Misdemeanor Prosecution*

Amanda Agan† Jennifer L. Doleac‡ Anna Harvey§

April 11, 2021

Abstract

Communities across the United States are reconsidering the public safety benefits of prosecuting nonviolent misdemeanor offenses. So far there has been little empirical evidence to inform policy in this area. In this paper we report the first estimates of the causal effects of misdemeanor prosecution on defendants' subsequent criminal justice involvement. We leverage the as-if random assignment of nonviolent misdemeanor cases to Assistant District Attorneys (ADAs) who decide whether a case should move forward with prosecution in the Suffolk County District Attorney's Office in Massachusetts. These ADAs vary in the average leniency of their prosecution decisions. We find that, for the marginal defendant, nonprosecution of a nonviolent misdemeanor offense leads to large reductions in the likelihood of a new criminal complaint over the next two years. These local average treatment effects are largest for first-time defendants, suggesting that averting initial entry into the criminal justice system has the greatest benefits. We also present evidence that a recent policy change in Suffolk County imposing a presumption of nonprosecution for a set of nonviolent misdemeanor offenses had similar beneficial effects: the likelihood of future criminal justice involvement fell, with no apparent increase in local crime rates.

*We thank Suffolk County District Attorney Rachael Rollins and the Suffolk County District Attorney's Office for their cooperation, and especially thank Bobby Constantino for his tireless efforts in helping us access and understand the data. Thanks to Rebecca Regan for excellent research assistance, to Manudeep Bhuller for code and discussions about calculating the complier means, and to Martin Andresen for extensive help with the estimation of marginal treatment effects via `mtefe`. We thank Paul Goldsmith-Pinkham, Peter Hull, Michal Kolesár, Emily Leslie, Justin McCrary, and Sam Norris for conversations that improved the paper. We appreciate feedback from Jim Greiner, Steve Lehrer, James MacKinnon, Steven Raphael, and Megan Stevenson; seminar participants at American University, Bentley University, Boston University, the Federal Reserve Bank of Minneapolis, Georgia State University, LSE-Centre for Economic Performance, Michigan State University, Queen's University, Rutgers University, University of Pittsburgh, University of Toronto, University of Virginia School of Law, and West Virginia University; and conference participants at the 2020 Emory University Conference on Institutions and Law Making, the 2020 APPAM Fall Research Conference, and the 2020 Duke University Empirical Criminal Law Roundtable. We gratefully acknowledge funding from the W.T. Grant Foundation and the Chan Zuckerberg Initiative.

†Department of Economics, Rutgers University, amanda.agan@rutgers.edu

‡Department of Economics, Texas A&M University, jdoleac@tamu.edu

§Department of Politics, New York University, anna.harvey@nyu.edu

1 Introduction

Every year approximately 13 million Americans are charged with misdemeanor offenses, and misdemeanor cases make up over 80 percent of the cases processed by the U.S. criminal justice system (Stevenson and Mayson, 2018). Many have expressed concern that prosecuting this volume of low-level offenses may do more harm than good (Natapoff, 2018). Some district attorneys across the country have begun to implement alternatives to misdemeanor prosecution, particularly for nonviolent defendants.[1] By allowing those charged with nonviolent misdemeanor offenses to avoid the potential negative consequences of a criminal prosecution (including time away from work and family, a criminal record of an arrest, and a possible criminal record of a conviction), alternatives to misdemeanor prosecution may decrease defendants' subsequent criminal justice contact. On the other hand, alternatives to misdemeanor prosecution may reduce specific deterrence, causing increases in future criminal behavior. The net causal effect of prosecution in marginal misdemeanor cases is, thus, an empirical question, but there is little evidence to guide prosecutors' policy choices. Simply comparing defendants who are prosecuted on misdemeanor charges with those who are not would be misleading, as these groups are likely different in ways that are both observable and unobservable to the researcher.

In this paper we use new data on the prosecution of nonviolent misdemeanor criminal complaints from the Suffolk County District Attorney's Office (SCDAO) in Massachusetts between 2004 and 2018 to estimate the impact of nonprosecution of nonviolent misdemeanors on future criminal justice system contact. In Suffolk County, an individual against whom a criminal complaint has been issued must appear at an initial arraignment hearing. Assistant District Attorneys (ADAs) assigned to arraignment courtrooms have the discretion to dispose of a complaint prior to or at the arraignment hearing, or to proceed with prosecution. We use the term "nonprosecution" to refer to cases that both a) do not proceed past the day of arraignment and b) do not result in a conviction or an "admission to sufficient facts."[2] "Prosecution" refers to all other cases.[3] Our empirical strategy exploits the as-if random assignment of cases to arraigning ADAs who vary in the leniency of their prosecution decisions.[4] This empirical design recovers the local average treatment effect (LATE),

[1] These alternatives can include, depending on the jurisdiction, declining to prosecute, diversion, dismissal, pretrial probation, and deferred adjudication.

[2] Under Massachusetts law, a defendant may "admit to sufficient facts to warrant a finding of guilty" (https://www.mass.gov/rules-of-criminal-procedure/criminal-procedure-rule-12-pleas-and-plea-agreements).

[3] The term 'nonprosecution" could also be defined to include cases that prosecutors eventually decline to prosecute well after the day of arraignment. Our empirical design only allows us to recover causal effects from nonprosecution that occurs no later than the day of arraignment. We thus use the term "nonprosecution" throughout to refer to nonprosecution that occurs no later than the day of arraignment.

[4] Arraigning ADAs and criminal complaints are assigned to arraignment courtrooms through separate

or the causal effect of nonviolent misdemeanor nonprosecution for individuals at the margin of nonprosecution, i.e., individuals for whom different arraigning ADAs might have made different prosecution decisions (Imbens and Angrist, 1994).[5]

We measure ADA leniency using the leave-out, residualized mean of an ADA's nonprosecution choices based on all other nonviolent misdemeanor cases that the ADA has arraigned not involving the current defendant.[6] The leave-out leniency measure is highly predictive of prosecution decisions, but uncorrelated with case and defendant characteristics. We find that going from the least lenient to the most lenient ADA increases a defendant's probability of nonprosecution by about ten percentage points (the mean rate of nonprosecution in our sample is 21%). Importantly, in Suffolk County cases that move forward with prosecution after the arraignment are assigned to a different ADA, not the ADA at arraignment, allowing us to separately identify the effects of being assigned to a lenient arraigning ADA as opposed to a lenient ADA in other phases of the case.

We first estimate the impacts of misdemeanor nonprosecution on subsequent criminal complaints (arrests). We find that the marginal nonprosecuted misdemeanor defendant is 33 percentage points less likely to be issued a new criminal complaint within two years post-arraignment (58% less than the mean for "complier" defendants who are prosecuted; $p < 0.01$).[7] We find that nonprosecution reduces the likelihood of a new misdemeanor

processes that we describe in detail in the text. Throughout we use the term "assignment" as a shorthand reference to the processes by which arraigning ADAs come to represent SCDAO at arraignment for a set of criminal complaints.

[5]Other papers using similar research designs based on as-if random assignment of cases to decision makers (generally judges) include Kling (2006), Loeffler (2013), Aizer and Doyle (2015), Mueller-Smith (2015), Aneja and Avenancio-León (2019), Eren and Mocan (2019), and Bhuller et al. (2020) to estimate the impacts of incarceration; Bhuller et al. (2018), Dobbie et al. (2018), Arteaga (2020) and Norris, Pecenco and Weaver (2020) to estimate the impact of parental incarceration on children; Collinson and Reed (2019) and Humphries et al. (2019) to estimate impacts of eviction; Di Tella and Schargrodsky (2013) to estimate the impacts of electronic monitoring; Dahl, Kostøl and Mogstad (2014) to estimate the impact of parental disability insurance receipt on children's usage of the program; Maestas, Mullen and Strand (2013) and French and Song (2014) to estimate impacts of disability insurance receipt on labor supply; Doyle Jr (2007) to estimate the impact of foster care on child outcomes; Diamond, Guren and Tan (2020) to estimate impacts of foreclosures; and Dobbie, Goldin and Yang (2018), Stevenson (2018), and Leslie and Pope (2017) to estimate the impacts of pretrial detention; and Dobbie, Goldsmith-Pinkham and Yang (2017) to estimate the impacts of consumer bankruptcy. The only other papers we know of exploiting as-if random assignment of cases to prosecutors are Sloan (2020*b*) and Sloan (2020*a*), describing the variation in prosecutor leniency and estimating the impact of other-race prosecutor bias on charging decisions in the office of the District Attorney of New York County.

[6]We also explore robustness to other estimation strategies for the first stage, including using all the ADA dummies directly in 2SLS or LIML, UJIVE estimation, and using lasso to choose among the many instruments. See Appendix C.1 for more discussion.

[7]See Appendix C.3 for details on the calculation of average outcomes among prosecuted compliers. For all two-stage least squares (2SLS) estimates reported in the paper, we also compute Anderson-Rubin confidence intervals by inverting the weak-instrument test of Anderson and Rubin (1949), as suggested by Andrews, Stock and Sun (2019) and Lee et al. (2020). The Anderson-Rubin confidence intervals have correct size and optimal power even with weak instruments when models are just-identified; our main estimates remain

complaint by 24 percentage points (60%; $p < 0.01$), and reduces the likelihood of a new felony complaint by 8 percentage points (47%; not significant). Nonprosecution reduces the number of subsequent criminal complaints by 2.1 complaints (69%; $p < .01$); the number of subsequent misdemeanor complaints by 1.2 complaints (67%; $p < .01$), and the number of subsequent felony complaints by 0.7 complaints (75%; $p < .05$). We see significant reductions in subsequent criminal complaints for violent, disorderly conduct/theft, and motor vehicle offenses. Our primary estimates follow defendants for two years post-arraignment, but we show that our results are robust to one-year and three-year post-arraignment windows (effects appear to grow over time). We also see similar declines in the probabilities of subsequent criminal prosecution and subsequent criminal record acquisition.

Misdemeanor prosecution may have negative effects for marginal nonviolent defendants because it may pull some defendants into the criminal justice system who otherwise would remain outside that system. If this is the case, then we would expect to see larger effects for first-time defendants, relative to defendants who appear in our data at least once before. In line with this hypothesis, we find that our estimates are larger (particularly in percentage terms) and more precisely estimated for first-time defendants.[8]

We run a battery of additional checks to support the validity of our instrumental variable design. We demonstrate that our first stage is positive and significant at conventional thresholds across all covariate subgroups, supporting the assumption of average monotonicity required to interpret our estimates as LATEs (Frandsen, Lefgren and Leslie, 2019). Our results are also robust to interacting our ADA leniency instrument with offense types, victim/victimless offenses, and ADA experience; these interacted instruments allow the effect of leniency to vary across groups. While we are unable to test the exclusion restriction directly, we consider whether the leniency of the ADA at arraignment affects other case outcomes among defendants who are prosecuted. If arraigning ADAs only affect the decision to prosecute, then ADA leniency should not affect these other outcomes when we restrict attention to the prosecuted group. As expected, we find that our ADA leniency measure is uncorrelated with case outcomes for the set of prosecuted defendants. We interpret this as support that the exclusion restriction holds.[9] Estimates from reduced form models are

statistically significant when these confidence intervals are used. For more on the just-identified assumption see Appendix C.1. To calculate the effect of going from the least lenient to the most lenient ADA, readers should multiply these estimates by 0.1, the estimated change in the likelihood of nonprosecution.

[8] We use three different definitions of "first-time defendants": those without a prior criminal complaint in Suffolk County, those without a prior DCJIS criminal record of a Suffolk County complaint, and those without a prior Suffolk County conviction. Results are very similar across all three definitions.

[9] Arraigning ADAs can also request bail in cases that they choose to prosecute. In practice this happens infrequently in this sample of cases, and we show that accounting for the bail decision does not affect our results.

consistent with our 2SLS estimates.

We also confirm the robustness of our results to several approaches to imputing missing data. One drawback of our administrative data is that demographic characteristics (gender, age, and race/ethnicity) are missing for some defendants, and are missing more often for nonprosecuted defendants. Because the availability of these variables is correlated with treatment, we do not include them in our main analyses. However, we show that our results are nearly identical when we include age and gender (which are only slightly missing) and predicted race/ethnicity. The identity of the ADA at arraignment is also missing for many of our cases. Our main analysis is done within the sample of cases not missing this ADA information. However, we also construct several progressively larger samples for which missing arraigning ADA information is imputed based on patterns of observed arraigning ADAs by court/day and court/week. Our results are again nearly identical even in our largest imputed sample, which increases the percentage of cases with arraigning ADA information to 76% of all cases.

We consider possible causal mechanisms that could be generating our findings. Cases that are not prosecuted by definition are closed on the day of arraignment. By contrast, the average time to disposition for prosecuted nonviolent misdemeanor cases in our sample is 185 days. This time spent in the criminal justice system may disrupt defendants' work and family lives. Cases that are not prosecuted also by definition do not result in convictions, but 26% of prosecuted nonviolent misdemeanor cases in our sample result in a conviction. Criminal records of misdemeanor convictions may decrease defendants' labor market prospects and increase their likelihoods of future prosecution and criminal record acquisition, conditional on future arrest. Finally, cases that are not prosecuted are at much lower risk of resulting in a criminal record of the complaint in the statewide criminal records system.[10] We find that nonprosecution reduces the probability that a defendant will receive a criminal record of that nonviolent misdemeanor complaint by 55 percentage points (56%, $p < .01$). Criminal

[10] Formal criminal records, called criminal record information (CORI) in Massachusetts, are maintained by the Massachusetts Department of Criminal Justice Information Services (DCJIS) and are available to criminal justice agencies in future cases, as well as to employers under certain conditions. Criminal records of convictions are generally available with defined exceptions. Daycare and preschool employers in Massachusetts may access criminal records of complaints that did not result in convictions, even if defendants take action to have those records sealed. Schools, nursing homes, assisted living facilities, and other employers working with certain vulnerable populations may access criminal records of complaints that did not result in convictions unless defendants take action to have those records sealed (https://www.gbls.org/sites/default/files/2019-11/booklet-10-applying-for-jobs-housing-or-other-opportunities-after-sealing-criminal-records.pdf). Cases dismissed prior to formal arraignment do not receive DCJIS records in Massachusetts. Cases that proceed to and past formal arraignment do receive DCJIS records. In our data we can only identify day of disposition, not whether a disposition that occurs on the day of arraignment occurs before or after formal arraignment.

records of misdemeanor arrests may also damage defendants' labor market prospects and increase their likelihoods of future prosecution and criminal record acquisition, conditional on future arrest. All three of these mechanisms may be contributing to the large reductions in subsequent criminal justice involvement following nonprosecution.

We consider the characteristics of the marginal defendants for whom ADA assignment matters (the "compliers"); this helps us to interpret the LATEs that we estimate. We find that compliers make up about 10% of our sample. The observable characteristics of this group differ on a few dimensions from the sample as a whole. Compliers are less likely to have been charged with a drug offense, to have been charged with a "serious misdemeanor" (punishable by more than 100 days in jail), to have misdemeanor or felony convictions within the prior year, and to be non-citizens.

Our 2SLS estimates are larger in absolute value than our OLS estimates. This difference could be due to selection bias in the OLS estimates and/or the composition of the complier sample. We reweight our OLS estimates to match the sample of compliers using two different reweighting schemes, finding that the reweighted OLS estimates are very similar to the unweighted OLS estimates (Dahl, Kostøl and Mogstad, 2014; Bhuller et al., 2020). This implies that the differences between our OLS and 2SLS estimates are likely being driven by negative selection: arraigning ADAs are more likely to prosecute defendants at lower risk of subsequent criminal justice contact. This negative selection biases OLS estimates toward finding that prosecution has less detrimental effects than it does.[11] Our 2SLS estimates remove this selection bias.

We estimate marginal treatment effects (MTEs) to explore heterogeneity in the LATE (Heckman and Vytlacil, 2005; Heckman, Urzua and Vytlacil, 2006).[12] Consistent with the hypothesis of negative selection, we find larger negative treatment effects as ADA leniency increases.

The results of our analysis imply that if all arraigning ADAs acted more like the most lenient ADAs in our sample when deciding which cases to prosecute, Suffolk County would likely see a reduction in criminal justice involvement for these nonviolent misdemeanor defendants. Because nonviolent misdemeanor defendants in Suffolk County are disproportionately Black, reducing the prosecution of nonviolent misdemeanor offenses would disproportionately benefit Black residents of the county.

We conclude by analyzing the effects of a policy change in Suffolk County establishing

[11] We report evidence suggesting that youth is an example of a characteristic that may contribute to this negative selection effect, since ADAs are more likely to not prosecute younger defendants, despite their higher risk of future criminal justice contact.

[12] See Doyle Jr (2007), Maestas, Mullen and Strand (2013), French and Song (2014), Arnold, Dobbie and Yang (2018), and Bhuller et al. (2020) for empirical examples of MTE estimation in leniency designs.

a presumption of nonprosecution for a list of nonviolent misdemeanor charges. In a series of event study, difference-in-differences (DD), and 2SLS DD models using the date of the policy change as an instrument for nonprosecution, both with and without the inclusion of nonviolent felony cases as a comparison group, we find results that are consistent with our main results: increasing nonprosecution reduced the likelihood of subsequent criminal complaints within a one-year post-arraignment window. In addition, there does not appear to have been an increase in reported crime due to the policy change.

Most closely related to our work is Mueller-Smith and Schnepel (2019), which studies the impact of deferred adjudication for felony cases in Texas. Deferred adjudication is an alternative to prosecution in Texas whereby a defendant pleads guilty to a charge, but the charge is later dismissed after successful completion of a period of probation. Mueller-Smith and Schnepel (2019) exploit two policy changes which changed the probability of deferred adjudication for felony defendants. They find that marginal felony defendants who received deferred adjudication had significantly lower probabilities of subsequent conviction and higher probabilities of subsequent employment.

In our view this study contributes to the literature in several ways. First, we provide the first evidence on the causal effects of the decision to prosecute misdemeanor defendants—a topic of tremendous policy interest. Second, our findings contribute to the literature on the net costs and benefits of criminal justice intervention, and the diminishing marginal returns to such interventions. In this context, it appears that prosecuting defendants for nonviolent misdemeanor offenses has substantial costs for those individuals without any evidence of public safety benefits (and suggestive evidence of public safety costs). Finally, we add to a growing literature that uses as-if randomization of cases to decision makers (in this case arraigning ADAs) to measure the causal effects of their decisions. There has been substantial recent work refining this econometric method, and we apply this method in a new context, using current best practices.

2 Background

The large volume of misdemeanor cases in the United States arises from the criminalization of relatively common behaviors. These behaviors may include (depending on the jurisdiction) disorderly conduct, disturbing the peace, possession of small quantities of prohibited substances, trespassing, and driving without a valid license/registration/insurance. Some have argued that, in fact, the criminalization of relatively minor unwanted behaviors may increase their incidence, relative to other strategies (Natapoff, 2018). Many such behaviors may stem from root causes such as poverty, mental illness, or substance abuse; for these, alternative strategies such as directing social services to those engaged in these behaviors could potentially be more successful. Misdemeanor convictions can decrease employment prospects, increasing the likelihood that those with misdemeanor conviction records turn to illegal forms of economic activity (Uggen et al., 2014; Leasure, 2019). Even when misdemeanor prosecutions do not result in convictions, lengthy prosecutions of misdemeanor arrests both disrupt defendants' work and family lives, and increase the probability that the arrest goes onto the defendant's criminal record, again potentially increasing the likelihood of subsequent offending behavior. Natapoff (2018) argues that, for these reasons, misdemeanor prosecution "makes our entire country less safe."

Misdemeanor prosecutions may also change post-arraignment law enforcement behavior, even with no post-arraignment changes in employer or defendant behavior. If during subsequent criminal justice contact decision makers see previous criminal charges or convictions, they may be more likely to move forward with arrest and/or prosecution.[13]

On the other hand, misdemeanor prosecution may increase "specific deterrence" (Becker, 1968) by increasing the punitiveness of misdemeanor defendants' post-arrest experience. By increasing specific deterrence, misdemeanor prosecution could decrease defendants' likelihood of engaging in post-arraignment behavior that would increase the risk of new complaints, prosecutions, and criminal records. Existing empirical work provides little guidance on the potential specific deterrence impacts of misdemeanor prosecution. There is evidence that sanctions (or more severe sanctions) for driving violations or DUIs decrease subsequent infractions for individuals who experience the sanction (Hansen, 2015; Gehrsitz, 2017; Dusek and Traxler, 2020). But there is also evidence that felony prosecutions increase the likelihood of future felony convictions and decrease future net employment and earnings (Mueller-Smith and Schnepel, 2019).[14]

It is unclear if the downstream effects of misdemeanor prosecution are more likely to be similar to those that follow from driving infractions, or to those that follow from felony prosecution. The net effect of misdemeanor prosecution is an empirical question—one that is being hotly debated in many cities and counties around the United States. So far we know next to nothing about the causal impacts of alternatives to misdemeanor prosecution

[13] As reported in Column (1) of Table 2, for example, individuals against whom nonviolent misdemeanor criminal complaints have been issued in Suffolk County are more likely to be prosecuted if they have records of prior convictions within the past year.

[14] Most work on specific deterrence has focused on incarceration (which also imposes an incapacitation effect), and the findings are mixed. Some have found that incarceration or longer periods of incarceration decrease future crime; others have found increases in future crime (Chen and Shapiro, 2007; Drago, Galbiati and Vertova, 2009; Hjalmarsson, 2009; Di Tella and Schargrodsky, 2013; Aizer and Doyle, 2015; Mueller-Smith, 2015; Bhuller et al., 2020; see Nagin, Cullen and Jonson, 2009, Raphael and Stoll, 2014, Chalfin and McCrary, 2017, and Doleac, 2020 for reviews).

on defendants' criminal justice outcomes.[15]

2.1 Misdemeanor Prosecution in Suffolk County, Massachusetts

In this paper we study the effects of nonviolent misdemeanor prosecution in Suffolk County, Massachusetts (which includes the cities of Boston, Chelsea, Revere, and Winthrop). In Suffolk County, misdemeanor charges are processed in one of nine municipal or district courts.[16] Each of these courts has a geographically defined jurisdiction.

Applications for misdemeanor criminal complaints (typically made by police officers) are brought to the court with geographic jurisdiction for the location at which the alleged offense occurred. A Clerk Magistrate within each court first reviews each application to determine if there is probable cause to issue a criminal complaint. This determination is generally based simply on whether there is a statement of alleged facts attached to the application for complaint; a police report suffices to establish probable cause. After a criminal complaint has been issued, the complaint is assigned an arraignment date and the individual named in the complaint is issued a notification to appear at the arraignment hearing. Within each court, arraignment hearings are scheduled to be heard in designated arraignment courtrooms by the court with jurisdiction over the case, without regard to the identity of the ADA assigned to arraign cases in that arraignment courtroom on that day.

