Why Does Education Reduce Crime?

Brian Bell*, Rui Costa and Stephen Machin*****

January 2019

* *King's Business School and Centre for Economic Performance, London School of Economics*

** *Centre for Economic Performance, London School of Economics*

*** *Department of Economics and Centre for Economic Performance, London School of Economics*

Abstract

Prior research shows reduced criminality to be a beneficial consequence of education policies that raise the school leaving age. This paper sets up a unifying empirical framework to study reasons why crime reductions occurred due to a sequence of state-level dropout age reforms enacted between 1980 and 2010 in the United States. The reforms changed the shape of crime-age profiles, and in so doing generate both a short term incapacitation effect together with a more sustained, longer run crime reducing effect. In contrast to previous research looking at earlier US education reforms, crime reduction does not arise solely from education improvements, and so the observed longer run effect is interpreted as dynamic incapacitation. Additional evidence based on longitudinal data combined with an education reform from a different setting in Australia corroborates the finding of dynamic incapacitation underpinning education policy induced crime reduction.

Keywords: Crime age profiles; School dropout; Compulsory schooling laws.

JEL Classifications: I2; K42.

Acknowledgements

We are grateful to numerous seminar and conference participants for helpful comments and feedback on earlier drafts of this paper. This research was partly funded by the ESRC at the Centre for Economic Performance. We also thank Tony Beatton and Michael Kidd for sharing their Queensland data extract file used in the Beatton, Kidd, Machin and Sarkar (2018) study.

1. Introduction

For most crime types and in different settings, an established research finding is that education lowers criminality. In the causal crime-education literature, this finding frequently emerges in studies when school dropout age increases resulting from changes made to compulsory school leaving (CSL) laws simultaneously boost education and reduce crime.[1] What is currently less well understood is how this education policy-induced crime reduction comes about. This paper makes the argument that additional insight is gained by zooming in on the dynamics of policy-induced shifts in the age structure of criminality that occur from the enactment and implementation of CSL laws. In particular, the scope for law changes to differentially affect crime-age profiles is studied as a means to better understand empirically the reasons why education lowers crime.

To date, the impact of CSL laws on the age structure of crime has featured in two strands of research in the economics of crime. The first of these argues that the crime reduction from CSL law changes reflect an incapacitation effect which keeps children in the classroom to an older age (and so off the streets not committing crime) – see Anderson (2014) for US research on this. Other studies of juvenile crime by Jacob and Lefgren (2003) and Luallen (2006) look at teacher strikes and calendar year changes respectively to show that changes in the requirement to be in school on a particular day have effects on crime that can plausibly be considered as incapacitation.[2]

A second strand asks the question whether extra time spent in the education system induced by CSL law changes has a longer term effect on an individual's productivity. The extra education can enhance future labour market prospects, and so deterring individuals affected by

---

[1] Such law changes have been studied in a range of settings to show that a beneficial unintended consequence of them is reduced criminality - see, inter alia, Lochner and Moretti (2004), Machin, Marie and Vujic (2011), Hjalmarsson, Holmlund and Lindquist (2015) and Bell, Costa and Machin (2016).

[2] Other research has considered different forms of incapacitation, for example conscription (Galiani, Rossi and Schargrodsky, 2011), teen pregnancy (Black, Devereux and Salvanes, 2008) and violent movie screenings (Dahl and DellaVigna, 2009).

the policy change from entering a life of crime. Indeed, evidence of longer term benefits of crime reduction are provided by papers that study the causal impact of education on crime working through schooling laws for people who are old enough to have left the education system (Lochner and Moretti, 2004; Machin, Marie and Vujic, 2011; Bell, Costa and Machin, 2016).

To date, most existing research has focused on one or the other of these by studying either direct incapacitation effects or longer-term effects, rather than examining both within the same empirical setting. In this paper, we look at both in a unifying framework, and use this to draw implications from the findings about the means by which education reduces crime. This is done in practice by developing a research design that focusses in detail on the way in which CSL law changes alter the shape and structure of crime-age profiles.

This research approach generalises the way in which an education impact on crime can be empirically studied. It directly tests whether crime-age profiles adapt in the face of policy-induced changes in the compulsory school leaving age. The focus is on a more flexible specification of the crime reduced form than has generally been used by researchers in the causal crime literature. This reduced form is modified to study the changing nature of crime-age profiles in a multiple regression discontinuity framework studying US state-level changes in the compulsory school leaving age.

Evidence from the school dropout age reforms enacted in the last four decades in the United States shows that these policies have significantly altered crime-age profiles. This change in the shape is shown to be consistent with there being both a temporary incapacitation effect and a more sustained, post-incapacitation age crime reducing effect. These combine to generate sizable crime reductions from school dropout age policy reforms.[3] In contrast to the

[3] Without placing as much focus on the scope to affect crime-age profiles Chan (2012) also studies crime reduced forms using US data. A related paper, based upon Danish register data, is Landerso, Nielsen and Simonsen (2017) which studies the crime impact of reforming age of school entry.

previous research on earlier US reforms, this does not arise solely as a result of education improvements, and so the evidence of a longer run effect is interpreted as dynamic incapacitation. Additional evidence based on longitudinal data from a different setting, in Queensland, Australia, which permits the analysis of intensive and extensive crime participation by individuals corroborates the finding of dynamic incapacitation underpinning education policy-induced crime reduction.

The rest of the paper is structured as follows. Section 2 first discusses crime-age profiles and then outlines a framework for thinking about how changes in school leaving ages may have scope to shift and alter the profile structure and shape. Section 3 describes the data that are used, offers some initial descriptive data analysis on compulsory school leaving laws and presents the research design used in the empirical work contained in the paper. Section 4 reports the main results on the impact of dropout age reforms on crime-age profiles. Section 5 provides further discussion, examines evidence on the mechanisms by which dropout reforms reduce criminality and presents corroborative evidence based on Australian longitudinal data. Section 6 offers conclusions.

2. Theoretical Considerations

Crime-Age Profiles

The crime-age profile is a well-established empirical regularity. Almost two hundred years ago, Adolphe Quetelet presented evidence that crime in early nineteenth-century France peaked when individuals were in their late teens (Quetelet, 1831). Subsequent research has confirmed the existence of a strong crime-age pattern in many settings, with crime peaking in the late teens and declining quite rapidly thereafter.[4]

[4] Sullivan (2012) offers a theoretical review and Siennick and Osgood (2008) present a review of empirical work and findings.

Figure 1 shows this for US males using arrest rates, with a peak rate at age 18 and declines thereafter. In a well-known study, Hirschi and Gottfredson (1983) conjecture that crime-age profiles are broadly invariant over time and space. They suggest criminals can be identified by their lack of self-control, which is determined well before adolescence, and subsequently persists throughout life. At first sight, such a hypothesis would seem to imply that the crime-age profile should be reasonably flat. To avoid this conclusion, Hirschi and Gottfredson suggest that offenders burn-out over time – maturation – and that exposure to criminal opportunities decline as activity patterns change with age. By contrast, Sampson and Laub (1993, 2005) focus on the life-course of criminal activity and highlight how events such as family, relationships, schooling and employment change as one ages. These life cycle dynamics of crime generate the crime-age profile, with the inverse U-shape coming about from patterns of crime onset, specialisation and desistence that occur as individual's age.

A large body of evidence in criminology has tried to assess the relative merits of these different views. Greenberg (1985) presents evidence that both the peak crime age and the rate of subsequent decline differs across crime types, localities, race and gender, whilst Hansen (2003) shows that the crime-age profile differs for those who leave school at the compulsory school leaving age and those who remain in education. Further discussion and additional evidence is given in Cohen and Vila (1996), and the meta-study of Pratt and Cullen (2000).

In the economics literature, Grogger (1998) examines how changing returns to legal activity can affect the shape of the crime-age profile, whilst Lochner (2004) uses a human-capital model to show that crime should indeed peak at around the time of labour market entry. More recently, Bindler and Hjalmarsson (2017) consider convictions from 19th century London to show that there was a U-shaped trend in the average age of male convicts over the century. They suggest that increased use of prison sentences, as opposed to the death penalty and penal transportation, may have raised the average age of conviction as a result of a rise of recidivism.

Economic Models of Education Policy and Crime-Age Profiles

Since Becker (1968) formalized the economic approach to studying criminal behaviour, a variety of models have been developed in attempts to better help understand what lies behind individual criminality. Work by Ehrlich (1973), Witte (1980), and Witte and Tauchen (1994) thinks of engagement in crime as an allocation of time decision. More recently, dynamic aspects have been included to more clearly represent life-course profiles of crime. The notion of criminal capital being a substitute for human capital, which can improve an individual's prospects in the crime market vis-a-vis the labour market, has been a central feature (see, for example, Lochner, 2004, and Mocan, Billups and Overland, 2005).

Appendix B of this paper contains a simple model that incorporates this dynamic feature into the basic time allocation structure of Witte and Tauchen's framework. The model offers a useful means to show how crime-age profiles have scope to be shifted by changes in the mandatory dropout age. Individuals choose how to allocate time between the legal and illegal sectors, depending on the relative returns in each sector. Schooling constrains the amount of time individuals can allocate to either activity when they are aged below the compulsory school leaving age. The key feature is therefore that whilst younger individuals may commit some crime, because they are kept in school this acts as an incapacitation effect preventing them from engaging in as much crime as those older than the dropout age who have more available time for such activity.[5] As individuals age, the returns to legal activity rise with experience and crime rates fall.

The framework can be used to help understand some of the mechanisms behind the crime-age profiles that are observed in the data. The fact that optimal crime is decreasing in age matches the desistance stage (e.g. in the life course approach of Sampson and Laub, 1993,

[5] Despite the model focusing on individual's allocation of time independently of others individuals (peers), work by Patacchini and Zenou (2009), Deming (2011), Billings et al (2014) and Billings et al (2019) has shown evidence that peers effect the crime behavior of juveniles in school. In the context of the analysis present in this paper, peer effects would act to reinforce the effect of compulsory school laws on individuals of the same or similar age.

2005) that typically starts in the late teens or early twenties. The onset age, with an increasing crime-age profile, can be thought of in light of this framework as a case in which the net return to crime is actually increasing in very early ages given the reduced level of sanctions commonly imposed on juveniles and relatively low legal wage opportunities.

The key practical dimension is the way the simple model can be used to elaborate on how the crime-age profile may change when the minimum dropout age is increased. In the model this first implies a strengthening of the age constraint at younger ages[6] that will limit the allocation of available time to engage in criminal activities because of incapacitation. Second, there is a medium-term effect at later ages as the return to crime will be lower due to less investment being made in a criminal career (even if we make the extreme assumption that education is non-productive). A higher dropout age implies entrance into the unconstrained optimal crime time allocation at an older age, in which the dominant role is played by the net return to crime. If the latter is decreasing in age, by the time the individual is free to allocate his/her time the return will now be lower than it would under a counterfactual lower dropout age as the individual concerned faces potentially tougher sanctions and the loss of the criminal experience premium that would have otherwise been accumulated.

Figure 2 shows a simulation of the model where the minimum dropout age is increased.[7] It focusses on the age range 15-24, as will the empirical work discussed below, and considers a rise in the dropout age from 17 to 18. As previously described the effects of incapacitation (short-term) and criminal premium loss (medium-term) on time allocation are easily identified at the respective ages. These effects are congruent with the empirical crime-age profile shown in Figure 1, with the simulated reform showing strong effects of incapacitation, followed by a permanent damping down of the crime-age profile at subsequent ages.

[6] The magnitude of the effect of the higher dropout age will depend on its enforceability and on the extent of truancy.
[7] The calibration parameters of the model are provided in Appendix B.

Various statistical characterizations of the shift in the profile can be described. For example, in Figure 2 the mean offending age rises a little, going from 19.36 to 19.42, the mode from 17 to 18, and the median from 18.72 to 18.85. As we shall see, the empirical results presented later in the paper broadly match these moment changes, though one difference is that the incapacitation change is less pronounced in the data, whilst the longer-term benefits of education policy-induced crime reduction are stronger. Of course, the model is highly simplified so it should not be expected to perfectly reflect the empirical evidence – it is merely a tool for exposition that reveals possible theoretical mechanisms that may underpin shifts in crime age profiles generated by changes in the dropout age.

3. Data Description and Empirical Approach

Arrest Data

The crime data used in the analysis is provided by the FBI Uniform Crime Report (UCR) which compiles yearly arrest data by age and sex at local police enforcement agency level. This is currently available from 1974 to 2015. As most crime is committed by men and at younger ages and the compulsory school laws also apply to these ages we choose to conduct our analysis on males aged 15 to 24 years old. For these ages, arrests are reported by single year of age.

For the purpose of the analysis, the geographical level of aggregation is at the county level as in Anderson (2014). This may initially seem odd since the reforms we are focused on occur at the state-level and it would therefore seem natural to analyse the impact at that level of aggregation. The problem we face in doing that is the substantial non-reporting of arrests by individual agencies to the UCR. This non-reporting changes over time and across states. To generate annual state-level arrest data therefore requires some method of imputation.[8] To give

[8] One alternative approach is to only use the yearly observations on state-level arrest data when at least a minimum, say 95%, of the state population is reported on by the relevant agencies (see Bell, Bindler and Machin, 2018). But this generates an unbalanced panel and is therefore not appropriate within our framework.

the most extreme example, consider the CSL reform in Illinois that became effective from 2006 and increased the compulsory attendance age to 17. If we use a five-year window around the reform, only 1 of the 102 counties in Illinois consistently report arrest data every year – fortunately at least it is Chicago.

In the context of our research, using any such imputation would be inappropriate. We are seeking to exploit the discontinuity between cohorts over a short window and this requires consistent data to be available both pre- and post-reform. We therefore aggregate all agencies within a county and only include the county in our analysis if all the agencies report for all relevant years (or at most miss one year) around the reform window. Table A1 of Appendix A presents more detail on the numbers of covered and missing counties for each reform, together with information on the percentage of the state population covered. Detailed county-level population numbers by sex, age and race are matched to arrest data and adjusted to the covering standards so as to produce precise age arrest rates and demographic composition controls.[9]

Compulsory Schooling Laws

We have updated the compulsory schooling laws used in Bell, Costa and Machin (2016). Over time in empirical research using CSLs, the choice of how to measure the binding compulsory school age has been open to scrutiny and some disagreement. For example, Stephens and Yang (2014) propose a refined version of the Goldin and Katz (2008) measurement combining start age, dropout age, grade requirement and child labour laws, whereas Oreopoulos (2009) and Anderson (2014) focus only on the dropout age enacted in the laws. It is important to take account of grade exemptions as they often make up part of recent laws. Therefore, for a given birth cohort $(t-a)$ where t denotes year and a is age, the measure of binding school age in state s is then given by:

$$DA_{s,(t-a)} = min\{Dropout\ Age_{s,t(t-a)}, Grade\ Required\ to\ Dropout_{s,(t-a)}\} \qquad (1)$$

[9] Unfortunately, the UCR data does not include a racial breakdown of arrests, making it impossible to evaluate the effect of the policies along a racial dimension.

Figure 3 maps how changes in the dropout age enacted between 1980 and 2010 occurred between different states in the US. The map makes clear that some regions - such as the West South Central (Arkansas, Texas and Louisiana) and West Pacific (California and Washington) - have been more active over this period in introducing legislative changes.

