Diversion in the Criminal Justice System

Michael Mueller-Smith* Kevin T. Schnepel†

Draft date: January 17, 2019

Abstract

This paper provides the first causal estimates on the popular, cost-saving practice of diversion in the criminal justice system, an intervention that provides offenders with a second chance to avoid a criminal record. We exploit two natural experiments in Harris County, Texas where first-time felony defendants faced abrupt changes in the probability of diversion. Using administrative data and regression discontinuity methods, we find robust evidence across both experiments that diversion cuts reoffending rates in half (-32 p.p.) and grows quarterly employment rates by 53 percent (+18 p.p.) over 10 years. The change in trajectory persists even 20 years out and is concentrated among young black men. An investigation of mechanisms indicates that stigma associated with a felony conviction plays a key role in generating these results. Other possible mechanisms including changes in incarceration, universal adjustments in policy or practice, and differences in criminal processing are ruled out empirically.

Keywords: felony records, criminal justice, race, recidivism, labor market

JEL classification codes: J24, K14, K42

*mgms@umich.edu, Department of Economics, University of Michigan

†kevin_schnepel@sfu.ca, Department of Economics, Simon Fraser University & School of Economics, The University of Sydney

Acknowledgements: We thank Amanda Agan, Martha Bailey, Steve Billings, John Bound, Charlie Brown, Jennifer Doleac, Phil Cook, Ben Hansen, Sara Heller, David Jacks, Jens Ludwig, Justin McCrary, Aurelie Ouss, Emily Owens, Krishna Pendakur, Becky Pettit, David Phillips, Steve Raphael, Jeffrey Smith, Mel Stephens, Glen Waddell, and Abigail Wozniak for helpful comments and suggestions; conference participants at the 2015 Southern Economic Association Annual Meetings, 2016 NBER Summer Institute, the 8th Transatlantic Workshop on the Economics of Crime, the 11th Annual Conference on Empirical Legal Studies, the 22nd Annual Meeting of the Society of Labor Economics, and the 2017 UM-MSU-UWO Labor Day Conference; and numerous seminar participants. We also thank Christopher King, Greg Kumpton and Patty Rodriguez at the Ray Marshall Center, and Serena Renner for editorial assistance. This research was supported in part by NICHD center and training grants to the Population Studies Center at the University of Michigan (R24 HD041028 and T32 HD007339).

The U.S. criminal justice caseload, whether defined in terms of arrestees, defendants, or detainees, has expanded dramatically over the last half century. As of 2010, more than 8 percent of the adult population in the U.S. had a prior felony conviction, compared with just 3 percent in 1980 (Shannon et al. 2016). As caseload growth has challenged system capacity, legal and procedural strategies have emerged to focus efforts on the most serious offenders. Often this means showing greater leniency toward individuals posing minimal risk to conserve resources. One popular practice is a class of interventions referred to as *diversion* wherein public officials choose to pause, terminate, or divert someone's progression through the justice system. By 2013, thousands of jurisdictions across the U.S. had implemented some form of diversion whether at the arrest, prosecution, or institutional stage (Center for Health and Justice 2013). While such programs differ in terms of timing, eligibility, and requirements for successful completion, they often provide first-time offenders a second chance to avoid the lifelong stigma of a criminal conviction record.

Our current understanding about how diversion impacts future behavior is limited.[1] Recent findings document a range of ways in which the criminal justice system may harm employment, education, and reoffending outcomes (Dobbie et al. 2018, Stevenson 2018, Mueller-Smith 2015, Aizer and Doyle 2015, Di Tella and Schargrodsky 2013, Raphael 2014, Lovenheim and Owens 2014, Finlay 2009, Pager 2008, 2003). To the extent that diversion protects individuals from the negative consequences of the criminal justice system, it may improve an offender's self-sufficiency and minimize further criminal activity. Recent evidence on the effectiveness of specific deterrence (Bhuller et al. 2016, Hansen 2015, Owens 2009)—the prediction that a past experience of a more severe sanction can decrease the probability of future offending—however, suggests impacts in the opposite direction. These ambiguous theoretical predictions demonstrate the need for convincing empirical evidence.

In this paper, we present the first causal evidence on the effect of diversion in the criminal justice system on recidivism and labor market outcomes. To accomplish this, we use a regression discontinuity (RD) design to evaluate two discrete changes in the use of diversion by criminal courts in Harris County, Texas. The first discontinuity follows a Texas penal code reform in 1994 that reduced diversion opportunities for offenders charged with certain reclassified drug and property offenses. The second discontinuity follows the unexpected failure of a 2007 ballot initiative that resulted in an immediate increase in the diversion rate for low-risk defendants. These policy shifts are attractive from a research perspective due to

[1] Chiricos, Barrick, Bales and Bontrager (2007) find that two-year recidivism rates among offenders with a court deferral agreement are significantly lower compared with convicted offenders using data from Florida. While the authors control for offender and county characteristics, estimates may suffer from bias since variation in conviction status is endogenous.

their rapid implementation, which caused defendants to experience abruptly different case decisions depending on whether they were charged before or after the change in policy.

Our findings indicate that diversion substantially improves behavioral outcomes over a 10-year follow-up period among first-time felony defendants. The probability of any future conviction declines by 48 percent (-32 percentage points) and the total number of future convictions falls by 66 percent (-1.45 new convictions). Quarterly employment rates improve by 53 percent (18 percentage points) and total earnings grow by 64 percent ($61,540).[2] In the 1994 sample, we can show that these positive effects persist and expand even 20 years out, suggesting that diversion permanently changes one's lifetime trajectory. This finding likely relates to the composition of our main study sample: individuals charged with, but not yet convicted of, their first felony offense.

In the local context of our study, criminal court diversion is typically administered as a *deferred adjudication of guilt*, meaning that the defendant avoids a formal felony conviction by undergoing a period of community supervision that operates just like felony probation. Unlike an expungement approach—a legal process resulting in the removal of an existing conviction from record—the defendant has never been legally convicted, which may help him or her avoid the many restrictions and penalties associated with prior felony convictions (National Reeentry Resource Center 2018). In contrast to other potential court interventions like specialized drug or veterans courts, deferred adjudication requires minimal additional public resources since it relies on an established community supervision program (e.g. probation).

While diversion can be used to reduce the incidence of incarceration, our evidence indicates that is not the case for our marginal defendants. In 1994, defendants who are not diverted as a consequence of the penal code reform are instead convicted and sentenced to community supervision. Consistent with this, we observe no differences in time served in prison or jail during the initial 5 years of follow-up data. In 2007, we do observe a change in incarceration sentences associated with the uptick in diversion, but this does not translate into overall differences in actual time served during our primary outcome windows. A detailed, week-by-week analysis of time-served in the first follow-up year shows a short-lived temporal displacement of incarceration in 2007 associated with the change in sentences; however, cumulative exposure is equalized within 19 weeks and thus differential exposure to incarceration does not appear to explain our estimated impacts of diversion on reoffending and labor market outcomes.

Instead, we believe the results are driven largely by the protective effect of avoiding a

[2]Notably, the two independent experiments yield estimated treatment effects that are highly consistent, often within several percentage points.

felony conviction record. Nearly half of the defendants who receive deferrals successfully complete their term of community supervision, receive a case dismissal, and maintain a clean criminal record. Moreover, approximately one third of the marginal defendants from the 2007 experiment are diverted with an outright case dismissal, guaranteeing no felony record at least initially. We re-run our analysis on repeat offenders, who are already saddled with a felony record regardless of diversion status, and fail to replicate our findings from the first-time offender caseload. These results build on an extensive literature, largely centered around audit and correspondence research designs, that demonstrates the negative repercussions of criminal convictions in the U.S.

We believe using diversion to study the effect of felony convictions overcomes three important limitations of prior research. First, using a natural experiment instead of an audit design allows our study to generate estimates based on final labor market outcomes rather than callback rates and net of any sorting behavior to non-discriminating employers present in the labor market (Becker 1971). Second, in contrast to work considering ex-post interventions like record clearing or expungement (Selbin et al. 2017), research evaluating diversion can measure the change in trajectory of future outcomes caused by disposition status before potential irreversible damage has accrued to labor market experience or compliance with the law. Finally, unlike research examining the impact of offense reclassification (e.g. reducing a felony charge to a misdemeanor) (Bird et al. 2018), this study examines variation in conviction status that holds the effective level and type of correctional supervision requirements constant.

Although we believe our analysis advances the literature on felony convictions, we acknowledge the possibility that other mechanisms may contribute to our estimated impacts. For instance, among those with a deferred adjudication, re-sentencing deterrence for one's original offense—similar to the mechanism observed in Drago et al. (2009)—may also discourage criminal activity during the initial follow-up period. In our sample, successfully completed deferral agreements last 2.86 years on average, and while we observe significant impacts during this period, our effects persist well beyond these first few years that carry an additional incentive to refrain from criminal activity. As such, we cannot rule out re-sentencing deterrence, yet we believe it functions in a secondary capacity.

Another possible mechanism is that a prosecutor may be more likely to eschew new charges and/or convictions since revoking an existing deferral agreement may allow the prosecutor to impose a more immediate punishment that does not require the same evidentiary burden as a new conviction. This behavior could account for the lower conviction rates observed in our findings. Our results, however, are consistent across many different measures of recidivism including jail bookings (which proxy for arrests) and including non-technical probation

revocations, indicating that mechanical changes in criminal processing play a minimal role in practice.

Furthermore, we do not measure significant discontinuities in observable demographic characteristics, prior criminal histories, or the density of the criminal caseload, allaying concerns regarding endogenous sorting or sample imbalance. We conduct a number of robustness tests to demonstrate that our findings do not rely on any particular RD specification or bandwidth. We show that there are no changes in behavioral outcomes in response to the discontinuities for other parts of the felony caseload to ensure our results are not driven by caseload-wide policy changes unrelated to diversion. We demonstrate the robustness of labor market outcomes to specifications that account for missing earnings records as well as those that eliminate potential incapacitation effects arising from incarceration. Finally, we rule out artifacts of seasonality based on placebo exercises that apply our sample criteria and RD strategies to defendants charged around the same discontinuity dates shifted forward or back by one year to Sept. 1, 1993; Sept. 1, 1995; Nov. 7, 2006; and Nov. 7, 2008.

The largest treatment effects are measured for young black men with a misdemeanor record—a group that exhibits the highest likelihood of future criminal involvement. Because the criminal justice system disproportionately affects this demographic in the United States (Shannon et al. 2016), expanding the practice of diversion could help reduce the broad socioeconomic inequalities recently documented by Chetty et al. (2018).

The remainder of the paper is structured as follows: Section 2 describes the two discontinuities in diversion rates in Harris County, Texas that form the basis of our research design as well as describes the process of diversion studied in detail; Section 3 addresses the administrative datasets used to perform the analysis; Section 4 outlines our regression discontinuity empirical strategy; Section 5 presents our results and provides evidence that supports our identification assumptions; Section 6 discusses identification threats and robustness checks; Section 7 highlights mechanisms that explain our pattern of empirical results; and Section 8 provides concluding remarks. Two appendices (one print "A", one online "B") provide results from additional robustness exercises, alternative specifications, and subgroup analyses.

2 Institutional Background

2.1 The 1994 Penal Code Reform

In 1993, the Texas Legislature enacted a reform of its penal code that affected defendants who had committed offenses on or after Sept. 1, 1994.[3] The reform sought to "get smart" on incarceration through reducing sanctions for low-risk offenders while increasing time served before parole eligibility for aggravated violent offenders.

A major component of the reform was the creation of a probation-before-incarceration requirement for low-risk, first-time felony offenders. The specific legal wording of this provision, however, disincentivized prosecutors' use of diversion. At issue was the removal of the ability to threaten incarceration to ensure compliance with a diversion agreement, which prosecutors felt severely undermined its effectiveness. Under the new penal code, a second round of probation was required for defendants who violated their diversion agreements since they had not yet been legally convicted. The fact that the language of this provision weakened the credibility of diversion was raised by attorneys in October 1993, yet no action was taken to amend the statutory language before the law was implemented (Fabelo 1997).

Not all crimes were impacted by the probated incarceration requirement. The change affected only the lowest felony drug and property crimes, specifically possession of less than 1 gram of a controlled substance[4] and property offenses totaling less than $20,000 in damages. Diversion rates in the affected caseloads dropped significantly with the new penal code. In lieu of diversion, defendants were typically convicted and sentenced to probation as required by the statute. Both before and after the Sept. 1, 1994 cutoff, marginal defendants were subjected to community supervision; the main difference from the defendant's perspective was that before the cutoff, one could avoid a felony conviction whereas afterwards, a felony conviction was non-negotiable.

2.2 The 2007 Failed Ballot Initiative

During the 2000s, overcrowding in the Harris County Jail was a widespread concern. This local jail—which houses inmates with shorter sentences and serves several other functions

[3] Two pieces of legislation accomplished this overhaul: Senate Bills 1067 and 532. In general, prosecutors in Texas are required to file charges within 48 hours of arrest, which limited the opportunity for manipulation around the introduction of this penal code reform.

[4] Controlled substances covered under these statutes include crack cocaine and cocaine as well as heroin, methamphetamine, and other serious drugs. Marijuana possession was covered by a separate part of the penal code and was not impacted by this specific change.

including detention of individuals either waiting for trial or to be transferred to the state prison system—had up to 1,900 inmates sleeping on mattresses on the floor by 2005 (Hughes 2005). To address overcrowding, the county sought to expand capacity by 2,500 beds with $195 million to be raised through county bonds for construction of a new jail facility. Before issuing the bonds, the county first sought permission from local voters on the Nov. 6, 2007 ballot.[5]

A local campaign against the jail expansion combined with a large voter turnout led to a narrow and unexpected defeat of the initiative by a vote of 50.6 to 49.4 percent. This outcome was particularly surprising given that all the other local bonds were approved, and a $1 billion state-wide bond to expand state prison capacity was overwhelmingly approved (58.2 to 41.8). The local campaign against the jail expansion proposition suggested that the intended location of the new jail would be bad for local economic development and that existing infrastructure could be more efficiently used with less reliance on incarceration. However, some commentators explicitly placed the responsibility of the overcrowding problem on the courts in Harris County, citing an over-use of incarceration at the cost of taxpayer funds (Snyder 2007, Henson 2007a,b).

Most Harris County criminal courts exhibited a discernible change in their caseload outcomes immediately following the election. Diversion among first-time felony defendants increased by 18 percentage points, from less than 50 percent of the caseload to roughly 65 percent (Figure 2). It appears that the courts were responding to their critics given that both the district attorney and criminal court judges are publicly elected officials and would face the electorate in the coming year.[6]

The fact that the defeat of the ballot initiative was unanticipated limits the likelihood of systematic sorting. Furthermore, disposition dates are typically scheduled well in advance,

[5]The proposed jail expansion (Proposition 3) was part of a broader bond package put to voters in 2007 in response to the county's fast growing population. Together Harris County and the Port of Houston Authority added six local bond propositions to the Nov. 6, 2007 election ballot at a combined total of $880 million in potential bonds. The projects included upgrading roads and parks, expanding capacity at the port, building a new forensic lab, and constructing a new family law center.

[6]We have also considered an alternative theory for the 2007 change in conviction rates. On October 31, 2007, the Harris County District Attorney received a subpoena associated with the *Ibarras v Harris County Texas* civil lawsuit regarding improper coaching of defense witness testimony. On Nov 5, 2007, an inventory of his official e-mail was collected and some messages were improperly released to the public in December 2007 showing the DA was having a romantic affair with his executive secretary (Rogers et al. 2007). Additional emails were released on January 8, 2008, which showed the DA had exchanged dozens of pornographic, racist and political emails on his office computer (Rogers et al. 2008), which led to his resignation on February 15, 2008 (Lozano 2008). While the initiation of these events in the DA's office coincide with the 2007 cutoff we examine, it is unclear how they could lead to lasting changes in diversion practices in the courts extending out to 2009.

constraining the ability to sort. As shown in Section 5, it does not appear that widespread sorting occurred around the election. Instead, it appears that some defendants were simply lucky to have been scheduled to be disposed after the election rather than before.

2.3 Criminal Process of Deferred Adjudications in Texas

Following a felony charge in Harris County, a defendant can avoid a conviction through a deferred adjudication or a case dismissal. While some defendants have their cases dropped with the intent of diversion, deferred adjudication agreements are the most frequently used diversion option. Deferred adjudication is broadly available across less serious felony offenses,[7] and usually targets first-time felony defendants. Charges relating to aggravated violence or sexual conduct are typically ineligible (Tex. Code Crim. Proc. §42A.101(a)). While defendants can face charges for multiple offenses, it is very rare in this context that a defendant would be convicted of one charge and deferred for another. In fact, we only observe this situation for 0.002% of our estimation sample in 1994 (less than 50 individuals) and do not observe any such instances in 2007.

Deferred defendants are required to undergo community supervision. A maximum length of 10 years of probation can be imposed for a deferred adjudication, but, on average, defendants are sentenced to 2.35 (3.87) years of community supervision in our 1994 (2007) study sample.[8] They join a larger community corrections caseload supervised by the Harris County Community Supervision and Corrections Department that includes felony convicts who have received a probation sentence. Observationally, deferred defendants' experience of community supervision does not differ from that of their peers with official convictions.

All individuals sentenced to community supervision may be subject to judge-ordered conditions, including community service, a short period of confinement in a county jail, or assignment to a substance abuse treatment facility (Tex. Code Crim. Proc. §42A.301).[9] One basic condition for all community supervision is that the defendant does not commit any new criminal offense. When a defendant violates any condition of the community supervision—either through a technical violation or a non-technical (new offense) violation—a prosecutor can file a motion to revoke community supervision, which triggers an administrative hearing.[10]

[7] Driving while intoxicated charges are ineligible for deferred adjudication by state law.

[8] The actual amount of time on community supervision can be reduced ex-post at the court's discretion, which is why the average length of completed agreements is lower than the noted sentence ranges.

[9] The short periods of confinement used in conjunction with an order of community supervision (sometimes referred to as *shock probation*) is our best understanding for why exposure to incarceration equalizes within several months in the 2007 sample in spite of the stark difference in sentencing. See Section 7.2 for further discussion.

[10] The implication of revocations in our analysis is further discussed in Section 7.2.

If a defendant successfully completes the community supervision, a judge must dismiss the original offense (Tex. Code Crim. Proc. §42A.701(e)). The defendant's criminal history at no point includes a felony conviction unless the deferral has been revoked. We plot the standing of deferral agreements in the years following the case disposition for each of our samples in Figure 1. In the 1994 sample, approximately 40% of the deferral agreements are successfully completed and, in the 2007 sample, over 50% are successfully completed. Once the agreements are completed, there is no way to reverse the case dismissal.

3 Sample Construction and Measurement of Outcomes

3.1 Administrative Data Sources

We rely on several county and state-wide sources of administrative data in this study. From the criminal justice system, these include criminal court records from the Harris County District Clerk, jail entry and exit data from the Harris County Sheriff's Department, state prison data from the Texas Department of Criminal Justice, and the Texas Department of Public Safety's Computerized Criminal History (CCH) database, which tracks state-wide convictions in Texas.

The Harris County criminal court records contain charges (felony and misdemeanor) and court outcomes for all adults between 1980 and 2017 regardless of the final verdict.[11] The records contain longitudinal case notes on major events in the court proceedings, including initial and revised disposition and sentencing outcomes. We use this data to construct the core analysis sample, measure our source of identifying variation (date of charge or disposition), document first-stage diversion outcomes and their associated sanctions, and track future criminal activity. To assess the robustness of our recidivism measure, we also examine Harris County Jail bookings as a proxy for arrests and the state-wide CCH conviction database to evaluate spatial spillovers. While each data source has potential limitations,[12] the findings remain remarkably consistent.

We are also able to observe time spent in Harris County Jail and the Texas prison system. However, we cannot reliably link time served to a specific charge or conviction. As a

[11] Cases sealed to the public by order of the court, which account for less than half of one percentage point of the overall caseload, and criminal appeals were not included in the data.

[12] The Harris County Jail records can also contain multiple bookings associated with the same offense that may inflate the degree of criminal activity, and previous audits have found the CCH data to have incomplete state-wide coverage, particularly prior to the early 2000s, due to the voluntary nature of reporting to the CCH.

consequence, any changes in realized incarceration could be the result of initial sentencing disparities as well as future criminal behavior.

The two Harris County datasets are linked using a unique county identifier tied to an individual's fingerprint known as the "system person number" (SPN). We match this data to state-wide data (CCH and the state prison records) using a defendant's full name and date of birth.[13]

To evaluate the impact of these different sanctions on labor market outcomes, we also match offenders to administrative earnings and employment data drawn from quarterly unemployment insurance wage records between 1994 and 2017 from the Texas Workforce Commission. Wage and employment records were matched to criminal justice records using a Social Security Number (SSN).[14] To reduce the influence of extreme outliers, we cap quarterly earnings at $50,000.[15]

3.2 Sample Restrictions

We restrict our sample to first-time felony defendants charged two years before and two years after the Sept. 1, 1994 threshold and defendants disposed two years before and after the Nov. 7, 2007 threshold. Through imposing the first-time offender restriction, we ensure that each defendant will only appear once in our estimation sample. In addition, individuals without a felony record are generally the target of diversion programs in the U.S., therefore providing evidence for this group is particularly important from a policy perspective.

In 1994, we further limit the analysis to consider those charged with affected statutes as described in Section 2. In 2007, we focus on low-risk defendants defined as those facing a sentence of county jail, felony probation, or no correctional supervision, who we believe would be most likely to be impacted by the 2007 ballot result.[16] We also apply a donut procedure in 2007 and drop all observations with disposition dates between October 30, 2007 and Nov. 9, 2007 to account for a transition period in diversion rates during which some courts appear to

[13] In 1994 (2007), 73.8% (81.8%) of the first-time felony offender sample ever match to a valid jail spell and 46.4% (30.0%) match to a valid prison spell.

[14] Approximately 77% (75%) of the 1994 (2007) sample has a recorded SSN in the case management system. Individuals without a SSN on file were dropped from the labor market analysis. The unemployment insurance wage records do not contain name and date of birth preventing any form of probabilistic matching for those defendants without SSNs. Table A.6 probes the robustness of our findings to imputing missing labor market outcomes.

[15] This affects less than 0.025% of person-quarters observed in the data.

[16] There is no significant change in the likelihood of being in the high or low-risk group across the 2007 threshold, which should assuage concerns of endogenous sample selection.

be responding contemporaneously to their public critics.[17] Table B.5 shows that eliminating this donut procedure reduces the size of the first-stage relationship in 2007 by 33 percent.

3.3 Imputation for Missing Follow-up Data

Although this paper leverages a wealth of administrative records, the timing of the available data makes it infeasible to observe outcomes over a full decade for each individual in the analysis. In the 1994 sample, we are missing initial labor market outcomes for those charged prior to Jan. 1, 1994; in the 2007 sample, we cannot observe end-of-decade criminal justice and labor market outcomes for those disposed after Oct. 1, 2007 and Jan. 1, 2008 respectively. While the missing data represents a minor fraction of the cumulative 10-year outcomes (see Table A.1), we do not want to ignore this issue as it could generate misleading trends in the graphical analysis as well as impact the bandwidth selection and bias-correction procedure in our RD analysis.

We impute the missing data using two types of panel models. In 1994, we backwards "forecast" missing quarters of earnings data using a Tobit model based on observed earnings starting in the first quarter of 1994.

$$Earnings_{i,t} = \begin{cases} Earnings^*_{i,t} & \text{if} \quad Earnings^*_{i,t} > 0 \\ 0 & \text{if} \quad Earnings^*_{i,t} \leq 0 \end{cases}$$

$$Earnings^*_{i,t} = \alpha^{94} + \sum_{q=1}^{6} \beta^{94}_q Earnings_{i,t+q} + \sum_{q=1}^{6} \gamma^{94}_q Work_{i,t+q} + \epsilon_{i,t}$$

In 2007, we utilize an equivalent procedure except we forecast earnings based on lagged values.

$$Earnings^*_{i,t} = \alpha^{07} + \sum_{q=1}^{6} \beta^{07}_q Earnings_{i,t-q} + \sum_{q=1}^{6} \gamma^{07}_q Work_{i,t-q} + \epsilon_{i,t}$$

The imputed values are iteratively computed by: (1) applying the fitted values for the α, β's and γ's to generate a linear prediction of $Earnings^*_{i,t}$, (2) adding a random shock $\hat{\epsilon}_{i,t} \sim N(0, \hat{\sigma}^2)$ based on the model's estimate of the variance of the error term to avoid overstating the precision of the estimates, and (3) censoring the results $\widehat{Earnings}^*_{i,t}$ at zero to generate $\widehat{Earnings}_{i,t}$.[18]

[17] An examination of the micro-data shows that different courts appear to have reacted to the pre-election debate and ballot outcome at different points in time which together create a transition period of approximately two weeks from one practice to another in the overall caseload. This may also relate to the events occurring in the District Attorney's office described in Footnote 6.

[18] In the case of the imputed criminal justice outcomes in 2007, the model is slightly modified due to

Table A.1 shows the share of observations for each sample, outcome variable, and follow-up year that includes imputed values. To assess whether these imputations play an outsized role in our main results, we compute the contemporaneous impact of diversion on future convictions and employment separately by follow-up year in Figure A.2. We find that the level and trend of the contemporaneous impacts are consistent whether or not the follow-up data includes imputed values demonstrating the robustness of our findings to the exclusion of the imputation procedure.

