The Deterrence Effect of Prison: Dynamic Theory and Evidence*

David S. Lee
Princeton University, NBER
Firestone Library A-16-J-1
Industrial Relations Section
Princeton University
Princeton, NJ 08544
E-Mail: davidlee@princeton.edu

Justin McCrary
UC Berkeley, NBER
586 Simon Hall
Berkeley Law
University of California
Berkeley, CA 94720
E-Mail: jmccrary@berkeley.edu

August 2016

Abstract

Using administrative, longitudinal data on felony arrests in Florida, we exploit the discontinuous increase in the punitiveness of criminal sanctions at 18 to estimate the deterrence effect of incarceration. Our analysis suggests a 2 percent decline in the log-odds of offending at 18, with standard errors ruling out declines of 11 percent or more. We interpret these magnitudes using a stochastic dynamic extension of Becker's (1968) model of criminal behavior. Calibrating the model to match key empirical moments, we conclude that deterrence elasticities with respect to sentence lengths are no more negative than -0.13 for young offenders.

JEL Classification: D9, K4

Keywords: Prison, deterrence, self-control

*Earlier drafts of this paper circulated as NBER Working Paper # 11491, "Crime, Punishment, and Myopia" and Princeton Industrial Relations Section Working Paper # 550, "The Deterrence Effect of Prison: Dynamic Theory and Evidence." Brett Allen, Diane Alexander, and Yan Lau 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 UC Berkeley, the Center for Advanced Study in the Behavioral Sciences, Chicago GSB, Columbia, London School of Economics, the NBER Summer Institute, MIT, Northwestern Law, the University of Illinois at Urbana-Champaign, the University of Rochester, the University of Toronto, University College London, 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.

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 2006, criminal justice system expenditures as a share of national income increased 112 percent, and the ratio of criminal justice employees to the population grew 107 percent (LEAA 1972, U.S. DOJ 2006). Over the same period, the incarcerated fraction of the population increased 373 percent, making the U.S. incarceration rate the highest in the world (Maguire and Pastore, eds. 2001, Chaddock 2003).

Perhaps because of these trends, the economics of crime has emerged as an active area of research. Recent studies have investigated a wide range of potential factors, including the effects of police and incarceration (Levitt 1997, Di Tella and Schargrodsky 2004, Klick and Tabarrok 2005, Chalfin and McCrary 2016, Levitt 1996, Levitt 1998, Helland and Tabarrok 2007, Drago, Galbiati and Vertova 2009), conditions in prisons (Katz, Levitt and Shustorovich 2003), gun ownership (Lott 1998, Cook and Ludwig 2000, Duggan 2001, Ayres and Donohue 2003) parole and bail institutions (Kuziemko 2007, Abrams and Rohlfs 2011), 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, Bayer, Hjalmarsson and Pozen 2009) 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 2006, Mueller-Smith 2015) the impact of wages and unemployment rates on crime (Grogger 1998, Raphael and Winter-Ebmer 2001), the strategic interplay between violent and property crime (Silverman 2004), the effect of incarceration on the supply of crime in the economy (Freeman 1996, Freeman 1999), and optimal law enforcement (Polinsky and Shavell 2000, Eeckhout, Persico and Todd 2010).

One of the key questions for both the literature as well as policymakers is the extent to which more punitive criminal justice sanctions can deter criminal behavior. This notion is at the core of the seminal model of Becker (1968) and of many less formal treatments, such as the classic works by Beccaria (1764) and Bentham (1789), widely cited in the law and economics and criminology literatures.

However, despite a good deal of empirical research, magnitudes of the deterrence effect of criminal justice sanctions remain somewhat uncertain. Our own focus is on the deterrence effect of prison. The most prominent papers addressing this question yield a somewhat wide range of elasticities. On the one hand, assuming exogeneity of changes in the punitiveness of criminal sanctions across U.S. states, Levitt (1998) finds crime elasticities with respect to punishments as large as -0.40, and assuming exogeneity of sentence enhancements given to early prison releasees, Drago et al. (2009) find elasticities as high as -0.74 for a population of Italian offenders. On the other hand, Helland and Tabarrok (2007) compare convicted defendants who become exposed to the threat of the California "three-strikes" statute to observationally equivalent acquitted defendants, and find elasticities of -0.06, an order of magnitude smaller.

In this paper, we use a different identification strategy to isolate the deterrence effects of long prison sentences. 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] If the offense is committed on or after his 18th birthday, however, 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 cost of participating in crime. We argue that while other determinants of criminal offending may change rapidly with age, they do not change discontinuously at 18. This allows us to attribute any discontinuous drop in offense rates at 18 to a behavioral response to adult criminal sanctions, relative to juvenile criminal sanctions.

Our identification strategy is conceptually distinct from a standard Regression Discontinuity (RD) design, which relies on assumptions of imprecise control, and hence the smoothness of the density of the RD running variable to identify treatment effects (Lee 2008, Lee and Lemieux 2010). Here, it is precisely the measured discontinuity in the density of age at offending that we are attributing to a deterrence effect.

Our analysis takes advantage of a large, person-level, longitudinal dataset covering the universe of felony arrests in Florida between 1995 and 2002. This is a major improvement over publicly available data sets, which only capture age in years and are not longitudinal. Publicly available data sets thus make it difficult to distinguish deterrence from incapacitation: a lower arrest rate for 18-year-olds could be entirely driven by more offenders being imprisoned, and thus being unable to commit crimes.[3] In contrast, our data are longitudinal and furnish information on exact dates of birth and offense. We use the precise timing of arrests relative to 18 to isolate a pure deterrence effect, as opposed to an incapacitation effect.

Our discontinuity analysis yields small, precise point estimates: a 1.8 percent decline in the log-odds

[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 below in Section V.

[2]In Florida, our study state, the age of criminal majority is 18. This is typical of U.S. states, but some states have legislated cutoffs at 16 and 17 (Bozynski and Szymanski 2003).

[3]Deterrence refers to the behavioral reduction in crime due to offender anticipation of punishment. Incapacitation refers to the mechanical reduction in crime that occurs when offenders are incarcerated and unavailable to commit additional crimes (against the unimprisoned).

of offending at 18, with standard errors that can statistically rule out declines of 11 percent or more. Our estimates are consistently small across crime categories and across types of jurisdictions within Florida.

These findings can be interpreted directly as an evaluation of at least three different types of policy reforms: (1) reducing the age of criminal majority, (2) increasing the rate at which juveniles are transferred to the adult criminal court, thereby increasing the expected sanction that a juvenile faces, or (3) increasing adult sentences, leaving juvenile sentences fixed. The first two of these policy reforms have been enacted in recent years in multiple states and are currently part of ongoing criminal justice policy discussions (National Research Council 2001, Snyder and Sickmund 1999), and the third policy reform dominated the criminal justice landscape during the 1980s (Levitt 1998). Our results suggest that these reforms have limited deterrence effects for youthful offenders.

To extrapolate our results to evaluate a broader range of policy reforms of interest, and to connect our empirical results to structural parameters of offender behavior, we develop a stochastic dynamic model of crime. The model retains at its core the essence of the static Becker model of criminal behavior, but places the individual in a dynamic setting. The model makes precise the link between our discontinuity estimates of the intertemporal behavioral response to an anticipated increase in punishments and other policy-relevant deterrence elasticities. We calibrate the model to match easily attainable empirical quantities. This calibration exercise leads us to three additional conclusions.

First, our model and estimates are inconsistent with "patient" time preferences (e.g., annual discount factors of 0.95) and suggest much smaller discount factors. Essentially, the small change in behavior at 18 is consistent with offenders having short time horizons, leading them to perceive little difference between nominally long and short incarceration periods. Second, a moderately-sized (e.g., -0.25) elasticity of crime with respect to the probability of apprehension is consistent with our estimates, but requires the discount factor to be extremely small, which in turn implies a small elasticity with respect to sentence lengths. This result follows because our model leads to implied bounds for deterrence elasticities for police and prison. Finally, the most negative elasticity with respect to sentence lengths that is consistent with our estimates and model is -0.13.

In addition to estimating deterrence parameters, we also present RD evidence on the incapacitation effect of adult sanctions. We find that being arrested just after 18 leads to an immediate (within 30 days) subsequent reduction in offending, relative to being arrested just before 18. We interpret this as evidence of the incapacitation effect of adult sanctions. Overall, from our main analysis as well as this supplementary evidence on incapacitation, we conclude that if lengthening prison sentences leads to significant crime reduction, it is likely operating through a direct, "mechanical" incapacitation effect, rather than through a behavioral response to the threat of punishment.

The remainder of the paper is organized as follows. Section II discusses related studies on deterrence. Section III describes our identification strategy and approach to estimation, while Section IV describes our data in detail. Section V presents our main results. In Section VI, we develop our stochastic dynamic model, calibrate it, and interpret our reduced-form magnitudes through the lens of this model. Section VII concludes. Two appendices provide further detail on our data and theoretical model.

II. Existing Literature

This section briefly describes a number of prominent studies that are most related to our analysis. These recent studies attempt to address an important problem that arises in isolating the causal impact of more punitive criminal sanctions on criminal behavior: crime control policies (e.g., policing, sentencing) are often suspected to be endogenous responses to current crime levels and trends. The problem of policy endogeneity has long been recognized in the literature (Ehrlich 1973, Ehrlich 1987, Nagin 1998, Levitt 2004, Donohue and Wolfers 2008). To circumvent policy endogeneity, recent studies exploit arguably exogenous variation in the punitiveness of criminal sanctions (i.e., prison sentences).[4] The elasticities resulting from these analyses range from somewhat small (-0.06) to somewhat large (-0.74).

The largest magnitudes come from the analyses of Levitt (1998) and Drago et al. (2009). The highly-cited study of Levitt (1998) is the most closely related to our own. The analysis uses a state-level panel, regressing juvenile crime rates on a measure of punitiveness of the juvenile justice system (number of delinquents in custody per 1000 juveniles), including state and year effects, as well as other control variables. It also disaggregates the data by age cohort, and implements a difference-in-difference strategy, whereby the changes in adult-juvenile relative crime rates are compared between states that had higher and lower increases in adult-juvenile relative punitiveness over time.[5] Levitt (1998) concludes that juveniles are significantly responsive to criminal sanctions, reporting effects that imply an elasticity of -0.40 for violent crime.[6]

[4] See, for example, Nagin's (1978) criticisms of much of the older literature due to its failure to recognize problems with endogeneity. See also the discussion in Freeman (1999) and Levitt and Miles (2007).

[5] Relative punitiveness is defined as the number of adult prisoners per adult violent crimes, relative to the number of delinquents per juvenile violent crimes.

[6] On p. 1181, it is noted that "[b]etween 1978 and 1993, punishment per crime fell 20 percent for juveniles but rose 60 percent for adults. Over that same time period, rates of juvenile violent and property crime rose 107 and 7 percent, respectively. For adults, the corresponding increases were 52 and 19 percent. On the basis of the estimates of Table 2, if juvenile punishments had increased proportionally with those of adults, then the predicted percentage changes in juvenile violent and property crime over this period would have been 74 and 2 percent." In other words, if punishments rose 60 percent instead of falling

Another study finding large elasticities of the deterrence effect of prison is that of Drago et al. (2009), which examines the crime impact of a 2006 Italian clemency act. This statute led to the release of prisoners whose crimes were committed prior to May of 2006, subject to some minor exceptions. Importantly, the statute contained a sentence enhancement provision: for releasees who were re-arrested within five years of release and subsequently sentenced to more than two years, their sentence would be augmented by the amount of time that was remaining on their first sentence at the time of clemency. The paper documents the empirical relation between re-arrest within 7 months of release and the time remaining on the sentence length at time of clemency, controlling for other observables, such as the original sentence length. The results suggest an elasticity of crime with respect to sentence length of -0.74 at 7 months follow-up and approximately -0.45 at 12 months follow-up.[7] The identification strategy is based on sentence enhancements, and this quantity is interpreted as representing a deterrence effect. Given that upon release, the difference in age between those with little or much time remaining on their sentences are similar, the key identifying assumption here is that those who are arrested earlier in life have the same propensity to commit crime as those arrested later in life.

At the other end of the range of elasticities, Helland and Tabarrok (2007) and Iyengar (2008) find generally small behavioral responses of criminals to California's three-strikes law. This statute requires that those previously convicted of two "strikeable offenses" who are subsequently convicted of any felony must receive a prison sentence of 25 years to life, with a further requirement that parole can begin no earlier than completion of 80 percent of the sentence.[8] Helland and Tabarrok (2007) assess the impact of the three-strikes law using data on prisoners released in 1994, the first year the statute was in effect. The paper compares the recidivism pattern between those previously convicted of two strikeable offenses and those previously tried for two strikeable offenses, but convicted for only one of those offenses. The estimates suggest that three-strikes lowered the incidence of crime by about 20 percent. Since the increase in expected sentence lengths associated with three-strikes is at least 300 percent, this implies an elasticity estimate on the order of -0.07.[9] Using a different data set and approach Iyengar (2008) also studies the three-strikes statute and arrives at a set of estimates that imply similar elasticity estimates, on the order of -0.10.[10]

20 percent (a difference of 80 percent) crime would have risen by only 74 percent rather than 107 percent (a difference of 33 percent), implying an elasticity of about $33/80 \approx 0.41$.

[7]See Drago et al. (2009, pp. 273–274).

[8]The statute details those offenses which are strikeable. Essentially, a strikeable offense is a felony that is serious or violent.

[9]Helland and Tabarrok (2007, pp. 327–328) argue that three-strikes increased expected sentences at 3rd strike eligibility by 330 percent, or from about 60 months to at least 260 months.

[10]Iyengar (2008) estimates a 16 percent decline in the incidence of crime at 2nd strike eligibility (associated with a statutorily required doubling of the sentence) and a 29 percent decline in crime at 3rd strike eligibility.

These papers differ in several different ways, making it hard to know precisely why the elasticity estimates are so different. The papers analyze data from different jurisdictions (the U.S. as a whole, Italy, and California) and for different time periods. They also employ different methods for measuring crime. Levitt (1998) measures crime using crime aggregates, while Drago et al. (2009), Helland and Tabarrok (2007), and Iyengar (2008) measure crime using recidivism. However, perhaps the most important difference in these papers is the assumptions employed. Arguably, the most plausible set of assumptions invoked are those in Iyengar (2008), who assumes that the *ordering* of criminal history is idiosyncratic—that is, the criminal propensity of individuals with a theft then a burglary on their record equals that of those with a burglary then a theft on their record. In our view, this assumption is still stronger than the smoothness assumptions we employ, however is more plausible than the difference-in-difference approaches employed by the other three papers.

As noted, our approach differs from those in the previous literature by exploiting smoothness assumptions, which, combined with our high-frequency data, allows us to separate deterrence from incapacitation effects of sanctions. The use of annual data (cf., Levitt (1998)) has the potential to conflate deterrence and incapacitation effects. The scope for conflation of these concepts is large when differences between adult and juvenile incarceration rates appear within a year of the 18th birthday. In Section V, we show empirically that differences in incapacitation rates emerge rapidly and are evident within even 30 days after the 18th birthday.

It is important to note that our estimates are most relevant for individuals arrested, but not necessarily convicted, on felony charges in Florida prior to age 17. These juveniles are likely to have high propensities for crime, and thus our sample is likely to be more comparable to the Italian ex-prisoners in Drago et al. (2009) or the Californian repeat offenders in Helland and Tabarrok (2007), than to the aggregate data utilized in Levitt (1998), which also includes first-time and low-propensity offenders.

Note that our approach is conceptually distinct from the standard RD design, which relies on an assumption that individuals do not precisely manipulate their running variable around the threshold of interest. In the standard RD context, a discontinuity in the density of the running variable constitutes evidence against the validity of the RD design (McCrary 2008). In our main analysis, the discontinuity in the density of age at offense measures deterrence and is itself the object of interest. We utilize a more standard RD analysis in our measurement of incapacitation, as discussed in Section V.

The final distinctive feature of our analysis is that we interpret our reduced-form estimates through the lens of a stochastic dynamic model of criminal behavior. We thus expand on Becker's (1968) model and explicitly consider the dynamic nature of the problem, much in the spirit of the theoretical work of Lochner (2004).[11]

III. Identification and Estimation

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 ways. A significant difference is that the expected length of incarceration is significantly longer when the offender is treated as an adult, rather than as a juvenile.[12] We quantify this difference in Section V, below. To identify the deterrence effect of adult prison, we follow a cohort of youths from Florida longitudinally, and examine whether there is a discontinuous drop in their offense rates when they turn 18. Our data contain exact dates of birth and offense, allowing us to analyze the timing of offenses 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 responses to sanctions.

Our approach does not require the determinants of criminal behavior to 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. We are arguing that on the day of the 18th birthday, there is no discontinuous change in the ability of law enforcement to apprehend an offender; no "jump up" in wages; no discontinuous change in the distribution of criminal opportunities; and so on. There are some exceptions to this (e.g., the right to vote and the right to gamble change discontinuously at 18), but we deem these to be negligible factors in an individual's decision whether to commit crime.

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). Indeed, the determinants of criminal behavior change significantly at a year-to-year frequency, as is apparent from the age distribution of arrestees at annual frequencies (see Appendix Figure 1).[13] In the age range of 17, 18, or 19, for example, youth are

[11] See McCrary (2010) for a review of the literature on dynamic models of criminal behavior. An important early paper in this literature is Flinn (1986), and a recent paper is Sickles and Williams (2008).

