ABSTRACT

Title of dissertation:	ESSAYS ON GOVERNMENT RESPONSES TO CRIME AND RACIAL INEQUALITY
	Cody Nathanael Tuttle Doctor of Philosophy, 2020
Dissertation directed by:	Professor Melissa Kearney Department of Economics

In this dissertation, I study government responses to crime and racial inequality to provide evidence on roots of economic and social disparities and the effectiveness of policies aimed to address those disparities.

In the second chapter, I estimate the effect of access to Food Stamps on criminal recidivism. In 1996, a federal welfare reform imposed a lifetime ban from Food Stamps on convicted drug felons. Florida modified this ban, restricting it to drug traffickers who commit their offense on or after August 23, 1996. I exploit this sharp cutoff in a regression discontinuity design and find that the ban increases recidivism among drug traffickers. The increase is driven by financially motivated crimes, suggesting that the cut in benefits causes ex-convicts to return to crime to make up for the lost transfer income.

In the third chapter, I test for racial disparities in the criminal justice system by analyzing abnormal bunching in the distribution of crack-cocaine amounts used in federal sentencing. I compare cases sentenced before and after the Fair Sentencing Act, a 2010 law that changed the 10-year mandatory minimum threshold for crack-cocaine from 50g

to 280g. First, I find that after 2010, there is a sharp increase in the fraction of cases sentenced at 280g (the point that now triggers a 10-year mandatory minimum), and that this increase is disproportionately large for black and Hispanic offenders. I then explore several possible explanations for the observed racial disparities, including discrimination. I analyze data from multiple stages in the criminal justice system and find that the increased bunching for minority offenders is driven by prosecutorial discretion, specifically as used by about 20-30% of prosecutors. Moreover, the fraction of cases at 280g falls in 2013 when evidentiary standards become stricter. Finally, the racial disparity in the increase cannot be explained by differences in education, sex, age, criminal history, seized drug amount, or other elements of the crime, but it can be almost entirely explained by a measure of state-level racial animus. These results shed light on the role of prosecutorial discretion and potentially racial discrimination as causes of racial disparities in sentencing.

In the final chapter, I estimate the effect of school desegregation on long-run economic outcomes by studying a natural experiment in Jefferson County, KY. In 1975, the district, under a court order, developed a unique busing assignment plan to merge the majority-white County district and the majority-black City district. Under this plan, students were assigned to be bused to new schools (versus stay at their home school and have new students bused in) based on their race and the first letter of their last name. Using this plausibly conditional random assignment and confidential data from the US Census Bureau, I find black students assigned busing to former County schools live in better neighborhoods (e.g. neighborhoods with higher tract-level income) at adulthood than black students assigned to remain in former City schools. This effect is strongest for

students bused in earlier grades and is increasing in the total number of years a student is assigned busing. Busing assignment has small to zero effect on white students. I explore the implications of white disenrollment from the district (i.e. "white flight") by using a novel dataset of archival yearbook records. I find the effect for white students remains small even after accounting for disenrollment. These results suggest that school desegregation in this setting had positive long-run effects for black students by giving them access to better schools (e.g. schools with more capital investment, more credentialed teachers, lower drop-out rates, etc.).

ESSAYS ON GOVERNMENT RESPONSES TO CRIME AND RACIAL INEQUALITY

by

Cody Nathanael Tuttle

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Professor Melissa Kearney, Chair/Advisor
Professor John Haltiwanger
Professor Judith Hellerstein
Professor Brian Johnson
Professor Ethan Kaplan

Dedication

To my family: Marisa, my parents, my siblings, and my nieces and nephews.

Acknowledgments

I am deeply grateful to Melissa Kearney, my adviser, for her invaluable guidance and support. Her ability to be precise, clear, and rigorous without losing sight of the big picture has been inspiring and immensely helpful. As a mentor, she is generous, encouraging, and kind. I have learned a lot from her about the type of researcher that I want to be and the type of adviser that I want to be.

I also owe a great debt of gratitude to Ethan Kaplan and Lesley Turner, who encouraged me to start developing research ideas early and gave frequent feedback on those ideas. Without their support, I would not be where I am now.

I thank Judy Hellerstein for giving detailed comments on my drafts, asking difficult questions during presentations, and challenging me to continue to "think hard" about the issues I study. I am also grateful for the opportunity to work with John Haltiwanger in my position at the US Census Bureau. As a mentor and a co-author, he has shown me the value of asking important questions and answering them with very careful attention to detail. I thank Brian Johnson for contributing his time and expertise to my dissertation committee. I have been fortunate to receive feedback from him and John Laub on my research at the intersection of economics and criminology.

Numerous other faculty in the Economics department at the University of Maryland have contributed insightful comments and feedback to the chapters in this dissertation. For this help, I thank Katharine Abraham, Allan Drazen, Jessica Goldberg, Nolan Pope, John Shea, and Sergio Urzúa. For their help with countless hurdles, I thank Terry Davis, Vickie Fletcher, Jessica Gray, Amanda Statland, and Mark Wilkerson. I thank Cindy

Clement for the guidance and advice she gave me when I taught my first course. I am also grateful for the generous financial support from the Roger and Alicia Betancourt Fellowship in Applied Economics.

I am extremely grateful to my friends and classmates, including Joonkyu Choi, Prateik Dalmia, Alejandro Graziano, Seth Murray, Veronika Penciakova, Fernando Saltiel, Matthew Staiger, Mateo Uribe-Castro, Riley Wilson, and other fellow graduate students at the University of Maryland. My experience over the past six years and my work would not be the same without them.

I also acknowledge helpful comments from Stan Veuger, referees and editors, participants and discussants at several conferences, and my late uncle, Judge Robert Breetz. I thank Jimmy Grant, Alessio Ruvinov, Drew White, and my nephew, Ryan Willoughby, for assistance with data collection. I am also grateful to Todd Gardner, Shawn Klimek, and many other employees at the US Census Bureau for their advice and help with obtaining access to data. I owe many thanks to my cousin, Julie Breetz, for sharing her story of busing in Louisville, KY with me.

A series of opportunities led me to the Economics program at the University of Maryland, and I have a deep appreciation for the people who gave me those opportunities. First and foremost, John Perry, who hired me as a research assistant at Centre College and oversaw my first foray into economics research. His belief in me and his encouragement of my interests set me on the path I'm on today. Second, Chris Rodger, Aaron Marden, and the researchers I worked with at Mathematica Policy Research, who taught me a great deal about programming and conducting rigorous, policy-relevant research. Finally, Danny Shoag and Erich Muehlegger, who taught me how to carry out an

applied microeconomics research project and who were incredibly supportive as I applied to graduate school.

Finally, I cannot say thank you enough to my family. My parents, Donna and Steve, who made all of this possible and who have believed in me constantly. My siblings, Kim, Tab, Nick, Steph, and Brandon, who have been loving and supportive since day one. And my partner, Marisa, who has experienced the ups and downs of graduate school alongside me. She has helped me turn over new ideas, listened to me practice research presentations, and has been a constant source of emotional support. It is with great love and appreciation that I thank them all.

Table of Contents

List of Tables

List of Figures

Disclaimer

The second chapter of this dissertation was previously published as: Tuttle, Cody. 2019. ”Snapping Back: Food Stamp Bans and Criminal Recidivism.” *American Economic Journal: Economic Policy* 11 (2): 301-327. It is reproduced here with the following notice. Copyright American Economic Association; reproduced with permission of the *American Economic Journal: Economic Policy*.

The fourth chapter of this dissertation uses confidential data from the US Census Bureau, as such, the following disclaimer applies. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the US Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed (DAO/DRB approval number: CBDRB-FY19-CMS-7678).

Chapter 1: Introduction

Recent work finds that even after controlling for parental income, black men earn less in adulthood than white men, are less likely to complete high school and college, and are far more likely to be incarcerated (Chetty et al. 2020). In this dissertation, I study questions at the intersection of crime, race, and economics to understand the root of these disparities and provide evidence on government policies that aim to address crime and racial inequality.

I begin with two chapters that focus on the US criminal justice system, an area where racial disparities abound. As of 2017, approximately 1.4 million people were incarcerated in state or federal prisons (BJS 2019b). Non-Hispanic whites, despite accounting for approximately 60% of the US population, represented only 30% of that incarcerated population (BJS 2019b). Every year hundreds of thousands of prisoners are released and charged with the difficult task of successfully re-integrating into their community (BJS 2019b). In the second chapter of this dissertation, I study the role of financial assistance in easing that transition. Then, in the third chapter, I examine disparities in the criminal justice system themselves by studying racial differences in federal sentencing. Finally, in the last chapter, I shift focus to disparities in the provision of education and study the long-run economic effects of school desegregation.

In 2017 alone, state and federal prisons released over 620,000 individuals (BJS 2019b). Upon release, these individuals encounter numerous challenges in the re-entry process, including but not limited to finding legal work. Only 55 percent of releasees have any reported earnings in the year after their release (Looney and Turner 2018). Among those with reported earnings, the median earnings is only $10,090 (Looney and Turner 2018). In fact, Mueller-Smith (2015) estimates that an extra year in prison *causes* a 30 percent decrease in formal earnings post-release and a 3.6 percentage point decrease in formal employment. Despite these labor market challenges, many releasees are shut out from traditional forms of public support, such as public housing, unemployment insurance, cash assistance, and Food Stamps (CEA 2016). These barriers may contribute to the high rates of recidivism in the US, with 50% of released prisoners returning to prison within five years (BJS 2018).

In the second chapter, I study the effect of access to Food Stamps on criminal recidivism. In 1996, the Federal government passed a major welfare reform that, among other changes, imposed a lifetime ban from Food Stamps on people convicted of drug felonies. Florida modified this ban to apply only to people convicted of committing a drug trafficking offense on or after August 23, 1996. To identify the causal effect of the ban, I compare recidivism rates for people committing their offense just before August 23, 1996 to recidivism rates for those committing it just after. I find that the ban increases recidivism, specifically recidivism caused by banned individuals committing new financially motivated crimes (e.g. selling drugs and theft) to make up for the lost transfer income.

While this study evaluates one consequence of a government response to drug crime, it also speaks broadly to the relationship between public assistance and illegal labor supply. The results suggest that providing financial assistance to people with felonies could ease re-entry and reduce re-offending. Even more, the findings stress that policymakers should consider the illegal labor supply margin when designing policies that could encourage work.

The Food Stamps ban disproportionately affects black men and women because of their disproportionate contact with the criminal justice system. In the third chapter of this dissertation, I study one aspect of that disproportionate contact: racial disparities in prison sentences. Specifically, I study racial disparities in federal sentencing by examining a sudden change in the sentencing rules for crack cocaine offenses. Incarceration has high economic and non-economic costs, and a racial disparity in sentencing means those costs disproportionately impact black and Hispanic offenders (Haney 2001; Mueller-Smith 2015; The Hamilton Project 2016; BOP 2020). In addition, the federal government spends billions of dollars per year on the criminal justice system (BJS 2019a). If a government actor in that system is making decisions that produce racial disparities, we should understand what those decisions are, who is making them, and why they are being made.

In federal drug cases, a mandatory minimum sentence is triggered based on the amount of drugs involved in the offense. The amount "involved," however, is not necessarily equal to the amount seized, and it is subject to legal discretion from police and prosecutors.[1] In 2010, the Fair Sentencing Act changed that trigger amount for crack

[1] This discretion does not pose problems for the analysis of the SNAP ban in Chapter 2. State prosecutors, in general, have less discretion over drug amounts than federal prosecutors. In order for state prosecutor discretion to impact the results in Chapter 2, it must be the case that prosecutors discontinuously change their behavior at August 23, 1996. Chapter 2 provides contextual and empirical evidence that this does not occur. Finally, there is no reason for prosecutors to change their behavior concerning drug amount around August 23, 1996. Florida did not modify the ban to apply only to drug traffickers until May 1997,

cocaine from 50 grams to 280 grams. Using data on the universe of federal drug cases sentenced from 1999-2015, I find that the fraction of cases right at and above 280 grams suddenly increases after 2010. Moreover, the increase in cases at 280 grams is disproportionately large for black and Hispanic offenders. This phenomenon alone can explain up to 7% of the racial sentencing gap in crack cocaine cases after 2010.

To determine where the "bunching" at 280 grams first occurs, I use several additional datasets that permit a detailed view of the criminal justice system. I find that the increase in cases at 280 grams is the result of a subset (about 30%) of federal prosecutors exercising wide discretion over the total amount that is charged in these cases. To investigate mechanisms even further, I identify a pivotal Supreme Court case from June 2013 that tightens evidentiary standards specifically in mandatory minimum cases (Bala 2015). The fraction of cases at 280 grams falls immediately after the Court's decision in that case, suggesting that those cases would not have held up under the stricter evidentiary standards.

Finally, I explore several explanations for this racial disparity, including discrimination. Among other things, I show that the disparity is not driven by observable correlates of race and that it is seemingly unrelated to costs to the prosecutor of developing a case (e.g. costs imposed by defense counsel or judges). In addition, I highlight that within the fairly narrow geography of a federal district, some prosecutors bunch cases at 280 grams and some do not, a fact that rejects a simple model of statistical discrimination. I then use a measure of state-level racial animus developed from Google Trends data by Stephens-Davidowitz (2014), and I find that the racial disparity in bunching at 280 grams is much

meaning that prosecutors, as of August 23, 1996, should expect all drug amounts to trigger the ban.

larger in states with higher levels of racial animus. These analyses suggest the disparity is due to taste-based discrimination.

Mandatory minimum sentences were designed in the 1980s with the goals of reducing drug crime and making sentencing an objective process. The Fair Sentencing Act in 2010 relaxed the crack cocaine thresholds to address a long-standing disparity between crack and powder cocaine sentences. I highlight the limitations of these seemingly objective policies by documenting the use of discretion in amounts charged and the racial disparity in the use of that discretion. More broadly, this paper sheds light on the role of prosecutorial discretion and potentially discrimination as causes of racial disparities in sentencing. Prosecutors are given significant latitude over charging decisions and are thus an important actor in the criminal justice system. Understanding the link between their decisions and sentencing disparities is critical for determining how those disparities can be reduced.

While the third chapter studies how similar defendants are treated differently by race *after* entry into the criminal justice system, the fourth chapter in this dissertation studies how opportunities and education policy lead to the myriad racial inequalities present outside of the criminal justice system. Using a novel empirical strategy and confidential data from the US Census Bureau, I study the long-run effects of school desegregation, one of the most extensive government efforts to reduce racial inequality in US history.

In 1954, the Supreme Court ruled that the segregation of schools is unconstitutional. That decision in *Brown v. Board of Education* and a slate of subsequent rulings triggered hundreds of court-ordered desegregation plans over the next three decades. By 1988, 44% of black children were attending majority white schools (Orfield et al. 2014). Despite their success in terms of integrating students, many of these plans were dissolved after a suite of Supreme Court decisions in the early 1990s. As a result, the released school districts have re-segregated (Lutz 2011). Today, only 23% of black children are attending majority white schools (Orfield et al. 2014).

This re-segregation is alarming because several studies show that school segregation has detrimental short-run effects for black students (e.g. Lutz 2011; Billings, Deming, and Rockoff 2014; Cook 2016; Bergman 2018). In addition, prior work finds that school desegregation has short-run benefits for black students (e.g. Guryan 2004; Reber 2009; Johnson 2015). Despite the widespread use of busing in the 70s and 80s, we know little about its effects on long-run outcomes like final educational attainment, earnings, employment, or residential location in adulthood (aside from Johnson 2015). Even more, since prior work primarily studies district-level changes that occur after court orders are handed down, their estimates capture the net effect of a school desegregation policy. Absent within-district evaluations of desegregation, it is challenging to disentangle the effect of changing peers versus changing school resources.

I focus on Jefferson County, KY, a large school district in the US, and the unique busing assignment plan it adopted to merge its majority-white County schools with its majority-black City schools. Under the plan, students were assigned to be bused to new schools (versus stay at their home school and have new students bused in) based on their race and the first letter of their last name. I use this busing assignment scheme and administrative records on individual place of birth linked with responses to the 2000 Decennial Census to estimate the long-run economic effects of busing assignment.

I find that black students assigned busing to former County schools live in neighborhoods with higher tract-level income and higher tract-level education in adulthood than black students assigned to remain in former City schools. For white students, busing assignment has a small negative or zero effect. To understand how "white flight" affects these estimates, I use a newly collected dataset of nearly 100,000 student records from archival school yearbooks and commencement programs. Results from these records suggest that the effect of busing on white students is small and statistically indistinguishable from zero even when white disenrollment is considered. Since the former City and former County schools are equally integrated once busing is in effect, these results suggest that the desegregation in Jefferson County improved the long-run outcomes of black students by sending them to better schools.

Overall, this dissertation explores government responses to crime and racial inequality by focusing on two critical policy areas, criminal justice and education. First, I study a government response to drug crime that prevents people with drug felonies from receiving Food Stamps. I show that this ban increases recidivism rates, suggesting that financial assistance for released offenders may reduce re-offending. Then, in the third chapter, I study racial disparities in federal sentencing. Specifically, I examine abnormal bunching in the distribution of drug amounts charged before and after the Fair Sentencing Act. I provide new evidence that prosecutors in federal court treat black and Hispanic defendants differently than similar white defendants. Finally, in the last chapter, I turn to inequalities in the provision of education. I show that a school desegregation plan in Jefferson County, KY had long-run positive effects on black students but no negative effect on the economic outcomes of white students. Together, these papers shed light on the criminal justice system, the disparities therein, and racial inequalities in education.