Defining the precise initial cohort that is affected by these changes in compulsory schooling laws is not always as mechanical as subtracting the new dropout age from the year the law was enacted. Some of the more recent laws studied in this paper contain a degree of complexity that is significantly higher than those enacted in the first three quarters of the twentieth century that have been considered in most previous research.[10] In particular, some of the more recent law changes also feature employment exemptions, parental consents, mitigating circumstances and different effective dates. These all have some scope to add potential sources of measurement error to any attempt to code the laws.[11]

Table 1 lists the 30 laws between 1974 and 2010 that are studied in the empirical analysis, together with detail on various relevant features of them including the particular dropout age change and new dropout age, and whether they feature exemptions by school grade.

Research Design

We study crime evolution before and after changes in compulsory school leaving laws based on arrest rates by individual year of age a for men in county c located in state s in time period t. A baseline crime reduced form is as follows:

$$Arrest_{acst} = \beta Reform_{s(t-a)} + \gamma X_{acst} + \alpha_a + \alpha_c + \alpha_t + \varepsilon_{acst} \qquad (2)$$

where *Arrest* is the log arrest rate, *Reform* is a dummy variable (to begin with) indicating whether or not there was a dropout age reform affecting birth cohort $(t-a)$ in state s, X is a

[10] See Goldin and Katz (2014) for a careful discussion about the evolution and complexity of the compulsory schooling laws in the United States.

[11] When the time lapse between enactment and effective date of the law is more than 9 months, the changes have been cross-validated empirically by analysing the data around the potential discontinuity to assert the binding date and cohorts affected.

set of county level controls and α_a, α_c and α_t respectively are fixed effects for age, county (also subsuming state fixed effects) and time, and ε is the equation error term.

When structured as in equation (2), this crime reduced form is essentially the one that has been estimated in much of the existing work examining the causal impact of schooling laws by pooling together data across states which did and did not change their schooling laws over time. This has been done for a number of outcomes of interest: for wages, see for example, Acemoglu and Angrist (2001) and Oreoupoulos (2009); for crime, see Lochner and Moretti (2004) and Bell, Costa and Machin (2016); and for a range of outcomes probing robustness of the approach in detail see Stephens and Yang (2014).

We begin by presenting estimates this way for comparison, but then move on to treat each of the reforms listed in Table 1 as a separate regression discontinuity (RD) around which we can examine what happens to crime before and after the reform takes place. To motivate the RD analysis, Figure 4 shows the discontinuity for the arrest rate for the pooled reforms (centred at t = 0). It shows a significant reduction in the arrest rate of 4 arrests per 1000 population (or 4.6 percent of the pre-reform mean of 0.086) relative to the earlier cohorts who were unaffected by the reform.

More formally, for a given school dropout reform in a particular state, the following specification for different time windows (w) around the dropout age policy changes can be estimated:

$$Arrest_{acst} = \beta Reform_{s(t-a)} + f_s(t-a) + \gamma_s X_{acst} + \alpha_{sa} + \alpha_c + \alpha_{st} + \varepsilon_{acst}$$
$$for \quad (t-a) - w \leq t-a \leq (t-a) + w, \qquad w = \{5, 7, 10\} \tag{3}$$

where the forcing variable in the classic RD design (see Imbens and Lemieux, 2008; Lee and Lemieux, 2010) is birth cohort $(t-a)$ and the general function $f_s(.)$ allows for various functional forms that can be adopted for estimation.

To study the manner in which the policy change induces shifts in crime-age profiles, we further amend the RD design, allowing heterogeneity by age in the policy reform. This is precisely what the theoretical framework we described in Section 2 above argued needs to be done to see: a) how crime-age profiles may alter for different dropout ages; and b) to pin down the nature of incapacitation effects that occur when young people stay in school to later ages.

In practice, we estimate separate before/after policy effects in the crime reduced form for each age fixed effect, so that a more general estimating equation follows:

$$Arrest_{acst} = \theta_a(Reform_{s(t-a)}) + f_s(t-a) + \gamma_s X_{acst} + \alpha_{sa} + \alpha_c + \alpha_{st} + \varepsilon_{acst} \tag{4}$$

$$\left.\frac{\partial Arrest_{acst}}{\partial Age_a}\right|_{a=j} = [\theta_j \times Reform_{s(t-a)}] + \alpha_{sj}$$

where the partial derivative shows the impact of the reform for age j (j = 15, 16….24).

Controls

We match in a set of control measures that according to existing evidence (e.g. Levitt, 1997; Card and Krueger, 1992) may relate to both arrests and educational attainment and progress. Some of Card and Krueger's (1992) school quality measures (pupil-teacher ratios, average teacher salary, number of schools) were updated at county-level using Common Core Date (CCD) data. Police numbers were recovered from the FBI Law Enforcement Officers Killed and Assaulted (LEOKA) database and socio-demographic indicators were collected from the Local Area Personal Income (LAPI) data from Bureau of Economic Analysis. More details are provided in the Appendix A.

4. Crime-Age Profiles and Dropout Age

Baseline Estimates of Crime Reduced Forms

Although the primary focus of the paper is on the crime-age profile, the empirical analysis begins by estimating the effect of the dropout reforms on the overall arrest rate. This is both because an overall effect is a necessary condition for the reforms to also alter the shape of the profile – since it is hard to think how the reform could increase the crime rate for those affected at any point in the profile – and because the prior literature has focused on such reduced forms and so it is useful to demonstrate that the reforms considered in this paper, which are more recent, generate similar effects as those examined previously.

Table 2 reports the baseline estimates of the crime reduced form. At this stage, all reforms across time and space are treated as equivalent and thus have a single indicator for reform. Later in this section separate estimates for each type of reform are presented (e.g. an increase in the school-leaving age from 16 to 17, 16 to 18, 17 to 18 etc.). It turns out that the results are robust to allowing each type of reform to have separate estimates and it is therefore more straightforward to start with to present estimates for the weighted-average effect of all types of reforms, which is what Table 2 does. All standard errors are clustered at the state-cohort level, which is the dimension along which each reform occurs.

Implicit in the discussion thus far has been the assumption that each reform can be considered as exogenous to the parameters of interest. Crucially, we assume that school-leaving reforms where not instigated at a particular time and in a particular state in response to crime concerns related to the precise cohorts that would be affected by the reform. This seems unlikely to us because crime outcomes are generally viewed as an unintended consequence of school leaving age reforms. However, one way of assessing this is to consider balancing tests that compare observables between cohorts on either side of the discontinuity that the reform creates. Such tests are presented in Appendix Table A2 and there is no evidence to suggest any pattern around the discontinuity.

The first column in Table 2 presents estimates that simply turn on a reform dummy for particular cohorts in particular states using the dating provided in Table 1. This is therefore equivalent to the typical type of estimates that are presented in the reduced-form economics of crime literature such as Lochner and Moretti (2004) and given as equation (2) above. They do not explicitly take advantage of the discontinuity that each reform generates. The impact of the reform is significantly negative at the 1% level, and as such shows a strong crime reducing effect from higher dropout ages.[12]

The preferred estimates are those in the subsequent columns of the Table that are equivalent to equation (3) above and exploit the discontinuity across cohorts. They include a full set of state interactions with all the control variables and estimates are presented for different parametric forms for the running variable and for the length of the window around which we estimate the discontinuity. The first three estimates use a 10-year window around each discontinuity and each allows the running variable to have different parametric form on either side of the reform. It matters little what the functional form for the running variable is, so we proceed from now on with using a simple linear function.[13]

The discontinuity estimates are roughly half the size of the estimates presented in column (1), but remain strongly significant. In the final two columns we experiment with a narrower window around the discontinuity. Again, there is not much to choose between these various specifications, so we proceed with a 5-year window on the basis that this more tightly focuses on the discontinuity.[14] This final estimate suggests a 6% fall in arrest rates for these young adults as a result of the dropout reform.

---

[12] We have also estimated column (1) allowing for quadratic or cubic terms in the running variable and these produce coefficients very similar to the -0.099 reported in column (1): to be precise -0.086 and -0.091 respectively.

[13] All subsequent results are robust to using a quadratic or cubic function for the running variable, though such forms are computational feasible only for the longer windows around the discontinuity.

[14] The results in Table 2 follow the standard approach in the RDD literature of assuming that the reform is exogenous. We presented balancing tests on observables in Table A2 that are supportive of this assumption, but we recognise that this is a far from exhaustive list of observables and in any case one can never prove that all unobservables are balanced. A key concern may be that states decided to implement a reform at exactly the time it might have the most beneficial effect on crime. To examine this further we adopt a synthetic control approach and essentially combine the RDD design with a diff-in-diff approach. For each reform, we consider all other states as potential controls and use a five-year window prior to the reform to generate a synthetic control. Consider for example the reform in California that raised the leaving age in 1987. We use the average arrest rate for 15-24 year olds from 1982-1986 and match on arrest rate, percent black, percent young, personal income per head, employment-population rate and police officers per head. This then generates a set of weights for all other states that best matches the California arrest rate for 15-24 year olds in the pre-reform period. If we re-estimate the final column of Table 2 using this approach we obtain a coefficient estimate (and associated standard error) on the reform of -0.040 (0.007).

Different Types of Reform

The estimates presented in Table 2 pooled all the reform types together to estimate an average effect across the 30 reforms studied. In Table 3 estimates are reported separately for the 5 year window specifications for the five different reform types: the 29 reforms that featured an increase, either from 16 to 17, 17 to 18, 16 to 18, or any other increase; and the one reform in Texas in 1985 where the rewriting of the law lowered the dropout age from 17 to 16.

Column (1) begins by reporting an estimate for the weighted-average of the 29 reform types that involved an age increase. The estimated reduction in the Log(Arrest Rate) is -0.062 which, of course, is very similar to the column (6) specification of Table 2 of -0.060, although is slightly more negative. In column (2), the Texas increase attracts a significant positive coefficient of 0.090. That the effect is positive in the case of a dropout age reduction offers a useful robustness test in line with crime reducing impacts of dropout age increases (and the opposite for this single case of lower dropout age) for the analysis.

Estimates for the four different groups of dropout age increases are presented in columns (3) through (6) of Table 3. There were respectively 8 reforms raising the dropout age from 16-17, 6 from 17 to 18, 8 from 16 to 18 and 8 in the catch-all 'Other' group.[15] In all four groups, there is a significant crime reduction effect, and the estimates are in a quite tight range between -0.041 and -0.071.

The use of county-level panel data means it is also possible to estimate the discontinuity for each reform separately. Estimates produced from doing this are presented in Table A3, but it is easier to visualise the various estimates as they are presented in Figure 5. Each point represents a separate reform labelled along the horizontal axis, and 95% confidence bands for each estimate are shown. Only one of the 30 reforms generates a significantly positive effect

[15] The reforms in the 'Other' group are listed in the notes to Table 3.

on arrest rates – the 1985 Texas reform. Of the other 29 reforms, 16 are significantly negative, and all but 4 have a negative estimate.

Different Crime Types

Table 4 present estimates for the 29 pooled reforms involving age increases that distinguish between different crime types (total, violent, property and drug arrests).[16] It also presents estimates that differ by two broad age groups (15-18 and 19-24). This second set of estimates offers a first indication as to whether the crime-age profile is altered by the reforms. The results of the Table suggest a fairly consistent pattern across crime types, though the effect is larger in magnitude (in absolute terms) for drug arrests than the other types of crime. Focusing on the age groups, in all cases the effect is larger for those contemporaneously affected by the reforms (i.e. in the younger 15-18 age range) than for those who were affected in the past. However this latter group still experiences a significantly lower arrest rate as a result of the reform that they were subject to when at school.[17]

The Impact on Crime-Age Profiles

Having demonstrated the crime-reducing effect of the reforms overall, and first identified some variation by broad age group, the focus is now directly placed on the effect on the entire crime-age profile, with an aim of studying the extent to which its shape may change in response to the education reforms. To begin, the specification for the 5 year window is generalised to have different reform effects at each single age – corresponding to equation (3) above. This then allows examination of the key question of the paper – can policy reforms alter the entire shape of the crime-age profile?

Consistent with the theoretical simulation presented in section 2, the results reported in Table 5 show that reforms have the largest effect for those directly incapacitated as a result of

[16] For the remainder of the empirical analysis, the focus is placed only upon the 29 dropout age increases, excluding the Texas increase. Results are however robust to including the increase.
[17] For the total arrests specification in column (1) of the Table, the null hypothesis that the two age groups have the same arrest response to the reform can be rejected, with a p-value of 0.004.

school attendance. However, they also show a significantly negative effect for later age groups that are not incapacitated in school as a result of the reform. These two findings emerge to varying degrees for different crime types.

Figure 6 shows the estimates, with 95% confidence bands, for each crime type. To highlight the effect on the crime-age profile overall, Figure 7 shows the estimated profiles pre- and post-reform by crime type. It is clear how the reforms are reducing crime at all stages of the life-cycle, though generally more heavily in the early years. Thus there is evidence of both a temporary incapacitation effect – when the young people are locked up in school – and a longer term crime reducing effect.

Closer inspection of Figure 7 does reveal some differences in the balance between crime reductions at younger and older ages across crime types. When pooled, the total crime figure shows larger incapacitation effects. The same is true for property and drug crimes, and in the case of the former there is little in the way of an effect at older post-incapacitation ages. For violent crimes, the opposite holds: little in the way of incapacitation, but some crime reduction at older ages.

As we noted in the discussion of the theoretical model that in part motivates the empirical work, we can also look at how reforms affect various moments of the crime-age profile. Table A4 in the Appendix presents estimates for total crime and for the three sub-categories. All the moments are significantly shifted by the reform. The measures of central tendency (mean, mode) are shifted to the right as predicted and the standard deviation falls - thus the crime-age profile becomes more compressed after the dropout age is raised.

5. Mechanisms and Discussion

The reported results considered so far show a strong negative effect on arrest rates from school leaving age reforms. This operates both at the time an individual's behaviour is directly impacted by the policy, and also in subsequent years when they are not. The former effect is likely to be a result of incapacitation – a young person is constrained to remain in school, so they have less free time to allocate to crime. In this section, some potential mechanisms that may explain the latter longer run effect are considered.

Education and Employment Outcomes

There is by now a large literature that examines the causal effect of education on crime.[18] A natural interpretation of the dropout reform reducing criminality is that, in addition to the direct incapacitation effect that occurs from requiring students to remain in school for an additional year, the additional year also generates a productive educational benefit for those on the margin of criminal behaviour. This then raises their human capital, wages and employment and reduces the probability of committing crime in the future. This would be consistent with the theoretical model outlined in Section 2, and with the earlier US research studying the impact of the earlier compulsory school leaving reforms from the 1960s, 1970s and 1980s.[19]

To assess this explanation of the results, the empirical connection between the reforms and different measures of education and work are considered. First of all, looking at the incapacitation side of things, we explore whether school attendance did in fact increase by utilising Current Population Survey (CPS) data on 16-18 year olds between 1974 and 2015 (see the Data Appendix for more details). Panel A of Table 6 shows the estimates, structured in the same way as the earlier baseline results for arrests. There is significant evidence of incapacitation, with the 5-year window specification in column (4) showing a 5 percentage point rise, or a 6.6 percent increase relative to the pre-reform mean. This reaffirms that school incapacitation effects were a key dimension of the dropout age reforms.

[18] Many of these studies were cited earlier, but see also the review in Lochner (2011).

