The effects of DNA databases on the deterrence and detection of offenders

Anne Sofie Tegner Anker[*], Jennifer L. Doleac[†], and Rasmus Landersø[‡]

April 2020[§]

[*]Rockwool Foundation Research Unit, Copenhagen, Denmark & Department of Sociology, University of Copenhagen, Denmark. Email: asa@rff.dk.

[†]Department of Economics, Texas A&M University, College Station, TX. Email: jdoleac@tamu.edu.

[‡]Rockwool Foundation Research Unit, Copenhagen, Denmark. Email: rl@rff.dk.

[§]Thanks to David Eil, Naci Mocan, Richard Myers, Emily Owens, John Pepper, and Jay Shimshack; seminar participants at UPenn Law, UVA, Aarhus University, Grinnell College, UC Davis, Indiana University Criminal Justice, Williams College, UCLA Luskin, IDC Herzliya, Tel Aviv University, Hebrew University, Ben-Gurion University, the Inter-American Development Bank; and conference participants at the 2016 Conference on Empirical Legal Studies, the 2017 DC HELD Policy Day, the 2018 STATA TX Empirical Micro Conference, the UCL/CReAM Topics in Labour Economics Workshop, and the Economics of Violence Against Women workshop at the Ifo Institute for helpful comments.

Abstract

This paper studies the effects of adding criminal offenders to a DNA database. Using a large expansion of Denmark's DNA database, we find that DNA registration reduces recidivism within the following year by up to 42%. It also increases the probability that offenders are identified if they recidivate, which we use to estimate the elasticity of crime with respect to the detection probability and find that a 1% higher detection probability reduces crime by more than 2%. We also find that DNA registration increases the likelihood that offenders find employment, enroll in education, and live in a more stable family environment.

1 Introduction

Surveillance technologies have the potential to improve public safety by increasing the probability that offenders are caught for their crimes, thereby deterring criminal behavior. They may also take serial offenders who are not deterred off the streets faster. While the existence and direction of these effects have much support in the literature, we currently know very little about precisely how much deterrence we achieve for any given increase in the likelihood that an offender is apprehended. Furthermore, crime deterrence may have additional benefits through effects on labor market attachment, education, and family life. Understanding these effects is essential for determining how best to use scarce law enforcement resources.

This paper addresses these issues by studying the causal effects of DNA registration of criminal offenders. The goal of DNA registration is to deter offenders and increase the likelihood of detection for future crimes by enabling matches of known offenders with DNA from crime scene evidence. We consider the effects of this intervention on deterrence from subsequent crime and the likelihood that recidivism is detected by law enforcement, and we also provide the first causal estimate of the elasticity of crime with respect to detection probability, a central parameter in the economics of crime first formalized by Becker (1968).[1]

To do this, we measure the effects of a 2005 Danish reform that increased offenders' probability of being added to the DNA database from 4% to almost 40%. The change allowed police to add anyone charged with what is roughly equivalent to a felony in the U.S. (which is the relevant policy margin for most U.S. states considering database expansions), increasing offenders' average probability of being included in the DNA database dramatically.[2] Using the database expansion as an exogenous shock to the likelihood of DNA registration, we

[1] Becker (1968) on pp. 11: "an increase in p_j [detection probability], would reduce the expected utility, and thus the number of offenses, more than an equal percentage increase in f_j [sanctions], if j has preference for risk."

[2] All offenders are subject to improvements in forensic technology throughout this period – including law enforcement's ability to collect DNA evidence from crime scenes and compare them with DNA from suspects. This might have a deterrent effect on everyone. However, being added to the database increases an offender's likelihood of being identified in cases where he would not otherwise be a suspect. The effect of this database-specific increase in the probability of detection is what we estimate in this paper.

estimate that being added to the DNA database reduces recidivism by 6.5 percentage points (42%) in the first year ($p < 0.01$) – a deterrence effect persisting for at least three years.

Using the rich Danish register data, we are further able to explore heterogeneity in effects of DNA registration by previous criminal history, age, and family structure. We find statistically significant deterrence effects for all groups except older offenders. The effects of DNA registration are larger for first time offenders, offenders with children, and offenders initially charged with violent crime, while DNA databases prevent subsequent property, weapon, and violent crime, which supports the hypothesis that offenders frequently commit multiple types of crime instead of specializing in one specific type.

In addition, we find that DNA registration has beneficial effects on subsequent employment, education, and family life. Young offenders are more likely to enroll in education while older offenders are more likely to be employed if they are in the DNA database. Also, first-time offenders are more likely to be married after they are added to the DNA database, and recidivists are more likely to be with the same partner and to live with their children, at least in the short run. These findings are consistent with the hypothesis that keeping people out of trouble (and out of prison) can put their lives on a more positive track. We also report a variety of balancing, robustness, and placebo tests, which support the causal interpretation of our findings.

Quantifying the effects of surveillance tools on crime is often difficult because we only observe that someone offends if he is identified by police. Like many surveillance tools, DNA databases work by increasing the likelihood of such detection. That is, conditional on the same amount of criminal behavior, we will identify offenders more frequently in our data if they are in a DNA database. Improvements in detection thereby lead to an upward bias when we estimate effects on crime. Yet, in this setting, most crimes are solved based on other evidence (such as eyewitness accounts or catching the offender in the act) before DNA evidence could be used to identify the offender. The net effects of DNA registration described above, therefore, provide estimates of the true deterrence effects with only a small upward bias.

However, using the detailed register data, we show how to separately estimate the detection and deterrence effects of DNA registration, which also allows us to provide the first estimates of the elasticity of criminal behavior with respect to detection probability. We exploit the fact that it takes time to analyze and process crime scene DNA evidence, together with the rich Danish register data on the timing of all subsequent reported offenses and charges. We distinguish new charges that might have been aided by the DNA database from charges that were filed so quickly after the offense that this could not be a result of a database match. The first set of (slow) charges is affected by both the deterrence and detection effects of DNA databases, but the second set of (fast) charges provide a clean estimate of deterrence, which we use to separate the two effects.

We estimate a statistically significant detection effect implying that police identify the offender of a crime 3-4 percentage points more often due to DNA registration. The magnitude of the detection effect suggests that economically meaningful deterrence effects could be missed if the two effects of surveillance technologies are not separately identified. These separate estimates of the deterrence and detection effects imply an elasticity of crime with respect to detection probability of -2.7 over three years.

We foremost contribute to the literature on detection and deterrence of crime by showing that DNA registration of offenders increases detection probability, thereby deterring offenders from future crime.[3] To our knowledge, we are the first to estimate an elasticity of criminal behavior with respect to *the probability of detection*. Previous work on this topic focuses on the elasticities of crime with respect to specific inputs such as *police hiring* (these estimates range between -0.1 and -2; see e.g., Chalfin and McCrary, 2017a; Evans and Owens, 2007; Levitt, 1997).[4] Our estimates are consistent with this literature's findings, but we show that the underlying elasticities of overall detection are larger than what is previously reported for

[3]See Chalfin and McCrary (2017b) for a review of this literature.

[4]Chalfin and McCrary (2017a) provide the most precise estimated elasticities of -0.67 for murder, -0.56 for robbery, and -0.23 for burglary.

specific inputs, which is what we would expect if increasing inputs (such as police capacity) by 1% increases offenders' detection probability by less than 1%. We also contribute to this literature by showing that detection not only deters potential offenders from crime – it may also improve their life-trajectories more generally.

Furthermore, the effects of DNA databases on crime have only been analyzed once before. Doleac (2017) uses U.S. data to estimate the net deterrence effect (i.e., a combination of the deterrence and detection effects) based on state variation in DNA databases for recently-incarcerated felons. We build on this by studying a much wider array of outcomes, separating detection from deterrence effects, and analyzing the effect of DNA databases using a cleaner identification design and highly detailed data for a much broader group that is at the current policy frontier in the U.S. (those *charged* with any felony, instead of only those convicted of a felony). We find substantial deterrence effects for this set of less-serious offenders and that effects are larger for first-time offenders, suggesting that the marginal benefits of adding people to a DNA database is largest early in their criminal trajectory.

The large public safety benefits found here are also related to the existing evidence on other high-tech surveillance tools' effectiveness. For instance, electronic monitoring has been found to reduce recidivism (Di Tella and Schargrodsky, 2013; Marie, 2015; Henneguelle et al., 2016). While electronic monitoring has been used as an alternative to pre-trial detention or incarceration and operates through different mechanisms, the results are consistent with our findings that surveillance can provide a substantial, low-cost deterrent for individuals who might otherwise be prone to commit crime.

Finally, our results contribute to a large literature on how to encourage desistance from crime (Doleac, 2019). While much of that literature shows that many popular interventions do not have their intended effects, we show here that DNA databases are effective at reducing recidivism for many groups of offenders.

The paper proceeds as follows: Section 2 describes the background and Section 3 details the empirical strategy. Section 4 describes the data, Section 5 presents the results, and

Section 6 presents estimated deterrence and detection effects separately. Section 7 concludes.

2 Background and the reform of the DNA database

Both before and after the DNA database was created, police solved crimes using a variety of other evidence, such as eyewitness testimony and collecting fingerprints from a crime scene. Police investigators would not need a DNA database to lead them to a suspect in cases where the person was caught in the act, where the victim knew the offender, or where the offender was an obvious suspect (for instance, a husband would be an obvious suspect if his wife was murdered). Thus, while DNA databases are a powerful tool that enables police to find new leads in cases where their standard investigative techniques fall short, there were and still are a variety of other ways that police can solve crimes. The empirical question is whether and how much DNA databases add value above those pre-existing investigative methods.

The Danish Central DNA Database was introduced on July 1, 2000, in order to (i) ease police detection work by identifying offenders and (ii) deter offenders by increasing an offender's probability of getting caught for any subsequent crimes (Justitsministeriet, 1999). The database consists of a person-specific section with DNA samples from suspects, and an evidence-specific section with DNA samples collected at crime scenes or from a victim (Lov om oprettelse af et centralt dna-profilregister, 2000). At the time the database was created, however, only suspects of a limited number of the most serious offenses (e.g., murder, robbery, arson, major violence, incest, and rape) could be included in the person-specific section, and only when the DNA profile was essential to a specific criminal investigation. Likewise, police only collected crime scene evidence from other types of cases if they were suspected to be linked to cases of serious crime and could aid in the apprehension of such offenders.

The process of examining DNA evidence goes through several steps. First, the crime scene is investigated or the offender is sampled and the DNA sample is transferred to the forensic lab at the University of Copenhagen where two independent analyses are initiated to ensure the validity of the result. Next, the DNA sample is 'copied' to ensure even microscopic samples can be analyzed several times.[5] This process takes between one and two weeks. Once the DNA sample is fully analyzed, it is first matched against a database of the staff involved in the DNA collection and analysis to rule out contamination. Then the quality of the sample is used to estimate a likelihood score with <1/1,000,000 as the most precise. These results are summarized in a report, which is sent to the police. According to the Forensic Institute at the University of Copenhagen, the police should expect to wait four weeks for a DNA sample to be processed (95% of samples are processed within four weeks) (Retsmedicinsk Institut, 2014).

2.1 The 2005 reform

The Danish DNA database was expanded on May 24, 2005.[6] The bill introduced two major changes surrounding DNA registration. First, the list of crime types that qualify for DNA registration was vastly expanded to include all offenses where the maximum penalty is a prison sentence of 18 months or more.[7] This is roughly equivalent to adding anyone charged with a felony in the United States. Examples of newly-qualifying offenses include burglary and simple violence/assault. Second, prior to the reform, DNA profiles were only collected and added to the database if they were deemed to be essential to a specific criminal investigation. Thus, offenders who confessed were not obliged to have their DNA added to the database, nor were individuals charged in cases with no DNA evidence (Det Etiske Råd, 2006). The reform eliminated these requirements. Furthermore, the reform also made it easier and cheaper to obtain DNA samples for the database, as it authorized the police to collect the DNA sample instead of requiring medical personnel.

The changes in 2005 had a substantial impact on the likelihood that a charge would result

[5] Conducted via a polymerase chain reaction.

[6] The law was proposed on February 22nd 2005, passed on May 24th 2005 and enacted on May 25th 2005 (Lov om ændring af lov om oprettelse af et centralt dna-profilregister og retsplejelove, 2005).

[7] The law also added possession of child pornography as a qualifying offense, even though the maximum penalty for that particular crime is a prison sentence of 1 year (Justitsministeriet, 2005).

in DNA registration.[8] Figure 1a shows the likelihood that a charged individual was added to the DNA database (see Section 4.1 for more on the sample description). In our sample, the likelihood of being registered in the DNA database increased from 4% in May 2005 to almost 40% in October 2005. In the subsequent years, DNA registration becomes gradually more prevalent and by 2007 almost 60% of charged offenders had their DNA registered. Yet, Figure 1a also suggests that there was a lag in law enforcement's implementation of the new rules for DNA registration in 2005, which we will discuss in detail in Section 3.

For DNA registration to ease the police's detection work and deter offenders, DNA evidence from crime scenes must be collected. Figure 1b shows the evolution of the total number of cases where crime scene evidence is included in the DNA database. The figure shows that the collection of crime scene evidence is steadily increasing through this period. Furthermore, as Figure 1c shows, the reform also coincides with a large increase in the number of matches ("hits") between the offender and evidence sides of the DNA database. This provides preliminary evidence that the reform increased the likelihood of detection for registered offenders.

3 Empirical strategy

To identify the causal effect of DNA registration on an individual's (observed) crime, we need exogenous variation in who is added to the DNA database. We exploit the 2005 expansion of Denmark's DNA database, which introduced a large shock to the probability that someone charged with a crime is added to the DNA database. Offenders charged within a period around the reform are effectively randomized into control and treatment groups based on the precise timing of their charges. Yet, the full policy implementation was delayed until October 2005 due to police officers' summer vacations: police departments were short-

[8] In Denmark, the process of criminal prosecution starts with the police pressing charges if it is assessed that an individual has committed a crime, which then subsequently can lead to a formal indictment and a court case if the state prosecutors believe that the case can lead to a conviction.

staffed during the summer, and the work required to stock the extra DNA collection kits was delayed. Therefore, while the reform motivates an RD strategy, we will treat the change as a 'regular' instrumental variable, excluding months June through September (the summer months immediately after the reform) while conditioning on running variables that count months before May and after October 2005 (a strategy often referred to as 'donut RD').

We estimate effects using two-stage least squares, with a binary instrument Z indicating whether the offender was charged before or after the reform. The first stage is:

$$DNA_i = \gamma Z_i + \mu_1 g(x_i) + X_i \beta_1 \tag{1}$$

where DNA_i is a binary indicator of DNA registration, $g(x_i)$ a flexible running variable counting the months before May and after October 2005, and X_i a set of observable covariates.[9] The second stage for outcome $\tilde{y}_i$ is:

$$\tilde{y}_i = \beta^{IV} \widehat{DNA}_i + \mu_2 g(x_i) + X_i \beta_2 \tag{2}$$

Observations i are at the charge level and we cluster standard errors by individual offender. We argue that the reform satisfies the standard IV / LATE conditions (Imbens and Angrist, 1994): the instrument strongly predicts DNA registration, the exclusion restriction holds, and the reform did not reduce the detection probability for any offenders.

Based on this strategy, we present the estimated effects of DNA registration on subsequent crime in Section 5.2; in Section 5.2.2 we also replicate the main results using a Difference-in-Difference (DiD) strategy. The IV approach assumes that the exclusion restriction holds (being charged after May 2005 is related to subsequent recidivism only through the charge date's effect on DNA registration), while the DiD approach (based on an intensity of treat-

[9] In our main specifications we define $g(x_i)$ as a linear function of the running variable where slopes may differ from pre- to post-reform to capture different trends in crime across time. We also present robustness tests using more flexible functions for $g(\cdot)$. X includes: age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, offense type, and month fixed effects.

ment measure) instead assumes that offenders who were less affected by the reform are a good counterfactual for those who were more affected by the reform. The finding of qualitatively similar effects using the two different approaches strengthens our causal interpretation of the estimated effects.

In Section 6 we show how we separate the deterrence and detection effects, and estimate the elasticity of crime with respect to detection probability. All of the results in Section 5 should be thought of as the net of deterrence and detection – that is, the deterrence effect with an upward bias (though in this setting the bias is small because in practice most crime is solved without the aid of the DNA database).

4 Data

We focus solely on adult offenders, for whom the judicial system bears close resemblance to those in other OECD countries.[10] We use Danish full population register data with information on all residents. Unique individual identifiers allow us to merge information on involvement with the criminal justice system and demographic characteristics among others, and the identifiers also allow us to link each individual to family members and partners.

4.1 Sample definition

We construct the sample from two main data sources: (i) the charge register, which contains information on the crime date, charge date, and crime type, and (ii) records of all the individuals in the person-specific section of the DNA database. Both data sources contain unique personal identification and record numbers allowing us to merge them and identify the cases for which offenders were added to the DNA database.

In our main sample we include charges that occurred between June 2003 and September

[10]While Denmark differs from the U.S. in many respects, average crime rates are overall similar across the two countries: See pp. 207 in OECD, 2005 and http://www.oecdbetterlifeindex.org/topics/safety/. Substantial differences exist for specific crime-types as for example gun-violence or homicide.

2007.[11] Due to the lag in police practice in terms of implementing the new rules concerning DNA registration, we exclude the months of June-September of 2005, and use a 24 months sampling window on either side of that period. We choose the bandwidth of 24 months on the basis of a cross-validation (CV) procedure (as described in Lee and Lemieux, 2010, and Ludwig and Miller, 2005) in order to minimize prediction error close to the reform.[12]

Besides the time frame, the charges included in the sample have to fit four criteria: (i) the charge has to be for an offense against the Penal Code or Weapons Act; the latter mainly consists of illegal possession of explosives, firearms and other weapons (see Table A-1 for Danish crime categorizations). These include the vast majority of criminal offenses, and so we only discard individuals charged with traffic offenses, small-scale drug possession and offenses such as Health Code and Tax Law violations. (ii) Individuals have to be a resident of Denmark.[13] (iii) We only include charges against males aged 18-30 at the time of the charge. This group is the most criminally active and is the most relevant for estimating effects on criminal behavior. (iv) To avoid giving individuals who are charged with several crimes within the time frame disproportionately-high weight in the analyses, we only include charges against men who at the time of charge have had a maximum of 10 previous charges.[14]

Our unit of observation is a charge. To illustrate how we handle multiple charges against the same individual, suppose individual i is charged initially at time t_0. This will enter as one observation with any subsequent recidivism in the following years $t_0 + 1, t_0 + 2, ...$ recorded as outcomes linked to that observation. A subsequent charge to individual i at, for example,

[11] Only one charge per person per day is included to avoid having crimes that violate several different laws disproportionately represented in the data.

[12] The cross-validation procedure consists of two steps. First, we estimate the reduced form regressions with a dummy variable indicating before/after June-September of 2005 and running variables measuring months before or after (+ covariates), but leave out observations in the 1-3 month preceding June and following September. Second, we use the estimates to predict the outcome for the observations in the excluded window around the reform, and calculate the mean prediction error (or CV functions) for each outcome which we finally aggregate across the outcomes and across the 1-3 month prediction windows. We have done this for bandwidths from 10—40 months before/after the reform. Figure A-1 shows that a bandwidth of 24 months yields the best prediction.