[12] Further, taking into account the conditions of confinement as well as the duration of confinement surely amplifies the deterrence effect of the adult system. As noted by Parsell (2012), "Congressional findings in the Prison Rape Elimination Act of 2003 posited that juveniles were five times as likely to be sexually assaulted in adult rather than in juvenile facilities--often within their first 48 hours of incarceration. Youth advocacy groups report that juveniles housed in adult facilities are 36 times more likely to commit suicide."

[13] Appendix Figure 1 plots the frequency distribution of age (measured in years), for those arrested for an index crimes (murder, rape, robbery, assault, burlgary, larceny, and motor vehicle theft), as computed using the Federal Bureau of

graduating from high school, starting new jobs, and developing physiologically and psychologically in ways that could affect underlying criminal propensities (Lochner 2004, Wilson and Herrnstein 1985). Since this research design relies heavily on the continuity of all factors aside from sanctions, in Sections IV and V, we assess a number of alternative ways in which data issues or police discretion in reporting could affect our interpretation of the discontinuity in arrest rates as a response to more punitive sanctions.

Implementing our research design is straightforward. Here we describe the basic idea behind our approach, and later describe minor adjustments associated with implementation. Suppose we have a sample of N individuals, and we can track their subsequent offending behavior, starting at age 17. Then for each week, we can calculate the number of individuals arrested for the first time since 17, as a fraction of those who are still at risk of doing so. If n_1, n_2 , and n_3 are the number of individuals who are arrested in the first, second, and third weeks after their 17th birthday, respectively, then nonparametric estimates of the hazard of arrest are given by $\widehat{h}(1)=n_1/N$, $\widehat{h}(2)=n_2/(N-n_1)$, $\widehat{h}(3)=n_3/(N-n_1-n_2)$. We refer to these as local average estimates of the hazard of offense. The empirical question to be addressed is whether the hazard drops off precipitously at age 18, as seems intuitive and as is predicted by deterrence theory, as we discuss below.

It is convenient to summarize these averages and the corresponding discontinuity with a flexible parametric form. To do this, we estimate a panel data logit model, as suggested in Efron (1988). Specifically, we organize the data set into an unbalanced panel, with N observations for the first period, $N-n_1$ for the second, $N-n_1-n_2$ for the third, and so on. We then estimate the logit model

$$\mathbb{P}(Y_{it}=1|X_t,D_t) \quad = \quad F(X_t'\alpha+D_t\theta) \tag{1}$$

where Y_{it} is the indicator for arrest for persion i in period t, $X_t \equiv (1,t-t_0,(t-t_0)^2,...,(t-t_0)^q)'$, q is the order of the polynomial, t_0 is the week of the 18th birthday, D_t is 1 if $t \geq t_0$ and is 0 otherwise, and $F(z)=\exp(z)/(1+\exp(z))$.[14]

Below, we compare the predicted values from the logit model to the local average hazard estimates $\widehat{h}(t)$, and report a likelihood ratio test for the restrictions imposed by the logit form.[15] As will be clear from the empirical results below, these models do a good job of providing a parsimonious but accurate

Investigation's Uniform Crime Reports (UCR). Because our administrative data pertain to Florida during the period 1995-2002, we show the data for Florida and the U.S. for 1995 and for the U.S. in 2002.

[14] The vector X_t can also include interactions of the polynomial with the indicator for adulthood. Practically, because the regressors only vary at the level of the group, we estimate the model at the group level and avoid the construction of the (large) micro data set.

[15] To do so, we estimate a logit model with functional form $F(W_t'\pi)$, where W_t is a series of indicators for each week, and compare the fit to the (nested) parametric approximation.

fit to the functional form suggested by the local averages. 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.

In a later section, we estimate the incapacitation effect of adult sanctions. For this, we focus on the timing of the second arrest since the 17th birthday, viewed as a function of age at first arrest since 17.[16] Specifically, we compare how quickly a re-arrest occurs after being arrested just before age 18, to how quickly it occurs if arrested just after 18. If the marginal adult takes longer to re-offend, this provides some indirect evidence of an incapacitation effect at work. While this is not the main target of our analysis, it nevertheless provides some information on the incapacitation mechanism. This is particularly useful for quantifying the sentence length elasticity and for calibrating the economic model that we develop below. The implementation of this analysis follows a more standard RD design, where the running variable is the age at first arrest since 17, the "treatment" is whether that arrest occurred before or after age 18, and the dependent variable is whether the second arrest occurs within, for example, 30 days.

IV. Data and Sample

In this Section we describe our data, our main analysis sample, and how various features of our data affect our estimation. We also discuss why our sample is unlikely to be affected by differential reporting at age 18. For ease of exposition, we defer to Section V a more detailed discussion of expungement of juvenile records.

A. Main Analysis 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, the 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. Importantly, the data are longitudinal: each arrest incident is linked to a person-level identifier.[17] The raw data, therefore, can be described as a database of individuals, each with an associated arrest history.

From this database, we focus on three key arrest events that define our sample and outcomes of interest:

[16] Note that this approach to estimating incapacitation effects of prison is accurate to the extent that there are small to no deterrence effects of prison.

[17] A more detailed description of the database and its construction is provided in the Data Appendix.

(1) the first arrest recorded in our administrative data, which we refer to as the "baseline arrest", (2) the first arrest since the 17th birthday (we call this the "first arrest"), and (3) the second arrest since the 17th birthday (we call this the "second arrest"). Our main analysis sample is defined by those whose baseline arrest occurs prior to the 17th birthday, for a total of $N = 64{,}073$ individuals.[18] To examine deterrence we examine the incidence of the "first arrest" at age 18, and to provide some evidence on incapacitation, we examine the time between the "first arrest" and "second arrest" as a function of the age at "first arrest."

More specifically, for our analysis on deterrrence, we implement the estimation approach described in Section III, with the following modifications and considerations:

- Our last date of observation is December 31, 2002, leading to some individuals being censored. The standard way to adjust for this is to compute hazards as $\widehat{h}(1) = n_1/N$, $\widehat{h}(2) = n_2/(N - n_1 - m_1)$, $\widehat{h}(3) = n_3/(N - n_1 - n_2 - m_1 - m_2)$, and so on, where m_1 and m_2 are the numbers of individuals who are 1 and 2 weeks into their 17th year on December 31, 2002 (i.e., censored), respectively. Following Efron (1988), we construct the unbalanced panel for the logit in an analogous way.
- We note that $N = 64{,}703$ is larger than the true "at risk" population, since at any age $t > 17$, an individual could still be incarcerated for the "baseline arrest" or could be incarcerated because of a non-Index crime arrest that occurred between age 17 and t.[19] This fact by itself will not generate a discontinuity in our estimated hazards, as long as the true "at risk" population is evolving smoothly in age, particularly at age 18. There is an institutional reason for a discontinuity in the number "at risk" at age 21; Florida law mandates that no individuals above age 21 can be held in a juvenile correctional facility. However, there does not appear to be such a reason for an effect at age 18.
- In terms of the estimated magnitude of our discontinuity, this means our logit parameter θ estimates

$$\ln\left(\frac{\rho h}{1-\rho h}\right) - \ln\left(\frac{\rho h_J}{1-\rho h_J}\right) = \ln(h) - \ln(h_J) + \ln(1-\rho h) - \ln(1-\rho h_J)$$

where h and h_J are the arrest hazards for an 18.02 and 17.98-year-old, respectively, and $\rho \in (0,1)$ reflects the over-estimation of the risk set. Thus if $h_J > h$, as would be consistent with deterrence, then

$$\ln\left(\frac{h}{1-h}\right) - \ln\left(\frac{h_J}{1-h_J}\right) \leq \theta < \ln(h) - \ln(h_J) < 0$$

In practice, the difference between the upper and lower bounds is scant. Consequently, there is little

[18] The baseline arrest is allowed to be any arrest. The first and second arrests since 17 are restricted to be index arrests unless otherwise specified.

[19] Also, in our main analysis we focus on the incidence of Index felonies, and so at any post-17 age, the individual could be incarcerated from a non-Index felony arrest.

loss in thinking of θ as estimating the percentage change in the hazard as a youthful offender transitions to adulthood.

- Our main analysis sample will include individuals who may be more likely be affected by the increase in sanctions. In particular, it seems likely that for this group, there is a perceived positive net benefit to criminal activity. After all, an individual in our sample has already been arrested at least once by 17, indicating that at least one crime was deemed worthwhile. 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. For these near-certain law-abiders, it would be mechanically impossible for their criminal activity to decline after 18.
- 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 "baseline arrest." Glassner, Ksander, Berg and Johnson (1983) provide anecdotal evidence to support this viewpoint.[20]
- Our main analysis 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. But our sample, which requires having at least one juvenile record, necessarily consists of those who have not expunged their entire juvenile arrest record. We discuss this in greater detail in Section V.

Table 1 reports some summary statistics for our main analysis sample. We begin with 64,073 individuals whose baseline arrest occurred prior to 17. 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 two columns present information on the baseline arrest and the first arrest since 17. Age at the baseline arrest is about 15, and the most common category of offense is property crime, followed by violent crime. Individuals are distributed evenly among small, medium, and large counties.[21] Slightly less than half of our main analysis sample is observed reoffending after age 17. The sex and gender composition of arrestees is similar at baseline arrest and first arrest since 17, as is the county-size distribution. Offenses are distributed somewhat more evenly among the four crime types described at first arrest since 17 than at baseline arrest. For comparison, the final column reports the same

[20] For example, responding to a question regarding how he knew that sanctions were more punitive after the age of criminal majority, one twelve-year-old interviewed by those 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 me that" (p. 220).

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

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 main estimation sample is broadly representative of this larger arrest population.

B. Measurement and Reporting Discontinuities

Our approach requires that criminality be measured in a smooth fashion near 18. There are some potential threats to continuous measurement, which we now discuss. One issue is that crimes are not necessarily defined similarly for juveniles and adults (e.g., truancy, corruption of a minor, statutory rape, and so on). To avoid definitional problems, we consider only crimes that are well-defined for both juveniles and adults, such as burglary.

Even for crimes that are defined the same way for juveniles and adults, it is still possible that police officers could exercise discretion in executing an arrest that would lead to discontinuous measurement. For example, a police officer might view possession of small amounts of marijuana as forgivable for youth and might overlook the incident upon learning that the individual was still a minor. This would result in a discontinuous increase at 18 in the probability of observing criminality in arrest records.[22] Alternatively, one could imagine that the officer would view possession of marijuana as forgivable, but would want to "teach a lesson" to the offender, as long as the cost of the lesson were not too great. This would result in a discontinuous decrease at 18 in the probability of observing criminality.

To avoid such problems with discontinuous measurement of criminality at 18, we focus on arrests for so-called Index crimes: murder, rape, robbery, assault, burglary, and theft, including motor vehicle theft. We are confident that these felonies are sufficiently serious that an arresting officer would not overlook an offense. An additional benefit of focusing on Index crimes is that these are the crimes most commonly studied in the literature.

Finally, even if police officers themselves do not exercise discretion regarding making an arrest, administrative records may be more complete for adult arrests than juvenile arrests. The data from Florida do not suffer from this problem for the post-1994 period, due to an important criminal justice reform, the 1994 Juvenile Justice Reform Act, which requires that police departments forward records of serious juvenile felony arrests to the FDLE (see Section V).

[22]Another example that leads to the same prediction of an increase at 18 in the probability of observing criminality is from Levitt (1998). There, a hypothetical officer desires to see an arrestee punished and deems that the punishment accorded a juvenile is not worth the paperwork required to complete the arrest.

V. Evidence on Deterrence and Incapacitation

This section presents our main empirical results. We address two issues. First, we quantify the deterrence effect of adult criminal sanctions, relative to juvenile criminal sanctions, by studying the pattern of offending around the 18th birthday. Second, we quantify the incapacitation effect of adult criminal sanctions, relative to juvenile criminal sanctions, by studying the change around the 18th birthday in durations between arrests.

A. Evidence on Deterrence

Our main result is summarized by Figure 1. The top panel of this figure is an empirical hazard function for being arrested, between ages 17 and 19, among those arrested at least once before their 17th birthday. Each open circle represents those arrested in a given week as a proportion of those not yet arrested. For example, the first circle shows that about 0.005 of the main sample are arrested within a week of their 17th birthday. Of those who still not had been arrested by their 18th birthday, almost 0.0025 are arrested in the week of their 18th birthday. The solid line gives predicted probabilities of arrest, based on maximum likelihood estimates of the logit in equation (1).[23] 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 typical week-to-week differences.

For comparison, the bottom panel of Figure 1 plots the analogous weekly arrest probabilities for those who were arrested at least once before their 19th birthday. We track the arrest records for individuals in this "placebo" sample for two years, from 19 to 21. The arrest probabilities are smooth in age, as would be expected since this age range is well past 18, and there is no legal significance to being 20 years old. The top and bottom panels of the figure are quite similar.

Table 2 reports estimated discontinuities in arrest probabilities at 18, based on the logit model of Equation (1). These estimates support the inference suggested by Figure 1: the drop in arrests at 18 is small in magnitude and statistically insignificant. The estimated discontinuity is roughly -0.018, with an estimated standard error of about 0.047.[24] These estimates imply that we can statistically rule out values

[23] Here, X_t is a cubic polynomial in t. These predictions correspond to estimation of the model in column (1) of Table 2.

[24] We report in Table 2 standard errors based on the observed information matrix (OIM). These are accurate if the logit model is the correct parametric model within our neighborhood of one year. We note that the standard errors are highly similar to those obtained from the outer product of the gradient (OPG) or (consequently) the robust or "sandwich" approach of White (1982). For example, to four digits, the estimate in column (1) is -0.0180 with an OIM standard error of 0.0474; the OPG standard error is 0.0472; and the robust standard error is 0.0475. As another example, the corresponding four numbers for column (7) are -0.0172, 0.0474, 0.0472, and 0.0475. The similarity of these estimated standard errors is consistent with the implications of the information matrix equality. One interpretation of this similarity is that the data contain little evidence against our parametric logit approximation within one year of the 18th birthday.

of θ more negative than -0.11 at the 5 percent level of significance.[25]

In Section VI, below, we present evidence that the expected sentence length facing an adult arrestee is roughly 230 percent greater than that facing a juvenile arrestee. The small decline in arrest rates at 18 thus suggests an elasticity of crime with respect to expected sentence lengths no more negative than -0.048. In Section VI, we compare these magnitudes to the predictions from a dynamic economic model of criminal behavior.

The estimated discontinuity is robust to changes in specification, corroborating the smoothness assumptions required for our approach. Moving from left to right in Table 2, we control for an increasing number of covariates measured as of the time of the baseline arrest. 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 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 significantly affect the estimated discontinuity.

Appendix Table 3 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.

Our approach to estimation and inference follows Efron (1988) and thus is rooted in parametric models applied to local neighborhoods. Arguing in favor of this method is that our parametric testing approach allows us to verify that the local averages are not inconsistent with the parametric model fitted to them, as described above. However, an alternative approach to estimation and inference in general RD contexts is linear estimation applied to potentially smaller local neighborhoods (e.g., Hahn, Todd and van der Klaauw 2001, Calonico, Cattaneo and Titiunik 2014).

A challenge to comparing our results to the nonparametric inference results of Calonico, Cattaneo and Titiunik (2014) is that their results are limited to linear models—and hence do not apply directly to nonlinear

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

models such as logit. Nonetheless, as is by now widely appreciated (c.f., Angrist and Pischke 2009), regression and logit marginal effects rarely differ substantially. Here, the parameter θ can be accurately approximated as $(h-h_J)/h_J$. This suggests a natural way to connect our approach to that of Calonico et al. (2014): (1) estimate h_J using local linear regression and the local averages from Figure 1 before the 18th birthday, (2) estimate the difference $h-h_J$ using local linear regression and all the local averages, and (3) scale the estimate of the difference $h-h_J$ by $1/h_J$ using the estimate from (1).[26] This approach can be thought of as an alternative approach to Column (1) of Table 2, since those estimates rely only on the local averages plotted in Figure 1.

In step (1) from above, we use a bandwidth of 135 days because that is what is implied by the bandwidth selector employed by `rdrobust` in step (2), below. Local linear estimation of the hazard with this bandwidth leads to a boundary estimate of $1/\widehat{h}_J = 1/0.00247 \approx 404$. In step (2), local linear estimation with quadratic models for bias adjustment leads to a point estimate (standard error) of -0.018 (0.065).[27] For comparison, our parametric point estimate (standard error) from Column (1) is -0.018 (0.047).

In the approach of Calonico et al. (2014), we adjust the point estimate for the possibility of bias; doing so leads to an adjusted estimate of -0.038. Those authors present an estimated standard error that additionally incorporates the uncertainty due to the fact that the bias is estimated. This flavor of standard error is somewhat higher: 0.075, or about 17 percent larger than the local linear standard error ignoring bias estimation. Consequently, and in summary, while the Efron (1988) parametric approach suggests a point estimate (standard error) of -0.018 (0.047), a local linear framework suggests -0.018 (0.075).[28]

These comparisons clarify that changes to estimation and inference approaches generally corroborate the magnitudes from our more parametric analysis. Point estimates are quite similar, with second-decimal differences arising only in the context of bias adjustment. Even those differences are not large relative to the range of sampling uncertainty. Changes to standard errors are more meaningful and yet are also qualitatively similar to what was obtained from the more parametric approach.