4 Research Design

We first estimate the impact of the 1994 and 2007 policy changes on diversion rates using a sharp RD design. We present both graphical evidence as well as statistical tests to confirm the reliability of our results. For our statistical tests, we follow the approach of Calonico et al. (2014, 2016a) to obtain bias-corrected point estimates using local linear functions, optimal bandwidths and valid confidence intervals. Formally, we estimate the discontinuity ($\hat{\tau}$) based on the following model:

$$\tau = \mu_+ - \mu_-,$$

where,

$$\mu_+ = \lim_{x \to 0^+} \mu(x), \quad \mu_- = \lim_{x \to 0^-} \mu(x), \quad \text{and} \quad \mu(x) \equiv E[Y_i | X_i = x].$$

The parameters μ_+ and μ_- represent the limit of the expectation of Y_i given X_i as it approaches the cutoff threshold from above and below respectively. As a result, τ should be thought to measure the magnitude of the jump in the outcome variable at the point of the discontinuity. In this notation, X_i is the running variable that has a cutoff threshold at $X_i = 0$, which generates a discontinuity in the outcome variable of interest (Y_i). Our running (or forcing) variable differs across the two quasi-experiments due to the nature of each change: the 1994 penal code reform affected offenders based on the date the charge was filed; the 2007 change in court behavior was based on the date the case was first disposed by the court (i.e., given a verdict).

We parameterize $\mu(x)$ using a local linear function:

$$\hat{\tau}(h_n) = \hat{\mu}_{+,1}(h_n) - \hat{\mu}_{-,1}(h_n), \quad \text{where,}$$

the relatively infrequent nature of criminal activity in the data. Instead, we estimate an annual data model based on the follow latent variable: $Total\,Convictions^*_{i,t} = \alpha^{94} + \sum_{y=1}^{3} \beta_y^{94} Total\,Convictions_{i,t-y} + \sum_{y=1}^{3} \gamma_y^{94} Any\,Convictions_{i,t-y} + \epsilon_{i,t}$.

$$\left(\hat{\mu}_{+,1}(h_n), \hat{\mu}^{(1)}_{+,1}(h_n)\right)' = \underset{b_0,b_1\in\mathbb{R}}{\arg\min} \sum_{i=1}^{n} 1(X_i \geq 0)(Y_i - b_0 - X_i b_1)^2 K(X_i/h_n), \quad \text{and}$$

$$\left(\hat{\mu}_{-,1}(h_n), \hat{\mu}^{(1)}_{-,1}(h_n)\right)' = \underset{b_0,b_1\in\mathbb{R}}{\arg\min} \sum_{i=1}^{n} 1(X_i < 0)(Y_i - b_0 - X_i b_1)^2 K(X_i/h_n).$$

K is the kernel function that determines the weighting scheme within a given bandwidth, while h_n represents the size of the bandwidth itself. We opt for a data-driven bandwidth selector that selects the median bandwidth from three mean squared error-optimal methods for the RD treatment effect estimator,[19] and utilize the uniform kernel function. Coefficients are estimated using a first-order local polynomial that has been bias-corrected using a second-order local polynomial with robust standard errors based on a heteroskedasticity-robust plug-in residuals variance estimator with HC_2 weights. Our primary specification adjusts for baseline covariates of age, gender, race/ethnicity, and prior number of misdemeanor convictions.[20] These parameterization decisions can be modified without impacting the findings of this study as documented in Tables B.3, B.4, and B.5.

After quantifying the "first-stage" relationship, we then turn to a fuzzy RD design to measure the causal effect of the diversion ($\hat{\phi}$) on court sanctions, future offending behavior, and labor market outcomes. In this last approach, the measured change in outcomes is scaled by the observed first-stage relationship to quantify the impact of diversion. In doing this, we assume that each threshold influences the outcomes measured only through its impact on diversion (exclusion restriction) and that there are not individuals for whom the threshold is associated with an increase in diversion in 1994 and a decrease in 2007 (monotonicity assumption). We believe the specific context and institutional details discussed in Section 2 support these assumptions. Further validation of the exclusion restriction assumption is provided in Section 6.

We also compute sample averages, which are presented with our estimates in brackets, to benchmark our results. For our sharp RD estimates, we present the overall caseload average in the bandwidth window prior to the cutoff. For our fuzzy RD estimates, we show an adapted control complier mean (CCM) (Abadie 2003, Kling et al. 2007, Heller et al. 2017) for the RD setting. The CCM is the average outcome for the *complier* population (i.e. those caused to be diverted due to the discontinuity) in the absence of diversion (i.e. convicted). Please see Appendix A.1 for a discussion of our application of this method to the RD context.

[19] We use the option *msecomb2* within the STATA *rdrobust* command described by Calonico et al. (2016a), which uses the median bandwidth from the following methods: one common MSE-optimal bandwidth selector for the RD treatment effect estimator; two different bandwidth selectors (below and above); and one common MSE-optimal bandwidth selector for the sum of regression estimates.

[20] See Calonico et al. (2016b) for notation of this methodology including baseline covariates.

5 Empirical Results

5.1 Caseload Density and Baseline Characteristics

To attribute a causal interpretation to our RD estimates, we must assume that defendants are randomly allocated before and after each threshold. Individuals charged immediately before Sept. 1, 1994 should be observationally equivalent to those charged after, and we should not see a discontinuity in the total number of cases; likewise, defendants disposed immediately on or after Nov. 7, 2007 should be observationally equivalent to those disposed before, and we should not see a discontinuity in the total number of dispositions.

A sharp general deterrence response to the change in diversion rates would violate this assumption. Additional threats to our empirical strategy include changes in policing practices and sorting of offenders by prosecutors/judges across the threshold dates to guarantee they face one punishment regime versus the other. This is of particular concern in the context of the 1994 reform since all relevant actors could fully anticipate the adoption of the new penal code. Each of these threats generate a testable prediction of discontinuous changes in either the size or composition of the criminal caseload across the discontinuity.

In spite of these potential threats to identification, we do not observe discontinuities in caseload densities or in baseline defendant demographic characteristics such as gender, age, race, and ethnicity or in misdemeanor criminal history. Balance across these factors is presented in Table 1, with supporting graphical evidence available for each estimate in Figure B.1. We do observe a marginally significant increase of 4 percentage points in the likelihood of being male after the 1994 cutoff, but no significant difference is observed in 2007. The third column of the table reports the combined estimate, $\frac{\tau^{07}-\tau^{94}}{2}$, which measures whether any systematic sorting toward a higher probability of diversion is detectable when averaging across the two experiments. Again, we do not see any evidence of endogenous sorting that contradicts our identification assumptions.

As an additional test, we calculate a predicted recidivism risk score for each defendant using the baseline characteristics and prior number of misdemeanor convictions.[21] Since no

[21] We calculate our measure of recidivism risk from the predicted dependent variables for each individual from the following OLS regression: Total Charges$_i^{10\text{ Years}} = \alpha + \mathbf{X}_i'\beta + \epsilon_i$, where $\mathbf{X}_i'$ is the set of the observable covariates (i.e., age, sex, race/ethnicity, and prior misdemeanor convictions — not the forcing variable) as well as the corresponding two-way interactions. The risk score is defined as $\hat{\alpha} + \mathbf{X}_i'\hat{\beta}$ and captures an offender's predicted rate of recidivism over ten years based on his or her observable characteristics. If police or prosecutors act in a discriminatory manner and monitor certain sub-populations at higher than average rates (e.g., African American men), then an alternative interpretation of this index would be having a higher or lower likelihood of involvement with the criminal justice system whether through differences in actual

information from the running variable or discontinuity is used in constructing this index, the assumptions of the RD research design would imply that no sharp changes in the predicted risk of recidivism should appear at the threshold, which is what we observe empirically in Table 1.

Continuity in the caseload density, demographic composition, and predicted recidivism risk all strongly support our identification assumption of continuity in unobserved determinants of our outcomes across the threshold. We do, however, estimate a significant difference in missing a SSN in the 2007 sample. Upon further investigation, this difference appears to be driven by an increase in the proportion of case dismissals following the Nov. 2007 ballot, which have a higher likelihood of missing a SSN. While any imbalance in observable characteristics across the threshold is a concern, this specific imbalance will only impact our labor market analysis. In Section 6, we discuss empirical strategies for dealing with these missing records.

5.2 Court Verdicts and Sentencing Outcomes

The exact meaning of and counterfactual to diversion varies between the 1994 and 2007 experiments. Panel A of Table 2 reports the impact of diversion on case dispositions, while Panel B shows the effect on court-ordered sanctions. These point estimates are calculated using a fuzzy RD design and should be interpreted as the causal impact of diversion on disposition and sentencing outcomes.

In 1994, a diversion is entirely characterized as a deferred adjudication of guilt. We find a modest decrease in the likelihood of receiving an incarceration sentence or being issued a fine, but do not estimate any difference in the likelihood of probation.[22] As can be seen from the CCMs, marginal defendants not diverted receive a felony conviction and a probation sentence, which is consistent with the probation-before-incarceration requirement created after the 1994 cutoff. While we observe a decrease in the likelihood of an incarceration sentence on file for those diverted, it is unlikely that these sentence differences reflect changes to sanctions in practice, which we confirm through our analysis of jail and prison records.

In 2007, what we term diversion is split roughly 60% and 40% between deferred adjudications of guilt and outright case dismissals respectively. While these are clearly different disposition outcomes, we have combined these rulings under the broader umbrella of diversion given that both avoid a criminal conviction and decrease further involvement in the criminal justice system. By 2007, the probation-before-incarceration requirement had been eliminated,

future behavior or differences in future monitoring.

[22]Fuzzy RD estimates using the dollar amount of fines as an outcome variable imply that diversion in 1994 was associated with an insignificant decrease of $82 in fines on the margin.

and the primary counterfactual to diversion is a felony conviction with an incarceration sentence. Here there is an unambiguous change in incarceration sentences as a consequence of diversion. Whether this translates into meaningful differences in time served depends on the length of sentence (which are relatively short by sample construction), parole practices (which we cannot directly observe), and other factors that could contribute to spending time in incarceration like pre-trial detention and future bookings. From Table 2, it is clear that diversion in 2007 leads to an increase in probation and/or a fine.[23] These estimated effects are less than a one-for-one exchange with incarceration due to the non-trivial share of defendants who received a dismissal-based diversion where no sentence would be applied.

Throughout our analysis, we test the average causal effect across 1994 and 2007 and equality between the two experimental estimates (the third and fourth columns respectively of tables reporting our RD results).[24] These exercises are valuable from two perspectives. First, we can strengthen our statistical precision to detect impacts on high-variance behavioral outcomes through combining the estimates. Second, testing the difference between the coefficients quantifies the differences and similarities between the 1994 and 2007 estimates, which helps us distinguish between potential mechanisms driving the causal effects. Given the differences in the implementation of diversion across 1994 and 2007, it is unsurprising that our tests of equality are rejected across the board for case dispositions and sanctions in Table 2. When pooling the two natural experiments, deferred adjudications of guilt are the predominant case disposition associated with diversion, although 16 percent of marginal defendants still end up with a case dismissal. The sentencing impacts are generally more muted in the combined estimate, and the opposite-signed impacts on fines cancel each other out.

5.3 Reoffending and Labor Market Outcomes

Our main recidivism analysis, which evaluates the 10-year impacts of diversion on future convictions by extensive and intensive margins, by type of crime, and by offense level, are presented in Table 3. As with the disposition and sanction outcomes, the coefficients are estimated using a fuzzy RD design and represent the causal impact of diversion.

In both 1994 and 2007, we find strong and consistent evidence that diversion leads to statistically significant and economically meaningful declines in recidivism (at both extensive

[23] Fuzzy RD estimates on the amount of fines show that diversion in 2007 is associated with an insignificant increase of $130 in fines on the margin.

[24] Given that the samples are independent and separately estimated, the tests of coefficient equality are conducted as follows: $Z = \frac{\phi_{1994}-\phi_{2007}}{\sqrt{SE^2_{1994}+SE^2_{2007}}}$. The tests of the average coefficient are conducted as follows: $Z = \frac{\phi_{1994}/2+\phi_{2007}/2}{\sqrt{(SE_{1994}/2)^2+(SE_{2007}/2)^2}}$

and intensive margins). Diversion decreases reconviction rates by slightly more than 30 percentage points in both samples, which reflects close to a 50 percent reduction relative to the CCM. On the intensive margin, diversion reduces the total number of future offenses by 1.3 to 1.6 new convictions (56 to 77 percent). The pooled coefficients are similar in magnitude but more precise allowing us to rule out small to modest impacts.[25] Not surprisingly, we cannot reject that diversion has the same causal effect in the two separate analysis samples.

The change in the types of crimes prevented differ somewhat between 1994 and 2007. In 1994, it appears that the decrease in convictions is driven by lower rates of drug and property offenses. In 2007, drug offenses show a more modest effect while violent offenses exhibit a more substantial and statistically significant decline. Despite these differences, we cannot statistically distinguish effects for specific types of crimes across the two natural experiments (as reported in the fourth column of Table 3). In fact, the pooled estimates show significant reduction for all types of crime.

By level of the offense, it appears more felony-level crimes were prevented through diversion in 1994 (91 percent decline) whereas more misdemeanor-level crimes were prevented through diversion in 2007 (83 percent decline). But, here again, we estimate significant impacts on both margins in the pooled sample, and statistical tests of the equality of the sample-specific coefficients fail to reject the null.

Our primary 10-year labor market results are presented in Table 4. This includes performance measured on three dimensions: employment rates (Panel A), total earnings (Panel B), and earnings stability (Panel C). Across both experiments and each group of outcomes, the evidence consistently shows that diversion has a positive causal impact on measures of labor market performance.

We find that diversion increases quarterly employment rates by 20 (16) percentage points in 1994 (2007), indicating more than a 50 percent increase over the CCMs. In other words, diversion increases the average number of quarters with positive earnings from 13.6 quarters to 20.8 quarters in the complier population over ten years. These employment gains are not confined to occasional, very-low-income work; the second row of Panel A indicates significant improvements (13 percentage points) in quarterly employment rates with earnings exceeding the federal poverty level for a single adult.[26] Table 4 also documents improved earnings overall as shown in Panel B. Averaging between the two experiments, we find that diversion increases

[25] In the pooled sample, we can reject that the impact of diversion is smaller than a 17 percentage point reduction in reconviction rates and a 0.79 reduction in total convictions based on a two-sided test with p-values less than 0.1.

[26] Although this result is encouraging, it is important to bear in mind that this represents a relatively modest target of just $3,035 per quarter in 2018.

log total earnings by 1.8 log points, and total earnings by $61,540. Taking into account that the CCMs are around $95,000 over this period, the results represent substantial improvements for UI-covered income over baseline rates.

It is possible that these large differences in earnings are a result of decreases in job churn or other types of employment disruption. To explore this, we create two outcome measures of employment stability: first, we calculate the maximum continuous spell (in quarters) of employment with a specific employer over our 10-year follow-up period; and, second, we measure continuous periods of earnings, a version of employment stability for the defendant that is not restricted to a single employer. Consistent with the prior labor market performance results, we find that diversion generates longer durations of continuous within-firm employment (3.7 quarters) and continuous periods of earnings (4.9 quarters). Increases in employer-specific experience or employment stability among those diverted likely contributes to the observed earnings growth over our follow-up period.

The 1994 experiment generally exhibits stronger statistical precision throughout Table 4. In particular, the impacts to employment rates in 2007 are marginally insignificant, which may raise questions about the robustness of these conclusions. Several pieces of evidence, however, stoke our confidence in these findings. First, across each of the labor market outcomes, the pooled estimates are highly precise with little doubt regarding the statistical significance. Because the pooled estimates integrate all of the identified exogenous variation into a single test, they are our best guess of the causal impact of diversion. Second, we fail to reject the null of treatment effect equality between 1994 and 2007 for all considered outcomes. Third, a number of alternative specification choices presented in our robustness exercises (see Tables B.3-B.5) do yield statistically significant improvements in employment rates as a result of diversion in 2007. Finally, the timing of the impacts presented in Figure 3, which will be discussed next, demonstrate significant effects on employment rates in 2007 between Years 1 and 8. This suggests the impact at Year 10 is somewhat less precise but unlikely to be the product of random noise.

Figure 3 plots the estimated cumulative impact of diversion on three outcomes (total future convictions, the quarterly employment rate, and total days incarcerated) over time. The first coefficient corresponds to the impact during the initial 365 days post-charge or disposition, and each sequential estimate adds another year of data to expand the total follow-up window up through 10 and 20 years for the 2007 and 1994 samples respectively.[27]

Both 1994 and 2007 plots show similar trajectories. We observe an immediate impact on

[27]Figure B.7 reports a summary of the visual evidence of the reduced-form discontinuities underpinning these results.

criminal activity and employment that starts in the first year and grows year by year as the follow-up window is expanded. The first five years show particularly rapid improvements in crime reductions and employment gains, yet the benefits continue to accrue as far out as we are able to track outcomes. Although our main analysis focuses on 10-year outcomes, the persistent and growing impact up to 20 years after the first felony charge in the 1994 sample is quite remarkable. We interpret this pattern to indicate that diversion, at least at the critical juncture of someone's first felony charge, has the potential to fundamentally alter an individual's trajectory in life.

The final row of Figure 3 depicts the cumulative effect on time served in jail or prison. Given the previous discussion on sentencing impacts, it is useful to evaluate the effects on actual time served to assess whether changes to incarceration is a key mechanism for achieving diversion's impacts (see also Section 7.2 for more exercises examining this hypothesis). The plots do not show any significant differences in cumulative time served until more than 10 years out in the 1994 sample; no significant effect is ever observed in the 2007 sample. Figure A.2, which shows the contemporaneous rather than cumulative effects on incarceration, documents an emergent trend starting in Year 6 for the 1994 sample. These impacts appear to be the byproduct of future criminal activity rather than a delayed implementation of differences in original sentencing (see Figure A.4 for the impacts on the accumulation of new sentences stemming from future crime). [28]

Our previous analyses consider criminal and labor market outcomes independently, however, the incidence of these impacts could fall on mutually exclusive subgroups, a common population or some combination thereof. To distinguish between these possibilities, we create four mutually-exclusive outcomes for the 10-year follow-up period. We consider whether a defendant's quarterly employment rate is above or below 50% during the follow-up period and whether they have zero or at least one new conviction during the follow-up period. The sample-specific and pooled results indicate substantial changes on the extreme outcomes (see Table 5). We estimate that diversion increases the fraction of individuals who jointly experience higher levels of employment and no future convictions by 25 percentage points, a gain of 232 percent over the CCM. At the same time, diversion reduces the fraction of defendants with low employment levels and at least one new conviction by 32 percentage points (66 percent). It's hard to know whether these results reflect a major improvement for a narrow share of the marginal caseload or an incremental improvement for a broader group. Regardless of the interpretation, however, the impacts demonstrate substantial improvements following modest (if any) costs to the justice system.

[28] In Section 7.1, we define this potential indirect mechanism of diversion as an amplification effect.

5.4 Heterogeneity Analysis

Certain subgroups may particularly benefit from diversion. To explore this possibility, we examine whether our estimated impacts differ over the spectrum of predicted recidivism risk.[29] We prefer this approach to the typical sub-group analysis because of the correlation between demographic traits in our sample. For example, black offenders are roughly two years younger than white offenders, so subgroup estimates by race may actually capture differences by age.[30]

We implement this analysis by separately estimating local polynomial RD specifications for each percentile in the risk score quantile function. Due to sample size constraints, we utilize a 40 percentile uniform bandwidth centered at the focal percentile when estimating these coefficients. Since the quantile function is not defined below zero or above one hundred, asymmetric bandwidths occur above the 80th and below the 20th percentiles.[31] Figure A.3 shows the prevalence of background characteristics over the smoothed risk score quantile function. Individuals at the highest predicted risk of recidivism are more likely to be young, African American, male, and have prior misdemeanor convictions.

The first row of Figure 4 presents the first-stage relationship between diversion and the threshold across the 1994 and 2007 sample thresholds. The first stage magnitude remains fairly constant throughout the distribution. The second and third rows report the 10-year impacts on total convictions and average employment rates. The largest point estimates appear at the top end of the distribution in both samples, although improvements in reoffending behavior also appear to show up among low-risk defendants as well. These results are striking in the fact that observable characteristics would predict that high-risk defendants would have worse future outcomes, yet we find that diversion effectively closes the employment gap between the 90th and median risk percentiles, and halves the re-conviction gap between these groups. Two related mechanisms could generate these results. First, young African American men may be over-targeted by law enforcement leading to a disproportionate treatment response in the marginal subpopulation. Second, the subpopulation may suffer more on average from a felony conviction record due to statistical discrimination or animus.

[29] See Section 5.1 for a description of the construction of this index. To account for the bias discussed in Abadie et al. (2016), the estimation of the risk score employs a leave-one-out or jackknife estimation procedure for the purpose of these exercises.

[30] We also report heterogeneous effects by sub-groups in Table B.1.

[31] This exercise requires the stronger assumption that defendants before and after the discontinuities exhibit similar risk score distributions. Figure B.8 shows a series of local polynomial RD estimates examining whether recidivism risk remains balanced within subsets of the risk score distribution. This provides a consolidated way to demonstrate balance throughout the distribution (as opposed to being balanced on average in the sample overall as was previously documented in Table 1 and Figure B.1).

6 Evaluating Potential Violations of the Exclusion Restriction and Robustness Checks

6.1 Potential Violations of the Exclusion Restriction

An important assumption underlying our research design is that no other factor that could impact defendants' future behavioral outcomes changed across our two thresholds other than the decrease or increase in the probability of diversion. We have observed and discussed how sentencing outcomes did respond, which could be thought to represent a violation of the exclusion restriction. We disagree with this interpretation, however, as we view these sentencing shifts as derivative of the initial change in case dispositions and therefore part of the overall treatment effect of diversion rather than violations of our identification assumptions.

In contrast, a more concerning pattern would be if we observed changes in behavioral outcomes in other parts of the felony caseload where there was no change in diversion rates. In this scenario it would be difficult to differentiate whether our study's findings were caused by the change in diversion practice or some other caseload-wide phenomena affecting defendants near the cutoff threshold. Table 6 and Figure B.9 reproduce our main results for all other felony defendants excluding those used in our main analysis.[32] For both the 1994 and 2007 samples, we find no evidence of a change in future behavioral outcomes among other parts of the felony caseload and can thus reject that the reduced form change is equal to the results for our main study sample.

Unobserved seasonal breaks that coincide with the timing of our cutoffs such as the start of the school year or holiday season could be another route to violating the exclusion restriction. While we would also expect these patterns to cause differences among the non-focal caseload outcomes in Table 6, we provide additional support in Table A.8, which presents estimated effects of a placebo experiment where we adjust the cutoff dates to one year earlier (Sept. 1, 1993 and Nov. 7, 2006) or one year later (Sept. 1, 1996 and Nov. 7, 2008) than our actual threshold dates. Overall, we do not find any meaningful discontinuities in initial court outcomes, future reoffending, or future labor market outcomes in these placebo exercises.

Because we observe two discontinuous changes in the likelihood of diversion, we can pool the two samples, estimate an instrumental variables model and test for over-identifying restrictions to evaluate whether our instruments are uncorrelated with the structural error term. Panel A in Table A.7 shows the estimated treatment effect of diversion using different specifications

[32] This combines repeat offenders as well as first-time offenders for non-impacted crime types. For results that specifically compare the estimates for first-time versus repeat offenders, see Table A.9.

including variables for sample-specific intercepts, sample-specific universal trends, and sample-specific post-trends.[33] Across all specifications, our point estimates replicate our main results, and more importantly, we see strong failure to reject correlation between the instruments and the error term, which should further address concerns regarding violations of the exclusion restriction.

6.2 Robustness Checks

The main reoffending analysis is based on Harris County conviction records. Tables A.2 and A.4 and Figures B.3 and B.4 demonstrate the robustness of our results to a variety of alternative recidivism measures: county jail bookings, county charges, county convictions including non-technical violations, and state-wide convictions. There does not appear to be any significant difference in the point estimates depending on the measure of recidivism. Given differences in the CCM by outcome, with bookings having the highest prevalence and convictions the lowest, the percent change does increase the closer one gets to conviction, but these changes are statistically indistinguishable.

To address concerns regarding differential mobility out of Harris County, we compare Harris County and non-Harris County reoffending outcomes in Texas using the state-wide Computerized Criminal History (CCH) Database (Table A.3). If diversion displaces criminal activity to other counties in Texas, we should expect to see a negative impact in Harris County and a positive impact elsewhere in the state. By contrast, however, the results show a decline in reoffending both within and outside Harris County. Although the Harris County-specific convictions make up the majority of the observed impact in the state-wide CCH data, the effects elsewhere in the state suggest that our county-specific measures, if anything, underestimate the total gains from diversion.[34]

We also challenge the robustness of our labor market analysis. First, we assess whether the missing earnings information for those without a SSN recorded in the case management system may bias our findings. This is particularly important for the 2007 analysis where we see a small but discrete jump in the likelihood of having a SSN recorded. Table A.6 shows the estimated impacts for employment rates, total earnings, and log total earnings under two assumptions. In the first column, we drop all observations that have a missing SSN, which is a

[33] The bandwidth is selected using a pooled RD model examining the sharp change in diversion.

[34] While these state-wide results partially alleviate mobility concerns, there could be differences in defendants being forced to leave Texas. If Hispanic immigrants are more or less likely to be deported, for example, this could bias our estimated effects. However, we expect this bias to go in the opposite direction of our estimated effects on reoffending outcomes since those receiving a felony conviction would be more likely to be deported compared with those with a diversion.

reproduction of our main estimates. In the second column, we impute labor market outcomes for defendants with missing SSNs based on their observable characteristics and the timing of their charge or disposition. The results show no meaningful change regarding labor market activity, suggesting the missing SSNs are not significantly confounding our findings.

Throughout our analysis, we present average treatment effects giving equal weight to the sample-specific coefficients. Another approach would be to pool the raw data and estimate a single unified RD, thereby generating an estimate for a combined treatment effect.[35] Table B.6 compares the results of this pooled approach relative to the simple averaging strategy presented in the main results. The results are highly consistent and cannot be rejected as being statistically indistinguishable. Due to its transparency, we prefer the simple averaging approach over the pooled estimation method.

