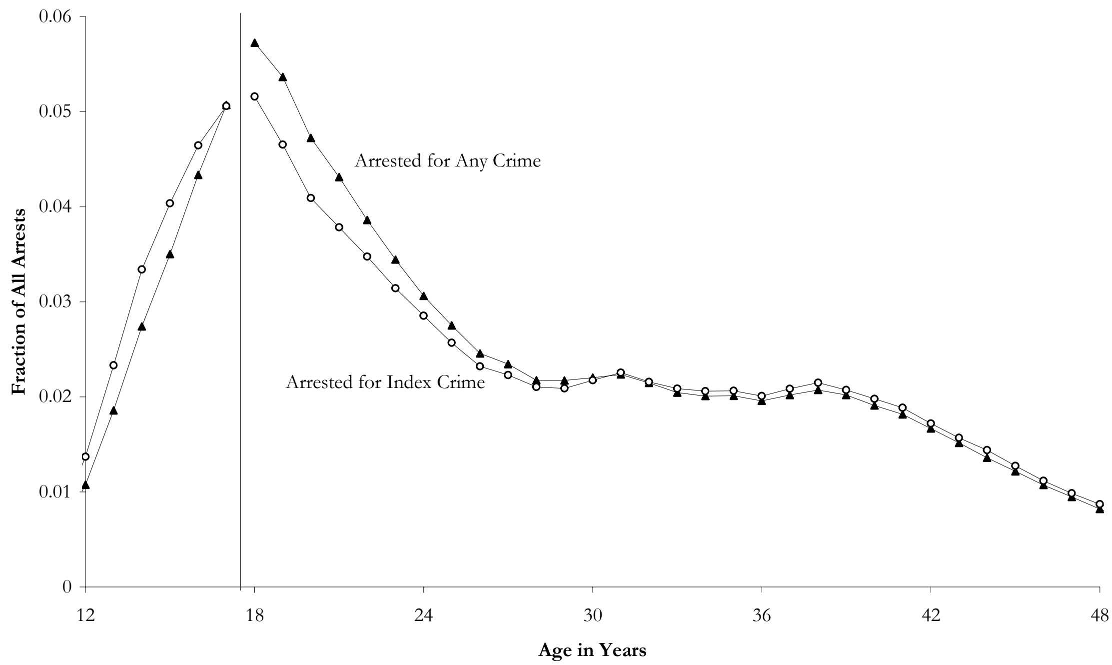
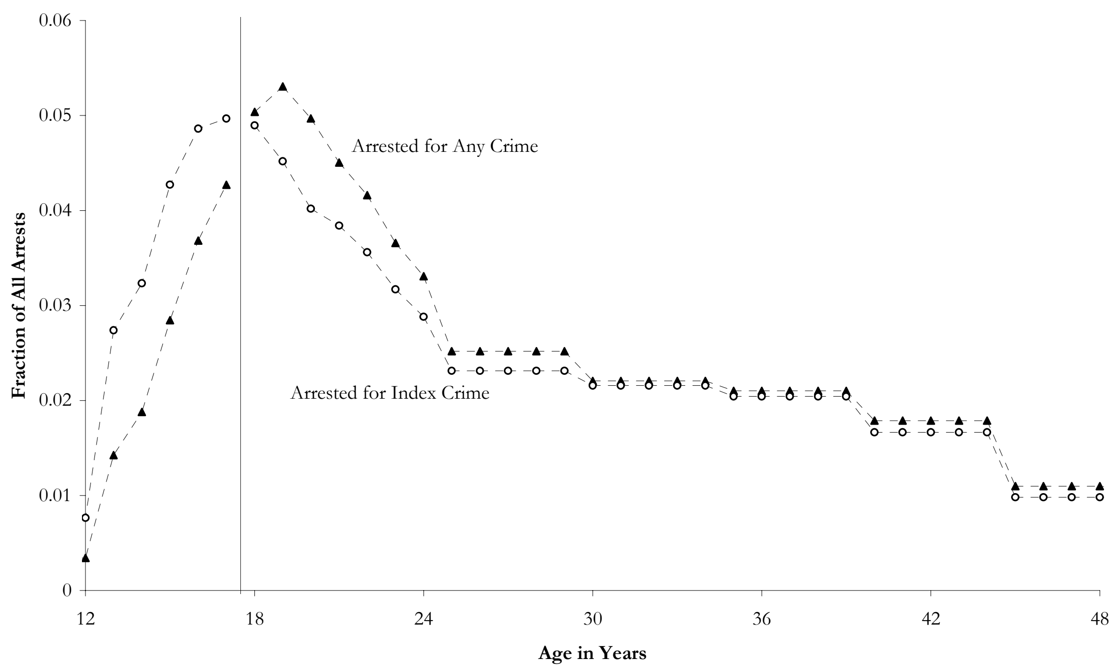