Below, we set aside econometric considerations to consider some potential threats to the validity of our interpretation of these discontinuities as reflecting deterrence effects.

[26] Practically, we implement this using the `rdrobust` package using the default options unless otherwise specified. The package was downloaded from `https://sites.google.com/site/rdpackages/rdrobust`.

[27] This standard error is similar to that advocated by Hahn, Todd and van der Klaauw (2001) and is described by `rdrobust` as "conventional."

[28] If we change to local cubic estimation with fourth-order models for bias adjustment, the point estimate (standard error) analogous to the local linear framework above increases in magnitude to -0.0498 (0.1091). Bias adjustment now reduces the magnitude of the point estimate, to -0.0424, and standard errors robust to bias estimation increase to 0.1220, or about 12 percent larger than the standard error ignoring bias estimation.

B. Transfers of Juveniles to the Adult Criminal Court

The first potential threat to the validity of our interpretation 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 (Government Accounting Office 1995).[29] In principle, prosecutors could be more likely to request that a juvenile 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 no "treatment."

Our data allow us to rule out this possibility empirically. The top panel of Figure 2 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 individuals against whom a formal prosecution was filed, expressed as a fraction of those who were arrested in that particular week.[30] There is a striking upward discontinuity at the age of 18.[31] Those who are arrested just before their 18th birthday have about a 20 percent chance of being formally prosecuted as an adult, while those arrested just after their 18th birthday have a 60 to 70 percent chance. The latter probability is not 1, because not all arrestees will have formal charges filed against them.

The bottom panel of Figure 2 provides further evidence of a discontinuity in the treatment, using a different measure of punishment. It 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 3, which reports coefficient estimates from different OLS regressions. In columns (1) through (5) the dummy variable for whether the individual was prosecuted is regressed on the juvenile/adult status dummy, a cubic polynomial in age at arrest, as well as its interaction with the juvenile/adult dummy, and other covariates. Columns (6) through (8) further include the age at the baseline arrest as an additional control. Across specifications, the discontinuity estimate of about 0.40 is relatively stable. Appendix Table 4 is an analogous table for the probability of being convicted

[29] In Florida, prosecutors have the discretion to try a juvenile arrestee in the adult criminal court. This typically results in a larger number of juveniles tried as adults than the other common form of juvenile transfer, in which the juvenile court judge retains the authority to transfer juveniles (Snyder and Sickmund 2006).

[30] Therefore, the sample for this figure is the same as that underlying the top panel of Figure 1.

[31] We note that, analogous to McCrary and Royer (2011), assessing the magnitude of the "treatment" discontinuity is possible here because there is little evidence of deterrence. Were there to be strong evidence of deterrence, the composition of the sample would potentially be quite different for those aged 17.98 and 18.02.

as an adult and sentenced to prison or jail.

Overall, the evidence strongly suggests that juvenile transfers to the criminal court are not prevalent enough to eliminate a sharp discontinuity in the punitiveness of criminal sanctions at 18.

C. Age-based Law Enforcement Discretion

Another possibility is 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. Suppose that 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.

There are a number of reasons why we believe this is 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. For example, all Index crimes 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 drug possession, it is more plausible that officers might exercise discretion in making the arrest.

Second, each individual in our main estimation sample already has 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: if a juvenile is apprehended just before his eighteenth birthday, 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 requires that felonies and some misdemeanors committed by juveniles be forwarded to the state for inclusion in the criminal history records maintained by the FDLE.[32] The impact of this law on the prevalence of juvenile records is shown in Appendix Figure 2. This figure show the number of juvenile arrests as a proportion of all arrests in the FDLE arrest data, by month from 1989 to 2002. There is a marked discontinuity in the ratio at October 1994, the month the JJRA took effect. This feature of

[32] The implication of this Florida law was summarized by a state 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).

Florida's criminal justice system was a major factor influencing us to use Florida as our study state.

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 3 provides evidence contrary to this prediction. The top panel of the figure disaggregates the arrest probabilities from the top panel of Figure 1 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 4 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 all zero in Table 2, and this holds for all specifications considered.

Although our analysis focuses on Index crimes, for completeness we show the results for all remaining offenses in the bottom panel of Figure 3. 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. Here, the cubic polynomial predictions do show a small perverse discontinuity, although the local averages do not reveal an obviously compelling jump at age eighteen. Still, if law enforcement discretion is of particular concern, a conservative approach would be to discount the results for non-Index crimes.

D. Expungement of Juvenile Records

A third possibility is that the ability of individuals to expunge and seal juvenile arrest records could generate downward biased estimates of arrest rates for juveniles, and hence mask any true deterrence effects. 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.[33] Our choice of sample, however, circumvents these two expungement provisions in the following ways. First, our estimation sample is restricted to those committing baseline crimes before age 17 but subsequent to January 1, 1995. Therefore, the individuals are not yet 24

[33]Florida creates an exception to this general rule when the individual has committed a serious offense as an adult. Juvenile records for "habitual offenders" are retained by the FDLE until the offender is 26 (Fla. Stat. 943.0585, 943.059).

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 observed arrest record prior to age 17. Thus, by construction, the individuals in our main estimation sample did not have their complete juvenile arrest history expunged.

To illustrate the importance of this sample choice to avoiding problems related to expungement, we present the time profile of arrests for the individuals who are excluded from our main analysis: those who were not observed as arrested prior to turning 17. Some of these individuals' 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 4. 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).[34]

E. Evidence on Incapacitation

Up to this point, we have examined the evidence for a deterrence effect of adult criminal sanctions, relative to juvenile criminal sanctions. We now ask: What is the incapacitation effect of treating an apprehended offender as an adult instead of a juvenile? To answer this, we use a standard RD design, where the "treatment" of adult status is a discontinuous function of the running variable of age at "first arrest." As described in Lee (2008), if there is imprecise sorting around the age of majority, then being treated as an adult has statistical properties analogous to a randomized experiment.[35] Indeed, in the analysis above we are unable to detect strong evidence of such sorting behavior. In essence, our results on deterrence imply that this design passes the test of manipulation of the running variable suggested in McCrary (2008), enabling credible assessment of incapacitation effects despite analyzing a selected sample of arrests as opposed to all individuals, arrested or not.

The RD design for detecting incapacitation effects is illustrated in Figure 5. The top panel of the figure plots the probability that the "second arrest" occurs within a specific window of time since the "first arrest," as a function of the age at "first arrest." Specifically, the leftmost open circle indicates that among those whose "first arrest" occurs the week after their 17th birthday, the probability of re-arrest within

[34]The dark bars represent the values used in the upper panel of Figure 1, except that Figure 1 normalizes each value with the at-risk population at each point in time to provide a probability value.

[35]By "imprecise sorting" we mean that for each individual, the density function of age at arrest is continuous at 18. See Lee (2008) for further discussion.

the subsequent 30 days is about 15 percent. Among those whose "first arrest" occurs just before their 18th birthday, the probability of re-arrest within the subsequent 30 days is about 20 percent.

There is a sharp discontinuity in the probability of re-arrest at age 18, with 18.02-year-olds having a probability of re-arrest within 30 days of about 10 percent. A natural explanation for this difference is that being handled by the adult criminal court 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 re-offend within any short time window, compared to the 18.02-year-old.

The solid circles and open triangles plot the same kind of graph, except that we examine the probability of a re-arrest occurring within 120 and 365 days after the "first arrest." The probabilities are higher for 120 and 365 days, since the probability of re-arrest increases with the window width.

The length of the follow-up period is arbitrary. The bottom panel of Figure 5 plots the profile of discontinuity estimates using follow-up lengths ranging from 1 to 500 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 already by 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.[36] This divergence continues to grow, slowing down at around 100 days after the initial arrest.[37]

VI. Predicted Effects from a Dynamic Model of Crime

In this Section we interpret the magnitudes of our estimated deterrence effects through the lens of an economic model of criminal behavior. We develop a dynamic extension of Becker's (1968) model of crime. We first consider how large of a discontinuity we should expect given our model, which can be calibrated to readily available sample means, and standard assumptions. We then use our model to draw a precise link between our discontinuity estimates, which identify a particular set of policy responses as discussed above, and other policy elasticities of interest that measure deterrence effects in a broader way. In doing so,

[36] It is worth noting that by Florida law, juvenile pre-trial detention cannot be longer than 21 days. No such restriction applies to adult pre-trial detention. Investigation of the juvenile hazard of re-offense indicates a strong spike near 21 days (results unreported).

[37] Note also that an incapacitation interpretation is not the only one consistent with the data. Alternatively, these data are also consistent with no difference between juvenile and adult lengths of incarceration. This could occur if being incarcerated in the adult criminal justice system had a negative causal effect on criminal propensities upon release. This notion is sometimes referred to as "specific deterrence," particularly in the criminology literature. See Cook (1980) for discussion of the concept and Hjalmarsson (2009b) for a recent empirical approach. This alternative interpretation requires a mechanism by which the quality of the experience of being incarcerated as an adult induces individuals to be more law-abiding upon release.

we employ both standard parametric assumptions as well as consider a nonparametric approach where we impose a monotonicity restriction on the distribution of criminal benefits. We show how that nonparametric restriction allows us to provide informative bounds on the broader policy elasticities that are not directly identified by our study. This nonparametric analysis delivers the most negative deterrence elasticities that are consistent with both our model and our reduced-form estimates.

A. The Model

The essence of Becker's model of crime is that an individual weighs the expected benefits (e.g., monetary or otherwise) against the expected costs (e.g., fines, disutility from being incarcerated). This basic notion can be captured in a discrete-time dynamic model, in which the individual faces a criminal opportunity every period and chooses between committing the offense and abstaining. If he commits the offense, there is a chance that he will be apprehended and incarcerated for a random number of periods before being released.

The model described below closely resembles the canonical job search model (McCall 1970). The elements of the model are as follows:

- In each period, the individual is presented with a criminal opportunity. The value of this opportunity, or benefit, is given by the non-negative continuous random variable B with finite mean, cumulative distribution function $F(\cdot)$, and density $f(\cdot)$.[38]
- After observing the value of the current period's criminal opportunity, the individual chooses to offend or abstain. In a given period t, if he abstains, he receives flow utility $u_t = a$. If he offends and is not caught, he obtains the flow utility $u_t = a + B$. If he is caught, however, he will be incarcerated beginning immediately and perhaps for some time. Apprehension occurs with probability p. While incarcerated, he receives flow utility $a - c$, where c is a positive per-period utility cost of being incarcerated.[39]
- The duration of incarceration is given by the non-negative discrete random variable S, which has probability function $\{\pi_s\}_{s=1}^{\infty}$ and finite mean $\mathbb{E}[S]$.
- The individual knows the structure of the problem as outlined, the probabilities p and $\{\pi_s\}_{s=1}^{\infty}$, the distribution of criminal opportunities $F(\cdot)$, and the parameters a and c. He chooses to offend or abstain in each period he is free, seeking to maximize the expected present discounted value of utility, $E_t\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}u_\tau\right]$. Here, E_t is the expectation operator conditional on information available as of period

[38] In an earlier draft, we considered an alternative model in which there was heterogeneity in p, the probability of apprehension (Lee and McCrary 2005).

[39] Thus, criminal opportunities cannot be acted upon while incarcerated.

t, δ is the discount factor, and u_t is either $a-c$, a, or $a+B$, depending on the agent's choices, whether he has been apprehended for any crimes committed, and whether he is currently detained.

The Bellman equation for this problem is

$$V(b)=\max\left\{a+\delta\mathbb{E}[V(B)]\ ,\ p\sum_{s=1}^{\infty}\pi_s\left[(a-c)\frac{1-\delta^s}{1-\delta}+\delta^s\mathbb{E}[V(B)]\right]+(1-p)\left[a+b+\delta\mathbb{E}[V(B)]\right]\right\} \tag{2}$$

The first argument of the max function is the payoff from abstaining: receipt of the flow utility a and the expected continuation payoff. The second argument of the max function is the expected payoff from committing the offense. If caught and incarcerated for s periods, the payoff is $(a-c)(1+\delta+\delta^2+\cdots+\delta^{s-1})+\delta^s\mathbb{E}[V(B)]$. If not caught, the individual's payoff is the best possible: $a+b+\delta\mathbb{E}[V(B)]$.

The individual's optimal strategy is characterized by a "reservation benefit," b^*, such that when $B>b^*$, he commits the crime, and when $B<b^*$ he abstains.[40] Of course, when $B=b^*$, the individual is indifferent. Consequently, we may calculate the reservation benefit by equating the two arguments of the max function above. Doing so shows that

$$b^* \quad = \quad c\frac{p}{1-p}\left[1+\sum_{s=1}^{\infty}\pi_s\frac{\delta-\delta^s}{1-\delta}\left(1+\frac{(1-\delta)\mathbb{E}[V(B)]-a}{c}\right)\right] \tag{3}$$

We make the following observations about this expression:

- When $\delta=0$ and the individual is entirely present-focused, the expression reduces to $b^*=c\frac{p}{1-p}$. In that special case, crime will be committed whenever the ratio of the net benefit to net cost, B/c , exceeds the odds of apprehension, $\frac{p}{1-p}$. This is precisely the notion put forth in a standard (and static) Becker model of crime. It is intuitive that when the individual completely discounts the future, the length of incarceration, S, is irrelevant.
- Similarly, $b^*=c\frac{p}{1-p}$ if the punishment is only 1 period long (i.e., $\pi_s=0$ for $s=2,3,\ldots$).
- The second term inside the square braces thus captures the increased expected cost to offending due to a period of incarceration longer than 1 period. During incarceration, the offender bears the per-period cost c, but also is not able to exercise the option of committing offenses during this period. This lost value is summarized by the term $((1-\delta)\mathbb{E}[V(B)]-a)/c$.
- The Theory Appendix shows that the annuitized value of the dynamic program can be expressed as

$$(1-\delta)\mathbb{E}[V(B)] \quad = \quad a+(1-F(b^*))(1-p)\mathbb{E}[B-b^*|B>b^*] \tag{4}$$

[40] This mimics the standard reservation wage property of a job search model. See, for example, the textbook treatments in Adda and Cooper (2003) and Ljungqvist and Sargent (2004).

It is intuitive that the annuitized option value of the dynamic program, the second term in Equation (4), is the chances of encountering a worthwhile crime times the chances of avoiding apprehension times the expected benefit of those worthwhile crimes.

We obtain the following intuitive comparative statics (proof in Theory Appendix). The crime rate, which is given by $1-F(b^*)$, decreases with:

- Higher discount factor, δ. As the individual places more weight on future utility, the penalty of incarceration poses a higher cost.
- Higher per-period direct utility cost to incarceration, c.
- Higher apprehension rates, p.
- Longer sentence lengths. Any rightward shift in the distribution of the incarceration length S (i.e., the new distribution first-order stochastically dominates the old) decreases offending.

B. Benchmark Model Calibration and Predicted Values for θ: A Parametric Analysis

In this section, we calibrate the above model to produce predictions on the size of the reduced-form estimates of θ from our logit estimation. This exercise answers the question: "In this economic model of crime, how large of a θ would we have expected to observe?" In the calibration of our model, we let $a=0$ and $c=1$ and we take each period to be a single day.[41] Because each period is a day, the model's implied arrest probabilities for the boundary adult and juvenile, $p(1-F(b^*))$ and $p(1-F(b_J^*))$, respectively, where b_J^* is the juvenile analogue of b^*, correspond to our empirical estimate of θ as

$$\begin{aligned}\theta &= \ln\left(\frac{\mathbb{P}_{18.02}(\text{Arrest})}{1-\mathbb{P}_{18.02}(\text{Arrest})}\right)-\ln\left(\frac{\mathbb{P}_{17.98}(\text{Arrest})}{1-\mathbb{P}_{17.98}(\text{Arrest})}\right) \\ &= \ln\left(\frac{1-(1-p(1-F(b^*)))^7}{(1-p(1-F(b^*)))^7}\right)-\ln\left(\frac{1-(1-p(1-F(b_J^*)))^7}{(1-p(1-F(b_J^*)))^7}\right)\end{aligned} \tag{5}$$

where $\mathbb{P}_{17.98}(\text{Arrest})$ and $\mathbb{P}_{18.02}(\text{Arrest})$ are the predictions from our reduced-form empirical models for the probabilities of arrest one week prior to and after the 18th birthday, respectively. That is, the left-hand side of Equation (5) can be thought of as "measurement" and the right-hand side can be theory of as "theory."

To tighten this connection, we combine Equations (3) and (4) to obtain an implicit equation for b^*, and

[41] Setting $a=0$ and $c=1$ can be viewed as a normalization. Maximizing $\mathbb{E}_t[\sum_{\tau=t}^{\infty}\delta^{\tau-t}u_\tau]$ is equivalent to maximizing $\mathbb{E}_t[\sum_{\tau=t}^{\infty}\delta^{\tau-t}(u_\tau-a)/c]$. Thus, no matter the values of a and c, the solution will be equivalent to considering a problem where the flow utility takes the values of -1, 0, or B/c.

similarly obtain an implicit equation for b_J^*:

$$b^* = \frac{p}{1-p}\left[1+\sum_{s=1}^{\infty}\pi_s\frac{\delta-\delta^s}{1-\delta}\left\{1+(1-F(b^*))(1-p)\mathbb{E}[B-b^*|B>b^*]\right\}\right] \tag{6}$$

$$b_J^* = \frac{p}{1-p}\left[1+\sum_{s=1}^{\infty}\pi_s^J\frac{\delta-\delta^s}{1-\delta}\left\{1+(1-F(b^*))(1-p)\mathbb{E}[B-b^*|B>b^*]\right\}\right] \tag{7}$$

The sole difference in these two expressions is the distribution of sentence lengths. Intuitively, this is because the term $(1-F(b^*))(1-p)\mathbb{E}[B-b^*|B>b^*]$ is related to the expected continuation payoff and so pertains to adulthood for both the boundary juvenile and the boundary adult.