ADAs are assigned to arraignment courtrooms by SCDAO supervisors on a weekly or monthly basis based on availability and experience.[17] Typically more junior ADAs are more likely to be assigned to shifts in the arraignment courtrooms. In our data, ADAs are assigned to arraignment courtrooms in the municipal and district courts for an average of 85 days

[15] A few other studies consider variation in prosecutors' leniency. Rehavi and Starr (2014) and Tuttle (2021) both report evidence that federal prosecutors exhibit racial bias in their prosecution decisions. Sloan (2020*a*) and Sloan (2020*b*) use random assignment of cases to prosecutors in the office of the District Attorney of New York County to document variation in prosecutorial leniency and to test for other-race bias in prosecutors' decisions. All four studies highlight the discretion that prosecutors have about when and how to prosecute defendants, particularly in low-level, nonviolent cases. However, none of these studies estimates the causal impacts of prosecutors' decisions on defendants' subsequent outcomes.

[16] These courts are the Brighton Division, Boston Municipal Court; Central Division, Boston Municipal Court; Charlestown Division, Boston Municipal Court; Chelsea District Court; Dorchester Division, Boston Municipal Court; East Boston Division, Boston Municipal Court; Roxbury Division, Boston Municipal Court; South Boston Division, Boston Municipal Court; and West Roxbury Division, Boston Municipal Court.

[17] There are a few exceptions to this general practice, triggered by specific charge types. Cases with felony charges may receive additional scrutiny from supervising ADAs, strategic assignment to more experienced arraigning ADAs, and/or the involvement of ADAs from the Superior Court. We therefore exclude any cases in which defendants are charged with felony offenses, regardless of the final disposition of those felony charges. We also exclude cases with violent or firearm charges for similar reasons: in these cases more experienced ADAs may be called in to support or handle the arraignment. Our analysis sample thus consists of cases with only nonviolent misdemeanor charges, where the process for assigning arraigning ADAs is as described above.

dispersed across an average of 3.4 years.[18] Arraigning ADAs decide whether to proceed with the prosecution of a misdemeanor complaint. For cases that proceed past the day of arraignment, a different ADA, assigned through a different process, takes over subsequent case stages. All other court actors, such as judges and public defenders, are also assigned to cases through procedures that are independent of the process through which arraigning ADAs are assigned to arraignment courtrooms.

In practice, an arraigning ADA is given a large stack of paper files in the arraignment courtroom on the morning of arraignment, and needs to quickly work through how to proceed in each case. If an arraigning ADA is called away to a meeting during an arraignment session, another ADA may take over the role of arraigning ADA in that courtroom. Unless an arraigning ADA has a conflict of interest in a specific case (e.g., the ADA went to school with the defendant), SCDAO procedures do not allow for a defendant or an attorney to influence which arraigning ADA is overseeing a case, or for arraigning ADAs to select cases they prefer.[19]

During an arraignment hearing, the defendant is officially charged, the criminal complaint is read into the record, and a plea is entered on the defendant's behalf. The arraigning ADA has the discretion to dispose of a complaint prior to or during the arraignment hearing, or to proceed with prosecution. In our data, we are able to observe final charge dispositions and the dates of those dispositions. We define "nonprosecution" to include cases that receive final dispositions on all charges on the day of arraignment, such that no disposition results in a criminal conviction or an "admission to sufficient facts"; "prosecution" includes all other outcomes.[20] Nonprosecution dispositions, at the level of an individual charge, include dismissal/nolle prosequi (observed in 81% of all nonprosecuted cases in our sample) and pretrial probation (observed in 2% of all nonprosecuted cases in our sample).[21] In the case of dismissal or decline to prosecute, the charge is dismissed immediately; in the case of pretrial probation, the charge will be dismissed after successful completion of a period of pretrial probation.[22]

[18]This undercounts the duration of ADA arraignment shifts because the count is based only on the cases for which arraigning ADA information is recorded. Section 6.1 explores a variety of strategies to impute missing ADA information.

[19]When we asked current and former ADAs and supervisors if there is any way for a specific ADA to avoid handling a specific nonviolent misdemeanor case in the absence of a conflict of interest, or for a defendant to avoid a particular arraigning ADA, they emphatically said that was not possible in this context.

[20]Under Massachusetts law defendants may "admit to sufficient facts to warrant a finding of guilty." Admissions have some of the same adverse consequences for defendants as do guilty pleas; we do not include admissions in our definition of nonprosecution.

[21]No disposition details are recorded in 15.9% of nonprosecuted cases in our sample, although for these cases we see that there are no additional case events after the first event, and that no convictions are recorded.

[22]If pretrial probation is not successfully completed, the District Attorney's office may reopen the case. A final disposition of pretrial probation on the day of arraignment, with no further case events, implies that

Finally, Massachusetts statute stipulates that complaints disposed of prior to a defendant's formal arraignment do not become part of a defendant's criminal record, as maintained by the Massachusetts Department of Criminal Justice Information Services (DCJIS). Complaints disposed of at or after arraignment become part of a defendant's criminal record, even if the complaint does not result in a conviction.

3 Data

3.1 Sources and Sample

Our data are sourced from the Suffolk County District Attorney's Office, and include all criminal complaints issued in the county between January 1, 2000 and September 1, 2020. Cases are dated using the date of the first "event" recorded in a case; we refer to that date as the day of arraignment. Defendants are identified with unique IDs. In our main analysis, we follow each defendant in a case for a period of two years following the arraignment hearing. Our sample includes cases with arraignment dates between January 1, 2004 and September 1, 2018; we use data from January 1, 2000 to generate criminal histories, and we follow defendants up to September 1, 2020.[23] Arraigning ADAs are identified as those ADAs recorded as the ADA of record at the arraignment hearing. Cases heard in the nine district/municipal courts are identified by a court location code.

98.5% of charges are identified in the SCDAO data with an offense severity code indicating whether a charge is a misdemeanor, a felony, or a civil violation (e.g., a civil motor vehicle violation). We exclude any case with at least one charge identified as a felony charge. We use text extraction to identify charge types. As described previously, violent offenses may be treated differently during arraignment and thus we exclude cases with any charge for a violent offense—including assault, assault and battery, violating a domestic abuse prevention order, and criminal harassment—and those with any firearms-related charges. We sort the remaining charges into the following categories: motor vehicle, drug, disorder/theft, and

the case was not reopened. In our data, we only observe final dispositions. We do not observe dispositions at arraignment if cases are reopened post-arraignment due to failure to complete probation. We implicitly characterize as "prosecution" any (unobserved) pretrial probation dispositions at arraignment where the case is reopened post-arraignment. Our ADA leniency measure, described below, will code ADAs offering pretrial probation on the day of arraignment as more lenient if more of those cases are not reopened past arraignment (perhaps because of less restrictive probationary conditions). Based on our conversations with SCDAO staff we believe that few charges are resolved with pretrial probation at arraignment; the data we do have indicate that only 2% of the cases in our sample are resolved with pretrial probation on the day of arraignment and are not later reopened.

[23]We analyze one- and three-year followup periods in supplemental results. The one-year followup sample adds in one additional year of criminal cases (to September 1, 2019); the three-year followup sample subtracts one year of criminal cases (to September 1, 2017).

other. We refer to this final set of charges as "nonviolent misdemeanors."

Charges are associated with a variety of different final disposition codes. We characterize final dispositions at the charge level as resulting in a criminal conviction or no conviction. Final dispositions that result in convictions are pleas of guilty and guilty verdicts after bench or jury trials. Final dispositions that do not result in conviction are all other dispositions, including dismissal, pretrial probation, nolle prosequi, admission to sufficient facts, or a finding of not guilty after a jury or bench trial.

SCDAO also requested data from the criminal records database maintained by the Massachusetts DCJIS for the complete set of defendants in the SCDAO case management system. They matched these data to their case data by docket numbers and provided us with information about whether a case record matched to a DCJIS record. Not all cases in the SCDAO case management database are recorded in the DCJIS database. As noted earlier, cases that are disposed of prior to arraignment should not result in DCJIS records.[24] Other cases in the SCDAO case records may not match to a case record in the DCJIS case records because of human error in docket number entry.

3.2 Descriptive Statistics

Our main estimation sample includes cases whose arraignment hearings occur between January 1, 2004 and September 1, 2018; that do not include violent, firearms, or felony charges; that are arraigned in one of Suffolk County's nine district/municipal courts; and for which an ADA is identified at the arraignment hearing.[25] We further restrict our estimation sample to those nonviolent misdemeanor cases overseen at arraignment by an ADA who oversees at least 30 other nonviolent misdemeanor cases at arraignment hearings, and to those cases that are not "singletons" within our set of court-by-time fixed effects (defined below). Table 1 reports descriptive statistics for this sample. There are 67,553 cases in the SCDAO data that meet these criteria. Using our definition of prosecution, 21% of these nonviolent misdemeanor cases are not prosecuted; the remaining 79% are prosecuted. 74% of nonviolent misdemeanor cases that are prosecuted are eventually disposed of without criminal convictions.

Nonviolent misdemeanor cases that are not prosecuted are clearly different from cases

[24] We cannot identify whether a final disposition on the day of arraignment occurs before or after arraignment itself.

[25] Many cases are missing an identified arraigning ADA. Our primary estimation sample is restricted to the 33% of nonviolent misdemeanor cases arraigned between 2004 and 2018 for which an ADA is identified at arraignment. In Section 6.1 we explore the missingness of ADA information, finding that the missingness of arraigning ADA is not closely related to other case and defendant features, and that our findings are robust to several strategies for imputing missing data on arraigning ADAs.

that are prosecuted. Nonviolent misdemeanor defendants who are not prosecuted are issued criminal complaints that include fewer counts, fewer misdemeanor counts, and fewer "serious" misdemeanor counts (punishable by greater than 100 days incarceration). They are less likely to have had a misdemeanor or felony conviction within one year prior to the arraignment hearing in their case. They are more likely to be citizens.[26]

We include the following nonviolent misdemeanor charge types as covariates: motor vehicle charge, drug charge, disorder/theft charge, and other charge. Defendants who are not prosecuted are more likely to have been charged with a motor vehicle offense, and less likely to have been charged with a drug offense or a disorder/theft offense. For the purpose of assessing the monotonicity of our instrument, we also code offense types as "victimless" or "victim" offenses.[27] Defendants who are not prosecuted are more likely to have been charged with a "victimless" offense.

By construction, defendants who are not prosecuted have fewer days to disposition and are less likely to receive convictions, relative to prosecuted defendants. They are also less likely to acquire DCJIS records of their complaint, relative to prosecuted defendants. Defendants who are not prosecuted are then significantly less likely to receive a a new criminal complaint, to be prosecuted, and to receive a new DCJIS record within two years post-arraignment, relative to defendants who are prosecuted. However, because prosecution is clearly correlated with observable pre-treatment characteristics (and likely correlated with unobservable pre-treatment characteristics as well), we cannot draw conclusions about the effect of prosecution on the probability of post-arraignment outcomes from these data alone.

4 Research Design

We want to estimate the effect of misdemeanor nonprosecution on post-arraignment outcomes. Consider the following model, where Y_{ict} is the outcome of interest for individual i in case c in year t, $\mathbf{X}_{ict}$ is a vector of case- and defendant-level covariates, γ_{ct} are court-by-

[26] For some of the defendants in our estimation sample we are able to identify defendant date of birth, race/ethnicity, and gender. However, as we report in Section 6.1, these demographic features are systematically less likely to have been recorded for defendants who were not prosecuted. This problem is particularly acute for missing race/ethnicity information; race/ethnicity was also less likely to have been recorded for defendants who were not rearrested during our time period. In our main analyses we do not include these covariates. We consider robustness to the inclusion of gender and age (which are only slightly missing) and to predicted race/ethnicity in Section 6.1. Our results are robust to the inclusion of these demographic characteristics.

[27] "Victim" offenses include property offenses (e.g., larceny, shoplifting, burglary), threats, property damage, and leaving the scene of property damage or personal injury.

time fixed effects described later, and ε_{ict} is an error term:

$$Y_{ict} = \beta_1 Not\ Prosecuted_{ict} + \beta_2 \mathbf{X}_{it} + \gamma_{ct} + \varepsilon_{ict} \tag{1}$$

β_1 is our parameter of interest. The key problem for causal inference is that ordinary least squares (OLS) estimates of Equation 1 are likely to be biased by the correlation between prosecution and unobserved defendant characteristics that are correlated with the outcomes. This selection bias could be either positive or negative. For example, arraigning ADAs are more likely to prosecute misdemeanor defendants who have prior criminal convictions; defendants with prior convictions are more likely to have subsequent criminal justice contact. As reported in Section 7.1, arraigning ADAs are less likely to prosecute younger defendants; younger defendants are also more likely to have subsequent criminal justice contact. Unobservable characteristics could presumably be causing selection bias in either direction as well.

The as-if random assignment of misdemeanor cases to arraigning ADAs, described above, creates the opportunity to identify a source of variation in nonprosecution that does not depend on defendant or case characteristics. We estimate the causal impacts of misdemeanor nonprosecution by using the propensity of an as-if randomly assigned ADA to not prosecute a defendant as an instrument for nonprosecution. We interpret our causal effects in the local average treatment effect (LATE) framework (Imbens and Angrist, 1994). That is, if the assumptions discussed below hold, we are able to recover the local causal effects of misdemeanor nonprosecution decisions for defendants on the margin of being not prosecuted—those whose treatment status would be changed by switching from a less to a more "lenient" ADA at arraignment.

We construct a residualized leave-out ADA leniency measure for our instrument (French and Song, 2014; Dahl, Kostøl and Mogstad, 2014; Dobbie, Goldin and Yang, 2018). We leave out all cases involving a given defendant arraigned by a given ADA from the construction of the instrument for that defendant's cases, in order to avoid introducing the same estimation errors on the left- and right-hand sides of the regression. We residualize out court and time fixed effects to account for the systematic ways that ADAs in Suffolk County are assigned to misdemeanor cases. As-if randomization of cases to arraigning ADAs occurs within one of the nine municipal or district courts in Suffolk County, and within time periods. For example, multiple ADAs may be assigned to work in the Central Division of the Boston Municipal Court between February and November of 2015, and may rotate arraignment shifts across days of the week. Because misdemeanor case types may vary by court, month, and day of week, a simple leave-out measure of ADA leniency could be confounded by selection. We residualize out court and time fixed effects from our ADA leniency measure to address this potential source of bias.

More specifically, we include court-by-year-month and court-by-day-of-week fixed effects, γ_{ct}, in the construction of our instrument. The inclusion of these court-by-time fixed effects allows us to interpret variation in the instrument as variation in the tendency of an as-if-randomly assigned ADA to prosecute a nonviolent misdemeanor defendant, relative to the other nonviolent misdemeanor cases brought in that court in that month, and in that court on that day of the week.[28] Call this residual nonprosecution decision $Not\ Prosecuted^{*}_{ict}$.

As is standard to avoid the small-sample correlation between the ADA decision in this case and her average leniency, we then construct the leave-out mean measure of ADA nonprosecution (leniency) for each nonviolent misdemeanor case using these residual nonprosecution decisions:

$$\mathrm{Z}_{cta} = \left(\frac{1}{n_a - n_{ia}}\right)\left(\sum_{k=0}^{n_a}(Not\ Prosecuted^{*}_{ikt}) - \sum_{c=0}^{n_{ia}}(Not\ Prosecuted^{*}_{ict})\right) \tag{2}$$

where n_a is the number of nonviolent misdemeanor cases arraigned by ADA a and n_{ia} is the number of nonviolent misdemeanor cases involving defendant i arraigned by ADA a. This construction removes from the instrument the residualized nonprosecution decisions in all of a defendant's nonviolent misdemeanor cases arraigned by ADA a.

Figure 1 reports the distribution of our residualized ADA nonprosecution measure. As noted previously, we restrict the sample to exclude nonviolent misdemeanor cases overseen by arraigning ADAs assigned to fewer than 30 nonviolent misdemeanor cases, and cases that are "singletons" within our set of court-by-time fixed effects. After these restrictions, the sample includes 315 arraigning ADAs. The median number of nonviolent misdemeanor cases overseen by an arraigning ADA is 156 cases; the average is 214 cases. After residualizing out our set of court-by-time effects, the ADA measure ranges from -0.09 at the first percentile to 0.09 at the ninety-ninth percentile. That is, moving from the first to the ninety-ninth percentile of ADA leniency increases the rate of nonprosecution by 18 percentage points, an 86% change from the mean nonprosecution rate of 21 percentage points.

Our main analysis will be based on 2SLS estimates of second-stage Equation 1 (with and without case- and defendant-level covariates) and a first stage for individual i and case c assigned to arraigning ADA a at time t, using a linear probability model:

$$Not\ Prosecuted_{icta} = \alpha_1 Z_{cta} + \alpha_2 \mathbf{X}_{it} + \gamma_{ct} + \varepsilon_{ict} \tag{3}$$

[28]In Table A.15 we also consider a version of the instrument that uses a "raw" measure of ADA leniency based on the non-residualized nonprosecution rate and find similar results.

where, again, γ_{ct} are the court-by-month and court-by-day-of-week fixed effects, and $\mathbf{X}_{ict}$ includes case- and defendant-level covariates (number of counts, number of serious misdemeanor counts, any convictions for previous felonies or misdemeanors in the previous year, offense type, and citizenship). Z_{cta} are the leave-out measures of residualized ADA leniency described previously. Robust standard errors are clustered at both the defendant and ADA level. We also compute and report Anderson-Rubin confidence intervals by inverting the weak-instrument test of Anderson and Rubin (1949) (Andrews, Stock and Sun, 2019, Lee et al., 2020).[29]

Under the LATE assumptions of exogeneity, relevance, and monotonicity, we will estimate a weighted average of the causal impact of nonprosecution among the compliers—defendants whose prosecution decisions could have been different had their cases been (as-if randomly) handled by a different ADA at arraignment. We argue this particular LATE is also a policy relevant treatment effect (PRTE)—it estimates the local effect of policies that increase the leniency of prosecution decisions for marginal defendants (Heckman and Vytlacil, 2001).

In addition, we also estimate marginal treatment effects (MTEs) to explore heterogeneity based on unobservables and to understand the distribution of treatment effects (Björklund and Moffitt (1987), Heckman and Vytlacil (2005), Heckman, Urzua and Vytlacil (2006)). The MTEs are the average effects of nonprosecution for defendants on the margin between being prosecuted and nonprosecuted, where these margins correspond to percentiles of the distribution of the unobserved propensity to be not prosecuted. The MTEs are estimated by taking the derivative of the outcome with respect to the probability of nonprosecution.[30] Other treatment effect parameters can be calculated as different weighted averages of these MTEs if there is full support of the probability of nonprosecution over the interval [0,1] (Heckman and Vytlacil, 2005). Without full support, the weights can be rescaled so that

[29] It is somewhat of an open question how to evaluate the possibility of many-weak-instrument bias in leniency/examiner designs (Hull, 2017; Frandsen, Lefgren and Leslie, 2019; Bhuller et al., 2020). We report the robust first-stage F-statistic, which is large in our setting (Montiel Olea and Pflueger, 2013). Rather than rely on a threshold rule based on this first-stage F-statistic, we also construct Anderson-Rubin confidence intervals, which are of correct size and optimal power even with weak instruments when treating the leniency measure as a single non-constructed instrument (for more on these confidence intervals in over-identified models see Davidson and MacKinnon (2014)). In Appendix C.1 we further explore alternative IV specifications that account for potential biases from the construction of our leniency measure, including: using all the ADA dummies directly as instruments, using lasso to pick the most informative ADA dummies, and using the UJIVE estimation strategy proposed by Kolesár (2013). Across different estimation strategies we robustly find a negative relationship similar in magnitude to our baseline estimate.

[30] Estimating the MTEs requires the same assumptions as the LATE framework, including monotonicity, plus the additional assumption that there is additive separability between the observed and unobserved heterogeneity in the treatment effects, needed when the propensity score does not have full support, as ours does not (see e.g. Brinch, Mogstad and Wiswall, 2017, Mogstad and Torgovitsky, 2018, Andresen, 2018). For further details on the derivation of the MTEs in the potential outcomes framework see the citations in this paragraph, as well as Appendix C.2 and the citations therein.

they integrate to one over the region that does have common support between the treated and untreated (Carneiro, Heckman and Vytlacil, 2011). Using this strategy, we also use the MTE estimates to calculate overall average treatment effects, average treatment on the treated, and average treatment on the untreated.

4.1 Assessing the Instrument

4.1.1 Exogeneity

In order to be able to interpret our 2SLS estimates as the local average treatment effect (LATE) of misdemeanor nonprosecution, it must be the case that defendant and case characteristics do not covary systematically with arraigning ADA assignment. Table 2 reports the results of this randomization test.

First we consider the data without exploiting the as-if random assignment of arraigning ADAs. The first column of Table 2 reports linear probability estimates of the correlation between nonprosecution and case and defendant characteristics, after controlling for court-by-time fixed effects and clustering standard errors at both the defendant and the ADA level. Mirroring what we saw in the summary statistics, even with court-by-time fixed effects we see that the decision to not prosecute a particular defendant is not at all random. We show in Column (1) that defendants are less likely to be not prosecuted if they have more counts in the current case; have more serious misdemeanor counts; had a misdemeanor or felony conviction within the past year; have drug-related, disorder/theft, or other charges; and are not citizens.

In contrast, our measure of ADA leniency is not correlated with these observable characteristics. Column (2) reports estimates of the correlations between ADA leniency and these case and defendant characteristics, using the same specification. Consistent with our understanding that cases are allocated as-if randomly to arraigning ADAs, we find that arraigning ADAs with varying propensities to prosecute handle very similar nonviolent misdemeanor defendants (joint p-value = 0.23).

4.1.2 Instrument Relevance (First Stage)

The validity of our instrument also requires that our measure of ADA leniency be a strong predictor of nonprosecution decisions. Figure 1 plotted a local linear regression of nonprosecution on ADA leniency after controlling for our set of court-by-time fixed effects, for ADA leniency ranging from the first to the 99th percentiles. Nonprosecution is monotonically and approximately linearly increasing in ADA leniency.

Table 3 reports first stage results from Equation 3. Column (1) of Table 3 reports results only with court-by-time fixed effects. Column (2) adds case and defendant covariates: number of counts; number of misdemeanor counts; number of serious misdemeanor counts; whether the defendant had a prior misdemeanor conviction within the past year; whether the defendant had a prior felony conviction within the past year; indicators for whether the defendant is facing charges for a motor vehicle, drug, disorder/theft, or other crime; and defendant citizenship status. Consistent with Figure 1, our residualized ADA instrument is highly predictive of whether a defendant is not prosecuted. The estimated first stage result is robust to the inclusion of controls in Column (2), which is consistent with as-if random arraigning ADA assignment. With all controls, a nonviolent misdemeanor defendant assigned to an arraigning ADA who is 10 percentage points more likely to not prosecute a defendant is approximately 5.5 percentage points more likely to be not prosecuted.[31]

4.1.3 Exclusion Restriction

We also need it to be the case that arraigning ADA assignment only affects defendant outcomes through the probability of nonprosecution at arraignment. The exclusion restriction would be violated if ADA assignment were correlated with case and defendant characteristics that are also correlated with outcomes. As described above, Table 2 shows that assignment of cases to arraigning ADAs appears random after we condition on our court-by-time fixed effects.

The exclusion restriction would also be violated if arraigning ADAs affected future outcomes through channels other than the decision to prosecute or not prosecute. We cannot directly test the assumption that ADAs only systematically affect defendant outcomes through the prosecution decision. However, we believe that the exclusion restriction assumption is reasonable in our setting. Recall that after arraignment a different ADA, assigned through a different process, takes over the subsequent case stages. All other court actors, such as judges and public defenders, are also attached to cases through a different process. These institutional characteristics make it unlikely that the assignment of an arraigning ADA is correlated with post-arraignment actions that could independently affect defendant outcomes.