Figure III. Distribution of Age at Arrest: 2002 Public-Use Data

A. National Incident Based Reporting System Data

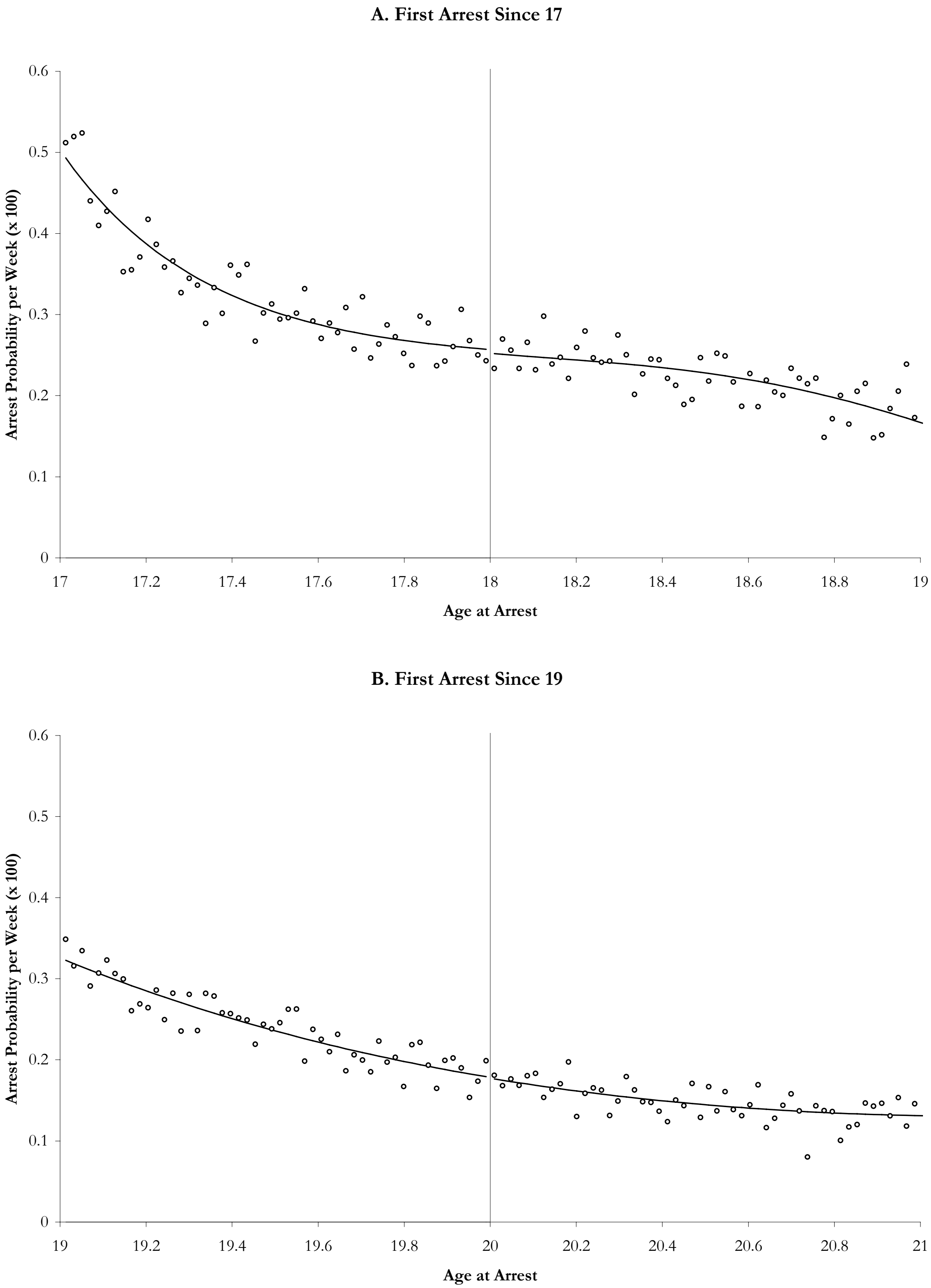


B. Uniform Crime Reports Data



Note: In top panel, "Arrested for Any Crime" = arrested for "Group A" NIBRS offense. In UCR data, some age ranges are aggregated by the FBI.

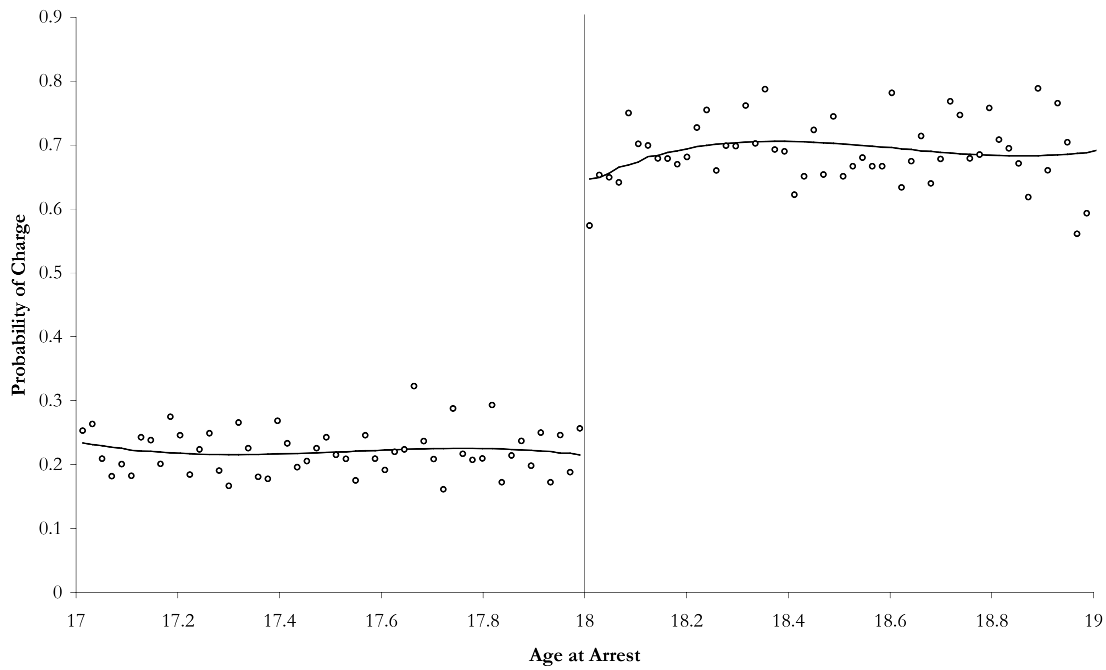
Figure IV. Criminal Propensity Estimates by Age



Note: Top (bottom) panel shows probability of first arrest since 17 (19), by age in weeks, for those arrested at least once prior to 17 (19).