In our benchmark calibration, we (1) let the adult and juvenile sentence lengths S and S_J be distributed as a discretized exponential (geometric), with $\mathbb{E}[S]=207$ (days) and $\mathbb{E}[S_J]=63$, (2) let the probability of apprehension be $p=0.08$ and the daily discount factor be $\delta=0.95^{1/365}$, and (3) let the criminal opportunity, B, be distributed exponentially with mean parameter $\lambda=4.56$. With these magnitudes and parameterizations, Equations (6) and (7) are used to solve for b^* and b_J^*, respectively, which yields the daily arrest hazards $p(1-F(b^*))$ and $p(1-F(b_J^*))$, respectively. These are then used to produce a predicted θ using Equation (5).[42]

The magnitudes we use for these parameters have some empirical backing, as we now describe:

- Our model suggests that we can use the evidence from Figure 5 to estimate the gap $\mathbb{E}[S]-\mathbb{E}[S_J]$. Specifically, we compute the time between "first arrest" and "second arrest" for 17.98- and 18.02-year-olds. According to our model, that difference is an estimate of $\mathbb{E}[S]-\mathbb{E}[S_J]$.[43] Using the same data and similar approach as the RD analysis in Figure 5, we estimate a Weibull duration model, where the duration is a function of a cubic in age and a dummy variable for adulthood.[44] This yields estimated durations of approximately 962 and 818 for adults and juveniles, respectively, and a gap of $\mathbb{E}[S]-\mathbb{E}[S_J]=144$ days.
- We obtain $\mathbb{E}[S_J]=9.05\times 7\approx 63$ from Appendix Table 1, which reports average incarceration lengths from a completely different data source (see Data Appendix for discussion). This approach suggests $\mathbb{E}[S]=63+144=207$. We note that this different data source yields $\mathbb{E}[S]=46.48\times 7\approx 325$ days, which implies a difference $\mathbb{E}[S]-\mathbb{E}[S_J]$ of 262. We opt for a conservative calibration by choosing the smaller gap

[42] As noted in Section IV, $\ln\left(\frac{h}{1-h}\right)-\ln\left(\frac{h_J}{1-h_J}\right)\leq\theta<\ln(h)-\ln(h_J)<0$, where h and h_J are the true arrest hazards for a boundary adult and juvenile, respectively. In our calibrations, we report the smaller (in magnitude) of the two, $\ln(h)-\ln(h_J)$, but as noted there is very little difference between the bounds in practice.

[43] According to our model, the average duration between arrests as an adult is $\mathbb{E}[S]+1/(p(1-F(b^*)))$; that is, once an adult is arrested, the time until the next arrest is given by the period of incarceration S plus the time between being released and the next arrest, which is exponentially distributed with hazard $p(1-F(b^*))$ and thus mean $1/(p(1-F(b^*)))$. Similarly, the time until next arrest for a boundary juvenile is $\mathbb{E}[S_J]+1/(p(1-F(b^*)))$.

[44] Specifically, we model the juvenile and adult durations as separate Weibull models, allowing both the shape and scale parameters to depend on a cubic in age at initial arrest.

of 144 days. The predicted θ would be yet larger if we were instead to increase this gap above 144 days.

- Appendix Table 2 provides estimates of clearance rates and reporting rates for Index crimes in 2002. These numbers suggest $p=0.08$.[45]
- Finally, we choose $\lambda=4.56$ so that the implied b^* (cf., Equation (6)) is consistent with our estimate of the adult arrest hazard of $p(1-F(b^*))=0.0013$, which is computed as follows. Using the same Weibull estimates from above, we have $1/(p(1-F(b^*)))=962-\mathbb{E}[S]=962-207=755$. The implied daily arrest hazard is $p(1-F(b^*))=1/755$, or about 0.0013.

The results of our calculations for the benchmark calibration are presented in the first row of Table 5, which shows that this set of parameters would predict a discontinuity estimate of -2.76. As noted above, our standard errors are small enough to statistically rule out values of θ more negative than -0.11. Thus, in the context of the model outlined, a benchmark calibration suggests that the data are very much at odds with long time horizons. The rest of Table 5 reports our exploration of the sensitivity of the predicted θ to different parameter values for the model. We vary each parameter one at a time, while maintaining the values for the other parameters in the benchmark calibration in the top row.

Inspection of Table 5 reveals that while the predicted θ increases in magnitude with increasing p, from -2 to -3.74 as p ranges from 0.025 to 0.400, none of the predicted values for θ are close to our point estimate of -0.018. Indeed, none are close to the outer edge of the two-sided confidence region for θ, -0.11. Note that the extreme values of p that we consider, 0.025 and 0.400, are at odds with the evidence summarized in Appendix Table 2. The conclusion that our estimates are much smaller than the benchmark prediction seems robust to assumptions regarding p.

Table 5 also presents estimates based on shifts to the distribution of sentence lengths. As noted, our benchmark calibration assumes sentences are distributed as a discretized exponential (geometric). The exponential distribution is a special case of the Weibull distribution, with the shape parameter of the Weibull equal to 1. To explore the sensitivity of the predicted θ to the assumed shape of the distribution, we vary the shape parameter for the juvenile and adult sentence length distribution over 0.25, 0.40, 2.0, and 4.0. For each such shape parameter, k_J and k_A, we adjust the Weibull scale parameters to match the means $\mathbb{E}[S_J]=63$ and $\mathbb{E}[S]=207$. To gain intuition about these changes, note that when the shape

[45] Clearance rates are the fraction of offenses known to police for which an individual has been arrested and handed over to prosecutors. These rates are estimated using the UCR data. Reporting rates are the fraction of Index offenses reported to police and are based on the NCVS. The aggregate clearance rate is 20 percent, and the aggregate reporting rate is 42 percent, for an approximate $0.2\times0.42=0.08$ probability of apprehension.

parameter is 0.25, the density is more convex and skewed, while when the shape parameter is 4, the shape of the distribution is more akin to that of the chi-square distribution. The table shows that the shape of the sentence length distribution affects our predictions only negligibly. For example, with the shape parameter for the juvenile distribution, k_J , equal to 0.25, the predicted θ is -2.95 and with a shape parameter of 4.0, the predicted θ is -2.76. Results for the impact of the adult shape parameter, k_A, are similar.[46]

The final row of Table 5 shows the effect on the predicted θ of changes to the discount factor. When the discount factor declines from 0.95 to 0.01 on an annual basis, the magnitude of the predicted θ falls from -2.50 to -1.57. This is a notable pattern, because—unlike the case of p, $\mathbb{E}[S]$, and $\mathbb{E}[S_J]$—we have no independent information on the discount factor. The relevant discount factor is not that of a typical individual, but that of a marginal offender—namely, those individuals for whom the economic and social environment is such that offending is contemplated but not assured—and thus could be much smaller than 0.95. We note that in some settings, such as the market for subprime loans, the literature has documented behaviors that are consistent with extremely small discount factors (e.g., Adams, Einav and Levin 2009).[47] Perhaps most importantly, discount factors could be much smaller when arrestees are drug-users, a common pattern. For example, urinanalysis evidence indicates that a large fraction of arrestees test positive for drug use. In 2000, 61.8 percent of Fort Lauderdale and 62.8 percent of Miami arrestees tested positive for one or more of the following: cocaine, marijuana, opiates, methamphetamine, or PCP (National Institute of Justice 2003). These considerations suggest the prevalence of short time horizons and low discount factors among the relevant subpopulation. In our model, as is intuitive, θ is small when the discount factor is. Present-oriented individuals may respond little to threatened future punishments.

Finally, we consider the impact of changes in the shape of the criminal benefit distribution, the other element of the model for which we have no independent information. In the benchmark calibration, this distribution was assumed to be exponential. Generalizing to a Weibull distribution, we vary the shape parameter, k, over 0.25, 0.40, 2.0, and 4.0. When $k=4.0$, many individuals are on the margin of the crime participation decision, but when $k=0.25$, few individuals are on the margin. We see that as the shape parameter falls from 4.0 to 0.25, the predicted θ falls in magnitude from -3.84 to -1.03.

We conclude from these analyses that a reasonable benchmark parametric model with standard "patient"

[46] This robustness to distributional assumptions on the sentencing side is due to the fact that the distribution of sentences only matters to the extent that it affects $\sum_{s=1}^{\infty} \pi_s \left(\delta - \delta^S\right)/(1-\delta)]$.

[47] As another example, Caskey (1996) shows that in states without interest rate regulations, pawnbrokers charge interest rates implying annual discount factors of 0.3.

discount factors predicts much larger discontinuities in offending at age 18 than we observe in our data. However, given that the predictions seem somewhat sensitive to the discount factor and the shape of the benefit distribution, when we consider the policy implications of our estimates below, we relax the parametric assumptions on the benefit distribution, and present a bounding analysis.

C. Connecting θ to Broader Policy Elasticities

In the introduction, we outlined a set of specific policies that are addressed by our reduced-form estimates above (e.g., changing the age of criminal majority). However, much of the literature focuses on broader policy parameters, such as elasticities of crime with respect to policing and expected sentence lengths. Intuitively, our reduced-form estimates are not *un*related to these policy parameters, but the nature of the connection is vague, imprecise, and qualitative. The feature of a behavioral model is that it makes these connections specific, precise, and even quantitative. In this subsection, we connect our empirical work to these broader policy parameters, with a particular focus on minimizing the assumptions necessary to make that connection.

In elasticity terms, the responsiveness of the crime rate to sentence length and police increases under our model can be shown to be

$$\begin{aligned} \eta_{\mathbb{E}[S]} &\equiv \frac{\partial(1-F(b^*)}{\partial \mathbb{E}[S]}\frac{\mathbb{E}[S]}{1-F(b^*)} = -\frac{pf(b^*)}{h}\mathbb{E}[S]\frac{\partial b^*}{\partial \mathbb{E}[S]} = -\left[\left(\frac{h_J-h}{h}\right)/\left(\frac{\nu-\nu_J}{\nu}\right)\right]\eta_\nu\left(\frac{1}{1+\nu h}\right)\frac{f(b^*)}{f(\bar{b})} \quad (8) \\ \eta_p &\equiv \frac{\partial(1-F(b^*)}{\partial p}\frac{p}{1-F(b^*)} = -\frac{pf(b^*)}{h}p\frac{\partial b^*}{\partial p} = -\left[\left(\frac{h_J-h}{h}\right)/\left(\frac{\nu-\nu_J}{\nu}\right)\right]\kappa\left(\frac{1}{1+\nu h}\right)\frac{f(b^*)}{f(\bar{b})} \quad (9) \end{aligned}$$

respectively, where $h=p(1-F(b^*))$, $h_J=p(1-F(b_J^*))$, $\nu=\sum_{s=1}^{\infty}\pi_s\frac{\delta-\delta^s}{1-\delta}$, $\nu_J=\sum_{s=1}^{\infty}\pi_s^J\frac{\delta-\delta^s}{1-\delta}$, $\eta_\nu\equiv\frac{\partial\nu}{\partial\mathbb{E}[S]}\frac{\mathbb{E}[S]}{\nu}$, $\kappa\equiv\left(c\frac{p}{(1-p)^2}(1+\nu)+\nu h\mathbb{E}[B-b^*|B>b^*]\right)/\left(c\frac{p}{1-p}\nu+\nu h\mathbb{E}[B-b^*|B>b^*]\right)$, and $\bar{b}\in(b_J^*,b^*)$ is such that $f(\bar{b})=\frac{F(b^*)-F(b_J^*)}{b^*-b_J^*}$ (see Theory Appendix for detailed derivation).

Since $\theta\approx-(h_J-h)/h$, Equations (8) and (9) show that our reduced-form discontinuity estimate is proportional to the broader policy elasticities $\eta_{\mathbb{E}[S]}$ and η_p. This is then the result that confirms the intuition that our reduced-form estimates are relevant to broader policy parameters. The additional factors in Equations (8) and (9) sharpen this connection and raise the possibility of quantifying the distinction between our reduced-form estimates and these broader policy parameters. We now provide discussion regarding these additional factors.

The second factor of Equations (8) and (9), $(\nu-\nu_J)/\nu$, gives the percent increase in the expected discounted number of periods of incarceration when crossing the age 18 threshold. This term reflects the influence of δ as well as the change of punishment regime from S_J to S. For patient individuals, $(\nu-\nu_J)/\nu$ is approximately equal to $(\mathbb{E}[S]-\mathbb{E}[S_J])/\mathbb{E}[S]$, but for present-oriented individuals it may be much smaller.[48]

[48] A slightly more accurate approximation is $(\mathbb{E}[S]-\mathbb{E}[S_J])/(1+\mathbb{E}[S])$ but since we measure time in days, this distinction

The third factors are unique to each elasticity. In the case of $\eta_{\mathbb{E}[S]}$, the third factor η_ν reflects the percentage change in ν brought on by a percentage change in the expected number of periods of incarceration. In the case of η_p, the third factor κ is a quantity that reflects the percentage change in p that has the equivalent impact as a percentage change in ν.

The fourth factor is an offsetting adjustment that accounts for the following effect: when the expected costs of offending rise, the option value of being free is reduced, which in turn dampens the per-period cost of incarceration.[49] Finally, the fifth term reflects the shape of the benefit distribution: $f(b^*)/f(\bar{b})=1$ if the density of B is flat between b_J^* and b^*, but more generally this ratio may differ from 1.

We can calibrate our benchmark model to ask what policy elasticities our estimates imply. For example, we can adjust the discount factor so that the predicted θ matches -0.11, the outer edge of our confidence interval. This yields a daily discount factor of 0.71, which predicts elasticities of $\eta_{\mathbb{E}[S]} = -0.049$ and $\eta_p = -4.44$. We could alternatively have kept δ fixed and attempted to match the predicted θ to the observed θ by adjusting the shape parameter k of the benefit distribution. Since it is arbitrary which parameter to adjust, in the next Subsection, we consider a wide range of discount factors, while relaxing the parametric restriction on the benefit distribution.

D. Nonparametric Bounds on Policy Elasticities

We now explore the range of policy elasticities that are consistent with our discontinuity estimates, relaxing the parametric restriction on the shape of the criminal benefit distribution.

We first note that it is always possible to *construct* a distribution for B that is both consistent with any negative value of θ, yet yields policy elasticities of zero or $-\infty$. One needs to simply set the density $f(b^*)$ to zero or arbitrarily large. This fact is not unique to our dynamic model: this will be the case with any model that has a distribution of criminal benefits and a threshold above which crime occurs.

It is thus more informative and plausible to consider some mild restrictions on the benefit distribution. We focus on the assumption that the density of benefits is weakly declining, or that more valuable opportunities are more rare. This assumption has some economic and conceptual justification, going back at least to Stigler (1970) and Viscusi (1986). Intuitively, as these authors have argued, opportunities that are valuable to the offender are typically costly for the victim. This provides incentives for victims to take greater precaution regarding more valuable criminal opportunities (e.g., deadbolt locks, car alarms, and mace).

is of little practical difference.

[49]See Theory Appendix for discussion.

It is thus plausible that more valuable opportunities are scarce, relative to less valuable opportunities.

In the Theory Appendix, we show that the nonparametric assumption of a declining density delivers an upper bound on the magnitude of the policy elasticities $\eta_{\mathbb{E}[S]}$ and η_p. In particular, we show that

$$\eta_{\mathbb{E}[S]} > \overline{\eta}_{\mathbb{E}[S]} \equiv -\left[\left(\frac{h_J-h}{h}\right)/\left(\frac{\nu-\nu_J}{\nu}\right)\right]\eta_\nu\left(\frac{1}{1+\nu h}\right) \tag{10}$$

$$\eta_p > \overline{\eta}_p \equiv -\left[\left(\frac{h_J-h}{h}\right)/\left(\frac{\nu-\nu_J}{\nu}\right)\right]\overline{\kappa}\left(\frac{1}{1+\nu h}\right) \tag{11}$$

where $\overline{\kappa}=\left(c\frac{p}{(1-p)^2}(1+\nu)+\nu h\underline{E}\right)/\left(c\frac{p}{1-p}\nu+\nu h\underline{E}\right)$ and $\underline{E}$ is given in the Theory Appendix.[50] Both $\overline{\eta}_{\mathbb{E}[S]}$ and $\overline{\eta}_p$ can be computed using information on discount factors, the distribution of sentence lengths, and our empirical findings on arrest rates for youthful offenders at age 17.98 and 18.02. Other than assuming a declining density, no further information regarding the distribution of criminal benefits is required to obtain these bounds; this is the sense in which this might be described as a nonparametric restriction.

Table 6 reports our estimates of the elasticity bounds $\overline{\eta}_{\mathbb{E}[S]}$ and $\overline{\eta}_p$ for our point estimate of $\theta=-0.018$ and for the edge of our (two-sided) 95 percent confidence interval, $\theta=-0.11$, using various values of the discount factor. Using the most negative value of $\theta=-0.11$, the sentencing elasticity ranges from -0.130 to -0.060, and the policing elasticity ranges from very large negative values to -0.136. If we use our point estimate, the sentencing elasticity $\overline{\eta}_{\mathbb{E}[S]}$ varies only slightly, from -0.013 to -0.010, while the probability of apprehension elasticity η_p varies from -0.117 to -0.044.