We first partially explore potential threats to the exclusion restriction by looking at the effect of arraigning ADA assignment on case outcomes other than through the initial decision to prosecute/not prosecute. We test whether our ADA leniency measure is predictive of the number of days to case disposition or the likelihood of a criminal conviction in a misdemeanor

[31]This table also reports robust first-stage F-statistics, which in the just-identified case are equivalent to the effective F-statistic of Montiel Olea and Pflueger (2013). Both of these F-statistics exceed the critical value of 23.11 they propose for just-identified models with τ =10% of worst case bias.

defendant's case within the sample of defendants who are prosecuted. If arraigning ADAs only have power over the prosecution decision, their leniency should not affect these outcomes when we restrict attention to prosecuted defendants.

Table A.1 reports these results. Consistent with the exclusion restriction, we find that our preferred leave-out instrument is not predictive of either days to disposition, or a criminal conviction in a misdemeanor case, conditional on a defendant being prosecuted. This supports our hypothesis that our main results operate through arraigning ADAs' effects on the initial decision to prosecute/not prosecute, not through other channels.

One other decision the arraigning ADA may make at the arraignment hearing is whether to request bail in cases that they decide to prosecute. Because these are relatively minor offenses, bail is typically not requested: the arraigning ADA requests bail in only 8% of cases that they choose to prosecute, and bail is set by the judge in only 6.6% of prosecuted cases.[32] In Section 5.3 we address the issue of bail requests in several ways and show that our main results still hold. Section 5.3 also reports reduced form estimates of the effects of our leave-out instrument on outcomes. If the exclusion restriction is violated, our reduced form estimates can still be interpreted as the causal effects of being assigned to a more or less lenient arraigning ADA.

4.1.4 Monotonicity

With constant treatment effects, the above assumptions on exclusion and randomization would be sufficient to recover causal effects. Under heterogeneous treatment effects, in order to recover the LATE—the causal impact for the compliers—we also need it to be the case that the impact of ADA assignment on the probability of nonprosecution is monotonic across defendants. This monotonicity assumption implies that defendants who are not prosecuted by stricter ADAs would also be not prosecuted by more lenient ADAs, and that defendants prosecuted by more lenient ADAs would also be prosecuted by stricter ADAs.

We cannot test this assumption directly, but there are several indirect tests we can pursue. Frandsen, Lefgren and Leslie (2019) provide a test for the joint null hypothesis that the exclusion and monotonicity assumptions hold. We calculate this test within the nine courts in our dataset, controlling for our main set of covariates and year-month and day-of-week fixed effects. In Table A.2 we show that within six of the nine courts in our data we fail to reject the joint null hypothesis that exclusion and monotonicity hold.

[32] Using data from Massachusetts Trial Courts and DCJIS, Bishop et al. (2020) show that across all cases—including violent misdemeanors and felonies—bail is imposed in 11-17% of arraignment hearings depending on the race of the defendant. Given that our analysis focuses on nonviolent misdemeanors, the low rates of bail assigned at arraignment we observe appear consistent with statewide data.

Frandsen, Lefgren and Leslie (2019) also show that one can relax the strict (pair-wise) monotonicity assumption of the original LATE framework to an average monotonicity assumption and still recover a weighted average of individual treatment effects. That is, average monotonicity is sufficient to interpret our 2SLS estimates as the causal effects of nonprosecution. This average monotonicity assumption implies that the covariance between a defendant's prosecutor-specific treatment and the prosecutor's overall propensity to not prosecute is weakly positive. One test of this assumption is that prosecutors' group-specific nonprosecution rates should be positively correlated with overall nonprosecution—prosecutors that are more lenient overall should be more likely to not prosecute people in any observable subgroup. This is equivalent to showing that the first stage should be positive in all subsamples of the data, as is common in the literature (Dobbie, Goldin and Yang, 2018; Bhuller et al., 2020).[33]

Table A.3 presents first stage results for a large variety of subsamples of our data: separately by number of counts, number of misdemeanor counts, number of serious misdemeanor counts, whether the defendant was convicted of a misdemeanor charge in the prior year, whether the defendant was convicted of a felony charge in the prior year, whether the defendant was a citizen, charge type, and victim/victimless offense, using the full sample of cases to calculate our measure of ADA leniency. Consistent with the average monotonicity assumption, we find that the relationship between our residualized measure of ADA leniency and nonprosecution is positive and significant in all subsamples.[34] In specification checks in Section 5.3, we also create versions of our instrument that are interacted with various ADA and case characteristics to relax these monotonicity assumptions.

5 Results

Tables 4 and 5 present OLS and 2SLS estimates of the impacts of misdemeanor nonprosecution on post-arraignment outcomes. We start with our main outcome of interest in Table 4:

[33] de Chaisemartin (2017) offers another way of relaxing the strict monotonicity assumption. Under a weaker condition he calls the "compliers-defiers" (CD) condition—for any pair of ADAs, there is a subset of compliers that is the same size as the subset of defiers (defendants who violate monotonicity for this pair) and that has the same local average treatment effect as the defiers—the 2SLS estimates are the LATE for the subgroup of the remaining compliers. The CD condition holds if the treatment effect has the same sign for both compliers and defiers, and if the treatment effect for compliers is greater than the treatment effect for defiers. We do not have strong reasons to believe that compliers and defiers would have differently signed treatment effects. This weaker "compliers-defiers" condition is also tested by the joint monotonicity-exclusion test of Frandsen, Lefgren and Leslie (2019), which we fail to reject across a large share of our sample.

[34] We show the equivalent first stage monotonicity checks for subsamples based on gender, age, and race/ethnicity in Section 6.1; these similarly all show positive and statistically significant relationships between ADA leniency and nonprosecution.

the likelihood of a subsequent criminal complaint within two years post-arraignment. We separately consider the likelihood of any subsequent complaint, the likelihood of subsequent misdemeanor charges, and the likelihood of subsequent felony charges. OLS estimates with controls in Column (2) imply that nonprosecution reduces the probability of a subsequent criminal complaint by 10 percentage points (a 27% decrease relative to the mean for prosecuted defendants). The 2SLS estimates with controls in Column (4) indicate that marginal nonprosecuted misdemeanor defendants are 33 percentage points less likely to receive a new complaint within two years ($p < 0.01$). This represents a 58% decrease relative to the mean for complier defendants who were prosecuted.[35] Nonprosecution in the initial case reduces the likelihood of a subsequent misdemeanor complaint by 24 percentage points (60%; $p < 0.01$), and of a subsequent felony complaint by 8 percentage points (47%; not significant). Reduced form estimates are presented in Table A.14. These reduced form estimates are also large, negative, and statistically significant.

Figure 2 shows how this effect evolves over the two-year followup period in three month bins. We see an immediate drop in the likelihood of a new complaint within the first three months after the arraignment hearing, and that effect remains steady through the first year. At that point the negative effect begins to grow larger over time.

Table A.4 shows similar results for the number of subsequent criminal complaints. With all controls (Column (4)), the 2SLS estimates indicate that nonprosecution reduces the number of subsequent criminal complaints by 2.1 complaints (69%; $p < 0.01$). The number of subsequent misdemeanor complaints is reduced by 1.2 complaints (67%; $p < 0.01$), and the number of subsequent felony complaints by 0.7 complaints (75%; $p < 0.05$).

Table A.5 reports 2SLS estimates with all controls for complaints within two years by subsequent crime type (violent, motor vehicle, disorder/theft, drug, and other). We find significant declines in all types of subsequent complaints except for drug and other charges. Nonprosecution reduces the rates at which nonviolent misdemeanor defendants are charged with subsequent violent offenses by 64%, with subsequent disorder/property offenses by 91%, and with subsequent motor vehicle offenses by 63%.

A subsequent complaint might not be consequential if it is not prosecuted or if it does not result in a criminal record. For this reason we also consider effects on future prosecutions and DCJIS records. Table 5 reports OLS and 2SLS estimates of the impacts of misdemeanor nonprosecution on the likelihoods of subsequent prosecution and subsequent acquisition of a DCJIS criminal record within a two-year interval post-arraignment. We see that nonprosecu-

[35] See Appendix C.3 for details on the calculation of average outcomes among prosecuted compliers. See Section 7.1 for more on the comparisons between OLS and 2SLS. Mueller-Smith and Schnepel (2019) find that felony deferred adjudication reduces subsequent convictions by 45%, a similar order of magnitude.

tion meaningfully decreases the probability that a marginal defendant will be prosecuted or receive a future DCJIS record. With the full set of controls (Columns (2) and (4)), the 2SLS estimates indicate that the marginal nonprosecuted misdemeanor defendant is 35 percentage points less likely to be prosecuted within two years post-arraignment (66%; $p < 0.01$), and 40 percentage points less likely to acquire a criminal record of a new complaint within two years post-arraignment (69%; $p < 0.01$).

Our main results use a two-year followup period for all defendants when considering the effects of nonprosecution on subsequent criminal justice contact. We also consider one-year and three-year followup periods. Our analysis sample becomes a bit larger when we look at a one-year followup window (N = 74,631); it becomes a bit smaller when we look at a three-year followup window (N = 63,655). Tables A.6 and A.9 show that our evidence for the randomization of ADA leniency across cases still holds in these samples; and Tables A.7 and A.10 show that our first stage results are strong and very similar in these samples. Our main results are replicated in these samples in Tables A.8 and A.11. Results are a bit smaller and less statistically significant with a one-year followup window, in line with what we saw in Figure 2. The results are a bit larger and more significant for all outcomes with a three-year followup window. Figure A.1 shows how these results evolve over time. As in Figure 2, we see a reduction in criminal complaints that holds steady through the first year, and then begins to grow over time.

5.1 Mechanisms

We consider three mechanisms that may be driving our findings. The first is that nonprosecution eliminates the possibility that defendants will spend a lengthy period of time in the criminal justice system with an open case. As reported in Table 1, cases for prosecuted nonviolent misdemeanor defendants in our sample take on average 185 days or approximately 6 months to resolve. Time spent in the criminal justice system—attending hearings and meetings with lawyers, for instance—can disrupt defendants' work lives, increasing the risk of reoffending.

The second mechanism that may be driving our findings is that nonprosecution eliminates the possibility that a defendant will receive a conviction in their case. As reported in Table 1, 26% of prosecuted nonviolent misdemeanor cases in our sample result in a conviction. Criminal records of misdemeanor convictions may damage defendants' labor market prospects, raising the risk of reoffending, and may increase the likelihood of future prosecution and conviction, conditional on future arrest.

The third mechanism that may be driving our findings is that nonprosecution reduces the probability that a defendant will receive a DCJIS criminal record of their misdemeanor arrest, and thus reduces the probability that law enforcement officers and employers will see a record of this arrest. Cases dismissed prior to formal arraignment do not receive DCJIS records in Massachusetts. Cases that proceed to formal arraignment do receive DCJIS records. In our data we can only identify day of final disposition, not whether a final disposition that occurs on the day of arraignment occurs before or after formal arraignment. We expect, however, that cases that are not prosecuted under our definition will have significantly lower rates of DCJIS record acquisition, relative to cases that are prosecuted. Table A.12 reports OLS and 2SLS estimates of the impacts of misdemeanor nonprosecution on the likelihood that a defendant receives a DCJIS criminal record in the case. With the full set of controls (Column (4)), the 2SLS estimates indicate that the marginal nonprosecuted misdemeanor defendant is 55 percentage points less likely to receive a misdemeanor complaint record in the DCJIS database (56%, $p < 0.01$). Prosecuted defendants who acquire DCJIS records of their criminal complaints may have decreased labor market prospects, increasing the risk of reoffending, and increased likelihood of future prosecution and conviction, conditional on future arrest.

5.2 Heterogeneity

Misdemeanor prosecution may have such negative effects for marginal nonviolent defendants because it pulls some defendants into the criminal justice system who otherwise would remain outside that system. If this is the case, then we would expect to see larger effects for first-time defendants.

In Table 6 we repeat our main analyses separately for first-time and repeat defendants, defined in three different ways. In Columns (1)-(2), we split defendants based on whether they had any previous complaints filed against them in Suffolk County. In Columns (3)-(4), we split defendants based on whether they had any previous complaints in Suffolk County that resulted in a DCJIS record. And in Columns (5)-(6), we split defendants based on whether they had any previous complaints in Suffolk County that resulted in a conviction.

We see suggestive evidence that first-time defendants realize greater benefits from nonprosecution than repeat defendants. The effects of nonprosecution on subsequent criminal justice contact are generally larger and more precisely estimated for first-time defendants. Because effects are imprecisely estimated for repeat defendants, we cannot generally reject the null hypothesis that the impacts for the two groups are the same. But the point estimates imply that marginal nonprosecuted defendants without previous SCDAO complaints are 80% less likely to receive a new criminal complaint within two years, relative to prosecuted compliers ($p < .01$), while the risk of subsequent criminal complaints for nonprosecuted repeat defendants is reduced by only 16%, relative to prosecuted compliers ($p > .10$).

In Section 6.1 below we discuss heterogeneity across demographic groups. We find no meaningful differences by gender or race/ethnicity, but suggestive evidence of differences by age group. In Section 8.1 we describe our marginal treatment effect estimates, which show heterogeneity by the (unobservable) probability of nonprosecution.

5.3 Specification Checks

In this section we pursue a variety of modifications to our primary specifications. First, as discussed earlier, ADAs can make bail requests at the arraignment hearing (although for the nonviolent misdemeanor cases in our sample bail is requested in only 8% of cases). We might worry that our leniency measure confounds two types of leniency: "nonprosecution leniency" and "no-bail leniency." In Table A.13 we address this in three ways, showing the 2SLS estimate for the effect on subsequent criminal complaints in each case. First, we create a "no-bail leniency" measure based on ADAs' propensity to request bail in other defendants' cases, and simply control for it in our regressions. Our results are nearly identical. Second, we use our no-bail leniency measure as an instrument for not receiving bail, and estimate the effect on subsequent complaints. We find a negative coefficient (which could be due to the correlation of the bail decision with the nonprosecution decision) but it is insignificant. Third, we use both nonprosecution leniency and no-bail leniency as instruments in the same regression, to measure the separate effects of the nonprosecution and no-bail decisions. Our estimate for nonprosecution is nearly identical to our main estimate, and the estimated effect of no-bail is near-zero and statistically insignificant. Based on these results, we conclude that arraigning ADAs' decisions about whether to request bail do not explain our results. These analyses again support our hypothesis that the only meaningful channel through which arraigning ADAs affect defendants' outcomes is through the decision of whether to prosecute the case.

Table A.14 reports reduced form estimates of the effects of our leave-out instrument on post-arraignment outcomes. If the exclusion restriction is violated, our reduced form estimates can still be interpreted as the causal effects of being assigned to a more or less lenient arraigning ADA. These reduced form estimates are very similar to the 2SLS estimates reported above, consistent with the strong first-stage relationship between the propensity of an arraigning ADA to not prosecute and a defendant's own arraignment outcome.

Table A.15 reports 2SLS estimates for receiving a DCJIS record in the initial case, and for subsequent complaints, prosecutions, and DCJIS records, for different versions of our instrument. Column (2) reports estimates using a version of our leave-out mean instrument that does not residualize out court-by-time fixed effects. This instrument is thus a raw measure of an ADA's leave-out nonprosecution rate. Results are slightly larger for most outcomes and remain significant between $p < 0.01$ and $p < 0.05$. Columns (3)-(5) report estimates for more flexible instruments constructed by interacting our main leave-out instrument with various ADA or case characteristics. This relaxes our monotonicity assumption and allows the effect of ADA leniency to vary with each of the following: (i) high versus low ADA experience (as measured by above- or below-median number of nonviolent misdemeanors arraigned as of the time of this case's arraignment), (ii) whether the crime is categorized as victimless, or (iii) several mutually-exclusive crime types. In all cases estimates are qualitatively similar to the main estimates presented above; coefficients maintain the same sign and are of similar magnitudes. With the ADA-by-crime-type instrument the impact on subsequent misdemeanor arrests is no longer statistically significant, although the overall effect is, and the impact on subsequent felony arrests gains significance.

In Appendix C.1 we further explore alternative IV specifications that account for potential biases from the construction of our leniency measure: including using all the ADA dummies directly as instruments, using lasso to pick the most informative ADA dummies, and using the UJIVE estimation strategy proposed by Kolesár (2013). Across different estimations strategies we robustly find a negative relationship similar in magnitude to our baseline estimate.

6 Missing Data

6.1 Missing Demographic Data

Both the institutional details of how cases are assigned to arraigning ADAs, as well as the data we observe, give us confidence that ADA leniency is not correlated with any defendant or case characteristics. However, as noted above, we do not observe age, gender, or race/ethnicity for all defendants in our sample. These demographic characteristics are more likely to be missing for cases that were not prosecuted, likely because SCDAO staff are time-constrained and may have deemed it less important to enter this information when cases were not moving forward. This selective missingness makes it problematic to include these variables in our main analysis; doing so could introduce bias.

This problem is less acute for gender and age. Only 1.4$ of cases are missing either gender or age (though 2.5% are missing either among nonprosecuted cases and 1.1% are missing either among prosecuted cases). However, 16.1% of cases are missing race/ethnicity information, with 27% of nonprosecuted cases missing race/ethnicity and 13% of prosecuted cases missing race/ethnicity. Moreover, as reported in Table A.16, the missingness of information on race/ethnicity is correlated not only with whether a case is prosecuted, but also with whether a defendant has subsequent criminal justice involvement. Defendants have at least one subsequent criminal complaint within two years post-arraignment in 26% of cases that are not prosecuted and not missing race/ethnicity information, but in only 10% of cases that are not prosecuted and missing race/ethnicity; and in 40% of cases that are prosecuted and not missing race/ethnicity information, but in only 18% of cases that are prosecuted and missing race/ethnicity.

The correlation of missing data on defendant race/ethnicity and rates of subsequent criminal justice contact is likely due to the way that defendant data are entered and stored in SCDAO electronic records. When a new case is entered into SCDAO's case management system, the administrator entering the case first searches for the defendant's name in the case database. If the administrator finds the defendant's name (possibly after further narrowing his selection by date of birth, social security number, and/or address), he selects the name to start a new case record. Any defendant demographic information already entered in the database will be auto-populated in the new case record. Any missing demographic information can be filled in, and the new information will be stored as part of the defendant's record. The likelihood that a defendant has race/ethnicity information associated with his case records is thus an increasing function of the number of times he is processed through SCDAO.

However, given the importance of age, gender, and race/ethnicity in predicting current case outcomes and future criminal justice involvement, we use the data we do have to (1) confirm the robustness of our analyses above to the inclusion of demographic characteristics, and (2) to consider the heterogeneity of our results across demographic groups. Because of the large and selective missingness of race/ethnicity, we construct an alternative measure using defendants' names to compute the probabilities that an individual is Black, white, and/or Hispanic.[36] These probability estimates will measure actual race/ethnicity with error, but the error will not be correlated with whether or not a case was prosecuted.

Table A.17 reports summary statistics including these demographic variables for the sample of cases not missing gender or age. 80% of defendants are male. We sort defendants into age group buckets to allow for measurement error in precise age. 23.1% of defendants are less than 24 years old; 24.6% are 24-30; 21.8% are 31-40; and 30.6% are over age 40.[37] In

[36] Specifically, we use Bayesian Improved Surname Geocoding (BISG) relying on the Census Bureau's surname list via the R package WRU. This is the same algorithm used by the Consumer Financial Protection Bureau for race/ethnicity prediction.

[37] Note these cutoffs represent the 25th, 50th, and 75th percentiles of age among those for whom we have

the smaller sample for which race information is recorded in the SCDAO database, 45.8% of defendants are coded as Black, 36.3% as white, and 16% as Hispanic. In the larger sample and using predicted race/ethnicity, 34.5% of defendants are most likely to be Black, 25.6% are most likely to be white, and 33.2% are most likely to be Hispanic.

Table A.18 shows the results of our randomization test when gender, age groups, and predicted race/ethnicity are included. Column (1) uses nonprosecution as an outcome, where we see that younger (age < 24 is the base group) and female defendants are more likely to be not prosecuted. When we use ADA leniency as the outcome, the p-value on the joint F-test of the significance of all the coefficients is 0.17. The only coefficient that is significant is that for male defendants, although the coefficient is small and we might expect one significant coefficient out of 17 coefficients. We also know that gender is one of the demographic characteristics that is selectively missing, so a significant difference is not surprising. Notably, the coefficients on age groups (which are also systematically missing) and predicted race/ethnicity are all near-zero and statistically insignificant. Table A.19 checks for monotonicity within each of these demographic groups. We sort defendants into race/ethnicity categories based on their highest predicted-race probabilities. The first stage is positive and significant for all subgroups. Table A.20 shows the first stage estimates when gender, age groups, and predicted race/ethnicity are included in the set of case and defendant covariates. Results are nearly identical to those in our main analysis.

Table A.21 shows our main results (effect of nonprosecution on the likelihood of a subsequent criminal complaint) with these demographic characteristics included as covariates. Our results are again nearly identical to those we described above. Of note, in the SCDAO administrative data approximately 46% of nonviolent misdemeanor defendants are Black. Only approximately 24% of Suffolk County residents are Black (U.S. Census Bureau, 2019 Population Estimates Program). Reducing the prosecution of nonviolent misdemeanor offenses would disproportionately benefit Black residents of the county.

We interpret these results as further support for our assumption of as-if random assignment of ADAs (and for the claim that ADA leniency is uncorrelated with case and defendant characteristics). While our results are robust to the inclusion of demographic characteristics, the selective missingness of these data leads us to prefer the specifications that do not include these variables. We thus highlight the main results presented earlier.

Table A.22 considers heterogeneity of the nonprosecution effect across different demographic groups. There is no difference by gender: effects for men and women are nearly identical. Estimated effects are larger for older defendants (24+) and smaller for defendants less than 24 years old. Effect sizes are similar across predicted race groups.

age.

6.2 Missing ADA Data

As noted earlier, 67% of cases meeting all our other sample criteria are missing information on the identity of the arraigning ADA. As reported in Table A.23, arraigning ADA information is missing in 64% of cases that are prosecuted, and in 75% of cases that are not prosecuted. However, unlike the missingness of race/ethnicity information, missingness of ADA information is not also strongly correlated with defendants' subsequent criminal justice contact, or with other case/defendant characteristics. The rates at which defendants have subsequent criminal justice contact are nearly identical for those with and without arraigning ADA information recorded, for both prosecuted and not prosecuted defendants.

The relative lack of correlation between missing ADA information and case/defendant characteristics is likely due to the way that ADA information is entered and stored in the SCDAO case management database. In contrast to the way that defendant race/ethnicity information is stored in SCDAO records and is auto-populated when a defendant receives a new criminal complaint, ADA information must be entered anew for each case. The missingness of ADA information is thus likely more idiosyncratic than the missingness of race/ethnicity information.