[19] For crime, see Lochner and Moretti (2004). For reviews of the sizable bodies of research on wage effects see Card (1999) and Oreoupoulos (2009). For a host of other non-wage outcomes variables – including health, voting behaviour and life satisfaction - see Oreopoulos and Salvanes (2011).

To explore what might lie behind the longer run crime reducing effects, the remainder of the Table reports results for education and job related outcomes for older individuals aged 19-60 in the American Community Survey (ACS) from 2006 onwards.[20] The outcomes are high-school dropout rates, whether or not an individual was in education or work, and log weekly real wages. Whilst there are statistically significant effects in the expected direction for a number of the specifications, the estimates are relatively small in magnitude. They do uncover education improvements that followed from dropout age reform, and an increased likelihood of being in school or work, but the effects are small – relative to the pre-reform mean, they respectively correspond to a 5.3 percent fall in high school dropout and a 0.4 percent increase in the likelihood of being in education or work. Unlike in the previous work on earlier reforms (e.g. Acemoglu and Angrist, 2001; Card, 1999), there is essentially no effect on wages in any specification.[21]

Interpretation

The positive effects of the reforms on economic and education outcomes are therefore modest, certainly in comparison to Lochner and Moretti (2004) who find education estimates that are quite a lot bigger than those reported in Table 6. However, our previous work (Bell, Costa and Machin, 2016) has demonstrated that the most recent reforms to compulsory schooling laws have substantially weaker effects on educational attainment than estimates identified using changes from dropout age reforms in the 1950s and 1960s. This is in line with the notion that the group of compliers – e.g. those that obtain a high-school diploma when the reform occurs who would not have done previously – are a smaller percentage of the eligible population for the period studied in this paper.

[20] ACS data is used because it is annual data that can be used to study the reforms across pooled birth cohorts. See the Appendix A for more detail.
[21] Lack of a wage effect from dropout age reforms is not unique to this paper. Pischke and Von Wachter (2008), for example, report no wage gains from German compulsory school leaving age reforms.

This interpretation makes sense as the high school dropout rate for those aged 16-24 fell from 27.2 percent in 1960 in Lochner and Moretti's data to 5.9 percent in 2015. This shrinks the group of potential compliers by a lot and makes it more likely that the dropouts are a hard core of individuals for whom such reforms are unlikely to have any effect (i.e. a higher share of never takers). This does not mean that there is no effect – after all a 0.5% percentage point (5.3 percent) fall in the dropout rate will certainly affect the criminal margin for some individuals. But it seems unlikely that the size of this change in educational attainment can explain the entire 3-4 percent reduction in arrest rates that we observe for 19-24 year olds.

If the reforms do not substantively boost educational attainment or wages, what other mechanisms can explain the lower crime rate further along the age distribution when direct incapacitation effects cannot be operative? One possibility is dynamic incapacitation. This is where the direct incapacitation from being kept in the school classroom causes changes that affect future crime participation, independent of whether there is any educational value to the incapacitation. For example, suppose that being kept in school during the day prevents an individual from being on a street corner dealing drugs. This reduces arrests at the time, but also potentially means that the individual leaves school without the criminal record they would otherwise have had. They now find it easier to pursue a life as a law-abiding citizen. Put another way, some individuals' crime onset is stopped by incapacitation and they never commit crime subsequently. For other individuals who may already have committed crime, the incapacitation reduces their crime intensity in the incapacitation period and this persists as they get older – the reform acts to reduce their criminal capital accumulation as compared to the counterfactual of no reform.

Other evidence also suggests that interventions at this crucial period of potential criminal development can alter the life course of criminality. Bell, Bindler and Machin (2018), for example, show that leaving high school in a recession can significantly increase the affected cohorts' arrest rates well into adult life – in a sense the recession generates crime scars that persist beyond the period of direct impact. In a different setting, looking at random assignment of judges in Chicago to identify the causal effects of juvenile incarceration, Aizer and Doyle (2015) show that incarceration both reduces the probability of high-school graduation and increases the likelihood of subsequent incarceration as an adult. Both these studies are consistent with a finding of dynamic incapacitation effects.

Evidence Based on Individual Panel Data

To further explore these effects, we also analysed a different education reform enacted in Queensland, Australia. The reason for presenting this evidence in addition to the US evidence is twofold. First of all, it shows very similar effects in a different setting. Second, the Queensland data has one clear advantage over the US cohort data as it follows the same individuals through time. Its disadvantage is that we can study only a single reform, rather than the multiple reforms studied in depth earlier in this paper.

In 2006 Queensland implemented an "Earning or Learning" reform that required all individuals aged 16 and 17 who previously could have left compulsory education at age 15 to participate in some form of education, training or work. The reform's impact on individual offending behaviour has been studied in depth by Beatton, Kidd, Machin and Sarkar (2018), but it is interesting in the context of the current paper to use their longitudinal data on all state school children matched to crime records when at school and up to age 23 after they have left school to further probe static and dynamic incapacitation effects.[22] The data on criminal behaviour refers to alleged criminal offences – either being arrested, cautioned or having a warrant for apprehension issued – and is therefore similar to the US arrest data.

[22] More detail on the Earning and Learning reform and on the data are given in Appendix A (sections A9 and A10).

To begin with, we examined the overall impact of the education reform on criminal behaviour, again exploiting the discontinuity between different cohorts. The first column of Table 7 shows that the overall effect is an 11% (coefficient of -0.006 on a pre-reform mean of 0.055) reduction in the treated cohorts' offending rate. This is somewhat larger than the average estimates for the US presented earlier – though not out of line with some of the more sizable individual state estimates. The next two columns of Table 7 shows that those affected by the reform are 13% less likely to offend whilst incapacitated as a result of the reform (age 15-17) and 10% less likely to offend in the post-incapacitation period (age 18-23). So we see a somewhat larger reduction in offending during the incapacitation period, as with the US results. In Figure 8 we present the crime-age profile shifts for the Australian case. Much like its US counterpart (Figure 7), we observe lower offending rates throughout the crime-age profiles for the different types of crime, again with larger effects occurring during the incapacitation period.

The longitudinal dimension of the Queensland data permits a more in-depth analysis of the dynamics of criminal behaviour following the education reform. In particular, we can study multiple offending behaviour for the same individuals over time and, in doing so, explore the extensive (whether individuals ever offend) and intensive (how much offending individuals do conditional upon being an offender) margins of crime. We start by exploring the impact of the reform on the extensive margin by looking at the individual's probability of ever being accused of committing a crime. Column (4) of Table 7 shows that the reform reduces this lifetime (at least up to age 23) probability by 18%. In other words, a key dimension along which these educational reforms change crime is by removing a set of individuals from the pool who ever commit a crime.

In columns (5) and (6) we return to consider the question of dynamic incapacitation, now using the longitudinal aspects of the data. We explore the change in the probability of offending for those aged 18-23, conditional on whether or not they have been accused of a crime when at school. There is a very substantial 24% decline in this probability for those who had no school-aged prior offence. We argue that this is the dynamic incapacitation effect at work – they have avoided trouble during their school age years and this then keeps them on the right track later in life.

In column (6) we see that for those who did not avoid being in trouble at a younger age, the reforms have less effect at older ages – the probability of reoffending after leaving education falls by 6.6%. So the key dimension of these reforms is to significantly reduce the number of individuals who are ever accused of committing a crime – and much of this reduction occurs among ages that are not directly incapacitated by the reforms.

In columns (7) and (8) we look at the intensive margin (as measured by number of total crimes accused of) for the same two sub-populations examined in columns (5) and (6). For both groups we see a reduction in the number of offences, suggesting that the reforms reduce the accumulation of criminal capital that enhances the productivity of criminals. Interestingly however the declines are broadly similar for the two groups (relative to their pre-reform means).

Shifts in the age of offending onset for individuals who are treated and not treated by the education reform are shown from hazard function estimates presented in Figure 9.[23] It shows a sizeable decrease in the probability of offending conditional on no prior offending across ages, with a lowering of magnitude for later ages. The estimates converge by age 23 – this should not be surprising since at that age the probability of being accused of an offence having never previously been accused of one by the age of 22 is extremely low. It would be odd if the reform affected marginal conditional probabilities so far away. But it is important to remember that this is not to say that the reform does not reduce the probability that an individual will have been accused of an offence by age 23 – our results show it certainly does.

[23] The hazard function describes the probability of first-time offending at given age a conditional on no prior offending record until age a, $\Pr(A = a \mid A < a)$, which is treated as a function of covariates year, reform and age that in this application stands for "time" as defined in classical survival analysis.

Cost-Benefit Calculation

In Table 8 we report back of the envelope cost-benefit calculations for the US educational reforms studied in this paper. Based on our earlier results we calculate the estimated foregone costs of crime as benefits using the same methodology as Lochner and Moretti (2004), and we additionally incorporate the costs of keeping students in high school for the additional school years. As reported by the cost-benefit ratio we find that by age 18 the policy almost breaks even, with 0.94 dollars being recovered for every dollar spent. This result is representative of the economic return to the incapacitation effect estimated in our analysis. Perhaps more interestingly, we conclude that by taking into account the effects of dynamic incapacitation for older ages (until age 24) the cost-benefit ratio shows a return of almost 2 (1.9) dollars per dollar spent on the policy. Whilst these estimates are only suggestive, they highlight how important the longer term effects of the policy are to an evaluation of the cost effectiveness of such reforms, in an environment in which there appear to be scant productivity-enhancing effects from the reforms.

6. Conclusions

By developing a more general way of modelling the impact of school dropout age reforms on crime, this paper presents the first evidence to show that compulsory schooling law reforms not only affect the overall level of crime, but they also re-shape crime-age profiles. This enables a better understanding of the reasons why education causally reduces crime, guided by the empirical observation that there is heterogeneity connected to age in the way in which CSLs reduce crime.

Focusing on changes in laws across US states since the 1980s, a multiple regression discontinuity framework is used to show that arrest rates for young men fall by around 6% on average as a result of these reforms. Whilst there is a larger negative effect for those in the age group that are directly constrained by the reforms – they are kept in school and incapacitated, hence having less time to devote to potential criminal activity – there is also a significant negative effect for those who are no longer directly constrained. The results are consistent with there being *both* an incapacitation effect *and* a longer-term beneficial crime reducing effect.

This longer run effect is interpreted as a dynamic incapacitation effect because further evidence we present shows that these same reforms at best had very modest effects on average educational attainment and wages, though somewhat more substantial effects on high-school dropout. These results are further corroborated by including a study of the intensive and extensive margins of age-related crime reductions due to similar education reforms exploiting longitudinal data on youths in Queensland, Australia. The key finding emerging from analysis of these data is that dropout age reforms both reduce the probability of offending whilst incapacitated in school but also very substantially reduce the probability that these individuals begin to offend after they leave school. The overall implication is that school dropout age reforms can have longer run crime reducing effects from incapacitation than those occurring just in the incapacitation period itself. This is important both from the perspective of calculating the social benefits that crime reduction due to CSLs generates and for generating a better understanding of how individual crime dynamics evolve over the life course.

References

Acemoglu, D. and Autor, D., (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings", in Ashenfelter, O. and Card, D. (eds.) *Handbook of Labor Economics*, Volume 4B, Amsterdam: Elsevier.

Acemoglu, D. and Angrist, J. (2001) "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws", in Bernanke, B. and Rogoff, K. (eds.) *NBER Macroeconomics Annual* 2000, Cambridge MA: MIT Press.

Aizer, A., and Doyle, J. (2015) "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges", *Quarterly Journal of Economics,* 130, 759-804.

Anderson, D. (2014) "In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime", *Review of Economics and Statistics,* 96, 318-31.

Beatton, T., Kidd, M., Machin, S. and Sarkar, D. (2018) "Larrikin Youth: New Evidence on Crime and Schooling", *Labour Economics*, 52, 149-59.

Becker, G. (1968) "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, 76, 169-217.

Bell, B., Bindler, A., and Machin, S. (2018) "Crime Scars: Recessions and the Making of Career Criminals", *Review of Economics and Statistics,* 100, 392-404.

Bell, B., Costa, R. and Machin, S. (2016) "Crime, Compulsory Schooling Laws and Education", *Economics of Education Review*, 54, 214-26.

Billings, S., Deming, D. and Rockoff, J. (2014) "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg", *Quarterly Journal of Economics*, 129, 435–76.

Billings, S., Deming, D. and Ross., S. (2019). "Partners in Crime", *American Economic Journal: Applied Economics*, 11, 126-50.

Bindler, A., and Hjalmarsson, R. (2017) "Prisons, Recidivism and the Age-Crime Profile", *Economics Letters,* 152, 46-49.

Black, S., Devereux, P. and Salvanes, K. (2008) "Staying in the Classroom and out of the Maternity Ward? The Effect of Compulsory Schooling Laws on Teenage Births", *Economic Journal*, 118, 1025-54.

Card, D. and Krueger, A. (1992) "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", *Journal of Political Economy*, 100, 1-40

Chan, S. (2012) "Not Until You're Older: The Minimum Dropout Age and the Crime-Age Profile", Working Paper, Unpublished doctoral dissertation, Boston College, Chestnut Hill, MA.

Cohen, L., and Vila, B. (1996) "Self-Control and Social Control: An Exposition of the Gottfredson-Hirschi/Sampson-Laub Debate". *Studies on Crime and Crime Prevention*, 5, 125-150.

Dahl, G. and DellaVigna. S. (2009) "Does Movie Violence Increase Violent Crime?", *Quarterly Journal of Economics*, 124, 677-734.

Deming, D. (2011) "Better Schools, Less Crime?", *Quarterly Journal of Economics*, 126, 2063–2115.

Ehrlich, I. (1973) "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation", *Journal of Political Economy*, 81, 521-565.

Galiani, S., Rossi, M., and Schargrodsky, E. (2011) "Conscription and Crime: Evidence From the Argentine Draft Lottery", *American Economic Journal: Applied Economics*, 3, 119-36.

Goldin, C. and Katz, L. (2008) *The Race Between Education and Technology*, Cambridge, Massachusetts: Harvard University Press

Gottfredson, M. (1985) ''Youth Employment, Crime, and Schooling'', *Developmental Psychology*, 21, 419–432.

Greenberg, D. (1985) "Age, Crime and Social Explanation", *American Journal of Sociology*, 91, 1-21.

Grogger, J. (1998) "Market Wages and Youth Crime", *Journal of Labor Economics,* 16, 756-91.

Hansen, K. (2003) "Education and the Crime-Age Profile", *British Journal of Criminology*, 43, 141-168.

Hirschi, T. and Gottfredsson, M. (1983) "Age and the Explanation of Crime", *American Journal of Sociology*, 89, 552-584.

Hjalmarsson, R., Holmlund, H., and Lindquist, M. (2015) "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data", *The Economic Journal*, 125, 1290-1326.

Imbens, G. and Lemieux, T. (2008) "Regression Discontinuity Designs: A Guide to Practice", *Journal of Econometrics*, 142, 615-635.

Jacob, B. and Lefgren, L. (2003) "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration and Juvenile Crime", *American Economic Review,* 93, 1560-77.

Landersø, R., Nielsen, H., and Simonsen, M. (2017) "School Starting Age and the Crime-Age Profile", *Economic Journal*, 127, 1096-1118.

Lee, D. and Lemieux, T. (2010) "Regression Discontinuity Designs in Economics", *Journal of Economic Literature*, 48, 281-355.

