Age-Dependent Court Sentences and Crime Bunching

Empirical Evidence from Swedish Administrative Data*

Mårten Palme[a], Mikael Priks[b] and Björn Tyrefors[c]

April 15, 2018

Abstract: We estimate the deterrent effect of incarceration by exploiting a rebate system in Swedish prison sentences before individuals' 21st birthday. Using detailed register data, we find evidence of bunching in the sense that more crimes were committed during the week prior to a 21st birthday, followed by a reduction in crime during the week after this birthday. The result is driven by the group of prolific offenders. We find no evidence, however, that harsher punishment permanently reduces crime rates.

* We thank Randi Hjalmarsson, Matthew Lindquist, Stephen Machin, Sten Nyberg as well as seminar participants at Bonn and Stockholm Universities for helpful comments. Financial support from the Jan Wallander and Tom Hedelius Foundations is gratefully acknowledged.

[a] Department of Economics, Stockholm University. E-mail: Marten.Palme@ne.su.se.

[b] Department of Economics, Stockholm University. E-mail: Mikael.Priks@ne.su.se.

[c] Research Institute of Industrial Economics (IFN) and Department of Economics, Stockholm University. E-mail: Bjorn.Tyrefors@ifn.se and Department of Economics, Stockholm University.

1. Introduction

The question of how punishment affects crime has been a long-standing debate in policy circles as well as in different fields of the social sciences. Following Becker (1968), many economists have argued that criminals exhibit rational behavior in the sense that they weigh expected costs and benefits before committing a crime. This implies that incentives matter and that relatively harsher punishment should reduce the crime rate in society. An alternative view is that the behavior of criminals is determined by emotional, psychological and social factors that are not directly affected by punishment. While discriminating between these alternative views is ultimately an empirical question, research has been restricted by a lack of high-quality data and empirical research methods.

In this paper, we apply a research design that exploits the age rebate in the Swedish legal system whereby punishments increase discontinuously at the 21st birthday. Life imprisonment, for example, cannot be used, and the convicted person is given a 25-percent reduction in the penalty. These rules create a discontinuity in the sentences around an individual's birthday that we use as identifying variation to study the probability and timing of committing crimes. Our research design requires day-to-day individual data on criminal behavior. We use the Swedish National Conviction Register, which includes individual information on all convictions throughout the Swedish legal system, linked with the Swedish Census, which provides demographic information on the included individuals.

By varying the bandwidths around the 21st birthday we are able to identify two separate potential effects of the age rebates in the Swedish judicial system. First, a general overall effect on the number of convictions of the harsher punishment after the 21st birthday. Second, a reallocation, or bunching, effect of crimes around the 21st birthday. Our theoretical motivation for the existence of the bunching of crimes before the 21st birthday is that waiting until an optimal opportunity arises is associated with reallocation costs for the offender (see, e.g., Felson and Clark 1998). Criminals have, for example, imperfect information about when, where and how a future crime will be committed, and preparations are likely to be more costly the more distant the crime is in the future. An economic model that takes these costs into account will generate front loading of crimes due to the rebate at the age threshold, but this effect will be weighed against and counteracted by the reallocation costs.

Our results show a large and significant bunching effect amounting to approximately a 15 percent increase in the number of crimes close to one week prior to this birthday, followed by a reduction the week after. Calculated from a reference point immediately prior to the estimated bunching behavior, the increase before and the decrease after the 21st birthday are about equal in size, suggesting zero overall effect. The latter result means that we are not able to conclude that the decreased deterrence effect of the more lenient punishments for those younger than age 21 has led to more crime, resulting in a cost for the society.

Our results differ from those obtained in the previous literature on US data (Hjalmarsson, 2009, and McCrary, 2017) in the sense that we show that criminals reallocate their criminal behavior as a response to harsher penalties at the age of criminal majority.[1] However, they concur with the previous results in the sense that we also find a zero overall effect of the youth rebate in the judicial system.