Chapter 2: Snapping Back: Food Stamp Bans and Criminal Recidivism

2.1 Introduction

Since the late 1990s, state and federal prisons in America have released over half a million prisoners every year (Council of Economic Advisors (CEA) 2016). Upon release, these offenders face a myriad of obstacles that inhibit a successful transition into a new life as law-abiding citizens.[1] To start, offenders have trouble finding work—survey evidence suggests over half are unemployed even a year after release (Schmitt and Warner 2010). Job searchers with a felony conviction are subject to extra scrutiny in the hiring process. Recent audit studies suggest that a felony conviction cuts probability of being called back by an interviewer in half (Pager, Western, and Sugie 2009). In addition, some occupational licensing rules bar felons from ever entering an occupation (Bushway and Sweeten 2007). Furthermore, offenders do not meet the requirements of the Unemployment Insurance program upon release and are frequently denied public housing by local Public Housing Authorities (CEA 2016). Finally, as a consequence of the 1996

[1] I use the terms “offender”, “ex-offender”, “former offender”, “prisoner”, “inmate”, “felon”, “releasee”, etc. frequently throughout this paper. These terms describe different groups. However, convicted and released drug traffickers (whom I also frequently refer to as simply “drug traffickers”), the focal group of this paper, belong to all of those groups or belonged to them at one point.

welfare reform, many offenders are now banned from receiving Supplemental Nutrition Assistance Program (SNAP, formerly named Food Stamps) and Temporary Assistance for Needy Families (TANF) benefits. With this in mind, it may not come as a surprise that half of releasees are back in prison within five years of their release and three-quarters are re-arrested within five years (CEA 2016). Recidivism in America may be at least partly the consequence of these barriers to reentry.

In this paper, I focus on one of those barriers, the SNAP ban, and ask how it affects recidivism outcomes, defining recidivism as a return to prison after release. It is particularly critical that we understand the effect of the SNAP ban because it is currently in effect in 27 states, and because survey evidence suggests SNAP is an important resource for offenders post-release (Wolkomir 2018). Approximately 70 percent of the former inmates in the Boston Reentry Study report receiving SNAP benefits even just two months after release (Western et al. 2015).[2] Even more, SNAP benefits are an important component of income for recipients. Based on a representative sample of adult male recipients (not limited to offenders), SNAP benefits make up approximately 20 percent of their reported gross income (see Table 2.2). Finally, to the extent that SNAP availability has insurance value, it may also affect the decisions of non-recipients.

To study the effect of the SNAP ban on recidivism, I use a federal policy change (as it was implemented in Florida) that imposed a lifetime ban from SNAP receipt on offenders who committed drug trafficking on or after August 23, 1996.[3] I will often refer

[2]Similar estimates of SNAP usage among households with an interaction with the criminal justice system can be found in the Fragile Families & Child Wellbeing Study (Sugie 2012) and in the Panel Study of Income Dynamics.

[3]I focus on Florida in this paper for a number of reasons, the foremost being that inmate-level data for all offenders released after October 1, 1997 is publicly available for download. Florida also has more people in prison or jail than all states but two (California and Texas) and has more people participating in

to this as "the cutoff date" in the remainder of the paper. Offenders committing drug trafficking on or after this date are also subject to a lifetime ban from TANF benefits. That said, over 85 percent of drug traffickers are male and less than 10 percent of TANF recipients are male—if TANF does play a role, it is likely to be small in comparison to SNAP, for which almost 40 percent of recipients are males aged 18-65 (U.S. Department of Health and Human Services (HHS) 2015). For this reason, I refer to the treatment only as "access to SNAP" or "the SNAP ban" in the remainder of the paper. To estimate the causal effect of the ban on recidivism, I employ a regression discontinuity design that compares outcomes for offenders who committed drug trafficking in a small window before the cutoff date to outcomes for offenders who committed it on or slightly after the cutoff. I find the SNAP ban has increased the probability of recidivism among drug traffickers.

Specifically, I find that drug traffickers subject to the ban are about 9 percentage points more likely to return to prison after release than drug traffickers who have access to SNAP. An increase of this size is large for drug traffickers in Florida. Among those offenders who commit their trafficking offense in the 240 days before the cutoff date, about 16 percent return to prison at some point post-release. This implies that the SNAP ban increased recidivism among drug traffickers by about 60 percent. However, this estimate is based on the small sample of about 1,000 drug traffickers committing an offense suffi-

SNAP than all states but two (again, California and Texas) (Kaeble and Cowhig 2016; Food and Nutrition Service (FNS) 2017). Finally, the discontinuity is well-functioning in Florida—I find no evidence of sorting, manipulation, or endogenous responses near the cutoff. I explored a similar policy discontinuity in North Carolina, but found evidence of sorting near the cutoff—offenders on the other side of the cutoff were older, more risky, and received higher sentences. In addition, a McCrary density test suggested a drop in crime right after August 23, 1996 in North Carolina. This invalidates the current approach in the context of North Carolina, and hence I focus on Florida.

ciently close to the cutoff date. Although I am able to reject a null effect of the ban, the estimate is noisy and the confidence interval is large. The 90 percent confidence interval on the main estimate is 1.7 percentage points to 17 percentage points, which implies the SNAP ban increased recidivism among drug traffickers by about 10 percent to 105 percent. Unfortunately, I do not have the statistical power to produce a more precise range of possible effect sizes.

Furthermore, the increase in recidivism is primarily driven by an increase in recidivism for financially motivated crimes (such as property crime and selling drugs). This result has important implications for state SNAP bans and for reentry policy in general. In fact, it is consistent with recent work by Munyo and Rossi (2015) showing that a disproportionate amount of recidivism happens on the first day of release and that first-day recidivism can be almost completely stifled by giving releasees a sufficient monetary stipend. Their work suggests that financial support can ease reentry. I provide further support for this idea by showing that recidivism increases after we decrease financial support to offenders by banning them from SNAP.

More broadly, this paper contributes to a literature in public economics that studies labor supply responses to transfer programs. Economic theory predicts that denying offenders SNAP benefits will incentivize work, encouraging offenders to reenter the labor force and earn the money necessary to put food on the table. For a number of reasons, however, finding employment in the legal sector is a challenge for ex-convicts. As such, the work incentives could drive offenders back into the illegal sector. The evidence in this paper is consistent with a model in which removing SNAP benefits does increase the labor supply of drug traffickers.

This relates to work by Hoynes and Schanzenbach (2012) that finds reductions in employment and hours worked for female-headed households after Food Stamps is introduced in a county. In this paper, I emphasize the importance of considering the illegal labor margin when designing policies that will affect work incentives, especially when those policies will be applied to people who have high attachment to the illegal labor market or high difficulty entering the legal labor market, both of which are true in the case of drug traffickers.

Finally, a number of papers have documented a long list of benefits from SNAP and safety net programs in general. First and foremost, SNAP relieves families of food insecurity and reduces poverty (Mabli and Ohls 2015; Short 2015). In addition, recent research suggests that SNAP receipt leads to a wide range of other positive outcomes, including improved adult health, improved child health in the long-run, better birth outcomes, and higher test scores for primary school students (Almond, Hoynes, and Schanzenbach 2010; Gassman-Pines and Bellows 2015; Gregory and Deb, 2015; Hoynes, Schanzenbach, and Almond 2016). I add another policy-relevant benefit to that list—access to SNAP decreases recidivism among drug traffickers.

Making a few crude but conservative assumptions about the cost of incarcerating an extra person and the social cost of crime, I can use the estimated effect of the ban on recidivism to calculate the societal cost of the SNAP ban in Florida. A more comprehensive cost-benefit analysis of the ban is beyond the scope of this paper, as it would require estimates of the effect on legal employment and the deterrence effect of the ban for would-be first-time traffickers. Rather, this cost estimate is intended to highlight the potential benefit of reducing recidivism by providing SNAP or other financial support post-release. I estimate the ban costs Florida about $3,700 per banned person. Given that Florida has approximately 19,000 people currently subject to the ban, this implies that the ban has cost the state over 70 million dollars to date, a number that grows with each drug trafficker shut out from SNAP.

The remainder of the paper is organized as follows. Section 2.2 recounts a short history of the SNAP ban, and Section 2.3 reviews the related literature. I describe the data in Section 2.4. Section 2.5 presents the methodology and Section 2.6 discusses the corresponding results. Section 2.7 concludes.

2.2 The Federal SNAP Ban

The passage of the Personal Responsibility and Work Opportunity Act (PRWORA) in 1996 dramatically changed welfare programs in America. Along with other major changes to welfare policy, PRWORA imposed a lifetime ban from SNAP on felony drug offenders. The ban was introduced as an amendment to the act by Senator Phil Gramm and passed through Congress with little opposition. Upon introducing the amendment, Senator Gramm argued, "if we are serious about our drug laws, we ought not to give people welfare benefits who are violating the Nation's drug laws." Based on remarks by Senator Connie Mack, it also appears that some believed that drug dealers should not receive benefits since, were their informal earnings counted, they would likely be ineligible (U.S. Congress 1996, S8498).

Since the passage of PRWORA, many states have modified or repealed the SNAP ban. Currently, 46 states have opted-out or modified the SNAP ban, up from only half of all states in 2002 (Gilna 2016; Wolkomir 2018). While some states have opted out entirely, many states have modified the ban to grant eligibility to people convicted of substance abuse crimes or to require enrollment in substance abuse treatment classes to become eligible (Wolkomir 2018). Florida quickly modified the ban such that it would only apply to people convicted of drug trafficking crimes committed on or after August 23, 1996.[4]

In Florida, drug trafficking constitutes the selling, manufacturing, or distributing of illegal drugs in large amounts. For example, a person is charged with "trafficking heroin" if they sell, manufacture, or distribute greater than 4 grams of heroin (FL Statute 893.135). Importantly, "selling, manufacturing, or distributing" (henceforth referred to as SMD) is a separate offense category that applies to people who sell, manufacture, or distribute illegal drugs in smaller amounts. People convicted of SMD or felony possession are eligible for SNAP benefits in Florida, regardless of when the offense was committed. I use these groups in placebo tests to emphasize that the increase in recidivism is specific to drug traffickers, the offenders who are banned from SNAP if they commit the offense after the cutoff date.

[4]The application for SNAP in Florida has a section that requires applicants to report whether or not they have been convicted of a drug trafficking offense that was committed on or after August 23, 1996. While the Florida Department of Families and Children does not have an automated system to cross-check applications with the Florida Department of Corrections, offender information is easily searchable online. The Office of Public Benefits Integrity in Florida has also partnered with the Florida Department of Corrections in the past to identify drug traffickers who were currently receiving or had received SNAP benefits. Florida estimates approximately $360,000 worth of SNAP and cash assistance benefits had been disbursed to ineligible individuals. Assuming those benefits were strictly SNAP benefits, that the average recipient stayed on SNAP for one year, and that the average benefit per month is $150, this implies only 200 drug traffickers were receiving benefits for which they were ineligible. Florida is home to approximately 19,000 drug traffickers who are subject to this ban, implying that only 1 percent of drug traffickers subverted the ban.

2.3 Related Literature

In this paper, I build on three literatures in economics and criminology by studying the effect of the SNAP ban on drug traffickers in Florida. To my knowledge, I provide one of the first empirical evaluations of a policy that currently affects former drug offenders in 27 states. This policy evaluation contributes broadly to the literature on prisoner reentry, specifically that which explores the effects of financial support for released offenders. Second, I contribute new evidence highlighting the relationship between financial need and criminal behavior. Finally, I add to an extensive literature in public economics that studies the effect of cash and in-kind transfers on labor supply.

For ex-offenders, finding legal work can be especially difficult. A large literature discusses the challenges that offenders face when looking for legal work, from occupational licensing restrictions to employer discrimination to the detrimental effects of incarceration itself. I provide a broad review of this literature and other work on prisoner reentry in Appendix B. The immense difficulty of successfully reintegrating into life outside of prison has spurred an interest in programs that can ease the transition and prevent offenders from returning to crime. In this paper, I examine one reentry strategy: providing financial support to offenders via SNAP. This builds on a growing literature on the effect of giving offenders financial support upon release.

In concurrent work, Yang (2017a) and Luallen, Edgerton, and Rabideau (2017) study the effect of the SNAP and TANF bans on criminal recidivism. Both papers contribute further evidence to this important policy question. Luallen, Edgerton, and Rabideau use data from the National Corrections Reporting Program (NCRP) which includes information about prison admissions and releases for several states. The authors also use the discontinuity in banned status at the cutoff date in addition to variation in state-level modifications of the SNAP ban. They find no effect on recidivism.

I depart from the analysis in Luallen, Edgerton, and Rabideau in two major ways. First, I focus on longer-run recidivism outcomes, while they study the effect on recidivism within 3 years. In this paper, I also find a small and statistically insignificant positive effect on recidivism within 3 years. Second, I use administrative data from Florida that includes the date each offense was committed. The NCRP data does not include the date the offense was committed, and thus, the authors must use conviction date (proxied by prison admission date) to identify treatment. Since the ban is actually determined by the date the offense was committed, the authors have a very noisy measure for treatment (convictions often take place months or years after the date the offense was committed). This measurement error will attenuate their results. In fact, I reestimate the main results from this paper using conviction date rather than offense date and also find a statistically insignificant effect on recidivism (results in Table A2.28).

Yang (2017a) exploits the extent to which states opt out of the Federal ban and the differential timing of opt-out. Yang uses state-by-time-by-crime variation in the application of the ban in a triple difference design. Using data from the NCRP, she finds that access to SNAP benefits decreases the probability of returning to prison within one year by about 2.2 percentage points or 13 percent from the mean. This result is consistent with my findings that access to SNAP decreases the probability of re-incarceration for drug traffickers.

My paper presents a more comprehensive analysis of the SNAP ban by examining long-run recidivism outcomes and the types of crimes offenders commit due to the ban. In addition, I focus on drug traffickers, a group of offenders who have ties to the illegal labor market and thus, may be most at risk to return to it. Also, several states that have partially opted out of the ban have, like Florida, maintained the ban for drug traffickers. Finally, the estimates from the triple difference design are biased if states enact policies that specifically affect drug felons in the same year that they opt out of the welfare ban. I approach the evaluation of this ban with a regression discontinuity design that is not subject to that concern.

There is an older literature in criminology and sociology that analyzes random experiments that allocate unemployment benefits to offenders and consistently finds that financial support decreases probability of re-arrest for property crimes (Mallar and Thornton 1978; Berk, Lenihan, and Rossi 1980). Specifically, these studies find that financial aid for ex-offenders reduces their likelihood of re-arrest for property crime by about 8-27 percent.[5] The effect of these programs on re-arrest in general is less clear, but the largest effects are concentrated in re-arrest for property crimes, which is both consistent with theory and with the results in this paper. Interestingly, Berk and Rauma (1983), in an early application of regression discontinuity design, also find that giving unemployment benefits to offenders decreases the likelihood of recidivism (defined as re-incarceration, parole revocation, or parole violation) by about 13 percent. As Raphael (2011) points out, the cash assistance programs studied in the 70s and 80s typically had benefit reduction rates from formal earnings of 100 percent, and as a result, led to a substantial drop in

[5] Berk, Lenihan, and Rossi do not find an effect of financial aid in their reduced form analysis of the experiment. They introduce a model that incorporates legal employment effects and report the results of that model.

formal labor supply that may have had an offsetting effect on recidivism.

Another compelling line of research documents an increase in crime two to three weeks after welfare disbursement, suggesting recipients are spending down the entire check and committing crimes until the next payment (Foley 2010). Similarly, Carr and Packham (2017) demonstrate that theft in grocery stores in Chicago fell dramatically after Illinois implemented a staggered disbursement schedule for SNAP. They leverage variation in benefit issuance based on first-letter of the recipient's last name and estimate similar effects from a shift in issuance dates in Indiana. This work further highlights the relationship between transfer programs and crime.[6] A more detailed review of the literature on financial need and criminal behavior is in Appendix B. The results in this paper, that the SNAP ban increases recidivism among released drug traffickers, provide further evidence that financial need is an important factor in the decision to commit crime.

The work cited above ties into a distinct literature in public economics about the effect of transfer programs on labor supply. Both theory and empirical evidence suggests that transfer programs discourage work. For SNAP, in particular, Hoynes and Schanzenbach (2012) use variation in county-level rollout of the Food Stamps program and find that the introduction of Food Stamps in a county decreases annual hours worked in those households most likely to be affected by the program (noneldery, female-headed households).[7] Their paper provides valuable evidence about the labor supply response of

[6]Studies of the effect of housing vouchers on crime tend to find a negligible or negative effect of voucher receipt on crime (Jacob, Kapustin, and Ludwig 2015; Carr and Koppa 2017). Carr and Koppa (2017) argue that vouchers free up financial resources to such an extent that they effectively subsidize spending on things that are complements to crime, like alcohol.

[7]For another example, Jacob and Ludwig (2012) exploit variation in housing voucher receipt from randomized placement on a waitlist in Chicago and find that voucher use decreases labor force participation by 6 percent. Also, Deshpande (2016) uses a policy discontinuity from PRWORA to demonstrate that children removed from SSI increase their labor supply but not by enough to offset the lost benefits.

female-headed households to Food Stamps, but evidence for the labor supply of males is necessarily limited, and there is no consideration of illegal labor supply. While I do not observe hours worked or wages, I do observe recidivism, which for many drug traffickers corresponds to participation in the illegal labor market.

In summary, public economic theory as well as empirical evidence suggests that decreasing transfer income may push workers back into the labor force. Yet other work highlights the difficulty offenders face in the legal labor market and the ease with which they can reenter the illegal labor market (see Appendix B). A strong incentive to return to work coupled with the difficulty of finding legal work may drive offenders back to the illegal sector (see Appendix C for a formal model of this phenomena). Existing research on the effect of financial support on recidivism typically focuses on short-run outcomes or considers financial support programs that differ markedly from SNAP in terms of benefit amount, potential length of receipt, and benefit reduction rate. The effect of the SNAP ban on recidivism speaks to labor supply responses to SNAP benefits, and even more, it directly relates to current prisoner reentry policy.