Levitt, S. (1997) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime", *American Economic Review*, 87, 270-90

Levitt, S., (1998) "Juvenile Crime and Punishment", *Journal of Political Economy*, 106, 1156-1185

Lochner, L. (2004) "Education, Work and Crime: A Human Capital Approach", *International Economic Review,* 45, 811-43.

Lochner, L. (2011) Non-Production Benefits of Education: Crime, Health, and Good Citizenship, Hanushek E., S. Machin and L. Woessmann (eds.) *Handbook of the Economics of Education*, Volume 4, Amsterdam: Elsevier.

Lochner, L. and Moretti, E. (2004) "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-reports", *American Economic Review,* 94, 155-89.

Luallen, J. (2006) "School's Out… Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes", *Journal of Urban Economics,* 59, 75-103.

Machin, S., Marie, O. and Vujic, S. (2011) "The Crime Reducing Effect of Education", *Economic Journal*, 121, 463-84.

McAdams, J. (2016) "The Effect of School Starting Age Policy on Crime: Evidence from U.S. Microdata", *Economics of Education Review*, 54, 227-41.

Mocan, N., Billups, S. and Overland, J. (2005) "A Dynamic Model of Differential Human Capital and Criminal Activity", *Economica*, 72, 655–681.

Oreopoulos, P. (2009) "Would More Compulsory Schooling Help Disadvantaged Youth? Evidence from Recent Changes to School-Leaving Laws", in J. Gruber (Ed.) *The Problems of Disadvantaged Youth: An Economic Perspective*, Chicago: University of Chicago Press for NBER.

Oreopoulos, P. and Salvanes. K. (2011) "Priceless: The Nonpecuniary Benefits of Schooling", *Journal of Economic Perspectives*, 25, 159-84.

Quetelet, A. (1831) [1984] *Research on the Propensity for Crime at Different Ages*, translated and introduced by Sawyer F. Sylvester. Cincinnati: Anderson.

Patacchini, E. and Zenou, Y. (2009) "Juvenile Delinquency and Conformism", *Journal of Law, Economics, and Organization,* 28, 1–31.

Pischke, J-S. and von Wachter, T. (2008) "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation", *Review of Economics and Statistics*, 90, 592–598

Pratt, T. and Cullen, F. (2000) "The Empirical Status of Gottfredson and Hirschi's General Theory of Crime: A Meta-Analysis", *Criminology*, 38, 931–964.

Sampson, R. and Laub, J. (1993) *Crime in the Making: Pathways and Turning Points Through Life*, Cambridge, Mass: Harvard University Press.

Sampson, R. and Laub, J. (2005) "A Life-Course View of the Development of Crime", *ANNALS of the American Academy of Political and Social Science*, 602, 12-45 and 73-79.

Siennick S. and Osgood, D. (2008) "A Review of Research on the Impact on Crime of Transitions to Adult Roles", *The Long View of Crime: A Synthesis of Longitudinal Research*. Springer, New York, 161–190

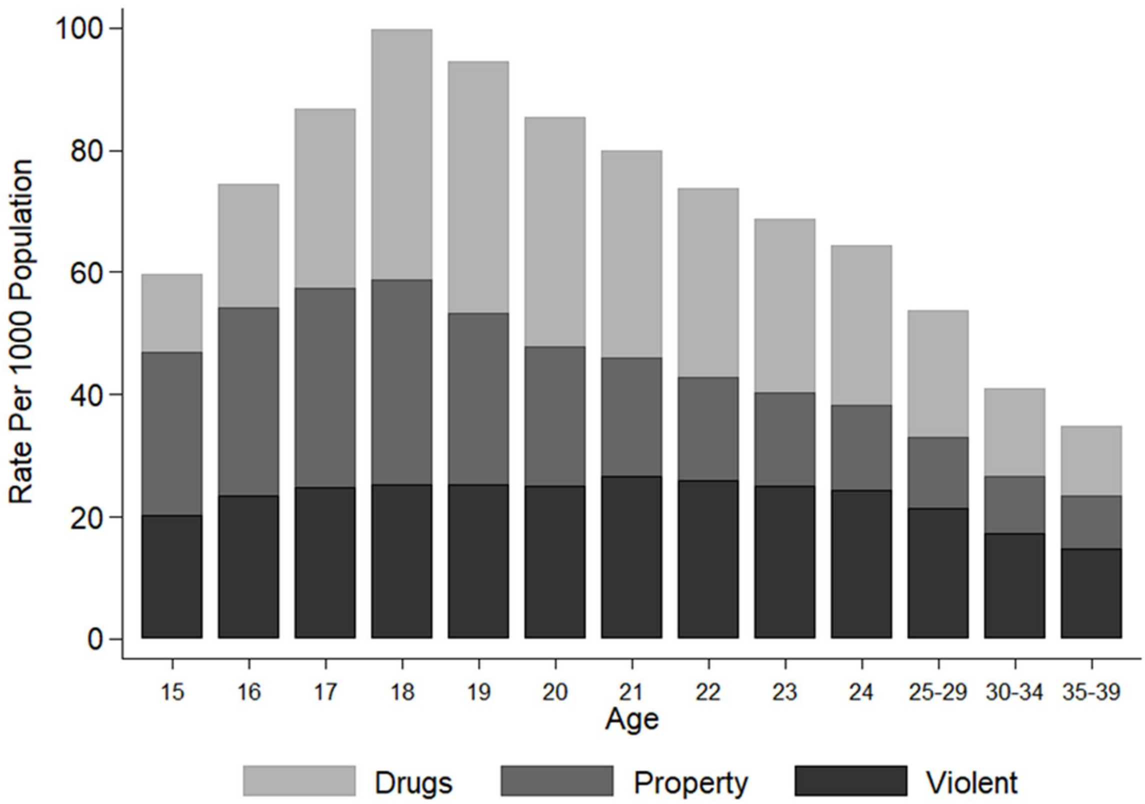
Stephens, M. and Yang, D-Y. (2014) "Compulsory Education and the Benefits of Schooling", *American Economic Review,* 104, 1777-92.

Sullivan, C. (2012) "Change in Offending across the Life Course", in Cullen, F. and Wilcox, P. (eds.) *Oxford Handbook of Criminological Theory*, Oxford University Press

Witte, A. (1980) "Estimating the Economic Model of Crime With Individual Data", *Quarterly Journal of Economics*, 94, 57–84.

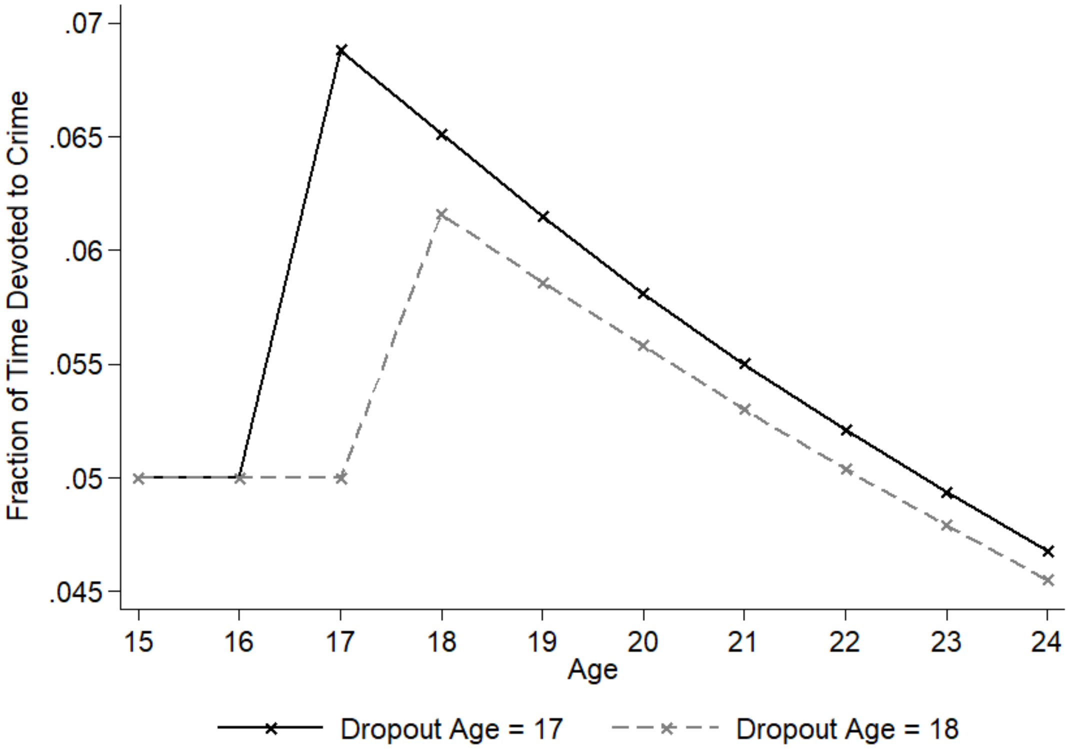
Witte, A. and Tauchen, H. (1994) ''Work and Crime: An Exploration Using Panel Data", *Public Finance*, 49, 155–167.

Figure 1: Male Offender Rates by Age, US



Notes: Male arrest rates by age, calculated for years 2000-2010 from UCR data. Only agencies reporting all years of the time period covered are included. The composition of the different type of crime is covered in the Data Appendix.

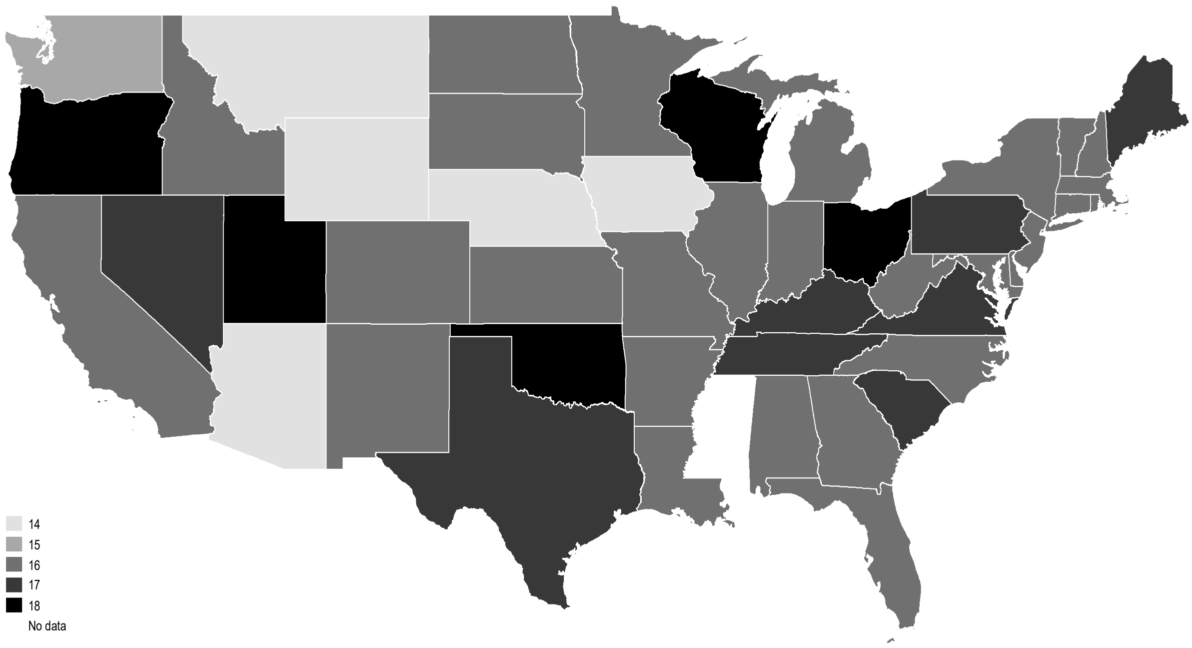
Figure 2: Model Simulation



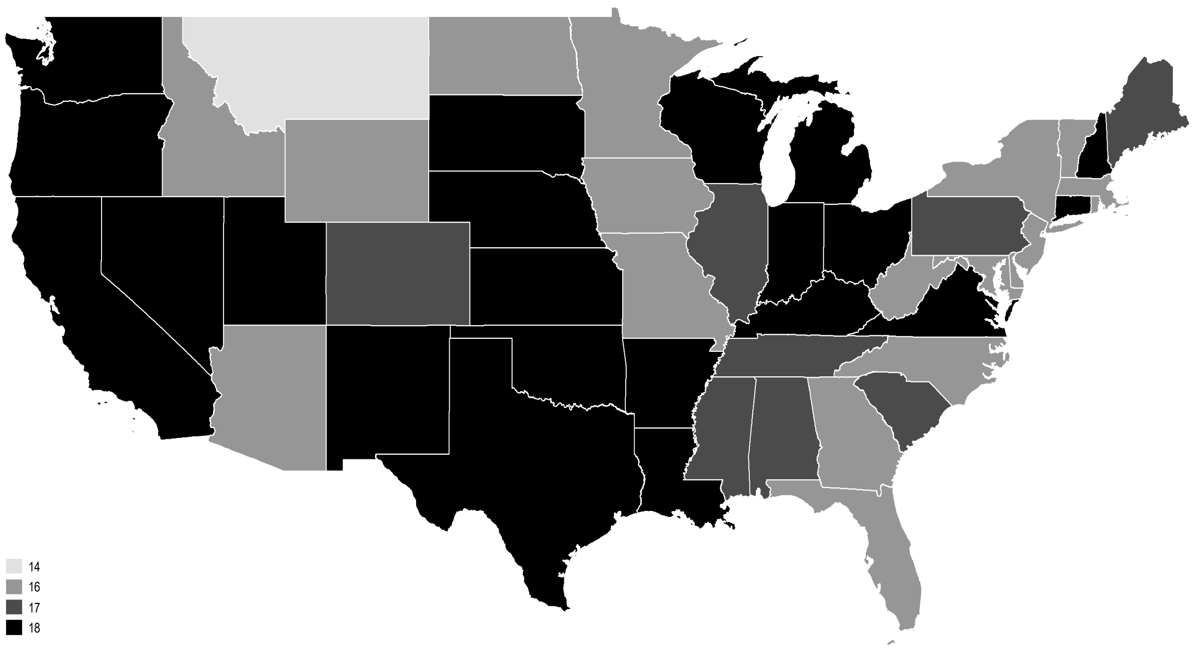
Notes: Details on the model simulation are presented in Appendix B. The ages 15 to 24 are those covered in the empirical analysis and the fraction of time devoted to crime is rescaled by a constant factor to (broadly) approximate the arrest rates observed in the data.