A potential reason to why we, as opposed to the previous literature, find evidence of intertemporal reallocation of criminal behavior is that the age limit we analyze is the 21st birthday rather than the 16th or 18th as in the previous studies. It is likely that youth criminals may be less informed, planned and structured in their criminal behavior than older ones. In fact, Hjalmarsson (2009) showed that juveniles underestimate the chance of jail when reaching the age of criminal majority. There are 15-percent youth rebates in the Swedish judicial system at each birthday from 15 to 20 (Jareborg, 2004). We are thus able to replicate the results from the previous studies on our data and we find no effects of youth rebates on crime at 16 and 17.[2] The hypothesis is further supported by the fact that our behavioral responses are attributed to repeated criminals that may

[1] In a seminal paper, Levitt (1998) used annual data and was therefore unable to study bunching.

[2] It is difficult to isolate the effect at 18 because not only penalties change discontinuously. In Sweden, 18 is the minimum age for obtaining a driver's license and for buying alcohol in restaurants. In the U.S., criminal majority tends to coincide with other changes potentially related to crime, such as laws regarding firearms, curfews, drivers' licenses, drop outs from school and gambling. In Florida, for example, at age 18, individuals are able to legally drop-out of school without parental consent. The deterrent effect at 20 in Sweden is also hard to isolated since this is the minimum age for buying alcohol in stores. At age 19, we find non-significant reductions in crime.

be more likely to act on incentives through punishment rules rather than youths that are more likely to commit unplanned crimes.

In addition to the literature on the effects of the increased punishment at the age of criminal majority, this study also relates to papers on intertemporal and spatial determinants of crime. Jacob et al. (2007) studied intertemporal displacement of criminal activity, using weather conditions as an instrumental variable for crime. They found that a ten percent increase in crime one week was followed by a 2-2.6 percent reduction in crime the following week. Draca et al. (2010) analyzed the police intervention that occurred in London in 2005 following the terror attacks in the city. They found direct effects of the intervention but no intertemporal displacement. Vollaard (2017) studied illegal discharges of oil from shipping exploiting that the probability of detection was reduced after sunset. He found clear evidence of intertemporal displacement of this type of crime since the timing of discharges moved back and forth with the time of sunset. While only a few papers have analyzed intertemporal displacement of crime, spatial displacement of crime has been suggested to be an important aspect of crime control. The previous empirical literature has, however, found only limited support for it (see e.g., Weisburd et al., 2006, and Braga and Bond, 2008).[3]

The outline of the paper is as follows. Section 2 presents a mechanism, and Section 3 discusses the Swedish court system and the structure of penalty reductions for juveniles in Sweden. Section 4 describes the data and Section 5 the empirical strategy. The results are presented in Section 6. Section 7 concludes.

2. Rules for Age of Criminal Majority, Reallocation Costs of Criminal Activities and Bunching of Crimes

[3] There is also related literature using various forms of exogenous variation to study the effect of punishment on crime but not intertemporal displacement. Kessler and Levitt (1999) and Vollard (2013), for example, exploited increases in sentence length for prolific offenders. Raphael and Ludwig (2003) and Abrams (2012) studied the effect of harsher weapon laws. Helland and Tabarrok (2007) examined the “three strikes and you’re out” reform in California. Kuziemko (2013) exploit discontinuities in parole-board guidelines; Hansen (2015) and Pettersson-Lidbom et al. (2017) use discontinuities in BAC levels; and Drago et al. (2009) analyzed random amnesties in Italy. Random assignment of judges as a source of exogenous variation in sentences has been used by, e.g., Mueller-Smith (2015), Aizer and Doyle (2015), Bhuller et al. (2016) and Dobbie et al. (2018).

An economic model where the agent compares the expected benefits with the expected costs of committing a crime would predict that a criminal reallocates crime that he or she has planned to commit after turning 21 to before that birthday, as the expected cost is lower at that point in time. If there were no costs of reallocating criminal activities in time, it would be rational for the criminal to commit all crimes he or she had planned throughout his or her criminal career before turning 21. The fact that we do not see such behavior in real life suggests that there are substantial reallocation costs associated with criminal activities.