2.4 Data and Descriptive Statistics

2.4.1 Offender Data

Florida Department of Corrections (FL DOC) makes data from its Offender Based Information System (OBIS) publicly available. These data include information about both active offenders and released offenders. I combine offense-level data, prison stay-level incarceration histories, and offender-level demographic data into a dataset where each observation is a unique prison stay. Using this data, I calculate recidivism for a given stay j as whether or not the offender ever has a prison stay occurring after stay j. Likewise, that recidivism is recorded as "financially motivated" if the offender was charged with a financially motivated crime for the prison stay occurring after stay j and that recidivism is recorded as "non-financially motivated" if the offender was not charged with a financially motivated crime for the prison stay after stay j.[8] In some analyses, I use a measure of time until recidivism–this is defined as the time between release from prison stay j and the earliest offense occurring after stay j.

I limit this data to offenses committed after October 1, 1995. First, Florida implemented a suite of criminal justice reforms that apply to offenders committing offenses on or after October 1, 1995. Most notably, offenders sentenced after October 1, 1995 are required to serve 85 percent or more of their sentence. Kuziemko (2013) shows that fixed-sentencing systems alter incentives for offenders while in prison, stifle the allocative efficiency of parole boards, and ultimately, increase recidivism. Restricting the sample to offenses committed after October 1, 1995 avoids including offenders that were sentenced under a drastically different system. Second, offenders are included in the publicly available OBIS data if they committed a felony, served time in a Florida prison for that felony, and were released after October 1, 1997. If an offender meets those three criteria, then all of their stays in FL prisons are included in the data. Limiting the sample mitigates sample selection problems arising from that restriction imposed by FL DOC.[9] Further details on

[8]FL DOC categorizes most offenses here: http://www.dc.state.fl.us/AppCommon/offctgy.asp#PC. I define financially motivated crimes as: property crimes (excluding property damage crimes such as vandalism), selling/manufacturing/distributing drugs, drug trafficking, fraud, forgery, racketeering, prostitution, counterfeiting, and crimes containing a "$", "sale", or "sell" in the charge description. I define non-financially motivated crimes as all crimes that are not categorized as financially motivated.

[9]Only six drug trafficking offenders in the data from October 1, 1995 to October 1,1997 are released

data construction are in Appendix E.

For the main results, I also remove individuals who are identified as Hispanic in the data (less than 7 percent of my sample). PRWORA restricted access to SNAP for documented and undocumented immigrants regardless of criminal history. In addition, non-citizen immigrants often face deportation after committing drug trafficking since it is classified as an "aggravated felony" under the Immigration and Nationality Act. For these reasons, many non-citizen immigrants will lose access to SNAP regardless of the date their offense is committed, thus including them in the sample will attenuate the estimated effect. Unfortunately, I do not observe immigrant status in the data. In the 2000 Census, about 41 percent of Hispanic individuals "institutionalized" in Florida are born outside of the US and less than 5 percent of Black or White individuals institutionalized in Florida report a birthplace outside the US. I report the main results on recidivism with Hispanics included in Table A2.4 to demonstrate that the results are qualitatively similar, but as expected, are attenuated slightly.

Summary statistics for offenders who committed offenses from October 1, 1995 to October 1, 1997 are reported in Table 2.1 for three groups: drug traffickers, all non-drug offenders, and offenders convicted of selling, manufacturing or distributing drugs (SMD offenders). I also report summary statistics for all drug traffickers released after October 1, 1997. Drug traffickers are quite different from offenders who commit other crimes. As Table 2.1 shows, recidivism is lower for drug traffickers than non-drug offenders or

---

prior to October 1, 1997. The results are not affected by the inclusion of these six offenders. Also, on average, drug traffickers are sentenced to approximately 4.6 years, and over 90 percent of drug traffickers are sentenced to 2 years or more. Finally, selection bias from the FL DOC restriction will bias all results downward since offenders in the control group (those committing an offense prior to August 23, 1996) are more likely to be released prior to October 1, 1997 and thus only observed in the event of recidivism.

SMD offenders in Florida. When benchmarking the recidivism results I find in Section 2.6, it is important to keep in mind the rates at which other criminals return to prison. A 9 percentage point increase in the recidivism rate of drug traffickers does not yield an unrealistic recidivism rate, rather, it yields a rate of recidivism that is still lower than the rates for non-drug offenders and other drug offenders.

2.4.2 SNAP Quality Control Data

Using the 1996-2014 SNAP Quality Control files provided by Mathematica Policy Research, I report summary statistics on the SNAP population in Florida in Table 2.2. I focus on male recipients aged 18-65 for this exercise since 89 percent of offenders are male. These statistics paint a picture of the male SNAP population in Florida and contain two key observations: (1) the SNAP benefit is an important source of income and (2) recipients do not have to be employed to receive SNAP benefits, despite the well-known work requirements of post-PRWORA SNAP.

Notably, the SNAP benefit men receive in Florida is around 20 percent of the total gross income they report. SNAP transfers are a sizable portion of gross income for this population. This statistic gives us a rough estimate of the toll of the SNAP ban on offenders. Assuming SNAP transfers would make up the same share of drug traffickers' reported gross income, then the SNAP ban effectively denies offenders this stream of income upon release. In other words, offenders who commit drug trafficking on or after August 23, 1996 are banned from SNAP and thus take home 20 percent less in gross income than offenders who commit drug trafficking just before August 23, 1996. Again, this is an estimate based on the SNAP benefits of male recipients aged 18-65 in Florida. SNAP transfers may represent more or less than 20 percent of former drug traffickers' gross income. In this light, it makes sense that there are potentially large effects of the SNAP ban on recidivism, especially since SNAP take-up among former offenders is high.

Approximately 70 percent of the former inmates in the Boston Reentry Study report receiving SNAP benefits even just two months after release (Western et al. 2015). Sugie (2012) also finds that about 70 percent of families in the Fragile Families & Child Well-being Study with a recent paternal incarceration report receiving SNAP in the past year. The Panel Study of Income Dynamics asks respondents in 1995 if they have ever been in the corrections system (jail, prison, youth corrections). Almost 50 percent of respondents who answered yes to that question were in families that reported receiving SNAP at some point from 1995-2013. Unfortunately, I cannot identify the subsample of these people who have been to prison (given prison, jail, and youth corrections are three very different populations).

PRWORA also introduced more stringent work requirements for SNAP recipients. Perhaps the requirement most relevant to this study is the work requirement for able-bodied adults without dependents (ABAWDs) since many offenders may be considered ABAWDs. The ABAWD work requirement states that able-bodied adults without dependents are limited to only 3 months of SNAP receipt every 3 years unless they: (1) work 20 or more hours per week, (2) participate in an employment and training program, or (3) participate in a workfare program (U.S. Department of Agriculture (USDA) 2016b).

First, note that ABAWDs do not have to be employed to meet the requirement; they can meet the requirement by enrolling in employment and training programs, many of which are actually targeted at ex-offenders (USDA 2016a). In fact, Table 2.2 shows that only 10 percent of single males receiving SNAP are employed and only 40 percent of men with families are employed. Second, when states face tough economic times, they can request to waive this requirement. This requirement was waived nationally from 2001-2003 and 2009-2016. In addition, the requirement was waived prior to 2009 for Labor Surplus Areas (counties in Florida with especially high unemployment) and for counties where Florida chose to apply a special exemption that allows states to exempt 15 percent of the state's caseload from the work requirement (USDA 2016b).

I exploit this variation in the ABAWD requirement and find that the SNAP ban does have the largest effect on recidivism when the ABAWD requirement is waived in Florida. The table below shows statistics broken down by years with and without nationwide ABAWD work requirement waivers. SNAP benefits are higher in years with nationwide ABAWD waivers, and single males represent a greater portion of the male SNAP population in Florida during those years.

2.5 Methodology

SNAP eligibility for drug traffickers is determined by a sharp cutoff date. Offenders who committed drug trafficking before August 23, 1996 are eligible for SNAP benefits, while offenders who committed drug trafficking on or after August 23, 1996 are permanently banned from SNAP. To estimate the effect of the SNAP ban on recidivism, I employ a regression discontinuity design that exploits this sharp policy rule. In general, the regression model is as follows:

$$Recidivism_{it} = \alpha + \beta_1 After_{it} + g(DaysFromCutoff_{it}) + g(DaysFromCutoff_{it}) \times After_{it} + \omega_{it} \quad (2.1)$$

where $Recidivism_{it}$ is equal to one if the offender i at time t ever returns to prison after being released and equal to zero if the offender does not return to prison.[10] $After_{it}$ is an indicator equal to one when the offense is committed on or after August 23, 1996 and equal to zero otherwise—this indicates whether the offender is subject to the SNAP ban or not. $g(DaysFromCutoff_{it})$ is a flexible function of offender i's offense date expressed as number of days from August 23, 1996 (centered at zero). The interaction term allows the relationship between the running variable (distance from August 23, 1996) and recidivism to vary before versus after the cutoff. No baseline covariates are included in this specification.[11]

My preferred specification for all results is the local linear regression discontinuity design with a rectangular kernel. I present the main results in this paper using two bandwidths. First, I show every result using the Imbens and Kalyanaraman (IK) (2012) optimal bandwidth chosen for that regression with polynomial of degree one and a rectangular kernel. This procedure yields different bandwidths for every dependent variable.

[10] Throughout the paper, I introduce a variety of "recidivism" measures. For example, I also estimate equation (2.1) on "financially motivated recidivism" and "non-financially motivated recidivism." Financially-motivated recidivism is equal to one if the offender returns to prison with any crime that is financially motivated and is equal to zero if the offender returns to prison only with crimes that are not financially motivated or if the offender does not return to prison. Non-financially motivated recidivism is equal to one if the offender returns to prison only with crimes that are not financially motivated and is equal to zero if the offender returns to prison with any crime that is financially motivated or if the offender does not return to prison.

[11] If covariates are orthogonal to the treatment and explain recidivism, including them should tighten my standard errors without changing the magnitude of my coefficients. I introduce controls for offender characteristics and offense day-of-week fixed effects in Table A2.5 and find that the results are similar but more precise.

For example, when examining the effect of the ban on any recidivism, the optimal bandwidth is ± 212 days from August 23, 1996 whereas the optimal bandwidth is ± 242 days when examining the effect of the ban on financially motivated recidivism. In addition, since I limit the data to offenses occurring after October 1, 1995, any bandwidth greater than ± 327 days will be asymmetric. For these reasons, I also include results based on a consistent bandwidth of ± 240 days.[12]

The choice to focus on the local linear design is motivated by Gelman and Imbens (2018) who suggest using lower-order polynomials. However, in a working paper, Card et. al (2014) argue that the optimal polynomial is dependent on the underlying data generating process, and in some cases, higher-order polynomials are indeed optimal. In addition, while I focus on the IK optimal bandwidth in this paper, other researchers have designed alternative algorithms for choosing a bandwidth (Ludwig and Miller 2007; Calonico, Cattaneo, and Titiunik 2014). I show that the main results are robust to higher order polynomials, alternative kernels, and many alternative bandwidths.

The main identifying assumption with the regression discontinuity design is that all unobserved determinants of recidivism are continuous with respect to the offense date (Imbens and Lemieux 2008). This assumption, although inherently untestable, does yield testable implications. First, the observable characteristics of offenders should be continuous across the threshold. Second, the density of drug trafficking offenses should also be continuous across the threshold. I test for discontinuous breaks in observed characteristics at the cutoff by estimating the following:

[12] The bandwidth is convenient because it corresponds to an even number of months (8 months before and after the cutoff) and is the average of the three IK optimal bandwidths for any recidivism, financially motivated recidivism, and non-financially motivated recidivism rounded to the nearest ten

$$Characteristic_{it}^{D} = \alpha + \beta_1 After_{it} + g(DaysFromCutoff_{it}) + g(DaysFromCutoff_{it}) \times After_{it} + \omega_{it} \tag{2.2}$$

where $Characteristic_{it}^{D}$ is an indicator for whether or not the offender i on day t is black, male, their age at intake, their total sentence length, the type of drug they are charged with trafficking, the number of prior offenses for which they have been convicted, and the number of concurrent offenses for which they were convicted. In addition, I test for a break in risk of recidivism. I calculate risk of recidivism using a logistic regression of recidivism on all characteristics and age-squared. I run this regression for those offenders not subject to the ban and not in the ± 212 day IK bandwidth (those committing drug trafficking from October 1, 1995 to January 24, 1996) and predict the "risk score" for offenders in my sample.

Results from the "risk score" test are presented in Figure 2.1, while Table A2.2 and Figures A2.1a-A2.1h show the results for each characteristic separately. If the identifying assumption is violated, we would expect to see a significant difference in observable characteristics after August 23, 1996 ($\beta_1 \neq 0$). I find no evidence of sorting around the cutoff on observable characteristics. I also run a regression of the dummy variable indicating the offense was committed after the cutoff on total years sentenced, race, age, number of concurrent offenses offense, type of trafficking, sex, and number of prior offenses. A joint significance test on the covariates in this regression further suggests no sorting occurred near the cutoff (p-value=0.9504). These results lend credence to the assumption that offenders, judges, police, and lawyers are not changing their behavior in response to the policy.[13]

In addition, I conduct a McCrary density test for excess mass in the number of drug trafficking crimes on either side of the discontinuity (McCrary 2008). A spike in the number of drug trafficking offenses after August 23, 1996 could suggest judges, police, or lawyers are manipulating the offense date or offense classification to subject more offenders to the SNAP ban. On the other hand, a significant drop in the number of drug trafficking offenses after August 23, 1996 could suggest offenders are decreasing drug trafficking activity once the policy goes into effect or that judges, police, or lawyers are manipulating offense date or offense classification to help offenders avoid the ban. In either case, this type of behavior would confound a causal estimate of the SNAP ban on recidivism. I do not find evidence that the number of drug trafficking offenses changes after August 23, 1996. These results, in Figure A2.2, provide further evidence that the identifying assumption is satisfied.

Although the tests reported in Figure 2.1, Table A2.2, and Figure A2.2 suggest no sorting is happening near the cutoff in Florida, it is worth discussing a few context-specific details that may further ease concerns about sorting. When PRWORA was introduced, it did not include the amendment that banned drug offenders from SNAP benefits—this amendment was introduced by Senator Phil Gramm on July 23, 1996, only a month before President Clinton signed the bill into law (U.S. Congress 1996, S8498). This leaves a very

[13] The break in probability an offender is black before versus after the cutoff is not significant, but it is large in the specification with the ±240 day bandwidth. Including a control for race in the main regression yields similar results in size and significance. Without controlling for race, the coefficient is 0.095. When I control for race, the coefficient is 0.103. In addition, I am testing several different characteristics with several different bandwidths. Importantly, when I combine these characteristics into a composite risk score, I find no break at the cutoff, and when I do a joint significance test of all characteristics, I find no evidence of a change in the characteristics of offenders at the cutoff.

short amount of time for information about the ban to disseminate to offenders, judges, police, prosecutors, or anyone else who could feasibly induce sorting. Even more, as the President had vetoed the previous two welfare reform bills, there was at least some uncertainty over whether or not the bill would become law (Haskins 2006). Finally, although PRWORA as a whole was widely covered by news outlets at the time, the ban on drug felons received little to no publicity.[14]

2.6 Results

2.6.1 Main Results

I begin by estimating the effect of the SNAP ban on any recidivism using the sharp cutoff date of the ban. Since I do not have access to SNAP administrative records, the effects estimated in this paper should not be interpreted as the average or local average treatment effect of SNAP receipt on recidivism. Rather, the results should be interpreted in one of two ways. First, as an intent to treat (ITT) effect, which can then be scaled up by the SNAP take-up rate among former offenders to estimate the local average treatment effect of SNAP receipt. Second, the ban itself may affect recidivism even apart from actual SNAP receipt. If the potential of receiving SNAP has insurance value, the ban may affect decision-making even among offenders who would not receive SNAP. In this case,

[14] Searches for the phrases “food stamps felon”, “food stamps crime” and “welfare felon” in LexisNexis return zero news articles from August 22, 1995 to August 22, 1997. The phrases “food stamps ban” and “food stamps drug” turn up only two articles—one about the PRWORA work requirements and the ban on noncitizens and the other detailing a case of Food Stamps fraud. In addition, a search of the Vanderbilt Television News Archive reveals 12 major news broadcasts over this period about “food stamps.” All of these segments are under 4 minutes long and based on the descriptions, they are broad discussions of the 1996 welfare reform. It does not appear that the ban on felony drug offenders was a particularly salient piece of the welfare overhaul in 1996.

the results should be interpreted as the local average treatment effect of the SNAP ban on recidivism.

The main results are in Table 2.3 below. In Panel A, I show results using the Imbens-Kalyanaraman (IK) optimal bandwidth and in Panel B, I show results using a bandwidth of ± 240 days. I will discuss results in terms of Panel B to make comparisons across analyses easy. Column (1) of Panel B shows the effect of the SNAP ban on any recidivism (ever returning to a Florida state prison). I estimate that the SNAP ban increased any recidivism among drug traffickers by about 9.5 percentage points on average. The baseline recidivism rate for drug traffickers committing their crime in the 240 days prior to the cutoff date is about 16.4 percent. This implies that the SNAP ban increased recidivism among drug traffickers by about 58 percent.

Admittedly, an effect of this magnitude is large and at first blush, might seem unrealistic. First, note that the 9.5 percentage point estimate is only the point estimate. Because the sample size is small, the estimates are noisy and the confidence interval is large. For example, the 90 percent confidence interval for the estimate in column (5) of Table 2.3 is (0.017, 0.172), which implies the SNAP ban increased recidivism among drug traffickers by about 10 percent to 105 percent.[15] Second, even large estimates may be reasonable when we consider that the SNAP benefit is a substantial chunk (about 20 percent) of gross income for men receiving SNAP in Florida.