[13] This implies that we exclude tourists and individuals from other EU countries moving freely within the EU without being registered with a Danish social security number.

[14] Different caps on maximum number of charges do not change our conclusions (see table A-2).

time $t_1 = t_0 + 1$, will enter as a new observation (if t_1 falls within the sample window) with recidivism in years $t_1 + 1, t_1 + 2, ...$ as outcomes. While this ensures that we do not select the sample on outcome variables, one might still worry that the sampling coincides with our instrument because we sample some individuals more than once (those who are charged several times within our sample window). We are confident that this is not affecting our results, for three reasons. First, our results hold when we focus on first-time offenders, which avoids repeated observations and any potential selection associated with this. Second, while our design is not formally a discontinuity, we estimate effects conditional on the running variable and effectively compare individuals charged within a small window of time; this reduces the number of charges per person. Finally, in Section 5 we implement placebo tests (placebo reforms in other years or using previous charges as outcomes) which all produce near-zero and insignificant estimates. Hence, there is nothing mechanical in our sampling generating spurious effects.

Our sample consists of 38,674 individuals who received a charge that fits the aforementioned criteria, with a total of 66,911 observations. As multiple charges against the same person are not independent observations, we cluster standard errors at the individual level.

4.2 Outcome variables

We use convictions for crimes committed after the charge in question as the outcome. Our main outcome is *all crime*, but we also consider violence, property crime, sexual offenses, other penal offenses (including serious drug offenses), and Weapons Act violations separately.

As our unit of observation is a criminal charge, individuals may appear several times in the data. We define the outcomes from the time of the charge for which an individual enters the sample. Counting from the day after this charge, we measure subsequent crime for which the individual is convicted within one, two, and three years. All crime measures are coded in a binary version indicating at least one conviction and in a version that counts the number of convictions within the one-, two- and three-year follow-up periods.[15]

In Section 6, we will distinguish between convictions for which the charge occurred three weeks or less after the crime date, and convictions for which the charge occurred more than three weeks after the crime date in order to separate the charges where prior DNA registration may have contributed to the detection of the offender.[16] Because the analysis of crime scene evidence takes time (cf. Section 2), it is not possible that a match in the DNA database led police to the offender if he was charged shortly after the crime. Any effect of DNA registration on the outcome measure during that window would come solely from a deterrence effect. Afterwards, DNA registration may have both a deterrent and a detection effect. While the processing time may take four weeks, we set the limit at three weeks as some samples may be processed faster.[17]

Although recidivism is our primary outcome, we also examine whether DNA registration affects labor market outcomes and family stability, which a large criminology literature identifies as one of the chief predictors of crime desistance (see e.g., Sampson et al., 2006). We use register data on labor market attachment to define the labor market outcomes by three mutually exclusive categories: (1) employed (i.e has a job), (2) in education or training, and (3) unemployed. We measure labor market attachment as the time during the first four years following the initial charge that the individual spends in each of these categories.

For measures of family stability, we use the timing of changes in marital status and home addresses to measure whether the individual is married, remains in the same relationship if he had a partner (by marriage or cohabitation) prior to the initial criminal charge, and lives with his child and the child's mother if he had children prior to the initial criminal charge.[18]

[15] Estimated long run effects may be attenuated, as those who are not added to the DNA database initially may be added (and treated) with increasing likelihood if they recidivate in subsequent years. Also, "number of crime convictions" is top-coded at a maximum of 10 convictions per follow-up year to limit outliers' impact.

[16] Overall, 80% of offenders are charged within three weeks of the crime (conditional on the offender being identified). For property crime, the fraction is around 75% while it is around 85% for violent crime.

[17] Results are robust to reducing the limit for fast charges to for example two weeks.

[18] We observe the unique individual identifier and home addresses of the full population, which allows us to identify whether a given offender lives with a partner and any children. The measure of the father living with his child and his child's mother is constructed for each of his children (born prior to the initial charge), and for this outcome the father appears in the sample once for each child and charge.

4.3 Data Descriptives

Table 1 shows average characteristics of the full sample and divided by whether the charge took place before or after the reform. Overall, individuals charged with crimes have 11 years of education, only slightly above the compulsory level in Denmark (9 years). Their annual incomes are low – about 112,000DKK ($17,500) – and nearly half are unemployed at the time of the charge. Most (86%) are single but a small share (12%) have children. Immigrants are heavily overrepresented, making up 21% of the sample (relative to less than 10% in the full population). Almost 40% live in one of the four largest cities. Table 1 also shows the sizes of the subgroups. For example, 24% enter the sample on their very first charge, whereas the rest have between 1 and 10 charges behind them (the overall mean is 3 previous charges).

Table 2 shows average complier characteristics – offenders whose DNA registration was induced by the reform – along with full sample means for comparison. The table shows that a larger share of the compliers belong to the younger age-category compared to the whole sample, fewer have children, and fewer enter the sample on their first charge. The compliers are also less educated and have a lower gross income, but are just as often unemployed. In terms of previous crime, the compliers are more often violent and sexual offenders compared to the overall sample. Still, most categories of offenders are well-represented within the complier group and our instrumental variable provides large and significant increases to the probability of DNA registration in all subsamples. This will allow us to consider heterogeneity of effects by offender characteristics while also supporting monotonicity of the IV.

Panel A in Table A-3 shows summary statistics of the crime outcomes by timing of the charge relative to the reform. On average, 15% and 11% of the pre and post reform groups, respectively, are convicted for another offense within one year. After three years these numbers are 38% and 34%, respectively, corresponding to 0.65 and 0.55 convictions for the pre and post reform groups. The most prevalent crime type is property crime, which constitutes approximately 55% of all recidivism. Almost 30% of recidivism is violent crime, while sexual offenses constitute less than 1%, and weapon-related and the residual 'other crime' (mainly drug-related offenses) each constitute around 7% and 8% of recidivism respectively.

Panel B presents summary statistics for labor market outcomes. During the first four years after the initial charge, on average around 1.9 years are spent in employment, 1.9 years in unemployment, and the remaining time spent enrolled in an education or training program. Panel C in Table A-3 summarizes marital status outcomes. Only 4.6% of the full sample are married by the time of the initial charge, a share that increases to 5% one year later and to 7% after three years. When looking at those who have a partner (by marriage or cohabitation) prior to the initial charge, 46% of them are with the same partner one year after the charge. For the offenders who have at least one child at the time of the initial charge, the probability that the father lives with the child and the child's mother is 30%.

5 Results

5.1 Validity of the reform as an instrument

Below we provide balancing tests showing that the reform provides a clean identification of the effects of DNA registration. As described above, we exclude June-September 2005 from our main analysis. Offenders charged between June 2003 and May 2005 make up our control group, those charged between October 2005 and September 2007 compose our treatment group, and our identifying assumption is that offenders' propensity to recidivate, conditional on their non-treatment characteristics, does not change between May 2005 and October 2005 (the summer months after the effective date of the 2005 DNA database expansion).

Following Pei et al. (2017), Table 3 shows results of regressions that test for discontinuities in the covariates by regressing each covariate on a dummy indicating whether the charge occurred after the reform, conditional on a running variable counting the number of months before and after the reform (and month fixed effects in column 2). According to the table, there are significant differences around the reform for a few covariates. Most striking is the estimates for crime type leading to the intitial charge. However, as the categories are mutually exclusive, one negative significant estimate must necesarily have a positive counterpart. Moreover, what matters for our analysis is whether those differences in individual characteristics are meaningful enough to affect offenders' propensities to reoffend (given our large sample size, we have sufficient statistical power to precisely estimate even differences that are not economically meaningful). Figure 2 shows offenders' predicted propensities to reoffend based on the pre-treatment relationship between observable characteristics and recidivism, for individuals charged before and after the reform.[19] Both the probability of committing any crime and the number of predicted crimes are smooth through the threshold.

Table 4 presents the regression equivalent of Figure 2, testing for discontinuities in outcomes predicted by the covariates. The distribution of predicted recidivism (based on observable characteristics) is indeed smooth through the threshold; we see no significant differences in this measure just around the reform, which makes it highly unlikely that the small differences in covariates seen in Table 3 bias our results. We will show that our first and second stage results are virtually unaffected by the inclusion of covariates, more evidence that these differences are not meaningful. We present a final balancing test at the bottom of Tables 6, 7, and 8: we regress pre-period outcomes on DNA registration and find no significant pre-period 'effects.' We furthermore conduct a McCrary test (McCrary, 2008) on the number of charges in our sample (excluding the summer months of 2005). Figure A-2 shows no significant discontinuities in the distribution at the threshold, allaying potential concerns that the timing of charges could have changed as a result of the reform.

Moreover, one might be concerned that the reform changed police behavior with respect to evidence collection or charges of suspects if, for example, the database expansion made police more aware of the value of DNA evidence and more careful to only charge defendants when such evidence was present. However, such a change in police behavior would affect all

[19] We use pre-reform data to regress recidivism on observable characteristics. Using the estimated coefficients, we predict recidivism based on observable characteristics for the full sample, and test for a discontinuity in this predicted measure. See e.g., Card et al. (2007) for a similar test and argumentation in relation to balancing of covariates and predicted outcomes in a discontinuity design.

active offenders (those with initial charges before or after the reform), regardless of whether they are in the database.[20] That said, we find no discontinuity in the likelihood of a charge leading to conviction across the reform, which shows that charges are not becoming more accurate as a result of the database expansion (Figure A-3). In addition, one consequence of the reform could also be that offenders avoid detection because they become increasingly careful not to leave DNA evidence at the crime scene. While such behavior would impede the reform's intended effects, it should not bias our results. If all offenders leave less DNA evidence behind then this will make the reform less effective, and we simply will not find any impact on crime rates or recidivism.[21]

Finally, the reform could have a general deterrence effect by making all would-be offenders aware that DNA registration could link them to past crimes via old crime scene evidence. This would imply that the reform not only changed the probability of being caught for those in the database, but also that the sanctions associated with being added to the DNA database in the first place (at which point they would be caught and punished for previous crimes, in addition to the new one). This could change who chooses to commit a crime after the reform, changing the composition of the sample across the timing of the policy change. None of the tests provided above suggest that this is the case as the treatment and control groups are balanced through the threshold defining our instrumental variable.[22]

[20] In particular, police can always get a warrant for a DNA sample from a suspect to compare with crime scene evidence; the difference for those in the database is that they might be matched to cases in which they would not otherwise have been a suspect.

[21] If only offenders in the database become more careful to avoid leaving DNA at the scene, this could bias our estimates downward, but we think that (1) this is less likely than everyone becoming more careful, and (2) that any effect on detection would be small. It is extremely difficult to avoid leaving DNA at a crime scene – humans shed skin cells constantly, so destroying DNA at a crime scene would require extensive effort and planning (e.g. bleaching the crime scene). As offenders frequently leave fingerprints at crime scenes, which is much easier to avoid by wearing gloves or wiping their prints off of surfaces they've touched, it seems unlikely that any but the most sophisticated offenders would take the actions necessary to eliminate their DNA from a crime scene. Also, Figure 1b shows the offender-evidence DNA-hit rate increases substantially after the reform, illustrating that offenders do not become careful enough to avoid detection by this technology.

[22] But even in the absence of compositional changes, if a share of new convictions due to database hits are for old cases, this would change the interpretation of our results substantially. If this is the case, we should see that the reform increased charges and/or convictions for crimes that were committed before but solved after DNA registration. In Table A-4 we test this by estimating the changes to charges and convictions for crimes that were committed *before* the specific charge that leads to DNA registration, but where charges were not pressed until *after* the DNA registration. All estimates are close to zero and insignificant showing

5.1.1 First stage results

Figure 1a illustrates the first stage effect of the DNA database expansion on the probability that a charge results in DNA registration. The summer months of 2005 are shown in grey. After excluding those months, there is a clear shock to the probability of DNA registration, with the the probability changing from 4% to almost 40%. Table 5 formally presents the first stage estimates. The reform increased the likelihood of DNA registration by 35 percentage points, which is a highly statistically significant increase ($p < 0.001$).

5.2 Main results

Figure 3 shows monthly averages relative to the sample mean of the probability of being convicted for a crime and the number of convictions for crimes committed within the first year following the initial charge (the excluded summer months of 2005 are shown in grey for transparency in the figure but are excluded from our regressions). The figure provides a first visualization of our main findings: recidivism decreases substantially following the reform.

Table 6 presents the estimated effects of DNA registration on subsequent convictions 1, 2 and 3 years after the initial charge, with standard errors in parentheses.[23] Columns 1–3 show effects on the probability of any subsequent conviction, and columns 4–6 show effects on the number of subsequent convictions. Columns 3 and 6 (our preferred estimates) show that DNA registration reduces the probability of a new conviction by 6.5 percentage points in year 1 (42%, $p < 0.001$), and the number of convictions by 0.093 (49%, $p < 0.01$). All estimates are economically meaningful and at least marginally significant.

Finally, the table presents placebo tests where we regress DNA registration on charges measured prior to the sampling charge in question. If we are isolating the causal effect of DNA registration on subsequent behavior, these estimates should be statistically insignificant.

that the increased DNA registration induced by the reform did not increase the likelihood that offenders were convicted for crimes committed before being added. Hence, our estimated effects are driven by a reduction in new crimes.

[23] As the inclusion of covariates does not affect point estimates but increase precision, all remaining tables present results conditional on covariates.

Indeed, the estimates are small and p-values range between 0.76 and 0.95.

We thus find that DNA registration reduces criminal recidivism substantially. Since the reform induced a very large share of criminals to be added to the database, effects of the sizes found here should be visible in the overall crime reports (if we have captured actual changes in crime and not just changes in factors such as offenders' precautionary measures). Figure A-4 shows exactly such a change in crime reports by plotting all reported crimes and reported burglaries (a common property crime) from January 2004 to December 2006. Relative to April-June 2005, the total number of reported crimes drops by around 5% while the number of reported burglaries drops by 5-10% following the reform.

5.2.1 Heterogeneity

A frequent topic of policy debate is which categories of offenders should be included in a DNA database. Should only serious offenders or violent offenders be added, once they have confirmed that they are a threat? Or is there value in including a broader set of individuals, in the hope of catching or deterring would-be serious offenders earlier in their criminal careers?

The left half of Table A-5 presents the estimated effect of DNA registration on subsequent crime convictions by the initial charge's crime type. Effects are strongest for violent offenders, where DNA registration reduces the probability of a subsequent conviction by almost 50% ($p < 0.01$) relative to the pre-reform mean. That effect persists through year 3. The table also suggests that offenders initially charged with property, weapons-related, or other penal offenses reduce crime following DNA registration, with some marginally significant estimates.

To examine the types of crime prevented by DNA registration, the right half of Table A-5 presents the effect of DNA registration by subsequent types of crime. DNA registration reduces the likelihood of a property crime conviction by 3.1 percentage points (34%, $p < 0.10$) and the likelihood of a violent crime conviction by 3.1 percentage points (63%, $p < 0.05$) during the first year. Both effects persist – at least in magnitude – for three years. The likelihood of a conviction for weapon offenses decreases by 1 percentage point (91%, p

< 0.10) during the first year, while sexual and other penal offenses (a small share of total crime) appear unaffected by DNA registration with estimates near zero.

Panel A in Table A-6 presents estimates of DNA registration on subsequent convictions separately for first-time offenders (those who enter our data for their first-ever charge) and recidivists (those who have at least one previous charge). The top half of the table shows effects on the probability of any subsequent criminal conviction. Overall, estimates for first-time offenders and recidivists are quite similar in magnitude. Yet, as pre-reform baseline recidivism rates differ between the two groups (7% of first-time offenders reoffend within one year compared to 23% for the rest of the sample) first-time offenders' 4.8 percentage point lower recidivism constitutes a 71% reduction, while recidivists' 6.8 percentage point decline constitutes a 30% reduction. The bottom half of Table A-6 shows effects on the number of subsequent convictions. Here the same pattern emerges, though we only see statistically significant effects for recidivists. Panel B in Table A-6 divides offenders by age. Effects are mainly driven by 18-23 year olds, particularly in year 1.

Panel C in Table A-6 shows effects separately for those who have at least one child by the time of the initial charge (12% of the sample) and those who do not. Deterrence from crime may be easier when offenders have children to serve as a role model for. Both groups reduce their crime, but the deterrence effects for fathers are especially strong when compared to the baseline recidivism rates, which are 20% lower than for those without children at the time of charge. For the fathers, all effects are consistently negative and large in magnitude.

5.2.2 Difference-in-Differences (DiD) estimates

The IV estimates above depend on the exclusion restriction assumption (being charged after May 2005 affects recidivism only through its effect on DNA registration). As we drop the summer months of 2005, our analysis rests largely on a comparison of criminal behavior in the spring and fall of 2005. To test the robustness of our findings, we next estimate the effects of the reform using a DiD design, which is based on the assumption that, in the absence of the policy change, the behavior of the treatment group would have evolved similarly to the behavior of a control group (this is alternatively referred to as a 'parallel trends' assumption).

To estimate the effect of the reform in a DiD framework, we need to define both a treatment group and a comparison group that provides a good counterfactual. This is not straightforward as DNA registration becomes more prevalent in all broad categories of crime following the reform.[24] At the same time, the few crime types that led to registration pre-reform, such as homicide, are too rare to provide sufficient statistical power. We therefore create a treatment-intensity measure based on the share of offenders in each crime type that were added to the database post-reform. We define a high-DNA (treatment) group as offenders of crime types where 75% or more were registered post-reform; offenders of crime types with less post-reform registration are in the low-DNA (comparison) group.

Figure A-5 shows the probability of a new conviction within one year for both groups, from 24 months before the reform until 24 months after the reform. Figure A-5a shows the raw levels while Figure A-5b shows the demeaned levels relative to average crime in the year preceding the reform. While the two groups have different levels of recidivism (Figure A-5a), Figure A-5b shows that the pre-period parallel trends assumption is met. Furthermore, both groups' recidivism drops following the reform (as offenders in both groups are significantly more likely to be added to the DNA database), but the crime reduction is larger for the high-DNA group. The gap between the high-DNA and low-DNA groups begins to widen at about six months post-reform, consistent with the reform's delayed implementation.

Table A-7 presents the DiD estimates of the reform on subsequent convictions 1, 2 and 3 years after the initial charge.[25] The estimates correspond to the difference between the high-DNA and low-DNA groups in the right part of Figure A-5a (or b) net of the difference between the two groups in the left part of the figure. The reform led to significantly less

[24] We cannot define as treatment and control groups crime grouped as "minor" offenses and "serious" offenses, as most minor offenses are categorized together with more serious offenses in the Penal Code. Shoplifting is, for example, simply "theft" in the Penal Code and hence also affected by the reform with rapidly increasing prevalence of DNA registration

[25] We estimate this as: $y_{it} = \alpha + \gamma_1 \mathbf{1}[post_i] + \gamma_2 \mathbf{1}[Treatment_i] + \gamma_3 \mathbf{1}[post_i] * \mathbf{1}[Treatment_i] + \epsilon_{it}$ where γ_3 is the DiD estimate.

crime 1, 2, and 3 years after the initial charge. As a last robustness check, Figure A-6 presents estimates using the DiD specification for different definitions of the pre and post period. The left part of the figure presents placebo estimates. All estimates are close to zero and insignificant. Only if we set the pre/post cut-off to 4, 5, 6, 7, or 8 months after the reform (the months where DNA registration has stabilized around 40–60%, see Figure 1a), the estimates are negative and significant. This corresponds well with the observed delayed implementation of the reform, which motivates the donut RD-approach.