Figure 3: State Dropout Ages, 1980 and 2010

1980

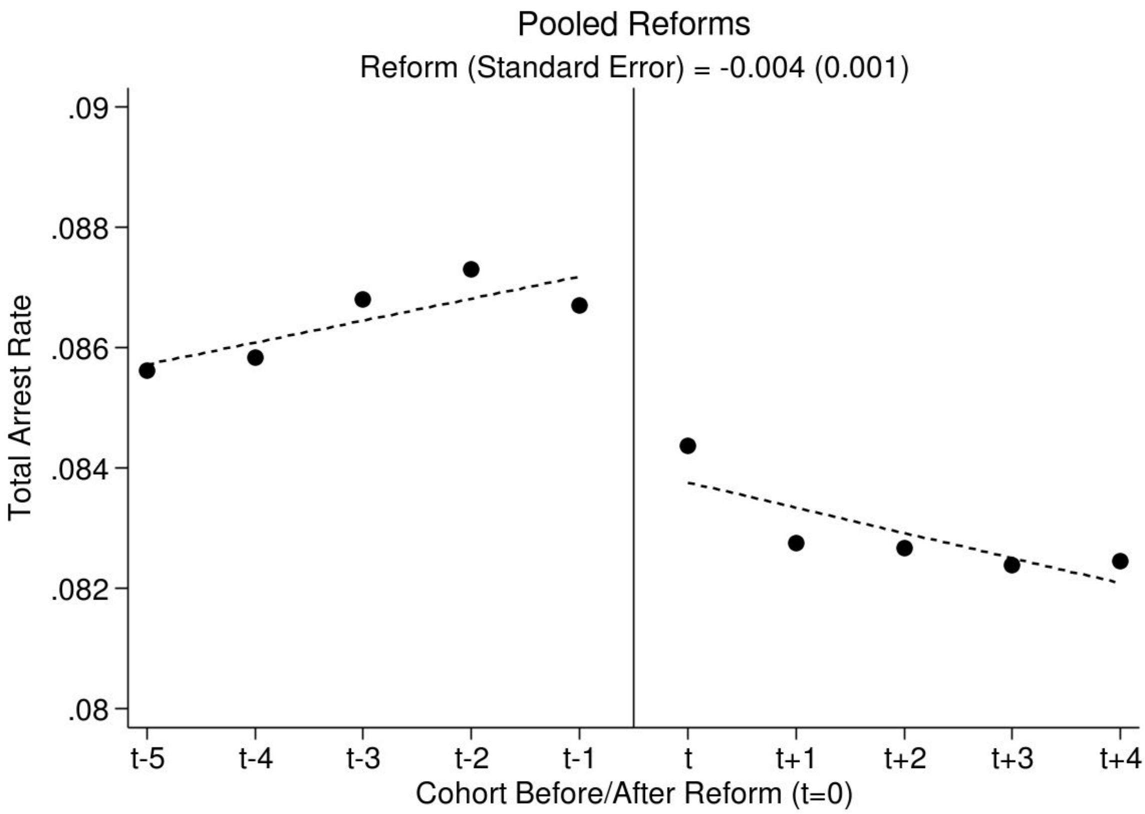


2010



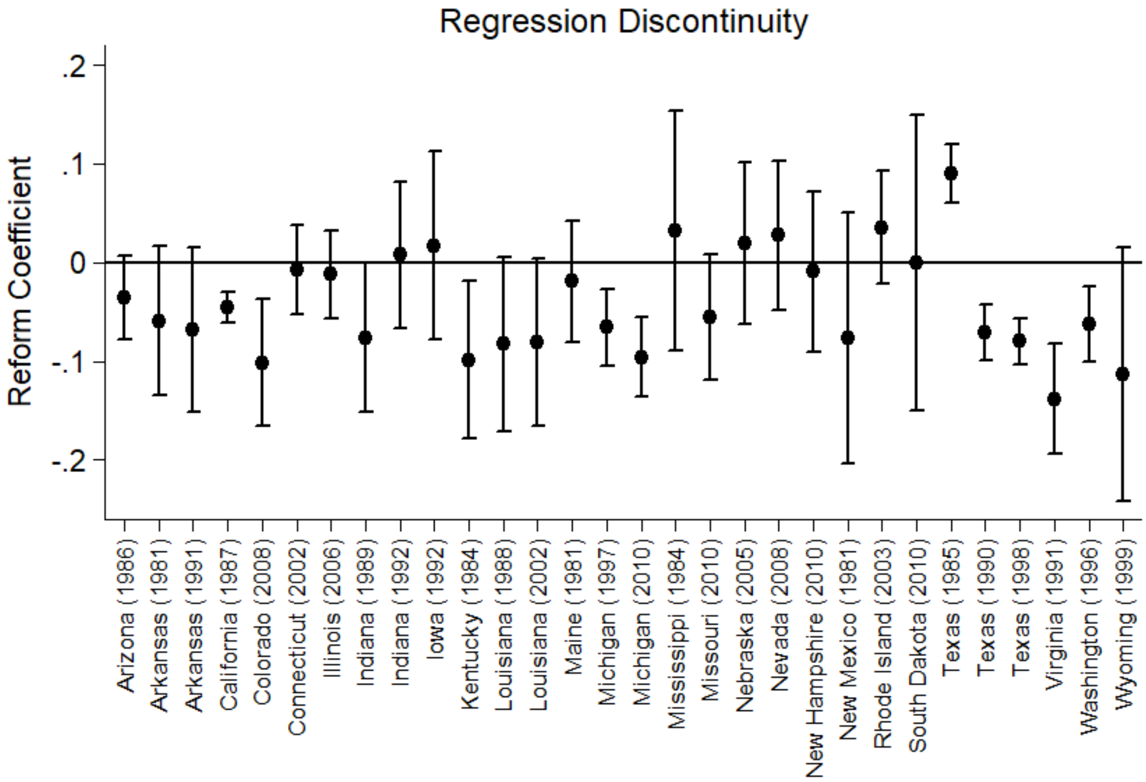
Notes: Dropout Age is calculated using the formula Dropout Age_{st} = min{Minimum Leaving Age_{st} , Grade $Exemption_{st}$ }

Figure 4: Arrest Rates Before/After Reforms



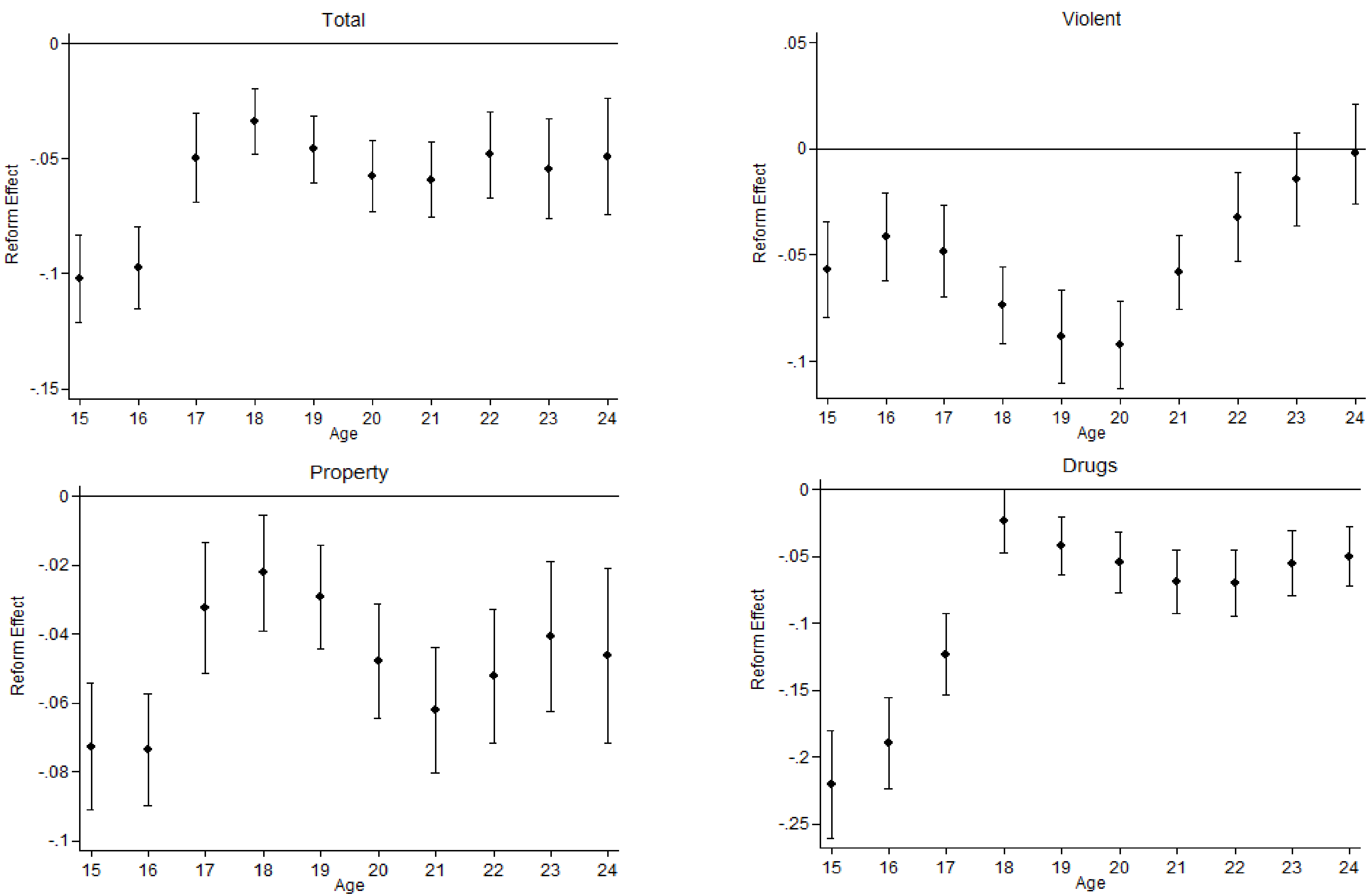
Notes: The reported discontinuity estimate (with associated standard error in parentheses) is the +/- 5 year mean difference pre and post-reform for the outcome variables.

Figure 5: Estimated Discontinuity Coefficients



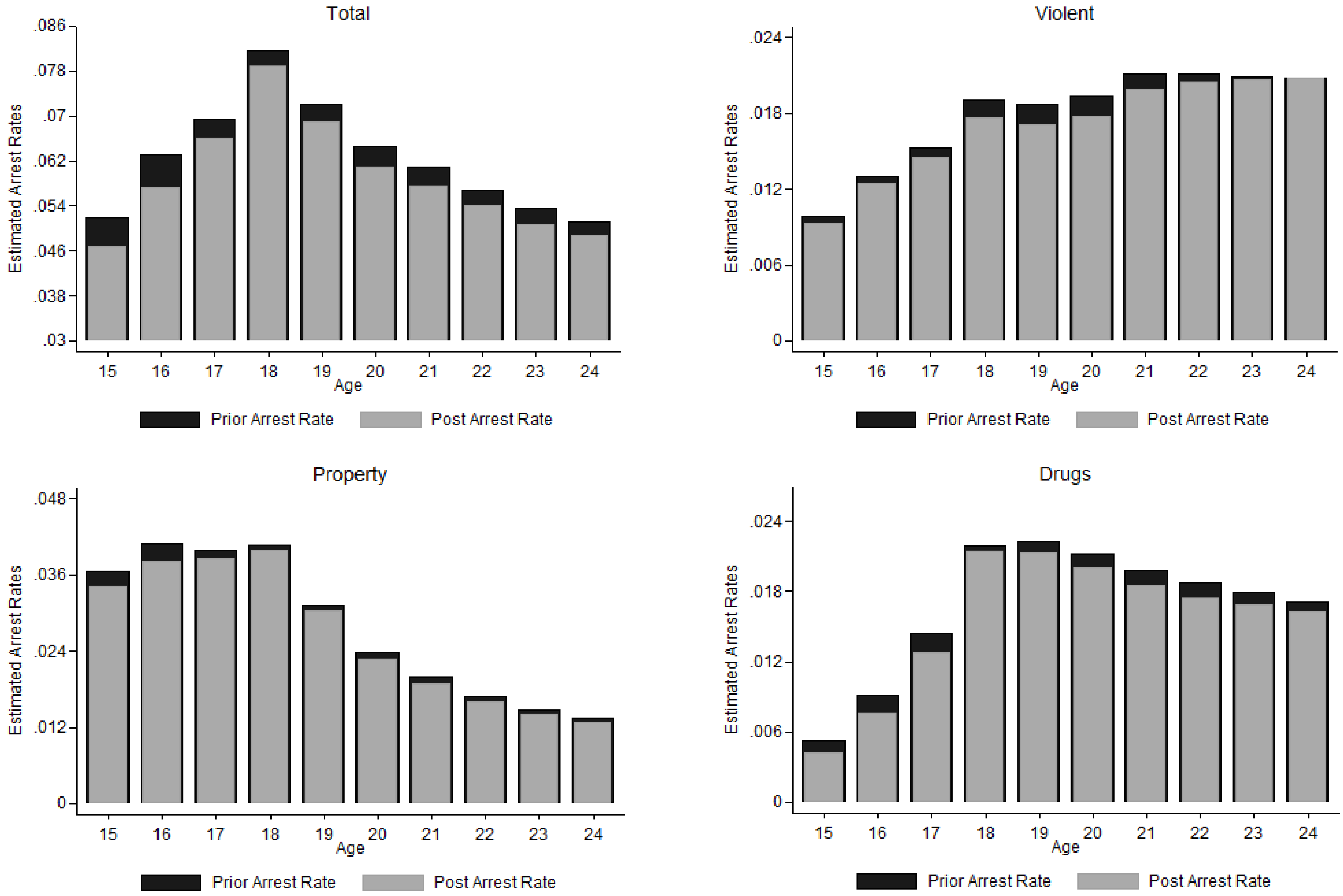
Notes: Coefficients from Table A4. 95% confidence intervals are shown.

Figure 6: Discontinuity Estimates by Age and Crime Type



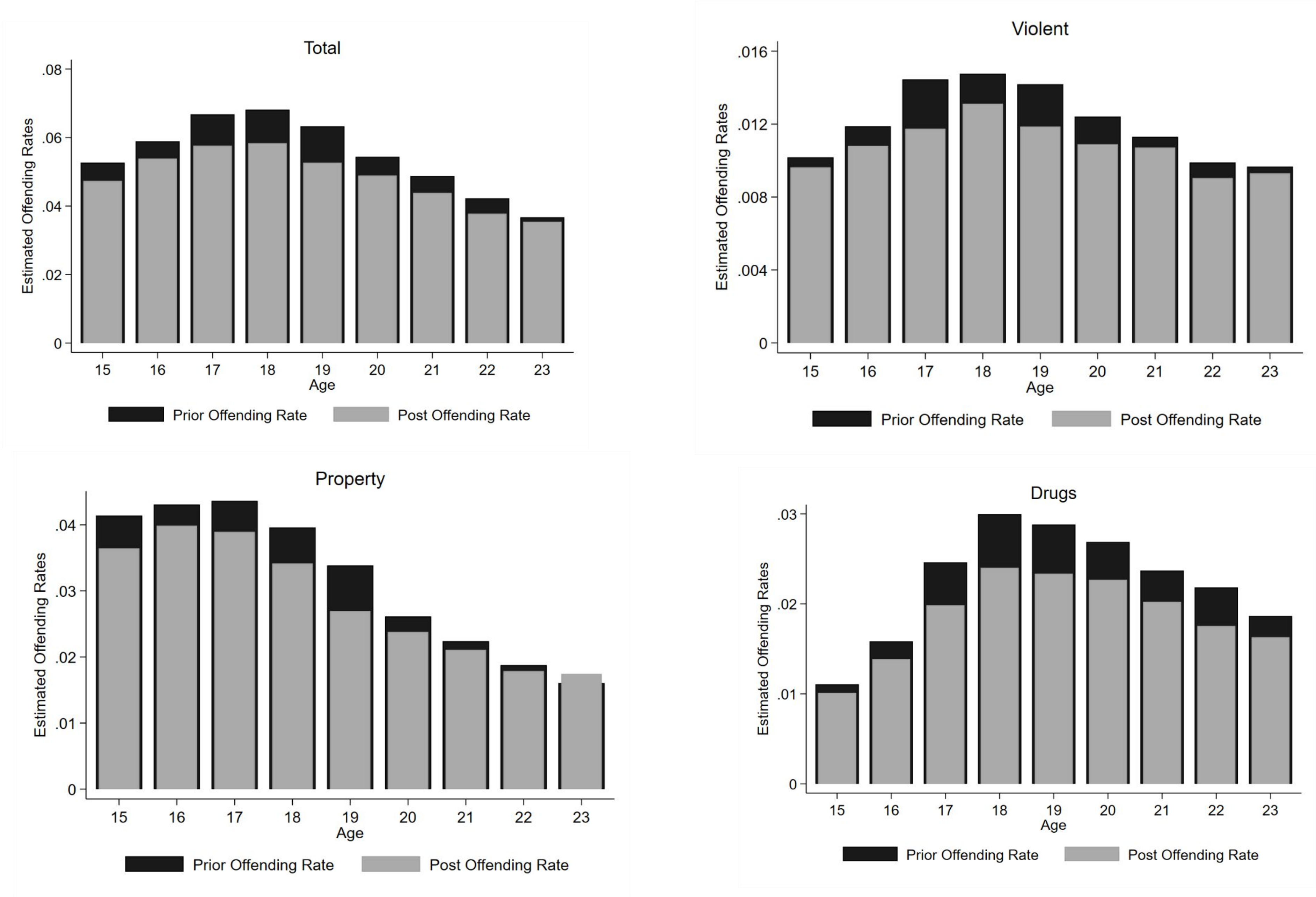
Notes: From estimates in Table 5. Texas (1985) is excluded from the estimation given that is a decrease in dropout age. Confidence intervals at 95% significance, standard errors clustered at reform-cohort level.