In addition, SNAP benefits are an important resource for ex-offenders. Recall that

[15] A 10 percent increase in recidivism is reasonable and in line with other papers in this field. Yang (2017a) finds that SNAP bans increase 1-year recidivism rates by about 13 percent. Yang (2017b) finds that a 5 percent increase in real wages due to local labor market opportunities decreases recidivism by about 2.3 percent–extrapolating this based on Table 2.2, a 25 percent increase in real wages due to SNAP receipt would decrease recidivism by 11.5 percent. Finally, several earlier papers found that giving unemployment assistance to released offenders decreased probability of re-arrest by 8 to 27 percent.

approximately 70 percent of the former inmates in the Boston Reentry Study report receiving SNAP benefits even just two months after release (Western et al. 2015). Sugie (2012) also finds that about 70 percent of families in the Fragile Families & Child Well-being Study with a recent paternal incarceration report receiving SNAP in the past year. The Panel Study of Income Dynamics asks respondents in 1995 if they have ever been in the corrections system (jail, prison, youth corrections). Almost 50 percent of respondents who answered yes to that question were in families that reported receiving SNAP at some point from 1995-2013.[16]

Finally, it is easy to assume that former drug traffickers are not reliant on SNAP because drug trafficking is potentially lucrative. However, when these offenders are released from prison, they do not automatically return to drug trafficking. The key idea in this paper is that former drug traffickers choose a number of hours to work in the illegal sector and that access to SNAP informs that choice. I argue that former drug traffickers who are banned from SNAP do choose to work more hours in the illegal sector, and thus, will be more likely to return to prison. In addition, it is not even clear that active drug traffickers earn a substantial income, on average. For example, a person is charged with trafficking heroin in Florida if they sell, manufacture, or distribute 4 grams of heroin. While 4 grams of heroin has a value of approximately $1,000 according to the Drug Enforcement Administration (2015), this does not imply that the trafficker nets a profit of $1,000. Work

[16]Also, a 58 percent increase in recidivism is not far from some others in the literature. Carr and Packham (2017) find that the timing of SNAP receipt alone decreases grocery store theft in Chicago by 32 percent. Di Tella and Schargrodsky (2013) find that electronic monitoring of inmates (relative to imprisonment) reduces rearrest by half of baseline. Hansen (2015) uses a discontinuity in driving under the influence (DUI) punishments and finds that being charged with an "aggravated DUI" reduces reoffending by 27 percent. Finally, Aizer and Doyle (2015) find that incarceration as a juvenile increases likelihood of adult incarceration (by the age of 25) by about 70 percent.

by Levitt and Venkatesh (2000) suggests that even "officers" (the position above "foot soldier" but below "gang leader") in a drug-selling gang earn approximately $1,400 per month (in 2010 dollars). Foot soldiers earn even less at around $200 per month (in 2010 dollars).[17]

In columns (2) and (3), I estimate the effect of the SNAP ban on probability of financially motivated recidivism and probability of non-financially motivated recidivism. I find the effect is completely driven by recidivism for financially motivated crimes. Column (2) of Panel B suggests that the SNAP ban increases financially motivated recidivism by 10 percentage points while column (3) suggests the ban had no detectable effect on non-financially motivated recidivism. The total increase observed in Column (1) was 9.5 percentage points. This implies that 100 percent of the increase in the probability of returning to prison comes from offenders committing crimes that have monetary compensation. Pre-existing differences in the types of crimes drug traffickers returned to prison for cannot account for this result. Drug traffickers who committed their offense in the 240 days prior to the cutoff date were equally likely to return to prison for both financial and non-financial crimes. Finally, the increase in recidivism for financially motivated crimes is significantly different from the change in non-financial crimes at the 5 percent level (p-value=0.0427).

Figure 2.2 and Figures 2.3a-2.3b present visual evidence of the results in Table 2.3.

[17] Levitt and Venkatesh also discuss legal sector employment, noting that around 80 percent of foot soldiers are employed in the legal sector at some point in a given year. However, these are not stable jobs (only 40-50 percent of foot soldiers are employed at any given time) and the jobs tend to be low-wage service-sector work. Levitt and Venkatesh further stress that both foot soldiers and officers report living with family because they cannot afford their own housing. Finally, to the extent that access to SNAP influences how much time (if any) to allocate to illegal work post-release, that decision should be reflected in the probability of recidivism.

The figures show linear polynomials (fitted on the underlying data) overlaid on scatter plots of recidivism outcomes collapsed to 30-day bin averages. In Appendix A, I include Figures A2.4a-A2.4f, which show both quadratic and kernel-weighted, smoothed polynomials versions of Figures 2.2-2.3b. To further demonstrate the robustness of the main results to choice of bandwidth and polynomial, I show the results of local linear, quadratic, and cubic regressions for bandwidths of 30-1080 days in Figures A2.5a-A2.5c. In Tables A2.6-A2.8, I report results from Probit, Logit, and Cox Hazard estimations, all of which are consistent with the main results in Table 2.3.

Since most drug traffickers in my sample never return to prison the data used in the analyses discussed above include many zeroes. To address concerns about over-dispersion, I collapse the data to 15-day bin averages, and redo the main analysis using OLS on the binned data (weighted by the number of observations in each bin). In these regressions, the dependent variable is the average recidivism rate for all offenders in a given 15-day bin. Likewise, the running variable, distance from August 23, 1996, takes on the average value of distance for all offenders in a bin. Binning also facilitates analyzing the data as count data in a Poisson model and as time-series data. I also control for the number of Fridays in each bin. These results are reported in Tables A2.9-A2.12 and Figure A2.6, and are also consistent with the findings in this paper. The evidence here and in Appendix A suggests that the SNAP ban increased the probability of recidivism for drug traffickers.

2.6.2 Heterogeneity Tests

The effect of the SNAP ban may be exacerbated by certain factors. The model in Appendix C predicts that when legal labor market opportunities are more scarce, the banned offenders will be more likely to turn to the illegal labor market. I test this in two ways. First, the effect of the SNAP ban should be smaller when ex-offenders face a tight labor market and increasing legal labor supply becomes more feasible. I interact the state-level unemployment rate at the month of the offender's release with all other variables in equation (2.1), and present the results in Table A2.16 (Bureau of Labor Statistics (BLS) 1996-2016). The effect is not statistically different from zero, but the point estimates imply the ban increases recidivism more for offenders released in poor legal labor markets. Second, evidence suggests that ex-offenders who are black face heightened discrimination in the legal labor market. If the SNAP ban does affect recidivism via work incentives, we should see stronger effects for black offenders. These results are in Table A2.17. Again, the estimates on the interaction between race and the cutoff are all positive, as expected, but they are not statistically different from zero.

I also investigate how the SNAP ban affects timing of re-incarceration. To do this, I estimate the effect of the ban on the probability the offender returns to prison in 0 to 5 years and the effect of the ban on the probability the offender returns to prison in 5 to 10 years. These results, presented in Table A2.18, suggest that the effect of the ban is slightly focused in earlier years rather than later years. Also, in Figure A2.7, I show the effect of the ban on recidivism within 1-year windows. Again, these results show that the increase in recidivism due to the ban is occurring in both earlier years and later years though more so in earlier years. It is difficult to interpret these results since time to re-incarceration is a function of both the time it takes for an ex-offender to re-enter the illegal labor market and the time it takes for an ex-offender to be caught once they re-enter. In addition, SNAP generosity and ABAWD waivers both vary over time.

Finally, I compare the effect of the SNAP ban on the probability an offender recidivates in a month (using month of offense) and county (using county of conviction) when the ABAWD work requirement is waived and the effect of the SNAP ban on the probability an offender recidivates in a month and county when the ABAWD work requirement is in effect (Florida Department of Children and Families (FL DCF) 1996-2016). When the ABAWD work requirement is waived, able-bodied adults without dependents who are not banned from SNAP can receive SNAP benefits even if they are unemployed and not enrolled in employment/training programs. Figure A2.9 displays the geographic variation in county-level ABAWD work requirement waivers for 1996, 1998, 2000, 2004, 2006, and 2008.[18] If the main results are due to SNAP receipt, then the increase in recidivism as a result of the ban should be driven by increased recidivism occurring in months and counties with ABAWD waivers. This is when the disparity in transfer income between the control group (not banned from SNAP) and the treatment group (banned from SNAP) is the greatest. In Table A2.19, I show that the increase in recidivism is concentrated in months and counties when the ABAWD work requirements are waived.[19,20]

[18] I do not show 2002 or years after 2008 because nationwide ABAWD waivers are in place. An animation showing the geographic variation in waivers from January 1996-December 2008 can be found here: https://www.dropbox.com/s/kufg1ieiwtjm0b6/Waivers%20by%20County-Month.gif?dl=0

[19] At a bandwidth of plus-or-minus 240 days from the cutoff date, the effect of the ban on recidivism when the ABAWD requirement is waived is statistically different from the effect on recidivism when the ABAWD requirement is in effect at the 5 percent level (p-value=0.0461) in the local linear model.

[20] I present alternative versions of this test in Tables A2.20 and A2.21.

2.6.3 Placebo Tests and Threats to Validity

Florida modified the Federal SNAP ban to exempt offenders convicted of drug possession or selling, manufacturing, and distributing (SMD) drugs; however state lawmakers did not pass legislation modifying the ban until May 1997 (Government Accountability Office (GAO) 2005).[21] If the results in this paper are driven by endogenous sorting around the cutoff, we should also find effects for offenders committing SMD since all available information as of August 23, 1996 indicated that the ban would apply to those offenders. These results are in Figure 2.4a and Table A2.22. I find no effect for SMD offenders, which further suggests that the effect for drug traffickers is not driven by endogenous sorting at the cutoff. I also estimate the effect of the SNAP ban with a regression discontinuity difference-in-differences design, using SMD offenders as a control group. Using the ± 240 day bandwidth, this strategy yields a coefficient estimate of about 9.5 percentage points.

Figure 2.4b and Table A2.23 display another placebo test examining recidivism for all non-drug offenders around the cutoff date. These offenders were never banned from SNAP as part of the federal policy, and thus their behavior should also be unaffected by the cutoff date. I find no change in recidivism for these offenders. I conduct additional placebo tests using all offenders convicted of a DUI, drug possession, property crime, and violent crime in Table A2.24 and Figures A2.11a-A2.11d. I find no evidence of increased recidivism after the cutoff date for these offenders.

One major concern with regression discontinuity designs that use time as the run-

[21] The sample of people who committed SMD or drug trafficking consists almost entirely of people who were incarcerated for over a year.

ning variable is that the policy cutoff date coincides with a seasonal pattern. If the results in this paper are driven by a general seasonal trend in the relationship between recidivism and date of offense or a trend specific to 1996, the placebo tests in Figures 2.4a-2.4b, Tables A2.22-A2.24, and Figures A2.11a-A2.11d should also recover positive estimates–they do not. However, it is possible that there is spurious seasonality around August 23 that is specific to drug traffickers. To rule out this explanation, I run 16 placebo regressions, one for each August 23rd from 1997-2012.[22] For example, in the 1997 regression, I code the variables $After_{it}$ and $g(DaysFromCutoff_{it})$ as if the cutoff date is August 23, 1997. I do not include years after 2012 since offenders committing crimes in those years have little time to recidivate. I use a bandwidth of ± 180 days in each regression to avoid overlapping observations in the tests. The distribution of coefficients from these regressions is in Figure 2.5. Standard regression discontinuity plots for all years from 1997-2012 are included in Figure A2.12. In addition, I estimate a regression discontinuity difference-in-differences design using all August 23rds from 1996-2012. I exclude August 23, 1998 and August 23, 1999 from this test because two criminal justice policies affecting drug traffickers were introduced in Florida in those years.[23] The results in Table A2.25 provide further evidence that seasonality in the relationship between offense date and recidivism cannot explain the findings in this paper.

[22]Ganong and Jäger (2018) suggest a similar exercise designed to test the significance of the estimated effect using randomization inference. Results from that test are plotted in Figure A2.14.

[23]To determine which years to exclude I refer to the document covering years 1980-2002 here: http://www.dc.state.fl.us/pub/history/index.html. For years after 2002, I search the phrase "'Florida' 'committed on or after' 'YYYY'" where "YYYY" is the year in question. I examine the first page of search results, and if a policy that affects drug traffickers is mentioned, I exclude that year. Through this process, I exclude 1998 and 1999. In October 1998, Florida overhauled their criminal justice system with a new "punishment code" that lowered the requirements necessary to receive a prison sentence. In July 1999, Florida instituted mandatory minimums for drug trafficking offenses.

Placebo tests using different crimes around August 23, 1996 rule out threats to validity that would affect multiple types of crime in 1996. Similarly, placebo tests using drug traffickers around other August 23rds rule out threats to validity that would affect drug traffickers in all years. Still, it is possible that some other event occurred near August 23, 1996 that affected only drug traffickers. While I cannot find any information about other potential treatments in Florida around this time, I also show results of a test designed to detect other significant breaks in my bandwidth. This test, designed by Card, Mas, and Rothstein (2008), detects August 29, 1996 as the true cutoff date. August 29, 1996 is only six days from the policy cutoff date. In fact, the fifteen placebo dates with the highest R-squared are all within 9 days of August 23, 1996, and August 23, 1996 yields the fourth highest R-squared. Dates near September 27, 1996 also return high R-squared. I check again in Florida and at the Federal-level for other polices enacted around September 27, 1996—I do not find any. These placebo results provide further evidence that the SNAP ban causally affects the recidivism outcome of drug traffickers.

I interpret the increase in financially motivated crimes as an increase in the illegal labor supply of ex-offenders. However, a more subtle interpretation is that ex-offenders not subject to the SNAP ban face a bigger deterrent to committing drug trafficking than ex-offenders subject to the ban—those not subject to the ban initially will lose access to SNAP if they commit drug trafficking after they are released since the ban applies to anyone who commits drug trafficking after August 23, 1996. This is an important concern for my analysis since these two interpretations yield different policy implications. If the ban increases the recidivism of banned offenders by pushing them into illegal work, that is a negative consequence that should be factored into policy discussions. If the ban decreases the recidivism of non-banned offenders by deterring them from drug trafficking, that is a positive consequence that should be considered in policy discussions.[24]

Fortunately, the deterrence hypothesis yields a testable implication. If the increase in recidivism is driven by non-banned offenders deterred from future drug trafficking, then the increase should be concentrated in an increase in recidivism for drug trafficking crimes. I find no detectable increase in recidivism due to future drug trafficking. However, I do find statistically significant increases in recidivism for other financially motivated offenses. The results in Figures 2.6a-2.6b and Table A2.26 indicate that banned offenders are 8.9 percentage points more likely to return to prison due to a financial crime that is not drug trafficking and only 1.1 percentage points more likely to return with a drug trafficking offense. Recall that the total effect on financial recidivism is a 10 percentage point increase. This suggests that only 11 percent of the total effect can be explained by the deterrence hypothesis.

2.7 Conclusion

SNAP provides valuable assistance to millions of low-income Americans. However, many ex-felons, a particularly needy and at-risk population, are excluded from SNAP. This paper provides evidence that denying drug offenders SNAP benefits has increased their likelihood of recidivism. Standard econometric tests for breaks in the data as

[24]A similar alternative explanation is that all offenders return to drug gangs upon release and that those gangs allocate their "banned" members to riskier crimes since they have less to lose if they are caught. Being assigned to carry out riskier crimes thus leads to increased recidivism for those subject to the SNAP ban. I also estimate the effect of the SNAP ban on recidivism for theft, a crime that I assume drug gangs are less likely to be in the business of committing (only 23 percent of offenders who have served time for selling, manufacturing or distributing drugs in the data have also served time for a theft charge). I find that offenders subject to the SNAP ban are indeed more likely to return to prison for theft.

well as institutional features of the policy change alleviate concerns about sorting threats to the regression discontinuity identification. Also, it does not appear that the ban was widely publicized in the year prior to August 23, 1996 or in the year following August 23, 1996. This main result speaks to an important policy discussion about state repeals of these bans.

Looking closely at the types of crimes that land these offenders back in prison, I find that the increase in recidivism is driven by crimes that have a monetary motive (property crimes, selling drugs, etc.) rather than crimes like drug possession or violent crimes. This result contributes to a literature on the labor supply effects of transfer programs, and highlights the importance of acknowledging the illegal labor margin when designing policies and programs that affect work incentives.

Using the estimate of the effect of the SNAP ban, I provide a back-of-the-envelope calculation of the cost associated with the increased recidivism. For every offender who recidivates because of the SNAP ban, Florida pays the cost to incarcerate that offender and the citizens of Florida suffer costs of victimization.[25] Using existing estimates of the marginal cost of incarceration and costs of victimization, I derive the cost of banning an extra drug offender. Cost per offender is defined as (Marginal Increase in Probability of Offending due to the Ban)$\times$(Marginal Cost of Year of Incarceration)$\times$(Mean Years Sentenced)$+$(Marginal Increase in Probability of Offending due to the Ban)$\times$(Victim

[25] The "marginal cost" of incarceration is a term used by the Department of Justice defined as "the direct care cost incurred [...] to house an inmate [...] includes the cost of feeding, clothing, and providing medical care for an inmate." This number is significantly lower than the "average cost" of incarceration which takes into account fixed costs, and using it in the cost-benefit analysis leads to a more conservative estimates of the costs. Also, in calculating the societal cost of the ban, I ignore the cost of providing released drug traffickers SNAP benefits. However, if we ignore the private benefit of SNAP to drug traffickers, taxpayers in general do save money by denying SNAP benefits to all drug traffickers.

Cost). More details on this calculation are shown in Appendix D. Assuming the ban increases recidivism by about 9 percentage points (the point estimate from the main results), I find the societal cost of the ban in Florida is about $3,700 per banned offender. With approximately 19,000 banned offenders, this implies the ban has cost Florida over 70 million dollars to date, a number that grows with every new trafficker who resorts to crime to make up for the lost benefits.

Ultimately, analysis of the SNAP ban speaks to prisoner reentry policy in general as well as the work incentives associated with transfer programs. Even more, analysis of the ban contributes to an active policy discussion about the repeal of these bans. In April 2016, Georgia's Governor Nathan Deal signed a law modifying the SNAP ban, joining Texas and Alabama, the two other states that modified the ban in 2016 (Phillips 2016). The SNAP ban continues to affect the day-to-day life of drug felons in 27 states, and it is certainly a relevant and important topic for future research.