Figure V. Transfer and Punitiveness by Age

A. Probability of Being Charged as an Adult



B. Probability of Sentence to Adult Confinement

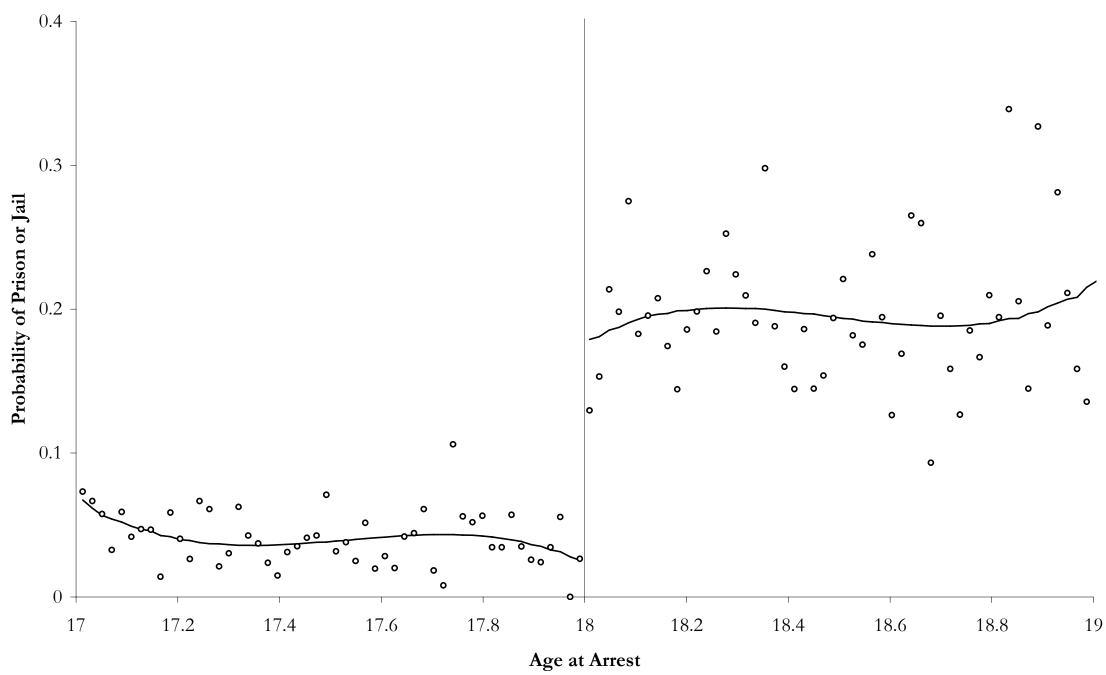
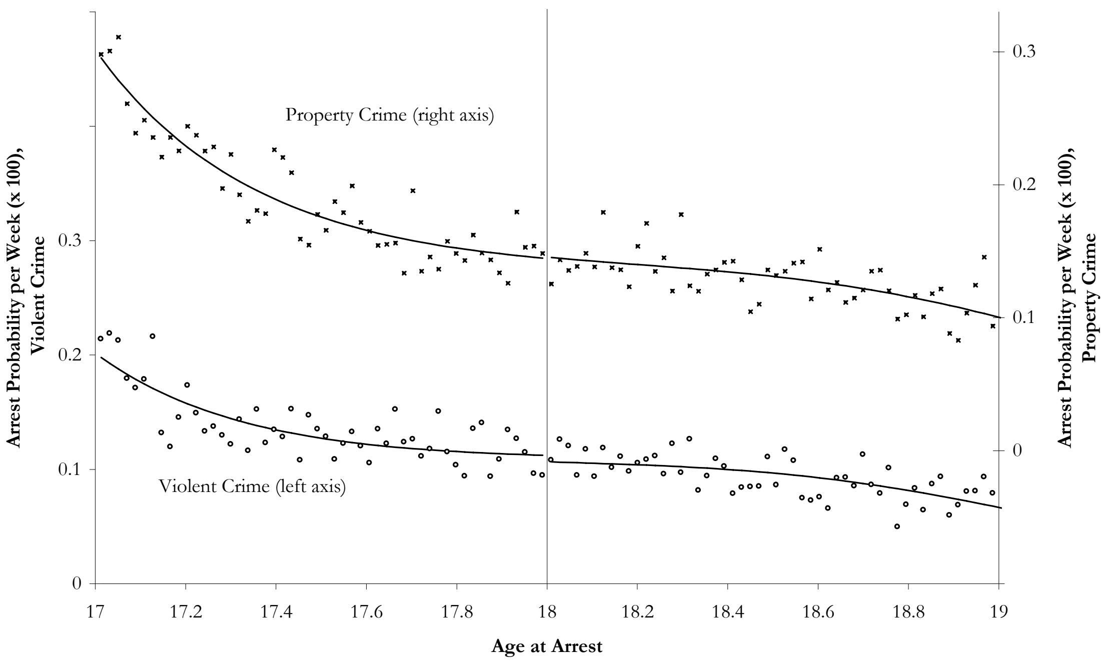
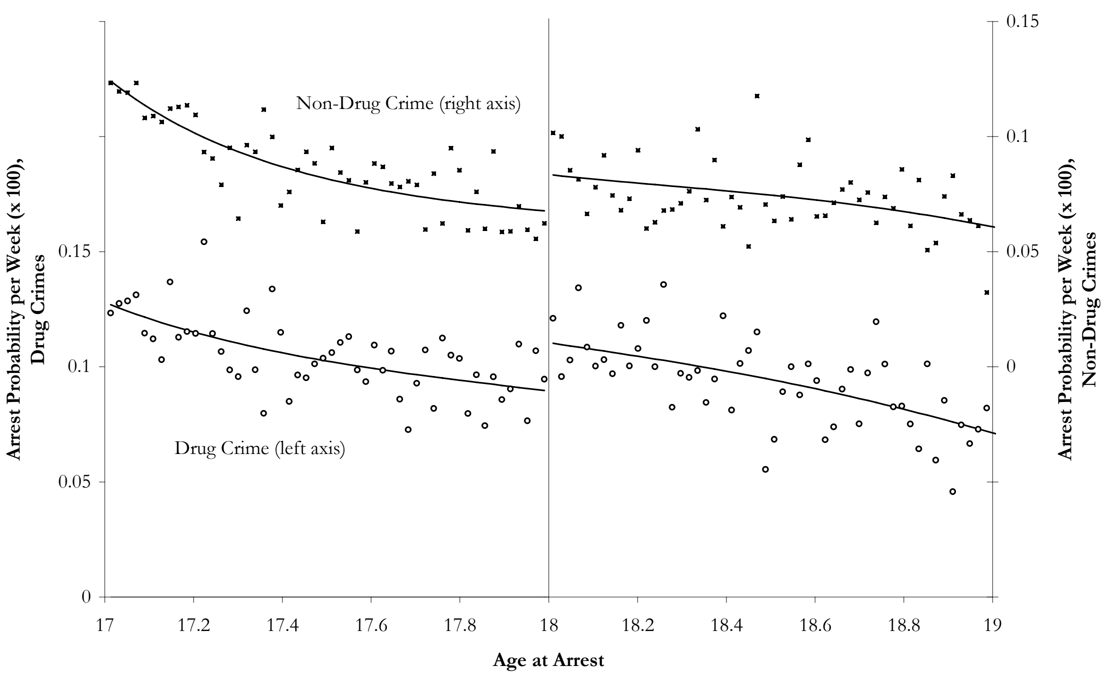