Nonetheless, the missingness of ADA information does appear to be correlated with prosecution/nonprosecution. To address the possible bias introduced by this selective missingness, we construct four alternative measures of imputed ADA assignment based on progressively more expansive criteria. Our first imputed measure of ADA assignment ("Imputation 1") identifies a) court/days with only one observed arraigning ADA, where that ADA either arraigns at least two cases or arraigns one case and there is only one other case missing ADA information, and assigns that ADA to any other cases on that court/day that are missing ADA information; and b) court/weeks with only one observed ADA, where that ADA arraigns at least four cases in total on at least two days, and assigns that ADA to any other cases in that court/week that are missing ADA information. Our second imputed measure of ADA assignment ("Imputation 2") additionally identifies a) court/days with only one observed arraigning ADA, and assigns that ADA to any other cases on that court/day that are missing ADA information; and b) court/weeks with only one observed ADA, and assigns that ADA to any other cases in that court/week that are missing ADA information. Our third imputed measure of ADA assignment ("Imputation 3") additionally identifies court/days with multiple observed arraigning ADAs but with a single modal arraigning ADA, and assigns that ADA to any other cases on that court/day that are missing ADA information. Our fourth imputed measure of ADA assignment ("Imputation 4") additionally identifies court/weeks with multiple observed arraigning ADAs but with a single modal arraigning ADA, and assigns that ADA to any other cases in that court/week that are missing ADA information.

Table A.24 reports the proportions of our main estimation sample and the four imputation samples that are missing arraigning ADA information. The imbalance on missingness of arraigning ADA information for prosecuted and not prosecuted cases progressively decreases as the imputation samples grow larger. In our largest imputed sample ("Imputation 4", containing 149,185 observations or 76.4% of the sample of nonviolent misdemeanor cases meeting all other sample criteria), 24% of prosecuted cases are missing arraigning ADA information, and 22% of not prosecuted cases are missing arraigning ADA information.

Table A.25 reports two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within two years, for alternative samples of cases for which ADA assignment has been imputed. All models instrument for nonprosecution using our main ADA leniency measure, estimated using only cases assigned to an observed arraigning ADA, and include all covariates included in Table 4. Because we are imputing ADA assignment by week, we implement court-by-week fixed effects instead of court-by-month fixed effects. Table A.25 also reports p-values from regressions that mimic our randomization tests in Table 2 Column (2) and first stage results that mimic Table 3 Column (2). Column (1) reports estimates for our primary estimation sample with court-by-week fixed effects. Columns (2) - (5) progressively expand the sample of cases to include those cases where arraigning ADA assignment has been imputed ("Imputation 1" - "Imputation 4"). Randomization p-values for the imputation samples range between 0.31 and 0.62. First stage coefficients are all large and positive, with F statistics ranging between 14 and 19. Second stage estimates for the imputation samples range between -0.50 and -0.65 and are all significant at $p < .01$. None of the second stage Anderson-Rubin confidence intervals contain zero.

The results from the imputation samples suggest that our primary estimates are not being driven by selection bias from missing arraigning ADA information. Our estimates are robust even in our most aggressively constructed imputation sample, containing 76.4% of the cases meeting all other sample criteria.

7 Understanding the LATE

Our 2SLS estimates represent the LATE for marginal defendants—defendants who would have received a different prosecution decision had their case been assigned to a different arraigning ADA. To better understand this LATE, we characterize the number of compliers and their characteristics following the approach developed by Abadie (2003) and Dahl, Kostøl and Mogstad (2014), and applied by Dobbie, Goldin and Yang (2018) and Bhuller et al.

(2020). Details of these calculations can be found in Appendix C.3.

In Table A.26 we estimate these shares in several ways. We define most and least lenient ADAs by their percentiles in the residualized leniency distribution, defining the least lenient ADAs as those at the ρ percentile and the most lenient as those at the $(100-\rho)$ percentile, where ρ varies between 1, 1.5, and 2. In the first three columns of Table A.26, we use a linear specification of the first stage, given by Equation 3. Under this linear specification, we find that ten percent of our sample are compliers, 73 percent are never takers, and 18 percent are always takers. The latter three columns use a local linear version of our first stage of nonprosecution on the residualized measure of ADA leniency, controlling for court-by-time fixed effects. Under this more flexible analog to our first stage equation, we find that nine percent of our sample are compliers, 72 percent are never-takers, and 18 percent are always-takers.

We then use a similar insight to describe observable characteristics of compliers by calculating the fraction of compliers in different subsamples (Abadie, 2003).[38] Table A.27 shows these results for various observable characteristics: in Column (1) we show the proportion of our sample represented by this subset of observable characteristics; in Column (2) we show the estimated proportion of this subsample composed of compliers; in Column (3) we show the ratio of how often the trait occurs in the estimated complier group, relative to the full sample. Compliers look similar to the full sample on many dimensions, but differ on others. In particular, compliers are less likely to have been charged with a drug offense, to have been charged with a serious misdemeanor (punishable by more than 100 days in jail), to have misdemeanor or felony convictions within the prior year, and to be noncitizens. We also estimate complier shares within demographic groups for the sample of data for which we have data on age and gender (see Section 6.1). Compliers are more likely to be younger (less than 24 years old) and also more likely to be female.

7.1 Comparing OLS and 2SLS Estimates

Our main OLS estimates are smaller in absolute value than our 2SLS estimates. OLS and 2SLS estimates can differ due to heterogeneity in the effect of nonprosecution on subsequent criminal justice contact for the compliers and/or due to selection bias. To explore possible heterogeneity, in Table A.28 we reweight our OLS estimates to match the sample of compliers using two different reweighting schemes (Dahl, Kostøl and Mogstad, 2014; Bhuller et al., 2020). Columns (1)-(3) use our main covariates; Columns (4)-(6) also include age, gender,

[38] For these calculations we use the linear specification of the first stage and a $\rho = 1\%$ cutoff to define most and least lenient ADAs.

and predicted race (and thus are estimated in the sample not missing age or gender).[39] We see that the reweighted OLS estimates are very similar to the unweighted OLS estimates under both reweighting schemes, implying that the differences between the OLS and 2SLS estimates are unlikely to be accounted for by heterogeneity in causal effects for compliers by observable characteristics (we cannot rule out heterogeneity on unobservable characteristics).

The differences we see, then, are likely driven by selection bias: arraigning ADAs are, on average, choosing to not prosecute defendants who have higher risk of subsequent criminal justice contact than marginal defendants. There are a variety of characteristics that ADAs might interpret as mitigating circumstances making defendants less culpable of their crimes or more worthy of a second chance, but that also increase the risk of subsequent criminal justice contact. Previously, we hypothesized that age could induce such negative selection. As reported in Table A.18 Column (1), older defendants are much less likely to be not prosecuted than defendants age 23 or younger (the base group in that regression); that is, younger defendants are less likely to be prosecuted. Arraigning ADAs may view younger defendants as less culpable or want to give them a second chance. However, younger defendants are at significantly higher risk of future criminal justice contact than are older defendants.[40] If we compare OLS and 2SLS estimates from Table A.22, which reports 2SLS and OLS estimates by demographic groups, we see that the point estimates for 2SLS and OLS are much closer within the age category of 18-23 years than for the overall sample, and that we cannot statistically distinguish between the two point estimates. This implies less room for selection bias within this age group. Given the larger differences between OLS and 2SLS for older defendants, other unobservable defendant characteristics are likely to be driving selection bias for other age groups. Because selection bias based on unobservable characteristics appears to be important in this context, we focus on our 2SLS estimates.

8 Policy Relevance and Moving beyond the LATE

Our 2SLS estimates give us a weighted average of the effect of nonprosecution among those defendants induced into nonprosecution by being (as-if randomly) assigned a more lenient ADA at arraignment—the LATE. The decision to prosecute or not prosecute a defendant against whom a criminal complaint has been issued is a decision that rests squarely with

[39] More details on these reweighting schemes can be found in the notes to Table A.28.

[40] It is true generally that younger people are more likely to have criminal justice contact, relative to older people (Laub and Sampson, 2001; Landersø, Nielsen and Simonsen, 2017). It is also true in the context of our data: in an unreported regression, defendants age 23 or younger are 3.5 pp (10%) more likely to have a new criminal complaint in the next two years than those age 24-30; 6 pp (20%) more likely than those age 31-40, and 8.8 pp (26%) more likely than those older than 40 (all comparisons $p < 0.01$).

the office of the District Attorney in that jurisdiction. Conditional on the set of behaviors that are considered criminal, and the behavior of police in arresting individuals suspected of committing those crimes, the only policy lever available to change nonprosecution rates is to increase the leniency of the individuals within a District Attorney's office who make the prosecution decision. This LATE estimate is thus also a policy-relevant treatment effect (PRTE) (Heckman and Vytlacil, 2001; Heckman and Urzua, 2010; Cornelissen et al., 2016).

As we increase leniency, we would presumably be drawing different marginal defendants into nonprosecution. Our LATE estimates do not tell us directly how these marginal defendants may differ in their treatment effects from defendants likely to be on the margin of prosecution for less lenient ADAs. We explore this question in two ways. First, we estimate marginal treatment effects (MTEs). These MTEs give insight into what might happen if we implemented a policy that increased ADA leniency. However, we cannot extrapolate beyond the data that we have: the MTEs are only estimated for predicted probabilities of nonprosecution for which we see both prosecuted and nonprosecuted individuals—the common support of the propensity score for nonprosecution.

Second, we consider the effects of a policy change. Several district attorney's offices around the country have begun to implement policies of presumptive nonprosecution for certain (usually low-level/misdemeanor) offenses. To the extent that such policies still allow room for ADA discretion, the presumption of nonprosecution may be applied largely to marginal nonviolent misdemeanor defendants, and may have similar effects as those estimated above. However, to the extent that the policies expand the set of marginal defendants beyond those in our sample, and/or are applied to non-marginal defendants, the policies may have different effects. We use the inauguration of Rachael Rollins as District Attorney of Suffolk County on January 2, 2019 as a natural experiment to explore such policy effects. During her 2018 election campaign, then-candidate Rollins pledged to establish a presumption of nonprosecution for a set of defined offenses. We use the transition to the Rollins administration to explore how increases in nonprosecution induced by the arrival of a new District Attorney impacted both subsequent criminal complaints and the rate of crimes reported to the Boston Police Department.

8.1 Marginal Treatment Effects

Because defendants are (as-if) randomly assigned to a large number of ADAs with different leniency rates, we can trace out the effects of nonprosecution along different margins by estimating marginal treatment effects (MTEs). These margins correspond to percentiles of the distribution of the unobserved propensity to be not prosecuted (Heckman and Vytlacil, 2005;

Heckman, Urzua and Vytlacil, 2006; See Appendix C.2 for more details on the derivation of the MTE in the potential outcomes framework).[41]

Marginal treatment effects are given by the derivative of the probability of a criminal complaint within two years with respect to the predicted probability of nonprosecution—the propensity score estimated using the leniency of the arraignment ADA as well as covariates and court-by-time fixed effects.[42] For a particular realization of the predicted probability of nonprosecution, the MTE captures the mean effect of nonprosecution on future criminal complaints for those who would not be prosecuted if assigned an ADA with a slightly higher leniency, and prosecuted if assigned to an ADA with a slightly lower leniency. The predicted probability of nonprosecution is increasing in ADA leniency. At the lowest levels of the predicted probability of nonprosecution, the MTEs report the effects of nonprosecution for defendants who would be not prosecuted by all but the least-lenient ADAs (i.e., defendants who are closer to the always-takers). At the highest levels of the predicted probability of nonprosecution, the MTEs report the effects of nonprosecution for defendants who would be not prosecuted by only more-lenient ADAs (i.e., defendants closer to the never-takers). The MTEs thus show how subsequent criminal justice contact varies across defendants who are induced into nonprosecution as the predicted probability of nonprosecution varies with the instrument. With these estimates we can explore heterogeneity by the unobservable propensity to be not prosecuted. As this propensity increases, the instrument is impacting different marginal defendants.

Figure A.2a shows the support of the predicted probability of nonprosecution (the propensity score) for prosecuted and non-prosecuted defendants. We can only trace out the MTEs along this range of common support. The estimates can become imprecise at the extreme ends of this distribution given smaller numbers of ADAs, so when estimating the MTEs we trim the top and bottom one percentiles of this common support distribution.[43]

Figure A.2b shows the estimated MTEs. Given the potential outcomes model outlined in Appendix C.2, we can interpret the x-axis as the predicted probability of nonprosecution. The MTEs decline monotonically as this probability increases, indicating that there is heterogeneity in the effect of nonprosecution on the probability of a future criminal complaint. The reduction in the probability of future criminal complaints is larger for defendants on the

[41]See Doyle Jr (2007), Maestas, Mullen and Strand (2013), French and Song (2014), Arnold, Dobbie and Yang (2018), and Bhuller et al. (2020) for empirical examples of MTE estimation in leniency designs.

[42]In practice, we use the Stata package `mtefe` (Andresen, 2019).

[43]Estimating and interpreting MTEs also requires a strict monotonicity assumption. Using the test of Frandsen, Lefgren and Leslie (2019) we could not reject the null of strict monotonicity holding in 6 out of 9 courts. The main MTEs presented here are estimated in our full data. In Figure A.2 we also repeat the analysis restricted to the 6 courts where we could not reject this null (see Table A.2) and find very similar results.

margin of nonprosecution for more-lenient ADAs (i.e., the defendants who are closer to the never-takers in our IV framework). Defendants on the margin of nonprosecution for less-lenient ADAs (i.e., defendants closer to the always-takers) experience a suggestive increase in the probability of a future criminal complaint as a consequence of nonprosecution, but the estimates at this more extreme end of the propensity score distribution are not statistically significantly different from zero and have wide confidence intervals.

We can express other treatment effect parameters as weighted averages of the MTEs, such as the (overall) average treatment effect (ATE), average treatment on the treated (ATT), and average treatment on the untreated (ATUT). We rescale the weights so that they integrate to one over the common support region shown in Figure A.2a and estimate these three treatment effects (Carneiro, Heckman and Vytlacil, 2011, Andresen, 2019). We report the estimates in the upper right corner of Figure A.2b. The estimated ATE is smaller than our LATE estimates but still large, negative, and statistically significant.

These estimates imply that increasing the leniency of ADA nonprosecution decisions—that is, not prosecuting more defendants who are currently never-takers in our IV framework—would, if anything, cause larger decreases in subsequent criminal justice contact than the average estimates reported earlier.

8.2 Effects of a Presumption of Nonprosecution

In this subsection we explore the impacts of the inauguration of Rachael Rollins as District Attorney of Suffolk County on January 2, 2019. During her 2018 election campaign, District Attorney Rollins pledged to establish a presumption of nonprosecution for 15 nonviolent misdemeanor offenses.[44] The proposed policy allowed for some ADA discretion in nonprosecution decisions, with supervisor approval. Figure B.1 reports monthly event study estimates of the effects of the Rollins inauguration on nonprosecution rates between January 1, 2018 and September 1, 2019 for a) cases with nonviolent misdemeanor offenses on the Rollins list, b) cases with nonviolent misdemeanor offenses not on the Rollins list, c) cases with any nonviolent misdemeanor offenses, and d) cases with nonviolent felony offenses. All models include fixed effects for court and day of week and all case-level covariates used throughout the paper; all plots report 90% confidence intervals based on robust standard errors clustered on defendants. There is seasonality in nonprosecution rates but no clear pre-inauguration directional time trends, for any category of case. After the inauguration of District Attorney Rollins, nonprosecution rates rose not only for cases involving the nonviolent misdemeanor

[44] https://rollins4da.com/policy/charges-to-be-declined/. Some charges on the Rollins list may be charged as felonies. We restrict our attention to changes in nonprosecution rates for nonviolent misdemeanor offenses (on and off the Rollins list), as defined earlier in the paper.

offenses on the Rollins list, but also for those involving nonviolent misdemeanor offenses not on the Rollins list (and for all nonviolent misdemeanor cases). However, nonprosecution rates did not rise for cases involving nonviolent felonies.[45]

Figure B.2 reports estimated effects of the Rollins inauguration on nonprosecution rates for these categories of nonviolent misdemeanor cases, with and without nonviolent felony cases as a control group, and estimated effects of nonprosecution on subsequent criminal complaints within one year of the current case, using the Rollins inauguration as an instrument for nonprosecution, again with and without nonviolent felony cases as a control group. The time period remains January 1, 2018 - September 1, 2019; defendants are followed for one year after the initiation of their current case, up to September 1, 2020. All models include fixed effects for court, month, and day of week and all case-level covariates used throughout the paper. Figure B.2 reports 95% confidence intervals based on robust standard errors clustered on defendants.

The top three coefficients reported in Panel A of Figure B.2 report OLS estimates of the average effects of the Rollins inauguration on nonprosecution rates for cases involving (i) nonviolent misdemeanors on the Rollins list, (ii) nonviolent misdemeanors not on the Rollins list, and (iii) all nonviolent misdemeanors. The bottom three coefficients reported in Panel A of Figure B.2 report difference-in-differences estimates of these effects, relative to nonprosecution rates for nonviolent felony cases. These latter estimates indicate that nonprosecution rates for cases involving nonviolent misdemeanors (on and off the Rollins list and overall) increased by five to eight percentage points on average between January 2, 2019 and September 1, 2019, or by 15 - 20% relative to the average 2018 nonprosecution rates for these cases (ranging between 34% for nonviolent misdemeanor offenses on the Rollins list to 38% for nonviolent misdemeanor offenses not on the Rollins list). All estimates are significant at $p < 0.01$.

The top three coefficients reported in Panel B of Figure B.2 report 2SLS estimates of the average effects of nonprosecution on the rates at which defendants are issued new criminal complaints within one year of the current case, for all three categories of nonviolent misdemeanors, using the Rollins inauguration as an instrument for nonprosecution. These second stage estimates are all negative and significant at $p < 0.01$. The bottom three coefficients reported in Panel B of Figure B.2 report 2SLS difference-in-differences estimates of these effects, relative to nonviolent felony cases.[46] These second-stage IV DD estimates indicate

[45] District Attorney Rollins issued a formal memo announcing the new nonprosecution policy on March 25, 2019, and SCDAO ADAs were trained in the new policy after that date. However, as can be seen in Figure B.1, increases in nonprosecution occurred between District Attorney Rollins' inauguration and March 25, 2019.

[46] Ouss and Stevenson (2020) use a similar IV difference-in-differences strategy in their study of a change

that the increases in nonprosecution after the Rollins inauguration led to a 41 percentage point decrease in new criminal complaints for nonviolent misdemeanor cases on the Rollins list (not significant), a 47 percentage point decrease in new criminal complaints for nonviolent misdemeanor cases not on the Rollins list ($p < .05$), and a 56 percentage point decrease in new criminal complaints for all nonviolent misdemeanor cases ($p < .05$).

Similar to our main estimation of impacts of increased ADA leniency at arraignment, these estimates suggest that policies introducing a presumption of nonprosecution for nonviolent misdemeanor offenses may have social benefits. The increases in nonprosecution of nonviolent misdemeanor offenses induced by the Rollins inauguration appear to have decreased the rates at which defendants are issued new criminal complaints within one year of the current case.

It is possible that future criminal complaints are not a good proxy for future criminal behavior, if police are less likely to arrest or cite individuals when their offenses are unlikely to be prosecuted. It is also possible that a change in policy to reduce prosecution of nonviolent misdemeanors could increase the number of crimes committed by other residents (who are not yet in our data) by reducing general deterrence. In both cases, looking directly at the number of crimes reported in Suffolk County would be helpful.

Figure B.3 shows the effects of District Attorney Rollins' inauguration on crimes reported to the Boston Police Department.[47] We focus on the types of offenses that could be affected by a presumption of nonprosecution—that is, crime types where the expected probability of prosecution has fallen. The data include crime reports from January 2017 through February 2020 (before COVID-19). We group incidents into the following categories: property damage, theft and fraud, disorder, drug, and other offenses.[48] Overall, we find significant reductions in reports of property damage and reports of theft/fraud. There is no evidence of an increase in any of these crime types.

Overall, we interpret these effects of District Attorney Rollins' inauguration and implementation of policies that reduced the prosecution of nonviolent misdemeanors as suggestive evidence that this policy shift reduced the future criminal justice involvement of this broader pool of defendants. Effects on reported crime are noisy, but there is certainly no evidence that this policy change had detrimental effects on public safety. It will be important to track changes over time in this setting and elsewhere, to more fully understand what trade-offs, if any, exist.

in bail policy in Philadelphia.

[47] Source: https://data.boston.gov/dataset/crime-incident-reports-august-2015-to-date-source-new-system

[48] There was a marked increase in the number of "verbal disputes" in the reported crime data beginning in October 2019—a near-doubling of such incidents. Because this seems to be driven by some other factor, we exclude this crime category from the analysis.

9 Discussion

Misdemeanor cases make up over 80 percent of the cases processed by the U.S. criminal justice system. Yet we know little about the causal impacts of misdemeanor prosecution or nonprosecution. We report the first estimates of the causal effects of misdemeanor nonprosecution on rates of post-arraignment criminal complaints, prosecutions, and criminal records. To do this, we leverage the as-if random assignment of nonviolent misdemeanor cases to arraigning ADAs in a large urban district attorney's office. Our findings imply that not prosecuting marginal nonviolent misdemeanor defendants substantially reduces their subsequent criminal justice contact, or, in other words, that prosecuting marginal nonviolent misdemeanor defendants substantially increases their subsequent criminal justice contact.

These findings are troubling, given the volume of misdemeanor prosecutions pursued in the United States. If prosecution of the marginal misdemeanor defendant increases the risk of post-arraignment criminal justice involvement, it is possible that misdemeanor prosecution also has negative labor market effects (Mueller-Smith and Schnepel, 2019). This in turn could lead to an increase in criminal activity. We may in fact be undermining public safety by criminalizing relatively minor forms of misbehavior.

The key policy question that motivated this study is whether scaling back the prosecution of nonviolent misdemeanor prosecution would enhance or reduce public safety. Our results reflect the combined effects of nonprosecution on defendant behavior and responses to that behavior from the criminal justice system, which makes the answer to this question less straightforward than it first appears. Since decisions to prosecute in future cases may be sensitive to a defendant's record of previous arrests/prosecutions/convictions, the effects of nonprosecution on future prosecution and criminal record acquisition may reflect, to large degree, the "ratcheting up" of criminal justice consequences for repeat offenders. That is, a repeat defendant may be more likely to be prosecuted than a first-time defendant, even when the criminal behavior they are accused of is the same. For this reason, a change in future prosecution and criminal record acquisition may not reflect a meaningful change in public safety. That said, future prosecutions and criminal records certainly incur a cost to the defendant (time, legal fees, court costs, and perhaps lost employment), and we should consider whether this cost is justified by any benefit to the community. If nonprosecution in an initial case helps defendants avoid future prosecutions and criminal records, with little to no change in their criminal behavior, that could easily yield net social benefits.

Future criminal complaints may provide a more objective measure of future criminal behavior, because police officers may be less likely to know/consider a person's criminal history when they are issuing complaints than prosecutors are to know/consider that history when deciding to prosecute. In this case, our finding that nonprosecution reduces the likelihood of future criminal complaints implies that scaling back the prosecution of nonviolent misdemeanor cases would indeed have social benefits due to a reduction in offending.

Our results suggest that inducing arraigning ADAs to be more lenient in their prosecution decisions could yield net social benefits. Preliminary evidence on the effects of a related policy change in Suffolk County—a presumption of nonprosecution for nonviolent misdemeanor offenses—supports this policy implication. We look forward to seeing future work on the longer-run effects of the SCDAO policy, and on the effects of similar prosecutor-led reforms in other contexts.