2.8 Tables and Figures

Table 2.1: Summary Statistics for Drug Traffickers & Other Offenders in Florida

	October 1, 1995 - October 1, 1997			Full Sample
	All Non-Drug Offenders	Sell/Mfg/Dist Offenders	Drug Trafficking Offenders	Drug Trafficking Offenders
Any Recidivism	0.399	0.564	0.178	0.112
	(0.490)	(0.496)	(0.382)	(0.224)
Financial Recidivism	0.246	0.364	0.113	0.087
	(0.431)	(0.481)	(0.317)	(0.195)
Non-Financial Recidivism	0.153	0.200	0.065	0.024
	(0.360)	(0.400)	(0.246)	(0.103)
Days Until Recidivism	1,330.189	1,204.634	1,615.329	1,075.090
	(1,237.552)	(1,187.955)	(1,269.476)	(899.813)
Black	0.455	0.850	0.486	0.377
	(0.498)	(0.357)	(0.500)	(0.485)
Age at Intake	30.952	31.031	33.181	33.910
	(10.114)	(9.155)	(10.226)	(10.164)
Time Sentenced (in Years)	4.438	3.006	5.163	4.116
	(4.040)	(2.649)	(3.563)	(5.159)
Observations	22,893	6,002	1,435	18,656

Notes: The first four rows present recidivism statistics: the fraction of offenders in each group who recidivate, recidivate with a financially motivated crime, recidivate with a non-financially motivated crime, and finally, the days until an offender recidivates (conditional on recidivating). Financially motivated crimes are: property crimes (excluding property damage crimes such as vandalism), selling/manufacturing/distributing drugs, drug trafficking, fraud, forgery, racketeering, prostitution, counterfeiting, and crimes containing a "$", "sale", or "sell" in the charge description. Non-financially motivate crimes are defined as all crimes that are not categorized as financially motivated. Financially motivated recidivism is thus defined as recidivism that involves a financially motivated crime whereas non-financially motivated recidivism is defined as recidivism that does not involve any financially motivated crime. The last three rows show the fraction of offenders who are black, the average age at intake, and the average sentence handed down by the court. Sell/mfg/dist offenders are those offenders convicted of selling/manufacturing/distributing drugs. Sell/mfg/dist is a separate offense from drug trafficking and those offenders were not ultimately included in the SNAP ban in Florida. An offender is tagged as a drug trafficking offender if they are convicted of a drug trafficking offense. An offender is tagged as a non-drug offender if they are not convicted of a drug crime. An offender is tagged as an SMD offender if they are convicted of SMD, but are not convicted of a drug trafficking offense. In addition, when calculating the summary statistics for all drug trafficking offenders, I collapse to the offender ID level since some offenders will have more than one stay for drug trafficking in this time period.

Table 2.2: Summary Statistics on Male SNAP Population in Florida

	No Nationwide ABAWD Waiver		Nationwide ABAWD Waiver	
	Single Male	Male with Family	Single Male	Male with Family
Fraction Black	0.310	0.168	0.292	0.141
	(0.463)	(0.374)	(0.455)	(0.348)
Age	45.280	39.932	43.077	40.994
	(11.502)	(11.969)	(12.693)	(11.644)
Fraction Unemployed	0.916	0.602	0.914	0.617
	(0.278)	(0.490)	(0.281)	(0.486)
SNAP Benefit (in 2010 $)	85.50	206.28	150.41	324.48
	(47.97)	(138.93)	(69.02)	(222.99)
Observations	1,587	1,656	1,962	1,188
Benefit as % of Gross Income	15.703	25.818	18.124	29.326
	(16.700)	(21.135)	(16.720)	(22.723)
Observations	1,027	1,347	924	968

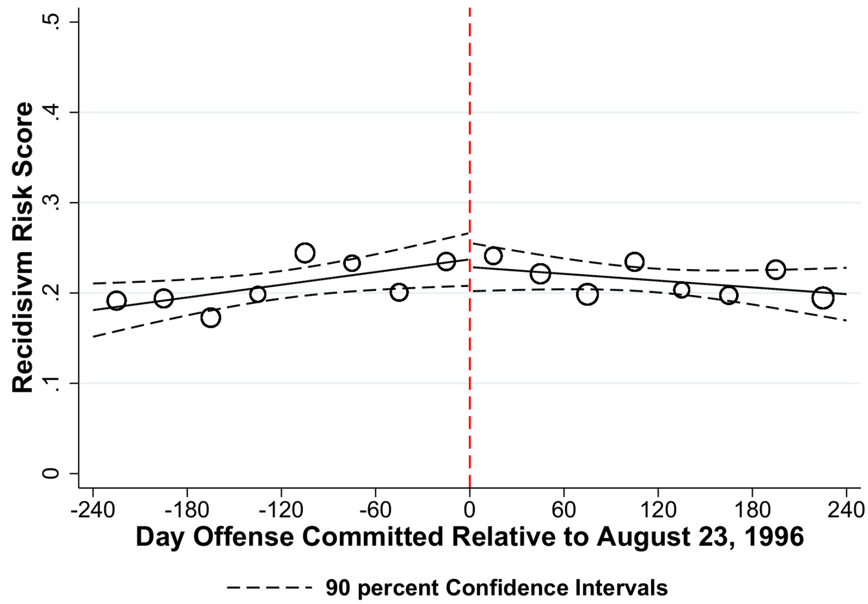
Notes: Summary statistics above are derived from the Mathematica Policy Research SNAP QC file from 1996-2014, which provide data on a sample of the SNAP population in each state. Mathematica Policy Research constructs the SNAP QC files to be representative at the state-level. I limit the sample to males aged 18-65 and listed as the primary or secondary recipient of the SNAP benefits. In calculating the benefit as a percentage of gross income, I remove zeroes in gross income and benefit-income ratios above one. In columns (1) and (2), I provide statistics for all years from 1996-2014 without nationwide ABAWD work requirement waivers (1996-2000, 2004-2008). In columns (3) and (4), I provide statistics for all years from 1996-2014 with nationwide ABAWD work requirement waivers (2001-2003, 2009-2014). The ABAWD work requirement states that able-bodied adults without dependents are limited to only 3 months of SNAP receipt every 3 years unless they: (1) work 20 or more hours per week, (2) participate in an employment and training program, or (3) participate in a workfare program (USDA, 2016b).

Table 2.3: Main Results, Effect of the SNAP Ban on Recidivism

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0803	0.1043***	-0.0100
	(0.0493)	(0.0398)	(0.0280)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	$\pm$212	$\pm$242	$\pm$254
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of $\pm$240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0950**	0.1003**	-0.0053
	(0.0467)	(0.0404)	(0.0286)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	$\pm$240	$\pm$240	$\pm$240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

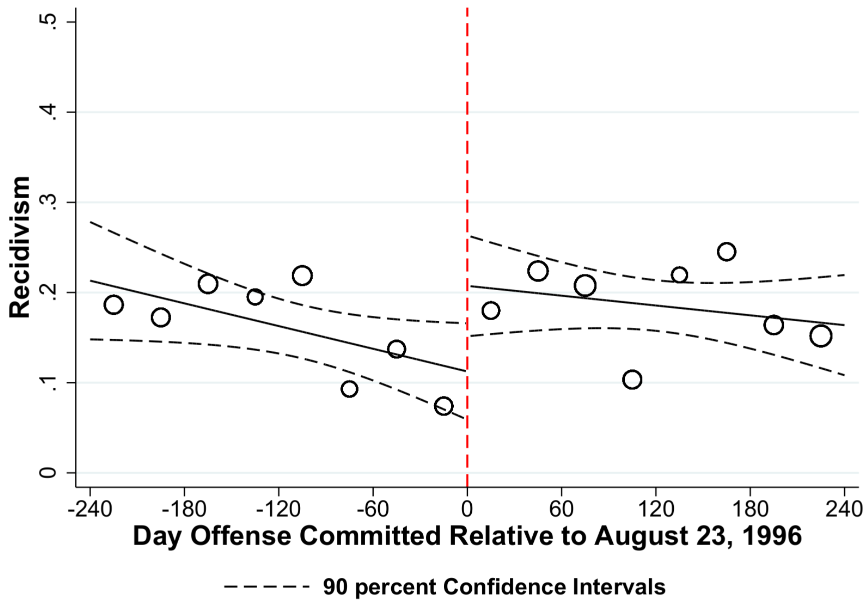
Notes: Standard errors clustered at the day of offense in parentheses. Number of days the drug trafficking offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender ever returns to a Florida prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table 2.1 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. In Panel A, the Imbens-Kalyanaraman optimal bandwidth is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of $\pm$240 days (or $\pm$8 months) is used with polynomial of degree one and a uniform kernel. Results are robust to these choices (see online Appendix Tables A2.13-A2.15). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 2.1: Smoothness Through Cutoff in Offender's Risk of Recidivism



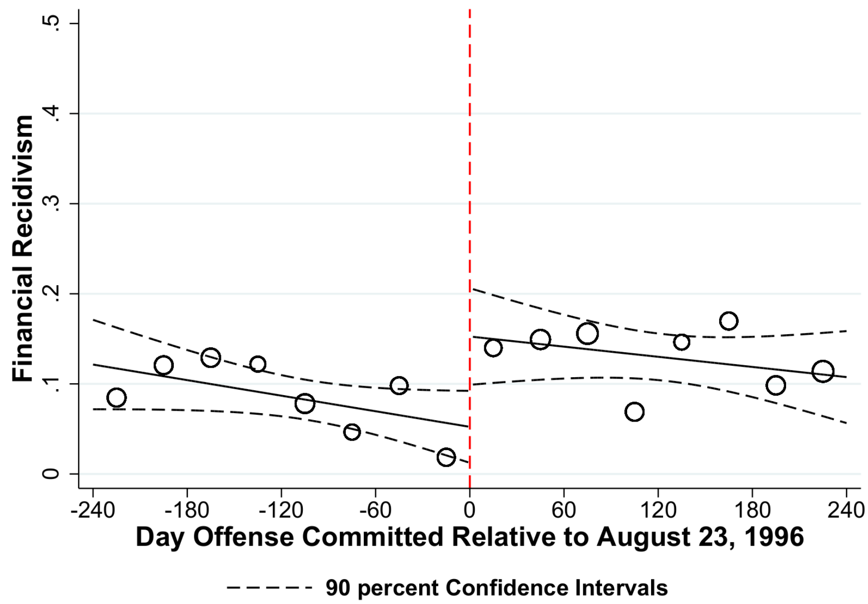
Notes: Recidivism risk score is calculated by: (1) estimating the relationship between offender characteristics and recidivism using a sample of pre-ban drug traffickers who are not included in the Imbens-Kalyanaraman (IK) optimal bandwidth and (2) applying those estimates to drug traffickers in the sample. The characteristics used to create this measure of offender risk are: age, age-squared, total years sentenced, total number of prior offenses, total number of concurrent offenses, sex, race, and type of drug trafficked. The figure above (and the following RD plots more generally) displays the lines from two local linear regressions, estimated separately on each side of the cutoff using the offense-level micro data. I also overlay a scatter plot of 30-day bin averages of the dependent variable weighted by the number of offenses in each 30-day bin. The dependent variable in this figure is offender risk score, and the figure shows that offender risk of recidivism (an index of several offender characteristics) is smooth through the cutoff date. Finally, the running variable in this figure (and the following RD plots) is the number of days between the offender's offense date and August 23, 1996 (the cutoff date that determines the offender's ban status). The running variable is centered at zero such that offenders committing an offense before August 23, 1996 have a negative distance from the cutoff date and offenders committing an offense after August 23, 1996 have a positive distance from the cutoff date.

Figure 2.2: Effect of SNAP Ban on Any Recidivism



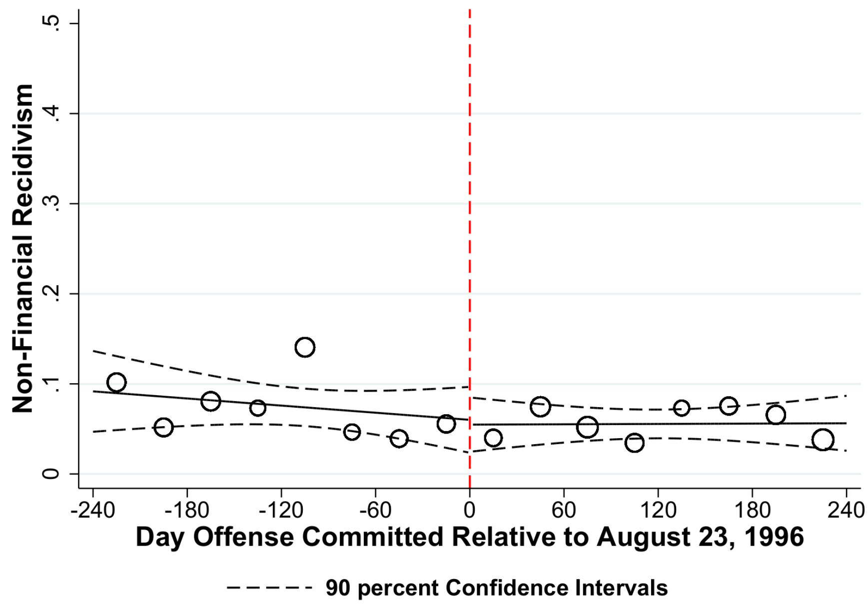
Notes: See Figure 2.1 notes for a general description of the creation of the RD plots, including information about bin size, estimation of the fit lines, and definition of the running variable. In this figure, the dependent variable is recidivism, defined as whether an offender ever returns to a Florida prison after release.

Figure 2.3a: Effect of SNAP Ban on Financial Recidivism



Notes: See Figure 2.1 notes for a general description of the creation of the RD plots, including information about bin size, estimation of the fit lines, and definition of the running variable. In this figure, the dependent variable is financial recidivism. See Table 2.1 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

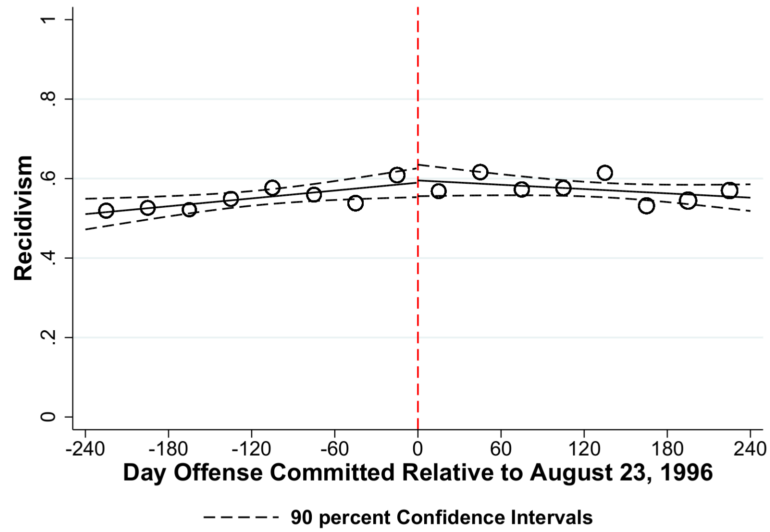
Figure 2.3b: Effect of SNAP Ban on Non-Financial Recidivism



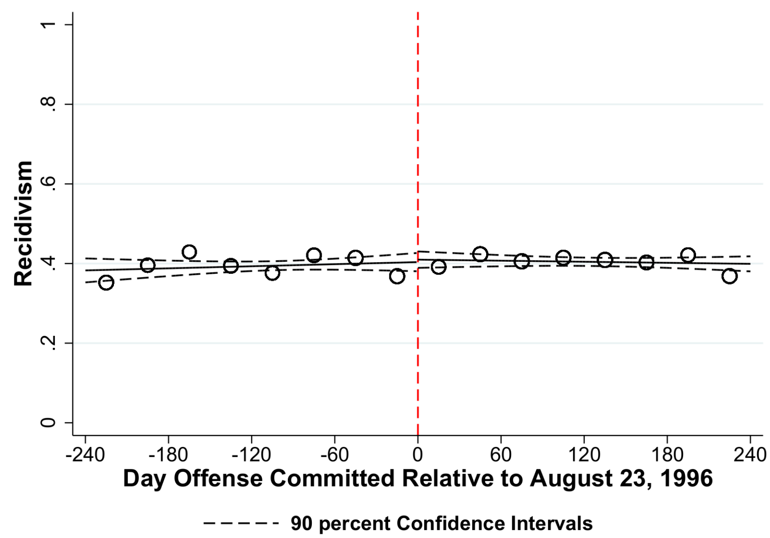
Notes: See Figure 2.1 notes for a general description of the creation of the RD plots, including information about bin size, estimation of the fit lines, and definition of the running variable. In this figure, the dependent variable is non-financial recidivism. See Table 2.1 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

Figure 2.4: Effect of SNAP Ban on Any Recidivism for Non-Banned (Placebo) Offenses

(a) Selling/Manufacturing/Distributing Drug Offenders

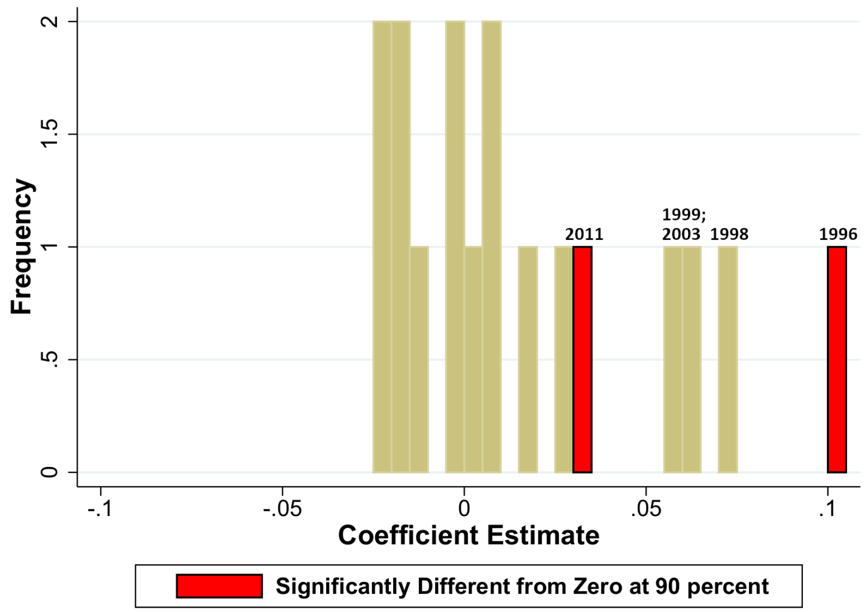


(b) Non-Drug Offenders



Notes: See Figure 2.1 notes for a general description of the creation of the RD plots, including information about bin size, estimation of the fit lines, and definition of the running variable. In both figures, the dependent variable is recidivism, defined as whether the offender ever returns to a Florida prison or not. Figure 2.4a displays this relationship for offenders convicted of committing the crime of selling/manufacturing/distributing (SMD) drugs. These offenders were exempted from the SNAP ban by the Florida legislature in May 1997. Thus, if the main results are driven by endogenous sorting around the cutoff, we should also observe an effect for SMD offenders. Figure 2.4b displays this relationship for offenders convicted of committing any non-drug crime. These offenders were never subject to the SNAP ban, and thus, their likelihood of recidivism should be smooth through the cutoff date. Both placebo tests show no change in recidivism for offenders committing their offense after the cutoff date.