Figure VI. Criminal Propensity by Type of Offense

A. Index Crimes

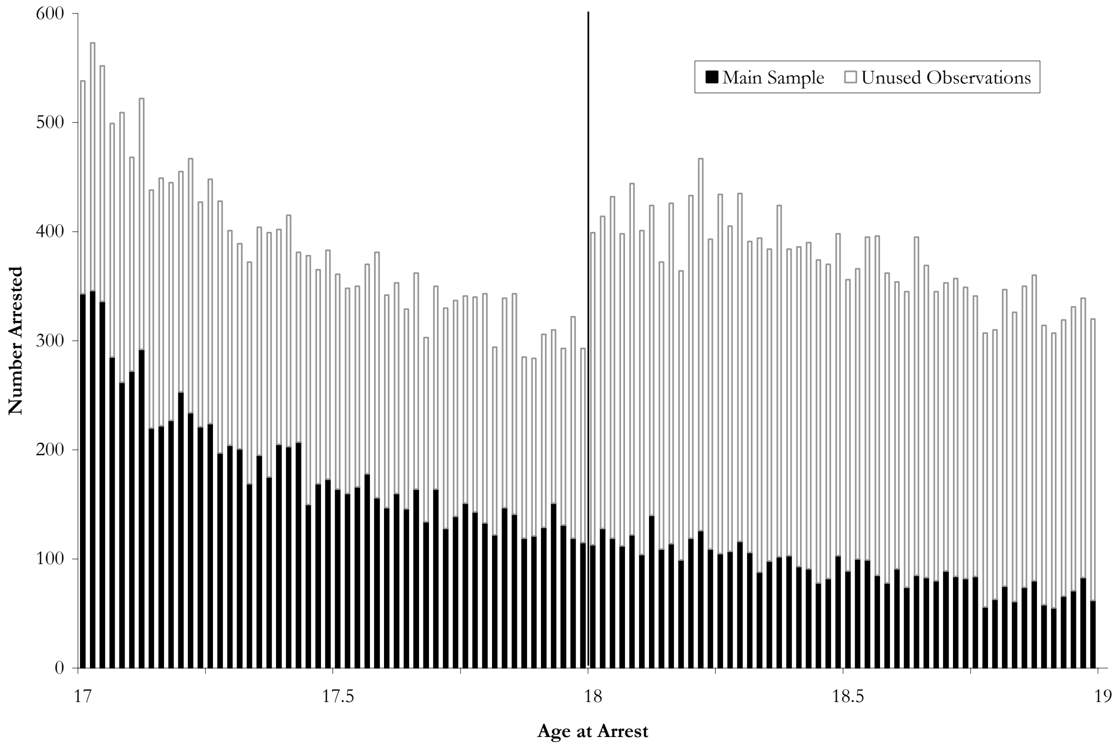


B. Non-index Crimes



Note: Top panel disaggregates estimates from top panel of Figure IV. Bottom panel presents estimates for non-index crimes.

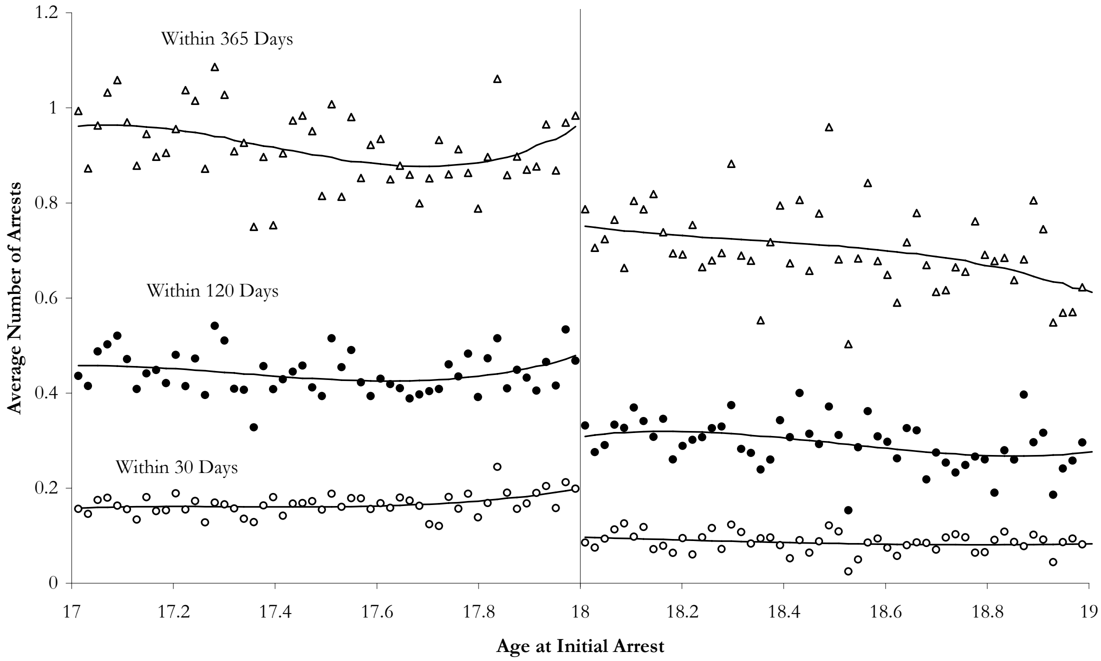
Figure VII. Number of Arrests by Age: Main Sample versus Unused Observations