Finally, our results are also robust to alternative data-driven bandwidth selection methods (Table B.3), which is not entirely surprising given the stability of the estimates over a range of fixed bandwidths starting from 90 days up to 720 days (see Figure A.1).[36] Table B.4 demonstrates robustness to the use of alternative variance estimation strategies, and Table B.5 reports results for several other modifications including models without the donut restriction, without including covariates, with alternative kernels, and without using the robust standard errors or bias correction as suggested by Calonico et al. (2014). We also provide estimates from models based on outcomes aggregated to the week level. Among these tables, there is no evidence to suggest our findings are driven by any choices we made in implementing the RD methodology.

7 Discussion of Potential Mechanisms

An important question remains: how does this low-cost intervention generate such large and long-term improvements? The magnitude of our estimates are particularly surprising given that prior evaluations of reentry services (as recently reviewed by Doleac (2018)) or summer jobs (Heller 2014, Gelber et al. 2015) have failed to yield impacts with similar magnitudes or degrees of persistence despite requiring substantially more resources.

[35] To accomplish this, both datasets are recentered around a cutoff date of zero, and the running variable from 1994 is multiplied by -1 to ensure that diversion rates consistently increase when moving from the left to the right over the cutoff threshold.

[36] For reference, the vertical red line in Figure A.1 indicates the position of the optimal bandwidth chosen by the Calonico et al. (2014) methodology for our primary results.

7.1 Mechanisms Supported by the Data

Growing evidence suggests that a felony conviction carries sharp penalties for employment opportunities (Raphael 2014, Finlay 2009, Pager et al. 2009, Pager 2003). A stigma effect may also exist within the criminal justice system if police, prosecutors, or judges react to prior felony convictions. This may operate mechanically due to penal code escalations or risk classification algorithms that depend on criminal conviction history.[37] Our analysis sample is composed of individuals with no prior felony charges, and so diversion can safeguard them against the lifelong mark of a felony record.

A natural way to evaluate this channel is to compare our estimated impacts with the corresponding coefficients for defendants already holding a felony record. Diversion, however, is rarely offered to those with a prior felony, which limits our ability to implement this strategy. Table A.9 compares our main RD estimates (both sharp and fuzzy) with the corresponding coefficients for defendants with 1 to 3 prior felony convictions. The exercise suggests that defendants without a felony record disproportionately benefit from diversion, which would be consistent with the stigma mechanism.[38]

The second direct channel that we find evidence for in the data is a form of deterrence that we term *re-sentencing deterrence*. If a defendant violates his diversion agreement, he can be convicted and sentenced at that time for the original offense.[39] This threat would discourage reoffending during the agreement period, which lasts on average 2.86 years in our sample, especially for minor non-felony crimes.[40]

The persistence of the estimated behavioral impacts of diversion remain well after the expiration of the deferral agreements (see Figures 3 and A.2), indicating that re-sentencing deterrence plays a secondary role to felony conviction stigma. In fact, Panel B in Table A.7

[37]For instance, certain types of prior convictions can trigger aggravated (elevated) charges upon re-offense which results in more severe and potentially more certain future convictions. Additionally, defendants under varying forms of supervision may face differing probabilities of arrest conditional on criminal activity. We are not aware of any empirical evidence that confirms this to be the case, however.

[38]This exercise is limited by two key factors. First, in 1994, there is no discernible first-stage relationship. Consequently, there is no way to compute treatment effect estimates for this cohort. In 2007, we do observe a statistically significant yet modest first-stage relationship in the repeat offender caseload. In fact, the point estimates suggest that diversion does not generate the same gains for this population, however, large standard errors, which are partially a consequence of the modest first stage, limit our ability to make this claim definitively.

[39]This only applies to individuals who receive a deferred adjudication of guilt disposition.

[40]A new minor misdemeanor conviction (e.g., shoplifting) would result in a comparatively outsized sanction (felony conviction record and potentially time in prison) due to the sentence imposed by the violation of the deferred adjudication agreement, whereas a new serious felony conviction (e.g., aggravated assault) would trigger a felony conviction and harsh sanctioning regardless of whether a deferred adjudication was active at the time.

shows that we cannot reject equality in the estimated treatment effect of deferred adjudication and case dismissals where no re-sentencing is possible due to double jeopardy.[41]

We also believe an indirect mechanism, which we term an *amplification effect*, contributes to the long-run impacts. Any initial changes to criminal or labor market activity may dynamically influence future outcomes. For instance, if a felony conviction causes additional criminal activity and this results in additional sanctions against the individual, the initial treatment differences between diversion and conviction dispositions may be reinforced or amplified. Stated differently, the dosage of justice system involvement may grow over time as a consequence of the future behavioral outcomes.

Figure A.4 shows direct evidence that amplification is plausible. This figure tracks the cumulative impact of diversion on new sanctions for future criminal activity and shows that diversion decreases the total number of future incarceration and probation sentences. The effects are larger for incarceration, but both impacts grow over the follow-up period which parallel the cumulative impacts to behavioral outcomes presented in Figure 3. The precise nature of an amplification effect is hard to pin down, though. Evolving differences in incarceration raise the possibility of incapacitation and post-release effects. Changing labor market opportunities alter the opportunity cost of criminal activity. A host of other interventions and incentives may also come into play. Without further structure imposed on the analysis, we are not able to disentangle these channels.

7.2 Potential Mechanisms Not Supported by the Data: Incarceration, Fines, and Probation Revocations

A number of alternative channels could theoretically explain our results but are not supported by the data. For instance, mechanisms implying a change in the composition of offenders at the threshold dates such as a general deterrence effect from the changes in expected punishment or any manipulation of charge or disposition dates are empirically ruled out in Section 5.1.

Our analysis of sentencing outcomes raises incarceration as a potential mechanism. Vetting this hypothesis is critical for an accurate interpretation of our findings. Figures 3 and A.2, previously discussed in Section 5.3, provide evidence that the change in incarceration sentences at our discontinuity plays a minimal role in generating our observed behavioral impacts. Figures B.10 and B.11 provide a number of alternative (scaled and reduced-form) timelines

[41] This exercise (described in further detail in Section 6.1) relies on the strong assumption that diversion exhibits a common treatment effect in 1994 and 2007 compared with case dismissal.

to flesh out this argument.

The null impact of diversion on time served in jail or prison for the 2007 sample is surprising given the clearly documented change in sentencing outcomes and the related overcrowding concerns in the Harris County jail. We further investigate this issue in Figure B.12 using high frequency (weekly) data on time served in the first year following disposition. The week-by-week results yield additional insight missed by the prior annual estimates: consistent with the change in sentencing outcomes for the 2007 sample, diverted defendants are significantly less likely to be incarcerated immediately following disposition. However, these defendants start experiencing higher rates of incarceration roughly two months after disposition, which effectively equalizes cumulative exposure to incarceration over time; in fact, there is no statistical difference in cumulative exposure to incarceration by the fourth month. While deferral revocations could partially contribute to this pattern, Figure 1 shows that only 5 percent of deferments are typically revoked by the three-month mark. Instead, this pattern is more consistent with the use of shock probation—short periods of incarceration scheduled in the beginning of community supervision or applied by a judge to punish minor technical violations without triggering a revocation. In summary, we observe only a temporal displacement of short incarceration spells from immediately after disposition to a few months following disposition and no statistical difference in cumulative exposure within 4 months of disposition for the 2007 sample.[42]

It is plausible that longer-term (>5yrs) employment and offending differences following the 1994 change in diversion are affected by incarceration patterns. This conjecture is consistent with the evidence on amplification previously discussed. An incapacitation channel would dampen the difference in reoffending outcomes but could exacerbate differences in employment and earnings. We quantify the potential role of incarceration-related incapacitation in our employment effect estimates in Table A.5. The table reports estimates where we recode employment status during quarters of incarceration to shut down the mechanical incapacitation channel. In column 2, we recode employment status to be equal to the average employment rate in the year prior to the current incarceration spell. In column 3, we recode employment status to be equal to 1 for all quarters with some incarceration. In this last exercise, we make the strong and arguably disproven assumption that all inmates in the absence of incarceration would have secured gainful employment in the formal sector (see Mueller-Smith (2015)). While

[42] A parallel exercise for the 1994 sample in Figure B.12 yields results that are consistent with our understanding that the marginal defendants received community supervision regardless of their diversion status. We find no statistical difference in exposure to incarceration in the weeks following the charge for those diverted due to the 1994 penal code reform and no statistical difference in cumulative exposure throughout the entire first year.

our results decrease if we assume the counterfactual to incarceration is guaranteed employment, our effects are very similar if we make the more realistic assumption that employment rates would have persisted at pre-incarceration rates in the absence of incarceration.

We also find that diversion significantly impacts financial penalties, yet the direction of the effects go in opposite directions in 1994 as opposed to 2007. Unless the causal effect of financial sanctions changed signs in the 13 years between the two experiments, it is unlikely this channel is of first order importance in generating the consistently signed behavioral impacts across the two experiments.

Another potential mechanism by which diversion decreases future convictions could be through the treatment of non-technical violations and the revocation process. For example, if a deferred defendant commits a new criminal offense while under community supervision, the prosecutor can file for a motion to revoke probation which may result in an incarceration sentence that overshadows any sanction from separately pursuing a conviction for the new criminal offense. This type of behavior could suppress future conviction rates for defendants under community supervision. To assess this concern, we incorporate case file data that captures probation and deferral revocations and evaluate whether our results hold after adding such revocations as "uncharged convictions" into our measure of recidivism. While we do not observe the underlying reason for revocation (technical violation or a new offense), we convert a fraction of revocations based on rates of non-technical (criminal) community supervision revocations provided by the Texas Legislative Board (TX Legislative Budget Board 2005, 2008).[43,44] Table A.4 reports consistent effects across the extensive and intensive margins of criminal activity compared with our main results, indicating revocations do not influence the findings.

8 Conclusion

This paper studies two discontinuities in criminal court diversion — a cost-saving strategy that offers defendants a second chance to avoid a felony conviction record — among a large population of low-risk offenders in Harris County, Texas. While these two changes occurred 13 years apart and originate from different contexts, we find large and consistent impacts

[43]This is a conservative approach given that many non-techincal revocations were actually separately prosecuted in practice and, as a result, would be double counted as both revocations and new convictions in this exercise.

[44]TX Legislative Budget Board (2008) reports that 36.8% of the 2007 revocations for felony community supervision arrangements were for non-technical (criminal violations). We were unable to find a similar statistic for Harris County in 1994 or 1995 and therefore use the state-wide average non-technical revocation rate of 45% between 1999 and 2005 as reported by TX Legislative Budget Board (2005).

from both experiments: future recidivism roughly halves and employment rates improve by around 50 percent. To the best of our knowledge, no prior research has found similarly large and long-lasting effects for such a low-cost intervention in the justice system.

An extensive empirical investigation of theoretical channels points to the stigma associated with a felony conviction as the focal mechanism. This finding contributes substantially to prior studies that discuss the impact of a prior conviction on re-offending and labor market outcomes. In particular, our estimates suggest that the negative impact of a conviction on employer callback rates, as found in the extensive audit-based literature, extends to actual employment and earnings outcomes and persists in spite of potential search behavior by convicts to match with non-discriminating employers.

A heterogeneity analysis demonstrates a striking pattern: those at the highest risk of recidivism gain the most from diversion. These individuals are typically young black men with one or more misdemeanor convictions, a group often discussed as over-policed in the United States. Our results indicate that intervening for this disadvantaged population at a critical moment (i.e., when they are being charged with their first felony offense) could significantly improve their life course. More broadly, these results suggest the criminal justice system likely plays a crucial role in the propagation of racial inequality in the U.S.

We consider the evidence presented to be highly rigorous. The research design addresses selection bias without relying on controlling for observables or propensity score matching. The results are shown to be robust to a variety of specification choices and bandwidths. A number of exercises support the viability of our exclusion restriction assumption. The findings are replicated in two independent samples which together exhibit both increasing and decreasing rates of diversion. As such, it is unlikely that our analysis is biased by some unobserved shock contemporaneous with the discontinuous changes in diversion.

In 2010, annual justice expenditures eclipsed $250 billion in the U.S. (Kyckelhahn 2010). Given the impact on federal and state budgets, substantial efforts have been made to reform criminal justice policy. The general trend has been toward more leniency, especially for first-time and low-risk defendants, with diversion emerging as a popular option. Our results suggest that these changes may lead to lower rates of reoffending and higher rates of rehabilitation in the coming years.

Some policymakers, however, have sought to reverse course (Sessions 2017). This could reflect an alternative, risk-averse perspective that seeks to minimize the possibility of high-cost rare events (e.g. inadvertently diverting a to-be murderer) even at the cost of average gains. Our study is not well-equipped to quantify the impacts of diversion on these rare events. We

do find a decline in violent crime overall, however, which should be attractive to those coming from this perspective.

While much has been recently written on ineffective criminal justice policy in the U.S., this paper provides compelling evidence on a successful intervention that both improves defendant outcomes and saves public resources. What we find most appealing about diversion is that it can be feasibly implemented without significant investments or changes to current infrastructure, making it a practical solution for criminal justice reform. While it is not clear whether individuals with a prior felony conviction would similarly benefit from diversion, the results are clear for those facing their first felony charge. For all these reasons, we believe diversion should be viewed as an promising option for jurisdictions seeking to reduce the fiscal cost and potential negative community impacts of the criminal justice system.

References

Abadie, A.: 2003, Semiparametric instrumental variable estimation of treatment response models, *Journal of Econometrics* **113**.

Abadie, A., Chingos, M. M. and West, M. R.: 2016, Endogenous stratification in randomized experiments. Working Paper, Retrieved from http://economics.mit.edu/files/11852 [Date Accessed: 11/29/2016].

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

Becker, G.: 1971, *The Economics of Discrimination*, University of Chicago Press.

Bhuller, M., Dahl, G. B., Löken, K. V. and Mogstad, M.: 2016, Incarceration, recidivism, and employment, *Working Paper, available at https://sites.google.com/site/magnemogstad/IncarcerationRecidivismEmployment_final.pdf?attredirects=0&d=1* .

Bird, M., Lofstrom, M., Martin, B., Raphael, S. and Nguyen, V.: 2018, The Impact of Proposition 47 on Crime and Recidivism, *Technical report*, Public Policy Institue of California.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R.: 2016a, *rdrobust*: Software for Regression Discontinuity Designs. University of Michigan Working Paper, Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2795795 [Date Accessed: 10/20/2016].

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R.: 2016b, Regression discontinuity designs using covariates. University of Michigan Working Paper, Retrieved from http://www-personal.umich.edu/~cattaneo/papers/Calonico-Cattaneo-Farrell-Titiunik_2016_wp.pdf [Date Accessed: 11/29/2016].

Calonico, S., Cattaneo, M. D. and Titiunik, R.: 2014, Robust nonparametric confidence intervals for regression-discontinuity designs, *Econometrica* **82**(6), 2295–2326.

Center for Health and Justice: 2013, No entry: A national survey of criminal justice diversion programs and initiatives. Retrieved from https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=268871 [Date Accessed:

05/23/2017].

Chetty, R., Hendren, N., Jones, M. and Porter, S.: 2018, Race and Economic Opportunity in the United States: An Intergenerational Perspective.

Chiricos, T., Barrick, K., Bales, W. and Bontrager, S.: 2007, The labeling of convicted felons and its consequences for recidivism, *Criminology* **45**(3), 547–581.

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

Dobbie, W., Goldin, J. and Yang, C. S.: 2018, The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges, *American Economic Review* **108**(2), 201–40.

Doleac, J. L.: 2018, Strategies to productively reincorporate the formerly-incarcerated into communities: A review of the literature. IZA Discussion Paper No. 11646.

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

Fabelo, T.: 1997, *Texas Criminal Justice Reforms: The Big Picture in Historical Perspective*, Criminal Justice Policy Council.

Finlay, K.: 2009, Effect of employer access to criminal history data on the labor market outcomes of ex-offenders and non-offenders, *Studies of labor market intermediation*, University of Chicago Press, pp. 89–125.

Gelber, A., Isen, A. and Kessler, J. B.: 2015, The effects of youth employment: Evidence from new york city lotteries, *The Quarterly Journal of Economics* **131**(1), 423–460.

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

Heller, S. B.: 2014, Summer jobs reduce violence among disadvantaged youth, *Science* **346**(6214), 1219–1223.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S. and Pollack, H. A.: 2017, Thinking, fast and slow? some field experiments to reduce crime and dropout in chicago, *The Quarterly Journal of Economics* **132**(1), 1–54.

Henson, S.: 2007a, Kuff: New jail building in Harris County "irresponsible to the point of negligence" [Blog Post], Grits for Breakfast. Retrieved from http://gritsforbreakfast.blogspot.com.au/2007/11/kuff-new-jail-building-in-harris-county.html [Date Published: 11/4/2007, Date Accessed: 10/21/2016].

Henson, S.: 2007b, Texans' Taxation Revulsion vs. their Incarceration Addiction: Which will prevail on county jail building? [Blog Post], Grits for Breakfast. Retrieved from http://gritsforbreakfast.blogspot.com.au/2007/10/texans-taxation-revulsion-vs-their.html [Date Published: 10/14/2007, Date Accessed: 10/21/2016].

Hughes, P. R.: 2005, Revised numbers show jail crowding is worse, *The Houston Chronicle* . Retrieved from http://www.chron.com/news/houston-texas/article/Revised-numbers-show-jail-crowding-is-worse-1525007.php [Date Published: 8/5/2005, Date Accessed: 10/21/2016].

Kling, J., Liebman, J. and Katz, L.: 2007, Experimental analysis of neigborhood effects, *Econometrica* **75**.

Kyckelhahn, T.: 2010, Justice expenditure and employment extracts. Retrieved from

http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5049 [Date Accessed: 08/30/2018].

Lovenheim, M. F. and Owens, E. G.: 2014, Does federal financial aid affect college enrollment? evidence from drug offenders and the higher education act of 1998, *Journal of Urban Economics* **81**, 1–13.

Lozano, J. A.: 2008, Houston DA resigns over e-mail scandal.

Mueller-Smith, M.: 2015, The criminal and labor market impacts of incarceration. Working Paper, Retrieved from http://sites.lsa.umich.edu/mgms/wp-content/uploads/sites/283/2015/09/incar.pdf [Date Accessed: 10/20/2016].

National Reeentry Resource Center: 2018, Nationall inventory of the collateral consequences of conviction, *Technical report*, U.S. Department of Justice's Bureau of Justice Assistance, https://niccc.csgjusticecenter.org/.

Owens, E.: 2009, More time, less crime? Estimating the incapacitative effect of sentence enhancements, *Journal of Law and Economics* **52**(3), 551–579.

Pager, D.: 2003, The mark of a criminal record, *American journal of sociology* **108**(5), 937–975.

Pager, D.: 2008, *Marked: Race, crime, and finding work in an era of mass incarceration*, University of Chicago Press.

Pager, D., Western, B. and Bonikowski, B.: 2009, Discrimination in a low-wage labor market a field experiment, *American sociological review* **74**(5), 777–799.

Raphael, S.: 2014, *The New Scarlet Letter?: Negotiating the US Labor Market with a Criminal Record*, WE Upjohn Institute.

Rogers, B., Stiles, M. and Murphy, B.: 2007, D.A. Rosenthal calls e-mail flap 'a wake-up call'.

Rogers, B., Stiles, M. and Murphy, B.: 2008, More e-mails emerge in Harris County DA scandal.

Selbin, J., McCrary, J. and Epstein, J.: 2017, Unmarked? Criminal Record Clearing and Employment Outcomes, *Journal of Criminal Law and Criminology* **108**(1).

Sessions, J.: 2017, Department charging and sentencing policy, Memorandum, Office of the Attorney General.

Shannon, S., Uggen, C., Schnittker, J., Massoglia, M., Thompson, M. and Wakefield, S.: 2016, The Growth, Scope, and Spatial Distribution of People with Felony Records in the United States, 1980-2010. Conditionally accepted for publication in *Demography*.

Snyder, M.: 2007, Picknickers may share Buffalo Bayou with inmates, *The Houston Chronicle* . Retrieved from http://www.chron.com/news/houston-texas/article/Picnickers-may-share-Buffalo-Bayou-with-inmates-1535996.php [Date Published: 10/10/2007, Date Accessed: 10/21/2016].

Stevenson, M.: 2018, Distortion of justice: How the inability to pay bail affects case outcomes, *Journal of Law, Economics & Organization* . Forthcoming, Available at https://ssrn.com/abstract=2777615 [Date Accessed: 8/14/2018].

TX Legislative Budget Board: 2005, Statwide criminal justice recidivism and revocation rates. Retrieved from http://www.lbb.state.tx.us/Documents/Publications/Policy_Report/Statewide%20Criminal%20Justice%20Recidivism%20and%20Revocation%20Rates2005.pdf [Date Accessed: 1/9/2019].

TX Legislative Budget Board: 2008, Texas community supervision revocation project: A comparison of revoked felons during september 2005 and september 2007. Retrieved from

http://www.lbb.state.tx.us/Documents/Publications/Policy_Report/Texas%20Community%20Supervision%20Revocation%20Project%20Comparison2005-2007.pdf [Date Accessed: 1/9/2019].

Figures

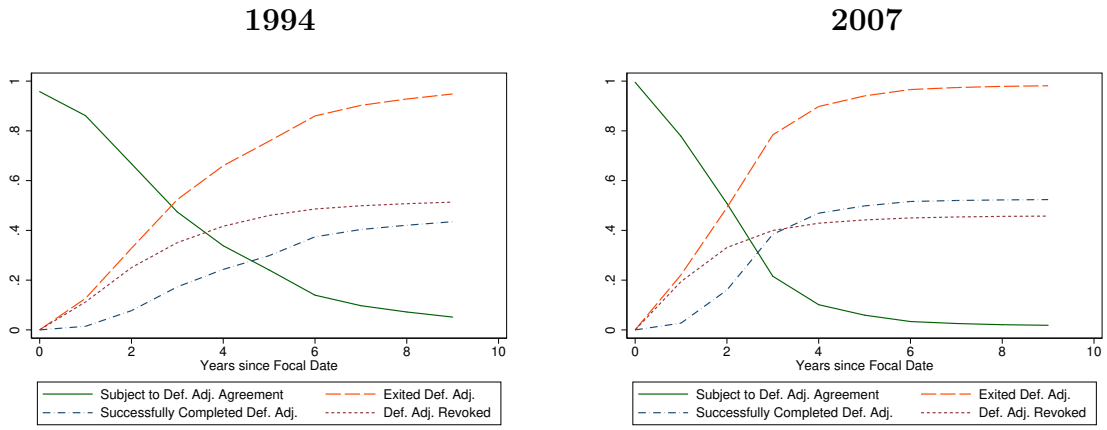


Figure 1: Duration of Deferred Adjudication Agreements

This figure describes the duration of diversion contracts (and associated community supervision) for those in our study sample whose case disposition is a deferred adjudication of guilt. The duration of completed agreements varies across individuals with an average length of 3.87 years for the 1994 sample and 2.35 years for the 2007 sample. Courts can provide early termination of agreements which is why durations can be shorter than sentenced probation lengths. The total proportion of active deferral agreements each year following the case disposition are depicted by the solid green line and, it's converse, the proportion of agreements that are no longer active is shown by the dashed orange line. We also plot the cumulative fraction of agreements that have successfully completed without violation (dashed blue line) and those revoked due to a violation (dotted red line) over time.

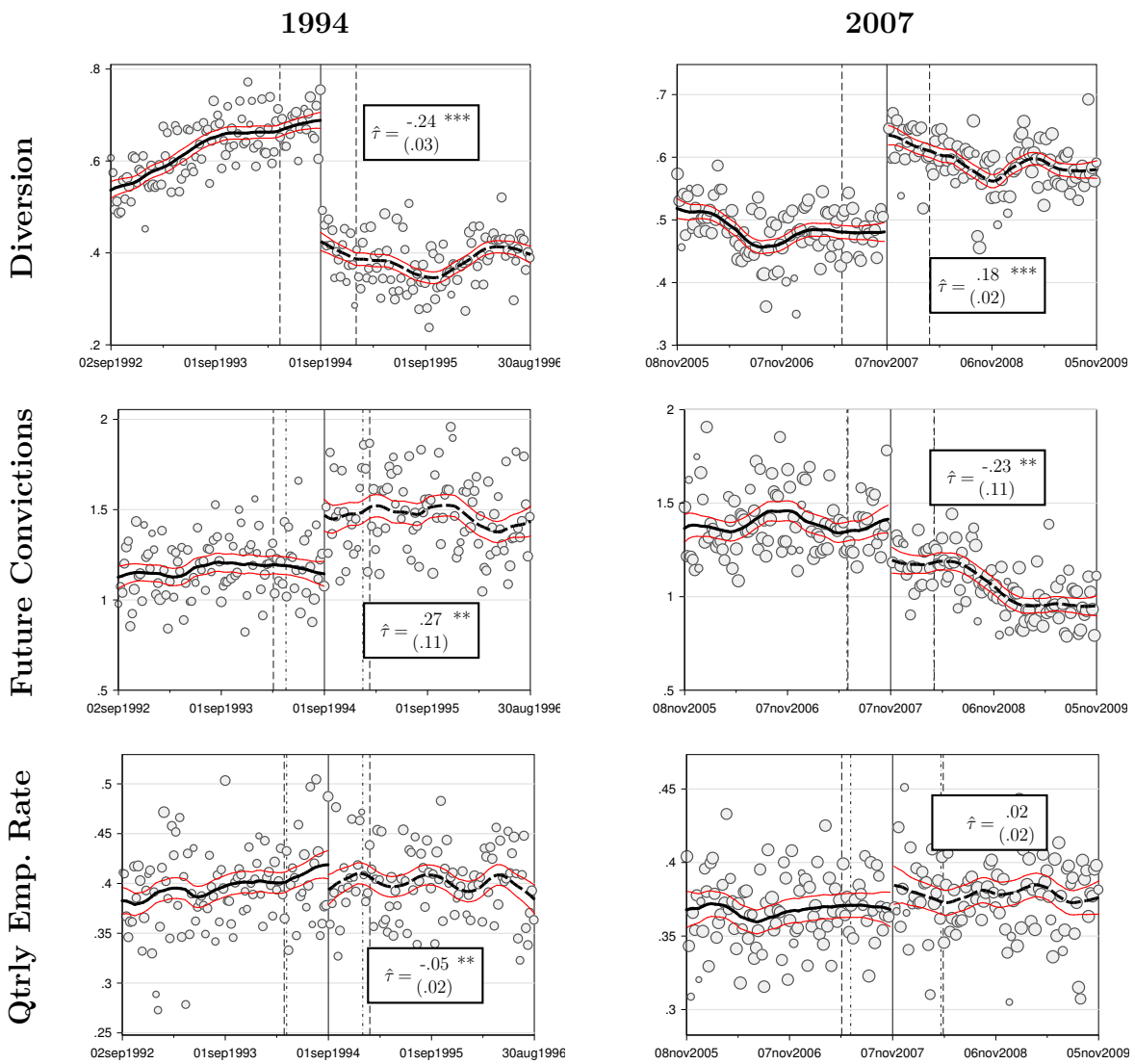


Figure 2: Graphical summary of the reduced form evidence in both the 1994 and 2007 samples

This figure presents reduced form graphical evidence of the changes in diversion, reoffending, and employment associated with the 1994 (left column) and 2007 (right column) natural experiments. The first row depicts the relationship between the running variable and diversion, defined as a court verdict of a deferred adjudication or guilt or a case dismissal. The next two rows plot our key outcomes measured over a ten-year follow-up period: the total number of future Harris County convictions; and, the average quarterly employment rate.