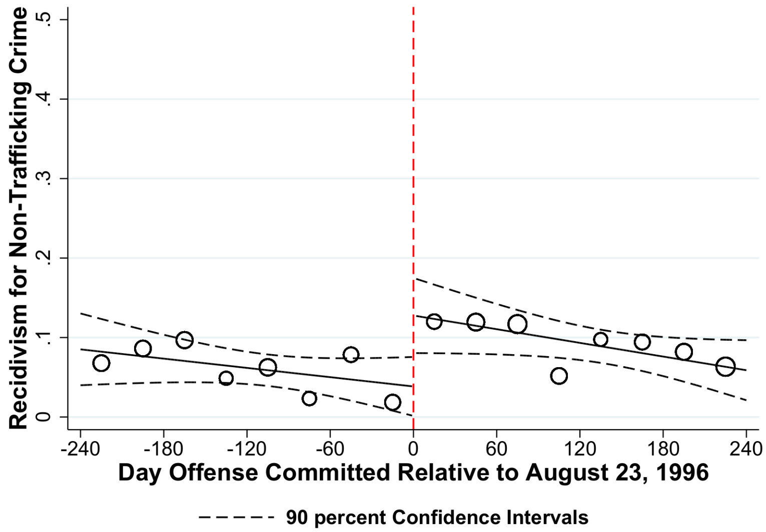
Figure 2.5: Distribution of Coefficients from Placebo Tests at August 23, 1997-2012



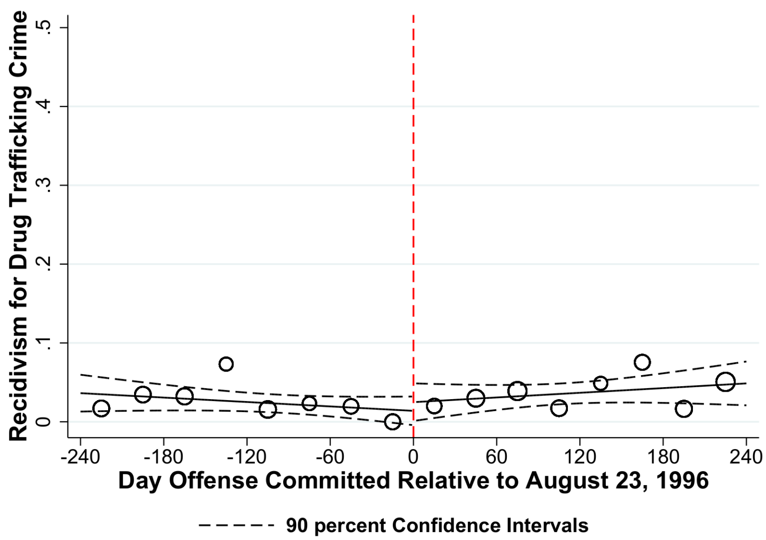
Notes: The figure above displays a histogram of the coefficient estimates from 16 placebo regressions (one at each August 23rd from 1997-2012) and the coefficient estimate from the main result (at August 23rd, 1996). The dependent variable in these placebo tests is recidivism, whether the offender ever returns to a Florida prison or not. In all regressions, I use a bandwidth of ± 180 days to avoid overlapping observations across tests. Only one estimate from the 16 placebo regressions is statistically different from zero, it is from the year 2011 and it is much lower in magnitude than the main result. In addition, there are three estimates that are larger than the 2011 placebo estimate. These correspond to years 1998, 1999, and 2003. In October 1998, Florida overhauled their criminal justice system with a new "punishment code" that lowered the requirements necessary to receive a prison sentence. In July 1999, Florida instituted mandatory minimums for drug trafficking offenses.

Figure 2.6: Effect of SNAP Ban on Type of Recidivism

(a) Recidivism due to Non-Trafficking Crimes



(b) Recidivism due to Drug Trafficking



Notes: See Figure 2.1 notes for a general description of the creation of the RD plots, including information about bin size, estimation of the fit lines, and definition of the running variable. In Figure 2.6a, the dependent variable is financial recidivism excluding recidivism for drug trafficking crimes. In Figue 2.6b, the dependent variable is recidivism for drug trafficking crimes only. See Table 2.1 for a definition of financially motivated crime and the associated recidivism measure. If the SNAP ban causes an increase in recidivism by reducing the drug trafficking activity of non-banned offenders (deterred by the threat of the ban after they are released), then recidivism for drug trafficking should be higher for banned offenders than non-banned offenders. Instead, recidivism for drug trafficking is similar for both banned and non-banned offenders while recidivism for non-trafficking crimes is higher for banned offenders. These figures imply that the main results are driven by increased criminal activity of banned offenders.

2.9 Appendix A. Additional Tables and Figure

Table A2.1: Additional Summary Statistics for Offenders in Florida

	October 1, 1995 - October 1, 1997			Full Sample
	All Non-Drug Offenders	Sell/Mfg/Dist Offenders	Drug Trafficking Offenders	Drug Trafficking Offenders
Recidivism -	0.216	0.288	0.102	0.072
ABAWD Waiver	(0.412)	(0.453)	(0.303)	(0.178)
Recidivism -	0.183	0.276	0.075	0.039
No ABAWD Waiver	(0.386)	(0.447)	(0.264)	(0.131)
# of Recidivism	0.994	1.635	0.413	0.502
Offenses	(1.715)	(2.133)	(1.146)	(1.232)
Trafficking Cocaine	-	-	0.789	0.410
	-	-	(0.408)	(0.468)
# of Prior Offenses	0.578	1.007	0.228	0.298
	(1.052)	(1.291)	(0.586)	(0.663)
# of Concurrent	1.578	2.134	1.502	1.629
Offenses	(0.929)	(1.080)	(0.894)	(0.871)
Male	0.928	0.917	0.885	0.868
	(0.258)	(0.276)	(0.319)	(0.339)
Observations	22,893	6,002	1,435	18,656

Notes: The first three rows present recidivism statistics: the fraction of offenders in each group who recidivate in a time and place (based on county of conviction) where ABAWD work requirements are waived, the fraction who recidivate in a time and place where the work requirements are not waived, and the number of offenses committed after prison stay j but before prison stay $j+1$ (coded as zero if there is no stay $j+1$ i.e. the offender does not recidivate). For the ABAWD recidivism measures, conviction county and date of earliest offense after stay j is used. The last four rows show: the fraction of offenders who were convicted of trafficking cocaine, the average number of prior offenses, the average number of concurrent offenses, and the fraction of offenders who are male.

Table A2.2: Evidence RD Identifying Assumption Holds: No Differences in Observable Characteristics

Characteristic:	# Other Offenses	Years Sentenced	Black	Age	# Prior Offenses	Male	Trafficking Cocaine	Risk Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Imbens Kalyanaraman Optimal Bandwidth								
Offense Committed After Aug. 23, 1996 (Banned)	0.0017	0.5552	-0.0563	-0.3157	-0.0988	0.0384	-0.0213	-0.0218
	(0.1276)	(0.3782)	(0.0692)	(1.3460)	(0.0764)	(0.0393)	(0.0527)	(0.0197)
Control Group Mean	1.5046	5.3285	0.4818	33.5553	0.2478	0.8631	0.8007	0.1952
Observations	944	1580	1281	1067	2290	1275	1317	1391
Bandwidth (in Days)	±246	±465	±338	±281	±802	±334	±349	±380
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1	1	1	1
Panel B. Consistent Bandwidth of ±240 Days								
Offense Committed After Aug. 23, 1996 (Banned)	-0.0108	0.3198	-0.1096	0.1294	-0.0072	0.0196	-0.0342	-0.0085
	(0.1305)	(0.4950)	(0.0830)	(1.4691)	(0.1046)	(0.0461)	(0.0626)	(0.0240)
Control Group Mean	1.5046	5.1615	0.4861	33.4352	0.2616	0.8611	0.8009	0.2083
Observations	918	918	918	918	918	918	918	918
Bandwidth (in Days)	±240	±240	±240	±240	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. Number of days the drug trafficking offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). In Panel A, the Imbens-Kalyanaraman optimal bandwidth is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. Since the data begins with offenses committed on October 1, 1995, the bandwidth is asymmetric for analyses where the bandwidth exceeds ±327 days. Column (1) shows no break in the number of other offenses for which the offender is currently being charged. Column (2) shows no break in the total number of years sentenced. Column (3) shows no break in racial composition and column (4) shows no break in age composition. Column (5) shows no break in the number of prior offenses the offender has been incarcerated in FL prison for. Column (6) shows no break in sex composition. Column (7) shows no break in the probability of trafficking cocaine. Risk of recidivism in Column (8) is calculated from a logistic regression of recidivism on all variables in columns (1)-(7) and age-squared for drug traffickers not subject to the ban and not in the IK sample window. The risk score is then predicted by applying the coefficients from that regression to the sample of drug offenders in my analysis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.3: Effect of the SNAP Ban on Time-Constrained Recidivism Rates

Outcome:	Recidivism within 10 Years	Financial Recidivism within 10 Years	Recidivism within 8 Years	Financial Recidivism within 8 Years	Recidivism within 5 Years	Financial Recidivism within 5 Years
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Imbens Kalyanaraman Optimal Bandwidth						
Offense Committed After Aug. 23, 1996 (Banned)	0.1099**	0.0965**	0.0950**	0.1026***	0.0436	0.0748**
	(0.0511)	(0.0403)	(0.0452)	(0.0336)	(0.0372)	(0.0301)
Control Group Mean	0.1652	0.0846	0.1393	0.0671	0.1046	0.0552
Observations	684	818	840	922	1028	972
Bandwidth (in Days)	±209	±242	±235	±256	±277	±259
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1	1
Panel B. Consistent Bandwidth of ±240 Days						
Offense Committed After Aug. 23, 1996 (Banned)	0.0894*	0.0914**	0.0909**	0.0998***	0.0581	0.0685**
	(0.0488)	(0.0411)	(0.0446)	(0.0357)	(0.0402)	(0.0316)
Control Group Mean	0.1649	0.0851	0.1386	0.0693	0.1071	0.0548
Observations	803	803	854	854	893	893
Bandwidth (in Days)	±240	±240	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Columns 1 and 2 estimate the effect of being banned from SNAP on whether or not the offender returns to prison within 10 years of being released and whether or not they return due to a financial crime within 10 years. Columns 3 and 4 estimate the effect on recidivism and financially motivated recidivism within 8 years of release. Finally, columns 5 and 6 estimate the effect on recidivism and financially motivated recidivism within 5 years of release. Financially motivated crimes are: property crimes (excluding property damage crimes such as vandalism), selling/manufacturing/distributing drugs, drug trafficking, fraud, forgery, racketeering, prostitution, counterfeiting, and crimes containing a "$", "sale", or "sell" in the charge description. Non-financially motivate crimes are defined as all crimes that are not categorized as financially motivated. Financially motivated recidivism is thus defined as recidivism that involves a financially motivated crime whereas non-financially motivated recidivism is defined as recidivism that does not involve any financially motivated crime. Time until recidivism is defined as the difference between the offender's release date for prison stay j and the next offense date before prison stay $j+1$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.4: Effect of the SNAP Ban on Recidivism Outcomes, Hispanic Individuals Included

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0787*	0.0922**	-0.0049
	(0.0456)	(0.0369)	(0.0258)
Control Group Mean	0.1525	0.0865	0.0704
Observations	867	1023	1067
Bandwidth (in Days)	±216	±248	±258
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0873**	0.0882**	-0.0010
	(0.0435)	(0.0380)	(0.0266)
Control Group Mean	0.1591	0.0882	0.0710
Observations	987	987	987
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2 for general notes about the RD estimation, including information about bandwidths and the running variable. Hispanic offenders are included in the sample for this analysis. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.5: Effect of the SNAP Ban on Recidivism Outcomes, Controls for Offender Characteristics & Day-of-Week Effects

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0922*	0.1064***	-0.0037
	(0.0492)	(0.0389)	(0.0289)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	±212	±243	±255
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.1053**	0.1043***	0.0010
	(0.0461)	(0.0395)	(0.0297)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. These analyses include controls for race, age, sex, type of trafficking, total years sentenced, number of prior offenses, number of concurrent offenses, and offense day-of-week fixed effects. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.6: Effect of the SNAP Ban on Recidivism Outcomes, Logit Model

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0776*	0.1022***	-0.0094
	(0.0460)	(0.0386)	(0.0261)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	±212	±243	±255
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0924**	0.0975**	-0.0055
	(0.0443)	(0.0389)	(0.0269)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. This table shows the main specifications estimated with logistic regressions. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.7: Effect of the SNAP Ban on Recidivism Outcomes, Probit Model

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0793*	0.1034***	-0.0092
	(0.0466)	(0.0388)	(0.0265)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	±212	±243	±255
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0940**	0.0986**	-0.0052
	(0.0449)	(0.0389)	(0.0273)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. This table shows the main specifications estimated with probit regressions. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.8: Effect of the SNAP Ban on Recidivism Outcomes, Hazard Model

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.5558	1.0959**	-0.1270
	(0.3415)	(0.4287)	(0.4681)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	±214	±271	±233
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.6419**	1.0710**	-0.0453
	(0.3200)	(0.4368)	(0.4804)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. This analysis employs a Cox survival model in which offenders enter the sample when they are released and exit when they return to prison. The coefficients are approximate semi-elasticities. For example, the coefficient in column (1) of Panel B indicates that the ban increased recidivism by approximately 60% from baseline. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.9: Results from Regression on 15-day Bin Averages of Recidivism

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0975*	0.1031**	-0.0092
	(0.0555)	(0.0463)	(0.0268)
Control Group Mean	0.1609	0.0880	0.0761
Observations	28	32	34
Bandwidth (in Days)	±212	±242	±254
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.1000*	0.1031**	-0.0031
	(0.0491)	(0.0463)	(0.0272)
Control Group Mean	0.1644	0.0880	0.0764
Observations	32	32	32
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at each 15-day bin in parentheses. In this analysis, the outcome variable is the average recidivism rate within each 15-day bin. Also, the average number of days the drug trafficking offenses in a bin were committed before or after Aug. 23, 1996 is the running variable (centered at zero). All models also control for the number of Fridays in each bin. Also, each regression is weighted by the number of offenders in each bin. In Panel A, the Imbens-Kalyanaraman optimal bandwidth (chosen from the micro data pre-aggregation) is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. Column 1 estimates the effect of the SNAP ban on recidivism rates. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

Table A2.10: Results from Regression on 15-day Bin Counts of Recidivism, Poisson Model

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After	0.6126*	1.0435*	-0.1639
Aug. 23, 1996 (Banned)	(0.3538)	(0.5407)	(0.3821)
Observations	28	32	34
Bandwidth (in Days)	±212	±242	±254
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After	0.6085**	1.0435*	-0.0713
Aug. 23, 1996 (Banned)	(0.3063)	(0.5407)	(0.3930)
Observations	32	32	32
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at each 15-day bin in parentheses. In this analysis, the outcome variable is the average recidivism rate within each 15-day bin. Also, the average number of days the drug trafficking offenses in a bin were committed before or after Aug. 23, 1996 is the running variable (centered at zero). All models also control for the number of Fridays in each bin. Also, each regression is weighted by the number of offenders in each bin. In Panel A, the Imbens-Kalyanaraman optimal bandwidth (chosen from the micro data pre-aggregation) is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. The coefficients are approximate semi-elasticities. For example, the coefficient in column (1) of Panel B indicates that the ban increased recidivism by approximately 60% from baseline. Column 1 estimates the effect of the SNAP ban on recidivism rates. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.11: Results from Time-Series Analysis of 15-day Bin Averages of Recidivism

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After	0.1131**	0.1191***	-0.0090
Aug. 23, 1996 (Banned)	(0.0504)	(0.0409)	(0.0273)
Control Group Mean	0.1609	0.0880	0.0761
Observations	28	32	34
Bandwidth (in Days)	±212	±242	±254
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After	0.1133**	0.1191***	-0.0031
Aug. 23, 1996 (Banned)	(0.0441)	(0.0409)	(0.0276)
Control Group Mean	0.1644	0.0880	0.0764
Observations	32	32	32
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at each 15-day bin in parentheses. Each regression includes one lag of the dependent variable (number of lags chosen based on model with highest AIC). In this analysis, the outcome variable is the average recidivism rate within each 15-day bin. Also, the average number of days the drug trafficking offenses in a bin were committed before or after Aug. 23, 1996 is the running variable (centered at zero). All models also control for the number of Fridays in each bin. In Panel A, the Imbens-Kalyanaraman optimal bandwidth (chosen from the micro data pre-aggregation) is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. The coefficients are approximate semi-elasticities. For example, the coefficient in column (1) of Panel B indicates that the ban increased recidivism by approximately 60% from baseline. Column 1 estimates the effect of the SNAP ban on recidivism rates. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.12: Results from Time-Series Analysis of 15-day Bin Counts of Recidivism, Poisson Model