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics* 113(2):231–263.

Ahrens, Achim, Christian Hansen and Mark Schaffer. 2019. "`ivlasso and pdslasso`: Programs for post-selection and post-regularization OLS and IV estimation and inference.".

Aizer, Anna and Jr. Doyle, Joseph J. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics* 130(2):759–803.

Anderson, Theodore W and Herman Rubin. 1949. "Estimation of the parameters of a single equation in a complete system of stochastic equations." *Annals of Mathematical statistics* 20(1):46–63.

Andresen, Martin E. 2018. "Exploring marginal treatment effects: Flexible estimation using Stata." *The Stata Journal* 18(1):118–158.

Andresen, Martin E. 2019. "`MTEFE`: Stata module to compute marginal treatment effects with factor variables.". This version November 2020.

Andrews, Isaiah, James H Stock and Liyang Sun. 2019. "Weak instruments in instrumental variables regression: Theory and practice." *Annual Review of Economics* 11:727–753.

Aneja, Abhay P and Carlos F Avenancio-León. 2019. No Credit For Time Served? Incarceration and Credit-Driven Crime Cycles. Working paper.

Angrist, Joshua and Brigham Frandsen. 2020. Machine labor. Working Paper 26584 National Bureau of Economic Research.

Angrist, Joshua D and Alan B Keueger. 1991. "Does compulsory school attendance affect schooling and earnings?" *The Quarterly Journal of Economics* 106(4):979–1014.

Angrist, Joshua David, Guido W Imbens and Alan B Krueger. 1999. "Jackknife instrumental variables estimation." *Journal of Applied Econometrics* 14(1):57–67.

Arnold, David, Will Dobbie and Crystal S Yang. 2018. "Racial bias in bail decisions." *The Quarterly Journal of Economics* 133(4):1885–1932.

Arteaga, Carolina. 2020. Parental Incarceration and Children's Educational Attainment. Working paper.

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

Bekker, Paul A. 1994. "Alternative approximations to the distributions of instrumental variable estimators." *Econometrica* pp. 657–681.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov and Christian Hansen. 2012. "Sparse models and methods for optimal instruments with an application to eminent domain." *Econometrica* 80(6):2369–2429.

Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies* 81(2):608–650.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Loken and Magne Mogstad. 2018. "Intergenerational effects of incarceration." *AEA Papers and Proceedings* 108:234–40.

Bhuller, Manudeep, Gordon Dahl, Katrine Loken and Magne Mogstad. 2020. "Incarceration, Recidivism and Employment." *Journal of Political Economy* 128(4).

Bishop, Elizabeth, Brook Hopkins, Chijindu Obiofuma and Felix Owusu. 2020. Racial Dispariies in the Massachusetts Criminal System. Technical report Criminal Justice Policy Program Harvard Law School.
URL: *https://hls.harvard.edu/content/uploads/2020/11/Massachusetts-Racial-Disparity-Report-FINAL.pdf*

Björklund, Anders and Robert Moffitt. 1987. "The estimation of wage gains and welfare gains in self-selection models." *The Review of Economics and Statistics* 69:42–49.

Bound, John, David A Jaeger and Regina M Baker. 1995. "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American statistical association* 90(430):443–450.

Brinch, Christian N, Magne Mogstad and Matthew Wiswall. 2017. "Beyond LATE with a discrete instrument." *Journal of Political Economy* 125(4):985–1039.

Carneiro, Pedro, James J Heckman and Edward J Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review* 101(6):2754–81.

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

Chao, John C and Norman R Swanson. 2005. "Consistent estimation with a large number of weak instruments." *Econometrica* 73(5):1673–1692.

Chen, M Keith and Jesse M Shapiro. 2007. "Do harsher prison conditions reduce recidivism? A discontinuity-based approach." *American Law and Economics Review* 9(1):1–29.

Collinson, Robert and Davin Reed. 2019. The effects of evictions on low-income households. Working paper.

Cornelissen, Thomas, Christian Dustmann, Anna Raute and Uta Schönberg. 2016. "From LATE to MTE: Alternative methods for the evaluation of policy interventions." *Labour Economics* 41:47–60.

Dahl, Gordon B, Andreas Ravndal Kostøl and Magne Mogstad. 2014. "Family welfare cultures." *The Quarterly Journal of Economics* 129(4):1711–1752.

Davidson, Russell and James G MacKinnon. 2014. "Confidence sets based on inverting Anderson–Rubin tests." *The Econometrics Journal* 17(2):S39–S58.

de Chaisemartin, Clement. 2017. "Tolerating defiance? Local average treatment effects without monotonicity." *Quantitative Economics* 8(2):367–396.

Di Tella, Rafael and Ernesto Schargrodsky. 2013. "Criminal Recidivism after Prison and Electronic Monitoring." *Journal of Political Economy* 121(1):28–73.

Diamond, Rebecca, Adam Guren and Rose Tan. 2020. The effect of foreclosures on homeowners, tenants, and landlords. Working Paper 27358 National Bureau of Economic Research.

Dobbie, Will, Hans Grönqvist, Susan Niknami, Mårten Palme and Mikael Priks. 2018. The intergenerational effects of parental incarceration. Working Paper 24186 National Bureau of Economic Research.

Dobbie, Will, Jacob Goldin and Crystal S. Yang. 2018. "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review* 108(2):201–40.

Dobbie, Will, Paul Goldsmith-Pinkham and Crystal S Yang. 2017. "Consumer Bankruptcy and Financial Health." *The Review of Economics and Statistics* 99(5).

Doleac, Jennifer L. 2020. Encouraging desistance from crime. Working paper.

Doyle Jr, Joseph J. 2007. "Child protection and child outcomes: Measuring the effects of foster care." *American Economic Review* 97(5):1583–1610.

Drago, Francesco, Roberto Galbiati and Pietro Vertova. 2009. "The deterrent effects of prison: Evidence from a natural experiment." *Journal of Political Economy* 117(2):257–280.

Dusek, Libor and Christian Traxler. 2020. "Learning from law enforcement.".

Eren, Ozkan and Naci Mocan. 2019. "Juvenile Punishment, High School Graduation, and Adult Crime: Evidence from Idiosyncratic Judge Harshness." *Review of Economics and Statistics* pp. 1–14.

Frandsen, Brigham. 2020. "TESTJFE: Stata module to perform test for instrument validity in the judge fixed effects design.".

Frandsen, Brigham R., Lars J. Lefgren and Emily C. Leslie. 2019. Judging Judge Fixed Effects. Working Paper 25528 National Bureau of Economic Research.

French, Eric and Jae Song. 2014. "The effect of disability insurance receipt on labor supply." *American Economic Journal: Economic Policy* 6(2):291–337.

Gehrsitz, Markus. 2017. "Speeding, Punishment, and Recidivism: Evidence from a Regression Discontinuity Design." *The Journal of Law and Economics* 60(3):497–528.

Hansen, Benjamin. 2015. "Punishment and deterrence: Evidence from drunk driving." *American Economic Review* 105(4):1581–1617.

Hausman, Jerry A, Whitney K Newey, Tiemen Woutersen, John C Chao and Norman R Swanson. 2012. "Instrumental variable estimation with heteroskedasticity and many instruments." *Quantitative Economics* 3(2):211–255.

Heckman, James J and Edward Vytlacil. 2001. "Policy-relevant treatment effects." *American Economic Review* 91(2):107–111.

Heckman, James J and Edward Vytlacil. 2005. "Structural equations, treatment effects, and econometric policy evaluation 1." *Econometrica* 73(3):669–738.

Heckman, James J and Sergio Urzua. 2010. "Comparing IV with structural models: What simple IV can and cannot identify." *Journal of Econometrics* 156(1):27–37.

Heckman, James J, Sergio Urzua and Edward Vytlacil. 2006. "Understanding instrumental variables in models with essential heterogeneity." *The Review of Economics and Statistics* 88(3):389–432.

Hjalmarsson, Randi. 2009. "Juvenile jails: A path to the straight and narrow or to hardened criminality?" *The Journal of Law and Economics* 52(4):779–809.

Hull, Peter. 2017. Examiner Designs and First-Stage F Statistics: A Caution. Technical report.

Humphries, John Eric, Nicholas S Mader, Daniel I Tannenbaum and Winnie L Van Dijk. 2019. Does eviction cause poverty? Quasi-experimental evidence from Cook County, IL. Working Paper 26139 National Bureau of Economic Research.

Imbens, Guido W and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* pp. 467–475.

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

Kolesár, Michal. 2013. "Estimation in an instrumental variables model with treatment effect heterogeneity." *Unpublished Working Paper* .

Landersø, Rasmus, Helena Skyt Nielsen and Marianne Simonsen. 2017. "School starting age and the crime-age profile." *The Economic Journal* 127(602):1096–1118.

Laub, John H and Robert J Sampson. 2001. "Understanding desistance from crime." *Crime and justice* 28:1–69.

Leasure, Peter. 2019. "Misdemeanor Records and Employment Outcomes: An Experimental Study." *Crime & Delinquency* 65(13):1850–1872.

Lee, David S., Marcelo J. Moreira, Justin McCrary and Jack Porter. 2020. *arXiv e-prints* . arXiv:2010.05058.

Leslie, Emily and Nolan G. Pope. 2017. "The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments." *Journal of Law and Economics* 60(3).

Loeffler, Charles E. 2013. "Does imprisonment alter the life course? Evidence on crime and employment from a natural experiment." *Criminology* 51(1):137–166.

Maestas, Nicole, Kathleen J Mullen and Alexander Strand. 2013. "Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt." *American Economic Review* 103(5):1797–1829.

Mogstad, Magne and Alexander Torgovitsky. 2018. "Identification and extrapolation of causal effects with instrumental variables." *Annual Review of Economics* 10:577–613.

Montiel Olea, José Luis and Carolin Pflueger. 2013. "A robust test for weak instruments." *Journal of Business & Economic Statistics* 31(3):358–369.

Mueller-Smith, Michael. 2015. The Criminal and Labor Market Impacts of Incarceration. Working paper. https://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2015/09/incar.pdf.

Mueller-Smith, Michael and Kevin T. Schnepel. 2019. Diversion in the Criminal Justice System. Working paper. https://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2019/01/Diversion.pdf.

Nagin, Daniel S, Francis T Cullen and Cheryl Lero Jonson. 2009. "Imprisonment and reoffending." *Crime and Justice* 38(1):115–200.

Natapoff, Alexandra. 2018. *Punishment Without Crime: How Our Massive Misdemeanor System Traps the Innocent and Makes America More Unequal*. New York: Basic Books.

Norris, Samuel, Matthew Pecenco and Jeffrey Weaver. 2020. "The effects of parental and sibling incarceration: Evidence from ohio." *Available at SSRN 3590735* .

Ouss, Aurelie and Megan Stevenson. 2020. "Bail, Jail, and Pretrial Misconduct: The Influence of Prosecutors." http://aouss.github.io/NCB.pdf.

Raphael, Steven and Michael A. Stoll. 2014. "A New Approach to Reducing Incarceration While Maintaining Low Rates of Crime.". Hamilton Project Discussion Paper 2014-03.

Rehavi, M. Marit and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122(6).

Sloan, CarlyWill. 2020*a*. "How Much Does Your Prosecutor Matter? An Estimate of Prosecutorial Discretion.". Working paper.

Sloan, CarlyWill. 2020*b*. "Racial Bias by Prosecutors: Evidence from Random Assignment.". Working paper.

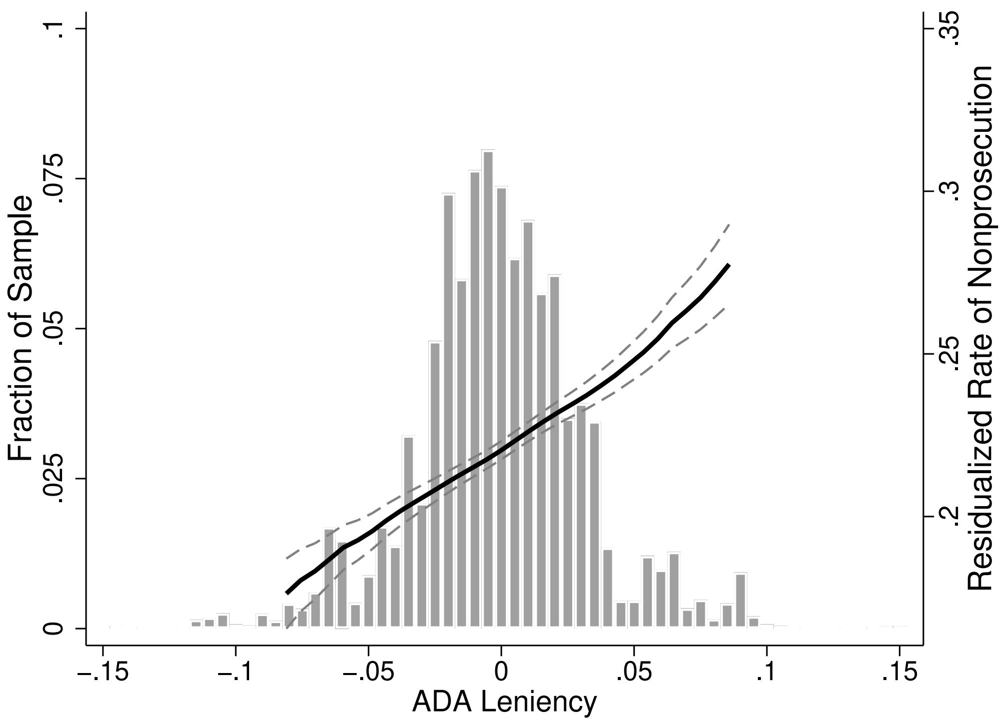
Stevenson, Megan and Sandra Mayson. 2018. "The scale of misdemeanor justice." *Boston University Law Review* 98:731.

Stevenson, Megan T. 2018. "Distortion of justice: How the inability to pay bail affects case outcomes." *The Journal of Law, Economics, and Organization* 34(4):511–542.

Tuttle, Cody. 2021. "Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums.". Working paper.

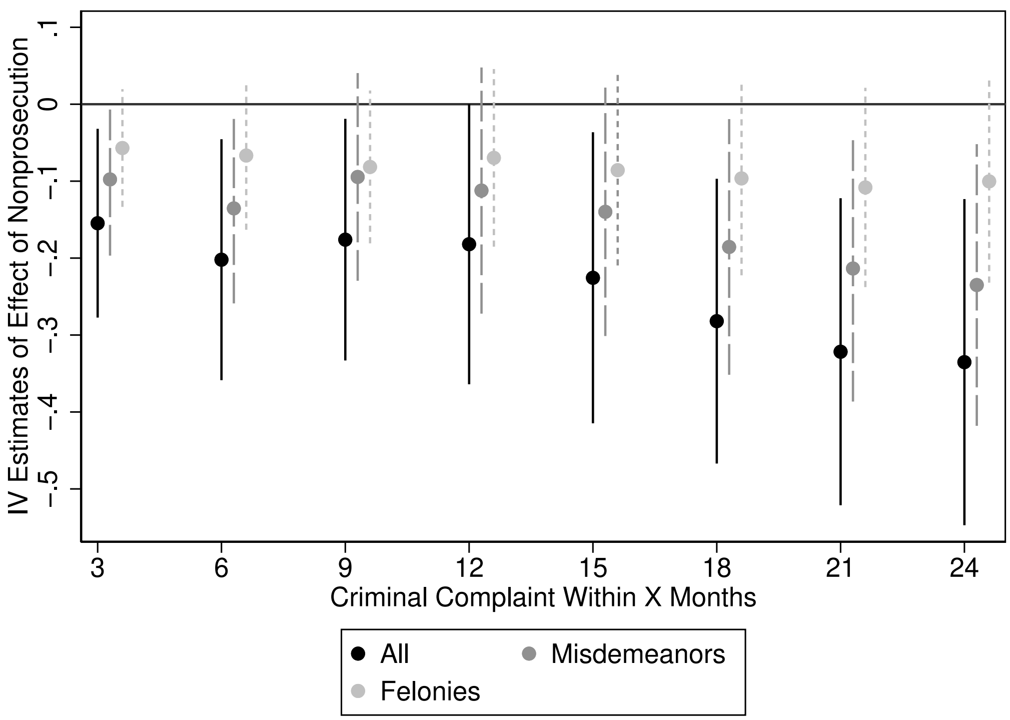
Uggen, Christopher, Mike Vuolo, Sarah Lageson, Ebony Ruhland and Hilary K. Whitham. 2014. "The Edge of Stigma: An Experimental Audit of the Effects of Low-Level Criminal Records on Employment." *Criminology* 52(4):627–654.

Figure 1: ADA Leniency and Nonprosecution



Note: This figure shows the distribution of our leave-out mean measure of ADA "leniency," residualized by court-by-month and court-by-day-of-week. More lenient ADAs have higher rates of not prosecuting nonviolent misdemeanor cases. The solid line is a local linear regression of nonprosecution on ADA leniency, along with the 95% confidence interval, estimated from the 1st to 99th percentiles of ADA leniency—a local linear version of our first stage. A case assigned to a more lenient ADA (computed using all cases except the current case and other cases with the same defendant) has a higher likelihood of being not prosecuted.

Figure 2: LATE Over Time (2 Years)

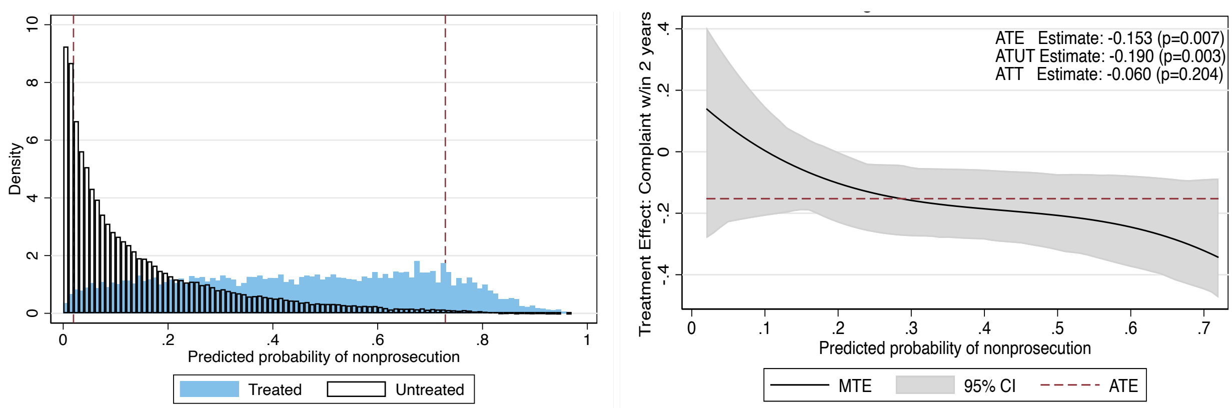


Note: This figure shows the local average treatment effect of nonprosecution on the likelihood of a new criminal complaint (y-axis) within a given number of months after the arraignment hearing (x-axis). Estimates are based on 2SLS regressions including covariates (the equivalent of Column (4) in Table 4). The circles show coefficients; the lines show 95% confidence intervals.

Figure 3: Marginal Treatment Effects

(a) Common Support

(b) MTE



Note: In (a) the dashed lines represent the upper and lower bounds on the common support of the propensity score (based on 1% trimming) used to estimate the MTEs. Propensity scores are predicted via a logit regression with all case- and defendant-level covariates included, including court-by-time FE. The MTE estimation is based on a local IV using a cubic polynomial specification in the sample with common support. The x-axis in Figure (b) is the predicted probability of nonprosecution estimated from the assigned ADA after residualizing out covariates and court-by-time fixed effects. Standard errors and resulting 95% confidence intervals are estimated using 100 bootstrap replications. The outcome of interest is the probability of a new criminal complaint within two years. The upper right corner on Panel (b) shows the estimated average treatment effect (ATE), average treatment on the untreated (ATUT), and average treatment on the treated (ATT). These were estimated by rescaling the weights on the MTEs for those parameters to integrate over the common support shown in (a) (Carneiro, Heckman and Vytlacil, 2011). All estimations were done via `mtefe` in Stata (Andresen, 2019).

Table 1: Summary Statistics

	(1) All	(2) Prosecuted	(3) Not Prosecuted
Baseline:			
Not Prosecuted	0.205	0.000	1.000
Number Counts	1.716	1.751	1.581
Number Misdemeanor Counts	1.319	1.365	1.141
Number of Serious Misdemeanor Counts	0.575	0.649	0.289
Misd Conviction within Past Year	0.085	0.099	0.030
Felony Conviction within Past Year	0.044	0.052	0.014
Citizen	0.765	0.744	0.849
Disorderly/Theft	0.284	0.307	0.194
Motor Vehicle	0.395	0.333	0.633
Drug	0.152	0.184	0.027
Other Crime	0.170	0.176	0.146
Victimless Crime	0.817	0.789	0.925
Case Outcomes:			
ADA Requested Bail	0.064	0.080	0.000
Bail Set at Arraignment	0.052	0.066	0.000
Amount Bail Set (Cond.)	401.759	401.759	.
Days to Disposition	146.667	184.510	0.000
DCJIS Record of Case	0.686	0.770	0.363
Any Conviction	0.209	0.263	0.000
Post-Case Outcomes:			
Criminal Complaint Within 2 Years	0.339	0.371	0.215
Misdemeanor Complaint Within 2 Years	0.225	0.243	0.155
Felony Complaint Within 2 Years	0.114	0.128	0.060
Prosecution Within 2 Years	0.305	0.341	0.164
Misdemeanor Prosecution Within 2 Years	0.193	0.216	0.105
Felony Prosecution Within 2 Years	0.112	0.126	0.059
DCJIS Record Within 2 Years	0.278	0.311	0.152
Misdemeanor DCJIS Record Within 2 Years	0.179	0.199	0.101
Felony DCJIS Record Within 2 Years	0.099	0.112	0.051
Observations	67553	53698	13855

Note: This sample includes cases with an arraignment hearing between January 1, 2004 – September 1, 2018, that have no felony or violent/gun misdemeanor charges, that are arraigned in one of Suffolk County's 9 district/municipal courts, that have an identified Assistant District Attorney (ADA) at arraignment, that are processed by an ADA who arraigned at least 30 nonviolent misdemeanor cases, and that are not "singletons" within our set of court-by-time fixed effects. Source: SCDAO.