Interestingly, for the case of $\theta=-0.018$, there are some values of δ that cannot be rationalized within the framework of the model. That is, as we show in the Theory Appendix, we find that if $\theta=-0.018$, then there exists no $\mathbb{E}[B-b^*|B>b^*]$ that is consistent with both the model and the restriction of a declining density.

Overall, the table points to two policy conclusions. First, if one believes that the probability of apprehension elasticity is large, then this implies very low discount factors, and consequently very low sentence length elasticities. Second, across the entire range of permissible discount factors, all of the sentence length elasticities are generally small in magnitude.

VII. Conclusion

Over the past 40 years, the incarceration rate in the United States has soared, from 93 persons in prisons per 100,000 in 1972 to 480 persons per 100,000 in 2012. Indeed, despite the declining crime rates of recent

[50] As we show in the Theory Appendix, our model and a declining density for B implies that $\mathbb{E}[B-b^*|B>b^*]$ can be no lower than $\underline{E}$, which is strictly positive as long as adults commit crime. See Appendix for the formula for $\underline{E}$.

years, prison and jail populations continue to climb, as the full impact of sentencing reforms from previous years becomes felt (Raphael and Stoll 2009).[51]

For economists, it is natural to wonder if more severe prison sentences deter crime by a sufficient amount to make paying for them worthwhile. Furthermore, can a lower level of crime be attained by an expenditure-neutral reallocation of funds from prisons to alternative uses, such as social programs (Donohue and Siegelman 1998) or policing (McCrary 2010, Chalfin and McCrary 2016)? These questions have recently acquired a renewed policy relevance, as state and local governments scramble to lower costs in the wake of the financial crisis.

A key input to these considerations is the magnitude of the deterrence effect of prison sentences. In this paper, we use a quasi-experimental approach to identify the deterrence effect of adult criminal sanctions, relative to juvenile criminal sanctions. Standard economic models of crime imply that participation in crime should drop discontinuously at 18, when sanctions become more punitive. Our central empirical finding is that the rate of criminal involvement of young offenders is generally a smooth function of age, with only a small change at 18. Our findings are based on a high frequency, longitudinal administrative data set on arrests in Florida.

Our focus on the pattern of criminal involvement around 18 years of age yields two key benefits. First, under mild smoothness assumptions, our estimated effects are unlikely to be tainted by omitted variables bias. Second, our estimated effect reflects purely a deterrence effect, rather than a conflation of deterrence and incapacitation—which may well be contributing to the larger estimated deterrence elasticities in the some of the existing literature. We document that incapacitation effects of prison are large and nearly immediate, at least for youthful offenders in our data.

The point estimates from our discontinuity analysis indicate an approximately 2 percent decline in the rate of criminal offending when a juvenile turns eighteen, when the expected incarceration length conditional on arrest jumps discontinuously by roughly 230 percent. This suggests a small "reduced-form" elasticity of -0.007.[52] This elasticity is directly relevant to three types of criminal justice policy reforms that impact youthful offenders when they cross the age of criminal majority: (1) reducing the age of criminal majority, (2) increasing the rate at which juveniles are tried in the adult criminal court, and (3) marginal increases in the juvenile-adult gap in punishments. These kinds of policy changes are not simply hypothetical, but are instead relevant to recent and ongoing criminal justice discussions: (1) four states have reduced the

[51] Since 1993, the most recent peak of criminal activity, through 2012, violent crime had fallen by 37 percent and property crime had fallen by 26 percent, whereas the prison population had increased 34 percent. See *Sourcebook of Criminal Justice Statistics Online*, Tables 3.106.2012, 6.1.2011, and 6.28.2012, and *Sourcebook of Criminal Justice Statistics*, 1975, Table 1.131.

[52] When we consider the smallest negative effects consistent with our (two-sided) confidence interval, -0.11, this elasticity is instead -0.047.

age of majority in recent years (Snyder and Sickmund 2006, Chalfin and Loeffler 2016), (2) all 50 states and the District of Columbia had by 1995 adopted provisions for juvenile transfer to the adult criminal court (Government Accounting Office 1995), and (3) during the 1980s and 1990s, many states increased the punitiveness of the sentencing regime for adults by much more than that for juveniles (Levitt 1998).[53] Our estimates imply that changes to these types of policy changes may affect the crime rate of youthful offenders through incapacitation channels, but have at best minor effects on deterrence.

Investigating the implications of our estimates for broader policies of interest requires imposing some structural assumptions on criminal behavior. Combining ideas from Becker (1968) and McCall (1970), we develop a stochastic dynamic model of crime and calibrate it to match key empirical quantities. We recognize that our forward-looking, rational-expectations, rational-agent model may indeed be too restrictive. For example, although aggregate statistics give us a reasonably objective estimate of both the average probability of apprehension and expected incarceration lengths, potential offenders may not accurately perceive, and could conceivably vastly underestimate, those risks and punishments (Lochner 2007, Hjalmarsson 2009a). Furthermore, it may well be that criminal offending is better described by a hyperbolic discounting model, or some alternative model of time preferences, rather than time-consistent exponential discounting.[54]

Nevertheless, with the caveat that our theoretical framework is one particular lens through which to view the results, we draw upon it to reach three key conclusions. First, the magnitudes of our point estimates are consistent with impatient or even myopic behavior on the part of criminal offenders. For example, combining our point estimate with the mild assumption of a decreasing density function of criminal benefits rules out annual discount factors of 0.022 or more.[55]

Second, within the range of discount factors consistent with both our empirical findings and our theoretical model, our predicted elasticities can be reconciled with previous estimates of the elasticity of crime with respect to the probability of apprehension (e.g., Evans and Owens (2007) or Klick and Tabarrok (2005)). Interestingly, the model indicates that large policing elasticities, such as the -0.75 estimate of Di Tella and Schargrodsky (2004), imply a very small discount factor—so small that it would imply an elasticity with respect to sentence lengths no more negative than -0.06. Generally, there is a trade-off between these two elasticities, if one stipulates to both our model and our empirical results.

[53]For further background on these institutional changes, see National Research Council (2001).

[54]For exploration of models of this type in this context, see Lee and McCrary (2005).

[55]Note, however, that the outer edge of our (two-sided) confidence region does not allow us to rule out any particular discount factor.

Third, no matter the discount factor, the most negative sentence length elasticity consistent with our data and model is -0.13. This is several times smaller than the larger elasticities from the literature, but is consistent with the magnitudes from Helland and Tabarrok (2007) and Iyengar (2008).

We conclude by noting that while our findings point to small deterrence effects of prison, this does not mean that prison is ineffective in reducing the overall incidence of crime. Indeed, we show that once arrested, the marginal juvenile transferred from the juvenile to adult criminal justice system is less likely to recidivate. We suspect that this is occurring through an incapacitation mechanism, which could be quite important, as suggested by the work of Jacob and Lefgren (2003). A fruitful avenue for future research would thus be an investigation of optimal sentencing structure and the optimal mix of police and prisons when incapacitation is the primary mechanism for crime reduction.

References

Abrams, David S. and Chris Rohlfs, "Optimal Bail and the Value of Freedom: Evidence from the Philadelphia Bail Experiment," *Economic Inquiry*, 2011, *49* (3), 750–770.

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, January 2009, *99* (1), 49–84.

Adda, Jerome and Russell W. Cooper, *Dynamic Economics: Quantitative Methods and Applications*, Cambridge: MIT Press, 2003.

Angrist, Joshua and Jörn-Steffen Pischke, *Mostly Harmless Econometrics*, Princeton: Princeton University Press, 2009.

Ayres, Ian and John J. Donohue, "The Latest Misfires in Support of the More Guns, Less Crime Hypothesis," *Stanford Law Review*, 2003, *55*, 1371.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen, "Building Criminal Capital Behind Bars: Peer Effects in in Juvenile Corrections," *Quarterly Journal of Economics*, February 2009, *124* (1).

Beccaria, Cesare, *On Crimes and Punishments*, Oxford: Clarendon Press, 1764.

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

Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation*, Oxford: Clarendon Press, 1789.

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.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

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.

Caskey, John P., *Fringe Banking*, New York: Russell Sage, 1996.

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

Chalfin, Aaron and Charles Loeffler, "Estimating the Crime Effects of Raising the Age of Majority: Evidence from Connecticut," *Criminology and Public Policy*, Forthcoming 2016.

——— **and Justin McCrary**, "Are U.S. Cities Underpoliced?: Theory and Evidence," *Review of Economics and Statistics*, Forthcoming 2016.

Cook, Philip J., "Research in Criminal Deterrence: Laying the Groundwork for the Second Decade," in Michael Tonry, ed., *Crime and Justice: An Annual Review of Research*, Vol. 2, Chicago: University of Chicago Press, 1980, pp. 211–268.

——— **and Jens Ludwig**, *Gun Violence: The Real Costs*, New York: Oxford University Press, 2000.

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 Justin Wolfers, "Uses and Abuses of Empirical Data in the Death Penalty Debate," in Steven D. Levitt and Thomas J. Miles, eds., *Economics of Criminal Law*, Northampton, MA: Edward Elgar, 2008.

——— **and Peter Siegelman**, "Allocating Resources among Prisons and Social Programs in the Battle against Crime," *Journal of Legal Studies*, January 1998, *27* (1), 1–43.

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

Drago, Francesco, Roberto Galbiati, and Pietro Vertova, "The Deterrent Effects of Prison: Evidence from a Natural Experiment," *Journal of Political Economy*, 2009, *117* (2), 257–280.

Duggan, Mark, "More Guns, More Crime," *Journal of Political Economy*, October 2001, *109* (5), 1086–1114.

Eeckhout, Jan, Nicola Persico, and Petra Todd, "A Theory of Optimal Random Crackdowns," *American Economic Review*, June 2010, *100* (3).

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.

——, "On the Issue of Causality in the Economic Model of Crime and Law Enforcement: Some Theoretical Considerations and Experimental Evidence," *American Economic Review*, May 1987, *77* (2), 99–106.

Evans, William N. and Emily G. Owens, "COPS and Crime," *Journal of Public Economics*, 2007, *91* (1), 181–201.

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

Flinn, Christopher, "Dynamic Models of Criminal Careers," in Alfred Blumstein, Jacqueline Cohen, Jeffrey A. Roth, and Christy A. Visher, eds., *Criminal Careers and "Career Criminals"*, Vol. 2, Washington, D.C.: National Academy Press, 1986, pp. 356–379.

Freeman, Richard B., "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*, 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 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.

Government Accounting Office, "Juvenile Justice: Juveniles Processed in Criminal Court and Case Dispositions," *Washington, DC: U.S. General Accounting Office*, August 1995.

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.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, February 2001, *69* (1), 201–209.

Helland, Eric and Alex Tabarrok, "Does Three Strikes Deter: A Non-Parametric Investigation," *Journal of Human Resources*, Spring 2007, *42* (2), 309–330.

Hjalmarsson, Randi, "Crime and Expected Punishment: Changes in Perceptions at the Age of Criminal Majority," *American Law and Economics Review*, 2009, *11* (1).

——, "Juvenile Jails: A Path to the Straight and Narrow or Hardened Criminality?," *Journal of Law and Economics*, November 2009, *52* (4).

Iyengar, Radha, "I'd Rather Be Hanged for a Sheep Than a Lamb: The Unintended Consequences of 'Three-Strikes' Laws," *NBER Working Paper #13784*, February 2008.

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.

Klick, Jonathan and Alex Tabarrok, "Using Terror Alert Levels to Estimate the Effect of Police on Crime," *Journal of Law and Economics*, April 2005, *48* (1), 267–280.

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

——, 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.

Kuziemko, Ilyana, "Going Off Parole: How the Elimination of Discretionary Prison Release Affects the Social Cost of Crime," September 2007. NBER Working Paper #13380.

LEAA, U.S., *Expenditure and Employment Data for the Criminal Justice System: 1969-1970*, Washington, D.C.:

U.S. Government Printing Office, 1972.

Lee, David S., "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, February 2008, *142* (2).

____ **and Justin McCrary**, "Crime, Punishment, and Myopia," *NBER Working Paper # 11491*, July 2005.

____ **and Thomas Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

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.

____ , "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.

____ **and Thomas J. Miles**, "Empirical Study of Criminal Punishment," in A. Mitchell Polinsky and Steve Shavell, eds., *Handbook of Law and Economics*, Vol. 1, North-Holland, 2007, chapter 7, pp. 455–498.

Ljungqvist, Lars and Thomas J. Sargent, *Recursive Macroeconomic Theory*, Cambridge: MIT Press, 2004.

Lochner, Lance, "Education, Work, and Crime: A Human Capital Approach," *International Economic Review*, August 2004, *45* (3), 811–843.

____ , "Individual Perceptions of the Criminal Justice System," *American Economic Review*, March 2007, *97* (1), 444–460.

____ **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.

Lott, John, *More Guns, Less Crime*, Chicago: University of Chicago Press, 1998.

Maguire, Kathleen and Ann L. Pastore, eds., *Sourcebook of Criminal Justice Statistics, 2000*, Washington, D.C.: U.S. GPO, 2001.

McCall, John J., "Economics of Information and Job Search," *Quarterly Journal of Economics*, 1970, *84* (1), 113–126.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, February 2008, *142* (2).

____ , "Dynamic Perspectives on Crime," in Bruce Benson and Paul R. Zimmerman, eds., *Handbook on the Economics of Crime*, Northampton, MA: Edward Elgar, 2010.

____ **and Heather N. Royer**, "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth," *American Economic Review*, February 2011.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," August 2015. Unpublished manuscript, University of Michigan.

Nagin, Daniel, "General Deterrence: A Review of the Empirical Evidence," in Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, eds., *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Washington, D.C.: National Academy of Sciences, 1978, pp. 95–139.

____ , "Criminal Deterrence Research at the Outset of the Twenty-First Century," in Michael Tonry, ed., *Crime and Justice: A Review of Research*, Vol. 23, Chicago: University of Chicago Press, 1998, pp. 1–42.

National Institute of Justice, "2000 Arrestee Drug Abuse Monitoring: Annual Report," April 2003.

National Research Council, *Juvenile Crime, Juvenile Justice*, Washington, D.C.: National Academy Press, 2001.

Parsell, T.J., "In Prison, Teenagers Become Prey," *New York Times*, June 5, 2012.

Polinsky, A. Mitchell and Steven Shavell, "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, March 2000, *38* (1), 45–76.

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

Sickles, Robin C. and Jenny Williams, "Turning from Crime: A Dynamic Perspective," *Journal of Econometrics*, 2008, *145* (1-2), 158–173.

Silverman, Dan, "Street Crime and Street Culture," *International Economic Review*, August 2004, *45* (3), 761–786.

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

____ **and** ____ , *Juvenile Offenders and Victims: 2006 National Report*, Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention, 2006.

Stigler, George J., "The Optimum Enforcement of Laws," *Journal of Political Economy*, May/June 1970, *78* (3), 526–536.

U.S. DOJ, "Justice Expenditure and Employment Statistical Extracts, 2006," *accessed online*, http://www.ojp.usdoj.gov/bjs/eande.htm 2006.

Viscusi, W. Kip, "The Risks and Rewards of Criminal Activity: A Comprehensive Test of Criminal Deterrence," *Journal of Labor Economics*, July 1986, *4* (3), 317–340.

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

White, Halbert, "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 1982, *50* (1), 1–25.

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

A. Data Appendix

1. Arrest-level Database

Our data set is constructed using four electronic files maintained by the 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.

i. Construction of Data Set

We construct our data set as follows. First, we begin with the arrests file, which contains a person identifier,[56] the offense date, the arrest date, the charge code, and the arresting agency. Each record in the arrest file pertains to a separate offense.[57] 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, using the FBI hierarchy (Federal Bureau of Investigation 2004, p. 10) There are 3,314,851 unique arrest-person observations.[58]

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. We use the race variable in the identifier file and birthday from the birth date file.

ii. 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, 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.

[56] This identifier is constant across the FDLE files.

[57] 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.

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

2. Average Incarceration Lengths: $\mathbb{E}[S]$ and $\mathbb{E}[S_J]$

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.

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 1 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 courtproceedings.[59] 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 the prison populations of 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, respectively) 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 penultimate 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.

The final column of Appendix Table 1 adjusts the figures in the penultimate column for juvenile transfer, assuming that 20 percent of juveniles are tried as juveniles, 50 percent are tried as adults, and that the remainder have no charges brought against them (cf., Figure 2). To understand the adjustment calculations we use, we introduce some new notation. We observe N_A individuals in adult prisons and jails, N_J individuals in juvenile facilities, M_A adult arrestees, and M_J juvenile arrestees. The only people in juvenile facilities are juveniles, but some of the individuals in adult prisons and jails are adults, and some are juveniles.

Decompose

$$N_A = N_A^A + N_A^J \tag{A.1}$$

where N_A^A is the number of adults in adult facilities and N_A^J is the number of juveniles in adult facilities, both of which are unobserved.