Criminologists Felson and Clark (1998) propose a model where optimal criminal opportunities are at the center, which could be used to shed light on this question. Depending on the type of crime, such opportunities depend on, for example, the day of the week, the time of the year, or whether an area is busy or not. In such a model, front loading or postponing of the crime comes at a cost, and it is likely that this cost is increasing in the time distance from the optimal opportunity. Committing many crimes in a short time period will, for example, come at large costs due to diminishing marginal productivity. It will therefore be cheaper to wait and commit most crimes later when more favorable opportunities arise. Moreover, since the number of crime opportunities are limited (there are only a limited number of cars to steal, houses to break into or victims to rob in an area, etc.), committing a crime early comes at an increased cost of committing a similar crime later, either because the opportunity has already been exploited or because of increased watchfulness on the part of potential victims. It is also costly to plan crimes ahead since it is difficult to receive information about the details of a crime that will occur in a distant future.

Consider a situation where rational prolific offenders committing crime repeatedly can take advantage of the best opportunity to commit crimes. Assume that an offender commits three different crimes and may select the time at which each one of the crimes is committed. There is an optimal opportunity to commit each crime, and front loading or postponing the crime comes at a cost, which is increasing in the time distance from the optimal opportunity.

Figure 1. Convex cost curves, youth rebates and intertemporal displacement

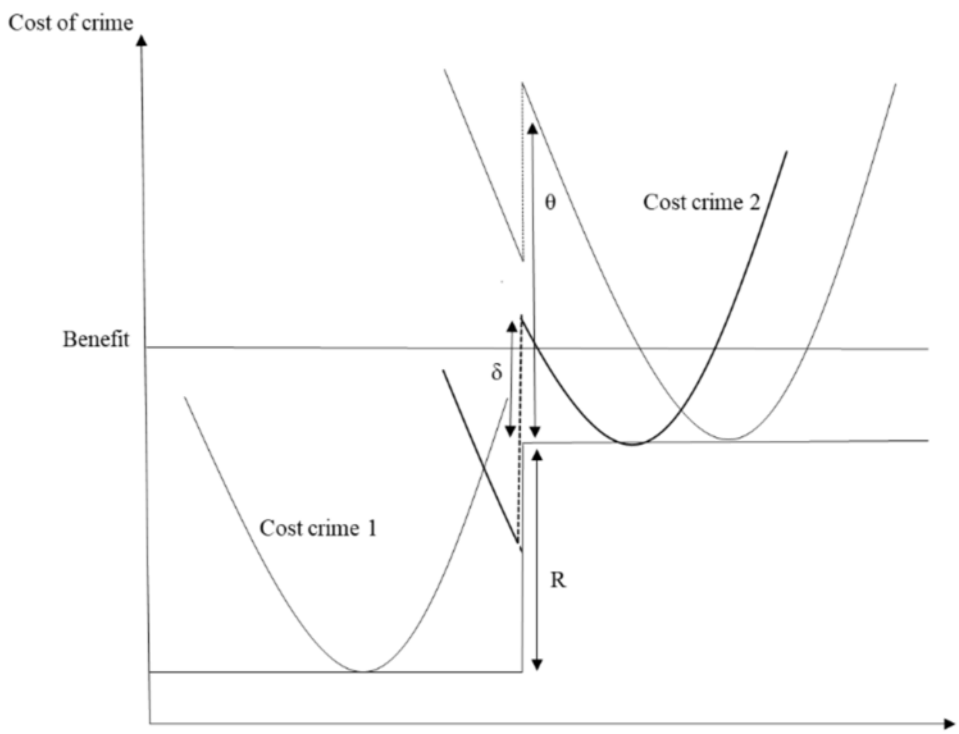


Figure 1 depicts a situation with three convex cost functions, one for each crime where time is measured from left to right. The utility of committing a crime is constant. There is a youth rebate in the punishment, which leads to an increase in the expected cost of committing the crime, R, at the 21st birthday. Therefore, the cost curves 2 and 3 jump up with R at the 21st birthday.