5.2.3 Additional robustness tests

We perform a series of additional robustness tests. We run a series of placebo tests (Table A-8), which artificially impose reforms in years other than 2005. Significant reduced form estimates occur only in the year of the actual reform. Hence, our estimation strategy does not attribute effects to arbitrary fluctuations in crime.

Table A-9 shows results while keeping the summer months of 2005 in the data. Across the board, the table replicates our main effects, although with less precision. Our results are also robust to different sample window definitions (Table A-10) and running variable specifications (Table A-11).

Finally, Table A-12 presents results where convictions are adjusted for the time incarcerated in the follow-up period to eliminate any bias that may occur if detection effects change incarceration rates and thereby also incapacitation. We divide the number of convictions by the proportion of the follow-up period where an individual was not incarcerated leading to estimates that are numerically larger but otherwise similar to the main results.

5.3 Non-crime effects of DNA registration

The consequences of crime have been linked to a variety of other outcomes that may in turn lead to even more crime, through effects on one's network, time available for investment in other activities, and because the stigma of a criminal record might limit future opportunities.[26] Deterrence from crime could in turn improve other outcomes. We therefore estimate the effect of DNA registration on labor supply, education, and family relationships.

Table 7 presents our estimates of the effects of DNA registration on years spent employed, in education or training, or unemployed during the four years after the initial charge (the categories are mutually exclusive). The first column shows effects for all offenders. While average time spent employed does not change, the number of years spent in education or training increases significantly by 0.098 years (1.2 months). This is a dramatic increase relative to the pre-reform mean. This education effect was driven by young offenders, as shown in the second column. They appear to shift from employment to education or training. This is consistent with their investing in human capital to have better legal employment options in the future. Older offenders' education is not affected, but they spend less time unemployed and spend four more months employed if they are added to the DNA database.

Table 8 shows the estimated effects of DNA registration on the likelihood of being married, the likelihood of remaining in the same relationship as before the initial charge (given that the offender was in a relationship), and the likelihood that the offender lives with his children and their mother (if the offender has children).

Columns 1–3 show effects for all offenders. Columns 4–6 show effects for first-time offenders only (less-hardened offenders, for whom lower recidivism may have a more substantial effect on other aspects of their lives), and columns 7–9 show effects for recidivists only.

We see no statistically significant effects for the full group of offenders, though the imprecisely-estimated coefficients imply economically meaningful effects. One year after their initial charge, offenders in the DNA database are 0.7 percentage points (12%) more likely to be married, 11.0 percentage points (24%) more likely to live with the same partner, and 12.4 percentage points (40%) more likely to live with their child and the child's mother.

For first-time offenders, the effect on the likelihood of marriage is a 3 percentage point (43%, $p < 0.05$) increase after the first year. This estimate grows in magnitude and remains

[26] For example incarceration (e.g., Aizer and Doyle, 2015), labor market outcomes (e.g., Grogger, 1998; Raphael and Winter-Ebmer, 2001; Mueller-Smith, 2015), and family formation (e.g., Laub et al., 2008).

statistically significant through the third year. Estimates of the effect of living with the same partner are initially near-zero, and remain statistically insignificant, though the relevant sample is small. For recidivists, we see no impact of DNA registration on the likelihood of being married (all coefficients are near-zero), but there is suggestive evidence that DNA registration increases the likelihood of living with the same partner as before DNA registration: offenders in the database are 13.1 percentage points (30%, $p < 0.10$) more likely to live with the same partner one year later, though that estimate falls to 4.4 percentage points by year 3. DNA registration increases the likelihood that an offender lives with his child and the child's mother by 15.3 percentage points (57%, $p < 0.05$) after one year, though that effect size again falls, to 6 percentage points after year 3.

Overall, these results point to criminal behavior – or desistance therefrom – often being interwoven with labor market attachment and family life. Our findings illustrate that policies affecting offenders' recidivism also have implications for a wider array of outcomes. The results also touch on the indirect consequences of criminal behavior. A disproportionate number of children with criminal fathers grow up with divorced parents and/or with an unemployed or absent father (see e.g., Wakefield and Wildeman, 2014) thereby strengthening intergenerational persistence of poverty, risky behavior, and crime. DNA registration may help to break elements of this vicious cycle via the effects on fathers' criminal behavior.

6 Deterrence, detection, and elasticities

6.1 Theoretical framework

Standard economic models suggest that the propensity to commit crime is a negative function of the expected punishment for that crime. As initially formulated by Becker (1968), an individual will commit crime when the expected benefits exceed the expected costs:

$$y_i = 1[\alpha_i - c_i > 0] \tag{3}$$

where α_i summarizes the expected benefits from crime (monetary and non-monetary payoffs) and c_i the expected costs (an increasing function, $f(p, s)$, of the detection probability, p, and sanction if convicted, s). This paper studies how changing p – by adding an offender to the DNA database – affects crime. DNA registration increases the detection probability from $p = \bar{p}$ to $p_i = \bar{p} + \gamma DNA_i$. Hence, crime in the two counterfactual states, y_i^0 and y_i^1, equals:

$$\begin{aligned} y_i^0 &= 1[\alpha_i - f(\bar{p}, s) > 0] \\ y_i^1 &= 1[\alpha_i - f(\bar{p} + \gamma DNA_i, s) > 0] \\ \Delta &= y_i^1 - y_i^0 \end{aligned} \tag{4}$$

We label Δ 'the deterrence effect of DNA registration'. Yet, we face two problems. First, we do not observe y_i^0 and y_i^1 for the same individual and we have to address the endogenous relationship between unobservable characteristics and DNA_i (we described our empirical strategy for this in Section 3). But we also face another problem common to studies of crime: we cannot link crime to offenders unless they are caught. Thus, we only observe crime with probability $\bar{p}$ and $\bar{p} + \gamma DNA_i$ without and with DNA registration, respectively:

$$\begin{aligned} \tilde{y}_i^0 &= \bar{p} * y_i^0 \\ \tilde{y}_i^1 &= (\bar{p} + \gamma DNA_i) * y_i^1 \end{aligned} \tag{5}$$

Hence, even if we observed an individual in both counterfactual states we would get:

$$\tilde{y}_i^1 - \tilde{y}_i^0 = \bar{p} * \Delta + \gamma DNA_i * y_i^1,$$

instead of the desired quantity, Δ. Hence, in addition to the endogenous relationship between offenders' unobservable characteristics and DNA registration the observed change in crime as a result of the DNA registration is attenuated because only a fraction of crime ($\bar{p}$) is observed, and because there may be an upward bias because DNA registration increases the fraction of crime that is observed (where offenders are caught) – that is the purpose of the technology. We define this latter source of bias as 'the detection effect':

$$\delta = \gamma DNA_i * y_i^1 \tag{6}$$

From the deterrence and detection effects, we define a central policy parameter: the elasticity of crime with respect to the detection probability, ϵ. We define this as (i) the percentage change in crime divided by (ii) the percentage change in the detection probability. As the deterrence effect, Δ, is estimated as the absolute and not relative reduction in crime, it is adjusted by the baseline level y^0 to be expressed in percentages as in point (i). Likewise, the detection effect, δ (the absolute change in detection rates) is adjusted by baseline crime levels y^0 and the baseline detection rate $\bar{p}$ to yield point (ii). Hence, the elasticity is expressed as:

$$\epsilon = \frac{\Delta / y_i^0}{\delta / y_i^0 * 1/\bar{p}} = \bar{p} * \frac{\Delta}{\delta.} \tag{7}$$

This result rests on offenders' ability to assess the detection probabilities. The key object for offenders' behavior is the perceived detection probability (Durlauf and Nagin, 2011). Offenders are clearly aware of DNA registration in the present context, as individuals observe and participate in the DNA sampling. Yet, if offenders' perceived risk of apprehension is biased, our estimates should instead be interpreted as the effects of changing the perceived detection probability, and the magnitude of the bias will determine the difference between the elasticities of crime with respect to actual versus perceived detection probability.[27]

[27] If offenders are overestimating DNA databases' effects on p, perhaps due to futuristic crime shows on television, then we would expect them to learn over time through personal experience or word of mouth what the true p is. At the same time, steady improvements in DNA technology have increased p over time. Going forward, net effects on behavior will depend on whether the technology improves faster than offenders adjust their biased perceptions.

6.2 Empirical strategy: separating detection and deterrent effects

The framework shows that the estimated effect of DNA registration consists of two underlying effects (for the compliers who are added to the DNA database as a result of the reform):

$$\begin{aligned} \beta^{IV} &= E(\tilde{y}_i^1 - \tilde{y}_i^0) \\ &= E(\bar{p} * \Delta + \gamma DNA_i * y_i^1) \end{aligned} \tag{8}$$

taking conditioning on covariates as implicit. There is a behavioral response to an increased detection probability after being added to the database (deterrence effect), and an increased probability of being apprehended due to a DNA match (detection effect). Separating the two effects will provide key information about how DNA registration affects criminal behavior. We do this by exploiting the Danish register data, which includes both when offenders are charged for a crime and the exact date of that crime. We divide observed crime $\tilde{y}_i$ into two categories: crime with a fast charge, $\tilde{y}_i^F$, and crime with a slow charge, $\tilde{y}_i^S$.

The former, $\tilde{y}_i^F$, denotes crime solved within three weeks from the date of the offense, before any DNA evidence from the crime scene could have been processed. The latter, $\tilde{y}_i^S$, denotes crime solved after three weeks from the date of the offense, at which point DNA evidence could have been processed and used to identify a suspect. Hence, changes in crime solved within three weeks from the date of the crime will only capture the deterrence effect, while changes in crime solved more slowly will be a composite of both the deterrence and detection effects (i.e. the combined effects on the likelihood that a crime occurs and that we observe it in the data). We are thereby able to identify both effects on criminal behavior:[28]

$$\begin{aligned}\text{Deterrence effect: } E[\Delta] &= (\beta_F^{IV})/(\pi\bar{p})\\ \text{Detection effect: } E[\delta] &= \beta_S^{IV} - \beta_F^{IV} * (1-\pi)/\pi \\ \text{Elasticity: } E[\epsilon] &= \beta_F^{IV}/(\pi\beta^{IV} - \beta_F^{IV})\end{aligned} \tag{9}$$

6.3 Results

Figure A-7 shows monthly averages of crime outcomes relative to the sample mean for crimes committed within the first year following the initial charge, separating crime into *fast charges* and *slow charges*. There is only a drop in crime for the former crimes leading to a fast charge, and the figure thereby gives a first visual impression of the different effects of DNA registration across time it takes to charge the offender.

Table A-13 shows the estimated effects of DNA registration on subsequent crime from *fast charges* and *slow charges*. From the table, we see that DNA registration reduced crime from fast charges. For example, Panel A shows that in year 1, DNA registration reduces the likelihood of recidivism by 5.7 percentage points (43%, p<0.01) and the number of new offenses by 0.076 (48%, p<0.01). For convictions following 'slow' charges, all estimates are

[28] Appendix B.1 shows how Equation (9) is derived. We assume that the baseline clearance rate of crime without the DNA database $\bar{p}$ occurs at a fixed rate and that it is uniform and invariant with offender characteristics that are not captured by the different crime types. Underlying this is three 'invariance' assumptions: (i) Procedures in the justice system did not change along with our IV except through the increased probability of detection $\bar{p} + \gamma DNA$. In support of this assumption, we find that there were not any changes in characteristics of charged offenders nor to the share of charges that lead to a conviction that coincide with our IV. We discuss this and provide balancing tests in Section 5.1. (ii) To compute $\pi\bar{p}$ and $(1-\pi)\bar{p}$, $\bar{p}$ must be invariant across crimes that are potentially solved 'fast' and 'slow'. Appendix B.2 relaxes this assumption and shows that this does not affect our estimated elasticities. In fact, the estimate we report in the main text can be thought of as a weighted average between the elasticities for potentially fast solved crime and potentially slow solved crime. If, for example, fast solved crimes are "low hanging fruit" committed by less skilled criminals and the underlying clearance rate is actually higher than for slow solved crime, then the elasticity of fast solved crime will be smaller (numerically larger). Yet, the average elasticity reported in the main text is unchanged. (iii) Our IV estimates are homogeneous between fast and slow solved crime. Appendix B.3 considers the consequences if this assumption is violated, and show that the resulting bias is not large. E.g., if the deterrence effects for potentially fast and potentially slow solved crimes differ by 20%, the estimated elasticity will be biased by approximately 10% (i.e. be either -2.9 or -2.4 instead of -2.7, depending on the gap's sign).

closer to zero and insignificant.[29]

We now use the distinction between convictions with charges filed within three weeks of the offense and those with charges filed more than three weeks after the offense, to separately identify the deterrence and detection effects of the DNA database. We will then use those estimates to calculate the implied elasticities of crime with respect to detection probability. Table 9 shows the estimated deterrence effects, detection effects, and elasticities as defined in Equation (9).[30] The table shows results separately for the main crime categories: all crime, property crime, and violent crime.[31]

The estimated deterrence effects are based on the above estimates for 'fast' charges (Table A-13), but scaled here by the inverse of the clearance rate. These estimates therefore show not only the change in convictions but the change in actual crimes committed. Table 9 adds further to our results by also estimating the detection effect. For all crime, we see that DNA registration increases the number of new crimes that are detected by approximately 0.077 crimes, and the probability of any subsequent detected crime by 3.6 percentage points. These effects represent 4–5% of pre-reform baseline crime. The results also show that the increasing number of matches between offenders and evidence in the DNA database (Figure 1c) did indeed reflect increased detection and not only that the DNA database served as a substitute for other detection work by the police.

Finally, the table shows estimated elasticities of crime with respect to detection probability. The estimated elasticity is -2.7 by year three, implying that a 1% increase in the likelihood of being caught reduces crime by 2.7%. While violence is more responsive to detection in absolute terms, the fact that the baseline clearance rate for violence is approximately 80% results in a lower elasticity with respect to detection probability (-2.7) in comparison

[29]The effects presented previously on 'all crime' confirm that the differences between crimes with 'fast' and 'slow' charges are not simply consequences of *shifting* charges where police delay investigations to wait for DNA evidence.

[30]As mentioned above, results are robust to using a two week threshold instead; see Tables A-14 and A-15.

[31]In the clearance rates for 'all crime' and 'property crime' we exclude minor crimes such as bike theft that are practically never solved and would drive the clearance rate towards zero. Table A-16 compares the main estimates with and without offenses with the lowest clearance rates. None of the results differ qualitatively.

with property crime (-3.2) where the baseline clearance rate is only 30%.

Crucial for the interpretation of these results, both from an academic and policy point of view, is whether our LATE estimates of the effects of DNA registration comprise heterogeneous responses across different treatment margins, which would imply that effects cannot be generalized beyond the common support we obtain from the reform. We test this in Table A-17 following Brinch et al. (2017). The table shows that the null hypothesis of homogenous treatment effects across our area of common support is rejected in 14 out of 18 tests across all crime, crime with fast charges, and crime with slow charges. A subsequent question is then whether our results cannot be generalized because the reform's compliers differ from always takers (i.e. the most hardened criminals who were in the DNA database already) or never takers (the least hardened criminals who were not even added to the database after the reform)? In Figure A-8 we use the decomposition from Black et al. (2015) to compute the difference between y^1 for always takers and compliers (the difference in crime given DNA registration) and differences between y^0 for never takers and compliers (the difference in crime given no DNA registration). The figure shows that compliers' crime only differs substantially from the least hardened criminals' crime. Thus, while the effects of the reform analyzed in this paper span across a wide range of offenders – approximately 35% of everyone charged with a crime – they cannot be generalized to the full population.

When weighing privacy costs of surveillance against public safety benefits, it is important to recognize that the effects for the criminal sample studied here may differ from the effects on other subpopulations. Our results indeed show that recidivism can be reduced effectively. The route towards desistance from crime is, however, not identical for all types of offenders.

7 Discussion

Governments around the world are taking advantage of improvements in technology to change their approaches to criminal justice and to introduce new policies to deter offenders from crime and to aid police in identifying offenders. One popular policy is the introduction and expansion of DNA databases allowing police to identify repeat offenders by matching previously-charged offenders with DNA samples collected at the scene of a crime. So far there has been relatively little analysis of the effects of DNA databases and similar technologies.

In this paper, we estimate the effects of DNA registration on subsequent convictions, using full population register data from Denmark. To obviate the non-random selection into the DNA database, we exploit a 2005 reform in Denmark, which increased the likelihood of being added to the database from approximately 4% to almost 40% for offenders charged with a wide range of crimes, to estimate offenders' responses to DNA registration.

We find that DNA registration has a deterrent effect on future crime. Reductions in the probability of conviction for violent, property and weapons-related crime drive this overall decline in recidivism. Both offenders who enter the DNA database for their first ever charge and individuals who have been charged before are deterred from committing subsequent crime, but when compared to their baseline recidivism rates DNA registration has the largest effect on first-time offenders.

Reducing criminal behavior should have beneficial effects on other aspects of deterred offenders' lives. Turning to non-crime effects of DNA registration, we find that DNA registration increases education for young offenders and employment for older offenders, and the likelihood of being married for first-time offenders. We also see indications that DNA registration leads to more stable relationships and decreases the risk of children of offenders growing up without their father present.

We exploit the nature of DNA databases to separate the deterrence and detection effects of this technology. We illustrate that the estimated effects of crime-prevention policies may be biased upwards if detection effects and clearance rates are not taken into account. We use our estimates of the deterrence and detection effects to provide the first causal estimate of a central theoretical and policy parameter: the elasticity of crime with respect to the probability of detection. Focussing on crime within a three year follow-up period, we estimate this elasticity to be -2.7. This implies that a 1% increase in the likelihood of being apprehended reduces crime by more than 2%, for those with a history of at least one felony charge. Our results thereby show that policies that increase the identification of criminal offenders are an effective tool to reduce crime and increase public safety.

References

Aizer, A. and Doyle, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, 130(2):759–803.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76:169–217.

Black, D., Joo, J., LaLonde, R., Smith, J. A., and Taylor, E. (2015). Simple tests for selection bias: Learning more from instrumental variables. IZA Discussion Paper 9346.

Brinch, C., Mogstad, M., and Wiswall, M. (2017). Beyond late with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Card, D., Chetty, R., and Weber, A. (2007). Cash-on-hand and competing model of intertemporal behavior: New evidence from the labor market. *The Quarterly Journal of Economics*, 122(4):1511–1560.

Chalfin, A. and McCrary, J. (2017a). Are U.S. cities underpoliced? Theory and evidence. *Review of Economics and Statistics*.

Chalfin, A. and McCrary, J. (2017b). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1):5–48.

Det Etiske Råd (2006). Et dna-profil-register, som omfatter alle borgere i danmark?: etiske overvejelser. Kbh.: Det Etiske Råd.

Di Tella, R. and Schargrodsky, E. (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy*, 121(1):28–73.

Doleac, J. L. (2017). The effect of DNA databases on crime. *American Economic Journal: Applied Economics*, 9(1):165–201.

Doleac, J. L. (2019). Encouraging desistance from crime. Working paper.