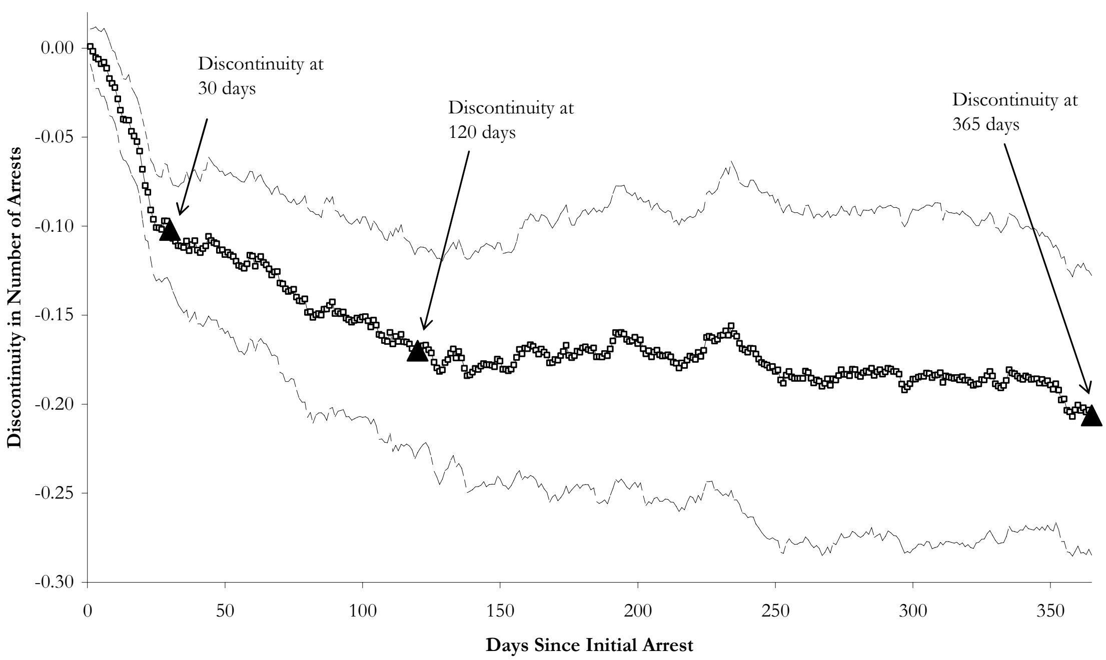
Note: Figure is a stacked histogram (shaded and light rectangles sum to total).

Figure VIII. Incapacitation Effects

A. Number of Follow-up Arrests, by Age at Initial Arrest

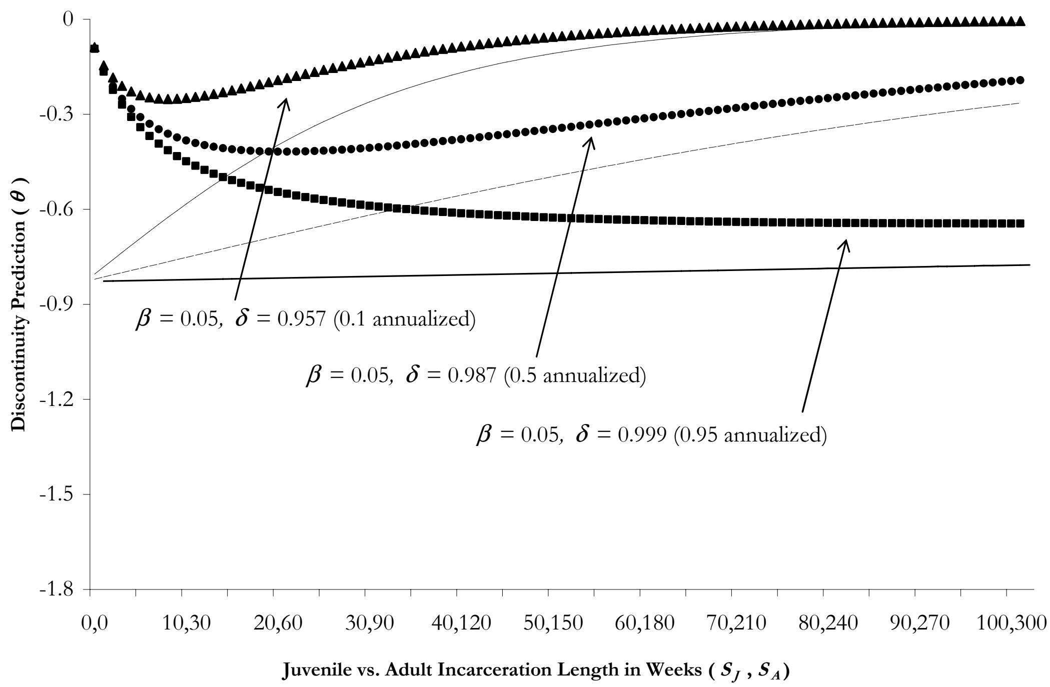


B. Discontinuity Estimates, by Length of Follow-up Window



Note: Initial arrest = first arrest since 17. Top panel gives average number of arrests *after* initial arrest, by age in weeks, for three different follow-up lengths. Bottom panel plots estimated discontinuity in number of follow-up arrests, for follow-up lengths ranging from 1 to 365 days, with twice standard error bands.

Figure IX. Predicted Response to Tripling of Incarceration Length, Hyperbolic Discounting



Note: Each scenario assumes $\overline{p}_t = 0.5$. Predictions under exponential discounting from Figure II shown as thin solid, dotted, and thick solid lines.

Figure X. Hypothetical Density of Criminal Opportunities: Marginal versus Inframarginal Offenders