Ideally, we would approximate expected sentences for adults and juveniles as

$$\widehat{E}[S] = \frac{P_A}{M_A} \tag{A.2}$$

$$\widehat{E}[S_J] = \frac{P_J}{M_J} \tag{A.3}$$

where $P_A = N_A^A$ is the number of adults incarcerated and $P_J = N_A^J + N_J$ is the number of juveniles incarcerated. Because N_A^A and N_A^J are unobserved, so are P_A and P_J.

To back out P_A and P_J, let the number of juveniles in adult facilities, relative to the number of

[59] 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 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.

incarcerated juveniles, be denoted by

$$\alpha \;=\; \frac{N_A^J}{P_J} \tag{A.4}$$

For example, if 20 percent of juveniles are tried as adults, and 50 percent are tried as juveniles, and if juveniles tried as adults receive adult sentences, then we can approximate $\alpha = 20/(20+50) \doteq 0.286$.

With α in hand, we adjust the observed quantities N_J and N_A to estimate N_A^J and P_J. Specifically, since

$$\frac{1}{1-\alpha} \;=\; \frac{P_J}{N_J} \tag{A.5}$$

$$\frac{\alpha}{1-\alpha} \;=\; \frac{N_A^J}{N_J} \tag{A.6}$$

we obtain

$$\widehat{E}[S] \;=\; \frac{N_A - N_J \frac{\alpha}{1-\alpha}}{M_A} \tag{A.7}$$

$$\widehat{E}[S_J] \;=\; \frac{N_J \frac{1}{1-\alpha}}{M_J} \tag{A.8}$$

These are the quantities reported in the final column of Appendix Table 1.

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.

3. Probability of Arrest

In our model, p is the expected probability of being apprehended, conditional on committing a crime. We provide an 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 the arrest of an alleged offender who is then charged with the offense.

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. This is likely an upwardly biased estimate of p, because not all criminal incidents are reported to the police, and therefore the denominator of Column (2) is likely 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 adjusted 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).

B. Theory Appendix

1. Expressions for $\mathbb{E}[V(B)]$, b^*, Comparative Statics, and ν

In this subsection, we derive $\mathbb{E}[V(B)]$, define and then characterize b^*, and provide some results for ν.[60] First, to develop $\mathbb{E}[V(B)]$, note that for opportunities $b<b^*$, the value is fixed at $V(b)=a+\delta\mathbb{E}[V(B)]$. For opportunities $b>b^*$, higher value opportunities are worth more: $V(b)$ is linear in b with slope $1-p$. These two segments of $V(b)$ connect at $b=b^*$. We thus have

$$V(b)=a+\delta\mathbb{E}[V(B)]+(1-p)(b-b^*)\mathbf{1}(b>b^*) \tag{B.1}$$

Taking expectations with respect to the distribution of B and rearranging yields

$$(1-\delta)\mathbb{E}[V(b)] \quad = \quad a+(1-F(b^*))(1-p)\mathbb{E}[B-b^*|B>b^*] \tag{B.2}$$

Second, to obtain b^*, substitute Equation (B.2) into Equation (3) from the main text. Doing so shows that b^* solves

$$G(p,\delta,c,b^*)\equiv c\frac{p}{1-p}\left[1+\nu\left(1+\frac{1-p}{c}\psi(b^*)\right)\right]-b^*=0 \tag{B.3}$$

where $\psi(b)=(1-F(b))\mathbb{E}[B-b|B>b]$ and we recall from the main text that $\nu=\sum_{s=1}^{\infty}\pi_s\frac{\delta-\delta^s}{1-\delta}$.

Third, to characterize how b^* changes with changes to δ, c, p, and $\mathbb{E}[S]$, we use the implicit function theorem. Specifically:

- $\frac{\partial b^*}{\partial \delta}=-\frac{\partial G/\partial\delta}{\partial G/\partial b^*}>0$, because $\frac{\partial G}{\partial \delta}=c\frac{p}{1-p}\frac{\partial \nu}{\partial \delta}\left(1+\frac{1-p}{c}\psi(b^*)\right)$ is positive and because $\frac{\partial G}{\partial b^*}=-p\nu(1-F(b^*))-1$ is negative. To obtain the sign for $\frac{\partial \nu}{\partial \delta}$, rewrite $\nu=\sum_{s=1}^{\infty}\pi_s(\delta+\dots+\delta^{s-1})$ and the result follows on inspection. To obtain the result for $\partial G/\partial b^*$, note that integration by parts shows $\psi(b)=\int_b^{\infty}(1-F(t))dt$. The Leibniz integral rule then shows that $\psi'(b)=-(1-F(b))$.
- $\frac{\partial b^*}{\partial c}=-\frac{\partial G/\partial c}{\partial G/\partial b^*}>0$ because as shown above $\frac{\partial G}{\partial b^*}$ is negative, and because $\frac{\partial G}{\partial c}=\frac{p}{1-p}(1+\nu)$ is positive.
- $\frac{\partial b^*}{\partial p}=-\frac{\partial G/\partial p}{\partial G/\partial b^*}>0$ because $\frac{\partial G}{\partial b^*}$ is negative, and because $\frac{\partial G}{\partial p}=c\frac{1}{(1-p)^2}(1+\nu)+\nu\psi(b^*)$ is positive.
- $\frac{\partial b^*}{\partial \mathbb{E}[S]}=\frac{\partial b^*}{\partial \nu}\frac{\partial \nu}{\partial \mathbb{E}[S]}$ because changes to $\mathbb{E}[S]$ affect b^* only through ν. Then note that $\frac{\partial b^*}{\partial \nu}=-\frac{\partial G/\partial\nu}{\partial G/\partial b^*}>0$ because $\frac{\partial G}{\partial b^*}$ is negative and $\frac{\partial G}{\partial \nu}=c\frac{p}{1-p}+p\psi(b^*)$ is positive. Consequently, $\frac{\partial b^*}{\partial \mathbb{E}[S]}$ is positive since $\frac{\partial \nu}{\partial \mathbb{E}[S]}$ is positive. This follows for any first-order stochastically dominating rightward shift in the distribution of S, which results in an increase in $\mathbb{E}[S]$ and an increase in ν (since $\frac{\delta-\delta^s}{1-\delta}$ is strictly increasing in s).

Fourth, to develop results for ν, observe that it is a shifted geometric series and thus both absolutely and uniformly convergent, allowing rearrangements and differentiation, respectively. Rearrangement then implies $\nu=\pi_1\times 0+\pi_2\times\delta+\pi_3\times(\delta+\delta^2)+\dots=\sum_{s=1}^{\infty}\delta^s\mathbb{P}(S>s)$.[61] Differentiation then provides confirmation that $\partial\nu/\partial\delta>0$ because the derivative can be written as $\partial\nu/\partial\delta=\sum_{s=1}^{\infty}s\delta^{s-1}\mathbb{P}(S>s)>0$.

2. Elasticities of Crime with Respect to p and $\mathbb{E}[S]$

In this subsection, we give expressions for the policing elasticity and the sentencing elasticity. In addition to the notation above, define the expected surplus associated with crimes worth doing, $E\equiv\mathbb{E}[B-b^*|B>b^*]$, the elasticity $\eta_\nu\equiv\frac{\partial\nu}{\partial\mathbb{E}[S]}\frac{\mathbb{E}[S]}{\nu}$, and the adult and juvenile arrest rates, $h\equiv p(1-F(b^*))$ and $h_J\equiv p(1-F(b_J^*))$, respectively, where b_J^* is defined analogously to Equation (B.3) with ν_J replacing ν.[62]

[60] See McCrary (2010) for details on some of the calculations described in this Appendix.

[61] If S is distributed geometric with mean $1/q$, then this formula quickly shows that $\nu=\delta(1-q)/(1-\delta(1-q))$.

[62] Note, however, that in the juvenile analogue to Equation (B.3), $\psi(b^*)$ retains its adult definition because the juvenile will be an adult in subsequent periods.

The *policing elasticity* is given by

$$\eta_p \equiv \frac{\partial(1-F(b^*)}{\partial p}\frac{p}{1-F(b^*)}=-\frac{f(b^*)p}{1-F(b^*)}\frac{\partial b^*}{\partial p} \tag{B.4}$$

By the mean value theorem, there is a $\bar{b}$ between b_J^* and b^* such that

$$f(\bar{b}) = \frac{(1-F(b^*))-(1-F(b_J^*))}{b_J^*-b^*} \tag{B.5}$$

Our model implies that the denominator in this expression is proportional to $\nu_J-\nu$. Too see this, note that Equations (6) and (7) in the main text imply that

$$b_J^*-b^* = \left(c\frac{p}{1-p}+p\psi(b^*)\right)(\nu_J-\nu)=\left(c\frac{p}{1-p}+hE\right)(\nu_J-\nu) \tag{B.6}$$

which in turn means that

$$f(\bar{b}) = \frac{h_J-h}{\nu-\nu_J}\frac{1}{c\frac{p}{1-p}+hE}\frac{1}{p} \tag{B.7}$$

Combining Equation (B.7) with Equation (B.4), we have

$$\eta_p = -\frac{f(b^*)}{f(\bar{b})}\frac{f(\bar{b})p}{1-F(b^*)}\frac{\partial b^*}{\partial p}=-\frac{f(b^*)}{f(\bar{b})}\frac{h_J-h}{\nu-\nu_J}\frac{1}{c\frac{p}{1-p}+hE}\frac{p}{h}\frac{\partial b^*}{\partial p} \tag{B.8}$$

$$= -\frac{f(b^*)}{f(\bar{b})}\frac{h_J-h}{h}\frac{\nu}{\nu-\nu_J}\frac{p}{c\frac{p}{1-p}\nu+h\nu E}\frac{\partial b^*}{\partial p} \tag{B.9}$$

Using the results from the prior section, we have

$$\frac{\partial b^*}{\partial p} = -\frac{\partial G/\partial p}{\partial G/\partial b^*}=-\frac{c\frac{1}{(1-p)^2}(1+\nu)+\nu\psi(b^*)}{-p\nu(1-F(b^*))-1}=\frac{c\frac{p}{(1-p)^2}(1+\nu)+\nu hE}{1+\nu h}\frac{1}{p} \tag{B.10}$$

Finally, combining Equations (B.9) and (B.10), we have

$$\eta_p = -\frac{f(b^*)}{f(\bar{b})}\left(\frac{h_J-h}{h}/\frac{\nu-\nu_J}{\nu}\right)\frac{c\frac{p}{(1-p)^2}(1+\nu)+\nu hE}{c\frac{p}{1-p}\nu+h\nu E}\frac{1}{1+\nu h} \tag{B.11}$$

Turning now to the *sentencing elasticity*, note that since changes in $\mathbb{E}[S]$ affect ν but no other parameter,

$$\eta_{\mathbb{E}[S]} = \frac{\partial(1-F(b^*))}{\partial\nu}\frac{\nu}{1-F(b^*)}\eta_\nu=-\frac{f(b^*)\nu}{1-F(b^*)}\frac{\partial b^*}{\partial\nu}\eta_\nu \tag{B.12}$$

Using the results from the prior section, we have

$$\frac{\partial b^*}{\partial\nu} = -\frac{\partial G/\partial\nu}{\partial G/\partial b^*}=-\frac{c\frac{p}{1-p}+p\psi(b^*)}{-p\nu(1-F(b^*))-1}=\frac{c\frac{p}{1-p}+hE}{1+\nu h} \tag{B.13}$$

Parallel to above, we now substitute in Equation (B.7) and then (B.13) into (B.12) to obtain

$$\eta_{\mathbb{E}[S]} = -\frac{f(b^*)\nu}{1-F(b^*)}\frac{\partial b^*}{\partial \nu}\eta_\nu \tag{B.14}$$

$$= -\frac{f(b^*)}{f(\bar{b})}\frac{f(\bar{b})\nu}{1-F(b^*)}\frac{\partial b^*}{\partial \nu}\eta_\nu \tag{B.15}$$

$$= -\frac{f(b^*)}{f(\bar{b})}\frac{h_J-h}{\nu-\nu_J}\frac{\nu}{c\frac{p}{1-p}+hE}\frac{1}{h}\frac{\partial b^*}{\partial \nu}\eta_\nu \tag{B.16}$$

$$= -\frac{f(b^*)}{f(\bar{b})}\left(\frac{h_J-h}{h}\Big/\frac{\nu-\nu_J}{\nu}\right)\frac{1}{c\frac{p}{1-p}+hE}\frac{\partial b^*}{\partial \nu}\eta_\nu \tag{B.17}$$

$$= -\frac{f(b^*)}{f(\bar{b})}\left(\frac{h_J-h}{h}\Big/\frac{\nu-\nu_J}{\nu}\right)\frac{1}{1+\nu h}\eta_\nu \tag{B.18}$$

3. Implications of Non-increasing Density $f(\cdot)$

Suppose now that the density of B is such that $f(\cdot)$ is non-increasing past b_J^*. The elasticity η_p is the most negative when the product $\frac{f(b^*)}{f(\bar{b})}\frac{c\frac{p}{(1-p)^2}(1+\nu)+\nu hE}{c\frac{p}{1-p}\nu+\nu hE}$ is as large as possible. Consider then the smallest possible E (since the second factor is positive and decreasing in E) consistent with the model and the assumption that $f(\cdot)$ be non-increasing. We have the following inequalities: (1) $E \geq \frac{1-F(b^*)}{2f(b^*)}$, and (2) $f(b^*) \leq \frac{F(b^*)-F(b_J^*)}{b^*-b_J^*}$.[63] Inequalities (1) and (2), together with Equation (B.4), yield

$$E \geq \frac{1-F(b^*)}{2f(b^*)} \geq \frac{(1-F(b^*))(b^*-b_J^*)}{2(F(b^*)-F(b_J^*))} = \frac{(1-F(b^*))(\nu-\nu_J)(c\frac{p}{1-p}+hE)}{2(F(b^*)-F(b_J^*))} \tag{B.19}$$

Minor rearrangement allows us to isolate E, and we obtain the bound

$$E \geq \frac{(1-F(b^*))(\nu-\nu_J)c\frac{p}{1-p}}{2(F(b^*)-F(b_J^*))-h(1-F(b^*))(\nu-\nu_J)} \equiv \underline{E} \tag{B.20}$$

Since $E>0$, we must have $\underline{E}>0$, which is true if and only if the denominator is positive, which in turn is true if and only if

$$\nu-\nu_J < \frac{2}{h}\frac{h_J-h}{h} \tag{B.21}$$

This means that since $\nu-\nu_J$ is increasing in δ, certain values of δ may be strictly inconsistent with the model and measured rates of arrest h and h_J.

When Inequality (B.21) is satisfied, the smallest E consistent with the model is given by the $\underline{E}$ defined in (B.20). When $E=\underline{E}$, Inequalities (1) and (2) are both binding, and if (2) is binding then the density between b_J^* and b^* is flat, which implies $f(b^*)/f(\bar{b})=1$. Thus, the most negative elasticity consistent with our modeling assumptions is given by

$$\overline{\eta}_p = -\left(\frac{h_J-h}{h}\Big/\frac{\nu-\nu_J}{\nu}\right)\frac{c\frac{p}{(1-p)^2}(1+\nu)+\nu h\underline{E}}{c\frac{p}{1-p}\nu+\nu h\underline{E}}\frac{1}{1+\nu h} \tag{B.22}$$

Similarly, under a non-increasing $f(\cdot)$, by setting $f(b^*)=f(\bar{b})$, we obtain the most negative sentencing

[63] The first inequality corresponds to a uniform distribution over the range b^* to $b^*+\frac{1-F(b^*)}{f(b^*)}$, which on inspection is the distribution which yields the smallest possible value of E and yet is consistent with a boundary adult crime rate of $1-F(b^*)$ and does not violate the assumption of a non-increasing density. The second inequality follows since if $f(\cdot)$ is non-increasing then the distribution function $F(\cdot)$ is concave, implying $f(b^*) \leq \frac{F(b^*)-F(b_J^*)}{b^*-b_J^*} \leq f(b_J^*)$ since $b_J^* \leq b^*$.

elasticity consistent with our estimates and model as

$$\overline{\eta}_{\mathbb{E}[S]} = -\left(\frac{h_J-h}{h}\Big/\frac{\nu-\nu_J}{\nu}\right)\frac{1}{1+\nu h}\eta_\nu \tag{B.23}$$

4. Intertemporal Elasticity

The sentencing elasticities discussed above incorporate two distinct mechanisms: (1) an increase in punishments reduces the immediate attractiveness of crime, and (2) an increase in punishments reduces the expected benefit from a future crime, and hence reduces the opportunity cost of being incarcerated (and therefore unable to commit crime). An elasticity that isolates the first mechanism would immediately change punishments today, while keeping future punishments constant. Call this notion of elasticity an "intertemporal elasticity."