General RD Figure Notes: In general, we present reduced form (sharp RD) results in panels of figures with the left and right columns representing the 1994 and 2007 experiments, respectively. The threshold dates for each experiment are 9/1/1994 (based on the date the charge is filed) and 11/7/2007 (based on the date the case is disposed) and we include two years of data on each side of the threshold date. Scatter points represent week-level bin averages and are weighted by the total number of individual cases. We overlay local polynomial lines and their associated 90 percent confidence bands weighted using an Epanechnikov kernal and 90-day bandwidth. We also report reduced form local-polynomial regression-discontinuity point estimates ($\hat{\tau}$) and standard errors in each plot. The data-driven bandwidth chosen for the reduced form estimates on each side of the discontinuity is depicted by the dashed vertical line and the bandwidth chosen for the fuzzy RD estimates is depicted by the dash-dotted vertical line. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

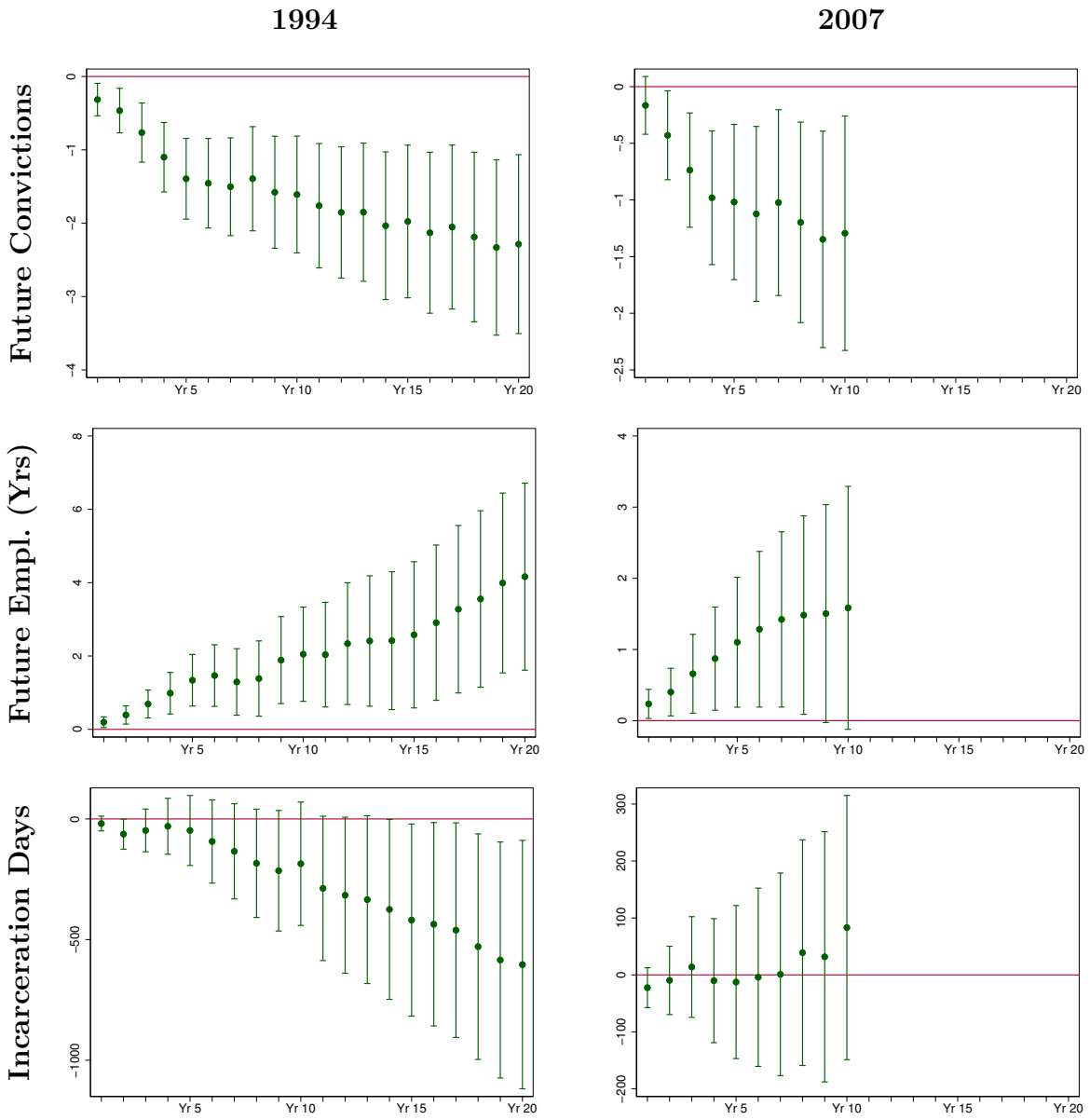


Figure 3: Timeline Impacts

This figure displays fuzzy RD estimates and associated 90% confidence bands for reoffending, employment, and incarceration outcomes that measure cumulative impacts of diversion after each year up through 20 years for the 1994 sample (left column) and 10 years for the 2007 sample (right column). The first two rows depict the estimated impact of diversion on our key outcomes up through each year indicated on the horizontal axis: the total number of future Harris County convictions; and, the average quarterly employment rate. The third row reports the estimated impact of diversion on incarceration days including both actual time served (as opposed to sentenced) in prison and jail. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. Contemporaneous (year-by-year) estimates are presented in Figure A.2, and visual, reduced-form evidence is available in Online Appendix Figure B.7.

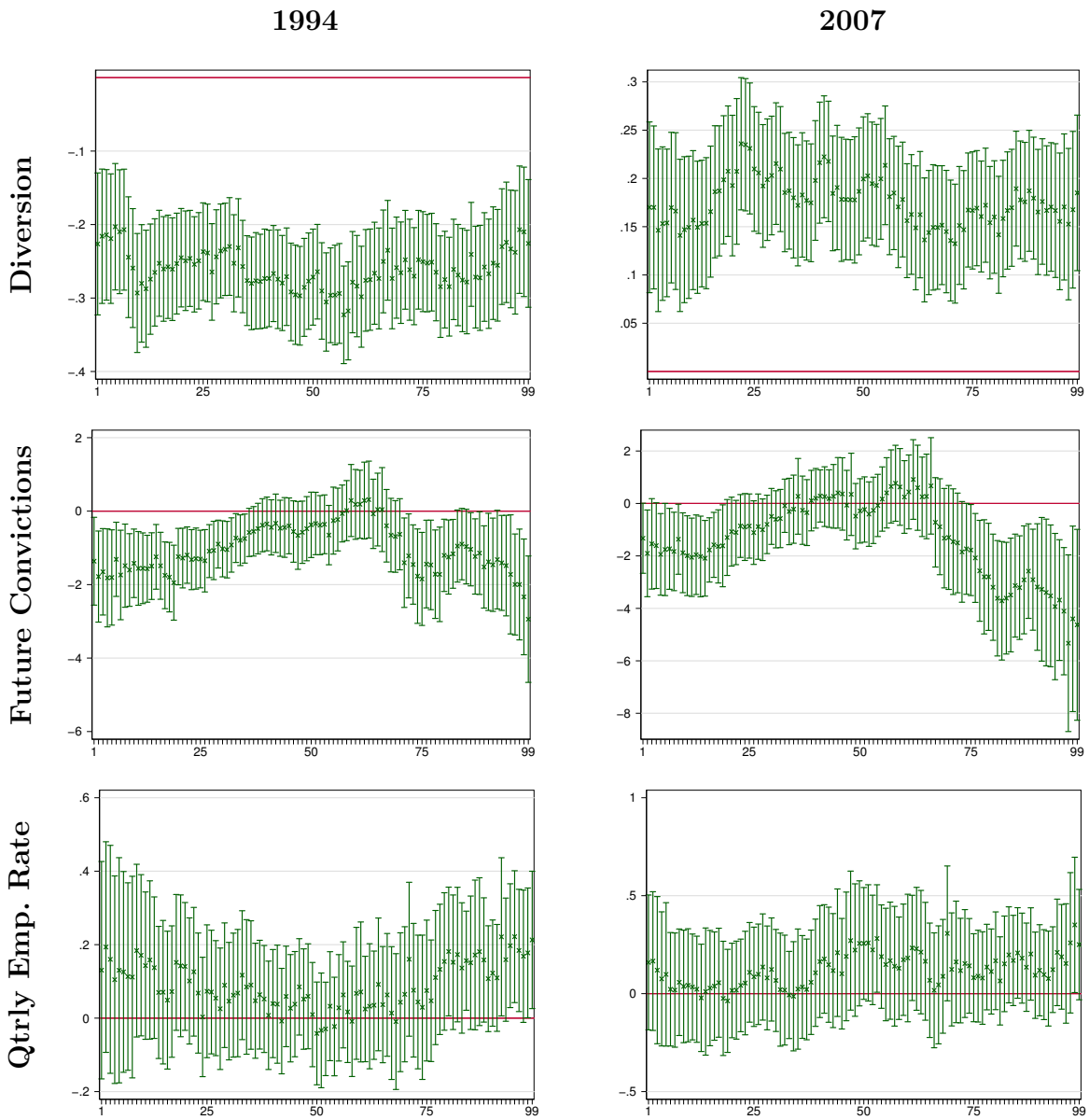


Figure 4: First-Stage and fuzzy RD Impacts over the Recidivism Risk Profile

This figure displays RD estimates and associated 90% confidence bands for diversion, reoffending, and employment outcomes calculated over the quantile function of the predicted recidivism risk score described in Section 5.4 for the 1994 experiment (left column) and 2007 experiment (right column). The first row presents first-stage estimates of the impact of the threshold date on the probability of diversion, defined as a court verdict of a deferred adjudication or guilt or a case dismissal. The next two rows plot fuzzy RD estimated impacts of diversion on our key outcomes measured over a ten-year follow-up period: the total number of future Harris County convictions; and, the average quarterly employment rate. Each coefficient reflects a distinct local polynomial RD estimate from regression centered at the focal percentile using a uniform kernel with a 40 percentile bandwidth. Estimates below the 20th and above the 80th percentiles will reflect narrower, asymmetric bandwidths. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

Tables

Table 1: Balance Tests

	τ^{94}	τ^{07}	$\frac{\tau^{07} - \tau^{94}}{2}$
Caseload Size	-3.73	-2.24	0.75
	(2.52)	(3.57)	(2.19)
	[28.6]	[50.4]	
Total Prior Misd. Conv.	-0.0064	0.033	0.020
	(0.058)	(0.052)	(0.039)
	[0.55]	[0.70]	
Age at Charge	-0.46	0.50	0.48
	(0.49)	(0.55)	(0.37)
	[28.8]	[29.4]	
Sex = Male	0.044*	0.023	-0.011
	(0.027)	(0.020)	(0.017)
	[0.68]	[0.73]	
Race/Ethn. = Black, not Hisp.	-0.036	-0.027	0.0043
	(0.027)	(0.022)	(0.017)
	[0.46]	[0.38]	
Race/Ethn. = Hispanic	-0.020	0.025	0.023
	(0.023)	(0.021)	(0.016)
	[0.21]	[0.31]	
Crime Type = Property	0.0031	0.0095	0.0032
	(0.027)	(0.022)	(0.017)
	[0.54]	[0.28]	
Crime Type = Drug	0.000089	-0.015	-0.0075
	(0.028)	(0.026)	(0.019)
	[0.46]	[0.45]	
Crime Type = Violent	–	-0.0064	–
		(0.014)	
		[0.10]	
Recidivism Risk Score	0.069	-0.029	-0.049
	(0.043)	(0.042)	(0.030)
	[1.27]	[1.29]	
Social Sec. Number Unrecorded	0.00061	0.047**	0.023
	(0.022)	(0.022)	(0.015)
	[0.22]	[0.22]	
Observations	31,254	52,701	.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents sharp RD estimates for pre-determined characteristics as recorded in the criminal court records from the Harris County District Court. The calculation of the recidivism risk score is described in Section 5. The first two columns report results from the 1994 and 2007 experiments, respectively. These estimates correspond to plots in Online Appendix Figure B.1. The third column calculates the average imbalance across the two experiments between the low-diversion and high-diversion period to assess whether there is any combined evidence of differential sorting. This calculation subtracts the 1994 point estimate since diversion was reduced in this experiment.

General RD Table Notes: Unless otherwise noted, all tables present local-polynomial regression discontinuity point estimates, standard errors in parentheses, and control-group means in square brackets. For all sharp RD estimates, control group means represent the average for defendants just to the left of the threshold dates within the data-driven bandwidth. For all fuzzy RD estimates, we report *complier control means* (CCM) as described in Section 4. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

Table 2: Impact of Diversion on Disposition and Sentencing Outcomes

	ϕ^{94}	ϕ^{07}	$\frac{\phi^{94}+\phi^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
Panel A: Case Disposition				
Def. Adjudication of Guilt	1.08***	0.59***	0.83***	0.49***
	(0.098)	(0.10)	(0.071)	(0.14)
	[0.00]	[0.00]		
Case Dismissal	-0.10	0.41***	0.16**	-0.52***
	(0.10)	(0.099)	(0.071)	(0.14)
	[0.00]	[0.00]		
Panel B: Sentencing				
Incarceration Sentence	-0.19**	-0.97***	-0.58***	0.79***
	(0.086)	(0.056)	(0.051)	(0.10)
	[0.16]	[1.00]		
Probation Sentence	-0.017	0.67***	0.33***	-0.69***
	(0.11)	(0.11)	(0.079)	(0.16)
	[1.00]	[0.00]		
Fined	-0.30**	0.46***	0.081	-0.76***
	(0.12)	(0.13)	(0.089)	(0.18)
	[0.96]	[0.00]		
Drug Treatment	0.032	0.083	0.058	-0.051
	(0.045)	(0.068)	(0.041)	(0.081)
	[0.030]	[0.00]		
Observations	31,131	52,792		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the effect of diversion on case dispositions and sentencing for the focal felony charge. The first two columns report results from the 1994 and 2007 experiments respectively. The third column calculates the average impact across the two experiments and the fourth column tests whether we can statistically distinguish our fuzzy RD estimates for the 1994 experiment from the corresponding estimates for the 2007 experiment.

Table 3: Impact of Diversion on Reoffending Behavior over 10 years

	ϕ^{94}	ϕ^{07}	$\frac{\phi^{94}+\phi^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
Panel A: Overall				
Any Convictions	-0.31***	-0.32**	-0.32***	0.013
	(0.12)	(0.14)	(0.091)	(0.18)
	[0.70]	[0.64]		
Total Convictions	-1.61***	-1.29**	-1.45***	-0.32
	(0.48)	(0.63)	(0.40)	(0.79)
	[2.09]	[2.31]		
Panel B: Crime Type				
Drug Convictions	-0.71***	-0.19	-0.45**	-0.52
	(0.26)	(0.24)	(0.18)	(0.35)
	[0.80]	[0.63]		
Property Convictions	-0.50**	-0.47	-0.48**	-0.039
	(0.22)	(0.34)	(0.20)	(0.41)
	[0.56]	[0.70]		
Violent Convictions	-0.11	-0.28*	-0.19**	0.17
	(0.099)	(0.15)	(0.089)	(0.18)
	[0.24]	[0.30]		
Panel C: Offense Level				
Misdemeanor Convictions	-0.64**	-1.09**	-0.87***	0.46
	(0.29)	(0.45)	(0.27)	(0.54)
	[0.94]	[1.32]		
Felony Convictions	-1.07***	-0.51	-0.79***	-0.56
	(0.30)	(0.37)	(0.24)	(0.47)
	[1.18]	[1.04]		
Observations	31,131	52,792		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the impact of diversion on reoffending outcomes measured using a 10-year follow-up period. The first panel reports results with the total number of convictions in the ten-year follow-up period at the extensive (*Any Conviction*) and intensive (*Total Convictions*) margins. Panel B splits total convictions by type: drug, property, and violent. Panel C reports results separating convictions by severity. *General RD Table Notes* from Table 1 apply.

Table 4: Impact of Diversion on Labor Market Outcomes over 10 years

	ϕ^{94}	ϕ^{07}	$\frac{\phi^{94}+\phi^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
Panel A: Employment				
Qtrly Employment Rate	0.20***	0.16	0.18***	0.047
	(0.078)	(0.10)	(0.065)	(0.13)
	[0.37]	[0.31]		
Earn $\geq$ 100% Fed. Pov. Level	0.14**	0.12	0.13**	0.020
	(0.069)	(0.095)	(0.059)	(0.12)
	[0.23]	[0.21]		
Panel B: Earnings				
Log Total Earnings	1.45	2.21*	1.83**	-0.76
	(1.02)	(1.34)	(0.84)	(1.68)
	[8.73]	[7.87]		
Total Earnings	82,262**	40,818	61,540*	41,444
	(38,649)	(53,793)	(33,119)	(66,238)
	[91,340]	[101,363]		
Panel C: Tenure				
Max Employment Spell (Quarters)	1.51	5.81**	3.66**	-4.31
	(1.57)	(2.76)	(1.59)	(3.18)
	[6.25]	[5.88]		
Max Cont. Earning Spell (Quarters)	3.44	6.40*	4.92**	-2.96
	(2.29)	(3.58)	(2.12)	(4.25)
	[9.85]	[8.18]		
Observations	24,042	39,674		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the impact of diversion on labor market outcomes measured using a ten-year follow-up period. The labor market outcomes vary across each panel. Panel A presents estimated effects on the quarterly employment rate in the first row and quarterly employment at or above the federal poverty level in the second row. All wage outcomes are adjusted to 2018 US dollars using the Houston MSA CPI index and the 2018 Federal Poverty Level for a single adult is ($12,140). Panel B presents results on a total earnings earnings dependent variable in log form (first row) and the raw level (second row). Finally, Panel C presents results for two dependent variables measuring employment stability by calculating the longest spell (measured in total quarters) of uninterrupted employment at a single employer (first row) or consecutive earnings (second row) during the follow-up period. *General RD Table Notes* from Table 1 apply.

Table 5: Impact of Diversion on Intersection Outcomes over 10 years

	ϕ^{94}	ϕ^{07}	$\frac{\tau^{94}+\tau^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
Qtrly Employment Rate $\geq$ 50%; No Convictions	0.29***	0.20*	0.25***	0.086
	(0.10)	(0.12)	(0.080)	(0.16)
	[0.14]	[0.084]		
Qtrly Employment Rate $\geq$ 50%; 1+ Convictions	-0.038	0.027	-0.0056	-0.065
	(0.088)	(0.11)	(0.071)	(0.14)
	[0.22]	[0.15]		
Qtrly Employment Rate $<$ 50%; No Convictions	0.098	0.084	0.091	0.014
	(0.081)	(0.13)	(0.077)	(0.15)
	[0.14]	[0.29]		
Qtrly Employment Rate $<$ 50%; 1+ Convictions	-0.22**	-0.41**	-0.32***	0.20
	(0.093)	(0.16)	(0.094)	(0.19)
	[0.50]	[0.47]		
Observations	24,042	39,674		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the impact of diversion on four mutually-exclusive categories considering labor market outcomes and criminal offending outcomes jointly over the ten-year follow-up period. The first row estimates the probability of more-than 50% employment *AND* no future convictions over the ten-year follow-up period while the fourth row reports estimates of the impact of diversion on a less-than 50% employment rate *AND* at least one future conviction. The other two rows report effects on different combinations of these employment and reoffending measurements. *General RD Table Notes* from Table 1 apply.

Table 6: Comparison of Reduced Form Estimates for Study Sample and other Felony Defendants in Harris County

	τ^{94}	τ^{07}	τ^{94}	τ^{07}
Court Deferral	-0.27***	0.12***	-0.018	0.020
	(0.027)	(0.025)	(0.016)	(0.013)
	[0.53]	[0.35]	[0.25]	[0.12]
Case Dismissal	0.0040	0.040**	0.0026	0.048***
	(0.020)	(0.016)	(0.019)	(0.016)
	[0.16]	[0.14]	[0.26]	[0.15]
Incarceration Sentence	0.047*	-0.17***	-0.014	-0.070***
	(0.025)	(0.025)	(0.018)	(0.016)
	[0.17]	[0.50]	[0.39]	[0.73]
Probation Sentence	-0.0044	0.12***	0.053**	0.032**
	(0.028)	(0.025)	(0.021)	(0.014)
	[0.66]	[0.37]	[0.35]	[0.14]
Fine Sentence	0.081***	0.066***	0.049**	0.015
	(0.031)	(0.025)	(0.020)	(0.013)
	[0.45]	[0.27]	[0.29]	[0.10]
Any Convictions	0.071***	-0.046**	0.013	0.0020
	(0.023)	(0.023)	(0.018)	(0.020)
	[0.46]	[0.50]	[0.50]	[0.59]
Total Convictions	0.27**	-0.23**	0.084	-0.0059
	(0.11)	(0.11)	(0.065)	(0.095)
	[1.14]	[1.42]	[1.11]	[1.92]
Qrtly Employment Rate	-0.051**	0.024	0.026	-0.0031
	(0.021)	(0.018)	(0.016)	(0.013)
	[0.42]	[0.37]	[0.34]	[0.25]
Total Earnings	-21,204**	6,179	7,583	-8,900
	(9,849)	(9,580)	(7,245)	(7,118)
	[113,077]	[108,923]	[93,087]	[70,135]
Observations	31,334	52,846	56,378	76,339

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents reduced form estimates quantifying how disposition, recidivism and labor markets change across the 1994 and 2007 discontinuities for our main study sample and all other felony defendants in Harris County. In order to avoid defendants repeatedly appearing in the analysis when examining behavior in the broader caseload, which is not constrained to first time defendants, we collapse observations associated with the same individual to their earliest date when multiple charges or dispositions occur within a 1 year period. *General RD Table Notes* from Table 1 apply.

A Print Appendix

A.1 Control Complier Means in the RD Design

For simplicity, let us assume that approaching the discontinuity from below represents the low diversion regime (i.e. control) and approaching the discontinuity from above represents the high diversion regime (i.e. treatment). Let C correspond to whether an observation is a complier or not. The CCM can thus be defined as:

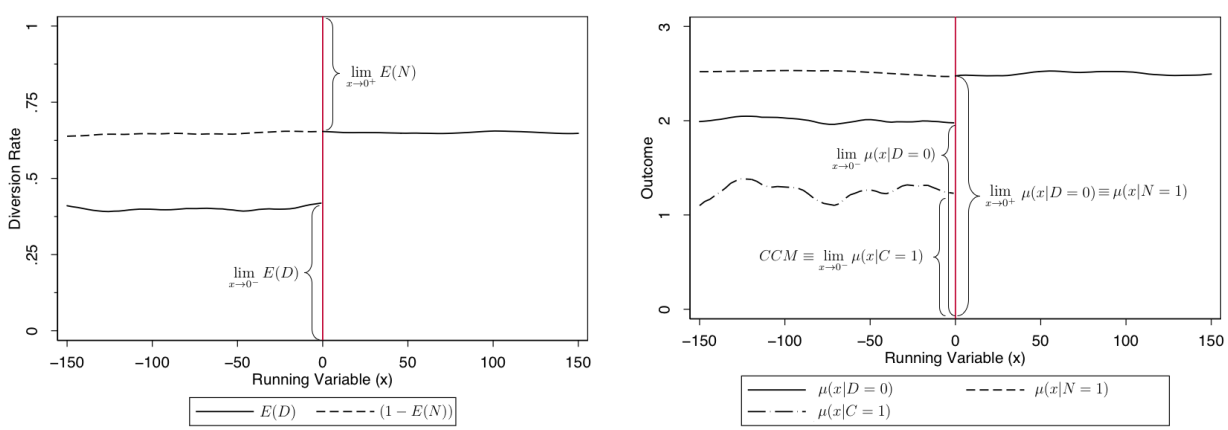
$$CCM \equiv \lim_{x \to 0^-} \mu(x|C = 1)$$

The challenge in calculating this estimator is that complier status is not directly observable. What is observable, however, is the average outcome among the non-diverted population (i.e. $D = 0$) approaching the discontinuity from below. This will be a weighted average of the outcomes among compliers and *never-takers* (i.e. those never diverted regardless of the discontinuity), which we define as $N = 1$.

$$\lim_{x \to 0^-} \mu(x|D = 0) = \lim_{x \to 0^-} \mu(x|C = 1)\,(1 - \omega) + \lim_{x \to 0^-} \mu(x|N = 1)\,(\omega)$$

The weight (ω) indicates the never-taker share of the non-diverted population: $\frac{\lim_{x \to 0^-} E(N)}{1 - \lim_{x \to 0^-} E(D)}$.