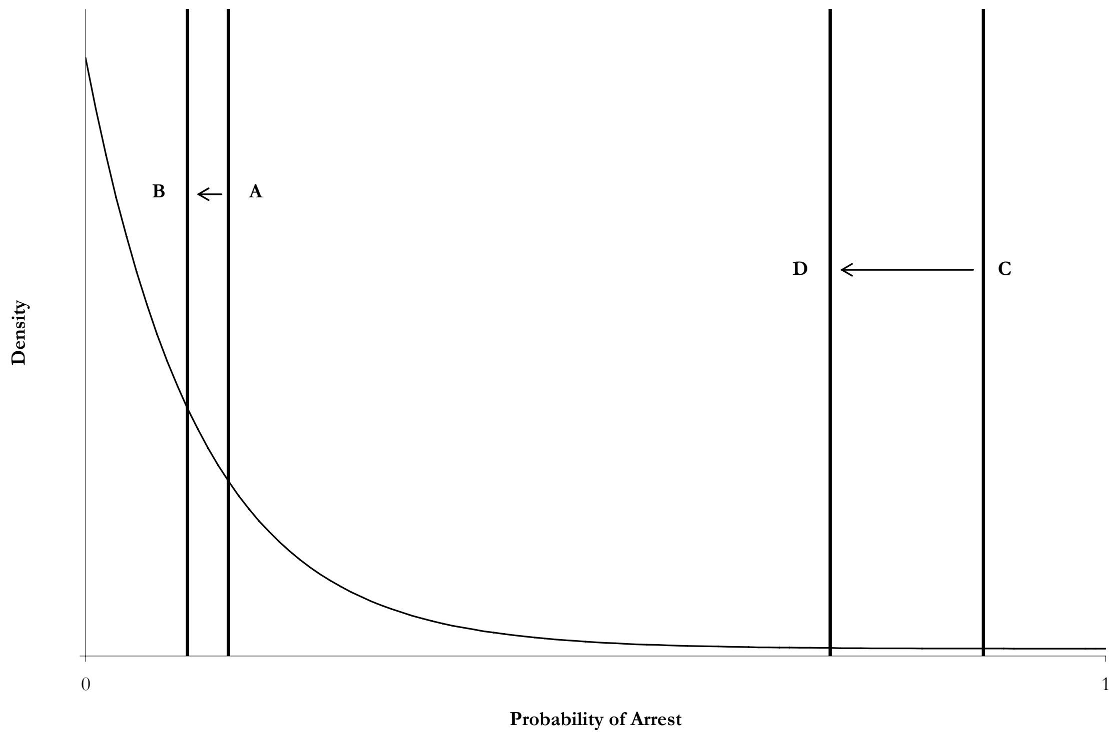


Table I. Summary Statistics, Estimation Sample

	Estimation Sample				Young Arrestees
Variable	First Arrest	Second Arrest	Third Arrest	First Arrest Since 17	All Arrests
	(1)	(2)	(3)	(4)	(5)
Non-white	0.42	0.49	0.55	0.52	0.48
Male	0.80	0.86	0.89	0.88	0.88
Age	14.93	16.64	17.47	18.39	18.14
	(1.39)	(1.97)	(2.06)	(1.24)	(0.56)
Arrested in Small County	0.28	0.30	0.31	0.31	0.31
Arrested in Medium County	0.36	0.36	0.36	0.35	0.35
Arrested in Large County	0.36	0.34	0.33	0.34	0.34
Index Crime, Violent	0.26	0.25	0.25	0.25	0.24
Index Crime, Property	0.51	0.42	0.37	0.33	0.39
Non-index Crime, Drug	0.07	0.15	0.18	0.22	0.20
Non-index Crime, Non-drug	0.15	0.18	0.20	0.19	0.17
Number of Persons	64,073	40,520	28,703	30,938	163,037

Note: Standard deviations in parentheses. Young arrestees are those arrested between 17 and 19. Number of arrests observed for young arrestees is 247,037.

Table II. Discontinuity Estimates of Deterrence

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Discontinuity	-0.0180	-0.0172	-0.0170	-0.0176	-0.0174	-0.0174	-0.0172	-0.0171
	(0.0474)	(0.0474)	(0.0474)	(0.0474)	(0.0474)	(0.0474)	(0.0474)	(0.0474)
Non-white		0.5448	0.5028	0.5758	0.5334	0.5425	0.5405	0.5405
		(0.0169)	(0.0173)	(0.0172)	(0.0176)	(0.0176)	(0.0176)	(0.0176)
Size of County of Baseline Arrest (Relative to Large County)								
Small			0.2501		0.2553	0.2468	0.2498	0.2503
			(0.0214)		(0.0215)	(0.0215)	(0.0215)	(0.0215)
Medium			0.0813		0.0875	0.0869	0.0860	0.0863
			(0.0209)		(0.0209)	(0.0209)	(0.0209)	(0.0209)
Type of Crime, Baseline Arrest (Relative to Non-index Non-Drug)								
Index Crime, Violent				0.0195	0.0087	0.0159	0.0159	0.0162
				(0.0283)	(0.0283)	(0.0283)	(0.0283)	(0.0283)
Index Crime, Property				0.1914	0.1896	0.2017	0.2048	0.2046
				(0.0255)	(0.0255)	(0.0256)	(0.0256)	(0.0256)
Non-index Drug				-0.1525	-0.1541	-0.2072	-0.2134	-0.2132
				(0.0419)	(0.0419)	(0.0421)	(0.0421)	(0.0421)
Controls for Age at Baseline Arrest?	N	N	N	N	N	Y	Y	Y
Order of Polynomial in Age at Baseline Arrest						1	3	5
Log-likelihood	-96,005.7	-95,491.1	-95,420.3	-95,411.8	-95,338.8	-95,236.6	-95,178.3	-95,174.8

Note: Standard errors in parentheses. Table presents coefficients from a logit model for being arrested for an index crime since 17. In addition to controls described, each model controls for a cubic polynomial in age at current arrest, relative to 18. Estimates are based on a panel of 4,928,226 observations pertaining to 64,703 persons.

Table III. Discontinuity in Probability of Being Charged in Adult Court

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated Discontinuity	0.4042	0.4035	0.4041	0.4037	0.4043	0.4051	0.4052	0.4054
	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)	(0.0235)
Log Discontinuity	*1.1056*	*1.1044*	*1.1054*	*1.1048*	*1.1058*	*1.1070*	*1.1073*	*1.1075*
Non-white		-0.0274	-0.0294	-0.0300	-0.0320	-0.0294	-0.0296	-0.0295
		(0.0057)	(0.0058)	(0.0058)	(0.0059)	(0.0059)	(0.0059)	(0.0059)
Size of County of Baseline Arrest (Relative to Large County)								
Small			0.0141		0.0140	0.0129	0.0128	0.0129
			(0.0072)		(0.0072)	(0.0071)	(0.0071)	(0.0072)
Medium			0.0199		0.0198	0.0195	0.0194	0.0195
			(0.0069)		(0.0069)	(0.0069)	(0.0069)	(0.0069)
Type of Crime, Baseline Arrest (Relative to Non-index Non-Drug)								
Index Crime, Violent				0.0055	0.0052	0.0060	0.0060	0.0060
				(0.0095)	(0.0095)	(0.0095)	(0.0095)	(0.0095)
Index Crime, Property				-0.0031	-0.0028	-0.0015	-0.0017	-0.0018
				(0.0085)	(0.0085)	(0.0085)	(0.0085)	(0.0085)
Non-index Drug				0.0228	0.0229	0.0131	0.0127	0.0127
				(0.0120)	(0.0120)	(0.0121)	(0.0121)	(0.0121)
Controls for Age at Baseline Arrest?	N	N	N	N	N	Y	Y	Y
Order of Polynomial in Age at Baseline Arrest						1	3	5
R^2	0.2144	0.2152	0.2155	0.2154	0.2157	0.2173	0.2174	0.2174

Note: Standard errors in parentheses. Log discontinuity is difference in log probabilities and is calculated from the presented difference estimate using a baseline rate of 0.2 for the marginal juvenile. Delta method standard errors for the log discontinuity are in each instance approximately 0.039.