The offender maximizes the expected utility minus the expected cost for each crime. In this framework, as shown in Figure 1, crime 1 will be committed at time t1 because this is when the distance between the utility and the cost of committing crime 1 is maximized. Crime 2 would similarly have been committed in period t2 if there were no rebate in punishment at age 21. However, because of the 25 percent rebate in the punishment, the cost curve of crime 2 shifts downwards by R if the crime is committed before the 21st birthday. In the figure, the expected cost reduction of the rebate R is larger than the cost increase that follows from moving the crime from t2 to t2', δ. This crime will therefore be frontloaded to time t2'. Crime 3 is considered too costly to move from t3 to t2' (the cost increase θ is larger than R) and will therefore be committed in period t3. The result is consequently a bunching effect around the 21st birthday.

What would the result be in a standard Becker model where criminals weight benefits against costs without the convex cost curves? If the utility of committing a crime is lower than the expected value of the rebate R, then all crimes after age 21 would be deterred. If instead, as in Figure 1, the utility of coming a crime is larger than the cost reduction R, then all crimes will be front loaded and committed before the 21st birthday (and no crimes will be deterred). Leaving the standard Becker model in the sense that we have both convex costs and a utility larger than the rebate generates the results that there are no long-term deterrent effects of the rebate at age 21 but rather a displacement effect around the birthday. The more convex the cost curves are, the shorter the bunching effect is, and for significantly convex curves, the youth rebate will have almost no effect at all on crime.

3. The Swedish Court System

3.1 The Processing of Crimes and Comparisons of Prisons in Other Countries

As in most other countries, the prosecutor decides whether the individual should be charged in Sweden's criminal justice system. Of charged cases, 77 percent result in a court trial. Of these court trials, 37 percent result in a fine at most, 24 percent in probation at most, and 23 percent result in at least some incarceration. The defendant is found not guilty only in 6 percent of cases.

The Swedish prison system is similar to other European countries, but it differs from the system in the U.S. Incarceration rates, for example, are similar to those of other European countries (Council of Europe Annual Statistics, 2015), but those in the U.S are approximately 10 times higher. Sentence lengths follow a similar pattern. Conditional on any incarceration, the median prison sentence in Sweden is 3 months (Dobbie et al. 2018), compared to approximately 7 months in Western Europe (Aebi, Tiago, and Burkhardt 2015) and 2.9 years in the United States

(Pew Center 2011).[4] The average annual cost per inmate is US$ 140 000 per year, which is much higher than that in the U.S., where the annual cost is only $ 35 000 (Institute of Public Affairs 2017). These high costs in Sweden are due to a high ratio of guards per prisoner and the greater resources spent on rehabilitation.

3.2 Juvenile Punishments

According to Swedish law, the age of criminals who commit crimes has to be taken into account when sentences are decided. The Swedish Penal Code, chapter 29, § 7 states that age "should be explicitly considered when determining the penalty if the crime is committed before the age of 21" and "no one may be sentenced to life in prison for crimes committed before one's 21st birthday". In the case of repeated offending, "the court must not sentence a higher penalty if the crime was committed before the age of 21" (The Swedish Penal Code, chapter 26, § 3). The youth rebate at age 21 is converted to the size of approximately 25 percent of the sentence. There is also a sharp distinction in Swedish law at an individual's 18th birthday because juveniles below that age can only be sentenced to prison if there are extraordinary circumstances (The Swedish Penal Code, chapter 30, § 5). Under age 15, an individual cannot be sentenced (The Swedish Penal Code, chapter 1, § 6). Apart from these laws, there is an informal practice, which is not part of the written law that grants smaller sentence reductions before juvenile birthdays from 15 to 21. For detailed descriptions of this rebate system, see Jareborg (2004), and in Swedish, Jareborg and Zila (2007) and Supreme Court Judge Borgeke (2008).