While we can calculate the proportion $(1 - E(D))$ and outcome averages $(\mu(x|D = 0))$ for the non-diverted population as we approach the threshold from below, we cannot disentangle the contribution of never takers from that of the compliers. However, we can instead use information from those who are not diverted above the threshold since, by the monotonicity assumption of our RD design, this is the population of never-takers. This simply involves replacing $\lim_{x \to 0^-} E(N)$ and $\lim_{x \to 0^-} \mu(x|N = 1)$ with $\lim_{x \to 0^+} E(N)$ and $\lim_{x \to 0^+} \mu(x|N = 1)$ respectively. For the sake of simplicity and transparency, we approximate the local average approaching the cutoff in the equations discussed using a fixed bandwidth of 2.5 months and a uniform kernel.



(a) Diversion rates in the overall caseload (b) Average outcomes in the non-diverted populations

An illustration of the control complier mean procedure

Our illustration on the next page depicts the CCM procedure applied to a RD setting as described in Section 4. It utilizes artificial data for clarity. Solid black lines represent information that is directly observable to the econometrician. Dashed and dotted lines represent statistics that are not observable, but can be inferred

based on the assumptions of the research design. Recall,

$$CCM \equiv \lim_{x \to 0^-} \mu(x|C=1) = \frac{\lim_{x\to 0^-} \mu(x|D=0) - \lim_{x\to 0^-} \mu(x|N=1)\left(\frac{\lim_{x\to 0^-} E(N)}{1-\lim_{x\to 0^-} E(D)}\right)}{\left(1 - \frac{\lim_{x\to 0^-} E(N)}{1-\lim_{x\to 0^-} E(D)}\right)}.$$

Two of the right-hand side terms can be directly estimated without assumptions: $\lim_{x\to 0^-} \mu(x|D=0)$ and $\lim_{x\to 0^-} E(D)$. The first term is the average outcome for those not diverted approaching the discontinuity from below. The second is the rate of diversion approaching the discontinuity from below. Both can be observed because these are caseload-wide statistics and do not rely on unobservable statuses like being a *complier* or *never-taker*.

The remaining terms, $\lim_{x\to 0^-} \mu(x|N=1)$ and $\lim_{x\to 0^-} E(N)$, cannot be directly observed in the data. We can, however, instead approximate their values at cutoff using observations approaching the discontinuity from above: $\lim_{x\to 0^+} \mu(x|N=1)$ and $\lim_{x\to 0^+} E(N)$ respectively. This is feasible because only never-takers will not be diverted after the threshold based on the monotonicity assumption.

The figure highlights the pieces of information that are used to execute the CCM calculation. Together, these present our best estimate of the potential outcome that compliers would have experienced in the absence of diversion. The approach yields similar estimates to methods proposed in Abadie (2003). Differences arise though in that our CCM calculation utilizes a stable bandwidth definition as the treatment effect estimate, which we find to be an attractive feature.

A.2 Appendix Figures

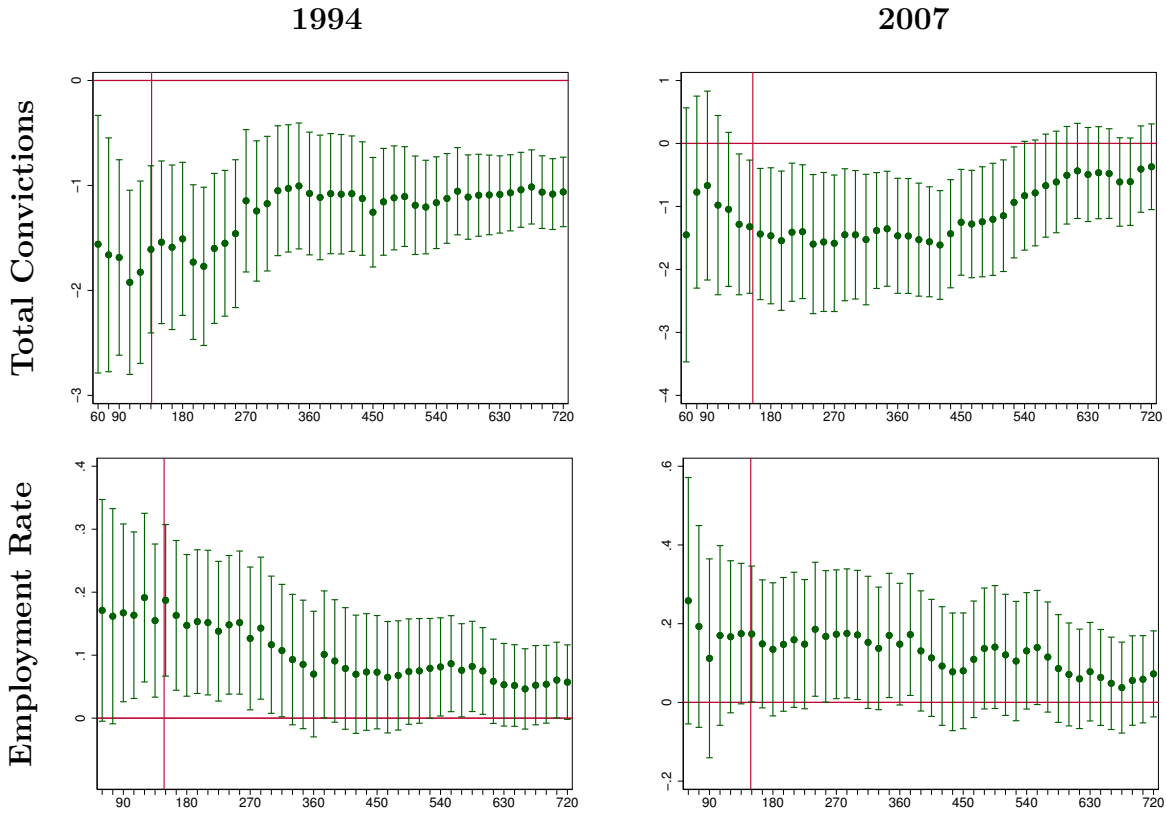


Figure A.1: Estimated effects by RD bandwidth

This figure displays fuzzy RD estimates and associated 90% confidence bands for future conviction and employment outcomes varying a fixed bandwidth in 15 day intervals from 60 days on each side of the threshold date to 720 days. The red vertical line in each plot indicates the average data-driven bandwidth used in our primary regression results in Tables 3 and 4 and reported in Table B.2.

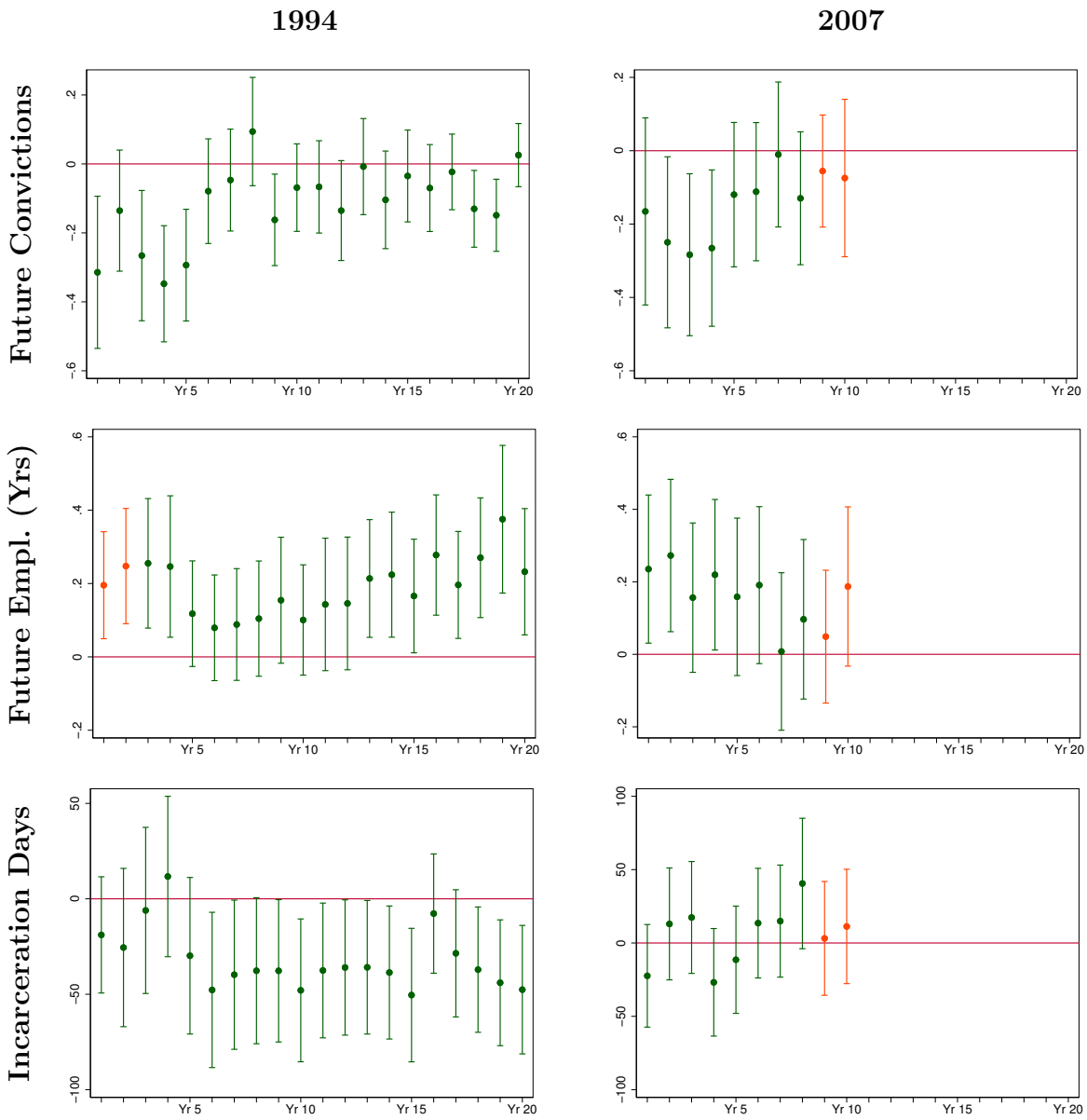


Figure A.2: Contemporaneous impacts by year of follow-up

This figure displays fuzzy RD estimates and associated 90% confidence bands for reoffending, employment, and incarceration outcomes that measure year-by-year (contemporaneous) impacts of diversion for each year period up through 20 years for the 1994 sample (left column) and 10 years for the 2007 sample (right column). The first two rows depict the estimated impact of diversion on our key outcomes up through each year indicated on the horizontal axis: the total number of future Harris County convictions; and, the average quarterly employment rate within each follow-up year. The third row reports the estimated impact of diversion on the number of incarceration days each year. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. Cumulative estimates are presented in Figure 3.

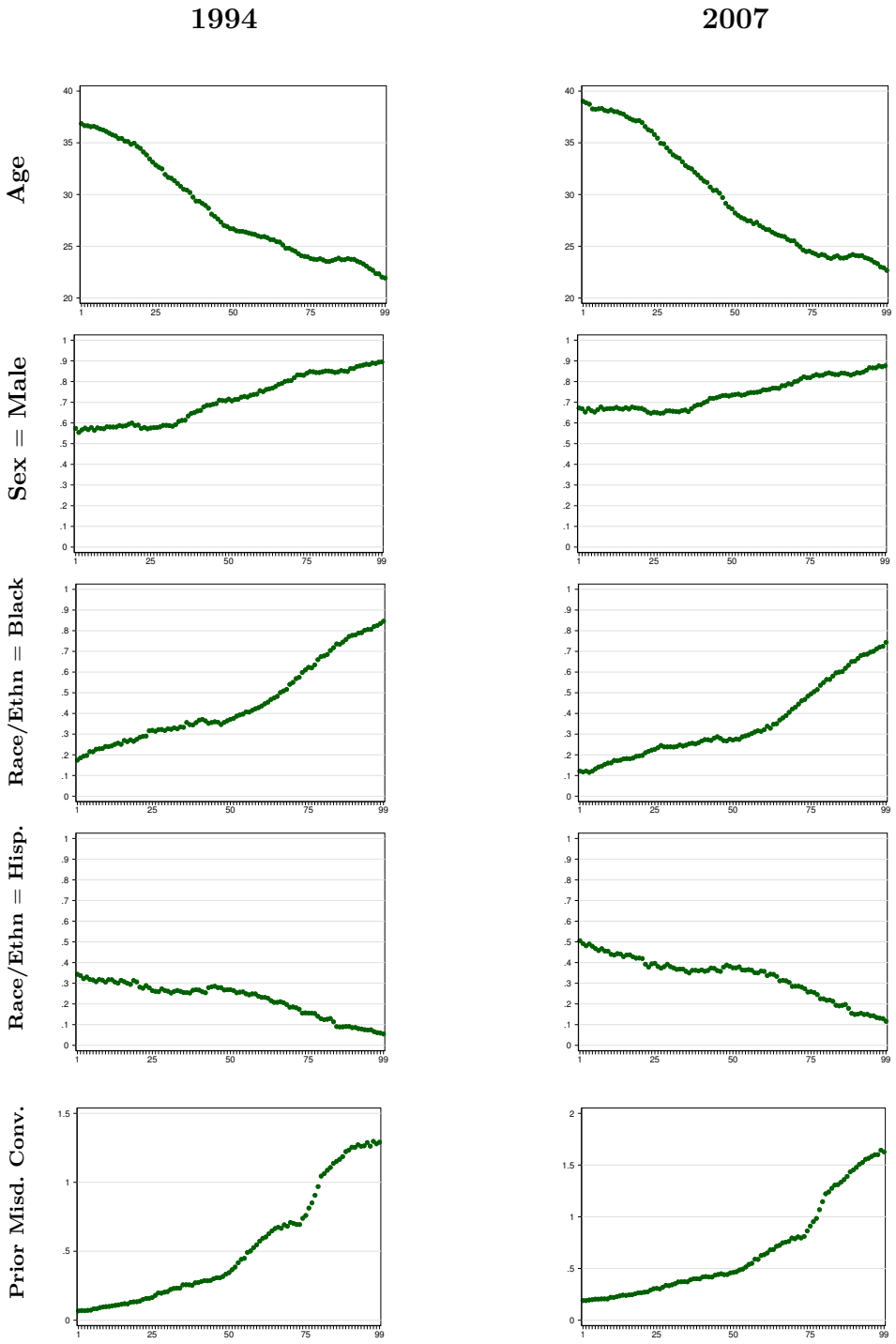


Figure A.3: Smoothed Defendant Characteristics over the Recidivism Risk Profile

This figure plots mean defendant demographic and criminal history characteristics over the quantile function of the predicted recidivism risk score described in Section 5.4. Estimated impacts of diversion across this distribution are presented in Figure 4.

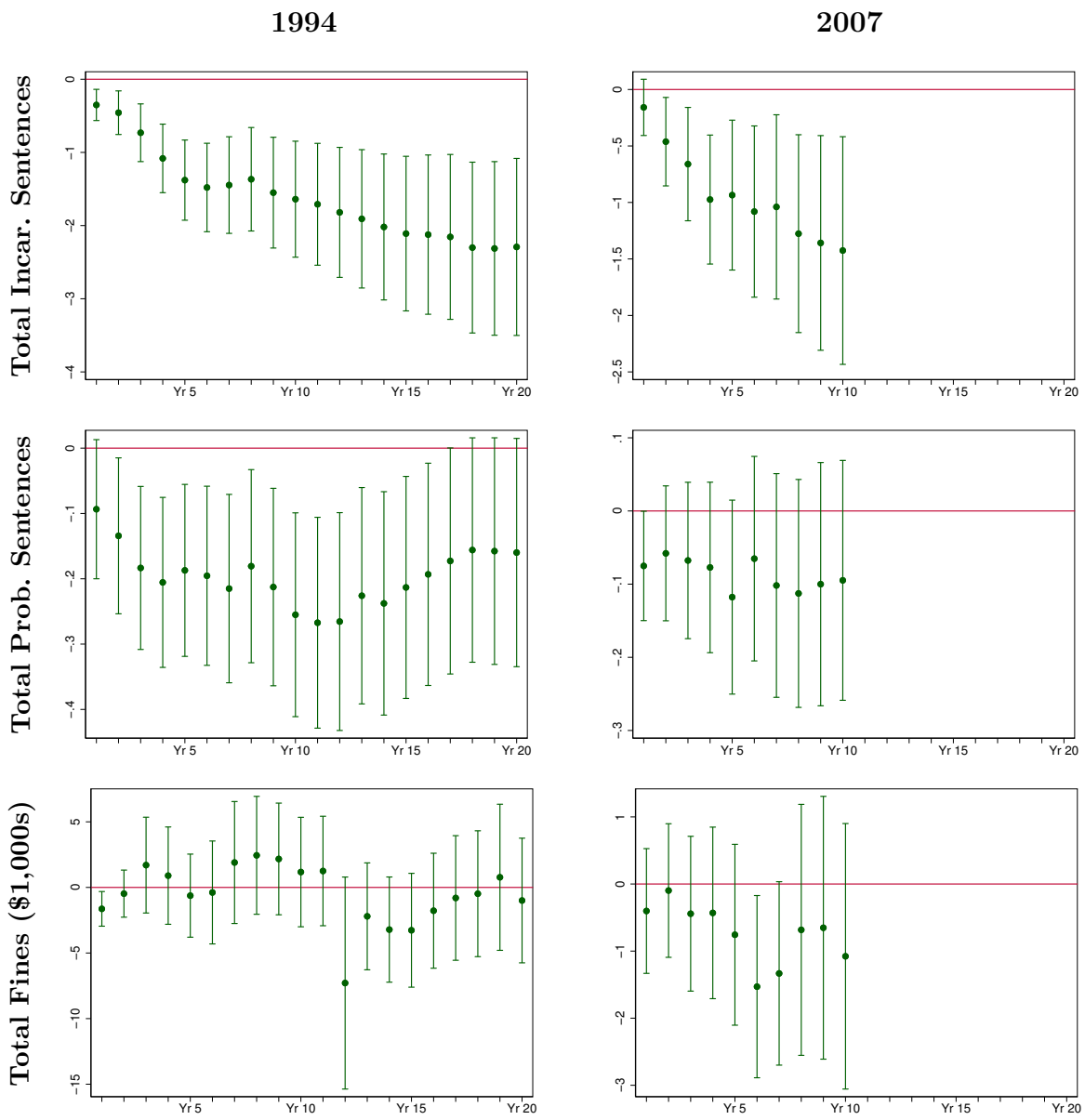


Figure A.4: Accumulation of new sanctions for future criminal activity

This figure displays fuzzy RD estimates and associated 90% confidence bands for incarceration, probation and fine sanctions associated with future criminal activity. The coefficients measure cumulative impacts of diversion for each year period up through 20 years for the 1994 sample (left column) and 10 years for the 2007 sample (right column). The pattern of results are consistent with the amplification effect mechanism described in Section 7. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

A.3 Appendix Tables

Table A.1: Imputation rates in outcome data by follow-up period

Outcome Year	Convictions (1994 Sample)	Earnings (1994 Sample)	Convictions (2007 Sample)	Earnings (2007 Sample)
1	0	0.38	0	0
2	0	0.15	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	0.23	0.16
10	0	0	0.48	0.42

This table describes the share of observations by outcome variable, follow-up year and estimation sample that include imputed values based on our methodology described in Section 3.2. We impute data due to the timing of available data. In the 1994 sample, we cannot observe initial labor market outcomes for those charged prior to Jan. 1, 1994; in the 2007 sample, we cannot observe end-of-decade criminal justice and labor market outcomes for those disposed after Oct. 1, 2007 and Jan. 1, 2008 respectively.

Table A.2: Alternative Measures of Recidivism

	ϕ^{94}	ϕ^{07}	$\frac{\phi^{94}+\phi^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
Total Bookings	-1.38**	-1.57*	-1.47***	0.19
	(0.65)	(0.81)	(0.52)	(1.03)
	[2.57]	[2.83]		
Total Charges	-1.64***	-1.60**	-1.62***	-0.039
	(0.53)	(0.70)	(0.44)	(0.88)
	[2.31]	[2.76]		
Total Convictions	-1.61***	-1.29**	-1.45***	-0.32
	(0.48)	(0.63)	(0.40)	(0.79)
	[2.09]	[2.31]		
Total TDPS CCH Conv.	-1.09**	-1.37***	-1.23***	0.28
	(0.48)	(0.40)	(0.31)	(0.62)
	[1.64]	[1.47]		
Observations	31,131	52,792		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the effect of diversion on alternative reoffending outcomes from the estimated impacts on Harris County convictions reported in the third row. The first three rows report estimates for various measures of recidivism available from Harris County: jail bookings, charges filed, and convictions (baseline result). The fourth row reports the impact of diversion on the total number of convictions recorded in the statewide CCH conviction database. *General RD Table Notes* from Table 1 apply.

Table A.3: Comparison of Harris and non-Harris County Recidivism in the TDPS CCH

	ϕ^{94}	ϕ^{07}	$\frac{\phi^{94}+\phi^{07}}{2}$	$H_0 : \phi^{94} = \phi^{07}$
TDPS CCH Conv. (All Counties)	-1.09**	-1.37***	-1.23***	0.28
	(0.48)	(0.40)	(0.31)	(0.62)
	[1.64]	[1.47]		
TDPS CCH Conv. (Harris County)	-1.07***	-0.99***	-1.03***	-0.083
	(0.41)	(0.35)	(0.27)	(0.54)
	[1.40]	[1.22]		
TDPS CCH Conv. (Non-Harris Counties)	-0.055	-0.49***	-0.27**	0.44*
	(0.19)	(0.17)	(0.13)	(0.26)
	[0.23]	[0.25]		
Observations	31,131	52,792		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents fuzzy RD estimates of the impact of diversion on the total number of convictions recorded in the statewide CCH conviction database in the first row and then separates outcomes into Harris County convictions and non-Harris County convictions to evaluate spatial spillovers. *General RD Table Notes* from Table 1 apply.

Table A.4: Quantifying potential role of revocations in recidivisms findings

	Extensive Margin		*Intensive Margin*	
	Any Convictions	Any Conv. or Revocations	Total Convictions	Total Conv. and Revocations
Diversion, 1994 Sample	-0.31***	-0.49***	-1.61***	-1.79***
	(0.12)	(0.12)	(0.48)	(0.50)
	[0.70]	[0.83]	[2.09]	[2.42]
Diversion, 2007 Sample	-0.32**	-0.21	-1.29**	-1.19*
	(0.14)	(0.14)	(0.63)	(0.64)
	[0.64]	[0.64]	[2.31]	[2.31]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table repeats the fuzzy RD estimates of the impact of diversion on our extensive and intensive-margin conviction outcomes from Table 3 in the first and third columns in order to compare these baseline estimates to those reported in the second and fourth columns where our outcome variable counts both convictions and a fraction of non-technical (criminal) probation revocations as recidivism. We use reports from the Texas Legislative Board to obtain proxies for the fraction of criminal revocations for each estimation sample. TX Legislative Budget Board (2008) reports that 36.8 % of the 2007 revocations for felony community supervision arrangements were for non-technical (criminal violations) which we use for the 2007 sample. We were unable to find a similar statistic for Harris County in 1994 or 1995 and therefore use the state-wide average non-techical revocation rate of 45% between 1999 and 2005 as reported by TX Legislative Budget Board (2005). This exercise allows us to quantify the potential role of prosecutors deciding not to pursue new convictions for offenses that violate a diversion agreement. This is a conservative approach given that many non-techincal revocations were actually separately prosecuted in practice and as a result would be double counted as both revocations and new convictions in this exercise. *General RD Table Notes* from Table 1 apply.

Table A.5: Quantifying potential role of incapacitation in employment findings

	Qtrly Employment Rate		
Diversion, 1994 Sample	0.20***	0.20**	0.13*
	(0.078)	(0.082)	(0.078)
Diversion, 2007 Sample	0.16	0.17*	0.081
	(0.10)	(0.10)	(0.10)
Specification	Empl_i not recoded (Baseline)	Recode $\text{Empl}_{i,q} = \overline{\text{Empl}}_i$ when $\text{Incar}_{i,q} > 0$	Recode $\text{Empl}_{i,q} = 1$ when $\text{Incar}_{i,q} > 0$

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table repeats the fuzzy RD estimates of the impact of diversion on our quarterly employment outcome from Table 4 in the first column in order to compare these baseline estimates to those reported in the second and third columns where our outcome variable recodes labor market results during quarters of incarceration to explore how employment might have evolved in the absence of incapacitation. The second column replaces employment status with the average employment rate in the year prior to their current incarceration spell. The third column recodes employment status to equal 1 for all quarters when we observe an individual incarcerated. This last exercise assumes that all individuals would have found employment in the formal sector if not for incapacitation, a scenario which is unlikely to occur given the observed pre-incarceration employment rates. *General RD Table Notes* from Table 1 apply.