Durlauf, S. N. and Nagin, D. S. (2011). Imprisonment and crime: Can both be reduced. *Criminology and Public Policy*, 10(1):13–54.

Evans, W. N. and Owens, E. G. (2007). COPS and crime. *Journal of Public Economics*, 91(1-2):181–201.

Grogger, J. (1998). Market wages and youth crime. *Journal of Labor Economics*, 110(1):756–791.

Henneguelle, A., Monnery, B., and Kensey, A. (2016). Better at home than in prison? the effects of electronic monitoring on recidivism in France. *Journal of Law and Economics*, 59(3):629–667.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467.

Justitsministeriet (1999). Forslag til lov om oprettelse af et centralt dna-profilregister, pub. l. no. lf 107. Retrieved from `https://www.retsinformation.dk/Forms/R0710.aspx?id=87870`.

Justitsministeriet (2005). Forslag til lov om ændring af lov om oprettelse af et centralt dna-profilregister og retsplejeloven. (udvidelse af dna-profil-registerets persondel, indikationskrav ved legemsundersøgelse m.v.)., pub. l. no. 2004/2 lf 14. Retrieved from `https://www.retsinformation.dk/Forms/R0710.aspx?id=101176`.

Laub, J. H., Nagin, D. S., and Sampsom, R. (2008). Trajectories of change in criminal offending: Good marriages and the desistance process. *American Sociological Review*, 63(2):225–238.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3):270–290.

Lov om ændring af lov om oprettelse af et centralt dna-profilregister og retsplejelove (2005). Lov om ændring af lov om oprettelse af et centralt dna-profilregister og retsplejeloven, pub. l. no. 369. Retrieved from `https://www.retsinformation.dk/Forms/R0710.aspx?id=2078`.

Lov om oprettelse af et centralt dna-profilregister (2000). Lov om oprettelse af et centralt dna-profilregister, pub. l. no. 434. Retrieved from `https://www.retsinformation.dk/Forms/R0710.aspx?id=836`.

Ludwig, J. and Miller, D. L. (2005). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *NBER working paper (11702)*.

Marie, O. (2015). Early release from prison on electronic monitoring and recidivism: A tale of two discontinuities. Working paper.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. Working paper, retrieved from `http://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2015/09/incar.pdf`.

Pei, Z., Pischke, J.-S., and Schwandt, H. (2017). Poorly measured confounders are more useful on the left than on the right. *NBER working paper (23232)*.

Raphael, S. and Winter-Ebmer, R. (2001). Identifying the effects of unemployment on crime. *Journal of Law and Economics*, 44(2):259–283.

Retsmedicinsk Institut (2014). Dna i straffesager. Retrieved March 17, 2016, from http://retsmedicin.ku.dk/om_instituttet/retsgenetik/Straffesager/til-journalister/.

Sampson, R. J., Laub, J. H., and Wimer, C. (2006). Does marriage reduce crime? a counterfactual approach to within-individual causal effects. *Criminology*, 44:465–506.

Wakefield, S. and Wildeman, C. (2014). *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality*. Oxford University Press.

Tables and Figures

Table 1: Mean characteristics and subgroup sizes by timing of initial charge

	Pre reform		**Post reform**		**All**	
	Mean	SD	Mean	SD	Mean	SD
In DNA database	0.043	0.202	0.488	0.500	0.256	0.436
Covariates						
Age	22.276	3.644	22.017	3.578	22.152	3.615
Immigrant background	0.211	0.408	0.216	0.412	0.214	0.410
Has children	0.129	0.335	0.113	0.317	0.121	0.326
Single	0.853	0.354	0.862	0.345	0.858	0.349
Lives in 1 of 4 biggest citites	0.368	0.482	0.380	0.485	0.374	0.484
Years of education	10.914	1.910	10.818	1.887	10.868	1.900
Gross income (10.000s)	11.671	9.477	11.655	12.181	11.663	10.858
In employment	0.523	0.499	0.581	0.493	0.551	0.497
# prior charges	3.122	2.997	3.143	2.981	3.132	2.989
Crime type						
Property	0.595	0.491	0.521	0.500	0.560	0.496
Violence	0.247	0.431	0.296	0.457	0.271	0.444
Sexual	0.023	0.148	0.025	0.155	0.024	0.152
Drugs (penal)	0.021	0.144	0.024	0.153	0.023	0.148
Other penal	0.058	0.233	0.069	0.254	0.063	0.243
Weapon	0.056	0.230	0.065	0.246	0.060	0.238
Observations	34829		32082		66911	
Subgroups	Share	N	Share	N	Share	N
Previous charges						
First-time offenders	0.244	8508	0.241	7718	0.243	16226
Recidivists	0.756	26321	0.759	24364	0.757	50685
Age group						
18-23	0.662	23053	0.693	22244	0.677	45297
24-30	0.338	11776	0.307	9838	0.323	21614

Note: The table shows means and standard deviations for all covariates for the full sample and for those charged before and after the reform separately. The table also shows the number and proportion of the sample belonging to specific subgroups used in the analysis. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table 2: Distribution of characteristics in the complier group

	Overall mean	Complier mean	Sig.
Covariates			
Aged 18-23	0.677	0.730	***
Aged 24-30	0.323	0.270	***
Imm. background	0.213	0.215	
Has children	0.121	0.114	***
Single	0.858	0.869	***
Lives in 1 of 4 biggest citites	0.374	0.389	***
Max. 10 years of educ.	0.474	0.495	***
Gross income above sample median	0.500	0.464	***
In employment	0.551	0.551	
First charge	0.243	0.198	***
Crime type			
Property	0.560	0.511	***
Violence	0.271	0.373	***
Sexual	0.023	0.036	***
Other penal	0.085	0.067	***
Weapon	0.060	0.016	***

$^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Note: The table shows the distribution of background characteristics in the complier group (column 2) following Angrist and Pischke (2009) and the overall sample (column 1). The final column indicates whether complier means are statistically significantly different from the overall sample mean (standard errors are calcuated on the basis of 100 bootstrapped samples).

Source: Own calculation based on Data from Statistics Denmark and the National Police.

Table 3: Unconditional balancing tests for each covariate

	(1)	(2)
Age	-0.052	-0.058
	(0.065)	(0.067)
Imm. background	0.008	0.012
	(0.008)	(0.008)
Single	-0.010	-0.013*
	(0.006)	(0.006)
Has children	-0.006	-0.007
	(0.006)	(0.006)
Lives in 1 of 4 biggest citites	0.005	0.011
	(0.009)	(0.009)
Years of education	-0.090**	-0.071*
	(0.034)	(0.035)
Gross income (10.000s)	-0.213	-0.058
	(0.180)	(0.187)
In employment	0.015	0.013
	(0.009)	(0.010)
Unemployed	-0.014	-0.009
	(0.009)	(0.009)
Enrolled in education	-0.001	-0.004
	(0.006)	(0.006)
# charges prior to the one in question	0.014	0.003
	(0.056)	(0.058)
Type of crime leading to initial charge:		
Violence	0.023**	0.023**
	(0.007)	(0.007)
Property	-0.036***	-0.039***
	(0.008)	(0.009)
Sexual	-0.001	-0.000
	(0.002)	(0.002)
Weapon	0.003	0.002
	(0.004)	(0.004)
Other penal	0.012**	0.014**
	(0.005)	(0.005)
Observations	66911	66911
Running variables	X	X
Month FE		X

Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Note: Table shows estimates from regressing each covariate on a dummy indicating whether charges occurred after the reform and running variables and month FE. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table 4: Test for discontinuities in predicted subsequent convictions

	P(convicted)			**# convictions**		
Years	All	Fast	Slow	All	Fast	Slow
1 year	-0.001	-0.000	-0.000	-0.001	-0.001	-0.000
	(0.002)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)
2 years	-0.000	0.000	-0.000	-0.001	-0.001	-0.001
	(0.003)	(0.002)	(0.001)	(0.005)	(0.004)	(0.001)
3 years	0.000	0.000	-0.001	-0.001	-0.000	-0.001
	(0.003)	(0.003)	(0.001)	(0.007)	(0.006)	(0.002)
Observations	66911	66911	66911	66911	66911	66911

Standard errors in parentheses. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows results from first regressing subsequent convictions on covariates measured before the initial charge (these covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month FE), and then regressing the predicted outcomes on the after-reform dummy and running variables. This is done to examine whether differences in covariates before and after the reform predict discontinuities in outcomes around the reform. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table 5: First stage estimation results

	DNA registration		
	(1)	(2)	(3)
Charged after reform	0.350***	0.347***	0.347***
	(0.007)	(0.007)	(0.007)
Observations	66911	66911	66911
Running variables	X	X	X
Covariates		X	X
Month FE			X

Standard errors in parentheses. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows estimates from first-stage regressions regressing DNA registration on timing of charge (before/after reform). Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table 6: Effects of DNA profiling on subsequent convictions (accumulated) by different conditioning sets

	P(convicted)			**# convictions**		
	(1)	(2)	(3)	(4)	(5)	(6)
A: All convictions						
1 year	-0.065**	-0.065***	-0.065***	-0.096**	-0.095***	-0.093**
	(0.020)	(0.019)	(0.019)	(0.029)	(0.028)	(0.029)
2 years	-0.074**	-0.075**	-0.075**	-0.174***	-0.173***	-0.163***
	(0.025)	(0.023)	(0.024)	(0.050)	(0.047)	(0.048)
3 years	-0.048+	-0.049*	-0.047+	-0.140*	-0.140*	-0.129*
	(0.026)	(0.024)	(0.025)	(0.065)	(0.060)	(0.061)
Pre-reform baseline						
1 year		0.153			0.189	
2 years		0.298			0.449	
3 years		0.375			0.652	
Placebo test						
Previous charges	-0.004	-0.004	0.001	0.039	0.059	0.048
	(0.019)	(0.019)	(0.019)	(0.160)	(0.155)	(0.159)
Observations	66911	66911	66911	66911	66911	66911
Running variables	X	X	X	X	X	X
Covariates		X	X		X	X
Month FE			X			X

Standard errors in parentheses + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$ Note: Table shows 2SLS estimates of regressing subsequent crime on DNA profiling (instrumented by timing of initial charge - before/after reform) by different conditioning sets. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table 7: Effects of DNA profiling on labor market outcomes, overall and by age group

# years after the first four years after charge	**All Offenders**	**Aged 18-23**	**Aged 24-30**
In employment	-0.011	-0.103	0.365*
	(0.076)	(0.088)	(0.161)
In education/training	0.098***	0.129***	0.010
	(0.027)	(0.034)	(0.036)
Unemployed	-0.087	-0.026	-0.375*
	(0.076)	(0.088)	(0.162)
Observations	66911	45297	21614
Pre-reform mean			
In employment	1.954	1.954	1.955
	(1.612)	(1.589)	(..657)
In education/training	0.120	0.159	0.044
	(0.453)	(0.516)	(0.280)
In unemployment	1.926	1.887	2.000
	(1.616)	(1.593)	(1.658)
Placebo test:			
	In employment	In education/training	Unemployed
Year -1	0.037	-0.014	-0.022
	(0.024)	(0.016)	(0.024)

Standard errors in parentheses. + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows IV estimates of regressing labor market outcomes on DNA profiling (instrumented by timing of initial charge - before/after reform). Covariates include age, immigrant background, has children, single, years of education, gross income, employment status (measured before charge), number of prior charges, crime type dummies, and month fixed effects. Panel labelled "Placebo test" shows estimates using labor market outcomes measured the year prior to the charge. Standard errors are clustered by personal identification number. Source: Own calculation based on Data from Statistics Denmark and the National Police.

Table 8: Effects of DNA profiling on family outcomes, overall and by previous charges

	All Offenders			**First charge**			**Recidivist**		
Year	Married (1)	Same partner (2)	Living with child and mother (3)	Married (4)	Same partner (5)	Living with child and mother (6)	Married (7)	Same partner (8)	Living with child and mother (9)
1 year	0.007	0.110	0.121+	0.030*	0.004	-0.050	0.002	0.131+	0.153*
	(0.006)	(0.069)	(0.068)	(0.013)	(0.163)	(0.167)	(0.007)	(0.074)	(0.072)
2 years	0.002	0.070	0.051	0.048*	0.138	-0.002	-0.009	0.051	0.063
	(0.008)	(0.067)	(0.070)	(0.019)	(0.166)	(0.172)	(0.009)	(0.071)	(0.074)
3 years	0.011	0.067	0.036	0.068*	0.155	-0.080	-0.002	0.044	0.060
	(0.011)	(0.068)	(0.071)	(0.024)	(0.166)	(0.176)	(0.011)	(0.071)	(0.074)
Placebo test	0.001			-0.006			0.002		
	(0.005)			(0.011)			(0.005)		
Observations	66911	9527	11767	16226	2532	2148	50685	6995	9619
Pre-reform baseline									
Year 1	0.058	0.467	0.307	0.069	0.551	0.484	0.054	0.436	0.268

Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows IV estimates of regressing family outcomes on DNA profiling (instrumented by timing of initial charge - before/after reform). Columns 1, 4, and 7 show results for the likelihood of being married. Columns 2, 5, and 8 show results for the likelihood of living with the same partner for those who had a partner prior to charge. Columns 3, 6, and 9 show results for the likelihood of father living with child's mother for all children born prior to father's charge (some of these children will not have been born by year -1). Panel labelled "Placebo test" shows estimates on marital status measured the year before the charge. Covariates include child's age and gender, and father's age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies, and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculation based on Data from Statistics Denmark and the National Police.

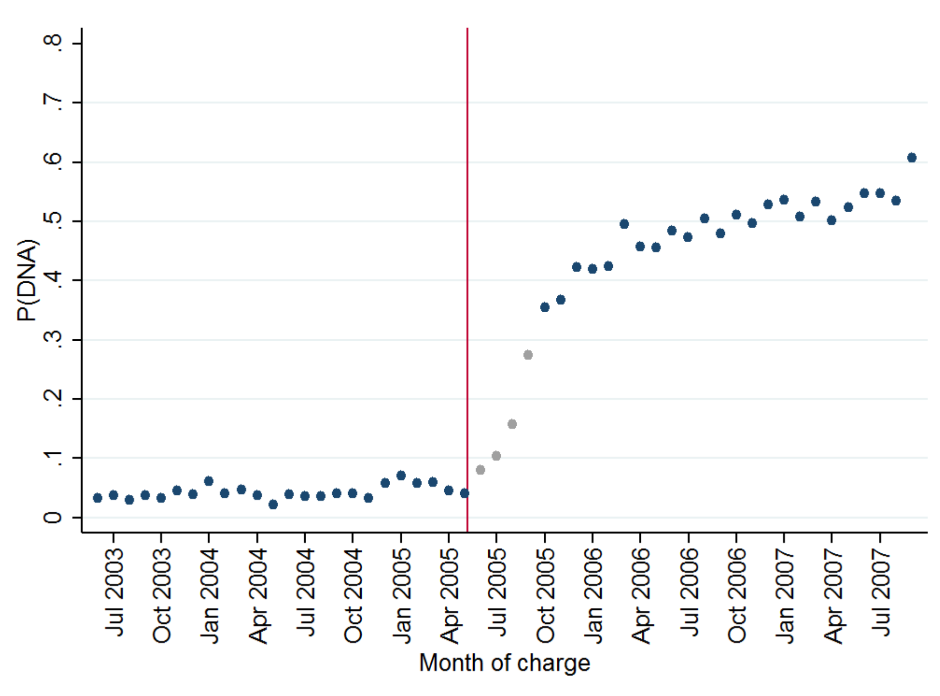
Table 9: Deterrence and detection effects on subsequent new crime

	Deterrence effect Δ		**Detection effect δ**		Clearance rate p	Clearance rate	Elasticity of #new crimes
	P(new crime)	#new crimes	P(new crime)	#new crimes		w. 3 weeks $p\pi$	with respect to p
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A:Any Crime							
3 years	-0.282 ***	-0.523 ***	0.036 *	0.077 **	0.399	0.249	-2.7
B: Property							
3 years	-0.258 ***	-0.305 *	0.024	0.029	0.305	0.174	-3.2
C: Violence							
3 years	-0.049 *	-0.056 +	0.017 *	0.017 +	0.820	0.616	-2.7
Pre-reform baseline / clearance rate ($\bar{p}$), 3 year							
	P(new crime)	# new crimes					
Any Crime	0.939	1.633					
Property	0.780	1.228					
Violence	0.172	0.216					

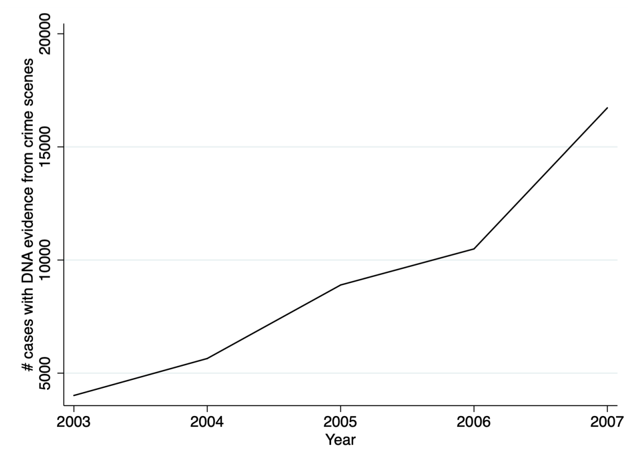
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Note: Table shows estimates of deterrence and detection effects calculated on the basis of IV-estimation (including covariates and month FE) from 100 bootstrapped samples. Clearing rates were calculated on the basis of all charges and all reported crime in 2005. In these measures we excluded crime types such as bicycle theft which is heavily reported (often for insurance purposes) but rarely solved and leading to a charge (<10% of the time) in order not to inflate estimates by an extremely low clearance rate. The fraction of crimes solved within 3 weeks is 0.623 overall, 0.569 for property crimes, and 0.752 for violent crimes. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Figure 1: DNA registration of offenders and DNA samples from crime scenes

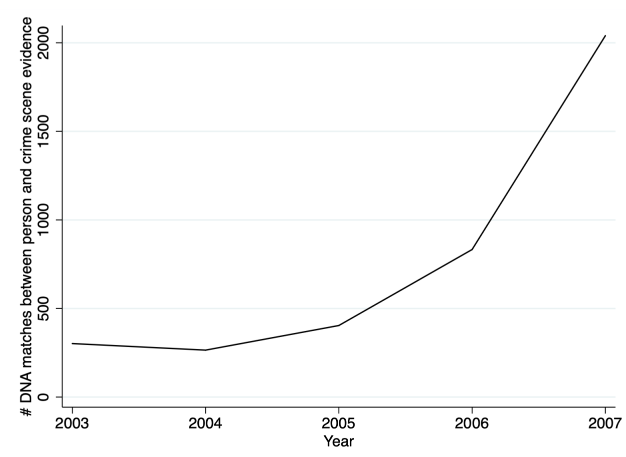
(a) Likelihood of DNA registration



(b) DNA samples collected from crime scenes

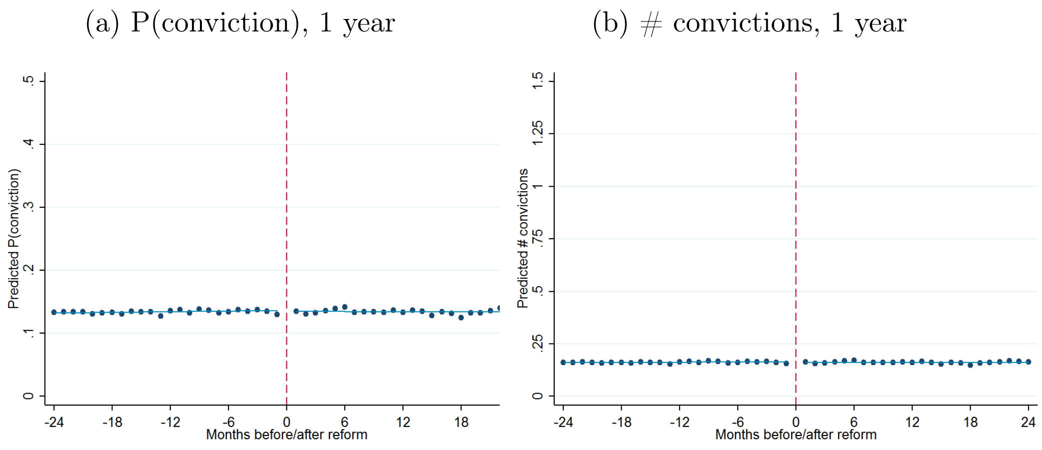


(c) Matches between offender profiles and crime scene evidence in the DNA database