Consider a generic response to a 1 percent increase in the juvenile's punishment $\mathbb{E}[S_J]$. Because the adult's punishment does not change, $\mathbb{E}[B-b^*|B>b^*]$ is held constant as $\mathbb{E}[S_J]$ changes. This elasticity is

$$\eta_{\mathbb{E}[S_J]} = -\frac{pf(b_J^*)}{h_J}\left[c\frac{p}{1-p}\nu_J+\nu_J hE\right]\eta_{\nu_J} \tag{B.24}$$

To isolate the intertemporal elasticity described above, we evaluate this elasticity assuming that the distribution of sentence lengths is equal for adults and juveniles. When this is the case, we have

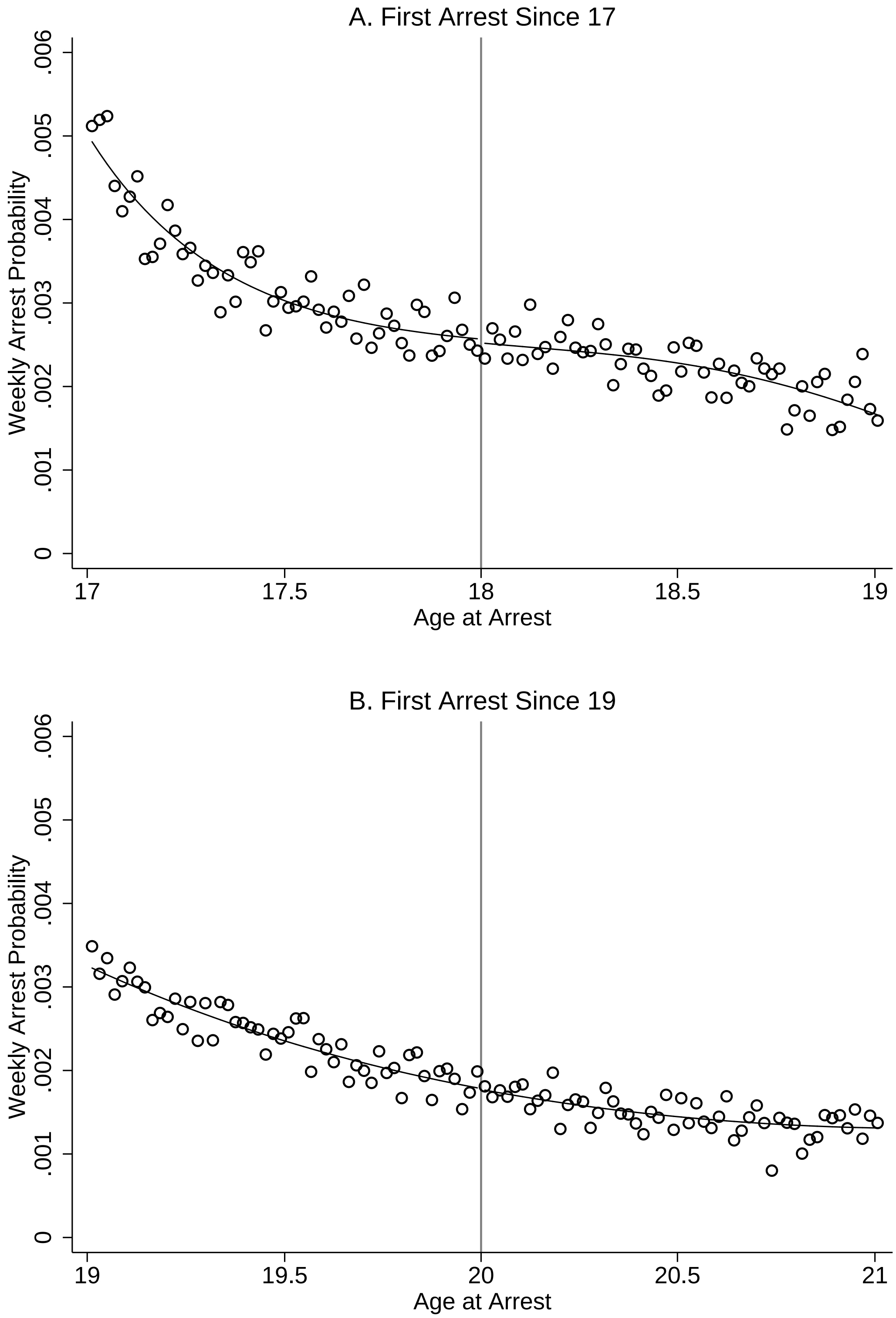
$$\eta_{\mathbb{E}[S_J]} = \eta_{\mathbb{E}[S]}(1+\nu h) \tag{B.25}$$

The resulting analogue to $\overline{\eta}_{\mathbb{E}[S]}$ becomes

$$\overline{\eta}_{\mathbb{E}[S_J]} = -\left(\frac{h_J-h}{h}\Big/\frac{\nu-\nu_J}{\nu}\right)\eta_\nu \tag{B.26}$$

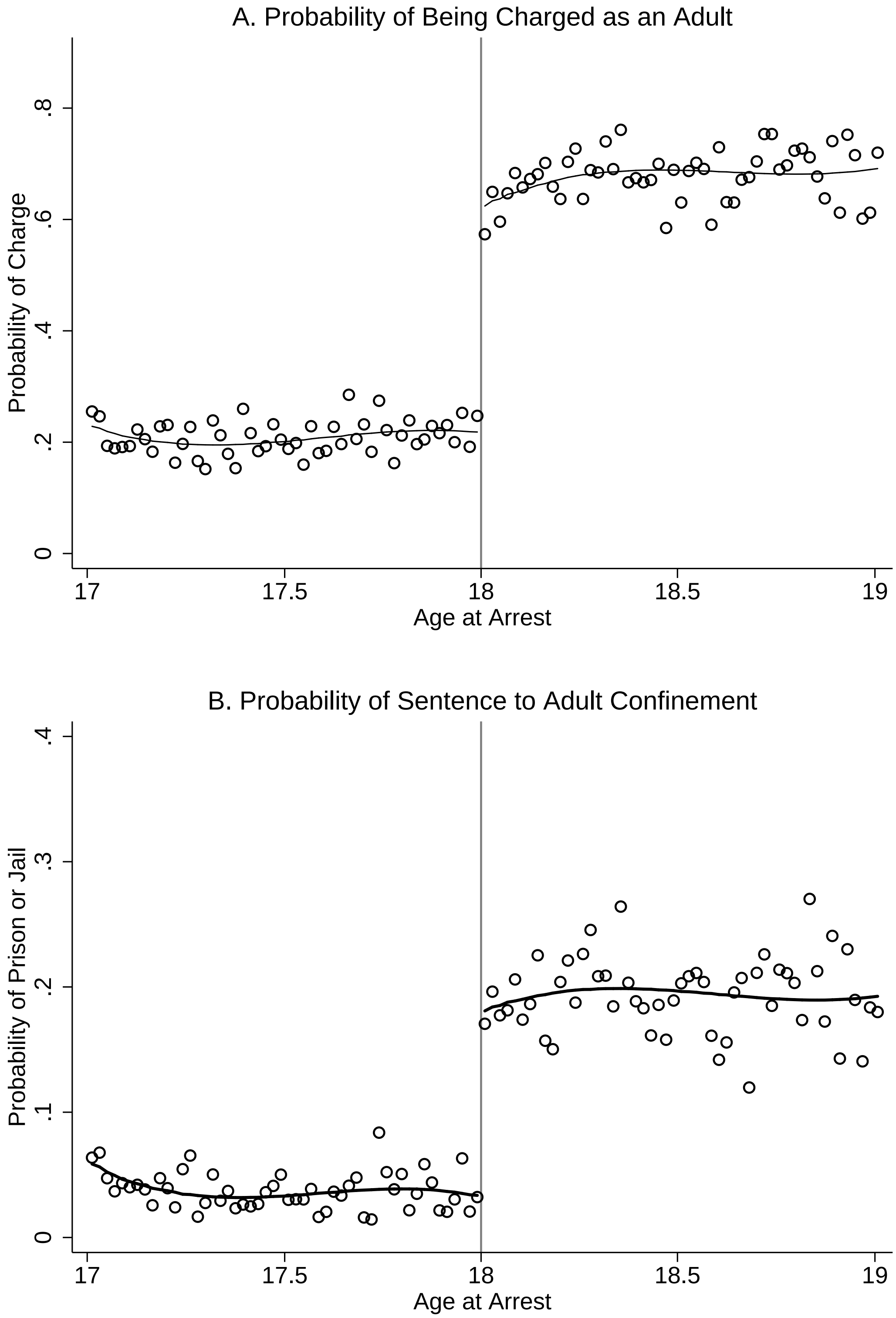
Both $\eta_{\mathbb{E}[S_J]}$ and $\overline{\eta}_{\mathbb{E}[S_J]}$ differ from the adult analogue only by the term $1/(1+\nu h)$. This discussion clarifies that this term reflects the fact that an increase in punishments reduces the expected benefit from a future crime, and hence reduces the opportunity cost of being incarcerated.

FIGURE 1. CRIMINAL PROPENSITY ESTIMATES BY AGE



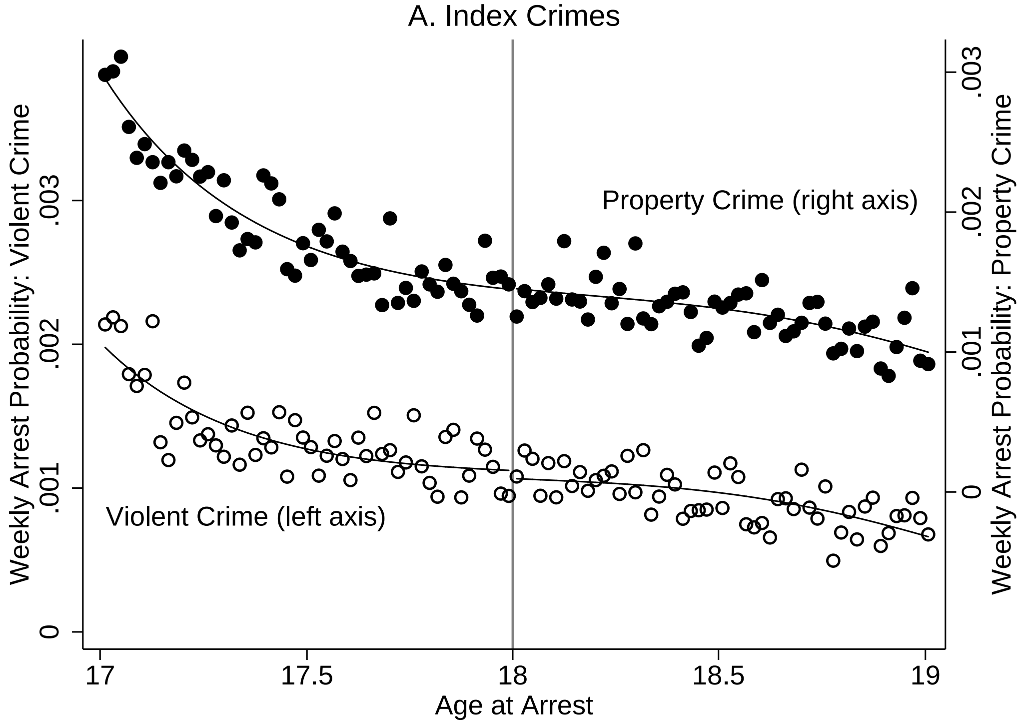
Note: Top panel of figure shows estimates of the hazard for index crime arrest for all those arrested at least once for any felony prior to 17 between 1995 and 1998. Open circles are weekly nonparametric estimates of the hazard, computed as the number offending in the given week, as a fraction of those who have neither reoffended nor been censored, as of the given week. Solid line presents a smoothed estimate based on a logit model, allowing for a jump at 18 (see text for details). Bottom panel of figure presents a falsification analysis pertaining to the arrest hazard for all those arrested at least once prior to 19 between 1995 and 1998.

FIGURE 2. TRANSFER AND PUNITIVENESS BY AGE



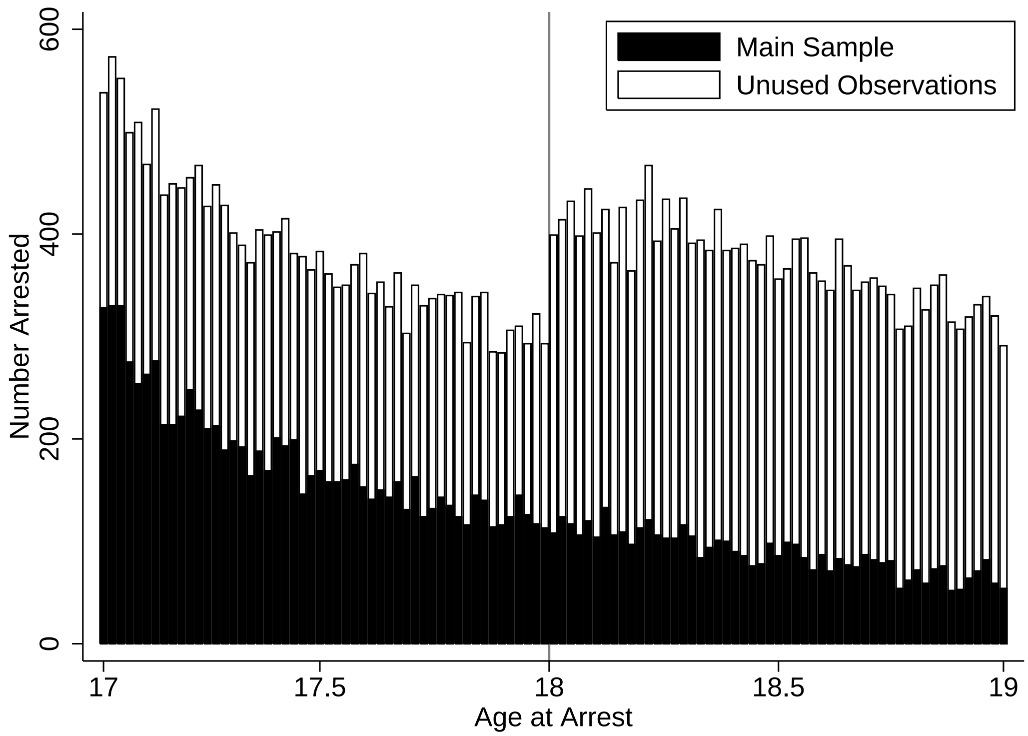
Note: Top panel of figure shows the probability of being charged as an adult, as a function of age at first arrest since 17. Sample is identical to that for Figure 1. Open circles are nonparametric local averages, computed as the sample proportion charged as an adult for individual arrested in the given week. Solid line is based on a flexible polynomial model (see text for details). Bottom panel shows analogous figure for whether arrestee was sentenced to adult confinement.

FIGURE 3. CRIMINAL PROPENSITY BY TYPE OF OFFENSE



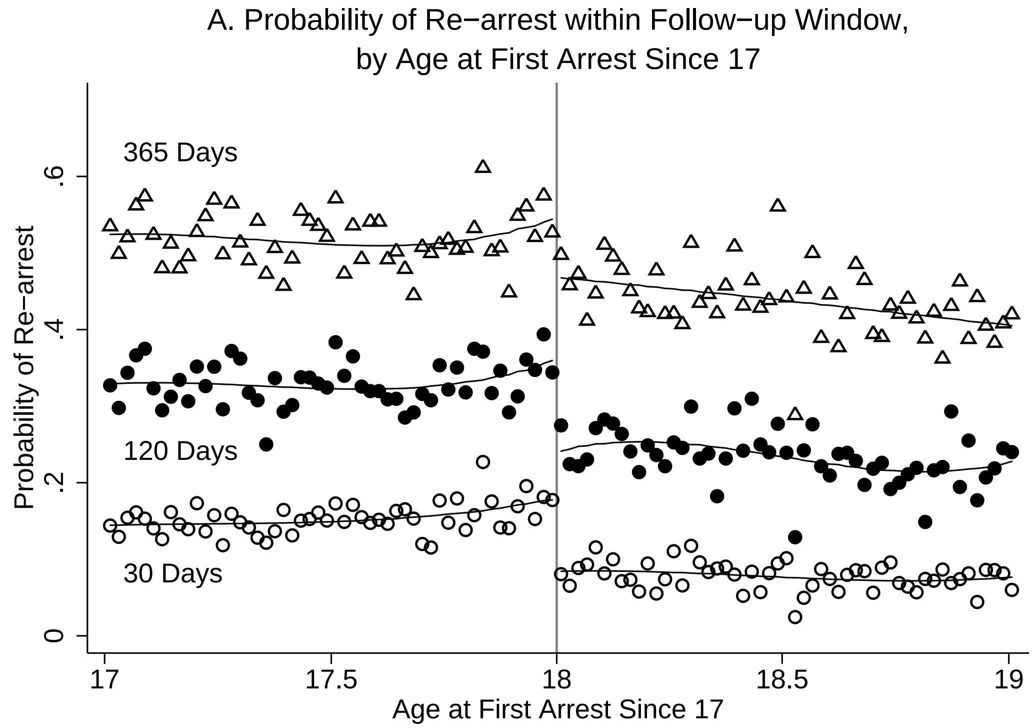
Note: Figure is analogous to Figure 1. Top panel shows re-arrest, disaggregated for murder, rape, robbery, and aggravated assault ("violent crime") and burglary, larceny, and motor vehicle theft ("property crime"). Bottom panel shows re-arrest for a non-index felony offense.