Table A.6: Quantifying potential role of missing SSNs in employment findings

	Missing Dropped	Missing imputed
Panel A: 1994 Sample		
Fuzzy RD: Qtrly Employment Rate	0.20***	0.16**
	(0.078)	(0.073)
Fuzzy RD: Total Earnings	82,262**	72,516**
	(38,649)	(34,666)
Fuzzy RD: Log Total Earnings	1.45	1.48*
	(1.02)	(0.88)
Panel B: 2007 Sample		
Fuzzy RD: Qtrly Employment Rate	0.16	0.13
	(0.10)	(0.079)
Fuzzy RD: Total Earnings	40,818	20,095
	(53,793)	(44,681)
Fuzzy RD: Log Total Earnings	2.21*	2.15*
	(1.34)	(1.11)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table repeats the fuzzy RD estimates of the impact of diversion on employment and earnings labor market outcomes from Table 4 in the first column in order to compare these baseline estimates to those reported in the second column where we impute earnings for defendants that we are unable to match to earnings records due to missing SSNs in the Harris County court records. The second column reports estimates where we impute earnings each quarter for defendants with missing SSNs based on a model that includes demographic characteristics, and prior criminal histories. We impute quarterly earnings separately for pre- and post-discontinuity observations. *General RD Table Notes* from Table 1 apply.

Table A.7: Testing Overidentification and Comparing Estimated Impacts of Def. Adj. and Case Dismissal

Panel A: Overidentification Tests			
	Total Future Convictions (N = 16,039)		
Diversion	-1.06***	-1.24***	-1.24***
	(0.15)	(0.32)	(0.33)
Hansen J statistic (P-value)	0.64	0.94	0.95
	Qtrly Employment Rate (N = 12,306)		
Diversion	0.038	0.16***	0.16***
	(0.026)	(0.052)	(0.053)
Hansen J statistic (P-value)	0.38	0.95	0.94
Panel B: Comparing impacts of Def. Adj. and Case Dismissals			
	Total Future Convictions (N = 16,039)		
Def. Adj. of Guilt	-1.04***	-1.25***	-1.25***
	(0.15)	(0.33)	(0.33)
Case Dismissal	-1.75	-1.13	-1.16
	(1.50)	(1.46)	(1.44)
H_0: Def. Adj = Dismissal (P-Value)	0.64	0.94	0.95
	Qtrly Employment Rate (N = 12,306)		
Def. Adj. of Guilt	0.032	0.16***	0.16***
	(0.028)	(0.053)	(0.054)
Case Dismissal	0.29	0.15	0.15
	(0.30)	(0.23)	(0.23)
H_0: Def. Adj = Dismissal (P-Value)	0.40	0.95	0.94
Specification Details:			
Sample-specific Intercepts	x	x	x
Sample-specific Trend		x	x
Sample-specific Post-Trend			x

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents IV estimates from various specifications using the pooled 1994 and 2007 samples to further evaluate the validity of our research design and to compare differential effects across the two types of diversion (deferred adjudication and case dismissals). In Panel A, we estimate the effects of our endogenous variable, diversion, using the threshold date indicators as two exogenous instruments. This specification allows us to assess the joint validity of the instruments through a Hansen J Statistic test. In Panel B, we use the two threshold-date indicators as exogenous instruments for two endogenous case disposition variables, deferred adjudication and case dismissal. The baseline specification in the first column includes our basic set of individual-level controls and an indicator variable for whether the observation is part of the 1994 or 2007 study sample; the second column allows for 1994 and 2007-specific trends in the running variable in addition to the controls from the first column; and the third column adds 1994 and 2007-specific post-threshold trends in the running variable in addition to the regressors previously noted. A single common bandwidth for these exercises is selected using a pooled RD regression quantifying the joint first stage relationship of the effect of the cutoffs on diversion.

Table A.8: Placebos Exercises using Alternate Year Discontinuities

	τ^{93}	τ^{95}	τ^{06}	τ^{08}
Sharp RD: Diversion	-0.031	-0.015	0.034	-0.028
	(0.031)	(0.034)	(0.026)	(0.026)
	[0.65]	[0.35]	[0.44]	[0.56]
Sharp RD: Total Convictions	-0.053	-0.012	-0.059	-0.017
	(0.12)	(0.15)	(0.14)	(0.10)
	[1.23]	[1.47]	[1.47]	[1.12]
Sharp RD: Qtrly Employment Rate	-0.065**	-0.024	0.0079	-0.020
	(0.031)	(0.027)	(0.019)	(0.022)
	[0.40]	[0.41]	[0.36]	[0.39]
Observations	15,430	15,942	26,240	26,606

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents sharp RD estimates of the impact of the discontinuity at four placebo threshold dates on diversion, total convictions and the average quarterly employment rate. We shift the true discontinuity dates back one year in the first and third column to 9/1/1993 and 11/7/2006, respectively; and forward one year in the second and fourth column to 9/1/1995 and 11/7/2008, respectively. We do not report fuzzy RD estimates since there is zero first-stage relationship. *General RD Table Notes* from Table 1 apply.

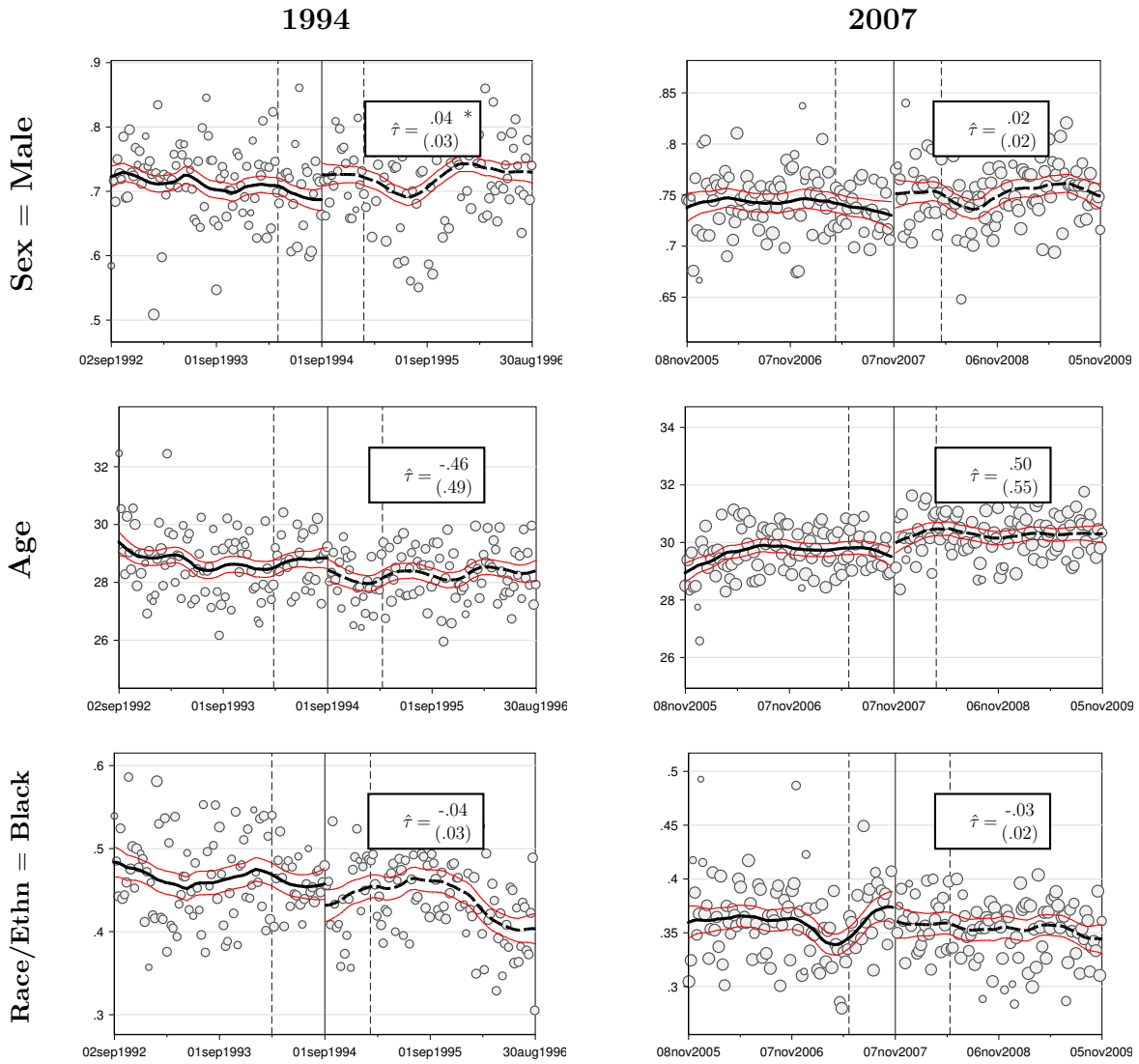
Table A.9: Comparison between First Time Felony Defendants and Repeat Felony Defendants

	1994 First Time	1994 Repeat	2007 First Time	2007 Repeat
Sharp RD: Diversion	-0.24***	-0.0061	0.18***	0.12***
	(0.028)	(0.034)	(0.025)	(0.025)
	[0.69]	[0.29]	[0.49]	[0.12]
Fuzzy RD: Total Convictions	-1.61***		-1.21*	0.31
	(0.48)		(0.63)	(2.18)
	[2.21]		[2.16]	[3.88]
Fuzzy RD: Qtrly Employment Rate	0.20***		0.16	0.10
	(0.078)		(0.10)	(0.14)
	[0.37]		[0.30]	[0.15]
Observations	31,334	15,780	52,846	26,865

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

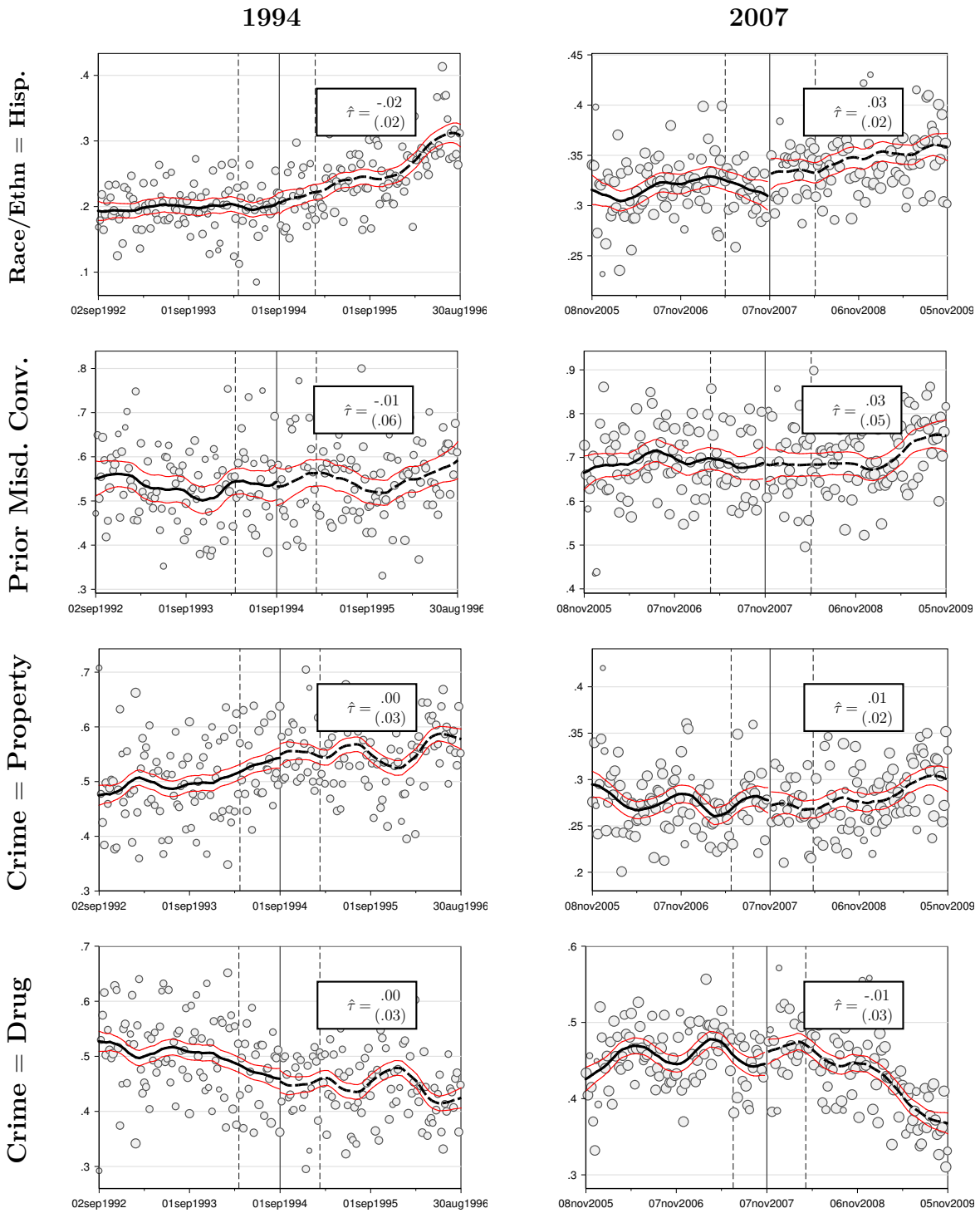
This table compares RD estimates for our estimation sample of first-time felony defendants with samples of individuals with prior felony convictions, but who meet all other sample requirements, for each experiment. Sharp RD estimates are reported for all groups in the first row. We then report fuzzy RD estimates for our primary reoffending and employment outcomes in the second and third rows. We do not report the second-stage fuzzy RD estimates for the repeat offender group in 1994 since effectively there is zero first-stage relationship. *General RD Table Notes* from Table 1 apply.

B Online Appendix (Not for Publication)

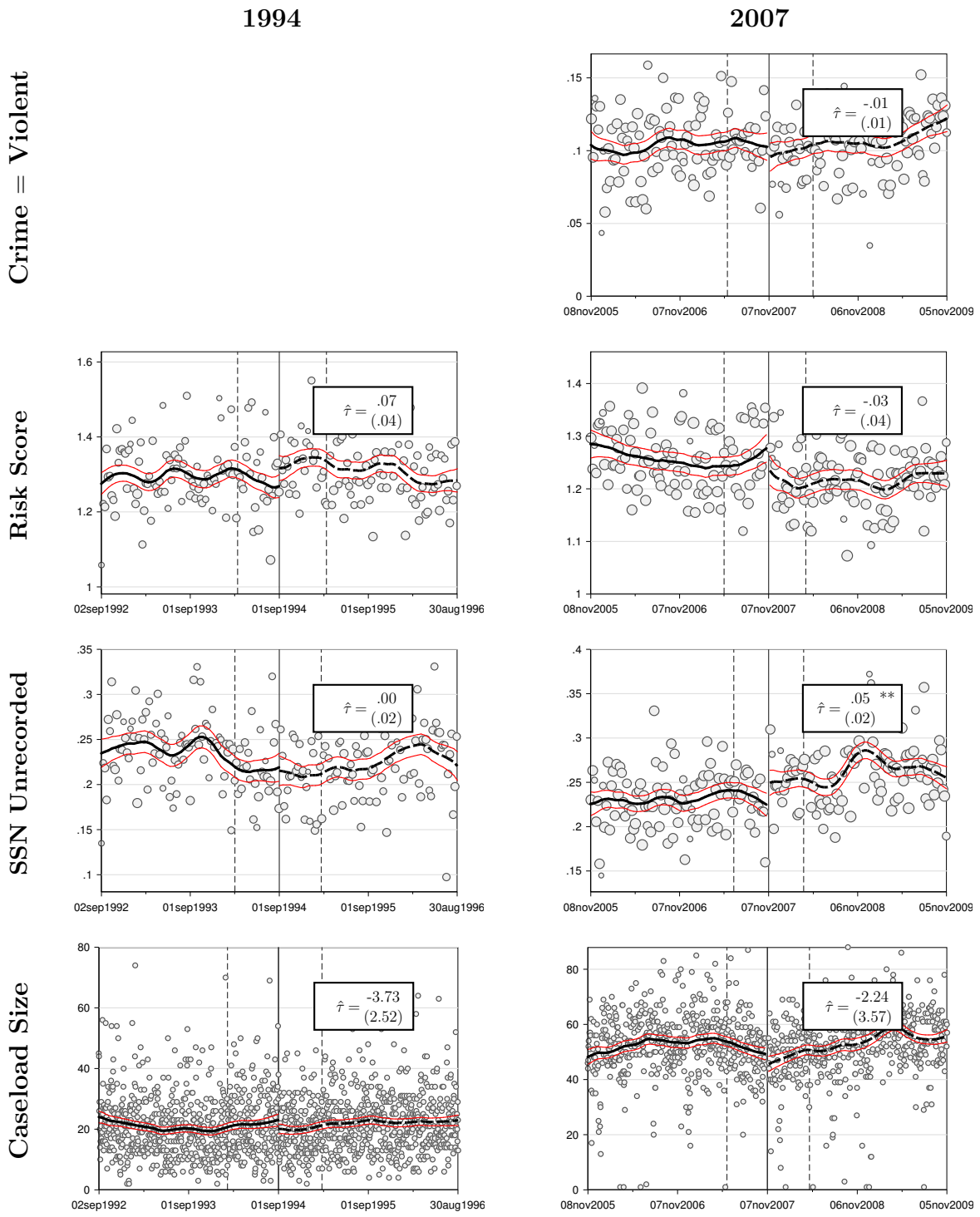


Online Figure B.1: Baseline comparison

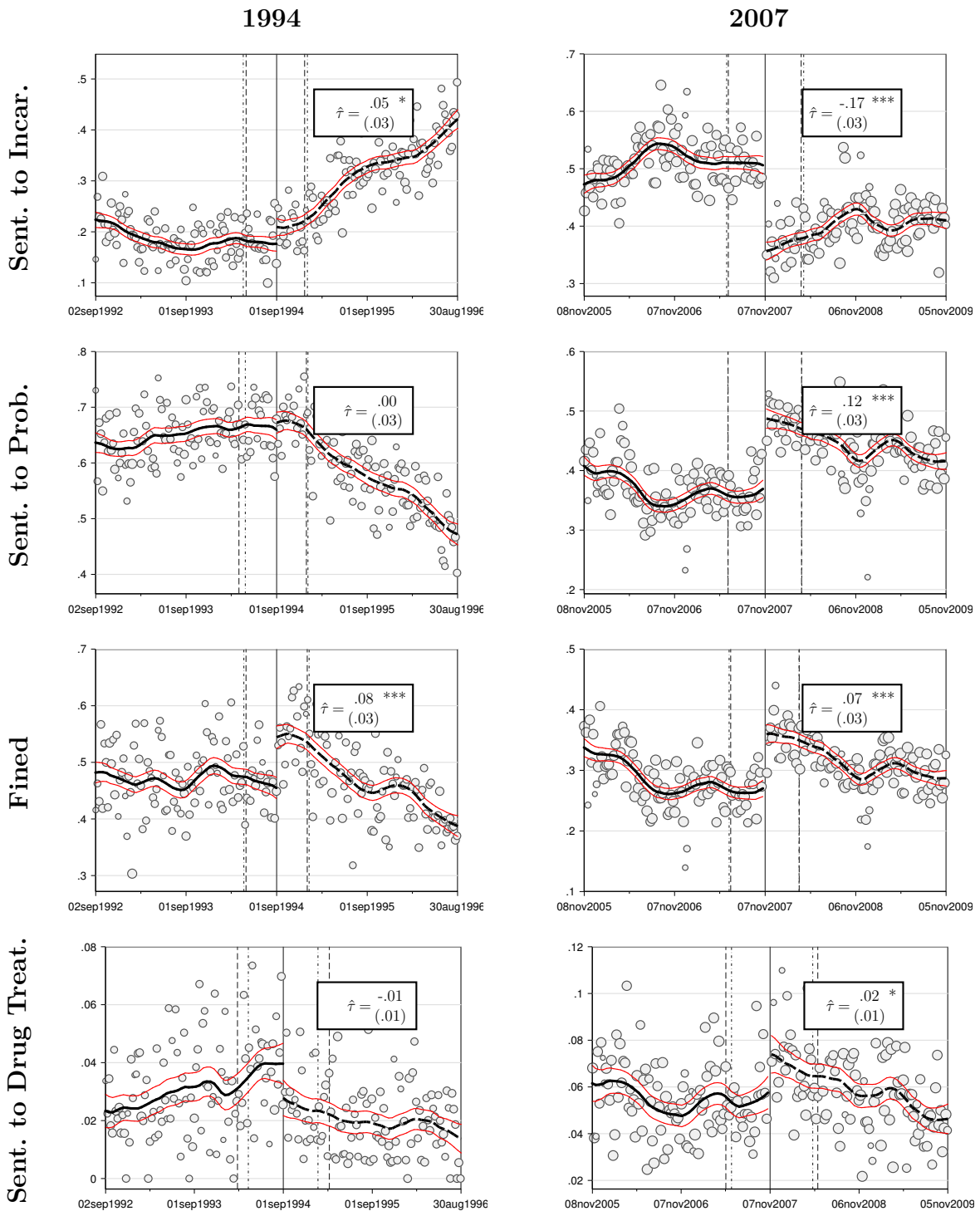
This figure presents reduced form graphical evidence of any changes in pre-determined characteristics as recorded in the criminal court records from the Harris County District Court associated with the 1994 (left column) and 2007 (right column) natural experiments. The calculation of the recidivism risk score is described in Section 5. SSN Unrecorded is an indicator variable for there being no recorded social security number in the court record. These estimates correspond to coefficients presented in Table 1. All *General RD Figure Notes* from Figure 2 apply.



Online Figure B.1: Baseline comparison (continued)

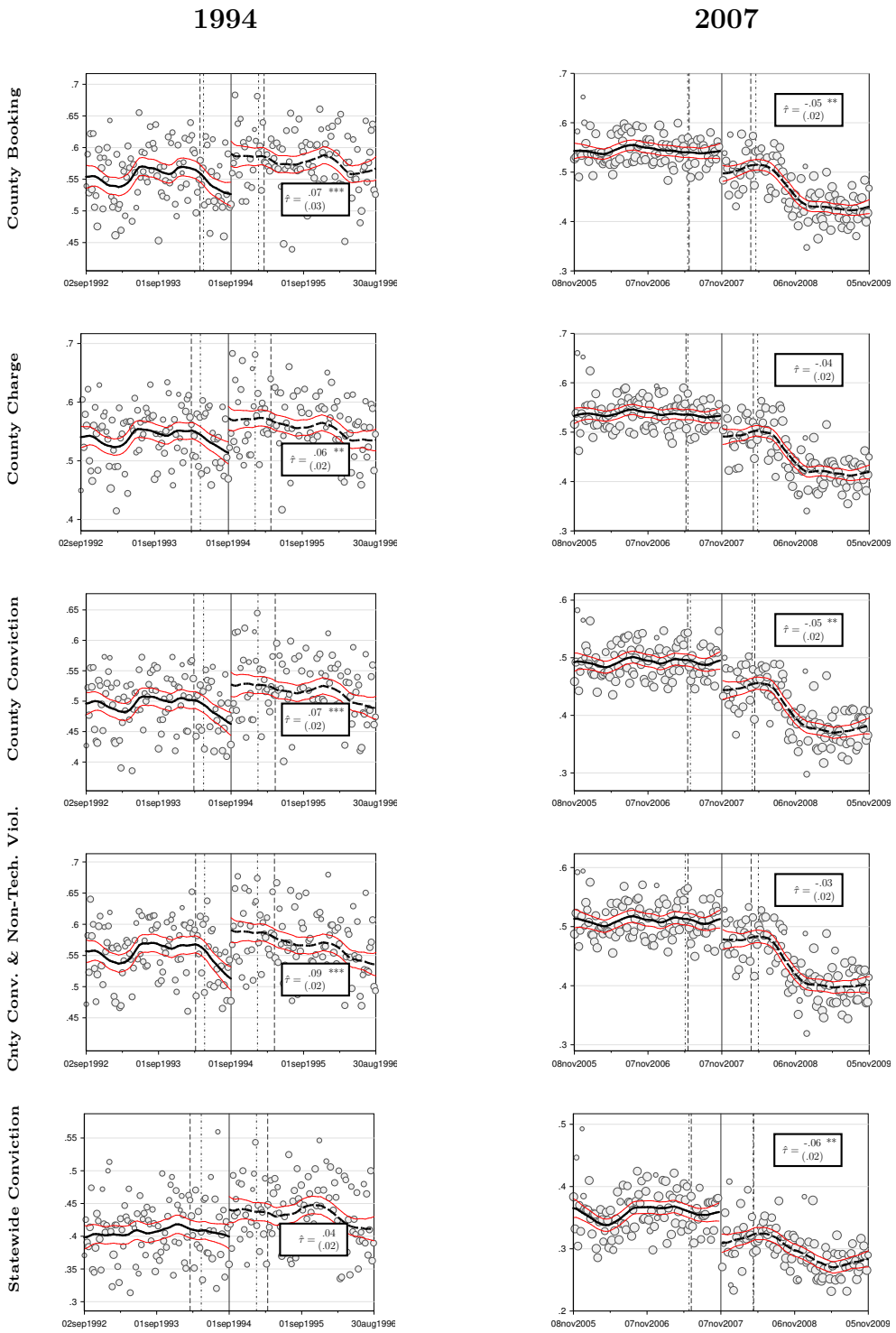


Online Figure B.1: Baseline comparison (continued)



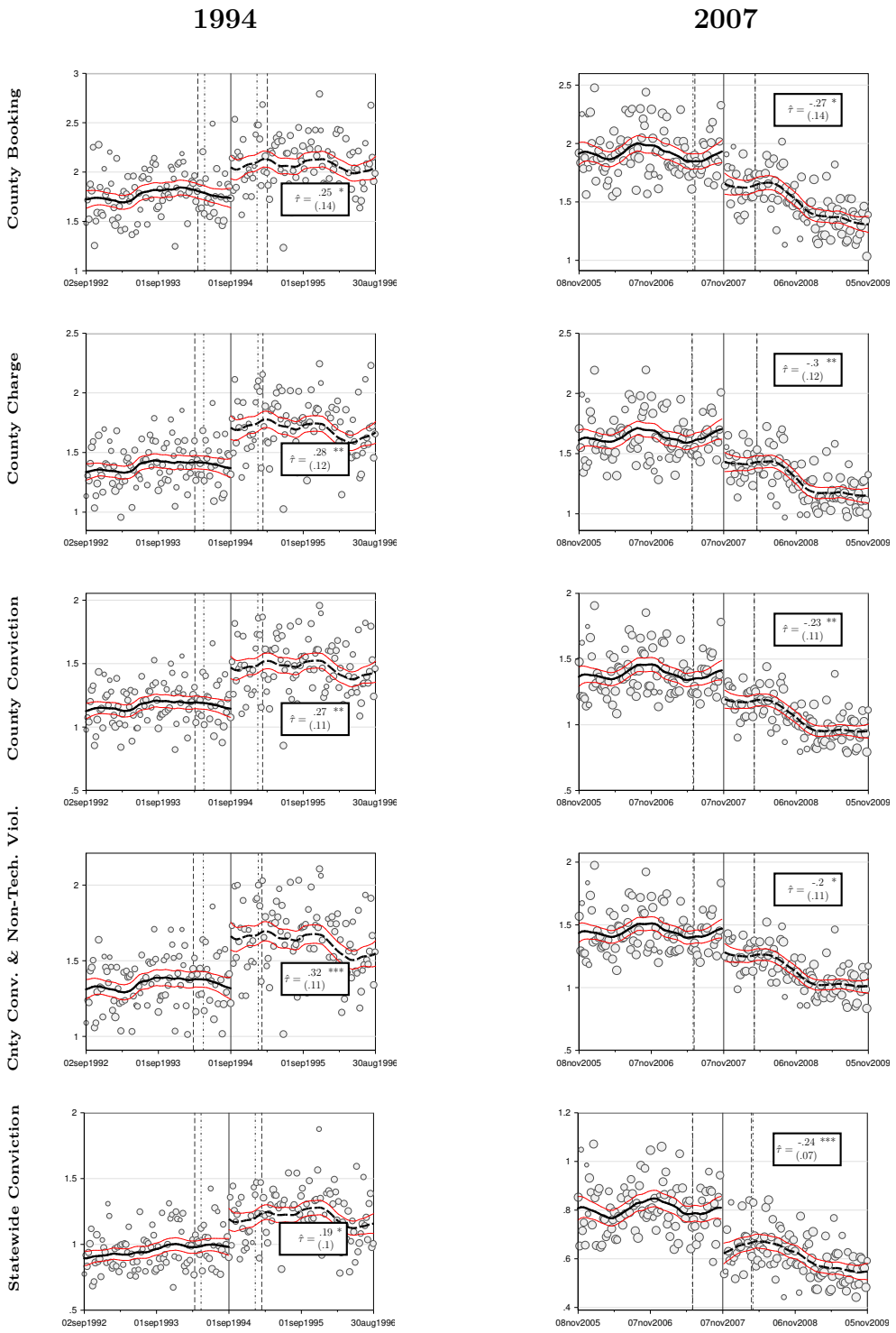
Online Figure B.2: Visual Evidence and Estimated Reduced Form Change - Sentencing Outcomes

This figure presents reduced form graphical evidence of any changes in sentencing outcomes recorded in the criminal court records from the Harris County District Court associated with the 1994 (left column) and 2007 (right column) natural experiments. These estimates represent the reduced form version of the fuzzy RD estimates presented in Table 2. All *General RD Figure Notes* from Figure 2 apply.