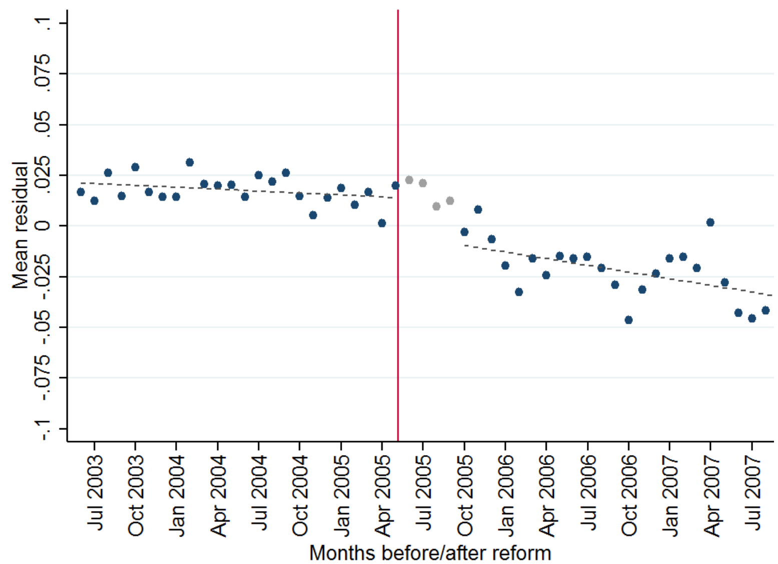
Note: Figure A shows the fraction of offenders in the sample who are registered in the DNA database by month of charge. The vertical line marks the timing of the reform. Figure B shows the number of cases with DNA evidence from crime scenes in the DNA database. Figure C shows the number of matches (hits) between offender DNA profiles and crime scene evidence. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Figure 2: Predicted probability of conviction and number of convictions from observable characteristics around the timing of the reform

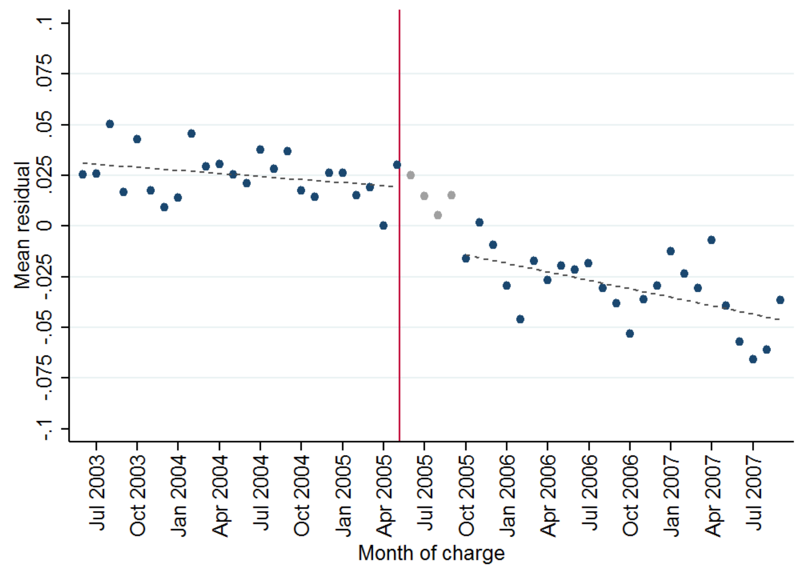


Note: Figures show predicted probability of any conviction and number of convictions for crimes that occurred within a year after a given crime charge, predicted from estimation results regressing outcomes on covariates, crime types and month FE. Figure A shows predictions for the binary outcomes and Figure B shows predictions for the number of subsequent convictions. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Figure 3: Monthly means of crime outcomes around the timing of the reform



(a) P(conviction), 1 year



(b) # convictions, 1 year

Note: Figures show monthly means of crime outcomes within one year. Figure A shows the probability of receiving at least one conviction and Figure B shows monthly means number of convictions within one year. We condition on covariates in all figures. Therefore the figures show deviations around the conditional sample mean and not absolute levels. Source: Own calculations based on Data from Statistics Denmark and the National Police.

A Appendix Tables and Figures

Table A-1: Crime categories

Criminal Code	Main categories of crime	Subcategories of crime	Our category
Penal Code	All sexual offenses	Incest	Sexual
		Rape	Sexual
		Pedophilia	Sexual
		Voyerism, flashing	Sexual
		Other sexual violations	Sexual
	Violent crime	Violence against public servant	Violence
		Disturbance of public peace	Violence
		Murder, manslaughter (+ attempted)	Violence
		Simple violence	Violence
		Major violence	Violence
		Threats	Violence
		Other violent assaults	Violence
	Property crimes	Fraud	Property
		Arson	Property
		Theft	Property
		Burglary	Property
		Robbery	Property
		Vandalism	Property
		Other property crime	Property
	Other crimes against penal code	Crimes against/as a public servant	Other (penal)
		Drug smuggling or sales	Other (penal)
		Obstruction of justice	Other (penal)
		Restrain orders	Other (penal)
		Other crimes, penal code	Other (penal)
Special Acts	Violations of Traffic Act	Accidents and speeding	-
		Traffic accidents w. alcohol	-
		Drunk driving	-
		Other traffic offenses	-
	Violations of Drug Act	Possession and or drugs sales -	
	Violations of Weapons/Arms Act	Explosives, firearms, knives	Weapon
	Smuggling, construction, health, social fraud, other		-

Table A-2: Effects of DNA profiling on subsequent convictions (accumulated) by different caps on prior charges

	P(convicted)			**# convictions**		
	Max. 5	Max. 10	Max. 15	Max. 5	Max. 10	Max. 15
1 year	-0.053**	-0.065***	-0.078***	-0.055*	-0.093**	-0.116***
	(0.019)	(0.019)	(0.020)	(0.026)	(0.029)	(0.031)
2 years	-0.058*	-0.075**	-0.080***	-0.098*	-0.163***	-0.190***
	(0.025)	(0.024)	(0.023)	(0.044)	(0.048)	(0.050)
3 years	-0.024	-0.047+	-0.053*	-0.061	-0.129*	-0.153*
	(0.027)	(0.025)	(0.024)	(0.057)	(0.061)	(0.064)
Observations	51550	66911	76531	51550	66911	76531

Standard errors in parentheses + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$ Note: Table shows 2SLS estimates of regressing subsequent crime on DNA registration (instrumented by timing of initial charge - before/after reform) by different caps on prior charges. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-3: Mean of crime and family outcomes, by timing of charge relative to the reform

	Pre reform	Post reform	Pre reform	Post reform	Pre reform	Post reform
A) Crime outcomes:	*P(conviction)*		*# convictions*			
Any Crime						
1 year	0.153	0.114	0.189	0.133		
2 years	0.298	0.246	0.449	0.341		
3 years	0.375	0.338	0.652	0.553		
Property						
1 year	0.091	0.058	0.111	0.067		
2 years	0.186	0.136	0.263	0.176		
3 years	0.238	0.198	0.375	0.292		
Violence						
1 year	0.049	0.044	0.053	0.047		
2 years	0.103	0.096	0.120	0.112		
3 years	0.141	0.138	0.177	0.170		
Sexual						
1 year	0.002	0.001	0.002	0.001		
2 years	0.003	0.002	0.003	0.002		
3 years	0.005	0.003	0.005	0.004		
Other penal						
1 year	0.012	0.008	0.012	0.009		
2 years	0.034	0.026	0.036	0.027		
3 years	0.051	0.047	0.055	0.050		
Weapon						
1 year	0.011	0.009	0.011	0.009		
2 years	0.026	0.023	0.027	0.024		
3 years	0.038	0.035	0.041	0.037		
Observations	34829	32082	34829	32082		
B) Labor Market outcomes:	*Employment*		*Education/training*		*Unemployment*	
Cumulated time year 1-4	1.954	1.878	0.120	0.212	1.926	1.910
Observations	34829	32082	34829	32082	34829	32082
C) Family outcomes:	*Married*		*Same partner*		*Living with child and mother*	
1 year	0.058	0.042	0.467	0.444	0.307	0.290
2 years	0.064	0.050	0.418	0.390	0.288	0.268
3 years	0.075	0.064	0.386	0.347	0.280	0.252
Observations	34829	32082	5106	4421	6614	5153

Note: The table shows means of crime, labor market and family outcomes for those charged before and after the reform separately. Source: Own calculations based on Data from Statistics Denmark.

Table A-4: Charges and convictions for crimes committed before DNA profiling

	P(charged)	# charges	P(convicted)	# convictions
3 years	-0.006	0.033	-0.009	-0.011
	(0.020)	(0.061)	(0.013)	(0.015)
Observations	66911	66911	66911	66911

Note: The table shows estimated changes in the probability of being charged, number of charges, probability of being convicted, and number of convictions for crimes committed *before* DNA profiling but where charges were not pressed until *after* the DNA profiling. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-5: Effects of DNA registration on subsequent accumulated probability of conviction and number of convictions, by crime type

	By initial crime type		**By subsequent crime type**	
	P(conviction)	# convictions	P(conviction)	# convictions
A: Property				
1 year	-0.053^{+}	-0.089*	-0.031^{+}	-0.051*
	(0.029)	(0.044)	(0.016)	(0.023)
2 years	-0.073*	-0.170*	-0.049*	-0.087*
	(0.035)	(0.074)	(0.021)	(0.038)
3 years	-0.036	-0.097	-0.037^{+}	-0.062
	(0.036)	(0.095)	(0.022)	(0.049)
B: Violence				
1 year	-0.067**	-0.077**	-0.031*	-0.035*
	(0.021)	(0.028)	(0.012)	(0.014)
2 years	-0.085**	-0.129**	-0.034*	-0.041^{+}
	(0.027)	(0.045)	(0.017)	(0.022)
3 years	-0.066*	-0.129*	-0.026	-0.027
	(0.029)	(0.059)	(0.020)	(0.028)
C: Sexual				
1 year	0.019	0.015	-0.000	-0.000
	(0.040)	(0.045)	(0.002)	(0.002)
2 years	0.063	0.023	0.001	0.001
	(0.064)	(0.089)	(0.003)	(0.003)
3 years	0.071	0.025	0.004	0.002
	(0.073)	(0.061)	(0.004)	(0.004)
D: Other penal				
1 year	-0.119*	-0.173*	0.001	0.004
	(0.060)	(0.081)	(0.006)	(0.006)
2 years	-0.035	-0.219^{+}	-0.013	-0.014
	(0.082)	(0.131)	(0.010)	(0.011)
3 years	-0.095	-0.362*	0.004	-0.008
	(0.087)	(0.163)	(0.013)	(0.014)
E: Weapon				
1 year	-0.319	-0.425	-0.010^{+}	-0.011^{+}
	(0.230)	(0.301)	(0.006)	(0.006)
2 years	-0.398	-0.862^{+}	-0.021*	-0.021*
	(0.286)	(0.502)	(0.009)	(0.009)
3 years	-0.102	-0.593	-0.031**	-0.034**
	(0.290)	(0.616)	(0.011)	(0.012)
Pre-reform baseline, 1 year				
Property	0.168	0.209	0.091	0.111
Violence	0.140	0.165	0.049	0.053
Sexual	0.042	0.043	0.002	0.002
Other penal	0.113	0.143	0.012	0.012
Weapon	0.159	0.193	0.011	0.011

Standard errors in parentheses. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows estimates of the effect of DNA registration by type of initial crime in the left half and the type of subsequent crime in the right half. Total number of observations: 66,991. Observations by initial crime type: property 37,443; violence 18,116; sexual 1,576; other penal 5,735; weapon 4,041. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-6: Effects of DNA profiling, heterogeneity by offender characteristics

	Panel A		Panel B		Panel C	
	First sample	Recidivist charge	Aged 18-23	Aged 24-30	Child	No child
P(convicted)						
1 year	-0.048^{+}	-0.068**	-0.090***	0.002	-0.048	-0.067**
	(0.025)	(0.023)	(0.023)	(0.033)	(0.050)	(0.021)
2 years	-0.081*	-0.071*	-0.085**	-0.050	-0.141*	-0.067**
	(0.035)	(0.028)	(0.028)	(0.043)	(0.066)	(0.025)
3 years	-0.030	-0.049^{+}	-0.050^{+}	-0.044	-0.115^{+}	-0.039
	(0.040)	(0.029)	(0.029)	(0.046)	(0.069)	(0.026)
Pre-reform baseline (1 year)	0.061	0.183	0.177	0.107	0.124	0.157
# convictions						
1 year	-0.037	-0.105**	-0.130***	0.004	-0.063	-0.097**
	(0.028)	(0.035)	(0.035)	(0.048)	(0.072)	(0.031)
2 years	-0.065	-0.182**	-0.189**	-0.093	-0.193^{+}	-0.159**
	(0.048)	(0.058)	(0.058)	(0.080)	(0.116)	(0.051)
3 years	-0.029	-0.147^{+}	-0.143^{+}	-0.090	-0.182	-0.122^{+}
	(0.063)	(0.075)	(0.075)	(0.099)	(0.146)	(0.066)
Pre-reform baseline (1 year)	0.068	0.228	0.219	0.129	0.154	0.194
Observations	16226	50685	45297	21614	8113	58798

Standard errors in parentheses. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows estimates of the effect of DNA profiling on subsequent crime. Separate estimates for subgroups are obtained by interacting the reform dummy with subgroup dummies. Subgroups in Panel A are first time offenders (sampling charge is their first charge) and redivists (has between1-10 prior charges). Subgroups in Panel B are offenders aged 18-23 and 24-30 at the time of the sampling charge. Subgroups in Panel C are those who have at least one child at the time of sampling and those that have none. Depending on the panel covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal id number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-7: Difference in difference estimates of the reform expanding DNA profiling on subsequent accumulated probability of conviction and number of convictions

	P(convicted)	**# convictions**
1 year	-0.018*	-0.019^{+}
	(0.007)	(0.010)
2 years	-0.022*	-0.028^{+}
	(0.009)	(0.017)
3 years	-0.024*	-0.038^{+}
	(0.010)	(0.022)
Observations	50267	50267

Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows Difference in Difference estimates of the reform. We estimate this as:
$y_{it} = \alpha + \gamma_1 \mathbf{1}[post] + \gamma_2 \mathbf{1}[Treatment_i] + \gamma_3 \mathbf{1}[post_i] * \mathbf{1}[Treatment_i] + \epsilon_{it}$
where γ_3 is the *DiD* estimate presented in the table.

Table A-8: Reduced form estimates predicting probability of convictions and the number of convictions from timing of initial charge in placebo samples

	P(convicted)	**# convictions**
2002, placebo reform	0.002	0.005
	(0.007)	(0.011)
2003, placebo reform	0.007	0.004
	(0.007)	(0.011)
2004, placebo reform	0.004	0.006
	(0.007)	(0.011)
2005, actual reform	-0.022***	-0.032**
	(0.007)	(0.010)
2006, placebo reform	-0.010	-0.007
	(0.006)	(0.009)

Standard errors in parentheses + p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Note: Table shows reduced form estimates from regressing subsequent convictions on a "after-reform"-dummy (along with running variables, covariates and month FE) in a series of placebo samples. The placebo samples mirrors the original sample except that the reform is artificially set to occur in e.g. 2002 instead of 2005, and as in the original samples the sampling window is defined as +/-24 months around the reform (except from June-September in the reform year). Standard errors are clustered on personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police

Table A-9: Effects of DNA profiling, including summer months

	Full sample	First charge	Recidivist	Aged 18-23	Aged 24-30	Child	No child
P(convicted), all crime							
1 year	-0.037	-0.041	-0.040	-0.077*	0.071	0.068	-0.050
	(0.031)	(0.042)	(0.038)	(0.038)	(0.055)	(0.089)	(0.033)
2 years	-0.063+	-0.099+	-0.058	-0.095*	0.021	-0.063	-0.062
	(0.038)	(0.059)	(0.046)	(0.045)	(0.072)	(0.114)	(0.040)
3 years	-0.042	-0.032	-0.049	-0.057	-0.008	-0.001	-0.046
	(0.039)	(0.065)	(0.047)	(0.046)	(0.077)	(0.119)	(0.041)
# convictions, all crime							
1 year	-0.075	-0.020	-0.094	-0.136*	0.093	0.126	-0.099*
	(0.047)	(0.048)	(0.058)	(0.057)	(0.083)	(0.148)	(0.050)
2 years	-0.147+	-0.042	-0.183+	-0.197*	-0.017	0.006	-0.164*
	(0.076)	(0.082)	(0.094)	(0.093)	(0.128)	(0.200)	(0.082)
3 years	-0.123	-0.019	-0.162	-0.137	0.104	0.069	-0.144
	(0.097)	(0.107)	(0.121)	(0.121)	(0.159)	(0.251)	(0.105)
First stage on probability of DNA profiling:							
Charged post reform	0.212***						
	(0.006)						

Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows estimates of the effect of DNA profiling on subsequent crime including the months that are excluded in the main analysis. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Observations: 72,338. Standard errors are clustered by personal id number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-10: Effects of DNA profiling on subsequent convictions (accumulated) by different bandwidth specifications

	P(convicted)				**# convictions**			
	BW:12	BW:18	BW:24	BW:30	BW:12	BW:18	BW:24	BW:30
A: All convictions								
1 year	-0.054^{+}	-0.064^{**}	-0.065^{***}	-0.074^{***}	-0.087^{*}	-0.106^{***}	-0.093^{**}	-0.108^{***}
	(0.030)	(0.022)	(0.019)	(0.017)	(0.042)	(0.032)	(0.029)	(0.025)
2 years	-0.041	-0.063^{*}	-0.075^{**}	-0.099^{***}	-0.107	-0.164^{**}	-0.163^{***}	-0.212^{***}
	(0.037)	(0.027)	(0.024)	(0.021)	(0.068)	(0.052)	(0.048)	(0.044)
3 years	-0.040	-0.061^{*}	-0.047^{+}	-0.070^{**}	-0.118	-0.175^{**}	-0.129^{*}	-0.183^{**}
	(0.038)	(0.028)	(0.025)	(0.022)	(0.088)	(0.067)	(0.061)	(0.057)
Observations	33509	50245	66911	83147	33509	50245	66911	83147

Standard errors in parentheses + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$ Note: Table shows 2SLS estimates of regressing subsequent crime on DNA profiling (instrumented by timing of initial charge - before/after reform) by different bandwidth specifications. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-11: Effects of DNA profiling on subsequent accumulated number of convictions using different running variable specifications

Years	(1)	(2)
1 year	-0.093**	-0.094*
	(0.029)	(0.041)
2 years	-0.163***	-0.125+
	(0.048)	(0.067)
3 years	-0.129*	-0.168+
	(0.061)	(0.087)
Observations	66911	66911
Running variables:		
Linear	X	X
Quadratic		X

Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows 2SLS estimates of regressing subsequent crime on DNA registration (instrumented by timing of initial charge - before/after reform) with the baseline specification of the running variable (linear, but flexible on each side or the reform) from Table 7, and a more flexible quadratic running variable (also flexible on each side of the reform). Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-12: Effects of DNA profiling on subsequent convictions - adjusted for time spent incarcerated

	# convictions		
	Adj. no cap	Adj. cap=0.5	Adj. cap=0.75
1 year	-0.098**	-0.104**	-0.101**
	(0.032)	(0.032)	(0.032)
2 years	-0.185***	-0.187***	-0.186***
	(0.053)	(0.053)	(0.053)
3 years	-0.155*	-0.156*	-0.156*
	(0.068)	(0.068)	(0.068)
Observations	66908	66911	66911

Standard errors in parentheses + p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Note: Table shows IV estimates of regressing subsequent convictions on DNA profiling (instrumented by timing of initial charge - before/after reform). Number of subsequent convictions have been divided by the proportion of the follow up period not spent incarcerated with different caps on the maximum proportion of time spent incarcerated. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies, and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculation based on Data from Statistics Denmark.