Table 2: Randomization

	(1) Nonprosecution	(2) ADA Leniency
Number Counts	-0.019***	-0.000
	(0.003)	(0.000)
Number Misdemeanor Counts	0.018***	0.000
	(0.004)	(0.001)
Number of Serious Misdemeanor Counts	-0.102***	-0.000
	(0.006)	(0.000)
Misd Conviction within Past Year	-0.068***	-0.001
	(0.005)	(0.000)
Felony Conviction within Past Year	-0.053***	-0.001
	(0.006)	(0.001)
Citizen	0.042***	-0.000
	(0.004)	(0.000)
Disorderly/Theft	-0.014*	-0.001
	(0.008)	(0.001)
Motor Vehicle	0.105***	-0.000
	(0.009)	(0.000)
Drug	-0.094***	-0.001
	(0.009)	(0.001)
Constant	0.224***	0.001
	(0.009)	(0.002)
Observations	67553	67553
Joint F-Test p-value	0	0.234

Note: This table reports regressions testing the random assignment of cases to arraigning ADAs. ADA leniency is estimated using data from other nonviolent misdemeanor cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports estimates from an OLS regression of nonprosecution on the variables listed and court-by-time fixed effects. Column (2) reports estimates from an OLS regression of ADA leniency on the variables listed and court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The p-value reported at the bottom of Columns (1) and (2) is for an F-test of the joint significance of the variables listed with standard errors two-way clustered at the individual and ADA level. ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

Table 3: First Stage: ADA Leniency and Nonprosecution

	(1)	(2)
ADA Leniency	0.60***	0.55***
	(0.07)	(0.07)
Observations	67553	67553
Court x Time FE	Yes	Yes
Case/Def Covariates	No	Yes
Mean Not Prosecuted	0.205	
First Stage F-Stat	67.38	58.94

Note: This table reports first stage results via a linear probability model where the outcome variable is nonprosecution. The regressions are estimated on the sample as described in the notes to Table 1. ADA leniency is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports results controlling for our full set of court-by-time fixed effects. Column (2) adds defendant and case covariates: number of counts; number of misdemeanor counts; number of serious misdemeanor counts; whether the defendant had a prior misdemeanor conviction within the past year; whether the defendant had a prior felony conviction within the past year; indicators for whether the defendant faces charges for a disorder/theft, motor vehicle, drug, or other offense; and defendant citizenship status. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Robust (Kleibergen-Paap) first stage F reported (which is equivalent to the effective F-statistic of Montiel Olea and Pflueger (2013) in this case of a single instrument). $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table 4: Probability of a Subsequent Criminal Complaint

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Criminal Complaint Within 2 Years				
Not Prosecuted	-0.14***	-0.10***	-0.34***	-0.33***
	(0.01)	(0.01)	(0.10)	(0.11)
			[-0.55, -0.13]	[-0.54, -0.10]
Mean Dep Var Prosecuted	0.37			
Mean Dep Var Prosecuted Compliers	0.57			
Panel B: Misdemeanor Complaint Within 2 Years				
Not Prosecuted	-0.08***	-0.06***	-0.24***	-0.24***
	(0.00)	(0.00)	(0.09)	(0.09)
			[-0.42, -0.06]	[-0.43, -0.05]
Mean Dep Var Prosecuted	0.24			
Mean Dep Var Prosecuted Compliers	0.40			
Panel C: Felony Complaint Within 2 Years				
Not Prosecuted	-0.06***	-0.04***	-0.10*	-0.08
	(0.00)	(0.00)	(0.06)	(0.07)
			[-0.22, 0.03]	[-0.21, 0.06]
Mean Dep Var Prosecuted	0.13			
Mean Dep Var Prosecuted Compliers	0.17			
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within two years. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table 5: Probability of a Subsequent Prosecution or Subsequent DCJIS Record

	Subsq. Prosecution		Subsq. DCJIS Record	
	(1) OLS	(2) IV	(3) OLS	(4) IV
Panel A: Any Within 2 years				
Not Prosecuted	-0.12***	-0.35***	-0.11***	-0.40***
	(0.01)	(0.11)	(0.01)	(0.09)
		[-0.55, -0.12]		[-0.59, -0.21]
Mean Dep Var Prosecuted	0.34		0.31	
Mean Dep Var Prosecuted Compliers	0.53		0.58	
Panel B: Misdemeanor Within 2 Years				
Not Prosecuted	-0.08***	-0.25***	-0.07***	-0.29***
	(0.00)	(0.09)	(0.00)	(0.08)
		[-0.43, -0.07]		[-0.46, -0.12]
Mean Dep Var Prosecuted	0.22		0.20	
Mean Dep Var Prosecuted Compliers	0.36		0.38	
Panel C: Felony Within 2 Years				
Not Prosecuted	-0.04***	-0.09	-0.04***	-0.11*
	(0.00)	(0.07)	(0.00)	(0.06)
		[-0.22, 0.05]		[-0.23, 0.01]
Mean Dep Var Prosecuted	0.13		0.11	
Mean Dep Var Prosecuted Compliers	0.16		0.20	
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent prosecution or a subsequent DCJIS record within two years. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the super-column headings combined with panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

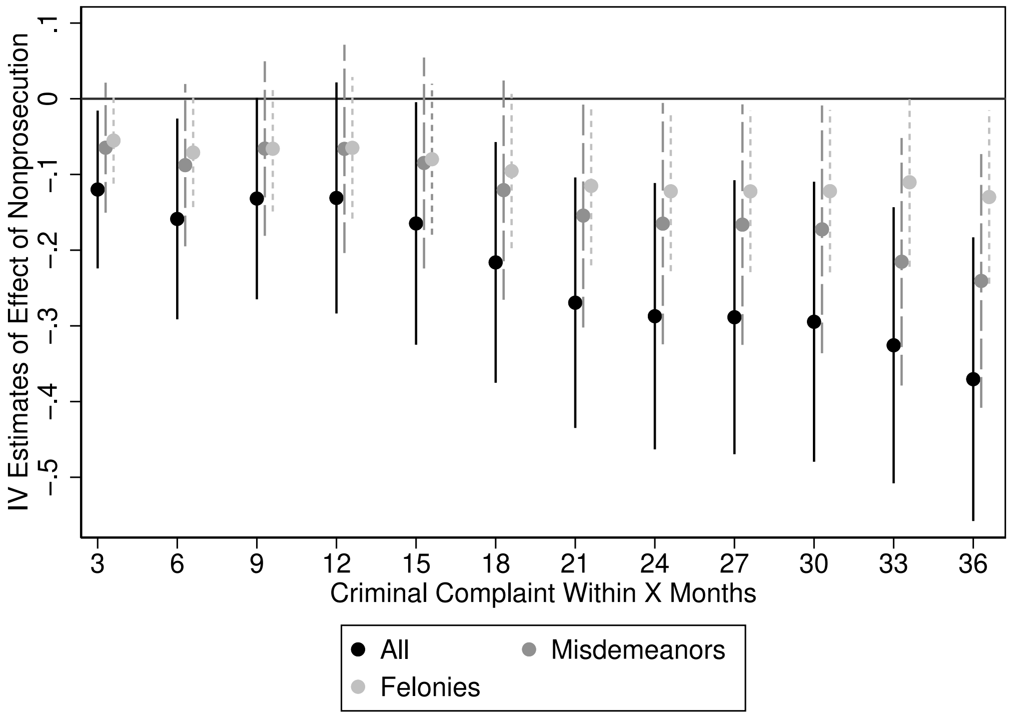
Table 6: Heterogeneous Effects: First Time Defendants versus Repeat Defendants

	Any Prev Complaint		Any Prev DCJIS		Any Prev Conviction	
	(1) No	(2) Yes	(3) No	(4) Yes	(5) No	(6) Yes
Panel A: Criminal Complaint Within 2 Years						
Not Prosecuted	-0.20*	-0.10	-0.29***	0.13	-0.26**	-0.11
	(0.12)	(0.23)	(0.10)	(0.30)	(0.11)	(0.37)
	[-0.42, 0.07]	[-0.56, 0.41]	[-0.50, -0.07]	[-0.44, 0.87]	[-0.48, -0.03]	[-0.85, 0.76]
Mean Dep Var Prosecuted	0.20	0.52	0.22	0.54	0.28	0.59
Mean Dep Var Prosecuted Compliers	0.25	0.62	0.32	0.61	0.40	0.86
Panel B: Prosecuted Within 2 Years						
Not Prosecuted	-0.25**	-0.11	-0.31***	0.09	-0.29***	-0.11
	(0.11)	(0.23)	(0.10)	(0.30)	(0.11)	(0.38)
	[-0.45, 0.01]	[-0.57, 0.41]	[-0.50, -0.08]	[-0.50, 0.81]	[-0.51, -0.06]	[-0.90, 0.76]
Mean Dep Var Prosecuted	0.18	0.48	0.20	0.51	0.25	0.56
Mean Dep Var Prosecuted Compliers	0.25	0.53	0.31	0.53	0.37	0.78
Panel C: DCJIS Record Within 2 Years						
Not Prosecuted	-0.20**	-0.40*	-0.27***	-0.22	-0.32***	-0.29
	(0.10)	(0.22)	(0.09)	(0.26)	(0.09)	(0.37)
	[-0.38, 0.02]	[-0.92, 0.03]	[-0.44, -0.09]	[-0.79, 0.33]	[-0.51, -0.13]	[-1.10, 0.52]
Mean Dep Var Prosecuted	0.15	0.45	0.15	0.49	0.22	0.52
Mean Dep Var Prosecuted Compliers	0.20	0.82	0.25	0.87	0.39	1.04
Observations	33803	33617	38929	28487	49752	17678
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports two-stage least squares estimates of the impact of nonprosecution on subsequent criminal justice involvement, for first-time and repeat defendants (defined in turn as (i) having any prior complaint in Suffolk County, (ii) having a prior complaint in Suffolk County that resulted in a DCJIS record, and (iii) having a prior complaint in Suffolk County that resulted in a conviction). The dependent variables are identified in the panel headings. Each panel reports the means of the dependent variable for prosecuted defendants, by subsample. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. The models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

A Additional Figures and Tables

Figure A.1: LATE Over Time (3 Years)



Note: This figure shows the effect of nonprosecution on the likelihood of a new criminal complaint (y-axis) within a given number of months after the arraignment hearing (x-axis), for a three-year post-arraignment window. Estimates are based on 2SLS regressions including covariates (the equivalent of Column (4) in Table A.11). The circles show coefficients; the lines show 95% confidence intervals.

Table A.1: First Stage Results for Other Case Outcomes

	Prosecuted Defendants	
	(1)	(2)
Panel A: Days to Disposition		
ADA Leniency	21.937	21.937
	(36.764)	(36.764)
Mean Not Prosecuted	184.398	
Panel B: Conviction		
ADA Leniency	-0.0390	-0.0428
	(-0.54)	(-0.61)
Mean Not Prosecuted	0.263	
Observations	53657	53657
Court x Time FE	Yes	Yes
Case/Def Covariates	No	Yes

Note: This table reports additional first stage results for days to disposition and whether a defendant receives a criminal conviction among defendants who are prosecuted. All specifications control for court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

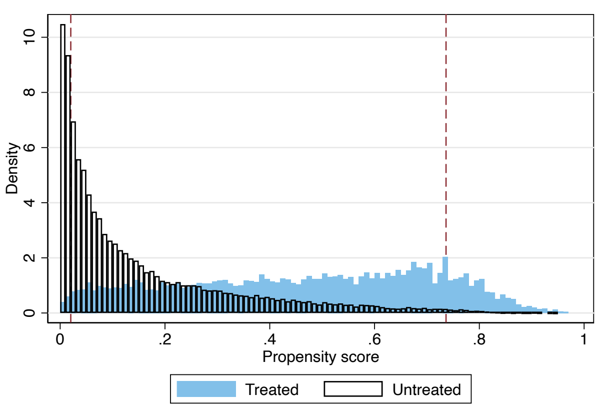
Table A.2: Frandsen, Lefgren, Leslie (2020) Test of Joint Null of Exclusion and Monotonicity, by Court

	Count	FLL p-value
Dorchester	15467	0.107
Roxbury	15055	0.167
Central	10335	0.000
West Roxbury	9514	0.460
East Boston	7757	0.639
South Boston	6385	0.517
Chelsea	1223	0.000
Brighton	1177	0.000
Charlestown	577	0.265

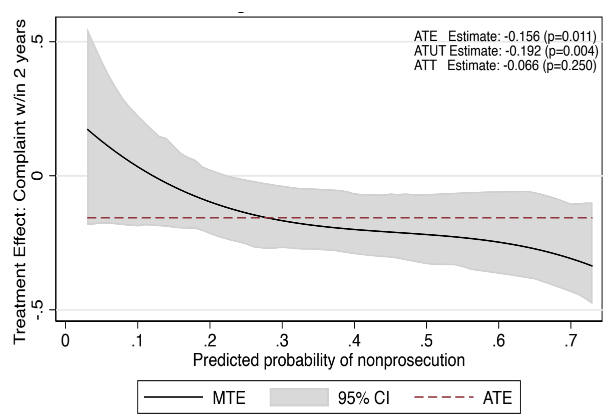
Note: This table presents results from the test proposed in Frandsen, Lefgren, and Leslie (2020) for the joint null hypothesis that the monotonicity and exclusion restrictions hold. We test this null within courts using day-of-week and year-month fixed effects along with our main covariates. A failure to reject the null implies that we cannot reject the hypothesis that the monotonicity and exclusion restrictions jointly hold. This test was implemented in Stata via the package `testjfe` (Frandsen, 2020).

Figure A.2: Marginal Treatment Effects: within "Strict Monotonicity" Courts

(a) Common Support



(b) MTE



Note: This figure repeats the analysis in Figure 3 but restricted to the 6 courts in which we could not reject the null of strict monotonicity holding (see Table A.2). In (a) the dashed lines represent the upper and lower bounds on the common support of the propensity score (based on 1% trimming) used to estimate the MTEs. Propensity scores are predicted via a logit regression with all case- and defendant-level covariates included, including court-by-time FE. The MTE estimation is based on a local IV using a cubic polynomial specification in the sample with common support. The x-axis in Figure (b) is the predicted probability of nonprosecution estimated from the assigned ADA after residualizing out covariates and court-by-time fixed effects. Standard errors and resulting 95% confidence intervals are estimated using 100 bootstrap replications. The outcome of interest is the probability of a new criminal complaint within two years. The upper right corner on Panel (b) shows the estimated average treatment effect (ATE), average treatment on the untreated (ATUT), and average treatment on the treated (ATT). These were estimated by rescaling the weights on the MTEs for those parameters to integrate over the common support shown in (a) (Carneiro, Heckman and Vytlacil, 2011). All estimations were done via `mtefe` in Stata (Andresen, 2019).

Table A.3: Monotonicity

	(1)	(2)	(3)	(4)	(5)	(6)
	1 Count	> 1 Count	Misd 1 Count	Misd > 1 Count	Ser Misd 1 Count	Ser Misd > 1 Count
Not Prosecuted	0.60***	0.57***	0.61***	0.47***	0.60***	0.48***
	(0.08)	(0.09)	(0.08)	(0.07)	(0.08)	(0.07)
Observations	37065	30341	48552	18869	36042	31336
	(7)	(8)	(9)	(10)	(11)	(12)
	No Misd Conv w/in 1 Year	Misd Conv w/in 1 Year	No Felony Conv w/in 1 Year	Felony Conv w/in 1 Year	Citizen	Non-Citizen
Not Prosecuted	0.62***	0.31***	0.60***	0.37*	0.26***	0.68***
	(0.08)	(0.12)	(0.07)	(0.20)	(0.08)	(0.08)
Observations	63610	3717	65862	1394	15739	51658
	(13)	(14)	(15)	(16)	(17)	(18)
	Any Disorder Crime	Any MV Crime	Any Drug Crime	Any Other Crime	Victimless Crime	Victim Crime
Not Prosecuted	0.64***	0.54***	0.10*	0.43**	0.65***	0.13*
	(0.13)	(0.11)	(0.06)	(0.21)	(0.08)	(0.07)
Observations	22225	30576	14132	8092	55164	12274
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	No	No	No	No	No	No

Note: This table reports first stage results by subsamples based on case and defendant characteristics, as listed in the column headers. The regressions are estimated on the sample as described in the notes to Table 1. All specifications control for court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.10$

Table A.4: Number of Subsequent Criminal Complaints

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Number Criminal Complaints Within 2 Years				
Not Prosecuted	-0.73***	-0.55***	-2.22***	-2.06***
	(0.04)	(0.04)	(0.65)	(0.67)
			[-3.60, -0.96]	[-3.48, -0.77]
Mean Dep Var Prosecuted	1.63			
Mean Dep Var Prosecuted Compliers	2.99			
Panel B: Number Misdemeanor Complaints Within 2 Years				
Not Prosecuted	-0.43***	-0.32***	-1.28***	-1.19***
	(0.02)	(0.02)	(0.37)	(0.37)
			[-2.06, -0.57]	[-1.99, -0.47]
Mean Dep Var Prosecuted	0.94			
Mean Dep Var Prosecuted Compliers	1.77			
Panel C: Number Felony Complaints Within 2 Years				
Not Prosecuted	-0.25***	-0.17***	-0.74**	-0.66**
	(0.02)	(0.02)	(0.32)	(0.33)
			[-1.41, -0.13]	[-1.36, -0.01]
Mean Dep Var Prosecuted	0.51			
Mean Dep Var Prosecuted Compliers	0.88			
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on numbers of subsequent criminal complaints within two years post-arraignment. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.5: Probability of a Subsequent Complaint by Complaint Type
2SLS Only

	(1) Violent	(2) MV	(3) Disorder/ Theft	(4) Drugs	(5) Other
Panel A: Any Complaint within 2 Years					
Not Prosecuted	-0.14**	-0.10*	-0.21***	-0.07	0.00
	(0.06)	(0.06)	(0.07)	(0.07)	(0.04)
	[-0.27, -0.03]	[-0.23, 0.01]	[-0.34, -0.07]	[-0.20, 0.07]	[-0.07, 0.08]
Mean Dep Var Pros.	0.08	0.10	0.16	0.08	0.04
Mean Dep Var Pros. Compliers	0.22	0.16	0.23	0.11	0.02
Panel B: Misdemeanor Complaint within 2 Years					
Not Prosecuted	-0.16***	-0.11*	-0.14**	-0.06	0.01
	(0.05)	(0.06)	(0.06)	(0.06)	(0.04)
	[-0.28, -0.06]	[-0.23, 0.00]	[-0.27, -0.02]	[-0.19, 0.07]	[-0.07, 0.09]
Mean Dep Var Pros.	0.05	0.09	0.12	0.06	0.04
Mean Dep Var Pros. Compliers	0.19	0.16	0.18	0.08	0.01
Panel C: Felony Complaint within 2 Years					
Not Prosecuted	-0.02	0.00	-0.05	-0.02	-0.01
	(0.04)	(0.01)	(0.04)	(0.04)	(0.01)
	[-0.10, 0.07]	[-0.02, 0.02]	[-0.14, 0.04]	[-0.09, 0.06]	[-0.04, 0.00]
Mean Dep Var Pros.	0.04	0.00	0.06	0.03	0.00
Mean Dep Var Pros. Compliers	0.07	-0.00	0.06	0.05	0.01
Observations	67553	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes

Note: This table reports two-stage least squares estimates of the impact of nonprosecution on rearrest probabilities by complaint type (i.e., the dependent variable in Column (1) is any violent complaint within two years). The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the column headings combined with the panel headings (i.e., in Panel B Column (1), the dependent variable is any violent misdemeanor criminal complaint within two years). Each panel reports the mean of the dependent variable for all prosecuted defendants and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

Table A.6: Randomization (1-Year Followup Sample)

	(1) Nonprosecution	(2) ADA Leniency
Number Counts	-0.019***	0.000
	(0.003)	(0.000)
Number Misdemeanor Counts	0.020***	-0.000
	(0.004)	(0.000)
Number of Serious Misdemeanor Counts	-0.110***	-0.000
	(0.006)	(0.000)
Misd Conviction within Past Year	-0.068***	-0.000
	(0.005)	(0.000)
Felony Conviction within Past Year	-0.062***	-0.001
	(0.006)	(0.001)
Citizen	0.042***	-0.000
	(0.004)	(0.000)
Disorderly/Theft	-0.015*	-0.001
	(0.008)	(0.001)
Motor Vehicle	0.113***	-0.000
	(0.009)	(0.000)
Drug	-0.093***	-0.001
	(0.008)	(0.001)
Constant	0.239***	0.001
	(0.009)	(0.002)
Observations	74631	74631
Joint F-Test p-value	0	0.325

Note: This table reports regressions testing the random assignment of cases to arraigning ADAs for the sample consisting of a one-year followup period post-arraignment. ADA leniency is estimated using data from other nonviolent misdemeanor cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports estimates from an OLS regression of nonprosecution on the variables listed and court-by-time fixed effects. Column (2) reports estimates from an OLS regression of ADA leniency on the variables listed and court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The p-value reported at the bottom of Columns (1) and (2) is for an F-test of the joint significance of the variables listed with standard errors two-way clustered at the individual and ADA level. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.7: First Stage (1-Year Followup Sample)

	(1)	(2)
ADA Leniency	0.57***	0.53***
	(0.07)	(0.07)
Observations	74631	74631
Court x Time FE	Yes	Yes
Case/Def Covariates	No	Yes
Mean Not Prosecuted	0.221	
First Stage F-Stat	65.55	58.05

Note: This table reports first stage results via a linear probability model where the outcome variable is nonprosecution. The regressions are estimated on the one-year followup sample. ADA leniency is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports results controlling for our full set of court-by-time fixed effects. Column (2) adds defendant and case covariates: number of counts; number of misdemeanor counts; number of serious misdemeanor counts; whether the defendant had a prior misdemeanor conviction within the past year; whether the defendant had a prior felony conviction within the past year; indicators for whether the defendant faces charges for a disorder/theft, motor vehicle, drug, or other offense; and defendant citizenship status. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Robust first-stage F reported (which is equivalent to the effective F-statistic of Montiel Olea and Pflueger (2013) in this case of a single instrument). $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.8: Probability of Subsequent Criminal Complaint (1-Year Followup Sample)

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Criminal Complaint Within 1 Years				
Not Prosecuted	-0.11***	-0.08***	-0.20**	-0.18*
	(0.00)	(0.00)	(0.09)	(0.09)
			[-0.36, -0.02]	[-0.35, 0.02]
Mean Dep Var Prosecuted	0.28			
Mean Dep Var Prosecuted Compliers	0.42			
Panel B: Misdemeanor Complaint Within 1 Years				
Not Prosecuted	-0.06***	-0.05***	-0.15**	-0.15*
	(0.00)	(0.00)	(0.08)	(0.08)
			[-0.30, 0.01]	[-0.30, 0.02]
Mean Dep Var Prosecuted	0.19			
Mean Dep Var Prosecuted Compliers	0.32			
Panel C: Felony Complaint Within 1 Years				
Not Prosecuted	-0.05***	-0.03***	-0.05	-0.03
	(0.00)	(0.00)	(0.05)	(0.06)
			[-0.15, 0.06]	[-0.14, 0.09]
Mean Dep Var Prosecuted	0.10			
Mean Dep Var Prosecuted Compliers	0.10			
Observations	74631	74631	74631	74631
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within one year. The regressions are estimated on the one-year followup sample. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes amongst prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

Table A.9: Randomization (3-Year Followup Sample)

	(1) Nonprosecution	(2) ADA Leniency
Number Counts	-0.017***	-0.000
	(0.003)	(0.000)
Number Misdemeanor Counts	0.015***	0.000
	(0.004)	(0.001)
Number of Serious Misdemeanor Counts	-0.098***	-0.000*
	(0.006)	(0.000)
Misd Conviction within Past Year	-0.066***	-0.000
	(0.005)	(0.000)
Felony Conviction within Past Year	-0.055***	-0.001
	(0.006)	(0.001)
Citizen	0.041***	-0.000
	(0.004)	(0.000)
Disorderly/Theft	-0.013	-0.001
	(0.008)	(0.001)
Motor Vehicle	0.099***	-0.000
	(0.009)	(0.001)
Drug	-0.094***	-0.001*
	(0.009)	(0.001)
Constant	0.218***	0.001
	(0.009)	(0.002)
Observations	63655	63655
Joint F-Test p-value	0	0.331