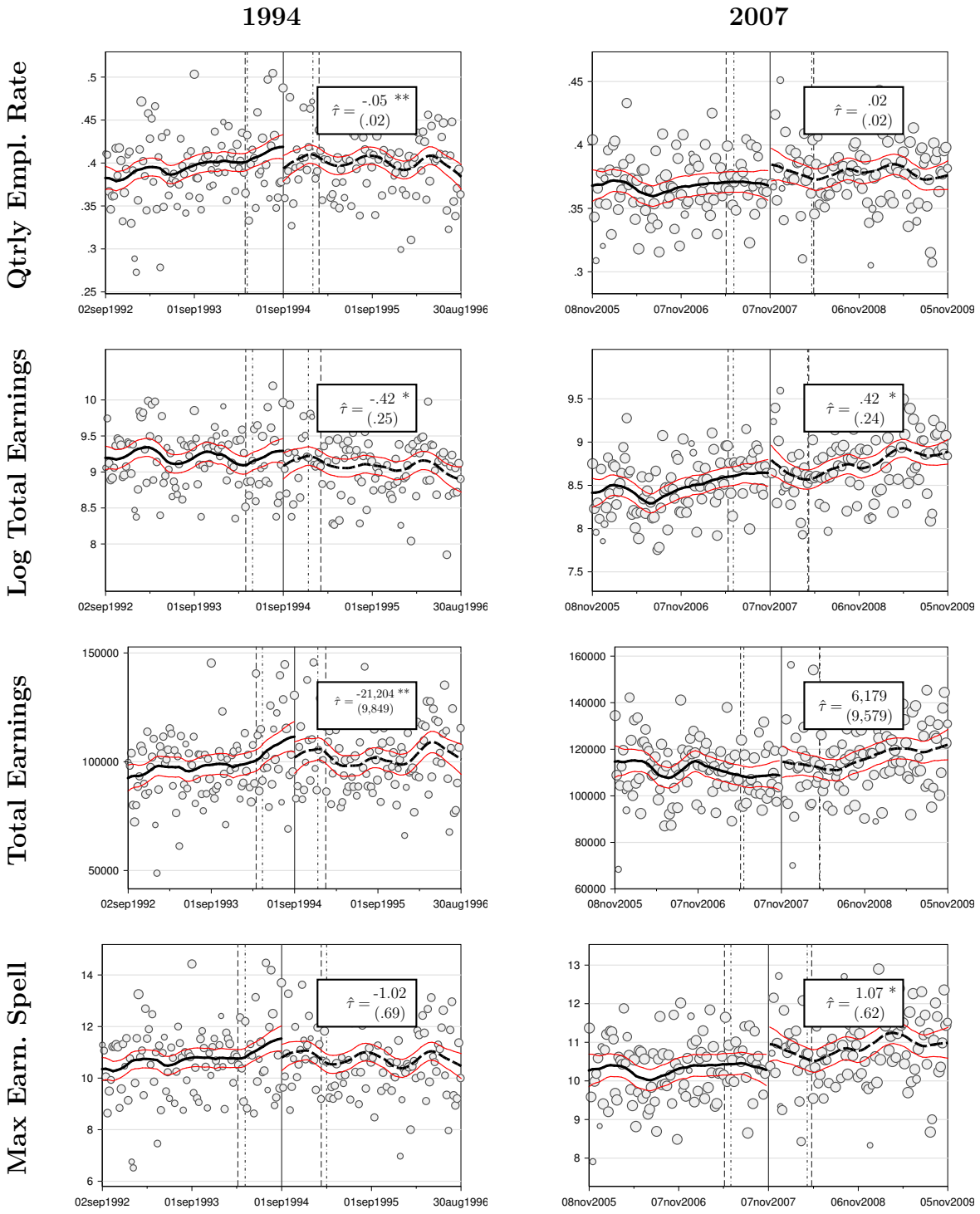
Online Figure B.3: Visual Evidence and Estimated Reduced Form Change - Recidivism (Extensive Margin)

This figure presents reduced form graphical evidence of any changes in reoffending outcomes at the extensive-margin associated with the 1994 (left column) and 2007 (right column) natural experiments. All *General RD Figure Notes* from Figure 2 apply.



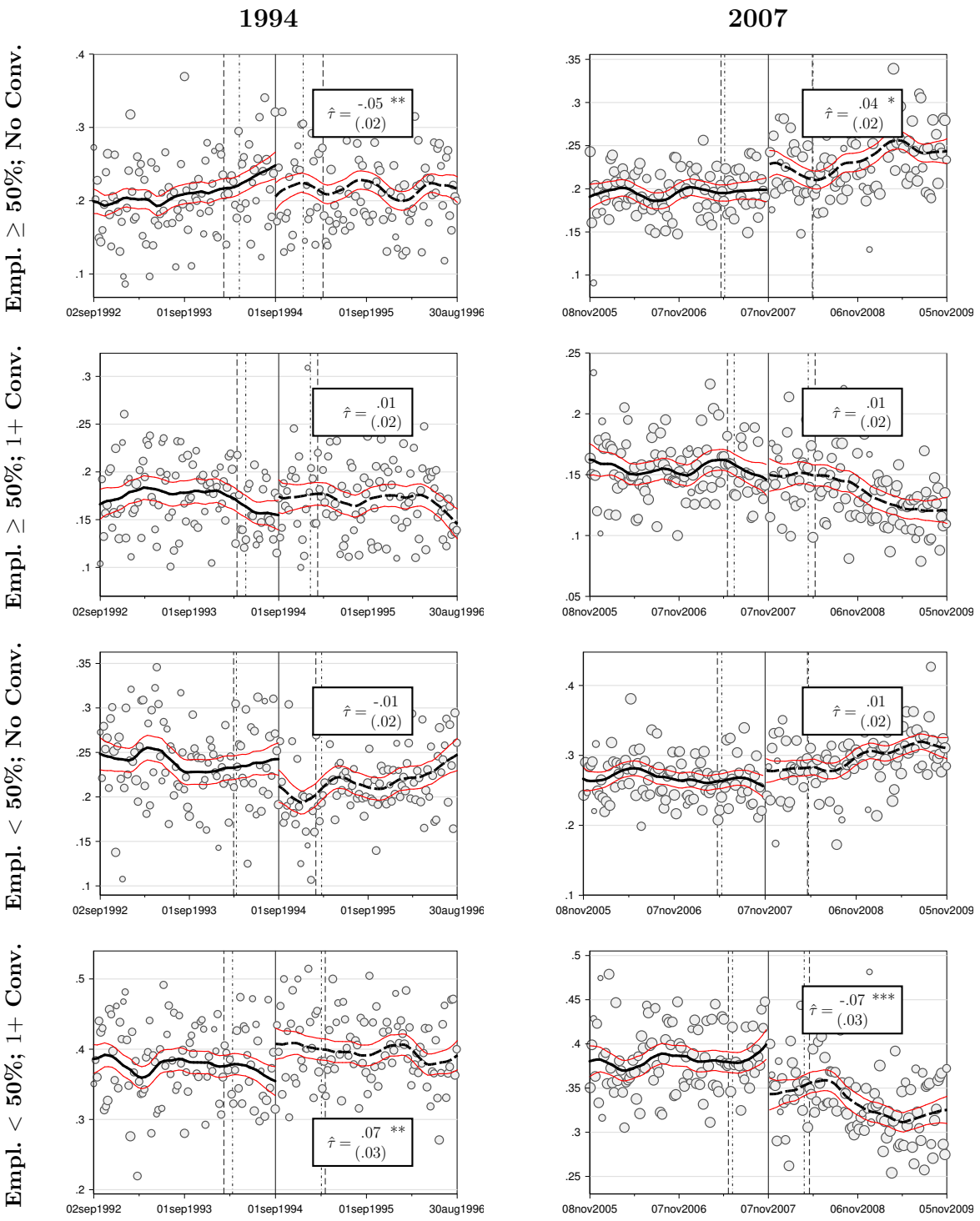
Online Figure B.4: Visual Evidence and Estimated Reduced Form Change - Recidivism (Intensive Margin)

This figure presents reduced form graphical evidence of any changes in reoffending outcomes at the intensive-margin associated with the 1994 (left column) and 2007 (right column) natural experiments. These results are reduced-form versions of the fuzzy RD reoffending results presented in Table A.2. All *General RD Figure Notes* from Figure 2 apply.



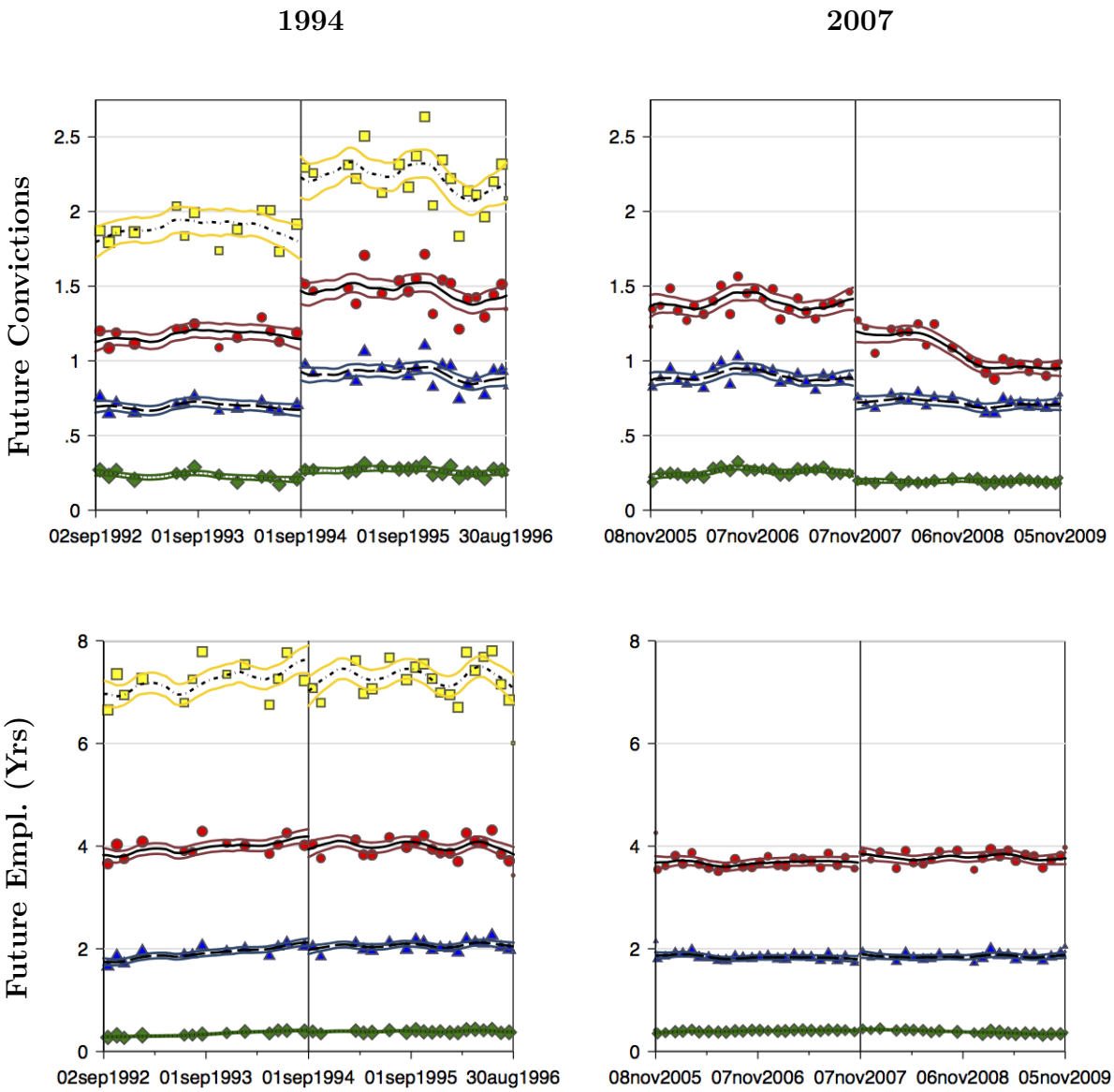
Online Figure B.5: Visual Evidence and Estimated Reduced Form Change - Labor Market

This figure presents reduced form graphical evidence of any changes in labor market outcomes with the 1994 (left column) and 2007 (right column) natural experiments. These results are reduced-form versions of the fuzzy RD reoffending results presented in Table 4. All *General RD Figure Notes* from Figure 2 apply.



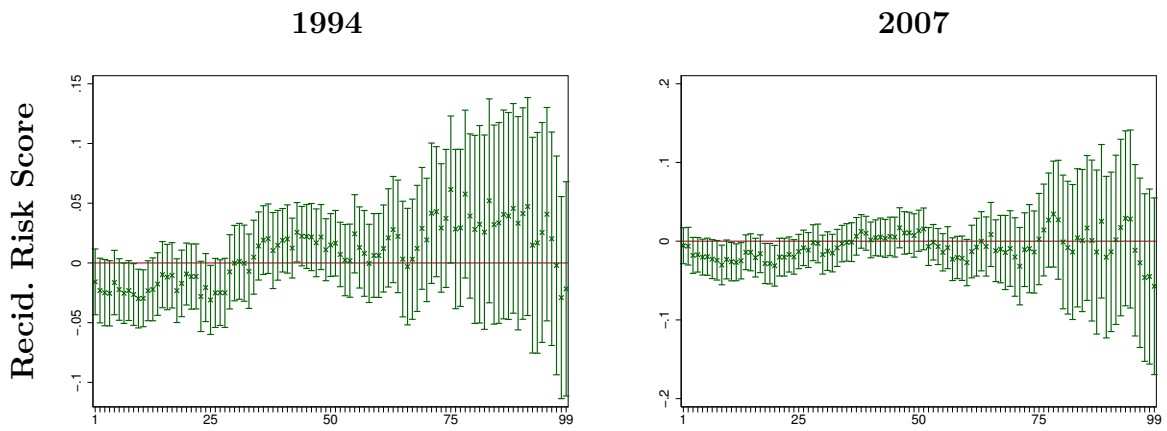
Online Figure B.6: Visual Evidence and Estimated Reduced Form Change - Combined Crime/Labor Outcomes

This figure presents reduced form graphical evidence of any changes in the offending/labor market intersection outcomes (defined in Table 5) associated with the 1994 (left column) and 2007 (right column) natural experiments. These results are reduced-form versions of the fuzzy RD results presented in Table 5. All *General RD Figure Notes* from Figure 2 apply.



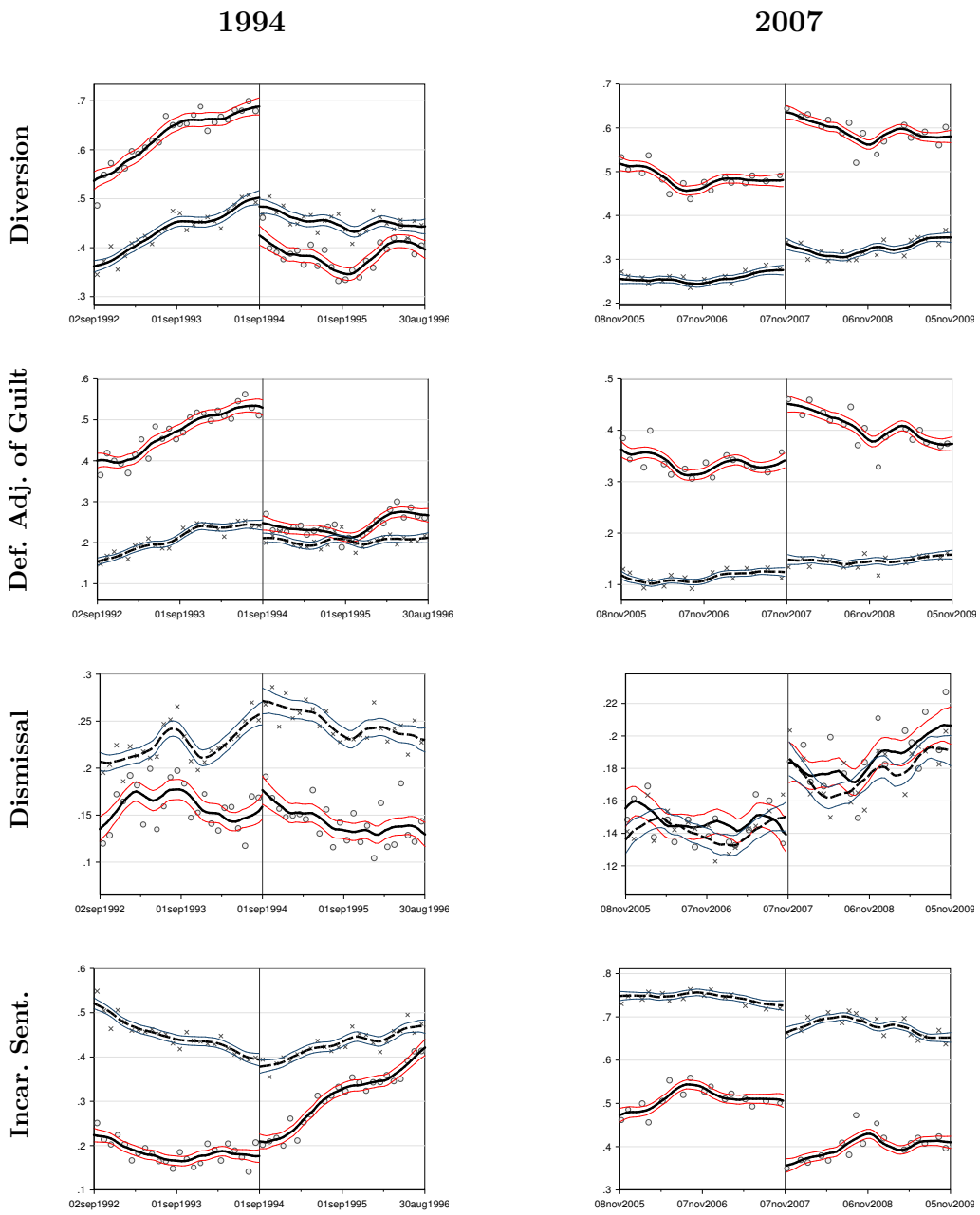
Online Figure B.7: Visual Evidence on the 1, 5, 10 and 20-year Cumulative Impacts

This figure presents reduced form graphical evidence on the cumulative effect of diversion on future convictions and employment associated with the 1994 (left column) and 2007 (right column) natural experiments. These estimates represent the reduced form version of the fuzzy RD estimates presented in Figure 3. To distinguish the different follow-up periods, we have used the following schema: 1 year (dotted regression line, green CI, diamond scatter), 5 year (dashed regression line, blue CI, triangle scatter), 10 year (solid regression line, red CI, circle scatter), and 20 year (dotted-dashed regression line, gold CI, square scatter). For readability, the scatter plots show monthly rather than weekly bins. Otherwise, all *General RD Figure Notes* from Figure 2 apply.



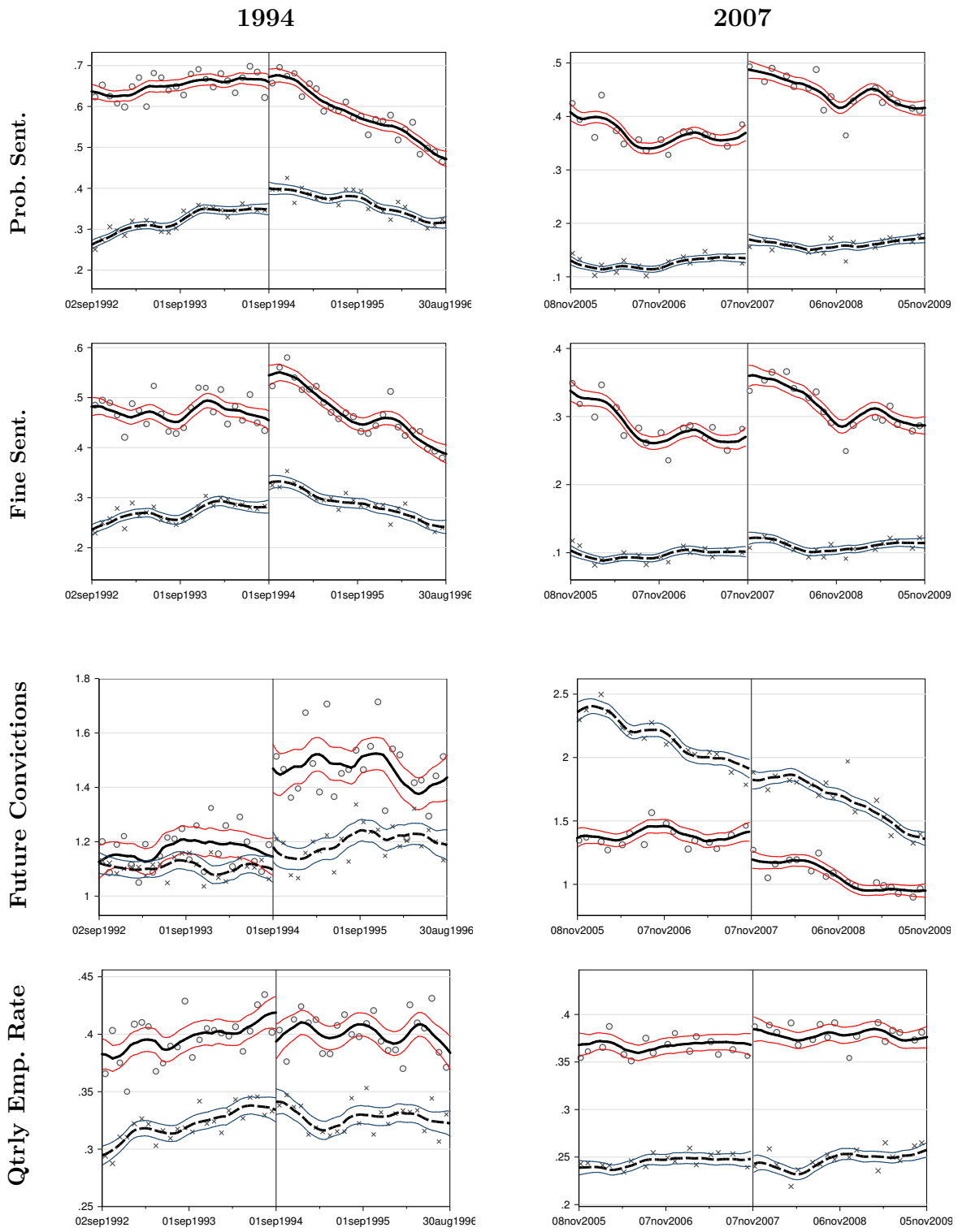
Online Figure B.8: Balance of Predicted Recidivism Risk over the Recidivism Risk Profile

This figure displays RD estimates and associated 90% confidence bands for evaluating balance in predicted recidivism risk score described in Section 5.4 calculated over the quantile function of the predicted recidivism risk score for the 1994 experiment (left column) and 2007 experiment (right column). Each coefficient reflects a distinct local polynomial RD estimate from regression centered at the focal percentile using a uniform kernel with a 40 percentile bandwidth. Estimates below the 20th and above the 80th percentiles will reflect narrower, asymmetric bandwidths. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. We do not adjust the regression estimates for observable covariates because this is a balance test.