Table A-13: Effects of DNA registration on subsequent convictions by time it takes to solve crime

	Panel A	**Panel B**		**Panel C**		**Panel D**		**Panel E**		**Panel F**	
	Full sample	Criminal trajectory		Age at charge		Parenthood		Initially charged for		Subsequently convicted of	
		First charge	Recidivist	18-23	24-30	Child	No child	prop. crime	viol. crime	prop. crime	viol. crime
P(convicted), solved fast											
1 year	-0.057**	-0.044$^+$	-0.060**	-0.079***	0.000	-0.058	-0.057**	-0.051$^+$	-0.058**	-0.033*	-0.025*
	(0.018)	(0.022)	(0.022)	(0.022)	(0.030)	(0.047)	(0.020)	(0.027)	(0.020)	(0.015)	(0.012)
2 years	-0.084***	-0.081*	-0.083**	-0.094***	-0.060	-0.177**	-0.073**	-0.084*	-0.082**	-0.047*	-0.030$^+$
	(0.023)	(0.032)	(0.027)	(0.027)	(0.040)	(0.062)	(0.024)	(0.034)	(0.026)	(0.020)	(0.016)
3 years	-0.067**	-0.055	-0.068*	-0.067*	-0.069	-0.188**	-0.053*	-0.060$^+$	-0.074**	-0.044*	-0.029
	(0.024)	(0.036)	(0.028)	(0.028)	(0.043)	(0.066)	(0.025)	(0.036)	(0.028)	(0.021)	(0.018)
P(convicted), solved slow											
1 year	-0.016	-0.004	-0.018	-0.026*	0.014	0.022	-0.020$^+$	-0.013	-0.012	-0.006	-0.006
	(0.010)	(0.011)	(0.012)	(0.012)	(0.019)	(0.027)	(0.011)	(0.016)	(0.010)	(0.009)	(0.005)
2 years	-0.030$^+$	-0.007	-0.035$^+$	-0.039*	-0.008	0.009	-0.035*	-0.038	-0.025	-0.018	-0.009
	(0.016)	(0.019)	(0.019)	(0.020)	(0.029)	(0.045)	(0.017)	(0.025)	(0.016)	(0.014)	(0.008)
3 years	-0.006	0.016	-0.010	-0.008	-0.000	0.050	-0.013	0.003	-0.011	-0.010	-0.006
	(0.019)	(0.023)	(0.023)	(0.023)	(0.033)	(0.053)	(0.020)	(0.029)	(0.020)	(0.016)	(0.011)
# convictions, solved fast											
1 year	-0.076**	-0.034	-0.085**	-0.101**	-0.009	-0.080	-0.075**	-0.073$^+$	-0.065**	-0.044*	-0.029*
	(0.026)	(0.025)	(0.032)	(0.032)	(0.041)	(0.063)	(0.028)	(0.040)	(0.025)	(0.021)	(0.013)
2 years	-0.124**	-0.061	-0.137**	-0.141**	-0.078	-0.204*	-0.114**	-0.122$^+$	-0.101**	-0.060$^+$	-0.032
	(0.041)	(0.042)	(0.050)	(0.050)	(0.067)	(0.099)	(0.044)	(0.063)	(0.039)	(0.033)	(0.020)
3 years	-0.127*	-0.046	-0.141*	-0.140*	-0.089	-0.216$^+$	-0.115*	-0.112	-0.122*	-0.052	-0.033
	(0.051)	(0.054)	(0.062)	(0.062)	(0.084)	(0.124)	(0.055)	(0.079)	(0.049)	(0.040)	(0.024)
# convictions, solved slow											
1 year	-0.017$^+$	-0.002	-0.021$^+$	-0.028*	0.013	0.017	-0.022$^+$	-0.016	-0.012	-0.007	-0.006
	(0.010)	(0.011)	(0.013)	(0.012)	(0.020)	(0.028)	(0.011)	(0.016)	(0.010)	(0.009)	(0.005)
2 years	-0.039*	-0.005	-0.046*	-0.048*	-0.015	0.010	-0.045*	-0.049$^+$	-0.028	-0.028$^+$	-0.009
	(0.019)	(0.020)	(0.023)	(0.024)	(0.031)	(0.049)	(0.021)	(0.030)	(0.019)	(0.016)	(0.008)
3 years	-0.002	0.017	0.005	-0.003	-0.001	0.033	-0.007	0.016	-0.007	-0.011	0.006
	(0.026)	(0.027)	(0.031)	(0.032)	(0.039)	(0.065)	(0.028)	(0.040)	(0.025)	(0.021)	(0.011)

Standard errors in parentheses. $^+$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Note: Table shows 2SLS estimates of regressing subsequent crime on DNA registration by time from date of crime to charge (solved fast < 3 weeks, solved slowly $\geq$ three weeks). The results are presented for the full sample (Panel A), by offender characteristics (Panels B-E), and by crime type (Panel F). Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-14: Effects of DNA profiling on subsequent accumulated probability of conviction and number of convictions, 2 week and 3 week cut-offs

	P(convicted)				**# convictions**			
	Fast charge		Slow charge		Fast charge		Slow charge	
	2w	3w	2w	3w	2w	3w	2w	3w
Years								
Main results								
1 year	-0.055**	-0.057**	-0.017	-0.016	-0.074**	-0.076**	-0.020	-0.017^{+}
	(0.018)	(0.018)	(0.11)	(0.010)	(0.025)	(0.026)	(0.012)	(0.010)
2 years	-0.077***	-0.084***	-0.040*	-0.030^{+}	-0.111**	-0.124**	-0.051*	-0.039*
	(0.023)	(0.023)	(0.018)	(0.016)	(0.039)	(0.041)	(0.022)	(0.019)
3 years	-0.065**	-0.067**	-0.016	-0.006	-0.113*	-0.127*	-0.016	-0.002
	(0.024)	(0.024)	(0.020)	(0.019)	(0.049)	(0.051)	(0.029)	(0.026)
Observations	66911	66911	66911	66911	66911	66911	66911	66911
Pre-reform baseline, 1 year								
Main outcomes	0.125	0.132	0.037	0.029	0.149	0.158	0.040	0.031

Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows 2SLS estimates of regressing subsequent crime on DNA registration (instrumented by timing of initial charge - before/after reform). Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-15: Deterrence and detection effects on subsequent new crime. 2 week cut-off

	Deterrence effect Δ		**Detection effect δ**		Clearance rate p	Clearance rate w. 3 weeks $p\pi$	Elasticity of #new crimes with respect to p
	P(new crime)	#new crimes	P(new crime)	#new crimes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A:Any Crime							
3 years	-0.301 ***	-0.507 ***	0.033 +	0.071 **	0.399	0.228	-2.9
B: Property							
3 years	-0.266 **	-0.264 +	0.023	0.017	0.305	0.159	-4.8
C: Violence							
3 years	-0.055 +	-0.061 *	0.018 +	0.021 +	0.820	0.549	-2.4
Pre-reform baseline / clearance rate ($\bar{p}$), 3 year							
	P(new crime)	# new crimes					
Any Crime	0.939	1.633					
Property	0.780	1.228					
Violence	0.172	0.216					

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Note: Table shows estimates of deterrence and detection effects calculated on the basis of IV-estimation (including covariates and month FE) from 100 bootstrapped samples. Clearing rates were calculated on the basis of all charges and all reported crime in 2005. In these measures we excluded crime types such as bicycle theft which is heavily reported (often for insurance purposes) but rarely solved and leading to a charge (<10% of the time) in order not to inflate estimates by an extremely low clearance rate. The fraction of crimes solved within 2 weeks is 0.572 overall, 0.520 for property crimes, and 0.670 for violent crimes. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-16: Effects of DNA profiling on subsequent accumulated probability of conviction and number of convictions, main results and excluding low clearance crimes

	P(convicted)			**# convictions**		
	All	Fast	Slow	All	Fast	Slow
Years	(1)	(2)	(3)	(4)	(5)	(6)
Main results						
1 year	-0.065***	-0.057**	-0.016	-0.093**	-0.076**	-0.017+
	(0.019)	(0.018)	(0.010)	(0.029)	(0.026)	(0.010)
2 years	-0.075**	-0.084***	-0.030+	-0.163***	-0.124**	-0.039*
	(0.024)	(0.023)	(0.016)	(0.048)	(0.041)	(0.019)
3 years	-0.047+	-0.067**	-0.006	-0.129*	-0.127*	-0.002
	(0.025)	(0.024)	(0.019)	(0.061)	(0.051)	(0.026)
Observations	66911	66911	66911	66911	66911	66911
Excluding low clearance crimes						
1 year	-0.071***	-0.065***	-0.016	-0.107***	-0.089***	-0.017+
	(0.019)	(0.018)	(0.010)	(0.027)	(0.024)	(0.010)
2 years	-0.076**	-0.087***	-0.030+	-0.161***	-0.122**	-0.039*
	(0.024)	(0.023)	(0.016)	(0.044)	(0.037)	(0.019)
3 years	-0.061*	-0.081***	-0.006	-0.138*	-0.135**	-0.002
	(0.025)	(0.024)	(0.019)	(0.057)	(0.046)	(0.026)
Observations	66911	66911	66911	66911	66911	66911
Pre-reform baseline, 1 year						
Main outcomes	0.153	0.132	0.029	0.189	0.158	0.031
Excluding low clearance crimes	0.141	0.122	0.029	0.174	0.144	0.031

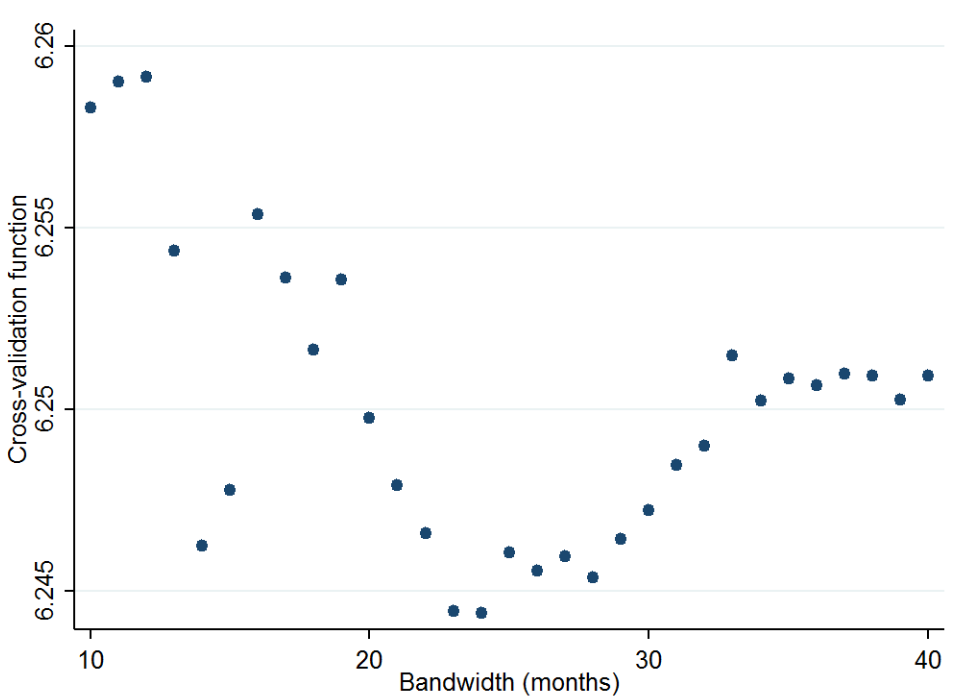
Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Note: Table shows 2SLS estimates of regressing subsequent crime on DNA registration (instrumented by timing of initial charge - before/after reform).The first panel reproduces our main results, but the second panel excludes crime types such as bicycle theft which is heavily reported (often for insurance purposes) but rarely solved and leading to a charge (<10% of the time), which corresponds to the crimes included when calculating the overall clearance rates. Covariates include age, immigrant background, has children, single, years of education, gross income, employment status, number of prior charges, crime type dummies and month fixed effects. Standard errors are clustered by personal identification number. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Table A-17: Test for external validity of LATE estimates

	P(convicted)			**# convictions**		
	All	Fast	Slow	All	Fast	Slow
Years	(1)	(2)	(3)	(4)	(5)	(6)
1 year	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$	$p=0.151$	$p<0.001$
2 years	$p<0.001$	$p<0.001$	$p=0.630$	$p=0.013$	$p=0.686$	$p<0.001$
3 years	$p=0.094$	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$	$p<0.666$

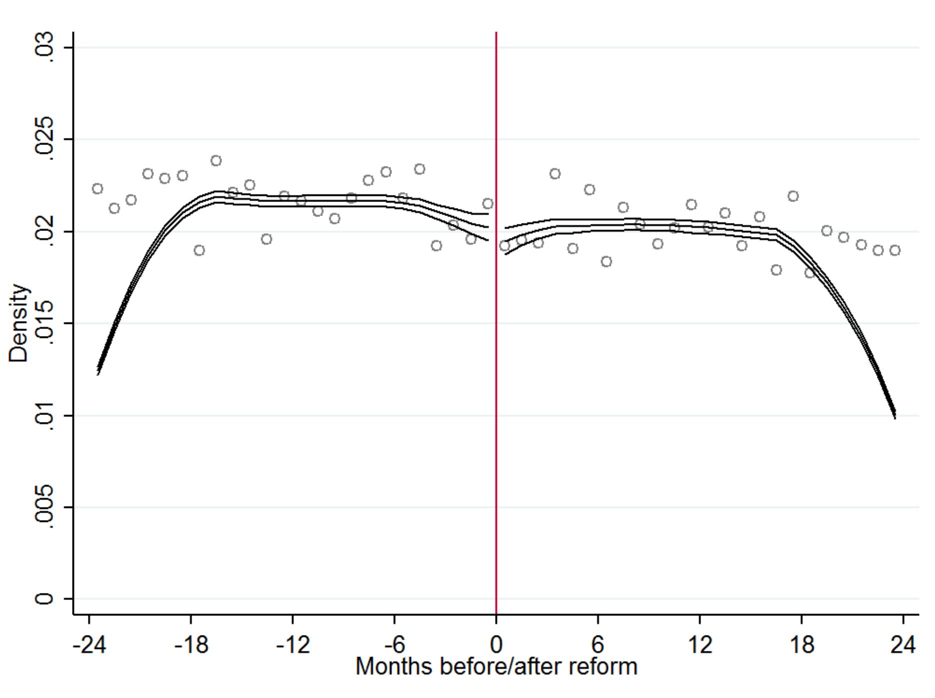
Note: Table shows tests for external validity of the IV estimates reported in Table 7 following Brinch et al. (2017):
$E(Y|DNA=0, Z=1) - E(Y|DNA=0, Z=0) =$
$E(Y|DNA=1, Z=1) - E(Y|DNA=1, Z=0)$
in the limit around the reform Z. The naught is that treatment effects are homogeneous and the alternative is that treatment effects are heterogeneous across the two treatment margins $Z=0$ (where approximately 5% are included in the DNA register) and $Z=1$ (where approximately 40% are included in the DNA register), see Figure 1a. Intuitively, this test correponds to testing whether there would be a significant slope if we estimated Marginal Treatment Effects between the two points of variation.

Figure A-1: Cross-validation function by bandwidth



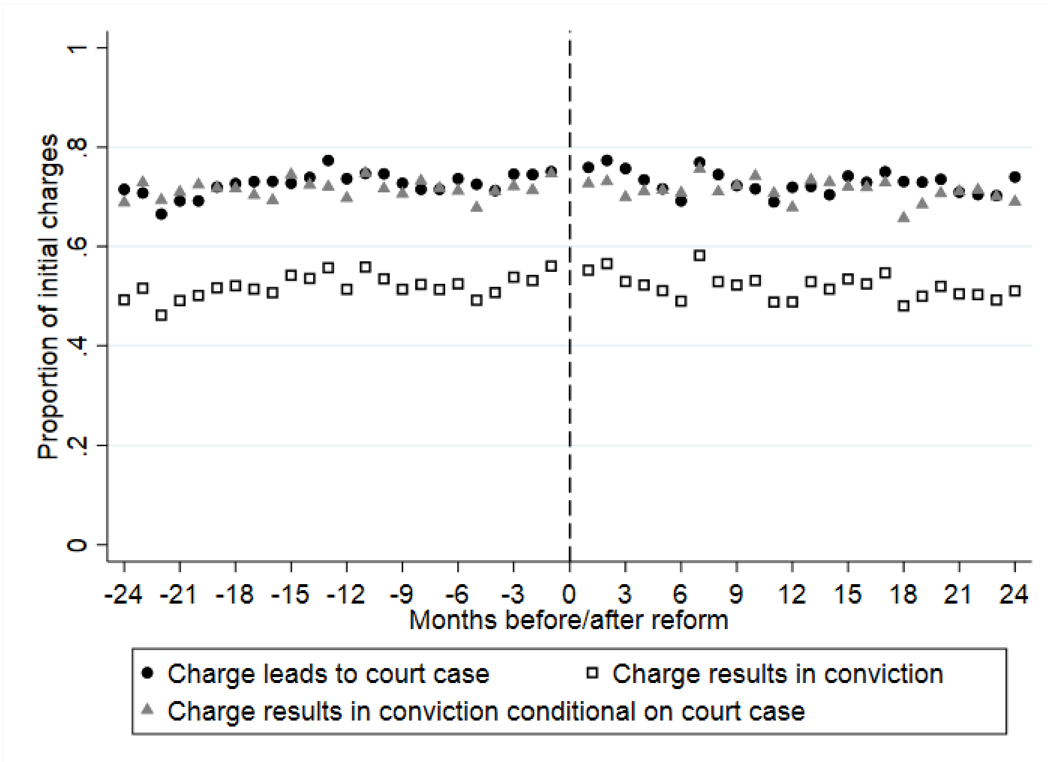
Note: The figures shows the cross-validation (CV) function plotted against different bandwidths. The CV function is calculated in two steps (as described in Lee and Lemieux (2010) and Ludwig and Miller (2005)). First, we estimated the reduced form estimates with a dummy variable indicating before/after reform and running variables measuring months before or after the reform (+ covariates), but leaving out observations in the 1-3 month preceding and following the reform. Second, we used the estimates to predict the outcome for the observations in the excluded window around the reform, and calculate the mean prediction error for each outcome. The prediction errors (CV functions) were then aggregated across the outcomes and across 1-3 month prediction windows. This was done for bandwidths ranging from 10 to 40 months before/after the reform. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Figure A-2: McCrary density test



Note: Figure shows density before and after reform in bins of one month. A McCrary test for discontinuity in density (with default bandwidth) gives a theta of -0.041 with standard error of 0.030 and a t-value of -1.339. Source: Own calculations based on Data from Statistics Denmark and the National Police.