Note: This table reports regressions testing the random assignment of cases to arraigning ADAs for the sample consisting of a three-year followup period post-arraignment. ADA leniency is estimated using data from other nonviolent misdemeanor cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports estimates from an OLS regression of nonprosecution on the variables listed and court-by-time fixed effects. Column (2) reports estimates from an OLS regression of ADA leniency on the variables listed and court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The p-value reported at the bottom of Columns (1) and (2) is for an F-test of the joint significance of the variables listed with standard errors two-way clustered at the individual and ADA level. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.10: First Stage (3-Year Followup Sample)

	(1)	(2)
ADA Leniency	0.64***	0.59***
	(0.07)	(0.07)
Observations	63655	63655
Court x Time FE	Yes	Yes
Case/Def Covariates	No	Yes
Mean Not Prosecuted	0.197	
First Stage F-Stat	83.24	74.31

Note: This table reports first stage results via a linear probability model where the outcome variable is nonprosecution. The regressions are estimated on the three-year followup sample. ADA leniency is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports results controlling for our full set of court-by-time fixed effects. Column (2) adds defendant and case covariates: number of counts; number of misdemeanor counts; number of serious misdemeanor counts; whether the defendant had a prior misdemeanor conviction within the past year; whether the defendant had a prior felony conviction within the past year; indicators for whether the defendant faces charges for a disorder/theft, motor vehicle, drug, or other offense; and defendant citizenship status. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Robust first-stage F reported (which is equivalent to the effective F-statistic of Montiel Olea and Pflueger (2013) in this case of a single instrument). ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

Table A.11: Probability of Subsequent Criminal Complaint (3-Year Followup Sample)

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Criminal Complaint Within 3 Years				
Not Prosecuted	-0.15***	-0.12***	-0.43***	-0.43***
	(0.01)	(0.01)	(0.10)	(0.10)
			[-0.63, -0.23]	[-0.63, -0.21]
Mean Dep Var Prosecuted	0.42			
Mean Dep Var Prosecuted Compliers	0.63			
Panel B: Misdemeanor Complaint Within 3 Years				
Not Prosecuted	-0.08***	-0.07***	-0.30***	-0.31***
	(0.00)	(0.00)	(0.09)	(0.09)
			[-0.48, -0.12]	[-0.49, -0.12]
Mean Dep Var Prosecuted	0.28			
Mean Dep Var Prosecuted Compliers	0.43			
Panel C: Felony Complaint Within 3 Years				
Not Prosecuted	-0.07***	-0.04***	-0.13**	-0.12*
	(0.00)	(0.00)	(0.06)	(0.07)
			[-0.26, -0.00]	[-0.25, 0.02]
Mean Dep Var Prosecuted	0.14			
Mean Dep Var Prosecuted Compliers	0.20			
Observations	63655	63655	63655	63655
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within three years. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.12: DCJIS Record Acquisition

	OLS		IV	
	(1)	(2)	(3)	(4)
Not Prosecuted	-0.37***	-0.32***	-0.58***	-0.55***
	(0.01)	(0.01)	(0.10)	(0.11)
			[-0.77, -0.37]	[-0.76, -0.31]
Mean Dep Var Prosecuted	0.77			
Mean Dep Var Prosecuted Compliers	0.99			
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability that a criminal complaint receives a DCJIS record (an official record of the criminal record in the criminal offender record information database). The regressions are estimated on the sample as described in the notes to Table 1. We report the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-month and court-by-day-of-week fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, confidence intervals based on inversion of the Anderson-Rubin test are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.13: Separating Nonprosecution and Bail Decisions

	(1) Main	(2) Control For No Bail Leniency	(3) IV No Bail Leniency	(4) Both IVs
Not Prosecuted	-0.335***	-0.338***		-0.341**
	(-3.11)	(-2.85)		(-2.56)
No Bail			-0.248	0.0204
			(-1.41)	(0.10)
Observations	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes

Note: This table reports 2SLS estimates that explore the role of bail requests on the probability that a defendant receives a new criminal complaint within two years post-arraignment. Column (1) reports our main estimates from Column (4) of Table 4, Panel A. Column (2) includes as a covariate a "no-bail leniency" measure based on ADAs' propensity to request bail in other defendants' cases. Column (3) uses the no-bail leniency measure as an instrument for not receiving bail. Column (4) includes both nonprosecution leniency and no-bail leniency as instruments in the same regression. All specifications control for court-by-time fixed effects and all case/defendant covariates. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.14: Reduced Form: Subsequent Criminal Justice Involvement

	Subq. Complaint		Subq. Prosecution		Subq. DCJIS	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Any						
ADA Leniency	-0.21***	-0.19***	-0.22***	-0.19***	-0.25***	-0.22***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)
Mean Dep Var Prosecuted	0.37		0.34		0.31	
Panel B: Misdemeanors						
ADA Leniency	-0.14**	-0.13**	-0.15***	-0.14**	-0.17***	-0.16***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Mean Dep Var Prosecuted	0.24		0.22		0.20	
Panel C: Felonies						
ADA Leniency	-0.07*	-0.06	-0.07*	-0.05	-0.08**	-0.06*
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Mean Dep Var Prosecuted	0.13		0.13		0.11	
Observations	67553	67553	67553	67553	67553	67553
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes	No	Yes

Note: This table reports reduced form OLS estimates of the impact of case assignment to a more "lenient" ADA. The regressions are estimated on the sample as described in the notes to Table 1. The dependent variables are identified in the column and panel headings. Each panel reports the mean of the dependent variable for prosecuted defendants. ADA leniency is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.15: Different IV Definitions

	(1) Main	(2) Non-Residualized	(3) Experience	(4) Victimless	(5) Crime Type
DCJIS Record	-0.55***	-0.87***	-0.51***	-0.60***	-0.61***
	(0.11)	(0.17)	(0.12)	(0.14)	(0.12)
	[-0.76, -0.31]	[-1.32, -0.55]	[-0.76, -0.24]	[-0.89, -0.29]	[-0.86, -0.37]
Complaint within 2 Years	-0.34***	-0.43***	-0.36***	-0.30**	-0.28**
	(0.11)	(0.16)	(0.12)	(0.14)	(0.11)
	[-0.54, -0.11]	[-0.80, -0.10]	[-0.63, -0.11]	[-0.57, 0.04]	[-0.51, -0.03]
Misd Complaint within 2 Years	-0.23**	-0.35**	-0.22**	-0.27**	-0.13
	(0.09)	(0.14)	(0.11)	(0.12)	(0.10)
	[-0.42, -0.04]	[-0.68, -0.07]	[-0.46, -0.00]	[-0.53, -0.01]	[-0.33, 0.11]
Felony Complaint within 2 Years	-0.10	-0.08	-0.13	-0.03	-0.15**
	(0.07)	(0.11)	(0.09)	(0.10)	(0.07)
	[-0.23, 0.04]	[-0.32, 0.18]	[-0.33, 0.05]	[-0.21, 0.21]	[-0.31, 0.00]
Prosecution within 2 Years	-0.35***	-0.40**	-0.39***	-0.34**	-0.26**
	(0.11)	(0.16)	(0.12)	(0.13)	(0.11)
	[-0.55, -0.12]	[-0.75, -0.06]	[-0.65, -0.14]	[-0.59, -0.02]	[-0.47, -0.01]
DCJIS Record within 2 Years	-0.40***	-0.39***	-0.40***	-0.37***	-0.31***
	(0.09)	(0.15)	(0.12)	(0.12)	(0.10)
	[-0.59, -0.21]	[-0.72, -0.08]	[-0.67, -0.16]	[-0.62, -0.09]	[-0.52, -0.11]

Note: This table reports two-stage least squares results of the impact of nonprosecution on the probability of the outcomes in the rows, for different definitions of the instrument. All models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA. All specifications control for court-by-time fixed effects and case/defendant covariates. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Column (1) repeats our main 2SLS results with covariates. In Column (2) we consider a version of our leave-out instrument that does not residualize out court-by-time fixed effects, and that is thus a raw measure of ADAs' leave-out nonprosecution rates. Columns (3)-(5) interact our main leave-out instrument with high/low ADA experience (as measured by above- or below-median number of nonviolent misdemeanors arraigned as of the time of this case's arraignment), with an indicator for whether the crime is categorized as victimless, and with our mutually exclusive crime types. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.16: Missing Race

	Not Prosecuted		Prosecuted	
	(1) Missing Race	(2) Not Missing Race	(3) Missing Race	(4) Not Missing Race
Outcomes:				
Criminal Complaint Within 2 Years	0.10	0.26	0.18	0.40
Prosecution Within 2 Years	0.06	0.20	0.16	0.37
DCJIS Record Within 2 Years	0.05	0.19	0.14	0.34
Baseline:				
Number Counts	1.45	1.63	1.73	1.75
Number Misdemeanor Counts	1.10	1.16	1.28	1.38
Number of Serious Misdemeanor Counts	0.28	0.29	0.71	0.64
Misd Conviction within Past Year	0.01	0.04	0.03	0.11
Felony Conviction within Past Year	0.00	0.02	0.02	0.06
Citizen	0.96	0.81	0.94	0.71
Disorderly/Theft	0.11	0.23	0.28	0.31
Motor Vehicle	0.67	0.62	0.41	0.32
Drug	0.01	0.03	0.15	0.19
Observations	3742	10113	7176	46522
Proportion Missing Race	0.270		0.134	

Note: This table reports summary statistics for the samples of nonviolent misdemeanor cases meeting all other sample criteria that do and do not have information on defendant race/ethnicity, as indicated by the column headers.

Table A.17: Summary Statistics (with Demographic Variables)

	(1) All	(2) Prosecuted	(3) Not Prosecuted
Baseline:			
Not Prosecuted	0.204	0.000	1.000
Number Counts	1.717	1.752	1.581
Number Misdemeanor Counts	1.320	1.365	1.142
Number of Serious Misdemeanor Counts	0.575	0.648	0.290
Misd Conviction within Past Year	0.086	0.100	0.031
Felony Conviction within Past Year	0.045	0.053	0.014
Citizen	0.764	0.743	0.848
Disorderly/Theft	0.284	0.307	0.196
Motor Vehicle	0.394	0.333	0.633
Drug	0.152	0.185	0.027
Other Crime	0.170	0.176	0.144
Victimless Crime	0.817	0.789	0.925
Male	0.799	0.814	0.739
Age $\leq$ 23	0.231	0.233	0.222
Age 24-30	0.246	0.245	0.250
Age 31-40	0.218	0.220	0.210
Age $\geq$ 41	0.306	0.302	0.318
Prob Hispanic	0.332	0.338	0.308
Prob Black	0.345	0.355	0.307
Prob White	0.256	0.242	0.312
Admin Race Data: N=56605			
Black	0.458	0.468	0.412
White	0.363	0.366	0.347
Hispanic	0.160	0.148	0.217
Case Outcomes:			
ADA Requested Bail	0.064	0.080	0.000
Bail Set at Arraignment	0.052	0.066	0.000
Amount Bail Set (Cond.)	403.995	403.995	.
Days to Disposition	146.921	184.477	0.000
DCJIS Record of Case	0.689	0.771	0.366
Any Conviction	0.210	0.263	0.000
Post-Case Outcomes:			
Criminal Complaint Within 2 Years	0.341	0.372	0.217
Misdemeanor Complaint Within 2 Years	0.226	0.244	0.156
Felony Complaint Within 2 Years	0.115	0.128	0.061
Prosecution Within 2 Years	0.307	0.343	0.167
Misdemeanor Prosecution Within 2 Years	0.194	0.216	0.107
Felony Prosecution Within 2 Years	0.113	0.126	0.060
DCJIS Record Within 2 Years	0.280	0.312	0.154
Misdemeanor DCJIS Record Within 2 Years	0.180	0.200	0.102
Felony DCJIS Record Within 2 Years	0.100	0.112	0.052
Observations	67064	53411	13653

Note: This sample includes cases with an arraignment hearing between January 1, 2004 – September 1, 2018, that have no felony or violent/gun misdemeanor charges, that are arraigned in one of Suffolk County's 9 district/municipal courts, that have an identified Assistant District Attorney (ADA) at arraignment, that are processed by an ADA who arraigned at least 30 nonviolent misdemeanor cases, that are not "singletons" within our set of court-by-time fixed effects, and that are not missing gender or age.

Table A.18: Randomization (with Demographic Variables)

	(1) Nonprosecution	(2) ADA Leniency
Number Counts	-0.017***	-0.000
	(0.003)	(0.000)
Number Misdemeanor Counts	0.016***	-0.000
	(0.004)	(0.001)
Number of Serious Misdemeanor Counts	-0.101***	-0.000
	(0.006)	(0.000)
Misd Conviction within Past Year	-0.061***	-0.001
	(0.005)	(0.000)
Felony Conviction within Past Year	-0.050***	-0.001
	(0.006)	(0.001)
Citizen	0.037***	-0.000
	(0.004)	(0.000)
Disorderly/Theft	-0.021**	-0.001
	(0.008)	(0.001)
Motor Vehicle	0.102***	-0.000
	(0.009)	(0.001)
Drug	-0.093***	-0.001
	(0.008)	(0.001)
Male	-0.060***	-0.001***
	(0.004)	(0.000)
Age 24-30	-0.019***	0.000
	(0.005)	(0.000)
Age 31-40	-0.024***	0.000
	(0.005)	(0.000)
Age $\geq$ 41	-0.011**	0.001
	(0.005)	(0.000)
Prob Hispanic	-0.075***	-0.001
	(0.014)	(0.002)
Prob Black	-0.073***	-0.002
	(0.012)	(0.002)
Prob White	-0.037***	-0.001
	(0.013)	(0.002)
Constant	0.349***	0.003
	(0.015)	(0.003)
Observations	67060	67060
Joint F-Test p-value	0	0.169

Note: This table reports regression results testing the random assignment of cases to arraigning ADAs, using the sample described in Table A.17. ADA leniency is estimated using data from other nonviolent misdemeanor cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports estimates from an OLS regression of nonprosecution on the variables listed and court-by-time fixed effects. Column (2) reports estimates from an OLS regression of ADA leniency on the variables listed and court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The p-value reported at the bottom of Columns (1) and (2) is for an F-test of the joint significance of the variables listed with standard errors two-way clustered at the individual and ADA level. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.19: Monotonicity (with Demographic Variables)

	(1) Black	(2) Hispanic	(3) White	(4) Male	(5) Female	(6) Age≤25	(7) Age>25
Not Prosecuted	0.65***	0.13*	0.54***	0.44***	0.51***	0.55***	0.74***
	(0.08)	(0.07)	(0.08)	(0.17)	(0.15)	(0.07)	(0.14)
Observations	55164	12274	25833	8949	20458	53550	13383
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	No	No	No	No	No	No	No

Note: This table reports first stage results by demographic group, as coded in the SCDAO data. The regressions are estimated on the sample as described in the "Admin Race Data" subsection of Table A.17. ADA leniency is estimated using all cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-time fixed effects. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$

Table A.20: First Stage (with Demographic Variables)

	(1)	(2)
ADA Leniency	0.59***	0.53***
	(0.07)	(0.07)
Observations	66612	66612
Court x Time FE	Yes	Yes
Case/Def Covariates	No	Yes
Mean Not Prosecuted	0.203	
First Stage F-Stat	63.19	55.90

Note: This table reports first stage results via a linear probability model where the outcome variable is nonprosecution, for the sample as described in Table A.17. ADA leniency is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. Column (1) reports results controlling for our full set of court-by-time fixed effects. Column (2) adds defendant and case covariates: number of counts; number of misdemeanor counts; number of serious misdemeanor counts; whether the defendant had a prior misdemeanor conviction within the past year; whether the defendant had a prior felony conviction within the past year; indicators for whether the defendant faces charges for a disorder/theft, motor vehicle, drug, or other offense; defendant citizenship status; gender; binned age; and predicted race/ethnicity. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Robust first-stage F reported (which is equivalent to the effective F-statistic of Montiel Olea and Pflueger (2013) in this case of a single instrument). $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.21: Probability of a Subsequent Criminal Complaint (With Demographic Variables)

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: Criminal Complaint Within 2 Years				
Not Prosecuted	-0.14***	-0.10***	-0.35***	-0.27**
	(0.01)	(0.01)	(0.11)	(0.11)
			[-0.56, -0.13]	[-0.48, -0.04]
Mean Dep Var Prosecuted	0.37			
Mean Dep Var Prosecuted Compliers	0.55			
Panel B: Misdemeanor Complaint Within 2 Years				
Not Prosecuted	-0.08***	-0.06***	-0.24***	-0.22**
	(0.00)	(0.00)	(0.09)	(0.09)
			[-0.42, -0.06]	[-0.41, -0.03]
Mean Dep Var Prosecuted	0.25			
Mean Dep Var Prosecuted Compliers	0.38			
Panel C: Felony Complaint Within 2 Years				
Not Prosecuted	-0.06***	-0.04***	-0.10*	-0.05
	(0.00)	(0.00)	(0.06)	(0.07)
			[-0.23, 0.03]	[-0.18, 0.10]
Mean Dep Var Prosecuted	0.13			
Mean Dep Var Prosecuted Compliers	0.17			
Observations	66612	66612	66612	66612
Court x Time FE	Yes	Yes	Yes	Yes
Case/Def Covariates	No	Yes	No	Yes

Note: This table reports OLS and two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent complaint within two years post-arraignment, using the sample as described in Table A.17. The dependent variables are identified in the panel headings. Each panel reports the mean of the dependent variable for all prosecuted defendants, and for prosecuted defendants within the set of compliers. See Appendix C.3 for details on the calculation of mean outcomes among prosecuted compliers. Two-stage least squares models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-time fixed effects. Where indicated they also include case and defendant covariates, which in this case include binned age and gender (as reported in the SCDAO data) and predicted race (see foonote 36. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses in Columns (1)-(4). For the IV estimates, Anderson-Rubin confidence intervals are shown in brackets. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

Table A.22: Heterogeneous Effects: By Demographic Group

	Male		Female		Age 18-23		Age 24-31	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Not Prosecuted	-0.10***	-0.30**	-0.08***	-0.26*	-0.10***	-0.16	-0.10***	-0.32
	(0.01)	(0.12)	(0.01)	(0.15)	(0.01)	(0.18)	(0.01)	(0.22)
Observations	53550	53550	13383	13383	12337	12337	16413	16413
	Age 31-40		Age $\geq$ 40		Pred White		Pred Black	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Not Prosecuted	-0.09***	-0.43	-0.09***	-0.28*	-0.10***	-0.28	-0.09***	-0.23
	(0.01)	(0.27)	(0.01)	(0.16)	(0.01)	(0.17)	(0.01)	(0.14)
Observations	14493	14493	20390	20390	24376	24376	26759	26759
	Pred Hisp		White		Black		Hisp	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Not Prosecuted	-0.09***	-0.26	-0.09***	-0.13	-0.09***	-0.24	-0.09***	-0.04
	(0.01)	(0.24)	(0.01)	(0.23)	(0.01)	(0.19)	(0.01)	(0.52)
Observations	14323	14323	20447	20447	25824	25824	8943	8943
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within two years, for the demographic groups specified in the column headings. The models instrument for nonprosecution using an ADA leniency measure that is estimated using data from other cases assigned to an arraigning ADA following the procedure described in the text. All specifications control for court-by-time fixed effects and case/defendant covariates (excluding race/ethnicity, age, and gender). Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. The groups White, Black, and Hispanic are based on race/ethnicity data as coded by SCDAO. Predicted White, Predicted Black, and Predicted Hispanic groups are based on race probability variables, described in Section 6.1. We label someone as "Predicted White" if the probability estimates suggest they are most likely to be white, "Predicted Black" if they are most likely to be Black, and "Predicted Hispanic" if they are most likely to be Hispanic. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.23: Missing ADA at Arraignment

	Not Prosecuted		Prosecuted	
	(1) Missing ADA	(2) Not Missing ADA	(3) Missing ADA	(4) Not Missing ADA
Outcomes:				
Criminal Complaint Within 2 Years	0.21	0.22	0.36	0.37
Prosecution Within 2 Years	0.15	0.16	0.32	0.34
DCJIS Record Within 2 Years	0.14	0.15	0.29	0.31
Baseline:				
Number Counts	1.62	1.58	1.71	1.75
Number Misdemeanor Counts	1.17	1.14	1.41	1.37
Number of Serious Misdemeanor Counts	0.29	0.29	0.72	0.65
Misd Conviction within Past Year	0.03	0.03	0.11	0.10
Felony Conviction within Past Year	0.01	0.01	0.06	0.05
Citizen	0.90	0.85	0.82	0.74
Disorderly/Theft	0.15	0.19	0.27	0.31
Motor Vehicle	0.64	0.63	0.44	0.33
Drug	0.02	0.03	0.13	0.18
Observations	42867	14184	97540	54748
Proportion Missing ADA	0.751		0.640	

Note: This table reports summary statistics for the samples of nonviolent misdemeanor cases meeting all other sample criteria that do and do not have information on the identity of the arraigning ADA, as indicated by the column headers.

Table A.24: Proportion of Samples Missing ADA at Arraignment

	(1) All	(2) Prosecuted	(3) Not Prosecuted
Main Sample	0.67	0.64	0.75
Imputation 1	0.58	0.56	0.63
Imputation 2	0.41	0.40	0.45
Imputation 3	0.31	0.31	0.31
Imputation 4	0.24	0.24	0.22

Note: This table reports the proportions of our main estimation sample and of each imputation sample (as described in the text) that are missing arraigning ADA information. See Section 6.2.

Table A.25: 2SLS Results with ADA Imputation Samples

	(1) Main (Court x Week FE)	(2) Imputation 1	(3) Imputation 2	(4) Imputation 3	(5) Imputation 4
Not Prosecuted	-0.531***	-0.501***	-0.652***	-0.553***	-0.535***
	(0.161)	(0.149)	(0.188)	(0.152)	(0.166)
	[-0.981, -0.233]	[-0.882, -0.220]	[-1.250, -0.339]	[-0.982, -0.291]	[-1.015, -0.250]
Observations	67123	85433	113148	134820	149185
Court x Time FE	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes
Mean Dep Var Prosecuted	0.371	0.371	0.369	0.370	0.371
Randomization p	0.491	0.316	0.313	0.383	0.623
First-Stage Coef	0.429	0.424	0.340	0.350	0.322
First-Stage F	15.20	18.85	13.81	18.40	16.82

Note: This table reports two-stage least squares estimates of the impact of nonprosecution on the probability of a subsequent criminal complaint within two years, for alternative samples of cases for which ADA assignment has been imputed (see Section 6.2 for details on the imputation samples). All models instrument for nonprosecution using our main ADA leniency measure, estimated using only cases assigned to an observed arraigning ADA following the procedure described in the text. All specifications include all covariates and court-by-week and court-by-day-of-week fixed effects. We use court-by-week rather than court-by-month fixed effects because the imputations are performed within a court-week (or a court-day; see text). Column (1) reports estimates for our primary estimation sample, using court-by-week (and court-by-day of week) fixed effects (and dropping singletons in court-by-week groups). Columns (2) - (5) progressively expand the sample of cases to include those cases where arraigning ADA assignment has been imputed following the strategies described in the text. Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

Table A.26: Sample Share by Compliance Type

	Linear			Local Linear		
	1%	1.5%	2%	1%	1.5%	2%
Compliers	0.10	0.09	0.08	0.09	0.09	0.08
Always Takers	0.18	0.18	0.19	0.18	0.19	0.19
Never Takers	0.73	0.73	0.73	0.72	0.73	0.73

Note: This table estimates the shares of our sample that are compliers, always-takers, and never-takers. The fraction of always-takers, π_a, is estimated by the share of the defendants who are not prosecuted by the least lenient ADA; the fraction of never-takers, π_n, by the share prosecuted by the most lenient ADA; and compliers as $1 - \pi_a - \pi_n$. Least lenient ADAs are defined by being at the 1st, 1.5, or 2nd percentile of the residualized ADA leniency distribution, and most lenient are defined as being at the 99, 98.5, or 98th percentile. The first three columns use a linear specification of our first stage as in equation 3; the latter three use a local linear specification.