Table IV. Heterogeneity in Discontinuity Estimates

Subsample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Arrestees	-0.0180	-0.0171						
	(0.0474)	(0.0474)						
White Arrestees			0.0240	0.0240				
			(0.0670)	(0.0670)				
Non-white Arrestees			-0.0592	-0.0590				
			(0.0670)	(0.0670)				
Baseline Crime was in Large County					0.0292 (0.0821)	0.0300 (0.0821)		
Baseline Crime was in Medium County					-0.0214 (0.0807)	-0.0208 (0.0807)		
Baseline Crime was in Small County					-0.0622 (0.0834)	-0.0619 (0.0834)		
Baseline Crime was Index, Violent							0.0553 (0.0930)	0.0556 (0.0930)
Baseline Crime was Index, Property							-0.0696 (0.0646)	-0.0683 (0.0646)
Baseline Crime was Non-index, Drug							0.0919 (0.1960)	0.0952 (0.1959)
Baseline Crime was Non-index, Non-drug							-0.0032 (0.1250)	-0.0025 (0.1250)
Log-likelihood	-96,005.7	-95,174.8	-95,490.1	-95,173.8	-95,835.6	-95,170.6	-95,954.9	-95,165.0
Controls for Race?	N	Y	Y	Y	N	Y	N	Y
Controls for County Size?	N	Y	N	Y	Y	Y	N	Y
Controls for Baseline Crime Type?	N	Y	N	Y	N	Y	Y	Y
Controls for Age at Baseline Arrest?	N	Y	N	Y	N	Y	N	Y
Order of Polynomial in Age at Baseline Arrest		5		5		5		5
Test for Equality of Discontinuity Estimates			0.771	0.770	0.612	0.620	1.590	1.582
degrees of freedom			[1]	[1]	[2]	[2]	[3]	[3]
p-value			0.380	0.380	0.736	0.733	0.662	0.663

Note: Standard errors in parentheses. Table presents discontinuity estiamtes for different groups estimated from logit models. Odd-numbered columns include only those controls appropriate to testing the treatment interaction of interest. For example, column (3) includes controls for (i) white, (ii) a cubic polynomial in age, as in Table II, (iii) the interaction of the same cubic polynomial with the indicator for white, (iv) the interaction of an indicator for being above 18 with the indicator for white (estimate for whites shown), and (v) the interaction of the indicators for being above 18 and non-white (estimate for non-whites shown). Even-numbered columns additionally include the richest set of controls considered in Table II. The final row of the table tests for the equality of the interacted treatment effects.

Appendix Table I. Juvenile and Adult Incarceration Length in Weeks

	Arrests	Jail	Average Duration	Prison	Average Duration	Total	Average Duration
A. Florida	**(1)**	**(2)**	**52*(2)/(1)**	**(3)**	**52*(3)/(1)**	**(2)+(3)**	**52*((2)+(3))/(1)**
All Offenses as Denominator							
Juvenile	131,330	1,482	0.59	5,211	2.06	6,693	2.65
Adult	766,259	32,585	2.21	89,730	6.09	122,315	8.30
Adult/Juvenile Ratio			**3.77**		**2.95**		**3.13**
Index Offenses as Denominator							
Juvenile	53,967	1,482	1.43	5,211	5.02	6,693	6.45
Adult	133,853	32,585	12.66	89,730	34.86	122,315	47.52
Adult/Juvenile Ratio			**8.86**		**6.94**		**7.37**
B. United States							
All Offenses as Denominator							
Juvenile	1,588,839	26,439	0.87	76,926	2.52	103,365	3.38
Adult	7,552,362	354,379	2.44	1,546,456	10.65	1,900,835	13.09
Adult/Juvenile Ratio			**2.82**		**4.23**		**3.87**
Index Offenses as Denominator							
Juvenile	420,543	26,439	3.27	76,926	9.51	103,365	12.78
Adult	1,091,530	354,379	16.88	1,546,456	73.67	1,900,835	90.55
Adult/Juvenile Ratio			**5.16**		**7.75**		**7.09**

Note: Table gives estimates of average incarceration length based on stock-flow comparisons. Adult and juvenile arrest counts are for the year 1999, from the FBI's Uniform Crime Reports, as reported in the Sourcebook of Criminal Justice Statistics. Adult jail population counts come from the Census of Jails and pertain to December 1999. Adult prison population counts come from the Census of Correctional facilities and pertain to June 2000. Juvenile jail and prison population counts come from the Census of Juveniles in Residential Placement and pertain to October 1999.

Appendix Table II. Arrest Probabilities

Crime Category	Offenses Known to Police	Fraction of Offenses Cleared by Arrest	Fraction of Victimizations Reported to Police	Upper Bound for $E[p_t \mid p_t < \overline{p_t}]$
	(1)	(2)	(3)	(2)*(3)
All Index Crimes	10,121,721	0.20	0.42	0.08
Violent Crime	1,184,453	0.47	0.48	0.23
Murder	13,561	0.64	0.77 *	0.49
Forcible Rape	80,515	0.45	0.54	0.24
Robbery	343,023	0.26	0.71	0.18
Aggravated Assault	747,354	0.57	0.46	0.26
Property Crime	8,937,268	0.17	0.40	0.07
Burglary	1,842,930	0.13	0.58	0.08
Larceny-Theft	6,014,290	0.18	0.33	0.06
Motor Vehicle Theft	1,080,048	0.14	0.86	0.12
Source:	UCR	UCR	NCVS, NCHS	Authors' calculations

Note: Figures pertain to 2002 and are taken from the 30th Online Edition of the Sourcebook of Criminal Justice Statistics. Figures labelled "UCR" are from the FBI's Uniform Crime Reports system (Table 4.19). Figures labelled "NCVS" are from the Census Bureau's National Crime Victimization Survey (Table 3.36). Asterisk indicates that the fraction of victimizations reported to police was estimated by taking the ratio of offenses known to police to the number of 2002 murders reported to the National Center for Health Statistics (Table E, National Vital Statistics Reports, Vol. 53, No. 17, 2005).