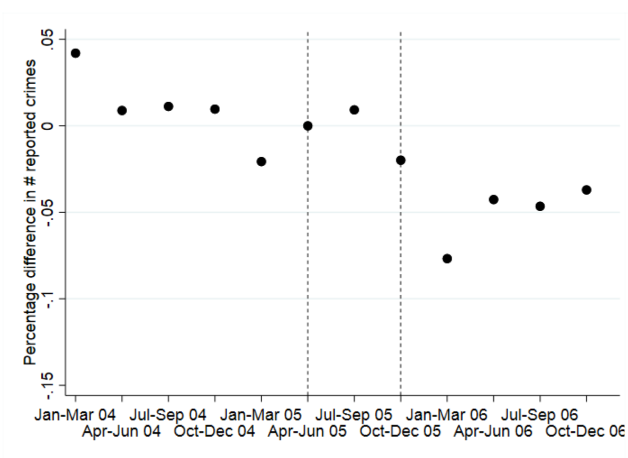
Figure A-3: Probability of charge leading to a court case and a conviction by date of charge



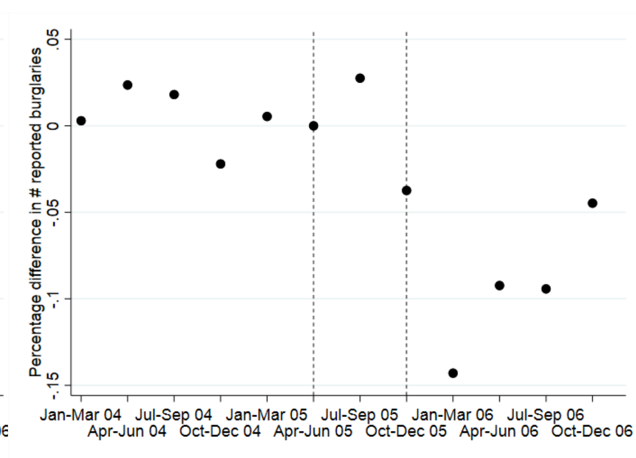
Note: Figure shows, by month of charge relative to the reform, the likelihood of charges leading to a court case, charges leading to a conviction, and charges leading to a conviction conditional on going to court.

Figure A-4: Reported crime relative to April-June 2005

(a) All crime



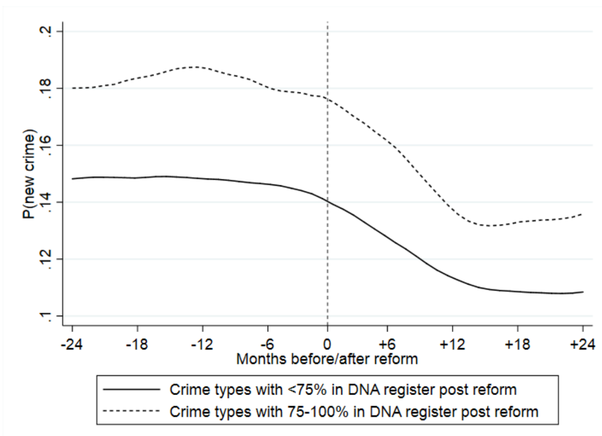
(b) Burglaries



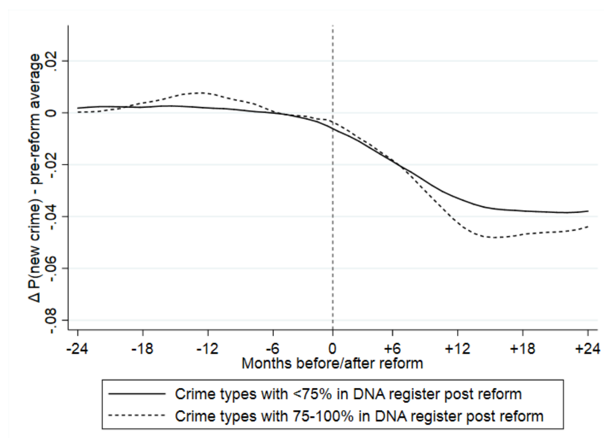
Note: Figure shows the number of reported crimes (/burglaries) relative to April-June 2005 level.

Figure A-5: Crime levels before and after the reform for the Difference in Difference control and treament groups

(a) P(conviction), 1 year, level

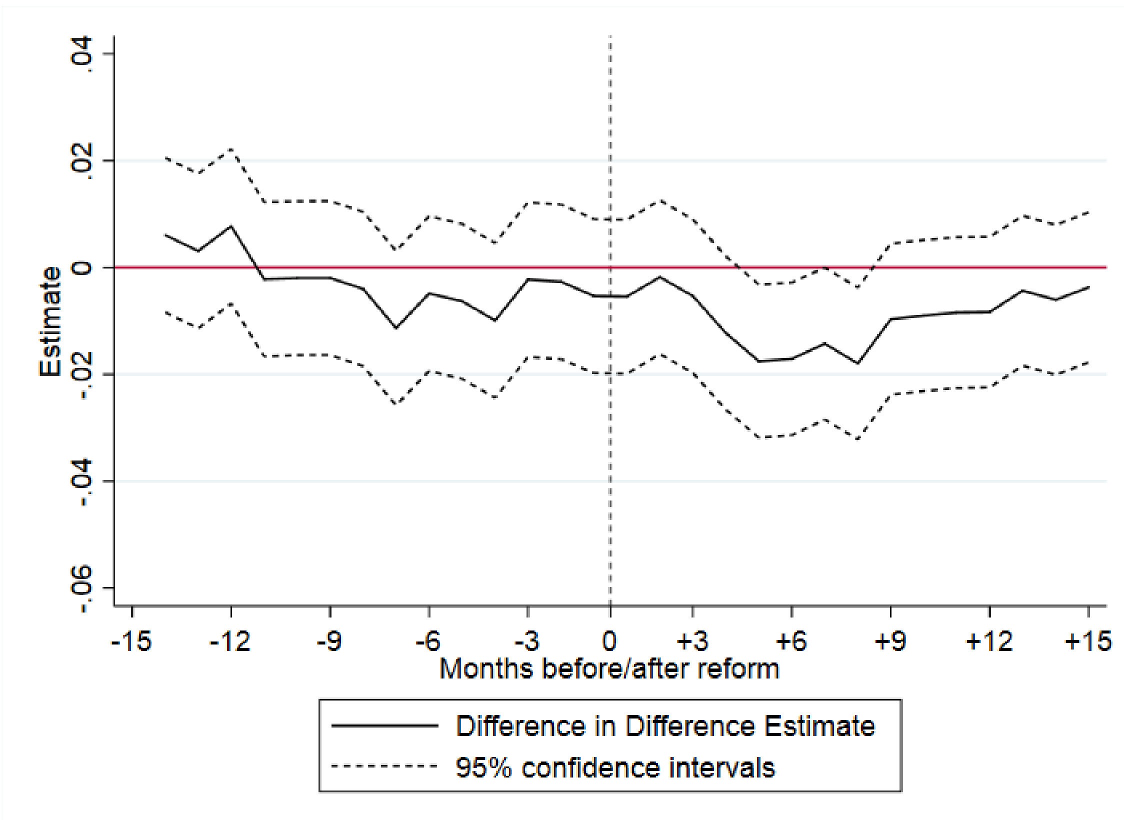


(b) P(conviction), 1 year, demeaned relative to pre-reform level



Note: Figure shows the probability of receiving a conviction for a new crime within the first year after an initial charge for charges pressed 24 months before the reform until 24 months after the reform. The crime levels are separated by treatment status, where the treatment group are those with crime types where at least 75% lead to DNA registration in the post reform period, and the control group are those with crime types where less than 75% were added to the database in the post reform period. The crime types where DNA registration was used pre-reform (homicide, rape, attempted murder, and very serious violence) are excluded from the figure as these groups' DNA registration was unaffected by the reform. Figure A shows the overall crime levels, Figure B shows crime demeaned such that pre-reform crime is equal to zero.

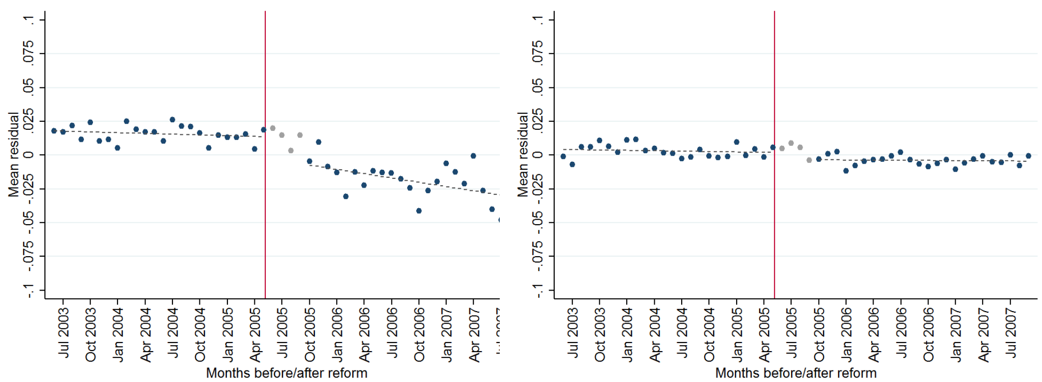
Figure A-6: Difference in Difference estimates using different cutoffs



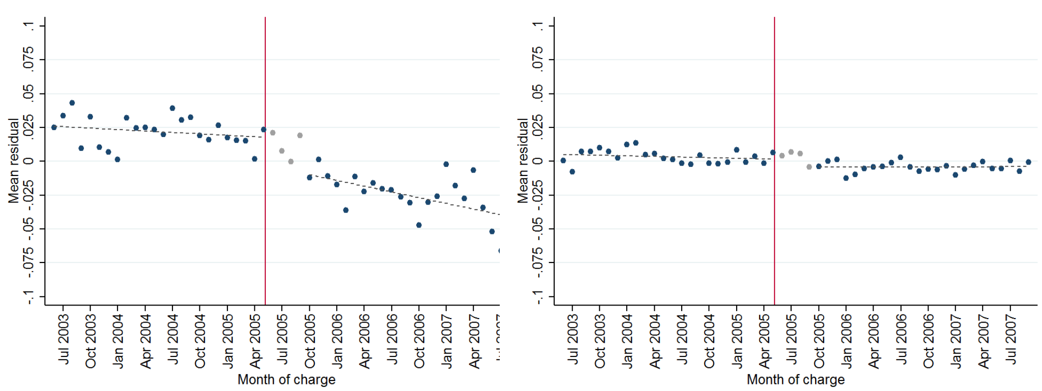
Note: Figure shows Difference in Difference estimates on the probability of receiving a conviction for a new crime within the first year after an initial charge varying the separation of pre and post periods from 15 months before the reform until 15 months after the reform.

Figure A-7: Monthly means of crime outcomes around the timing of the reform, by timing between date of crime and date of charge

(a) P(conviction w. fast charge), 1 year (b) P(conviction w. slow charge), 1 year



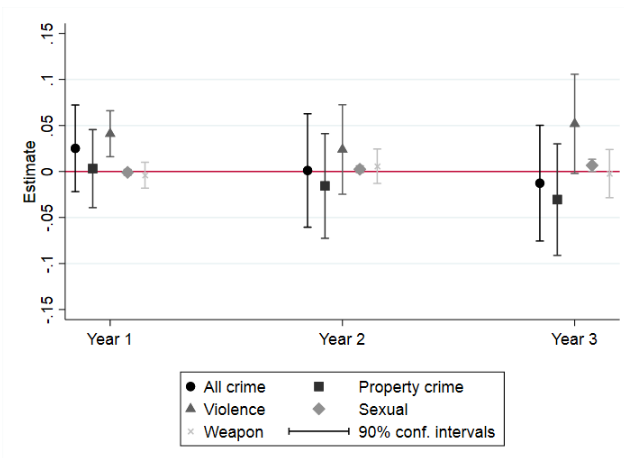
(c) # convictions w. fast charge, 1 year (d) # convictions w. slow charge, 1 year



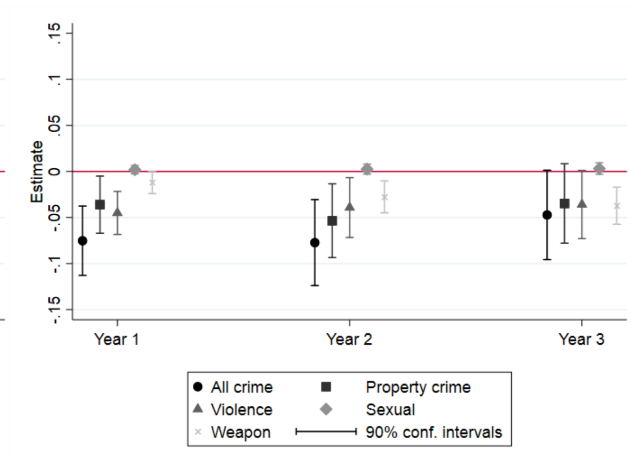
Note: Figures show monthly means of crime outcomes within one year by time it takes to charge the offender for crime. Figures A and B show the probability of receiving at least one conviction and Figures C and D show monthly means number of convictions. Figures A and C show means for charges filed within three weeks from the date of crime, and Figures B and D show results for crime charges filed after three weeks from the date of crime. We condition on covariates in all figures. Therefore the figures show deviations around the conditional sample mean and not absolute levels. Source: Own calculations based on Data from Statistics Denmark and the National Police.

Figure A-8: Differences in crime with and without DNA registration by compliance-status

(a) Difference in crime with DNA registration (Y^1) for always takes and compliers



(b) Difference in crime without DNA registration (Y^0) for never takes and compliers



Note: Figure shows estimated differences between Y^0 (i.e. crime for offenders who are not in the DNA database) for never-takers and compliers, and Y^1 (i.e. crime for offenders who are in the DNA database) for always-takers and compliers using the specification outlined in Black et al. (2015).

B Framework extensions

This section first derives the equations identifying the deterrence effect, the detection effect, and the elasticity of crime with respect to detection probability (Equation (9)) presented in Section 6.1. This section subsequently expands this framework and relaxes the assumptions on (i) invariance of the detection probability $\bar{p}$ across potentially fast and slow solved crimes, and (ii) homogenous deterrence effects across potentially fast and slow solved crimes.

B.1 Baseline identification of deterrence and detection effects

We identify the effects by exploiting the Danish register data. The data both include when offenders are charged for a crime and the exact dates of crime. We divide observed crime $\tilde{y}_i$ into two categories: crime that is solved fast, $\tilde{y}_i^F$, and crime that is solved slowly, $\tilde{y}_i^S$.

The former, $\tilde{y}_i^F$, denotes crime solved within three weeks from the date of crime, before any DNA evidence from the crime scene could have been processed. The latter, $\tilde{y}_i^S$, denotes crime solved after three weeks from the date of crime, at which point DNA evidence could have been processed and been used in the investigation. Hence, changes in crime solved within three weeks from the date of the crime will only capture the deterrence effect, while changes in crime solved more slowly will be a composite of both the deterrence and detection effects (that is, the combined effects on the likelihood that a crime occurs and the likelihood that we observe it in the data). In our main set of results, which we will present in Section 5.2, we will present estimates of DNA registration separately for all observed crime $\tilde{y}_i$, cases solved fast $\tilde{y}_i^F$, and cases solved slowly $\tilde{y}_i^S$, thereby making the different impacts of the deterrence and detection effects explicit. All estimates from $\tilde{y}_i$, $\tilde{y}_i^F$, and $\tilde{y}_i^S$ are attenuated as only a fraction of crime is linked to offenders. However, as estimates using $\tilde{y}_i^F$ do not contain a detection effect, they are not biased upwards and they, therefore, provide a lower bound of the deterrence effect.

We assume that the baseline clearance rate of crime without the DNA database $\bar{p}$ occurs at a fixed rate and that it is uniform and invariant with offender characteristics that are not captured by the different crime types (see footnote 28 in the main text for a further description of this assumption). Thereby, we express the fraction of solved crime that occurs within three weeks from the date of crime as $\pi\bar{p}$ both before and after the expansion of the DNA database. Therefore:

$$\tilde{y}_i^F = \pi\bar{p}y_i,$$

$$\tilde{y}_i^S = ((1-\pi)\bar{p} + \gamma DNA_i)y_i$$

DNA registration's effect on crime solved within three weeks using the reform as an IV equals:

$$\begin{aligned} \beta_F^{IV} &= \pi\bar{p} * E(\Delta) \Rightarrow \\ E(\Delta) &= (\beta_F^{IV})/(\pi\bar{p}), \end{aligned} \tag{B.1}$$

which is the deterrence effect. As we observe all crime reports and the share leading to a charge (the clearance rate) within three weeks from the crime date, we know $\pi\bar{p}$ and may estimate $E(\Delta)$. Turning to the effect on crime solved after three weeks from the crime date:

$$\beta_S^{IV} = E[\gamma DNA_i * y_i^1 + (1-\pi)\bar{p} * \Delta]$$

By subtracting the former estimate β_F^{IV} multiplied by $(1-\pi)/\pi$ from β_S^{IV} we arrive at:

$$\begin{aligned} \beta_S^{IV} - \beta_F^{IV} * (1-\pi)/\pi &= E[\gamma DNA_i * y_i^1 + (1-\pi)\bar{p} * \Delta] - \pi\bar{p} * E(\Delta) * ((1-\pi))/\pi \\ &= E[\gamma DNA_i * y_i^1] \\ &= E(\delta) \end{aligned} \tag{B.2}$$

which is the detection effect, and the elasticity of crime with respect to detection probability:

$$\begin{aligned} E[\epsilon] &= \bar{p} * [(\beta_F^{IV})/(\pi\bar{p})]/[\beta_S^{IV} - \beta_F^{IV} * (1-\pi)/\pi] \\ &= \beta_F^{IV}/[\pi * (\beta_S^{IV} + \beta_F^{IV}) - \beta_F^{IV}] \\ &= \beta_F^{IV}/(\pi\beta^{IV} - \beta_F^{IV}) \end{aligned} \tag{B.3}$$

B.2 Heterogenous baseline detection probability

In our data we observe the fraction of all crime where the offender is caught, and we label this $\bar{p}$. In the baseline framework we assume that $\bar{p}$ is invariant across the time it takes to apprehend the offender. However, it is plausible that the underlying clearance rate for the crimes that are potentially solved fast and slow, respectively, differ. If, for example, fast solved crimes are "low hanging fruits" committed by less skilled criminals and slow solved crimes are committed by more skilled criminals (note that we distinguish between (i) fast and slow solved crimes, and (ii) *potentially* fast and slow solved crimes. The former refers to what we actually observe in the data, the latter to underlying different types of crime).