Figure 7: Crime-Age Profile Shifts by Crime Type



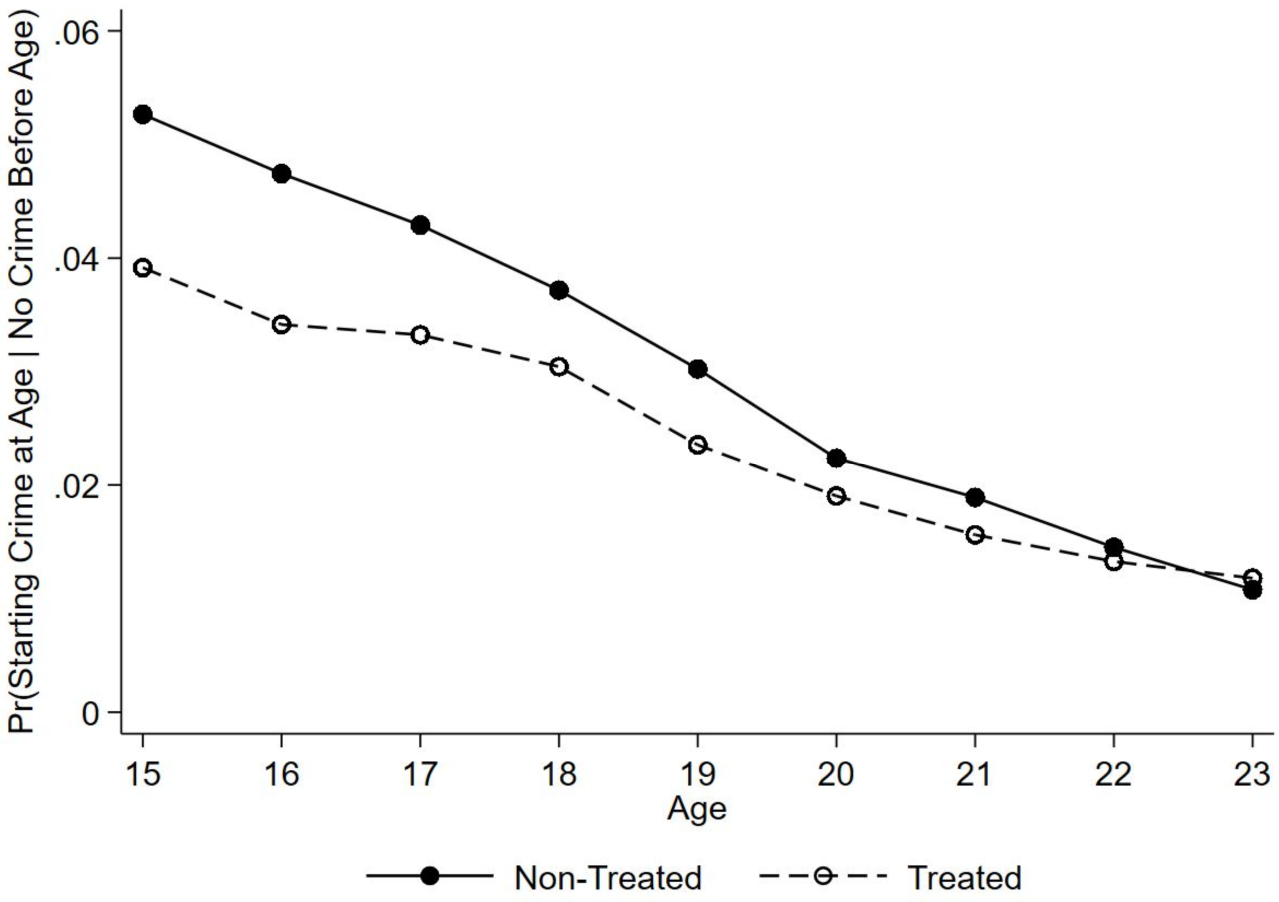
Notes: Prior arrest rate is the mean of arrest rate by age prior to discontinuity using a 5-year bandwidth. Post arrest rate is the calculated by estimated age effects of Figure 6.

Figure 8: Crime-Age Profile Shifts by Crime Type, Queensland - Australia



Notes: Prior offending rate is the mean of offending rate by age prior to discontinuity using a 3-year bandwidth. Post offending rate is the calculated by estimated age effects calculated in using the same model of Figure 7 applied to Queensland – Australia individual panel data.

Figure 9: Hazard Function from Duration Model, Queensland - Australia



Notes: The hazard functions are calculated based on the predicted values of a discrete time duration model assuming a log-logistic density with specification including age and year fixed effects and a full interaction set age*treatment.

Table 1: State Dropout Age Reforms

State	Effective School Year From Statute	Type	Change	New Dropout Age
Arizona	1985 and 1986	Exemption	8^{th} to 10^{th} grade	16
Arkansas	1981	Leaving Age	16 to 17	17
Arkansas	1991	Leaving Age	17 to 18	18
California	1987	Leaving Age	16 to 18	18
Colorado	2008	Leaving Age	16 to 17	17
Connecticut	2002	Leaving Age	16 to 18	18
Illinois	2006	Leaving Age	16 to 18	17
Indiana	1989	Leaving Age	16 to 17	17
Indiana	1992	Leaving Age	17 to 18	18
Iowa	1992	Exemption	8^{th} to 12^{th} grade	16
Kentucky	1984	Leaving Age	17 to 18	18
Louisiana	1988	Leaving Age	16 to 17	17
Louisiana	2002	Leaving Age	17 to 18	18
Maine	1981	Exemption	9^{th} to 12^{th} grade	17
Michigan	1997	Exemption	NA to 12^{th} grade	16
Michigan	2010	Leaving Age	16 to 18	18
Mississippi	1984	Leaving Age	Reenactment	17
Missouri	2009	Leaving Age	16 to 17	17
Nebraska	2005	Leaving Age	16 to 18	18
Nevada	2008	Leaving Age	17 to 18	18
New Hampshire	2010	Leaving Age	16 to 18	18
New Mexico	1981	Exemption	10^{th} to 12^{th} grade	18
Rhode Island	2003	Exemption	NA to 12^{th} grade	16
South Dakota	2010	Leaving Age	16 to 18	18
Texas	1985	Leaving Age	Rewriting of law	16
Texas	1990	Leaving Age	16 to 17	17
Texas	1998	Leaving Age	17 to 18	18
Virginia	1991	Leaving Age	17 to 18	18
Washington	1996	Exemption	9^{th} to 12^{th} grade	18
Wyoming	1999	Exemption	8^{th} to 10^{th} grade	16

Notes: Mississippi abolished its compulsory school law in 1956, and reenacted it 1983/84 with an initial leaving age of 7 with progressive raise until 17 by the school year 1989/90. Texas has written its laws of 1984 and 1989 in a different way, stating the minimum leaving age was to include the completion of school year in which the birthday occurred in effect decreasing/increasing the leaving age by some months. Two other reforms occurred during the same period – in South Carolina (1987) and Kansas (1996). Missing arrests data precludes them from this study.

Table 2: Baseline Estimates of Crime Reduced Forms

	Log(Arrest Rate), 1974 to 2015					
	(1)	(2)	(3)	(4)	(5)	(6)
	All States	10-Year Window	10-Year Window	10-Year Window	7-Year Window	5-Year Window
Reform	-0.099	-0.047	-0.065	-0.038	-0.062	-0.060
	(0.018)	(0.007)	(0.009)	(0.008)	(0.007)	(0.006)
Running Variable		Linear*Reform	Quadratic*Reform	Cubic*Reform	Linear*Reform	Linear*Reform
Reform Interactions		X	X	X	X	X
Sample Size	1,121,590	344,940	344,940	344,940	246,526	178,005
Number of States	48	24	24	24	24	24
Number of Counties	3,063	1,277	1,277	1,277	1,277	1,277

Notes: Sample includes males in each age group 15-24 inclusive for US counties. Estimates are weighted by population size and standard errors are clustered at state-cohort level (reform-cohort level for discontinuity windows). The dependent variable is the log of total arrest rate including violent, property and drug crimes. All specifications include age, year and county fixed effects. Covariates further include log of population, log of police force sworn and shares of female, black, non-white/non-black population. Reform Interactions means every covariate is made state-reform specific by adding an interaction with the state-reform indicator. Columns (2) to (6) include a centered running variable interacted with the dropout reform indicator as to allow differential trends at each side of the discontinuities.

Table 3: Estimates by Reform Type

	Log(Arrest Rate), 1974 to 2015, Discontinuity (+/- 5 years) Sample					
	(1)	(2)	(3)	(4)	(5)	(6)
	All Age Increase Reforms	17 to 16, Texas	16 to 17	17 to 18	16 to 18	Other
Reform	-0.062 (0.006)	0.090 (0.015)	-0.071 (0.015)	-0.068 (0.013)	-0.060 (0.011)	-0.041 (0.007)
Sample Size	156,517	21,488	47,943	46,984	34,209	27,381
Number of States	24	1	7	6	8	8
Number of Counties	1,242	222	533	487	374	282

Notes: As for Table 2. Same specification as column (6) of Table 2. Each column shows separate regression according to the relevant reform sample. "Other" include the following reforms: Arizona (1985), Iowa (1992), Maine (1981), Michigan (1997), Mississippi (1984), Rhode Island (2003), Washington (1996) and Wyoming (1999).

Table 4: Estimates by Crime Type and Age

	Log(Arrest Rate), 1974 to 2015, Discontinuity (+/- 5 years) Sample, All Age Increase Reforms			
	(1)	(2)	(3)	(4)
	Total	Violent	Property	Drugs
A. Overall Reform Effect				
Reform	-0.062	-0.056	-0.053	-0.099
	(0.006)	(0.008)	(0.007)	(0.010)
B. Reform Effects By Broad Age Groups				
Reform*Age 15-18	-0.064	-0.059	-0.053	-0.128
	(0.005)	(0.008)	(0.010)	(0.013)
Reform*Age 19-24	-0.041	-0.046	-0.043	-0.042
	(0.005)	(0.008)	(0.008)	(0.009)
Sample Size	156,517	156,517	156,517	156,517
Number of States	24	24	24	24
Number of Counties	1,242	1,242	1,242	1,242

Notes: As for Table 2. Same specification as column (6) of Table 2. Sample excludes Texas (1985) reform given that is a decrease in compulsory schooling.

Table 5: Age Varying Reform Impacts

	Log(Arrest Rate), 1974 to 2015, Discontinuity (+/- 5 years) Sample, All Age Increase Reforms			
	(1)	(2)	(3)	(4)
	Total	Violent	Property	Drugs
Reform*Age = 15	-0.102	-0.059	-0.078	-0.221
	(0.010)	(0.014)	(0.011)	(0.024)
Reform*Age = 16	-0.097	-0.043	-0.075	-0.190
	(0.009)	(0.013)	(0.010)	(0.020)
Reform*Age = 17	-0.050	-0.049	-0.035	-0.123
	(0.010)	(0.013)	(0.012)	(0.016)
Reform*Age = 18	-0.034	-0.076	-0.023	-0.023
	(0.007)	(0.011)	(0.009)	(0.013)
Reform*Age = 19	-0.046	-0.092	-0.032	-0.041
	(0.007)	(0.013)	(0.009)	(0.013)
Reform*Age = 20	-0.057	-0.095	-0.052	-0.057
	(0.008)	(0.012)	(0.009)	(0.013)
Reform*Age = 21	-0.059	-0.061	-0.063	-0.072
	(0.008)	(0.010)	(0.011)	(0.015)
Reform*Age = 22	-0.048	-0.033	-0.051	-0.074
	(0.010)	(0.013)	(0.012)	(0.016)
Reform*Age = 23	-0.054	-0.013	-0.053	-0.068
	(0.011)	(0.015)	(0.015)	(0.017)
Reform*Age = 24	-0.049	-0.003	-0.061	-0.060
	(0.014)	(0.017)	(0.018)	(0.016)
Sample Size	156,517	156,517	156,517	156,517
Number of States	24	24	24	24
Number of Counties	1,242	1,242	1,242	1,242

Notes: As for Table 2. Same specification as column (6) of Table 2. Sample excludes Texas (1985) reform given that is a decrease in compulsory schooling.

Table 6: Estimates for High School Attendance, Education, Employment and Wages

		(1)	(2)	(3)	(4)
	Pre-Reform Mean	All States	10-Year Window	7-Year Window	5-Year Window
A. High School Attendance (16-18)					
Reform	0.747	0.010	0.039	0.040	0.050
		(0.003)	(0.005)	(0.006)	(0.006)
B. High School Dropout					
Reform	0.109	-0.007	-0.004	-0.005	-0.006
		(0.001)	(0.001)	(0.001)	(0.001)
C. School or Work					
Reform	0.818	0.009	0.003	0.004	0.003
		(0.001)	(0.002)	(0.002)	(0.002)
D. Log Weekly Real Wages					
Reform	6.576	0.010	0.004	0.007	0.005
		(0.003)	(0.003)	(0.004)	(0.004)
Running Variable			Linear*Reform	Linear*Reform	Linear*Reform
Reform Interactions			X	X	X
Sample Size (Panel A)		1,026,804	254,257	181,689	131,001
Sample Size (Panels B and C)		6,816,430	1,716,601	1,201,659	861,019
Sample Size (Panel D)		4,854,245	1,272,952	893,008	640,527
Number of States (Panel A)		41	17	17	17
Number of States (Panels B to D)		48	24	24	24

Notes: CPS Basic Monthly (Panel A) sample includes all males, ages 16 to 18, from 1976-2015. Attendance in A is defined as an individual reporting to attend school full-time with education attainment lower than some college (See Appendix A).Panels B to D includes US born males in each age group 19-60 inclusive from 2006-2015 American Community Survey (ACS). Estimates are weighted by population weights and standard errors are clustered at state-cohort level. The dependent variables are: years of schooling, an indicator for high school dropout, an indicator for currently employed or attending school individuals (work or school) and log of real weekly wages. All specifications include age, year, black, hispanic and state of birth fixed effects (month fixed effects are added to row A). Reform Interactions means every covariate is made state-reform specific by adding an interaction with the state-reform indicator.