Appendix Table III. Robustness of Main Results

	Imposing Equality of Derivatives at 18					Allowing for Different Derivatives at 18				
A. Unconditional Estimates	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Estimated Discontinuity	0.174	0.135	-0.018	0.002	0.029	0.147	-0.040	0.014	-0.054	-0.065
	(0.035)	(0.036)	(0.047)	(0.048)	(0.057)	(0.036)	(0.053)	(0.072)	(0.090)	(0.109)
Order of Polynomial	1	2	3	4	5	1	2	3	4	5
Test Against Saturated Model	175.8	144.2	119.8	113.8	113.0	153.5	119.5	112.6	110.3	109.3
degrees of freedom	[102]	[101]	[100]	[99]	[98]	[101]	[99]	[97]	[95]	[93]
p-value	0.0000	0.0031	0.0865	0.1461	0.1421	0.0006	0.0789	0.1325	0.1351	0.1194
B. Controlling for Race, Size of County, and Type of Baseline Crime										
Estimated Discontinuity	0.173	0.134	-0.017	0.002	0.029	0.147	-0.039	0.014	-0.054	-0.065
	(0.035)	(0.036)	(0.047)	(0.048)	(0.057)	(0.036)	(0.053)	(0.072)	(0.090)	(0.109)
Black	0.534	0.533	0.533	0.533	0.533	0.534	0.533	0.533	0.533	0.533
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Small County	0.256	0.255	0.255	0.255	0.255	0.255	0.255	0.255	0.255	0.255
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Medium County	0.088	0.088	0.087	0.087	0.087	0.088	0.087	0.087	0.087	0.087
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Baseline Crime Violent	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Baseline Crime Property	0.190	0.190	0.190	0.190	0.190	0.190	0.190	0.190	0.190	0.190
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Baseline Crime Non-index Drug	-0.153	-0.154	-0.154	-0.154	-0.154	-0.154	-0.154	-0.154	-0.154	-0.154
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Order of Polynomial	1	2	3	4	5	1	2	3	4	5
Test Against Saturated Model	173.8	143.8	119.7	113.9	113.1	152.7	119.5	112.7	110.3	109.3
degrees of freedom	[102]	[101]	[100]	[99]	[98]	[101]	[99]	[97]	[95]	[93]
p-value	0.0000	0.0033	0.0873	0.1455	0.1414	0.0007	0.0791	0.1323	0.1351	0.1192

Note: Standard errors in parentheses. Table presents alternative parametrizations of the models given in Table II. Coefficients on the polynomial model are suppressed throughout. In each panel the bottom row tests the fit of the presented model against a saturated model which includes a series of exhaustive and mutually exclusive indicators for each possible age, in weeks.

Appendix Table IV. Discontinuity in Probability of Adult Incarceration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Main Estimates								
Estimated Discontinuity in Probability of Confinement	0.1475	0.1474	0.1459	0.1479	0.1464	0.1463	0.1459	0.1460
	(0.0155)	(0.0155)	(0.0154)	(0.0155)	(0.0154)	(0.0154)	(0.0154)	(0.0154)
Log Discontinuity	*1.7246*	*1.7238*	*1.7156*	*1.7264*	*1.7181*	*1.7176*	*1.7152*	*1.7162*
Non-white		-0.0053	0.0012	-0.0043	0.0020	0.0017	0.0017	0.0018
		(0.0037)	(0.0038)	(0.0038)	(0.0038)	(0.0039)	(0.0039)	(0.0039)
Size of County of Baseline Arrest (Relative to Large)								
Small			-0.0476		-0.0472	-0.0471	-0.0469	-0.0468
			(0.0047)		(0.0047)	(0.0047)	(0.0047)	(0.0047)
Medium			-0.0360		-0.0358	-0.0358	-0.0359	-0.0360
			(0.0045)		(0.0045)	(0.0045)	(0.0045)	(0.0045)
Type of Crime, Baseline Arrest (Relative to Non-index Non-drug)								
Index Crime, Violent				-0.0156	-0.0142	-0.0143	-0.0143	-0.0143
				(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0062)
Index Crime, Property				-0.0041	-0.0039	-0.0041	-0.0040	-0.0040
				(0.0056)	(0.0056)	(0.0056)	(0.0056)	(0.0056)
Non-index Drug				0.0029	0.0033	0.0044	0.0039	0.0040
				(0.0079)	(0.0079)	(0.0079)	(0.0079)	(0.0079)
Controls for Age at Baseline Arrest?	N	N	N	N	N	Y	Y	Y
Order of Polynomial in Age at Baseline Arrest						1	3	5
R^2	0.0649	0.0650	0.0695	0.0654	0.0699	0.0700	0.0702	0.0703
B. Decomposition of Probability of Incarceration into Prison and Jail:								
Estimated Discontinuity in Probability of Prison	0.0401	0.0400	0.0394	0.0401	0.0395	0.0395	0.0392	0.0394
	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)
Log Discontinuity	*1.4066*	*1.4054*	*1.3943*	*1.4067*	*1.3954*	*1.3953*	*1.3907*	*1.3930*
Estimated Discontinuity in Probability of Jail	0.1075	0.1074	0.1065	0.1078	0.1069	0.1068	0.1066	0.1067
	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)	(0.0120)
Log Discontinuity	*1.9417*	*1.9410*	*1.9339*	*1.9442*	*1.9370*	*1.9364*	*1.9350*	*1.9353*

Note: Standard errors in parentheses. Log discontinuity calculated from presented discontinuity estimate using a baseline rate of 0.032 for the marginal juvenile. Delta-method standard errors for the log discontinuity are in each case approximately 0.086. Bottom panel of table presents estimated discontinuity in probability of being sentenced to adult prison and to adult jail, respectively. Log discontinuities in bottom panel use a baseline marginal juvenile rate for prison and jail of 0.013 and 0.018, respectively. Delta-method standard errors for these log discontinuities are in each case approximately 0.15 and 0.086, respectively. Controls for regressions in bottom panel suppressed.