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.8792**	1.1973**	-0.1017
	(0.3898)	(0.4819)	(0.4455)
Observations	28	32	34
Bandwidth (in Days)	±212	±242	±254
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.7767**	1.1973**	0.0025
	(0.3237)	(0.4819)	(0.4623)
Observations	32	32	32
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at each 15-day bin in parentheses. Each regression includes one lag of the dependent variable (number of lags chosen based on model with highest AIC). The Stata command **arpois** is used to estimate this time-series Poisson model as illustrated in Schwartz et al. (1996). In this analysis, the outcome variable is the average recidivism rate within each 15-day bin. Also, the average number of days the drug trafficking offenses in a bin were committed before or after Aug. 23, 1996 is the running variable (centered at zero). All models also control for the number of Fridays in each bin. In Panel A, the Imbens-Kalyanaraman optimal bandwidth (chosen from the micro data pre-aggregation) is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. The coefficients are approximate semi-elasticities. For example, the coefficient in column (1) of Panel B indicates that the ban increased recidivism by approximately 60% from baseline. Column 1 estimates the effect of the SNAP ban on recidivism rates. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. As part of the time-series analysis, I conduct a Wald test for a known structural break at Aug. 23, 1996 and I reject the null that there is no break in the data. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.13: Effect of the SNAP Ban Robust to Alternative Optimal Bandwidths

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Calonico, Cattaneo, Titiunik (CCT) Bandwidth			
Offense Committed After	0.1454**	0.1458***	0.0462
Aug. 23, 1996 (Banned)	(0.0604)	(0.0539)	(0.0422)
Control Group Mean	0.1477	0.0605	0.0802
Observations	520	471	423
Bandwidth (in Days)	±139	±126	±111
Panel B. Half the Imbens, Kalyanaraman (IK) Bandwidth			
Offense Committed After	0.1678**	0.1454***	0.0281
Aug. 23, 1996 (Banned)	(0.0694)	(0.0545)	(0.0377)
Control Group Mean	0.1348	0.0613	0.0783
Observations	405	465	475
Bandwidth (in Days)	±106	±121	±127
Panel C. Ludwig, Miller Cross-Validation (CV) Bandwidth			
Offense Committed After	0.0616	0.0813**	-0.0196
Aug. 23, 1996 (Banned)	(0.0407)	(0.0341)	(0.0256)
Control Group Mean	0.1617	0.0887	0.0730
Observations	1252	1252	1252
Bandwidth (in Days)	±325	±325	±325

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about the running variable. In Panel A, the CCT optimal bandwidth is used with polynomial of degree one and a uniform kernel. In Panel B, the IK optimal bandwidth multiplied by one-half is used with polynomial of degree one and a uniform kernel. In Panel C, the CV optimal bandwidth is used with a polynomial of degree one and uniform kernel. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.14: Effect of the SNAP Ban Robust to Alternative Polynomials

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Imbens Kalyanaraman Optimal Bandwidth						
Offense Committed After Aug. 23, 1996 (Banned)	0.1249***	0.1379***	0.0032	0.1523**	0.1141*	0.0022
	(0.0483)	(0.0459)	(0.0320)	(0.0681)	(0.0674)	(0.0488)
Control Group Mean	0.1612	0.0884	0.0728	0.1612	0.0884	0.0728
Observations	2549	1549	1509	1813	1280	1259
Bandwidth (in Days)	±938	±451	±433	±583	±336	±326
Degree of Polynomial in Days from Aug. 23, 1996	2	2	2	3	3	3
Panel B. Consistent Bandwidth of ±240 Days						
Offense Committed After Aug. 23, 1996 (Banned)	0.1344*	0.1420**	-0.0076	0.1461	0.0971	0.0490
	(0.0703)	(0.0617)	(0.0414)	(0.0896)	(0.0784)	(0.0610)
Control Group Mean	0.1644	0.0880	0.0764	0.1644	0.0880	0.0764
Observations	918	918	918	916	916	916
Bandwidth (in Days)	±240	±240	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	2	2	2	3	3	3

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about the running variable. In Panel A, the Imbens-Kalyanaraman optimal bandwidth is used with polynomials of degree two and three and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomials of degree two (columns 1-3) and three (columns 4-6) and a uniform kernel. Columns 1 & 4 estimate the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Columns 2 & 5 and Columns 3 & 6 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.15: Effect of the SNAP Ban Robust to Alternative Kernels

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Imbens Kalyanaraman Optimal Bandwidth						
Offense Committed After	0.1061**	0.1069***	-0.0093	0.1046**	0.1077***	-0.0109
Aug. 23, 1996 (Banned)	(0.0483)	(0.0369)	(0.0285)	(0.0486)	(0.0372)	(0.0290)
Control Group Mean	0.1626	0.0904	0.0732	0.1614	0.0879	0.0733
Observations	1042	1201	1250	967	1109	1180
Bandwidth (in Days)	±270	±309	±324	±251	±287	±301
Kernel	Triangle	Triangle	Triangle	Epanechnikov	Epanechnikov	Epanechnikov
Panel B. Consistent Bandwidth of ±240 Days						
Offense Committed After	0.1108**	0.1164***	-0.0056	0.1064**	0.1156***	-0.0092
Aug. 23, 1996 (Banned)	(0.0513)	(0.0420)	(0.0324)	(0.0498)	(0.0408)	(0.0316)
Control Group Mean	0.1644	0.0880	0.0764	0.1644	0.0880	0.0764
Observations	918	918	918	918	918	918
Bandwidth (in Days)	±240	±240	±240	±240	±240	±240
Kernel	Triangle	Triangle	Triangle	Epanechnikov	Epanechnikov	Epanechnikov

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about the running variable. In Panel A, the Imbens-Kalyanaraman optimal bandwidth is used with polynomial of degree one and two kernels: (1) triangle (columns 1-3) and (2) Epanechnikov (columns 4-6). In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and two kernels: (1) triangle and (2) Epanechnikov. Columns 1 & 4 estimate the effect of being banned from SNAP after release on whether or not the offender returns to prison after being released. Columns 2 & 5 and Columns 3 & 6 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.16: Effect of SNAP Ban on Offenders Released During High Unemployment Months

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After	0.0442	0.0413	0.0509
Aug. 23, 1996 (Banned)	(0.1100)	(0.0839)	(0.0626)
Unemployment Rate (UR)	-0.0189*	-0.0161**	0.0002
	(0.0113)	(0.0066)	(0.0084)
UR X Banned	0.0070	0.0135	-0.0132
	(0.0198)	(0.0149)	(0.0102)
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	212	242	254
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After	0.0831	0.0346	0.0486
Aug. 23, 1996 (Banned)	(0.1019)	(0.0849)	(0.0612)
Unemployment Rate (UR)	-0.0188*	-0.0157**	-0.0031
	(0.0101)	(0.0066)	(0.0079)
UR X Banned	0.0026	0.0142	-0.0116
	(0.0179)	(0.0151)	(0.0099)
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates heterogeneity in the effect of being banned from SNAP on whether or not the offender returns to prison after being released by labor market conditions upon release. Column 2 and Column 3 estimate this heterogeneity in the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. Unemployment rate is the state-level unemployment rate in the month of the offender's release. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.17: Effect of SNAP Ban on Black Offenders

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After	0.0415	0.0694	-0.0299
Aug. 23, 1996 (Banned)	(0.0649)	(0.0474)	(0.0388)
Black	0.0604	0.0159	0.0327
	(0.0648)	(0.0408)	(0.0471)
Black X Banned	0.0987	0.0799	0.0460
	(0.1051)	(0.0797)	(0.0617)
Combined Effect:	0.1402	0.1492	0.0161
Banned+(Black X Banned)	0.0773	0.0640	0.0445
Control Group Mean	0.1587	0.0874	0.0764
Observations	791	936	980
Bandwidth (in Days)	212	242	254
Panel B. Consistent Bandwidth of $\pm$240 Days			
Offense Committed After	0.0294	0.0646	-0.0351
Aug. 23, 1996 (Banned)	(0.0602)	(0.0487)	(0.0386)
Black	0.0313	0.0136	0.0177
	(0.0630)	(0.0414)	(0.0483)
Black X Banned	0.1497	0.0846	0.0651
	(0.0976)	(0.0813)	(0.0625)
Combined Effect:	0.1791	0.1491	0.0300
Banned+(Black X Banned)	0.0734	0.0647	0.0461
Control Group Mean	0.1644	0.0880	0.0764
Observations	918	918	918
Bandwidth (in Days)	$\pm$240	$\pm$240	$\pm$240

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates heterogeneity by race in the effect of being banned from SNAP on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate heterogeneity by race on the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. The row "Combined Effect: Banned+(Black X Banned)" is the linear combination of the coefficients on "Banned" and "Black X Banned" and represents the total effect of the ban on black offenders. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.18: Effect of SNAP Ban on Timing of Re-Incarceration

Outcome:	Recidivism within 0-5 Years	Financial Recidivism within 0-5 Years	Recidivism within 5-10 Years	Financial Recidivism within 5-10 Years
	(1)	(2)	(3)	(4)
Panel A. Imbens Kalyanaraman Optimal Bandwidth				
Offense Committed After Aug. 23, 1996 (Banned)	0.0438	0.0716**	0.0438	0.0308
	(0.0372)	(0.0301)	(0.0310)	(0.0227)
Control Group Mean	0.1046	0.0536	0.0508	0.0304
Observations	1029	964	721	1042
Bandwidth (in Days)	277	256	219	305
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1
Panel B. Consistent Bandwidth of ±240 Days				
Offense Committed After Aug. 23, 1996 (Banned)	0.0581	0.0685**	0.0259	0.0184
	(0.0402)	(0.0316)	(0.0296)	(0.0264)
Control Group Mean	0.1071	0.0548	0.0497	0.0260
Observations	893	893	796	801
Bandwidth (in Days)	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison within 0-5 years of being released. Column 2 estimates the effect on financially motivated recidivism within 0-5 years of being released. Column 3 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison within 5-10 years of being released. Column 4 estimates the effect on financially motivated recidivism within 5-10 years of being released. See Table A2.3 for a definition of financially motivated crimes and the associated recidivism measure. Time until recidivism is defined as the difference between the offender's release date for prison stay j and the next offense date before prison stay $j+1$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.19: Effect of Ban when SNAP is Most Generous for Non-Banned Offenders

Outcome:	Recidivism in Time/Place with ABAWD Work Waiver	Recidivism in Time/Place with No ABAWD Work Waiver
	(1)	(2)
Panel A. Imbens Kalyanaraman Optimal Bandwidth		
Offense Committed After	0.0996**	-0.0039
Aug. 23, 1996 (Banned)	(0.0415)	(0.0292)
Control Group Mean	0.0874	0.0761
Observations	936	990
Bandwidth (in Days)	±242	±256
Degree of Polynomial in Days from Aug. 23, 1996	1	1
Panel B. Consistent Bandwidth of ±240 Days		
Offense Committed After	0.1037**	-0.0087
Aug. 23, 1996 (Banned)	(0.0418)	(0.0306)
Control Group Mean	0.0880	0.0764
Observations	918	918
Bandwidth (in Days)	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison with a crime that was committed in a time (based on earliest offense date after release) and place (based on county of conviction) where ABAWD work requirements were waived. Column 2 estimates the effect on recidivism with a crime that was committed in a time and place where ABAWD work requirements were in effect. The ABAWD work requirement states that able-bodied adults without dependents are limited to only 3 months of SNAP receipt every 3 years unless they: (1) work 20 or more hours per week, (2) participate in an employment and training program, or (3) participate in a workfare program (USDA 2016). Thus, when these requirements are waived, SNAP is especially generous for ABAWDs not subject to the ban. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.20: Effect of Ban when SNAP is Most Generous for Non-Banned Offenders, Using Release Plan Residence

Outcome:	Recidivism in Time/Place with ABAWD Work Waiver	Recidivism in Time/Place with No ABAWD Work Waiver
	(1)	(2)
Panel A. Imbens Kalyanaraman Optimal Bandwidth		
Offense Committed After Aug. 23, 1996 (Banned)	0.0997**	0.0002
	(0.0420)	(0.0317)
Control Group Mean	0.0833	0.0797
Observations	918	997
Bandwidth (in Days)	$\pm$240	$\pm$258
Degree of Polynomial in Days from Aug. 23, 1996	1	1
Panel B. Consistent Bandwidth of $\pm$240 Days		
Offense Committed After Aug. 23, 1996 (Banned)	0.0997**	-0.0048
	(0.0420)	(0.0335)
Control Group Mean	0.0833	0.0810
Observations	918	918
Bandwidth (in Days)	$\pm$240	$\pm$240
Degree of Polynomial in Days from Aug. 23, 1996	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison with a crime that was committed in a time (based on earliest offense date after release) and place (based on county of residence on release plan) where ABAWD work requirements were waived. Column 2 estimates the effect on recidivism with a crime that was committed in a time and place where ABAWD work requirements were in effect. See Table A2.19 for more information about the ABAWD work requirement. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.21: Effect of SNAP Ban on Offenders When ABAWD Work Requirements Waived, Hazard Model with Year Effects

Outcome:	Recidivism	
	(1)	(2)
Panel A. Imbens Kalyanaraman Optimal Bandwidth		
Offense Committed After	0.5680*	-0.7222
Aug. 23, 1996 (Banned)	(0.3413)	(0.8467)
Banned X ABAWD Waiver		1.6465*
		(0.9752)
Combined Effect:		0.9243**
Banned + (Banned X Waiver)		0.4105
Observations	117441	117441
Bandwidth (in Days)	±212	±212
Panel B. Consistent Bandwidth of ±240 Days		
Offense Committed After	0.6499**	-0.7483
Aug. 23, 1996 (Banned)	(0.3184)	(0.7009)
Banned X ABAWD Waiver		1.8310**
		(0.8301)
Combined Effect:		1.0827***
Banned + (Banned X Waiver)		0.3909
Observations	135733	135733
Bandwidth (in Days)	±240	±240

Notes: Standard errors clustered at the day of offense in parentheses. This analysis uses a Cox survival model in which offenders enter when they are released from prison and exit when they return to prison. Since the analysis includes time-varying covariates, the data was transformed to a format where every row is an offender-month-year observation for the time that they are out of prison. All specifications include year fixed effects. Number of days the drug trafficking offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison. Column 2 estimates heterogeneity in the effect by whether or not the offender is living in a county where ABAWD work requirements are waived. In Panel A, the Imbens-Kalyanaraman optimal bandwidth (chosen from the micro data pre-transformation) is used with polynomial of degree one and a uniform kernel. In Panel B, a bandwidth of ±240 days (or ±8 months) is used with polynomial of degree one and a uniform kernel. See Table A2.19 for more information about the ABAWD work requirement. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.22: Placebo Test: Recidivism for Sell/Mfg/Dist Drug Offenders (Not Banned)

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0109	-0.0235	0.0366
	(0.0294)	(0.0244)	(0.0255)
Control Group Mean	0.5534	0.3473	0.1934
Observations	4903	6103	3925
Bandwidth (in Days)	$\pm$302	$\pm$412	$\pm$239
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of $\pm$240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0056	-0.0313	0.0369
	(0.0326)	(0.0304)	(0.0254)
Control Group Mean	0.5510	0.3577	0.1933
Observations	3934	3934	3934
Bandwidth (in Days)	$\pm$240	$\pm$240	$\pm$240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths. Number of days the selling/manufacturing/distributing drugs (SMD) offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of committing an SMD offense on or after Aug. 23, 1996 on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. SMD offenses are not subject to the SNAP ban, and thus, committing one before versus after the cutoff date should not affect an individual's recidivism. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.23: Placebo Test: Recidivism for Non-Drug Offenders (Not Banned)

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0088	0.0065	-0.0002
	(0.0143)	(0.0126)	(0.0092)
Control Group Mean	0.3930	0.2425	0.1505
Observations	26375	29232	21928
Bandwidth (in Days)	±506	±595	±373
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.0062	0.0072	-0.0010
	(0.0187)	(0.0164)	(0.0109)
Control Group Mean	0.3933	0.2427	0.1506
Observations	15166	15166	15166
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths. Number of days the non-drug offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of committing an SMD offense on or after Aug. 23, 1996 on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. Non-drug offenses are not subject to the SNAP ban, and thus, committing one before versus after the cutoff date should not affect an individual's recidivism. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.24: Additional Placebo Tests: Recidivism Outcomes for Other (Not Banned) Offenders

Outcome:	Recidivism				
Offender Type:	All Non-Drug Offenders	DUI & Revoked License	Drug Possession	Property Crime	Violent Crime
	(1)	(2)	(3)	(4)	(5)
Panel A. Imbens Kalyanaraman Optimal Bandwidth					
Offense Committed After Aug. 23, 1996 (Banned)	0.0088	-0.0669	0.0040	-0.0238	-0.0085
	(0.0143)	(0.0756)	(0.0278)	(0.0195)	(0.0259)
Control Group Mean	0.3930	0.4264	0.5613	0.4756	0.3418
Observations	26375	798	5254	10523	7906
Bandwidth (in Days)	505	177	249	234	238
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1
Panel B. Consistent Bandwidth of ±240 Days					
Offense Committed After Aug. 23, 1996 (Banned)	0.0062	-0.0560	0.0082	-0.0225	-0.0085
	(0.0187)	(0.0648)	(0.0284)	(0.0191)	(0.0257)
Control Group Mean	0.3933	0.4077	0.5619	0.4756	0.3417
Observations	15166	1092	5103	10785	7965
Bandwidth (in Days)	±240	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths. Number of days the placebo offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of committing any non-drug offense after Aug 23, 1996 on recidivism. Column 2 estimates the effect of committing a DUI or driving with a revoked license after Aug 23, 1996. Column 3 estimates the effect of committing drug possession after Aug 23, 1996. Column 4 estimates the effect of committing a property crime after Aug. 23, 1996. Column 5 estimates the effect of committing a violent crime after Aug 23, 1996. None of these offenses are subject to the SNAP ban, and thus, committing one before versus after the cutoff date should not affect an individual's recidivism. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.25: Effect of the SNAP Ban on Recidivism with Seasonal Controls