Table 7: Reform Effects on Intensive and Extensive Margins of Crime, Queensland – Australia

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Offending at 15-23	Offending at 15-17	Offending at 18-23	Ever Offending at 15-23	Offending at 18-23 | No Prior Offense at 15-17	Offending at 18-23 | Prior Offense at 15-17	Number Offenses at 18-23 | No Prior Offense at 15-17	Number Offenses at 18-23 | Prior Offense at 15-17
Margin				*Extensive (Ever Offending)*	*Extensive (First Offending)*	*Extensive (Reoffending)*	*Intensive (First Offending)*	*Intensive (Reoffending)*
Effect at Relevant Age	-0.006 (0.001)	-0.008 (0.002)	-0.005 (0.001)	-0.043 (0.002)	-0.030 (0.002)	-0.038 (0.008)	-0.260 (0.065)	-0.757 (0.246)
Sample Size	1,251,478	433,309	818,169	151,197	134,456	16,741	14,755	9,249
Number of Individuals	151,197	151,197	151,197	151,197	134,456	16,741	14,755	9,249
Pre-Reform Mean of Dependent Variable	0.055	0.060	0.052	0.233	0.127	0.573	2.873	6.936

Notes: Sample includes 2002 to 2013, males, ages 15-23. Only observations with balanced observations conditional on the cohort-year frame are included within a -/+ 3 year cohort window. Every specification includes year and age fixed effects. Standard errors are clustered at individual level.

Table 8: Cost-Benefit Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Victim costs per crime	Property loss per crime	Incarceration costs per crime	Total costs per crime	Estimated change in arrests	Estimated change in crimes	Estimated change in incarcerations	Benefits (6)*(4)	Estimated change in school enrollment	Costs (9)*$3,910	Benefit/Cost Ratio (8)/(10)
Ages 15-18											
A. Violent Crimes	33,406	135	7,399	40,697	-12,325	-26,168	-6,240	$1,064,944,258			
B. Property Crimes	1,653	1,141	201	941	-18,498	-99,452	-9,295	$93,563,228			
C. Drug Crimes[a]	1,004	NA	6,431	7,435	-20,087	-24,922	-9,413	$185,292,728			
Total					-50,910	-150,541	-24,948	$1,391,301,234	367,046	$1,435,149,860	**0.94**
Ages 19-24											
A. Violent Crimes	33,406	135	7,399	40,697	-14,182	-30,110	-7,180	$1,225,398,740			
B. Property Crimes	1,653	1,141	201	941	-10,512	-56,516	-5,282	$53,169,892			
C. Drug Crimes[a]	1,004	NA	6,431	7,435	-11,088	-13,757	-5,196	$102,281,364			
Total					-35,782	-100,383	-17,658	$1,407,070,502	-	-	**1.90**

Notes: Costs of violent and property crimes are weighted averages of the breakdown costs from Lochner and Moretti (2003) using average share of crimes composing each of the categories as weights. Costs of drug crimes are based on the US Department of Justice (2011) victim costs and other crime costs, and incarceration costs are scaled in the same way as Lochner and Moretti (2003). Estimated change in arrests are calculated based on the results from Table 4 scaled using 1993 population within the age groups. Estimated crimes and incarcerations are calculated using 2009 clearances rates and conviction to incarceration rates respectively for each type of crime. Estimated change in enrolment is calculated using the results from Table 7 Column (4) scaled using the 1993 population within the age groups. Yearly costs per pupil (3,910) correspond to the average pupil costs (U.S. Department of Education, 2016) from 1974 to 2014. All figures are deflated to 1993 dollars.

FOR ONLINE PUBLICATION

Appendix A: Data Description

A1. Panel Data on Arrests

Panel data for the US come from the FBI Uniform Crime Reports (UCR). The measure of crime is arrests. The UCR reports the number of arrests by year, state, age, gender and type of crime. The original data identifies the number of arrests by law enforcement agencies within states. We construct a county-level panel on arrests by aggregating the number of arrests over law enforcement agencies within a county. Within the UCR, data for certain agencies is systematically missing. For example, New York City systematically does not report arrest numbers. For the agencies used in our estimation we impose a reporting pattern consistent with a maximum tolerance of one missing year per discontinuity window (i.e. for 10-year bandwidth the agency needs to report 18 out of the 20 years)[24].

In addition, the UCR reports the total population for each law enforcement agency in the reported year. Aggregating the UCR population count to the county-year level and comparing that number to official population counts allows us to identify county-year covering ratio. The weighted average county-level covering ratio is 89% for the 5-year bandwidth. When estimating the population per age-sex cell, we use the SEER*Stat population estimates[25] at county level and apply the yearly covering ratios homogeneously across different ages. The implicit assumption is that the missing population has the same age breakdown as the overall county-year population. The weighted average share of state population covered in the 5-year bandwidth is 81% for reform states.

We sample males aged 15 to 24 from 1974 to 2015. The UCR data are grouped by age category. From age 15 up to the age of 24, the number of arrests is measured by single age year.

Following the literature, we categorize arrests into property and violent crime using the UCR offense code variable as follows:

Violent crime:	Property crime:	Drug Crime:
01A = Murder and non-negligent manslaughter	05 = Burglary – breaking or entering	18 = Drug Violations (Possession, Sale and Manufacturing)
01B = Manslaughter by negligence	06 = Larceny – theft (except motor vehicle)	
02 = Forcible rape	07 = Motor vehicle theft	
03 = Robbery	09 = Arson	
04 = Aggravated assault		
08 = Other assaults		

24 Table 2, column (1) sample includes only agencies reporting at least 10 years over 1974-2015. The results are robust to a stricter reporting of all years in the bandwidth period.

25 See Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) 1969-2014.

In order to produce arrest rates, we aggregate the number of arrests for the above categories and divide the resulting number of arrests by the annual county-age-year population. Some cells report 0% arrest rates, however for those we assume the lowest arrest rate reported in the sample period. Cell reporting arrest rates above 40% are excluded from the sample to avoid outliers influencing the analysis.[26]

A2. Racial Breakdown Covariates

An analogous method to one used to obtain population age-sex cell in *A1* is performed to estimate the racial breakdown of each cell. We use the racial population estimates collected from SEER*Stat at county-year level and the county covering ratio to estimate the number of population by race for each sex-county-year cell.

A3. Police Numbers

The police numbers used are collected from FBI LEOKA (which is available from 1960-2015). This data reports several police enforcement measures yearly for each enforcement agency. We use the total number of sworn officers per county-year as the measure of police force present at the geographic area of interest.

A4. County-Year Economic and Employment Covariates

Information on economic and employment indicators at county-year level are collected from Local Area Personal Income (LAPI) from Bureau of Economic Analysis (BEA) 1960-2015. Measures of total employment, personal income, wage income and several others are available at county-year level from LAPI.

A5. School Quality Measures

We use the Local Education Agency (LEA) level data available in the Common Core Data (CCD) from National Center for Education Statistics (NCES), both fiscal and non-fiscal[27], to produce the school quality measures. By aggregating both the number of students, teachers and instruction salary expenses at county-year level, we are able to compute estimates of pupil-teacher ratio and average teacher salary. We interpolate missing years in the data, these are not frequent and do not affect the general results.

A6. ACS 2006-2015: Education and Work

We sample all males aged 19-60 from Integrated Public Use Microdata Series (IPUMS) for the American Community Survey (ACS) 2006-2015[28]. The sample is restricted to US born individuals as to ensure the individuals are directly affected by the compulsory schooling laws enacted.

[26] The results are robust to the use of 30% and 50% thresholds alike.
[27] Unfortunately, fiscal information at LEA level is only available since 1989.
[28] Despite the fact that ACS started being collected as a 1% sample in 2005, this year is not included in the sample analysis given the empirical break in the education variables between 2005 and 2006.

Years of schooling are coded according to Acemoglu and Autor (2011). High school dropout is defined as an individual who has less than a high school graduation diploma or equivalent. Work and school indicator function is defined for an individual who is either classified as currently employed or attending school (college or high school) both full-time or part-time. Weekly and hourly wages are coded for both part-time and full-time workers (excluding self-employed and unpaid family workers) according to Acemoglu and Autor (2011) with minor improvements on top coding by making the adjustments state-specific according to the ACS sample design.

Reforms are matched by state of birth, as it is assumed that an individual born in a given state has attended school in that same state at least until dropout age. When matching the reforms to the individual data from ACS, a one-year sliding on the reform year is observed and adjusted for. The previous arises due to the inability to precisely estimate the year of birth for a given individual as data is collected over different months for each survey year and state, and only age is provided in the ACS hence making year of birth an approximated variable.

A7. CPS Monthly Basic 1976-2015: School Attendance

We sample all males aged 16-18 from the National Bureau of Economic Research (NBER) archive of the Basic Monthly Current Population Survey (CPS) 1976-2015. Unlike with the ACS sample, we are not able to distinguish between US born and migrants in the CPS consistently through the sample period. Summer months (June, July and August) are excluded from the sample as they consistently report significantly low enrollment in high school or college.

High school attendance definition before 1984 is based on the answer to the question "What was your main activity last week?"[29] being "School" conditional on the individual not having any education attainment superior to high school graduation. After 1984, individuals are directly questioned about their enrollment status differentially between high school and college[30]. We, therefore, define an individual as attending high school if he/she declares to be enrolled in high school conditional not having any education attainment superior to high school graduation. We analyze the sample period when both questions are available (1984-1993) and conclude that, conditional on the individual not having any education attainment superior to high school graduation, 91% of the individuals stating to be enrolled in high school answered "School" as their main activity last week. This attests for the strong correlation between both measures, dissipating concerns on significant jumps in the variable of school attendance between periods.

In CPS individuals do not report their state of birth, hence reforms are matched by state of residence. Considering that school attendance is being measured contemporaneously, we have no strong reason to believe that individuals between 16 and 18 years of age would not be subject to the school dropout age of their state of residence. When matching the reforms to the individual data from ACS and CPS, a one-year sliding on the reform year is observed and adjusted for. The previous arises due to the inability to precisely estimate the year of birth for a given individual as data is collected over different months for each survey year

[29] This question was discontinued in 1994.

[30] This variable is only available for individuals 16 or older, hence the sample of the analysis starting at age 16 and not earlier despite a few reforms potentially affecting younger ages.

and state, and only age is provided in the ACS and CPS hence making year of birth an approximated variable.

A8. Compulsory Schooling Laws

Compulsory schooling laws are collected directly from official annotated statutes of each state in the Westlaw International Database for each of the corresponding years. When provided in the statutes, the effective date of the new law is taken as the year of reform otherwise enactment year is assumed to be the most sensible approximation.

The data retrieved includes maximum entry age, minimum leaving age and education grade which exempts a child from staying in school. The laws have historically increased in their complexity adding several exemptions including work permits and early age parental consent letters to exemplify the most common. The Labor Standards Act 1939 harmonized child labour laws across states in the US, recent changes were not of a comparable order of magnitude as the ones seen during that period. To be consistent we ignore the possibility of parental consent authorizations to leave school at an age below the minimum dropout age, as these are often seen as exceptions rather than the rule.

A9. Australia - Queensland Earning or Learning Reform

Analogously to the US case, the school system in Queensland consists of 12 years of education (grades 1 to 12). Prior to the enactment of the Earning or Learning reform of 2006, students were required to attend school until either completing grade 10 or turning 16 whichever occurred first, this structure of compulsory attendance is in every expect equivalent to US states with grade exemption clauses.

The Earning or Learning reform introduced a compulsory participation obligation. According to the reform, young people were mandated to participate in a range of activities broadly defined as earning or learning for up to an additional two years. Thus, the compulsory participation phase required youth to either stay on at school until obtaining a high school Senior Certificate; complete a vocational education Certificate III; or participate in paid employment for at least 25 hours per week until turning age 17.

All three dimensions covered by the Australian reform are present in different reforms analysed in the US. The toughening of grade exemption to high school completion status was enacted by Iowa, Maine, Michigan, New Mexico, Rhode Island and Washington. Vocational education as optional route was included in both California and Texas changes. Finally, employment permit and certificates as proof of full-time work are required in several states in order to request exception from compulsory school attendance when below the legal dropout age.

A10. Australia - Queensland Microdata

Australian analysis is based on Queensland administrative data matched at the individual level across state agencies, Department of Education and Training (DET) and the Queensland Police Service (QPS). The former enables to construct a panel of individual record data for the entire population of attendees at publicly funded schools matched with their individual criminal offence data for the period 2002 to 2013. The data used is the extract from Beatton, Kidd, Machin and Sarkar (2018) which was kindly made available for our analysis by Tony Beatton and Michael Kidd. In order to assure comparability with the US analysis, we focus on males ages 15 to 23.

Table A1. Coverage of Counties

State (Year)	Covered Counties / Total Counties	% Within County Coverage	% Overall State Coverage
Arizona (1985/86)	14 / 16	85.0%	91.3%
Arkansas (1981)	62 / 75	85.3%	83.2%
Arkansas (1991)	73 / 75	88.8%	90.6%
California (1987)	58 / 58	96.0%	95.1%
Colorado (2008)	55 / 64	87.9%	86.0%
Connecticut (2002)	8 / 8	71.6%	95.5%
Illinois (2006)[a]	1 / 102[a]	53.7%	22.2%
Indiana (1989)	37 / 92	51.5%	42.4%
Indiana (1992)	38 / 92	52.8%	43.6%
Iowa (1992)	89 / 99	86.1%	80.9%
Kentucky (1984)	94 / 120	74.7%	67.1%
Louisiana (1988)	36 / 64	72.7%	55.2%
Louisiana (2002)	42 / 64	70.4%	61.4%
Maine (1981)	16 / 16	94.3%	94.4%
Michigan (1997)	77 / 83	84.8%	85.0%
Michigan (2010)	78 / 83	91.9%	93.7%
Mississippi (1984)	24 / 82	40.3%	23.6%
Missouri (2009)	108 / 114	82.2%	90.4%
Nebraska (2005)	56 / 93	86.5%	84.2%
Nevada (2008)	15 / 17	95.8%	97.3%
New Hampshire (2010)	10 / 10	56.6%	55.6%
New Mexico (1981)	19 / 34	58.0%	50.8%
Rhode Island (2003)	5 / 5	98.6%	98.8%
South Dakota (2010)	29 / 66	84.3%	66.1%
Texas (1985)	223 / 254	89.1%	92.2%
Texas (1990)	235 / 254	94.9%	95.9%
Texas (1998)	230 / 254	95.1%	96.8%
Virginia (1991)	124 / 142	98.7%	93.4%
Washington (1996)	36 / 39	79.6%	51.9%
Wyoming (1999)	22/23	92.0%	94.1%

Notes: Coverage ratios are computed by dividing the population covered in the arrest data by the population estimated from the SEER*Stats for the respective geographies: county and state.