If incentives matter, we would expect reductions in crime subsequent to the abovementioned birthdays, but not afterwards. In Sweden, however, 18 is the age of majority when individuals are allowed to take their driver's license and buy alcohol in pubs and restaurants, which may affect criminal behavior. With respect to the informal rebate system at other age thresholds, there are other confounders that make an analysis intractable. At age 20, alcohol may be bought in stores and consumed outside restaurants, and the rebate effect can therefore not be isolated. Age 16 is also a problematic threshold to study since individuals are allowed to practice driving and obtain a license for some vehicles. Thus, there might be deterrent effects at thresholds below 21, as well,

[4] A reason for the short sentences in Sweden is that almost all prisoners are released to probation after two-thirds of the sentence.

but these effects cannot be isolated. The only birthday with a legally-based rebate and with no other confounding factors is the 21st birthday. Birthdays above age 21 can be used as placebos.

4. Data

Our dataset was constructed by matching several different national Swedish registers. The population is defined from consecutive years of the Swedish census. Data on criminality were obtained from the Swedish conviction register provided by the Swedish National Council for Crime Prevention (Brå). This register contains data on all convictions in the Swedish judicial system. The information we used is the date when a crime was committed, the type of crime and the length of the potential prison sentence.

Given our research design, the exact date of birth and its relation to the date when a crime was committed are of key importance. The data from the census only contain information on month of birth, which as will become clear when studying the results masks the sorting effect. It was possible, however, to use information from the Swedish birth register to obtain an almost perfect prediction on the exact date of birth. In Appendix A, we explain in detail how this was done and present results from an evaluation of our prediction procedure.

Abstracting from the discussion of confounding factors, we can analyze the threshold from age 16 and onwards.[5] Since our data end in 2010, we needed to ensure that we did not technically impose a jump in the density due to censoring. For example, say we analyze the number of convictions around the 20th birthday, and the window around the threshold is one year. Again, since our data end in 2010, we could only use cohorts born in 1989 or before. If we also used the cohort born in 1990, then we would have constructed a jump in the density of the number of convicted mechanically, since the last cohort could not have reached the age above the threshold. When pooling cohorts, we allowed only the cohort that had reached the specific age threshold analyzed plus one year. This means that our sample size decreased when studying thresholds for

[5] The age threshold of 15 cannot be used since crimes committed before age 15 are not recorded in the national conviction register, and our approach requires data both before and after the age threshold under study.

older ages. We used the cohorts from 1973 to 1993 when studying the age 16 threshold, 1973 to 1992 when studying the age 17 threshold, and so on. Table 1 shows the number of observations that could potentially be used for each threshold analysis.

Table 1. Potential sample depending on age threshold analyzed.

Threshold	16	17	18	19	20	21
Cohorts that can be used	1973-1993	1973-1992	1973-1991	1973-1990	1973-1989	1973-1988
Total number of convicted	909,291	897,153	879,606	857,946	831,935	803,832

Our main analysis focuses on crimes committed around the 21st birthday, where the maximum sample size we have for analyzing the 21st year threshold is 803,832 individuals as tabulated in Table 1. However, when merging the number of convicted individuals with other statistical sources, we lose observations. Since we matched conviction data with data from the birth register, we can only use Swedish-born convicted individuals. When matching country of birth with data from Statistics Sweden, 170,578 were born abroad, and 73 had an unknown country of birth. This leaves us with 633,181 observations, as documented in Table 2. [6]

When matching this dataset with the birth register, we lost 47,918 observations, so we ended up with 585,263 convictions, i.e., our attrition is approximately 8 percent defining Swedish-born as our population. However, since we had 75,658 empty cells with respect to date of crime, we are down to 509,605 convictions where we can determine the age at the date of crime. Since the population consists of the Swedish-born individuals, the final attrition is approximately 20 percent. Importantly, we have no reason to believe that the attrition is systematically related to exact birth dates.

[6] The pattern of attrition is similar for all thresholds. Please contact the authors for full information.

Table 2. Attrition when matching the different data sources for the maximum sample.