Therefore, we now expand the framework to allow for two different clearance rates $\bar{p}^F$ for fast solved crime and $\bar{p}^S$ for slow solved crime. As we will show below, the results presented in the main paper are a weighted average between the resulting detection and deterrence effects for potentially fast and slow solved crimes. If fast solved crimes are committed by less skilled criminals and slow solved crimes are committed by more skilled criminals, then the elasticity of crime with respect to the detection probability will be larger for fast solved crimes, because potentially fast solved crime is relatively more responsive to the DNA profilling.

The challenge is that we only observe the fraction of all crime that is solved, and whether this was within three weeks from the date of crime. If potentially fast and slow solved crime, y^F and y^S, are fundamentally different, we cannot separately determine the fraction of y^F and y^S that are not solved. Hence, while we observe $\bar{p}$ for all crime, we cannot distinguish between the underlying fractions of fast and slow solved crime (defined by π and $1-\pi$), and the specific heterogenous clearance rates $\bar{p}^F$ an $\bar{p}^S$. We can only observe that a given fraction of all cases leads to a fast charge, $\pi * \bar{p}^F$, and that another fraction of all cases leads to a slow charge, $(1-\pi) * \bar{p}^S$ where the overall clearance rate is the sum of the two:

$$\bar{p} = \pi * \bar{p}^F + (1-\pi) * \bar{p}^S \tag{B.4}$$

Below we show that heterogenous clearance rates do not change the overall elasticity of crime with respect to the clearance rate. In fact, the overall elasticity is simply a weighted average between the elasticity of fast solved and slow solved crime.

As a starting point, we will revisit how we measure one of the central moments in the baseline framework, the fraction of fast solved crime π. We measure this as the fraction of crime that is solved within three weeks from the date of the crime relative to all crime that is solved. Hence, this fraction implicitly involves the clearance rate. In the case with an invariant clearance rate $\bar{p}$ this will equal $\pi\bar{p}/\bar{p} = \pi$. Yet, if the underlying clearance rate differs across time it takes to solve a crime, then we actually use as π the term $\pi\bar{p}^F/\bar{p}$.

Next, we will expand Equation 5 with counterfactual crime with (y_1) and without (y_0) a DNA database to allow for differences between potentially fast and slow solved crime. Observed fast crime y^F and slow crime y^S are defined as:

$$\begin{aligned}
\tilde{y}_0^F &= \pi\bar{p}^F * y_0 \\
\tilde{y}_1^F &= \pi\bar{p}^F * y_1 \\
\tilde{y}_0^S &= (1-\pi)\bar{p}^S * y_0 \\
\tilde{y}_1^S &= ((1-\pi)\bar{p}^S + \gamma DNA) * y_1
\end{aligned} \tag{B.5}$$

DNA only enter observed slow solved crime, as fast solved crime is always solved before DNA evidence is available. Therefore:

$$\begin{aligned}
\tilde{y}_1^F - \tilde{y}_0^F &= \pi\bar{p}^F * \Delta, \\
\tilde{y}_1^S - \tilde{y}_0^S &= (1-\pi)\bar{p}^S * \Delta + \gamma DNA * y_1
\end{aligned}$$

From this we see that the deterrence effect Δ is identified from the fast solved crime, just as in the baseline framework where we had an invariant clearance rate:

$$\Delta = \frac{\tilde{y}_1^F - \tilde{y}_0^F}{\pi \bar{p}^F} \implies$$
$$E(\Delta) = \frac{\beta_F^{IV}}{\pi \bar{p}^F}$$

What has changed, however, is the identification of the detection effect δ:

$$\tilde{y}_1^S - \tilde{y}_0^S = (1-\pi)\bar{p}^S * \Delta + \gamma DNA_i * y_i^1 \implies$$
$$\tilde{y}_1^S - \tilde{y}_0^S - (1-\pi)\bar{p}^S * \Delta = \gamma DNA_i * y_i^1$$
$$= \delta$$

Inserting the result for the clearance rate from above yields:

$$E(\delta) = \beta_S^{IV} - \frac{1-\pi}{\pi}\frac{\bar{p}^S}{\bar{p}^F} * \beta_F^{IV}$$

This is identical to the baseline expression except for the fraction $\bar{p}^S/\bar{p}^F$, which for the homogenous $\bar{p}$ would have been cancelled out. Therefore, we can express the corresponding elasticities of crime with respect to the detection probability as done in Equation B.3 in the baseline framework:

$$\begin{gathered} \epsilon^F = \bar{p}^F * \frac{\Delta}{\delta} \qquad\qquad \epsilon^S = \bar{p}^S * \frac{\Delta}{\delta} \\ \Downarrow \\ E(\epsilon^F) = \bar{p}^F \frac{\beta_F^{IV}}{\pi \bar{p}^F \beta^{IV} - \bar{p}\beta_F^{IV}} \qquad\qquad E(\epsilon^S) = \bar{p}^S \frac{\beta_F^{IV}}{\pi \bar{p}^F \beta^{IV} - \bar{p}\beta_F^{IV}} \end{gathered} \tag{B.6}$$

From Equation B.6, it also follows that the weighted average between the two elasticities $\pi\epsilon^F + (1-\pi)\epsilon^S$ equals the overall elasticity, which we estimate in Table 9 in the main text:

$$\begin{gathered} \pi \bar{p}^F \frac{\beta_F^{IV}}{\pi \bar{p}^F \beta^{IV} - \bar{p}\beta_F^{IV}} + (1-\pi)\bar{p}^S \frac{\beta_F^{IV}}{\pi \bar{p}^F \beta^{IV} - \bar{p}\beta_F^{IV}} = \\ \bar{p}\frac{\beta_F^{IV}}{\pi \bar{p}^F \beta^{IV} - \bar{p}\beta_F^{IV}} \end{gathered} \tag{B.7}$$

which collapses to the elasticity from the baseline framework: $\frac{\beta_F^{IV}}{\pi\beta^{IV} - \beta_F^{IV}}$ if $\bar{p}^F = \bar{p}^S = \bar{p}$.

Recall from above that we in the baseline framework with an invariant $\bar{p}$ calculate the fraction π as $\pi\bar{p}^F/\bar{p}$. Inserting this into Equation B.3 from the main text, we get:

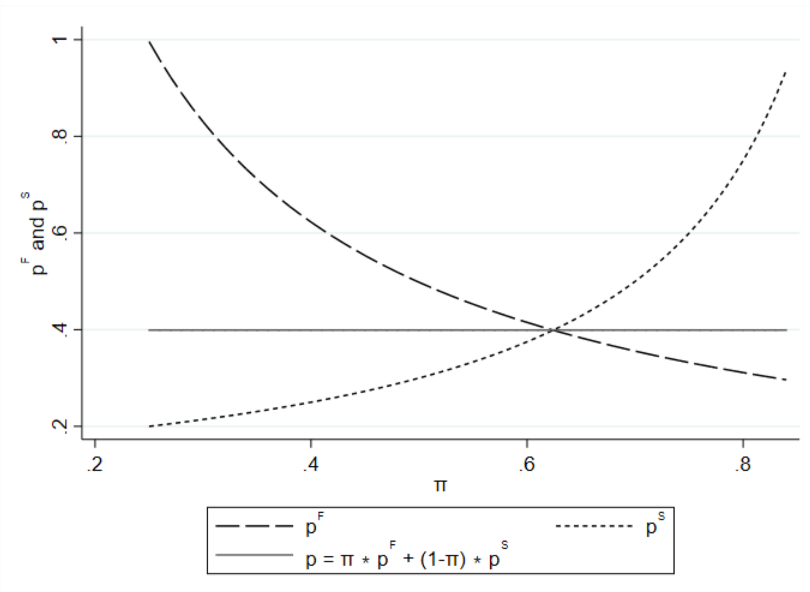
$$\begin{aligned} &\frac{\beta_F^{IV}}{\pi\frac{\bar{p}^F}{\bar{p}}\beta^{IV} - \beta_F^{IV}} = \\ &\bar{p}\frac{\beta_F^{IV}}{\pi\bar{p}^F\beta^{IV} - \bar{p}\beta_F^{IV}} \end{aligned} \tag{B.8}$$

which is exatly the expression from Equation B.7 above. To illustrate this, Figure B.1a shows values of clearance rates $\bar{p}^F$, and $\bar{p}^S$ across values of π and Figure B.1b shows the elasiticities for fast and slow solved crime, ϵ^F, and ϵ^S, across values of π. The figure shows that the weighted average between the fast and slow crime elasticities in Equation B.7 equals the elasticity we report in Table 9. Hence, the results reported in the paper are robust to different clearance rates across fast and slow solved crime.

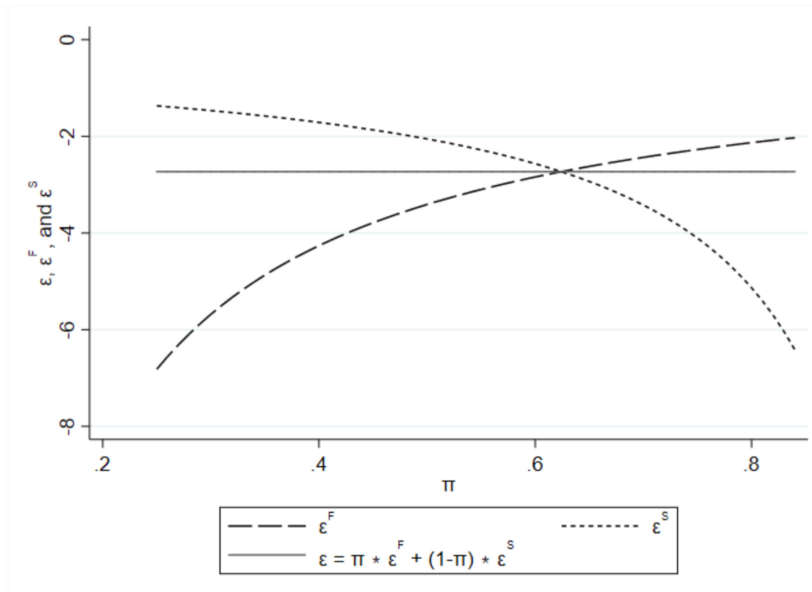
Going back to our initial example, if fast solved crimes are "low hanging fruits" committed by less skilled criminals and slow solved crimes are committed by more skilled criminals this suggests that the clearance rate for potentially fast solved crimes is larger than the clearance rate for potentially slow solved crimes ($p^F > p^S$). Figure B.1a shows that this implies that the underlying fraction of potentially fast solved crime, π, is smaller than suggested in the main text (if the fast solved crimes we actually link to offenders constitute a larger fraction of total potentially fast solved crimes, then π has to be smaller). Figure B.1b shows that the corresponding elasticity for fast solved crime with respect to the detection probability is thus larger whereas for slow solved crime it is smaller (as the actual response we observe for fast solved crime is now relatively larger because the fraction of fast solved crime is lower).

Figure B.1: Consequence of heterogeneous detection probability across fast and slow solved crime across the fraction of potentially fast solved crime π

(a) Detection probability



(b) Elasticity of crime with respect to detection probability



Note: Figure shows simulation results using the estimates from Table 9. The figure illustrates how hetergeneous detection probabilities across the time it takes to solve a crime would affect our results under the assumption of homogeneous deterrence effects (i.e. the only variation is in the detection probability). Figure A shows detection probabilities across share of crimes that are solved "fast" and "slow"; π. With hetergeneous detection probabilities, the overall detection probability remains unchanged, as it is a weighted average between the different detection probabilities. Figure B shows the corresponding elasticities of crime with respect to the detection probability. Again, the overall estimate is unchanged, as it is weighted average across π. However, the underlying elasticities change such that if we overestimate the share of fast solved crime (low π) then the elasticity of fast solved crime is underestimated and vice versa.

B.3 Heterogenous deterrence effects

In our baseline framework we identify the deterrence effect Δ from fast solved crimes, and use this together with the results for slow solved crimes to isolate the detection effect δ and thus also the elasticity of crime with respect to the detection probability ϵ. We now consider the case where there is not a uniform Δ for the two types of crime, but instead different deterrence effects for fast and slow solved crime Δ^F and Δ^S, respectively.

This complicates things to a larger degree than in the previous subsection. Different deterrence effects can arise for many different reasons as, for example, unobservable heterogeneity or nonlinearity. Hence, there is almost no limit to the possible deviations from our baseline framework. To make progress from this observation and study the consequences of different deterrence effects within our framework, we simply assume that the difference between the two deterrence effects are a scalar $\Delta^S - \Delta^F = d$.

We show below that this not only results in different elasticities of crime with respect to the detection probability for fast and slow solved crimes, it also changes the average estimate; what we report in Table 9 is biased. This bias will, however, be relatively small. If the two deterrence effects differ by 20%, the average elasticity will be biased by approxiately 10% (i.e. be either -2.9 or -2.4 instead of -2.7, depending on the direction of the difference).

Focussing first on fast solved crime, we will still identify the deterrence effect:

$$E[\Delta^F] = \frac{\beta_F^{IV}}{\pi \bar{p}} \tag{B.9}$$

However, we cannot identify the corresponding for slow solved crime. Instead, we now consider the consequence of different degrees of heterogeneity between Δ^F and Δ^S.

We can express Equation B.3 from the baseline framework as:

$$\beta_S^{IV} = (1-\pi)\bar{p}\Delta^S + \gamma DNAy_1 \tag{B.10}$$

As we here consider heterogeneity in the deterrence effect only, the detection effect, δ, will still be given by the last term $\gamma DNAy_1^S$. Furthermore, by inserting the difference between the deterrence effects for fast and slow solved crime, we get:

$$\begin{aligned}
\beta_S^{IV} = (1-\pi)\bar{p}(\Delta^F + d) + \gamma DNAy_1 \implies \\
E[\delta] &= \beta_S^{IV} - (1-\pi)\bar{p}(\Delta^F + d) \\
&= \beta_S^{IV} - (1-\pi)\bar{p}(\frac{\beta_F^{IV}}{\pi\bar{p}} + d) \\
&= \beta_S^{IV} - \frac{1-\pi}{\pi}\beta_F^{IV} - (1-\pi)\bar{p}d
\end{aligned} \tag{B.11}$$

Hence, if there are heterogeneous deterrence effects, our estimated detection effect will be biased by $-(1-\pi)\bar{p}d$. If the deterrence effect for fast solved crime is numerically larger than for slow solved crime ($d > 0$), we underestimate the detection effect and vice versa.

To see how this affects our estimated elasticity of crime with respect to the detection probability, we use the baseline relationship from Equation B.3 that $\epsilon = \bar{p}\Delta/\delta$, but expand it to allow for heterogeneous deterrence effects:

$$\epsilon^F = \bar{p}\frac{\Delta^F}{\delta}, \qquad \epsilon^S = \bar{p}\frac{\Delta^S}{\delta}$$

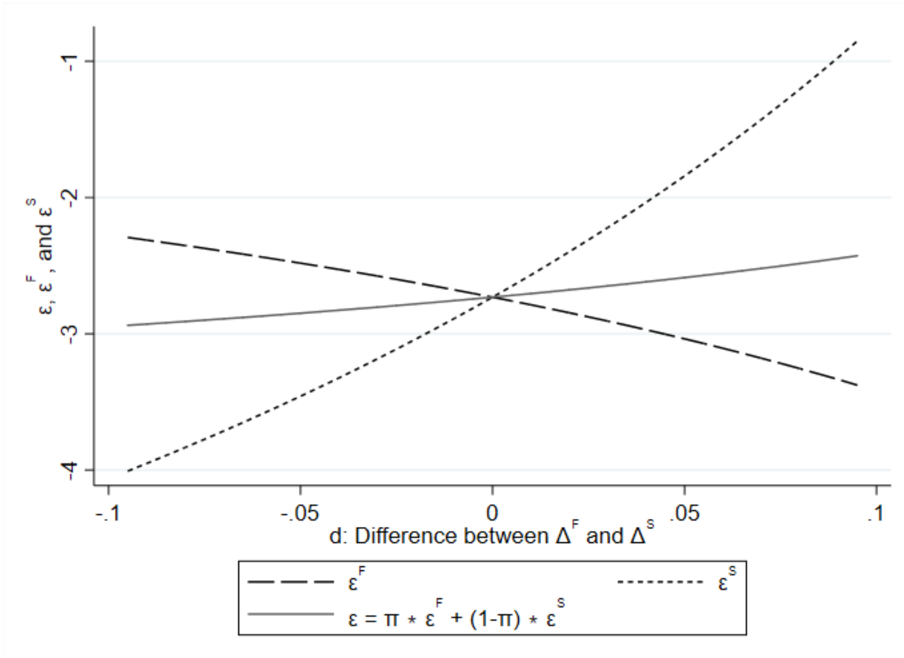
$$\Downarrow$$

$$E(\epsilon^F) = \bar{p}\frac{\beta_F^{IV}}{\pi\bar{p}\beta^{IV} - \beta_F^{IV} - (1-\pi)\bar{p}d} \qquad E(\epsilon^S) = E(\epsilon^F) + \frac{d}{\pi\bar{p}\beta^{IV} - \beta_F^{IV} - (1-\pi)\bar{p}d} \tag{B.12}$$

Figure B.2 shows the resulting elasticities along with the average elasticity across different levels of heterogeneity d.

The figure shows that heterogeneous deterrence effects would result in elasticities that differ subtantially from each other. There is an inverse relationship between the two elasticities across the heterogeneity d. The reason is that a higher d implies a lower deterrence effect for slow solved crime, and thereby also a lower detection effect. This decrease makes the elasticity for fast solved crime increase (because the numerator decreases), while for slow

Figure B.2: Elasticity of crime with respect to detection probability in the case of heterogeneous deterrence effects between fast and slow solved crimes



Note: Figure shows simulation results using the estimates from Table 9. The figure shows how heterogeneous deterrence effects across the time it takes to solve a crime would affect our estimated elasticity of crime with respect to the detection probability. The figure plots the resulting elasticities for all crime, and fast and slow solved crime across d, a scalar difference between the two deterrence effects.

solved crime the deterrence effect will decrease at a faster rate than the detection effect (by d and $(1-\pi)\bar{p}d < d$, respectively), thereby reducing the elasticity.

Yet, the figure also shows that the overall impact on the average elasticity of crime with respect to the detection probability is small. If there is a heterogeneity of ± 0.1 in deterrence effects (corresponding to $\pm 20\%$), then the average elasticity would only vary between -2.9 and -2.4, which corresponds to 10% relative to our main estimate of -2.7 from Table 9.