Table A2: Balancing Tests

	Balancing Covariates		
	-5 years	+5 years	Difference (Standard Error)
Share of Black	0.136	0.136	-0.001
	(0.006)	(0.006)	(0.009)
Share of Others	0.049	0.059	0.009
	(0.005)	(0.007)	(0.009)
Share of Female	0.484	0.483	-0.001
	(0.001)	(0.001)	(0.002)
Log Police	6.872	7.002	0.130
	(0.140)	(0.133)	(0.193)
Log Population	8.028	8.041	0.013
	(0.136)	(0.126)	(0.185)
Teacher-Pupil Ratio	18.32	17.73	-0.581
	(0.537)	(0.573)	(0.776)

Notes: Sample includes cohorts of males aged 15-24 for US counties over time. Means across all counties in the balanced sample for each of the 30 reforms (as in Table 1), on each side of the +/- 5 bandwidth. Estimates are weighted by population size and standard errors are clustered at reform-cohort level.

Table A3: Discontinuity Estimates by Individual Reform

State	Effective School Year	All Ages
Arizona	1985 and 1986	-0.036 (0.021)
Arkansas	1981	-0.059 (0.038)
Arkansas	1991	-0.068 (0.043)
California	1987	-0.046 (0.008)
Colorado	2008	-0.102 (0.033)
Connecticut	2002	-0.008 (0.023)
Illinois	2006	-0.012 (0.023)
Indiana	1989	-0.076 (0.039)
Indiana	1992	0.008 (0.038)
Iowa	1992	0.017 (0.048)
Kentucky	1984	-0.099 (0.041)
Louisiana	1988	-0.083 (0.044)
Louisiana	2002	-0.081 (0.043)
Maine	1981	-0.019 (0.031)
Michigan	1997	-0.066 (0.020)
Michigan	2010	-0.096 (0.021)
Mississippi	1984	0.032 (0.062)
Missouri	2009	-0.055 (0.033)
Nebraska	2005	0.019 (0.041)
Nevada	2008	0.027 (0.032)
New Hampshire	2010	-0.009 (0.041)
New Mexico	1981	-0.077 (0.065)
Rhode Island	2003	0.036 (0.028)
South Dakota	2010	-0.002 (0.077)
Texas	1985	0.090 (0.015)
Texas	1990	-0.071 (0.014)
Texas	1998	-0.080 (0.012)
Virginia	1991	-0.138 (0.029)
Washington	1996	-0.063 (0.019)
Wyoming	1999	-0.112 (0.066)

Notes: Same specification as column (6) of Table 2. Each row is estimated as separate regression for each reform with a 5-year window.

Table A4: Crime-Age Profile Summary Measures Before and After Dropout Age Changes

	Changes in Crime Age Profile Summary Measures							
	Mean	Standard Deviation	Skewness	Kurtosis	20^{th} Percentile	50th Percentile	80^{th} Percentile	Mode
Total	0.035	-0.024	-0.012	0.014	0.136	0.075	0.090	0.315
	(0.010)	(0.004)	(0.004)	(0.005)	(0.018)	(0.019)	(0.015)	(0.065)
Violent	0.044	-0.004	-0.022	-0.012	0.117	0.123	0.165	0.787
	(0.013)	(0.006)	(0.007)	(0.018)	(0.024)	(0.024)	(0.021)	(0.093)
Property	-0.003	-0.033	0.007	0.034	0.039	-0.015	-0.039	0.400
	(0.013)	(0.005)	(0.007)	(0.020)	(0.021)	(0.021)	(0.017)	(0.070)
Drugs	0.156	-0.011	-0.051	-0.015	0.178	0.202	0.218	0.438
	(0.014)	(0.006)	(0.007)	(0.015)	(0.021)	(0.024)	(0.020)	(0.064)

Notes: Calculated with population weights. Estimates are computed based on the residual arrests after compositionally adjusting at state-level for year, log police employed, log population and share of females, black and non-white/non-black population. Texas (1985) is excluded given that is a decrease in compulsory schooling. Discontinuities after 2008 are excluded given the unavailability of data to balance ages covered on both sides of the discontinuities.

Appendix B - A Model of Education Policy and Crime Age Profiles

The starting point is that an individual decides how to allocate time between the illegal sector, where they devote time to crime (c), and the legal sector, working ($t - c$), where t is his/her full-time endowment. However, time for crime is constrained whilst individuals are enrolled in school so the full-time endowment t will be a function of age, $t(a)$, which is reduced by a dropout age reform if an individual is in school and aged below the minimum age of school dropout a^d.

Normalizing c and t to the unit interval, one example of how the time endowment differs by age and dropout age is:

$$t(a) = \begin{cases} t_l \; for \; a < a^d \\ t_h \; for \; a \geq a^d \\ 0 < t_l < t_h < 1 \end{cases}$$

where the l and h subscripts index the amount of low and high free time available to allocate to crime. In this example $t_h < 1$ so 1 - t_h may be thought of as leisure time. The key feature of the model is that the younger individuals may do some crime (as $t_l > 0$) but because they are kept in school this acts as an incapacitation effect preventing them from engaging in as much crime as those older than the dropout age who have more available time for such activity.

The likelihood of these older individuals to do so depends on the relative returns to crime or work. The labour market returns to work are given by the wage, $w(e)$, which is a function of potential labour market experience, defined as $e = \max\{0, a - a^d\}$, and which reflects on-the-job-training and learning-by-doing. Each individual faces a rate of return to crime $r(e)$ and a sanction $s(a)$ if they are caught when doing crime.

Given the probability of being caught by law enforcement- defined $p(c)$ - and specifying a utility function $U(.)$ each individual maximizes his/her expected utility by choosing the optimal amount of time to spend on crime c in the following maximization problem:

$$\max_{\{c\}} \left(1-p(c)\right)U\left(r(e)c+w(e)(t(a)-c)\right)+p(c)U(r(e)c+w(e)(t(a)-c)-s(a)c)$$

$$s.t \;\; c \le t(a)$$

The following assumptions are required:

i) $U'(.) \ge 0$, $U''(.) \le 0$ – standard positive marginal utility and diminishing returns.

ii) $p_c' \ge 0$ – the probability of getting caught increases with time devoted to crime.

iii) $s_a' \ge 0$ – the sanction penalty increases as an individual approaches legal age and surpasses the extended age of the juvenile state court.

iv) $r_e' \ge 0$ – returns to criminal time increase as the individual gains potential experience and builds criminal capital.

v) $w_e' \ge 0$, $w_e'' \le 0$ – there are concave wage profiles that are particularly prevalent in young and low-educated workers.

vi) $t_a' \ge 0$ - as the individual gets older and is therefore older than the dropout age, the time endowment increases.

Imposing the assumptions (i) to (vi), and defining the relative return to crime over work as $n(e) \equiv r(e) - w(e)$, generates a solution to the individual's optimization problem to produce a level of crime c that satisfies the following first order condition:

$$\left(1-p(c)\right)U'(t(a)w(e)+n(e)c)n(e)+p(c)U'\left(t(a)w(e)+\left(n(e)-s(a)\right)c\right)\left(n(e)-s(a)\right)+p'(c)\left[U\left(t(a)w(e)+\left(n(e)-s(a)\right)c\right)-U(t(a)w(e)+n(e)c)\right]-\mu=0$$

$$\mu\left(c-t(a)\right)=0,\; c-t(a)\le 0,\; c\ge 0, \mu \ge 0$$

If $\mu \neq 0$ the constraint binds and $c = t(a)$ then the individual will use the full extent of his/her time endowment to engage in the illegal sector. On the other hand if $\mu = 0$, so the constraint does not bind, then the unconstrained problem is returned to. The optimality condition equalizes the marginal net benefit of crime to the marginal benefit of working as:

$$\left(1 - p(c)\right)U'(t(a)w(e) + n(e)c)n(e) + p(c)U'\left(t(a)w(e) + \left(n(e) - s(a)\right)c\right)\left(n(e) - s(a)\right) + p'(c)\left[U\left(t(a)w(e) + \left(n(e) - s(a)\right)c\right) - U(t(a)w(e) + n(e)c)\right] = 0$$

$$c - t(a) < 0, c \geq 0$$

Understanding how the optimal amount of crime varies with age, and so generates a crime-age profile, and in particular how changes in school dropout age can affect this, is the main feature of this model. The implicit derivative of the optimal crime choice with respect to age is given by $\frac{dc^*}{da} = -\frac{\partial F}{\partial a} / \frac{\partial F}{\partial c}$, where F stands for the first order condition.

The following proposition then emerges:

<u>Proposition</u>

If (i) individuals are risk averse, $k = -\frac{U''}{U'} \geq 0$[31], (ii) income is non-decreasing in age $t'(a)w(e) + t(a)w'(e) + \left(n'(e) - s'(a)\right)c \geq 0$, (iii) the net rate of return to crime is non-negative $n(e) \geq 0$, and decreasing in age, $n'(e) \leq 0$, and (iv) Lemma 1 below, then the slope of the crime age-profile will be decreasing in age $\frac{dc^*}{da} \leq 0$.[32]

[31] For simplicity, risk aversion is assumed constant across age so that dropout age change does not affect this parameter. Despite the scarce evidence on the relation between risk aversion and age, the consensus seems to point to a positive relationship. Assuming that older individuals are more risk averse, the results of the model simulation would be strengthen in terms of crime reduction as individuals would be constrained in their crime engagement until older ages. Furthermore, if education displays a similar relation with risk aversion the crime reducing effect over the age profile would be again more pronounced.

[32] The conditions stated are not exhaustive of all cases where $\frac{dc^*}{da} \leq 0$, however, they are the ones that are most in line with empirical evidence.

Proof:[33]

According to the model presented above, one can derive the following expressions:

$$\frac{\partial F}{\partial c} = (1-p(c))[U''(t(a)w(e)+n(e)c)n(e)^2] - 2p'(c)[U'(t(a)w(e)+n(e)c)n(e)] + p(c)\left[U''(t(a)w(e)+(n(e)-s(a))c)(n(e)-s(a))^2\right] + 2p'(c)[U'(t(a)w(e)+(n(e)-s(a))c)(n(e)-s(a))] + p''(c)[U(t(a)w(e)+(n(e)-s(a))c) - U(t(a)w(e)+n(e)c)]$$

$$\frac{\partial F}{\partial a} = (1-p(c))[U''(t(a)w(e)+n(e)c)(t'(a)w(e)+t(a)w'(e)+n'(e)c)n(e) + U'(t(a)w(e)+(n(e))c)n'(e)] + p(c)[U''(t(a)w(e)+(n(e)-s(a))c)(t'(a)w(e)+t(a)w'(e)+(n'(e)-s'(a))c)(n(e)-s(a)) + U'(t(a)w(e)+(n(e)-s(a))c)(n'(e)-s'(a))] + p'(c)[U'(t(a)w(e)+(n(e)-s(a))c)(t'(a)w(e)+t(a)w'(e)+(n'(e)-s'(a))c) - U'(t(a)w(e)+n(e)c)(t'(a)w(e)+t(a)w'(e)+n'(e)c)]$$

Let $k = -\frac{U''}{U'}$ be the coefficient of absolute risk aversion, $B \equiv t(a)w(e)+n(e)c$ and $B' = \frac{\partial B}{\partial a}$, then $\frac{\partial F}{\partial c}$ and $\frac{\partial F}{\partial a}$ can be rewritten as:

$$\frac{\partial F}{\partial c} = U'(B)[-k(1-p(c))n(e)^2 - 2p'(c)n(e)] + U'(B-s(a)c)[-kp(c)(n(e)-s(a))^2 + 2p'(c)(n(e)-s(a))] + p''(c)[U(B-s(a)c) - U(B)]$$

$$\frac{\partial F}{\partial a} = U'(B)\left[(B')\left(-kn(e)(1-p(c)) - p'(c)\right)\right] + U'(B)[n'(e)(1-p(c))] + U'(B-s(a)c)\left[(B'-s'(a)c)\left(-k(n(e)-s(a))p(c)+p'(c)\right)\right] + U'(B-s(a)c)[(n'(e)-s'(a))p(c)]$$

$\frac{\partial F}{\partial c}$ equals the second derivative of the objective function, assuming we have an interior solution, $\frac{\partial F}{\partial c} \leq 0$ to ensure the concavity of the objective function. The sign of $\frac{dc^*}{da}$ depends then on the sign of $\frac{\partial F}{\partial a}$.

[33] Individual risk aversion is a key feature of standard economic models, whilst the non-decreasing income as function of age (at least until retirement approaches) seems supported by existing empirical evidence, though for older individuals than considered here. The positive net rate of return to crime needs to hold if an individual is ever to engage their time in criminal activities - intuitively if the return was to be negative the individual would choose to engage all of his/her time in the legal sector. Perhaps the most challenging assumption is that of the net return to crime decreasing with age. We would argue that this is most plausibly the case for the later ages studied in this paper – where there are increasing sanctions due to the loss of juvenile and extended age status in court (Levitt, 1998) and no evidence of convex age returns to crime.

Lemma 1:

$$\left[U'(B)\left(B'\right)\left(-kn(e)(1-p(c))-p'(c)\right)\right]+\left[U'(B)n'(e)(1-p(c))\right]+\left[U'(B-s(a)c)(n'(e)-s'(a))p(c)\right]$$
$$\leq$$
$$-[U'(B-s(a)c)(B'-s'(a)c)(-k(n(e)-s(a))p(c)+p'(c))]$$

<u>Proposition</u>

If individuals are risk averse $k \geq 0$, income is non-decreasing in age $B'-s'(a)c = t'(a)w(e)+t(a)w'(e)+\left(n'(e)-s'(a)\right)c \geq 0$, the net rate of return to crime is non-negative and decreasing in age $n(e) \geq 0$, $n'(e) \leq 0$, and Lemma 1, then the crime age-profile will be decreasing in age $\frac{dc^*}{da} \leq 0$

Proof:

- $B' - s'(a)c \geq 0 \Rightarrow B' \geq 0$
- $n'(e) \leq 0 \Rightarrow n'(e) - s'(a) \leq 0$, as $s'(a) \geq 0$, $s(a) \geq 0$ and $c \geq 0$

It follows that:

$$U'(B)\left(B'\right)\left(-kn(e)(1-p(c))-p'(c)\right) \leq 0$$
$$U'(B)n'(e)\left(1-p(c)\right) \leq 0$$
$$U'(B-s(a)c)\left(n'(e)-s'(a)\right)p(c) \leq 0$$
$$U'(B-s(a)c)\left(B'-s'(a)c\right)\left(-k\left(n(e)-s(a)\right)p(c)+p'(c)\right) \geq 0$$

Combined with Lemma 1, this gives

$$\frac{\partial F}{\partial a} \leq 0 \Rightarrow \frac{dc^*}{da} \leq 0$$

Model Calibration

For the model simulation presented in Figure 2 in the main body of the paper, the following functional forms were used:

$$U(c) = \frac{(1+(t(a)-c)w(e)+(r(e)-s(a))c)^{1-\sigma}-1}{1-\sigma}, where\ \sigma = 2$$

$w(e) = log(1+e)$

$r(e) = r + (1 + r) * log(1 + e), where\, r = 0.55$

$s(a) = \mathrm{S} + e^{(sa)} - 1, where\, s = 0.3$

$t(a) = \begin{cases} 0.5\ for\ a < a^d \\ 1\ \ for\ a \geq a^d \end{cases}$

$a^d = 0.3,\ a'^d = 0.4\ , p(c) = c, 0 \leq c \leq 1\ , 0 \leq a \leq 1$