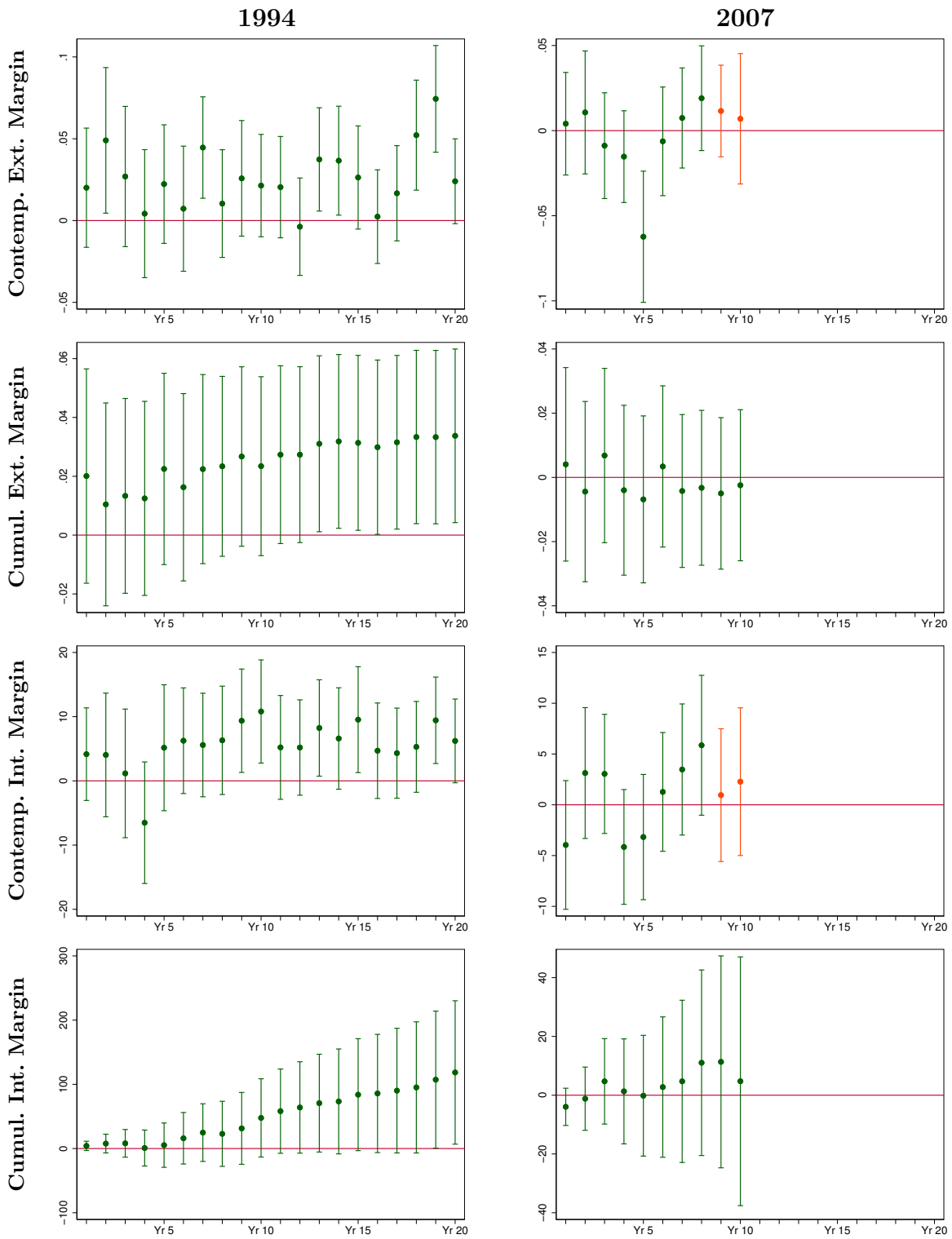


Online Figure B.9: Evaluating Changes in Disposition, Sentencing and Behavioral Outcomes for Study Sample and other Felony Defendants in Harris County

This figure presents reduced form graphical evidence of the changes in diversion, reoffending, and employment associated with the 1994 (left column) and 2007 (right column) natural experiments for our estimation samples (red/solid) and all other felony defendants in Harris County (blue/dashed) not included in our study sample due to the sample restrictions discussed in Section 3.2. In order to avoid defendants repeatedly appearing in the analysis when examining behavior in the broader caseload, which is not constrained to first time defendants, we collapse observations associated with the same individual to their earliest date when multiple charges or dispositions occur within a 1 year period. *General RD Figure Notes* from Figure 2 with the exception that the scatter points represent month-level (rather than week-level) bin averages.

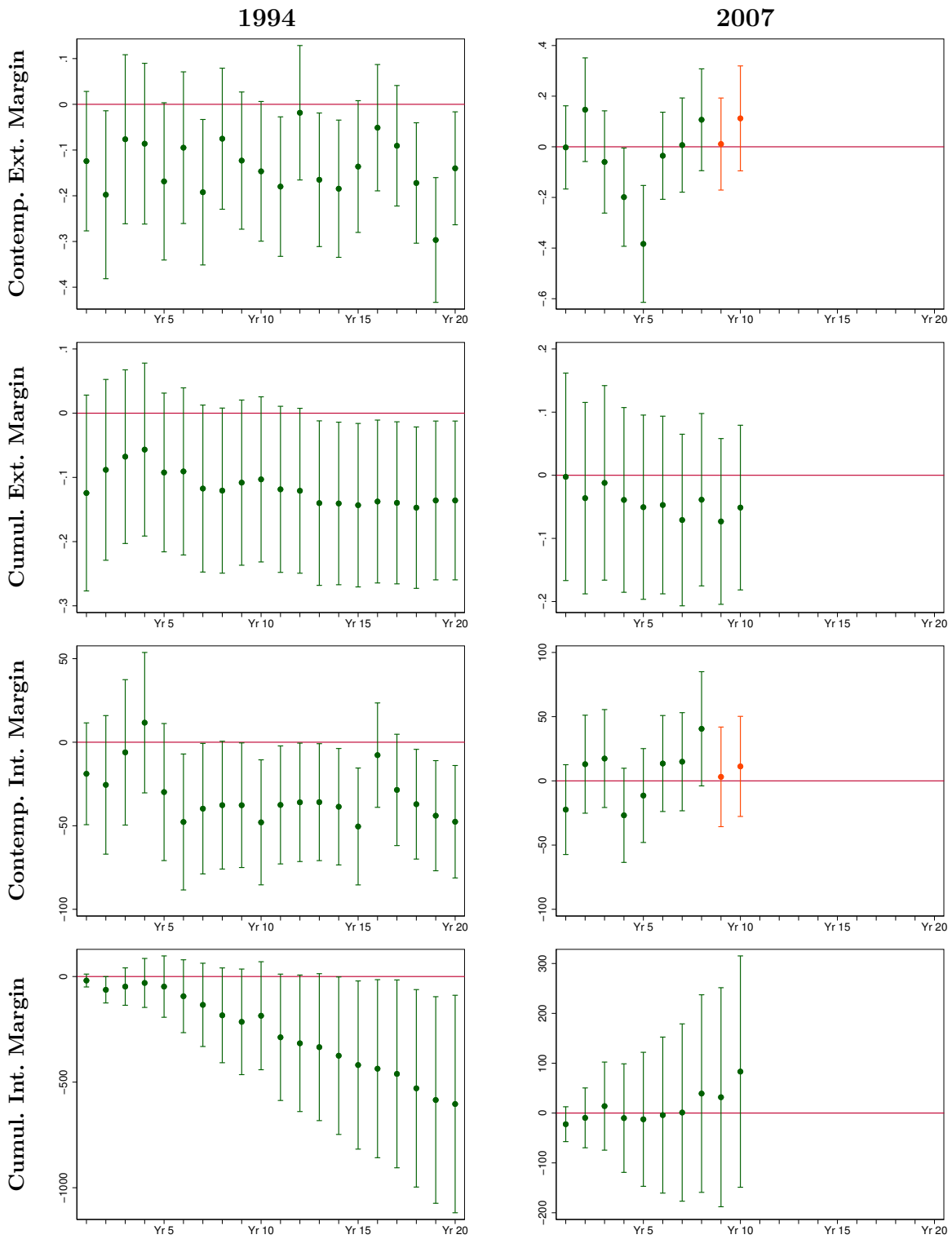


(Continuation of Figure B.9.)



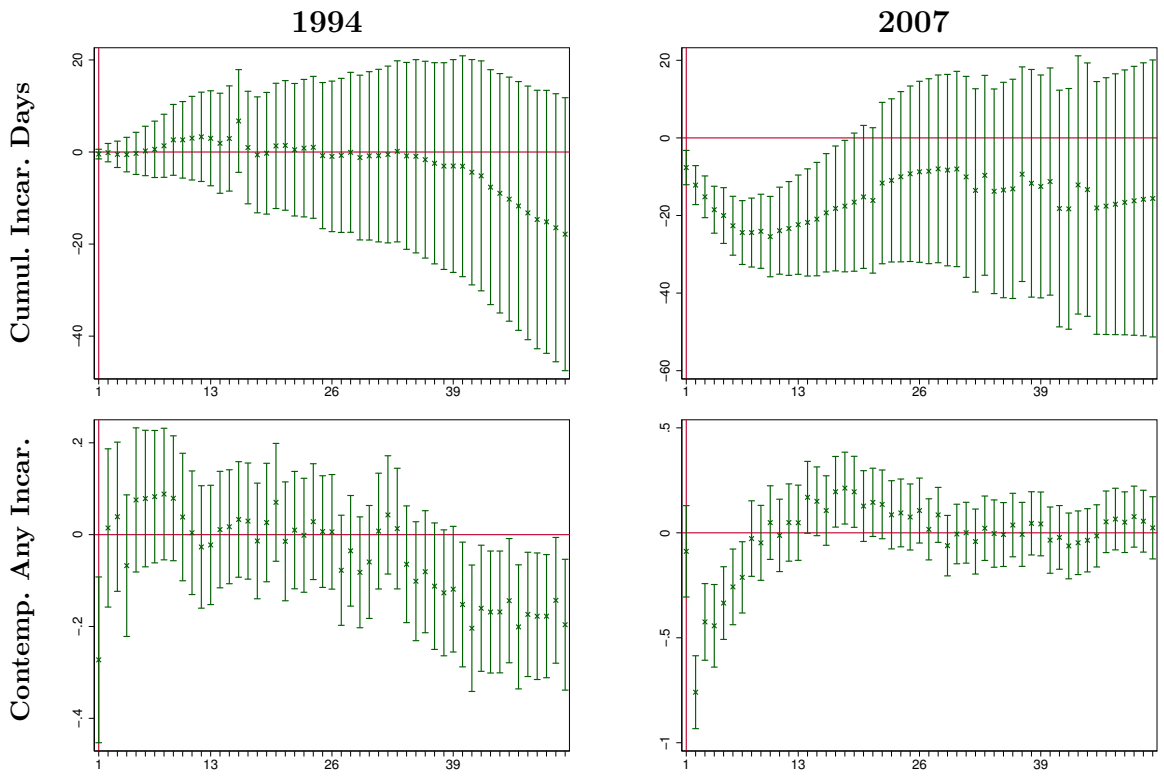
Online Figure B.10: Alternative Measures of Exposure to Incarceration over Time: Reduced Form

This figure displays reduced form estimates and associated 90% confidence bands for incarceration outcomes that measure year-by-year impacts of diversion for each year period up through 20 years for the 1994 sample (left column) and 10 years for the 2007 sample (right column). The first two rows depict the estimated impact of diversion on any incarceration both contemporaneously (first row) and cumulatively (second row). The third and forth rows repeat report the estimated impact of diversion on the total number of incarceration days both contemporaneously (first row) and cumulatively (second row) (repeating results previously provided in Figure 3 and Appendix Figure A.2. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. .



Online Figure B.11: Alternative Measures of Exposure to Incarceration over Time: Fuzzy RD Estimates

This figure displays fuzzy RD estimates and associated 90% confidence bands for incarceration outcomes that measure year-by-year impacts of diversion for each year period up through 20 years for the 1994 sample (left column) and 10 years for the 2007 sample (right column). The first two rows depict the estimated impact of diversion on any incarceration both contemporaneously (first row) and cumulatively (second row). The third and forth rows repeat report the estimated impact of diversion on the total number of incarceration days both contemporaneously (first row) and cumulatively (second row) (repeating results previously provided in Figure 3 and Appendix Figure A.2. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.



Online Figure B.12: Measures of Exposure to Incarceration by Week for First Year

This figure displays fuzzy RD estimates and associated 90% confidence bands for incarceration outcomes that measure week-by-week impacts of diversion for the first year following the case disposition for the 1994 sample (left column) and the 2007 sample (right column). The first row depicts the cumulative days incarcerated each week since the disposition date. The second row presents fuzzy RD estimates on the effect on any incarceration for the week indicated on the horizontal access. RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. .

Online Table B.1: Classic Heterogeneity Analysis

Panel A: 1994 Sample	Black	White	Hispanic	Crime = Property	Crime = Drug	
Sharp RD: Diversion	-0.31***	-0.25***	-0.16***	-0.15***	-0.37***	
	(0.044)	(0.045)	(0.055)	(0.047)	(0.043)	
Fuzzy RD: Total Convictions	-1.27**	-0.93	-1.07	-1.99**	-0.63	
	(0.63)	(0.64)	(1.19)	(0.82)	(0.48)	
Fuzzy RD: Qtrly Employment Rate	0.23**	0.27*	-0.20	0.36**	-0.0035	
	(0.11)	(0.15)	(0.26)	(0.16)	(0.068)	
Observations	12,456	7,035	4,313	12,433	11,696	
	Male	Female	< 30 Yrs Old	≥ 30 Yrs Old	No Misd. Conv.	1+ Misd. Conv.
Sharp RD: Diversion	-0.26***	-0.25***	-0.23***	-0.29***	-0.33***	-0.23***
	(0.034)	(0.052)	(0.034)	(0.053)	(0.042)	(0.038)
Fuzzy RD: Total Convictions	-1.04*	-1.00	-1.72**	-0.70*	-0.49	-1.81***
	(0.61)	(0.76)	(0.75)	(0.41)	(0.59)	(0.57)
Fuzzy RD: Qtrly Employment Rate	0.038	0.31	0.14	0.033	-0.014	0.25**
	(0.073)	(0.20)	(0.088)	(0.12)	(0.11)	(0.12)
Observations	22,419	8,915	18,477	12,857	8,988	22,346

Panel B: 2007 Sample	Black	White	Hispanic	Crime = Property	Crime = Drug	Crime = Violent
Sharp RD: Diversion	0.19***	0.17***	0.13***	0.088**	0.25***	0.18***
	(0.039)	(0.046)	(0.047)	(0.042)	(0.044)	(0.060)
Fuzzy RD: Total Convictions	-2.06**	-1.00	-1.57*	-2.74	-0.37	-1.16
	(0.88)	(1.02)	(0.91)	(1.80)	(0.63)	(1.40)
Fuzzy RD: Qtrly Employment Rate	0.21	0.10	0.19	0.024	0.14*	0.16
	(0.13)	(0.17)	(0.18)	(0.34)	(0.088)	(0.49)
Observations	16,804	11,757	10,683	11,287	17,267	4,261
	Male	Female	< 30 Yrs Old	≥ 30 Yrs Old	No Misd. Conv.	1+ Misd. Conv.
Sharp RD: Diversion	0.17***	0.16***	0.16***	0.20***	0.19***	0.15***
	(0.033)	(0.053)	(0.035)	(0.038)	(0.040)	(0.038)
Fuzzy RD: Total Convictions	-1.73**	-1.00	-2.38***	-0.39	-1.37	-1.27
	(0.70)	(1.21)	(0.78)	(0.77)	(0.98)	(0.80)
Fuzzy RD: Qtrly Employment Rate	0.15	0.048	0.093	0.11	0.082	0.15
	(0.11)	(0.17)	(0.12)	(0.13)	(0.11)	(0.16)
Observations	39,464	13,382	30,513	22,333	18,012	34,834

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents estimates of the impact of diversion for subgroups defined based on demographic characteristics or prior criminal histories as indicated in the title to each column. Panel A presents estimates across our key first stage (diversion) and second stage (total convictions and quarterly employment rate) for 11 subsamples. Panel B repeats this exercise for the 2007 experiment. *General RD Table Notes* from Table 1 apply.

Online Table B.2: Data-driven Selected Bandwidths for Main Results

	1994		2007	
	Left	Right	Left	Right
Table 1				
Caseload Size	214	182	165	171
Total Prior Misd. Conv.	165	160	220	185
Age at Charge	189	191	158	148
Sex = Male	152	145	206	168
Race/Ethn. = Black, not Hisp.	184	158	163	193
Race/Ethn. = Hispanic	165	145	182	187
Crime Type = Property	174	158	160	177
Crime Type = Drug	168	161	136	157
Crime Type = Violent			171	180
Recidivism Risk Score	171	193	183	152
Social Sec. Number Unrecorded	182	172	143	143
Table 2				
Court Deferral	128	121	144	133
Case Dismissal	127	117	146	145
Incarceration Sentence	135	124	155	155
Probation Sentence	126	126	150	148
Fined	132	131	145	137
Drug Treatment	144	142	158	175
Table 3				
Any Convictions	138	136	155	150
Total Convictions	136	136	154	153
Drug Convictions	133	128	158	159
Property Convictions	130	133	148	148
Violent Convictions	135	134	155	155
Misdemeanor Convictions	138	135	147	127
Felony Convictions	135	132	160	155
Table 4				
Qtrly Employment Rate	148	121	149	171
Earn $\geq$ 100% Fed. Pov. Level	147	149	161	172
Log Avg. Qrtly Earnings	126	103	151	154
Avg. Qrtly Earnings	141	102	165	169
Max Employment Spell (Quarters).	158	176	146	161
Max Cont. Earning Spell (Quarters)	148	183	153	158
Table 5				
Qtrly Employment Rate $\geq$ 50%; No Convictions	145	111	179	182
Qtrly Employment Rate $\geq$ 50%; 1+ Convictions	135	130	140	163
Qtrly Employment Rate $<$ 50%; No Convictions	173	175	174	174
Qtrly Employment Rate $<$ 50%; 1+ Convictions	172	184	146	146

This table presents the bandwidths chosen for each specification in our main tables of results (Tables 1 through 5). Based on Calonico et al. (2016a), we use a data-driven bandwidth selector that selects the median bandwidth from the following three mean squared error-optimal methods for the RD treatment effect estimator: one common MSE-optimal bandwidth selector for the RD treatment effect estimator; two different bandwidth selectors (below and above); and one common MSE-optimal bandwidth selector for the sum of regression estimates.

Online Table B.3: Robustness - Bandwidth Selection

	MSE1	MSE2	MSE3	CER1	CER2	CER3
Panel A: 1994 Sample						
Sharp RD: Diversion	-0.26***	-0.24***	-0.23***	-0.24***	-0.22***	-0.21***
	(0.027)	(0.028)	(0.029)	(0.031)	(0.033)	(0.034)
Fuzzy RD: Total Convictions	-1.70***	-1.49***	-1.61***	-1.61***	-1.79***	-1.63***
	(0.49)	(0.50)	(0.48)	(0.59)	(0.59)	(0.58)
Fuzzy RD: Qtrly Employment Rate	0.18**	0.20***	0.054	0.14	0.18*	0.14**
	(0.083)	(0.078)	(0.055)	(0.11)	(0.10)	(0.069)
Panel B: 2007 Sample						
Sharp RD: Diversion	0.17***	0.18***	0.17***	0.16***	0.16***	0.18***
	(0.025)	(0.026)	(0.019)	(0.031)	(0.031)	(0.024)
Fuzzy RD: Total Convictions	-1.16*	-1.21*	-1.37**	-0.81	-0.72	-0.79
	(0.63)	(0.65)	(0.63)	(0.91)	(0.93)	(0.95)
Fuzzy RD: Qtrly Employment Rate	0.16	0.16	0.20*	0.13	0.11	0.12
	(0.10)	(0.10)	(0.11)	(0.15)	(0.15)	(0.16)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents sharp RD (diversion) and fuzzy RD (total convictions and employment rate) estimates varying the optimal bandwidth selection method. The bandwidth selectors are coded as follows: MSE1 (one common mean squared error-optimal bandwidth selector for the RD treatment effect estimator); MSE2 (two different mean squared error-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator); MSE3 (one common mean squared error-optimal bandwidth selector for the sum of regression estimates); CER1 (one common coverage error rate-optimal bandwidth selector for the RD treatment effect estimator); CER2 (two different coverage error rate-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator); and, CER3 (one common coverage error rate-optimal bandwidth selector for the sum of regression estimates). All other RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2. Figure A.1 shows the robustness of the results to a range of non-data-driven, pre-specified bandwidths ranging from 90 to 720 days.

Online Table B.4: Robustness - Variance Estimators

	Day Cluster	Week Cluster	Nearest Neighbor
Panel A: 1994 Sample			
Sharp RD: Diversion	-0.24***	-0.24***	-0.24***
	(0.033)	(0.032)	(0.028)
Fuzzy RD: Total Convictions	-1.34**	-1.52***	-1.61***
	(0.53)	(0.50)	(0.48)
Fuzzy RD: Qtrly Employment Rate	0.16*	0.16	0.22***
	(0.088)	(0.11)	(0.080)
Panel B: 2007 Sample			
Sharp RD: Diversion	0.17***	0.17***	0.18***
	(0.032)	(0.027)	(0.025)
Fuzzy RD: Total Convictions	-1.43**	-1.04*	-1.21*
	(0.56)	(0.58)	(0.63)
Fuzzy RD: Qtrly Employment Rate	0.18*	0.21**	0.16
	(0.097)	(0.084)	(0.10)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents sharp RD (diversion) and fuzzy RD (total convictions and employment rate) estimates using alternative variance estimators as described with column titles. All other RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

Online Table B.5: Robustness - Alternative Specificaitons

	No Donut	No Covariates	Non-Robust	Non-Bias Corrected	Week Binned Running Var.	Epa. Kernel	Tri. Kernel
Panel A: 1994 Sample							
Sharp RD: Diversion	-0.24***	-0.24***	-0.24***	-0.25***	-0.19***	-0.26***	-0.25***
	(0.028)	(0.029)	(0.025)	(0.025)	(0.028)	(0.025)	(0.025)
Fuzzy RD: Total Convictions	-1.61***	-1.44***	-1.61***	-1.35***	-1.88***	-1.21***	-1.35***
	(0.48)	(0.53)	(0.42)	(0.42)	(0.60)	(0.39)	(0.42)
Fuzzy RD: Qtrly Employment Rate	0.20***	0.17**	0.20***	0.18***	0.062	0.18**	0.18**
	(0.078)	(0.077)	(0.069)	(0.069)	(0.082)	(0.076)	(0.078)
Panel B: 2007 Sample							
Sharp RD: Diversion	0.13***	0.17***	0.18***	0.18***	0.17***	0.17***	0.17***
	(0.023)	(0.025)	(0.021)	(0.021)	(0.024)	(0.025)	(0.024)
Fuzzy RD: Total Convictions	-1.56**	-1.69**	-1.21**	-1.14**	-1.44**	-1.30**	-1.31**
	(0.70)	(0.70)	(0.53)	(0.53)	(0.63)	(0.61)	(0.65)
Fuzzy RD: Qtrly Employment Rate	0.10	0.14	0.16*	0.14	0.16	0.15	0.16
	(0.11)	(0.11)	(0.089)	(0.089)	(0.10)	(0.10)	(0.10)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table presents sharp RD (diversion) and fuzzy RD (total convictions and employment rate) estimates that relax other baseline specification choices that are described with column titles. All other RD specification choices are described in Section 4 and the estimation sample for each experiment is described in Section 3.2.

Online Table B.6: Robustness - Comparing Strategies to Aggregate Estimates

	$\frac{\tau^{07}+\tau^{94}}{2}$, $\frac{\phi^{07}+\phi^{94}}{2}$	τ^{Pooled} , ϕ^{Pooled}
Reduced Form: Diversion	0.21***	0.21***
	(0.019)	(0.019)
Fuzzy RD:		
Court Deferral	0.83***	0.86***
	(0.071)	(0.064)
Case Dismissal	0.16**	0.15**
	(0.071)	(0.066)
Incarceration Sentence	-0.58***	-0.49***
	(0.051)	(0.065)
Probation Sentence	0.33***	0.22***
	(0.079)	(0.082)
Fine Sentence	0.081	-0.030
	(0.089)	(0.090)
Total Convictions	-1.45***	-1.25***
	(0.40)	(0.39)
Qtrly Employment Rate	0.18***	0.14**
	(0.065)	(0.058)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

This table compares estimated effects averaged across our two study samples as presented in our regression output tables with estimates from a pooled estimation approach where we create a unified running variable and combine the two samples. All other RD specification choices are described in Section 4 Section 4 and the estimation sample for each experiment is described in Section 3.2.