Stage in sample selection	Number of observations
All recorded convictions	803,832
Sample after matched with country of birth	633,181
Sample after matched with birth register	585,263
Sample after using only verdicts with known date of crime	509,605

More than one crime may be handled during one court session. If so, the most severe crime in terms of length of punishment solely determines the harshness of overall punishment. We therefore used the information on the most severe crime committed on the respective date only. Finally, since we focus on crimes committed close to intervals around the 21st birthday, the sample size actually used is smaller than the full sample of 509,605 using the maximum bandwidth.

5. Empirical Strategy

Following Lee and Lemieux (2010), we estimate a jump in the density of the number of verdicts following the different birthdays using the following model:

$$\log(Verdict_j) = \alpha + \beta Above_j + f(W_j) + \varepsilon_j \qquad (1)$$

where the outcome *Verdict* is the number of verdicts for a crime that was committed at a certain age *j*, measured in days relative to the age threshold; *Above* is an indicator variable taking the value of one if the crime was committed on the same day as the age threshold under study or later and zero otherwise. *W* is the forcing variable, age at the time of the crime measured in days, although normalized to be zero at the threshold, positive above and negative below. In other words, if the age threshold of 21 is studied, then W_j = Age (at crime) – 21. The parameter β is the policy variable measuring the percentage difference in the number of verdicts at the threshold, and ε is a random error term.

Equation (1) was estimated using local linear regressions (LLR) as suggested by Hahn et al. (2001) and Porter (2003). As Lee and Lemieux (2010) suggest, we use a rectangular kernel, which is equivalent to estimating a standard linear regression over the interval of the selected bandwidth on both sides of the cut-off point. There are many ways of choosing an optimal bandwidth (see, for example, Imbens and Kalyanaraman, 2012 and Calonico et al., 2014). We have chosen to be agnostic and report many different bandwidths. We also present results based on the higher order of the polynomial function, $f(W_j)$. Since the forcing variable is discrete (age in days at date of crime), we clustered the standard errors at the forcing variable following Card and Lee (2008).

The smallest bandwidth we present is only 0.02. This means that we compare the number of crimes carried out by criminals roughly one week before their birthday with crimes committed one week after since the bandwidth is measured as shares of a year in days (365.25 days). This gives us only 15 clusters, which means that the standard error in these regressions should be interpreted cautiously. However, following the logic of a standard regression discontinuity design, the point estimate is still informative and unbiased (Lee and Lemieux, 2010).

To be able to interpret the estimate β from Equation (1) as a causally valid estimate of the increased punishment following the age threshold on criminality, we need to assume that no criminal determinant other than the severity of punishments changes discontinuously at the birthday thresholds. To detect discontinuities in the density, high frequency data are required. As argued in Lee and McCrary (2017), "all other factors" are likely to be constant when examining offense rates in relatively short intervals. When pooling large intervals (such as annual comparisons across thresholds), many factors affecting criminal activity change in ways that could influence underlying criminal propensities. Compared to age thresholds used in previous studies, the threshold of age 21 in Sweden seems to be better suited for causal analysis of the effect of punishment on crime, as it is not linked to other eligibility ages other than voting in the old days.