Table A.27: Characteristics of Marginal Defendants

	(1) $\Pr[X = x]$	(2) $\Pr[X = x \mid Complier]$	(3) Ratio
Counts = 1	0.55	0.55	1.00
Counts > 1	0.45	0.42	0.92
Misd Counts = 1	0.72	0.74	1.03
Misd Counts > 1	0.28	0.22	0.78
No Serious Misd	0.53	0.57	1.07
Serious Misd	0.47	0.37	0.79
No Misd Conviction 1 Yr Prior	0.94	0.97	1.03
Misd Conviction 1 Yr Prior	0.06	0.03	0.57
No Felony Conviction 1 Yr Prior	0.97	0.98	1.01
Felony Conviction 1 Yr Prior	0.03	0.02	0.68
Not Citizen	0.23	0.10	0.43
Citizen	0.77	0.87	1.14
Any Disorderly/Theft Charge	0.33	0.38	1.16
Any Motor Vehicle Charge	0.45	0.41	0.91
Any Drug Charge	0.21	0.04	0.19
Any Other Charge	0.12	0.08	0.64
Demographics (within demographic sample)			
Age $\leq$ 23	0.23	0.29	1.26
Age 24-30	0.25	0.22	0.87
Age 31-40	0.22	0.18	0.82
Age $\geq$ 41	0.31	0.30	0.98
Male	0.79	0.73	0.92
Female	0.20	0.23	1.18
(Predicted) Black	0.40	0.42	1.05
(Predicted) White	0.36	0.34	0.94
(Predicted) Hispanic	0.22	0.22	1.01

Note: This table describes the observable characteristics of the complier sample, relative to the full sample. Column (1) shows the probability that an individual has a given characteristic in the full analysis sample. Column (2) shows the probability that someone in the complier group has that characteristic. Column (3) shows the ratio of the two (Column (2) divided by Column (1)). The estimates in Column (2) are constructed by calculating the shares of compliers within these various subsamples. The complier share calculations here rely on a linear first-stage estimation and a 1% cut-off to define ADA leniency. See Section 7 for more detail.

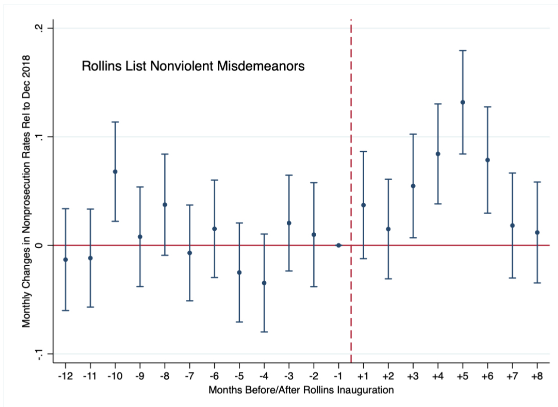
Table A.28: Reweighted OLS

	Main			With Demographics		
	(1) OLS	(2) Decile Weigts	(3) Quart x Prev. Charge Wts	(4) OLS	(5) Decile Weights	(6) Quart x Prev. Charge Wgts
Not Prosecuted	-0.104***	-0.102***	-0.101***	-0.099***	-0.094***	-0.093***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	67553	67553	67553	66612	66612	66612
Court x Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	No	No	No	Yes	Yes	Yes
Complier Weights	No	Yes	Yes	No	Yes	Yes

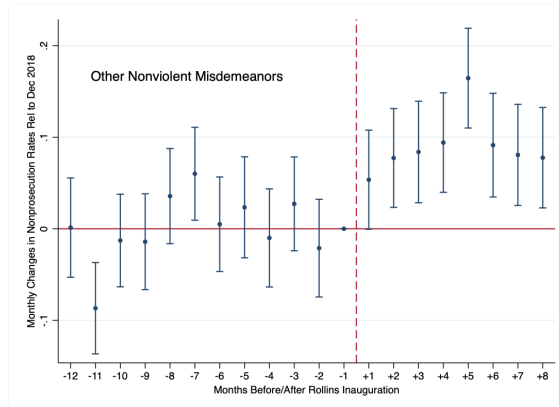
Note: Column (1) recreates our main OLS estimates, and Column (4) recreates our main OLS estimates in the sample of defendants for whom we have both age and gender. Columns (2)-(3) and (5)-(6) reweight those OLS estimates by splitting the sample into mutually exclusive subgroups, calculating the shares of compliers in each subgroup (as in Table A.27), and using the share of compliers relative to the share of the estimation sample in each subgroup as a weight. Columns (2) and (5) split the sample into 10 mutually exclusive subgroups based on deciles of the predicted probability of nonprosecution estimated with all available covariates. Columns (3) and (6) split the sample into 8 subgroups by quartiles of this propensity score and by whether the defendant has a previous complaint, an important source of hetereogeneity. ***$p < 0.01$,**$p < 0.05$,*$p < 0.10$.

B Presumption of Nonprosecution

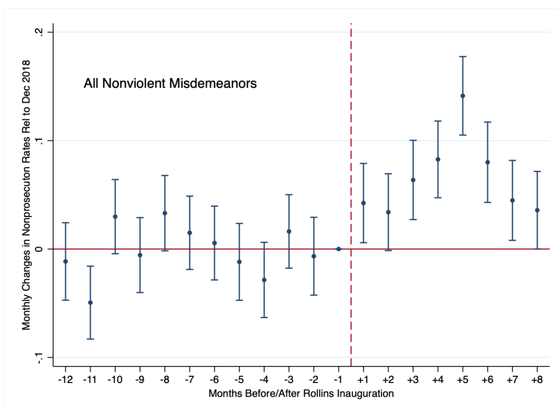
Figure B.1: Effects of Rollins Inauguration on Nonprosecution



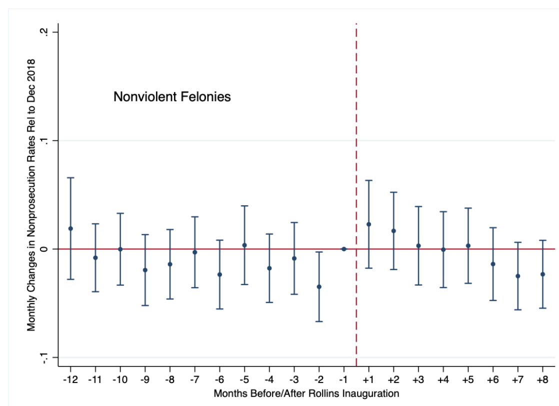
(a) Rollins List



(b) Other Nonviolent Misdemeanors



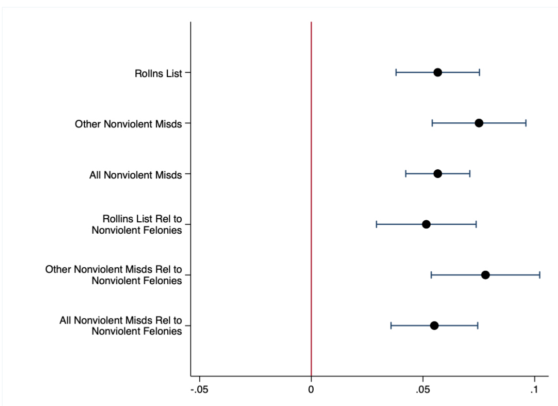
(c) All Nonviolent Misdemeanors



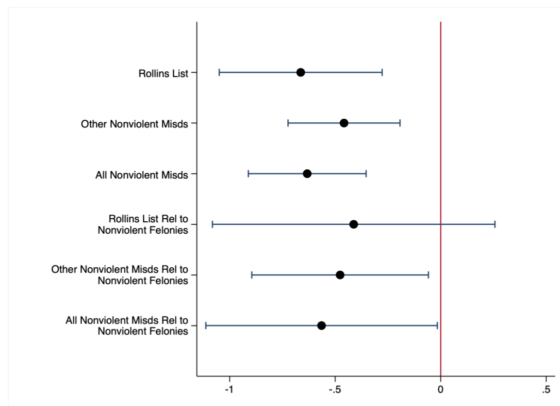
(d) Nonviolent Felonies

Note: This figure reports monthly event study estimates of the effects of the inauguration on January 2, 2019 of Rachael Rollins as District Attorney of Suffolk County. Each panel reports OLS estimates of monthly changes in nonprosecution rates for different categories of cases, relative to the baseline month of December 2018, between January 1, 2018 and September 1, 2019. All models include court and day of week fixed effects and all case-level covariates used throughout the paper. Robust standard errors clustered on defendant; 90% confidence intervals.

Figure B.2: Effects of Rollins Inauguration on Nonprosecution
and Subsequent Criminal Complaints



(a) Nonprosecution



(b) Subsequent Complaint

Note: This figure reports the effects of the inauguration on January 2, 2019 of Rachael Rollins as District Attorney of Suffolk County, for cases initiated between January 1, 2018 and September 1, 2019. Panel A reports OLS estimates of the average effects of the Rollins inauguration on nonprosecution rates for different categories of cases; the bottom three coefficients in Panel A report these estimates relative to nonprosecution rates for nonviolent felony cases. Panel B reports 2SLS estimates of the average effects of nonprosecution on the likelihood that a defendant is issued a new criminal complaint within one year of the current case, using the Rollins inauguration as an instrument for nonprosecution; the bottom three coefficients in Panel B report these 2SLS estimates relative to nonviolent felony cases. All models include court, month, and day of week fixed effects and all case-level covariates used throughout the paper. Robust standard errors clustered on defendant; 95% confidence intervals.

Figure B.3: Effects of Rollins Inauguration on Reported Crimes

(a) Property Damage

(b) Theft & Fraud

(c) Disorder

(d) Drug

(e) Other

Note: This figure reports the effects of the inauguration on January 2, 2019 of Rachael Rollins as District Attorney of Suffolk County, on the number of crime incidents reported to the Boston Police Department between January 1, 2017 and February 29, 2020. Each subfigure is a coefficient plot showing the effect on a particular category of reported crime. The y-axes show number of reported incidents; the x-axes show time (month) relative to January 2019. The coefficient for *t-7* includes months June 2018 and earlier; *t+8* includes months September 2019 and later. All regressions include month-of-year fixed effects. Robust standard errors; 95% confidence intervals.

C Technical Appendix

C.1 Comparisons to Alternative Instrument Estimation Strategies

Our main instrument is a residualized leave-out mean leniency measure that is estimated from the other nonviolent misdemeanor cases that an ADA has arraigned. For our main analyses we proceed by implementing this instrument in an 'as-if just-identified' manner: we report robust F-statistics that do not adjust for the fact that the instrument is estimated (although we also do not use these F-statistics directly in a threshold test), and we conduct inference in the second stage using confidence intervals based on inversion of the Anderson-Rubin test, which have correct size and optimal power even when instruments are weak in just-identified models (Anderson and Rubin, 1949; Andrews, Stock and Sun, 2019). Performing the estimation in this way is standard (Doyle Jr, 2007; French and Song, 2014; Dahl, Kostøl and Mogstad, 2014; Dobbie, Goldin and Yang, 2018; Bhuller et al., 2020) and has some attractive properties. We have 315 ADAs in our sample, each of whom could serve as a potential instrument. With this many instruments, estimation using the ADA dummies can suffer from bias from many (potentially weak) instruments (Bekker, 1994; Bound, Jaeger and Baker, 1995; Hausman et al., 2012). We also have many fixed effects (court-by-month and court-by-day of week) in the covariate set, necessary to identify the set of cases for which ADAs are as-if randomly assigned, which can also cause bias in jackknife instrumental variable estimators (JIVE) (Kolesár, 2013). It is also clear in the just-identified case how to handle inference in the second stage that is robust to (potential) weak instrument issues (Andrews, Stock and Sun, 2019). The continuous instrument also allows for the estimation of marginal treatment effects (Heckman and Vytlacil, 2005).

While this estimation strategy is convenient for the reasons mentioned above, it does not take into account that the instrument itself is constructed. In this subsection we explore the robustness of our main estimates to alternative estimation strategies in this setting. One alternative approach is to estimate our 2SLS model using the full set of 315 ADA dummy variables as instruments in the first stage. These results are shown in Table C.1, Column (2) (Column (1) repeats our main 2SLS results using the residualized leave-out mean leniency as an instrument). The estimated coefficient is -0.19, smaller in absolute value and closer to OLS than our main leave-out mean 2SLS estimate, which is unsurprising given that the bias from weak instruments moves estimates closer to the OLS estimate. Several strategies are suggested when IV estimates suffer from bias from many (weak) instruments. In Column (3) we estimate a limited information maximum likelihood (LIML) model with all the dummies as instruments (Bekker, 1994; Chao and Swanson, 2005; Angrist and Frandsen, 2020). The coefficient in this model is -0.27, closer to our main leave-out mean 2SLS estimate than the estimate in Column (2) with all the ADA dummy variables. LIML however is not consistent with heteroskedastic errors or with heterogeneous treatment effects (Hausman et al., 2012; Kolesár, 2013), which is the motivation for the UJIVE estimator we also use. In Column (4) we estimate the unbiased JIVE (UJIVE) estimator of Kolesár (2013).[49] JIVE estimators are generally suggested when the number of instruments is large (Angrist, Imbens and Krueger, 1999), although they can be biased with many covariates (Kolesár, 2013). The UJIVE estimator is consistent (for a convex combination of LATE estimates) with a large number of covariates. Here the coefficient is -0.26, again closer to our main leave-out mean 2SLS estimate than the estimate in Column (2) with all the ADA dummy variables. Another way to handle the potential bias from many (weak) instruments is to reduce the number of instruments by using lasso to pick the most informative ADA dummies in a 2SLS regression. We do this in Column (5) using a post-lasso first stage via the procedures of Belloni, Chernozhukov and Hansen (2014).[50]

In each case the coefficients we estimate with the alternative strategies are negative, large, and statistically significant, implying that nonprosecution decreases criminal complaints within two years post-arraignment between 33-70% relative to prosecuted compliers. For the most part, given the estimated standard errors, we cannot reject the null that these coefficients are the same as our main 2SLS estimate.[51]

[49] We thank Sam Norris for sharing the code to calculate the UJIVE instrument used in Norris, Pecenco and Weaver, 2020.

[50] In practice implemented via the user-written package `ivlasso` in Stata (Ahrens, Hansen and Schaffer, 2019), using the post-lasso results and using the `ivlasso` defaults with a plug-in penalty. The procedure retains three out of 315 instruments (similarly in the Angrist and Frandsen (2020) implementation of the plug-in penalty, lasso retains two instruments out of 180 in a re-estimation of the Angrist and Keueger (1991) QOB study). We also implemented a version of `ivlasso` with a cross-validated penalty; see Angrist and Frandsen (2020) for details on implementation. The algorithm with the CV penalty chooses more instruments, namely 173 out of the 315 in our case. The estimated post-lasso coefficient is smaller in absolute value (-0.22, se=0.049). The simulation results of Angrist, Imbens and Krueger (1999) and Belloni et al. (2012) imply that the plug-in penalty will have less bias although will also be less precise than the CV penalty estimates.

[51] The standard errors reported here for 2SLS or LIML using all the dummy variables have not been adjusted to take into account the potential for weak instruments; other inference strategies may imply larger confidence intervals.

Table C.1: Different IV Strategies

	(1) Main Leave-out Mean	(2) 2SLS All Dummies	(3) LIML All Dummies	(4) UJIVE	(5) lasso
Not Prosecuted	-0.34***	-0.19***	-0.27***	-0.26***	-0.40***
	(0.11)	(0.06)	(0.10)	(0.08)	(0.11)
Observations	67553	67553	67553	66809	67553
Court x Time FE	Yes	Yes	Yes	Yes	Yes
Case/Def Covariates	Yes	Yes	Yes	Yes	Yes
Mean Not Prosecuted	0.371				
Mean Not Prosecuted Compliers	0.570				

Note: This table reports two-stage least squares estimates using various estimation strategies for the instrument, as indicated in the column headers. All specifications control for court-by-time fixed effects and case/defendant covariates. The OLS estimate for this specification can be found in Table 4 Column (2), and is -0.10 (se=0.01). Robust standard errors two-way clustered at the individual and ADA level are reported in parentheses. Column (1) repeats our main 2SLS estimates using the residualized leave-out mean leniency measure with covariates (see Table 4, Column (4)). In Column (2) we use all 315 ADA dummy variables directly as instruments in the first stage. Column (3) uses limited information maximum likelihood estimation with all of the dummies as instruments. Column (4) uses the UJIVE estimator of Kolesár (2013). Column (5) uses post-lasso from Belloni, Chernozhukov and Hansen (2014) to choose the most informative ADA dummy variables; the algorithm chooses three of the ADA dummies as instruments. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.10$.

C.2 MTE Estimation

Re-orienting our framework to the potential outcome framework, let $Y_i(1)$ be the defendant's outcome Y if not prosecuted ($D_i = 1$) and $Y_i(0)$ the defendant's outcome if prosecuted ($D_i = 0$). An ADA makes a decision to prosecute or not prosecute a defendant based on characteristics both observable to the econometrician, X_i, and unobservable to the econometrician, ν_i. Define the latent propensity to be not prosecuted as: $D_i^* = \mu_D(Z_i, X_i) + \nu_i$. $D_i = 1$, or the defendant is not prosecuted, if $D_i^* = \mu_D(Z_i, X_i) + \nu_i \geq 0 \implies \mu_D(Z_i, X_i) \geq \nu_i \implies F_\nu(\mu_D(Z_i, X_i)) \geq F_\nu(\nu_i)$ and prosecuted otherwise, where F_ν is the (unknown) cumulative distribution function of ν. $F_\nu(\mu_D(Z_i, X_i)) = P(Z_i, X_i)$ is the propensity score: the probability of nonprosecution conditional on observables, X_i, and ADA leniency, Z_i. Call $F_\nu(\mu_i) = U_D$ quantiles of the distribution of the unobserved propensity to be not prosecuted. The marginal treatment effect is then defined as the treatment effect at a particular value of the unobservable propensity to be not prosecuted: $E(Y_i(1) - Y_i(0)|X_i = x, U_{Di} = u_D)$, that is the treatment effect for individuals on the margin of being not prosecuted when $P(Z_i, X_i) = u_D$. It can be estimated as the derivative of the average outcome conditional on X and $P(Z_i, X_i) = u_D$ with respect to the propensity score.

C.3 Understanding Compliers

To calculate the shares of compliers, never-takers, and always-takers, we use the insights of Abadie (2003) and Dahl, Kostøl and Mogstad (2014), and applied by Dobbie, Goldin and Yang (2018) and Bhuller et al. (2020).

Always-takers are defendants who would always be not prosecuted regardless of the ADA assigned to their case. Given our monotonicity and conditional independence assumptions, the fraction of always-takers can be calculated by the share of defendants not prosecuted by the most strict ADA(s):

$$\pi_a = Pr(\text{Not Prosecuted}_i = 1|Z_i = \underline{z}) = Pr(\text{Not Prosecuted}_i(\overline{z}) = \text{Not Prosecuted}_i(\underline{z}) = 1) \tag{4}$$

where $\overline{z}$ represents a maximum value of the ADA instrument (the most lenient ADA) and $\underline{z}$ represents a minimum value of the instrument (the most strict ADA).

Similarly, never-takers are defendants who would never be not prosecuted (always be prosecuted). We can estimate their fraction by the share of defendants who are prosecuted by the most lenient ADA(s):

$$\pi_n = Pr(\text{Not Prosecuted}_i = 0 | Z_i = \overline{z}) = Pr(\text{Not Prosecuted}_i(\overline{z}) = \text{Not Prosecuted}_i(\underline{z}) = 0) \tag{5}$$

Finally, compliers are defendants whose prosecution decisions would have been different had their case been assigned to the most lenient instead of the most strict ADA:

$$\pi_c = Pr(\text{Not Prosecuted}_i = 1 | Z_i = \overline{z}) - Pr(\text{Not Prosecuted}_i = 1 | Z_i = \underline{z}) = Pr(\text{Not Prosecuted}_i(\overline{z}) > \text{Not Prosecuted}_i(\underline{z})) \tag{6}$$

We can calculate this as $1 - \pi_a - \pi_n$. Under a linear specification of the first stage Equation 3, we can recover π_c as $\alpha_1(\overline{z} - \underline{z})$, π_a as $\alpha_0 + \hat{\alpha_1}\underline{z}$, and π_n as $1 - \alpha_0 - \hat{\alpha_1}\overline{z}$, where α_0 and α_1 are the estimated first stage coefficients. We also estimate these under more flexible local linear estimations of our first stage.

With these shares, we can then calculate average characteristics for complier defendants who were prosecuted: $E(Y_i(0) | \text{Not Prosecuted}_i(\overline{z}) > \text{Not Prosecuted}_i(\underline{z}))$. Among the prosecuted, average outcomes for defendants who were assigned to lenient ADAs are average outcomes for the never-takers:

$$E(Y_i | \text{Not Prosecuted}_i = 0, z_i = \overline{z}) = E(Y_i(0) | \text{Not Prosecuted}_i(\overline{z}) = \text{Not Prosecuted}_i(\underline{z}) = 0) \tag{7}$$

Among the prosecuted, outcomes for defendants who were assigned to strict ADAs are a weighted average of outcomes for compliers and never-takers, where the weights are their shares in the population:

$$\begin{aligned} E(Y_i | NotProsecuted_i = 0, z_i = \underline{z}) = {} & \frac{\pi_c}{\pi_c + \pi_n} E(Y_i(0) | \text{Not Prosecuted}_i(\overline{z}) > \text{Not Prosecuted}_i(\underline{z})) \\ & + \frac{\pi_n}{\pi_c + \pi_n} E(Y_i(0) | \text{Not Prosecuted}_i(\overline{z}) = \text{Not Prosecuted}_i(\underline{z}) = 0) \end{aligned} \tag{8}$$

Plugging Equation 7 into Equation 8, we can calculate average outcomes for compliers among the prosecuted for any outcome Y_i as:

$$
\begin{aligned}
E(Y_i(0)|\text{Not Prosecuted}_i(\overline{z}) > \text{Not Prosecuted}_i(\underline{z})) = {} & \frac{\pi_c + \pi_n}{\pi_c} E(Y_i|\text{Not Prosecuted}_i = 0, z_i = \underline{z}) \\
& - \frac{\pi_n}{\pi_c} E(Y_i|\text{Not Prosecuted}_i = 0, z_i = \overline{z})
\end{aligned}
\tag{9}
$$