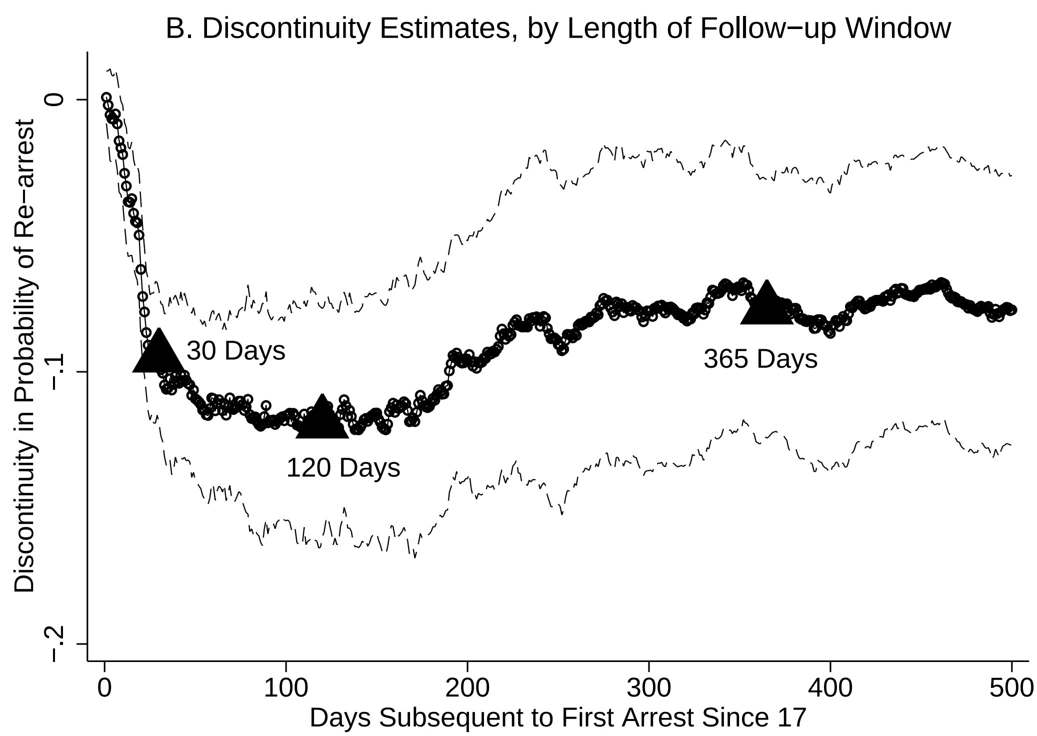
FIGURE 4. NUMBER OF ARRESTS BY AGE:
MAIN SAMPLE VERSUS UNUSED OBSERVATIONS



Note: Figure is a stacked histogram (shaded and light rectangles sum to total). Heights refer to the number of individuals arrested. "Main sample" refers to the sample underlying Figure 2 and pertains to those arrested at least once for any offense prior to 17 between 1995 and 1998. "Unused observations" pertains to all other felony arrests.

FIGURE 5. INCAPACITATION EFFECTS OF ADULT SANCTIONS





Note: Top panel presents probability of re-arrest within 30, 120, and 365 days of first arrest since 17, by age in weeks. Local averages by week, computed as sample proportions, are accompanied by a smoothed estimate based on a flexible polynomial model (see text for details). Bottom panel plots estimated discontinuity in probability of re-arrest, for follow-up lengths ranging from 1 to 500 days, with twice pointwise standard error bands.

Table 1. Summary Statistics, Estimation Sample

Variable	Main Analysis Sample: Baseline Arrest	First Arrest	Young Arrestees
	(1)	(2)	(3)
Non-white	0.42	0.52	0.48
Male	0.80	0.88	0.88
Age	14.93	18.39	18.14
	(1.39)	(1.24)	(0.56)
Arrested in Small County	0.28	0.31	0.31
Arrested in Medium County	0.36	0.35	0.35
Arrested in Large County	0.36	0.34	0.34
Index Crime, Violent	0.26	0.25	0.24
Index Crime, Property	0.51	0.33	0.39
Non-index Crime, Drug	0.07	0.22	0.20
Non-index Crime, Non-drug	0.15	0.19	0.17
Number of Persons	64,073	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 2. Discontinuity Estimates of Deterrence

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discontinuity Estimate	-0.018	-0.017	-0.017	-0.018	-0.017	-0.017	-0.017	-0.017
	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)
Non-white		0.545	0.503	0.576	0.533	0.542	0.540	0.540
		(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
County Size, Baseline (Relative to Large County)								
Small			0.250		0.255	0.247	0.250	0.250
			(0.021)		(0.021)	(0.021)	(0.021)	(0.021)
Medium			0.081		0.087	0.087	0.086	0.086
			(0.021)		(0.021)	(0.021)	(0.021)	(0.021)
Type of Crime, Baseline								
Index Crime, Violent				0.020	0.009	0.016	0.016	0.016
				(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Index Crime, Property				0.191	0.190	0.202	0.205	0.205
				(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Non-index Drug				-0.153	-0.154	-0.207	-0.213	-0.213
				(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
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	-96,006	-95,491	-95,420	-95,412	-95,339	-95,237	-95,178	-95,175

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 3. 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 4. 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	0.0300		
					(0.0821)	(0.0821)		
Medium County					-0.0214	-0.0208		
					(0.0807)	(0.0807)		
Small County					-0.0622	-0.0619		
					(0.0834)	(0.0834)		
Baseline Crime was:								
Index, Violent							0.0553	0.0556
							(0.0930)	(0.0930)
Index, Property							-0.0696	-0.0683
							(0.0646)	(0.0646)
Non-index, Drug							0.0919	0.0952
							(0.1960)	(0.1959)
Non-index, Non-drug							-0.0032	-0.0025
							(0.1250)	(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.

Table 5. Predicted Reduced Form Discontinuities: Parametric Calibration

A. Baseline Parameterization

Probability of Apprehension, p	Annual Discount Factor, $\delta^{1/365}$	Expected Sentences, in Weeks: Juveniles	Expected Sentences, in Weeks: Adults	Benefit Distribution Scale Parameter, λ	Predicted Reduced Form Discontinuity, θ
0.08	0.95	9.03	29.57	4.56	-2.76

B. Effects of Changes in Individual Parameters, Relative to Baseline

Probability of Apprehension		Juvenile Sentences Parameter		Adult Sentences Parameter	
p	θ	k_J	θ	k_A	θ
0.025	-2.00	0.250	-2.95	0.250	-2.22
0.050	-2.46	0.400	-2.78	0.400	-2.70
0.200	-3.33	2.000	-2.76	2.000	-2.77
0.400	-3.74	4.000	-2.76	4.000	-2.78

Annual Discount Factor		Benefit Distribution Parameter	
$\delta^{1/365}$	θ	k	θ
0.500	-2.50	0.250	-1.03
0.250	-2.26	0.400	-1.79
0.100	-2.01	2.000	-3.55
0.010	-1.57	4.000	-3.84

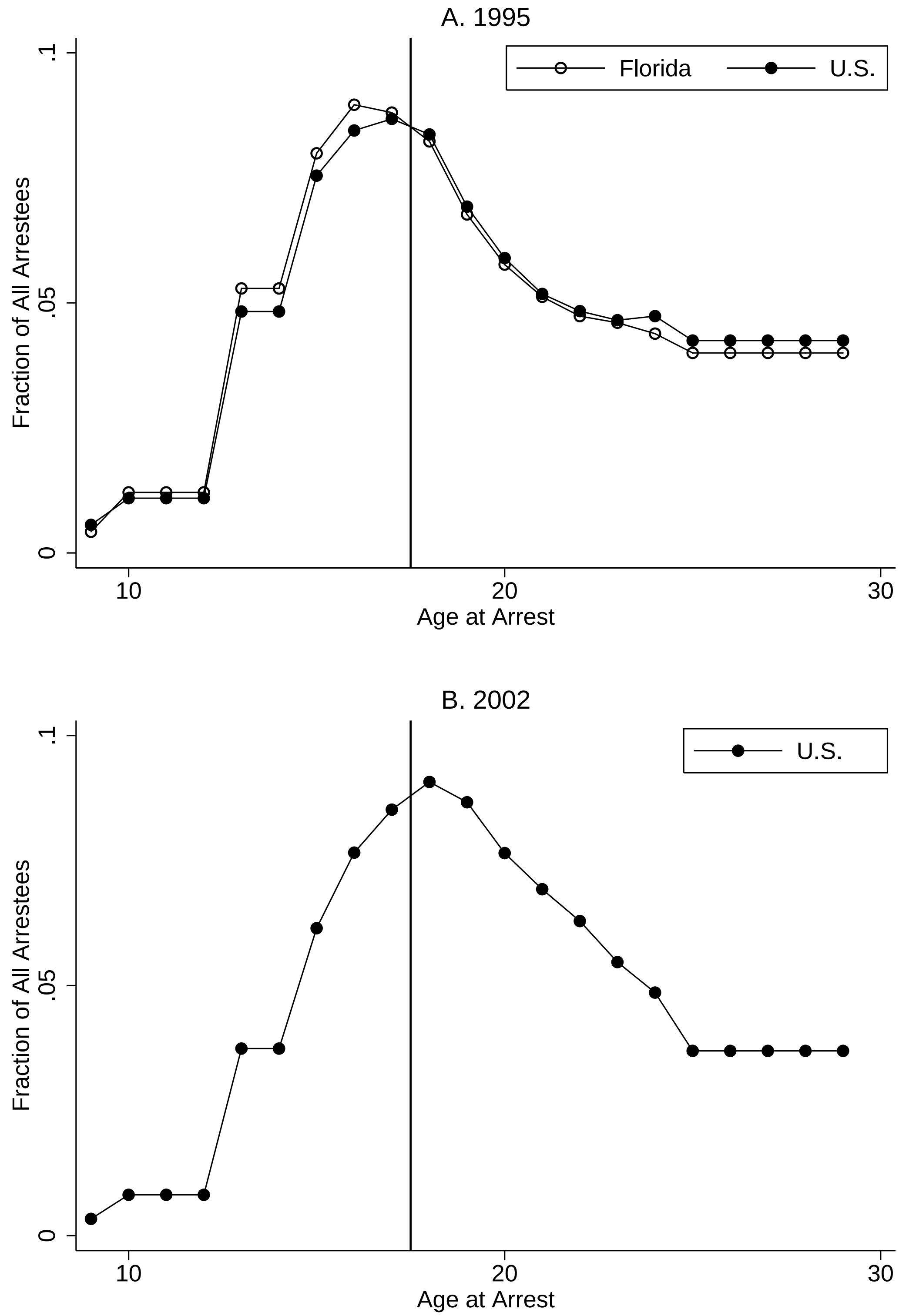
Note: In panel A, all Weibull shape parameters k, k_J, and $k_A$are equal to 1. In panel B, Weibull shape and other parameters vary, as specified. Throughout, when k_J varies, l_J is adjusted to hold E[S_J] constant, and when k_A varies, l_A is adjusted to hold E[S] constant. Each entry of the table matches the adult hazard by adjusting the benefit distribution scale parameter, λ. Details in text.

TABLE 6. BOUNDS ON POLICY ELASTICITIES

Annual Discount Factor	$\theta = -0.11$: $\overline{\eta}_p$	$\theta = -0.11$: $\overline{\eta}_{E[S]}$	$\theta = -0.018$: $\overline{\eta}_p$	$\theta = -0.018$: $\overline{\eta}_{E[S]}$
10^{-8}	-0.760	-0.060	-0.117	-0.009
0.0010	-0.379	-0.071	-0.057	-0.011
0.0025	-0.349	-0.073	-0.052	-0.011
0.0050	-0.326	-0.075	-0.048	-0.012
0.0100	-0.302	-0.077	-0.044	-0.012
0.0250	-0.271	-0.082	-	-
0.0500	-0.248	-0.086	-	-
0.1000	-0.223	-0.091	-	-
0.2500	-0.190	-0.101	-	-
0.5000	-0.163	-0.113	-	-
0.9500	-0.136	-0.130	-	-

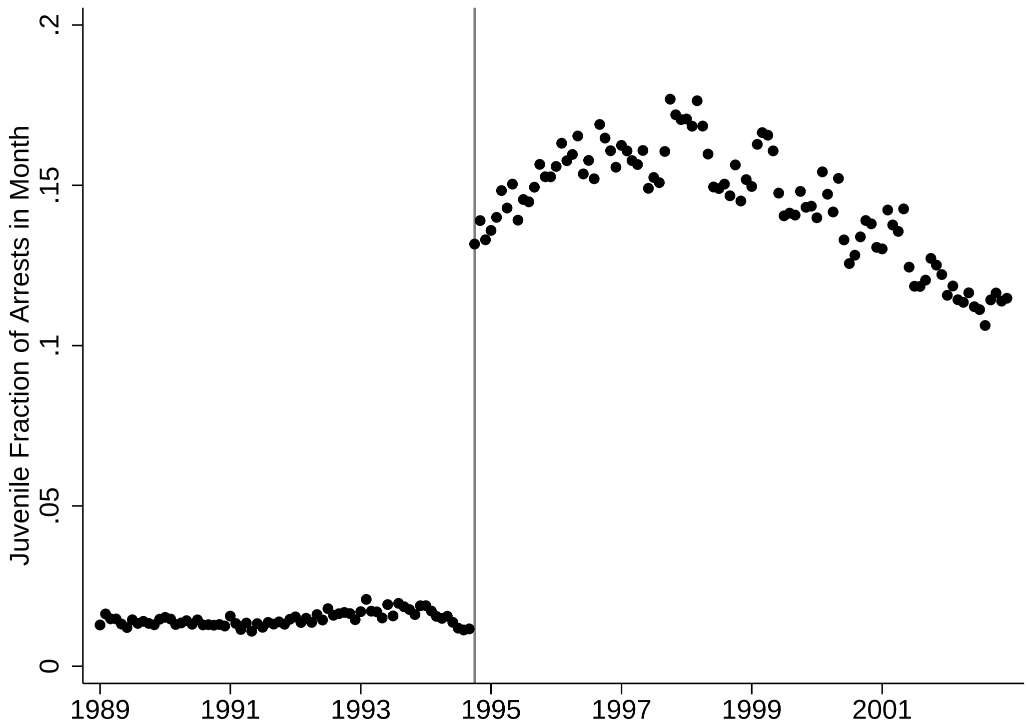
Notes: Table presents lower bounds on the elasticity of crime with respect to police, $\overline{\eta}_p$, and with respect to expected sentence lengths, $\overline{\eta}_{E[S]}$, for various values of the annual discount factor, δ, under two assumptions regarding the reduced form parameter ($\theta = -0.11$ and $\theta = -0.018$). Not all values of δ are consistent with the assumption of a declining density function for the distribution of criminal benefits; for values of δ inconsistent with this assumption, elasticity columns contain a dash.

APPENDIX FIGURE 1. AGE DISTRIBUTION OF CRIME: FBI UNIFORM CRIME REPORTS



Note: Figure presents the histogram of age at arrest, for 1995 and 2002, for Florida and the U.S, for cities of 10,000 population or more. Data pertain to arrests for murder, rape, robbery, aggravated assault, burglary, larceny, and motor vehicle theft ("index arrests") and are from the FBI Uniform Crime Reports. 1995 is the last year Florida provided the FBI with detailed age breakdowns for arrests.

APPENDIX FIGURE 2. IMPACT OF JUVENILE JUSTICE REFORM ACT OF 1994 ON FDLE COVERAGE OF JUVENILE ARRESTS



Note: Figure shows fraction of arrests in FDLE data set pertaining to juveniles, by month since 1989. Vertical line indicates effective date of Juvenile Justice Reform Act of 1994.

Appendix Table 1. Juvenile and Adult Incarceration Length in Weeks

	Arrests	Jail	Average Duration	Prison	Average Duration	Total	Average Duration	Average Duration, Adjusted for Juvenile Transfer[1]
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	3.71
Adult	766,259	32,585	2.21	89,730	6.09	122,315	8.30	8.12
Adult/Juvenile Ratio			**3.77**		**2.95**		**3.13**	**2.19**
Index Offenses as Denominator								
Juvenile	53,967	1,482	1.43	5,211	5.02	6,693	6.45	9.03
Adult	133,853	32,585	12.66	89,730	34.86	122,315	47.52	46.48
Adult/Juvenile Ratio			**8.86**		**6.94**		**7.37**	**5.15**
B. United States								
All Offenses as Denominator								
Juvenile	1,588,839	26,439	0.87	76,926	2.52	103,365	3.38	4.74
Adult	7,552,362	354,379	2.44	1,546,456	10.65	1,900,835	13.09	12.80
Adult/Juvenile Ratio			**2.82**		**4.23**		**3.87**	**2.70**
Index Offenses as Denominator								
Juvenile	420,543	26,439	3.27	76,926	9.51	103,365	12.78	17.89
Adult	1,091,530	354,379	16.88	1,546,456	73.67	1,900,835	90.55	88.59
Adult/Juvenile Ratio			**5.16**		**7.75**		**7.09**	**4.95**

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.

[1] Adjustment for juvenile transfer assumes 20 percent of juvenile arrestees tried in adult criminal court and 50 percent tried in juvenile court; see Appendix for details of calculations.

Appendix Table 2. Arrest Probabilities

Crime Category	Offenses Known to Police	Fraction of Offenses Cleared by Arrest	Fraction of Victimizations Reported to Police	Probability of Apprehension
	(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 3. Robustness of Main Results

	Imposing Equality of Derivatives at 18					Allowing for Different Derivatives at 18				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
A. Unconditional Estimates										
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 Dr	-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 4. Discontinuity in Probability of Adult Incarceration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Main Estimates								
Estimated Discontinuity in Probability of	0.1475	0.1474	0.1459	0.1479	0.1464	0.1463	0.1459	0.1460
Confinement	(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	0.0401	0.0400	0.0394	0.0401	0.0395	0.0395	0.0392	0.0394
Prison	(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	0.1075	0.1074	0.1065	0.1078	0.1069	0.1068	0.1066	0.1067
Jail	(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.