6. Results

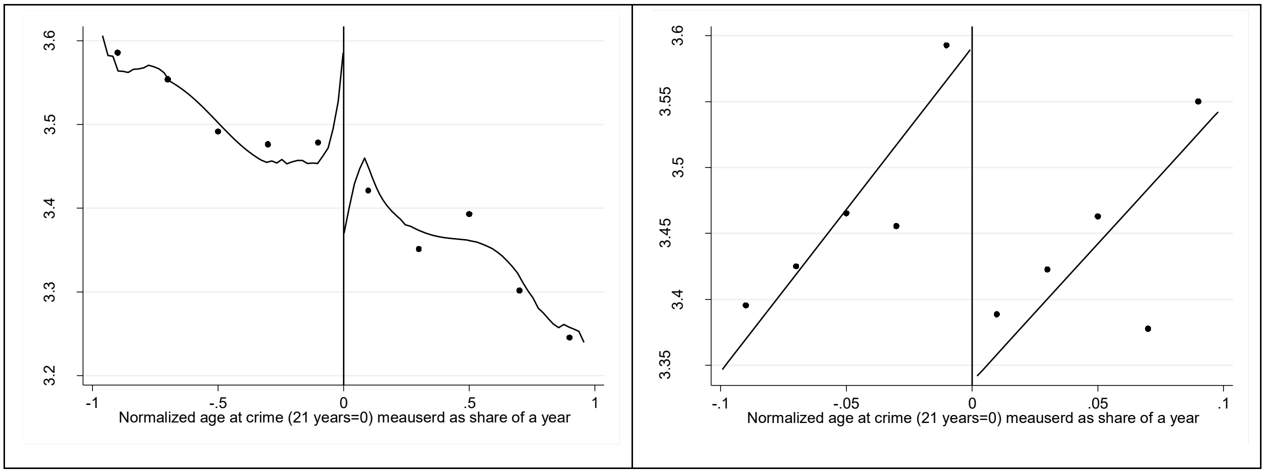
6.1 Main Results

Consider first the effects of the threshold of age 21 on all index crimes.[7] Figure 2A shows the logarithm of the number of crimes one year prior to and one year subsequent to the 21st birthday. The results show a local deterrence effect driven by bunching behavior close to the 21st birthday. The week before the 21st birthday, there is an increase in number of crimes, which is followed by a sharp reduction of approximately the same size the week following this birthday. The reduction in crime at the 21st birthday amounts to approximately 25 percent. Crime then reverts to the original trend. To prove that the results are not functional form specific and overshooting at the threshold due to no support, Figure 2B zooms in on the results roughly one month before and after the threshold of age 21. Again, a strong local 25 percent deterrence effect as well as a stable pattern is revealed when using data very close to the threshold only.

Figure 2. Number of index crimes (log scale) close to the 21st birthday.

A. One year before and after the 21st birthday	B. One month before and after the 21st birthday

[7] Murder and non-negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft and arson. There is no direct translation from index crimes to the Swedish criminal codes. See Table B1 in Appendix for the corresponding Swedish codes and our definitions.



Note: Zero denotes 21 years old at the date of the crime. The dependent variable is the logarithm of the number of sentenced individuals who committed the crime on a certain date. In (A), plus/minus 1.0 year around the threshold is used, and in (B), plus/minus 0.1 years around the threshold (roughly one month on both sides) is used. In (A), the plotted points are conditional means for a bin where the bandwidth is 0.2 year. In (B), the bandwidth is 0.02 year, roughly a week in width. The solid line in (A) is the predicted value of a fifth-degree smoother with a rectangular kernel and a bandwidth of 0.7, and in (B), the solid line is the predicted value of a local linear smoother with a rectangular kernel.

The regression results reported in Table 3 (without control variables) confirm a very large, significant and robust negative effect of sorting at the 21st birthday. For example, for the second bandwidth 0.06, where we have a relatively large size of the cluster and hence reliable standard errors, the effect is approximately 25 percent. Thus, 25 percent more crimes are committed the week before turning 21 compared to the week after. When using a wider bandwidth, the effects decrease, thereby indicating a non-linear effect. It is reassuring that the point estimates are similar when adding second- and third-order polynomials for the larger bandwidths. Thus, the regression results do line up well with the graphical evidence in Figure 2.[8] In Table B2 in Appendix B, we add the control variables birth month and year fixed effects and previous criminal history, and the results do not vary significantly compared to the results without controls.

Table 3. The 21st birthday rebate effect on the number of index crimes in percentages.

Order of polynomials	Bandwidths in share of a year measured as 365.25 days					
	0.02	**0.06**	**0.1**	**0.14**	**0.18**	**0.22**
1	-0.501**	-0.283**	-0.248***	-0.093	-0.089	-0.101*
	(0.180)	(0.106)	(0.083)	(0.076)	(0.063)	(0.059)

[8] We ran the STATA RD-package for optimal bandwidth, which gives a bandwidth of 0.117 cct. The search was made plus/minus 0.9 of the 21st birthday to exclude the 20th birthday.