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Offense Committed After	0.0986*	0.1046**	-0.0060
Aug. 23, 1996 (Banned)	(0.0532)	(0.0468)	(0.0316)
Offense Committed After	0.0040	0.0042	-0.0002
Any Aug. 23 1996-2012	(0.0094)	(0.0086)	(0.0062)
Control Group Mean	0.1587	0.0825	0.0762
Observations	16519	16519	16519
Bandwidth (in Days)	±180	±180	±180
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. Number of days the drug trafficking offense was committed before or after Aug. 23 in a given year is the running variable (centered at zero). Specifically, I estimate both a "seasonality effect" and a "true effect" of the ban, where the seasonality effect is the effect of committing a drug trafficking offense after Aug. 23 in general and the true effect is the effect of committing a drug trafficking offense after Aug. 23, 1996. In all specifications a bandwidth of ±180 days is used to avoid overlapping observations across years. Also, this estimation excludes the years 1998 and 1999 since those are two years in which Florida implemented criminal justice policies that would directly affect drug traffickers. Column 1 estimates the effect of being banned from SNAP on whether or not the offender ever returns to prison. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.26: Test of Deterrence Hypothesis: Effect of Ban on Type of Financially Motivated Recidivism

Outcome:	Recidivism for Drug Trafficking Crime	Recidivism for Non-Drug Trafficking Crime	Recidivism for Property Crime	Recidivism for Theft
	(1)	(2)	(3)	(4)
Panel A. Imbens Kalyanaraman Optimal Bandwidth				
Offense Committed After Aug. 23, 1996 (Banned)	0.0197	0.0919**	0.0415**	0.0549***
	(0.0161)	(0.0356)	(0.0192)	(0.0165)
Control Group Mean	0.0312	0.0621	0.0212	0.0123
Observations	1452	940	1232	1048
Bandwidth (in Days)	±411	±244	±317	±275
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1
Panel B. Consistent Bandwidth of ±240 Days				
Offense Committed After Aug. 23, 1996 (Banned)	0.0110	0.0892**	0.0526**	0.0586***
	(0.0181)	(0.0363)	(0.0218)	(0.0174)
Control Group Mean	0.0255	0.0625	0.0208	0.0116
Observations	918	918	918	918
Bandwidth (in Days)	±240	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths. Number of days the drug trafficking offense was committed before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison due to a drug trafficking crime. Column 2 estimates whether or not the offender returns to prison due to a financially motivated crime that is **not** drug trafficking. Column 3 estimates the effect on recidivism due to a property crime, and column 4 estimates the effect on recidivism due to theft. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.27: Effect of SNAP Ban on Recidivism for Crimes in Offender's History, Not in Offender's History, and Total Crimes

Outcome:	Recidivism with Only Crimes Not Convicted of Previously	Recidivism with a Crime Convicted of Previously	Total # of Crimes After Trafficking Conviction
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	-0.0071	0.1168**	0.3195**
	(0.0064)	(0.0504)	(0.1522)
Control Group Mean	0.0018	0.1600	0.3943
Observations	1225	735	735
Bandwidth (in Days)	±314	±197	±197
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	-0.0118	0.1067**	0.2464*
	(0.0087)	(0.0467)	(0.1374)
Control Group Mean	0.0023	0.1620	0.3866
Observations	918	918	918
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths and the running variable. Column 1 estimates the effect of being banned from SNAP after release on whether or not the offender returns to prison exclusively due to a crime that they have not committed before. Column 2 estimates the effect of being banned from SNAP on whether or not the offender returns to prison with a crime that they have committed before. Column 3 estimates the effect of being banned from SNAP on the total number of crimes the offender is convicted of in the future. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

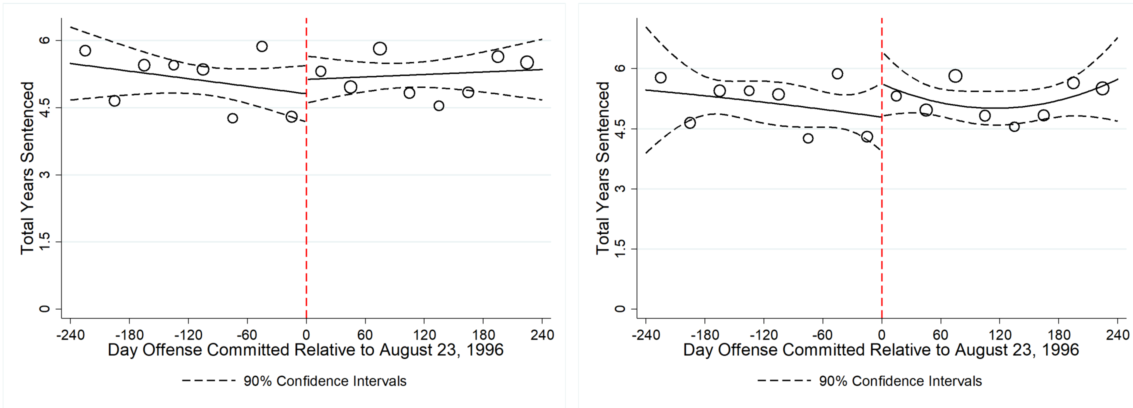
Table A2.28: Effect of SNAP Ban on Recidivism in Florida, Mis-Measuring Treatment by Using Conviction Date

Outcome:	Recidivism	Financially Motivated Recidivism	Non-Financially Motivated Recidivism
	(1)	(2)	(3)
Panel A. Imbens Kalyanaraman Optimal Bandwidth			
Offense Committed After Aug. 23, 1996 (Banned)	0.0195	-0.0599	0.0876**
	(0.0651)	(0.0423)	(0.0409)
Control Group Mean	0.1545	0.1048	0.0488
Observations	733	1147	702
Bandwidth (in Days)	±452	±687	±433
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1
Panel B. Consistent Bandwidth of ±240 Days			
Offense Committed After Aug. 23, 1996 (Banned)	0.1136	-0.0223	0.1359**
	(0.0791)	(0.0609)	(0.0552)
Control Group Mean	0.1570	0.1074	0.0496
Observations	387	387	387
Bandwidth (in Days)	±240	±240	±240
Degree of Polynomial in Days from Aug. 23, 1996	1	1	1

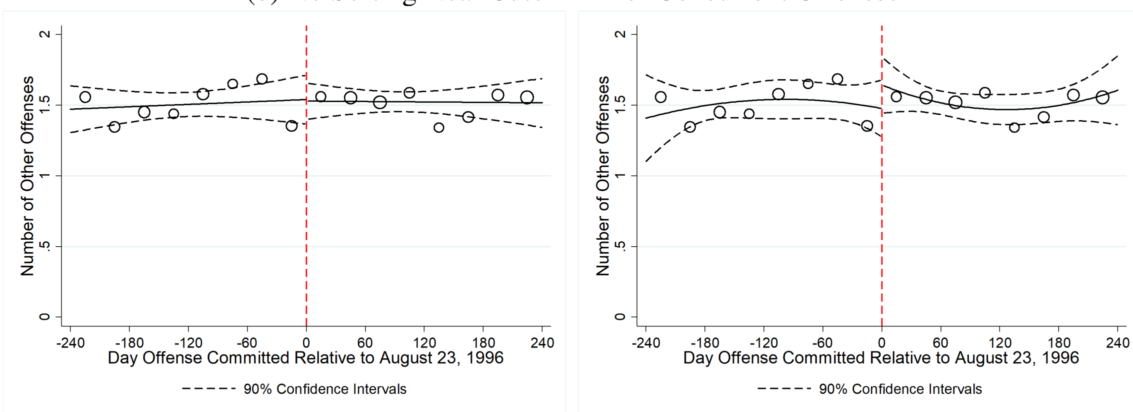
Notes: Standard errors clustered at the day of offense in parentheses. See Table A2.2 for general notes about the RD estimation, including information about bandwidths. Number of days the drug trafficker was convicted before or after Aug. 23, 1996 is the running variable (centered at zero). Column 1 estimates the effect of being banned from SNAP on whether or not the offender returns to prison after being released. Column 2 and Column 3 estimate the effect on recidivism with financially motivated crimes and the effect on recidivism with non-financially motivated crimes, respectively. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures. Since the ban is determined based on the date the drug trafficking offense is committed, estimating the effect based on date of conviction introduces measurement error into the model. Conviction dates are often months or years after the offense date. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A2.1: Visual Evidence that RD Identifying Assumption Holds

(a) No Sorting Near Cutoff in Total Years Sentenced

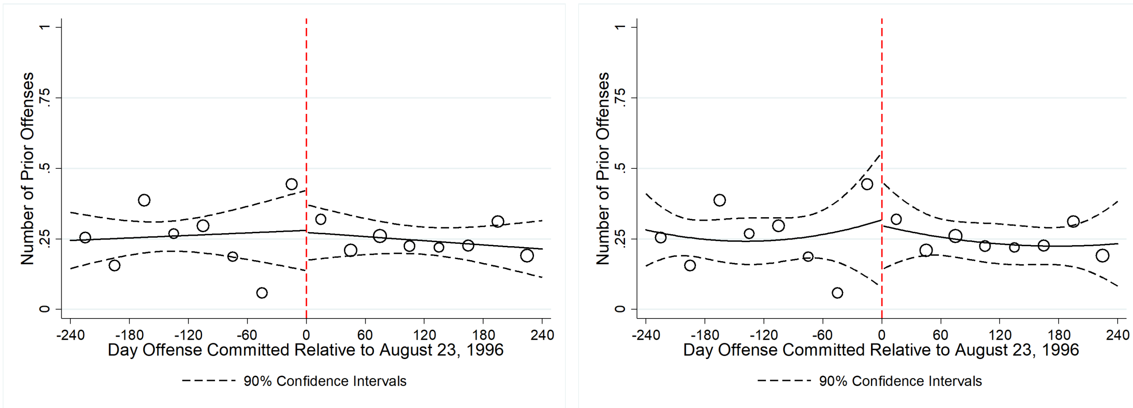


(b) No Sorting Near Cutoff in # of Concurrent Offenses

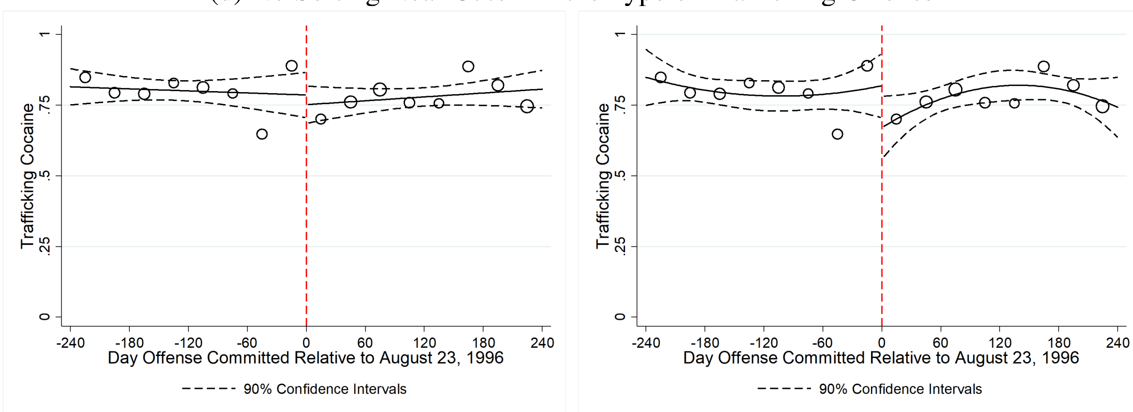


Notes: The figures in the first column display the lines from two local linear regressions, estimated separately on each side of the cutoff using the offense-level micro data. The figures in the second column display the lines from two local quadratic regressions, estimated separately on each side of the cutoff using the offense-level micro data. I also overlay a scatter plot of 30-day bin averages of the dependent variable weighted by the number of offenses in each 30-day bin. The running variable in these figures (and the following RD plots) is the number of days between the offender's offense date and August 23, 1996 (the cutoff date that determines the offender's ban status). The running variable is centered at zero such that offenders committing an offense before August 23, 1996 have a negative distance from the cutoff date and offenders committing an offense after August 23, 1996 have a positive distance from the cutoff date. The dependent variables in these figures are offender characteristics: total years sentenced and number of concurrent offenses.

Figure A2.1: Visual Evidence that RD Identifying Assumption Holds
(c) No Sorting Near Cutoff in # of Prior Offenses

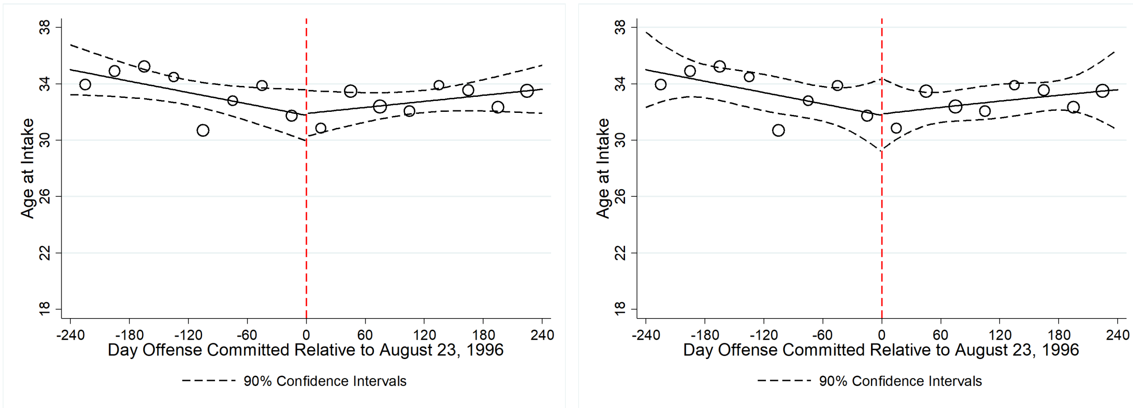


(d) No Sorting Near Cutoff in the Type of Trafficking Offense

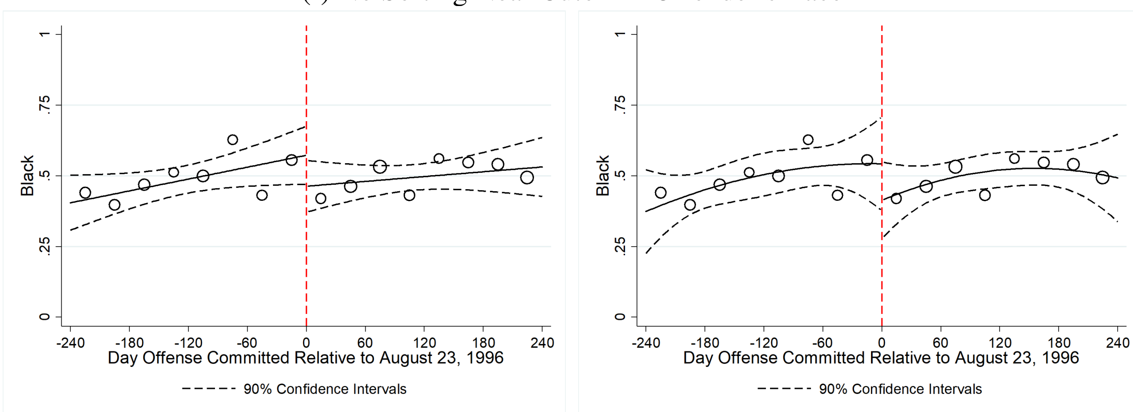


Notes: See notes from Figures A2.1a-b above. The dependent variables in these figures are offender characteristics: number of prior offenses and type of trafficking.

Figure A2.1: Visual Evidence that RD Identifying Assumption Holds
(e) No Sorting Near Cutoff in Offender Age at Intake



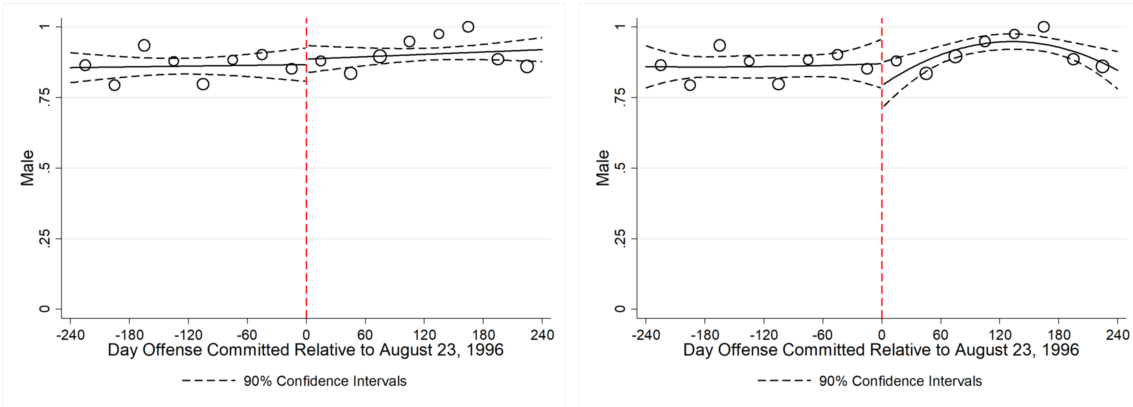
(f) No Sorting Near Cutoff in Offender's Race



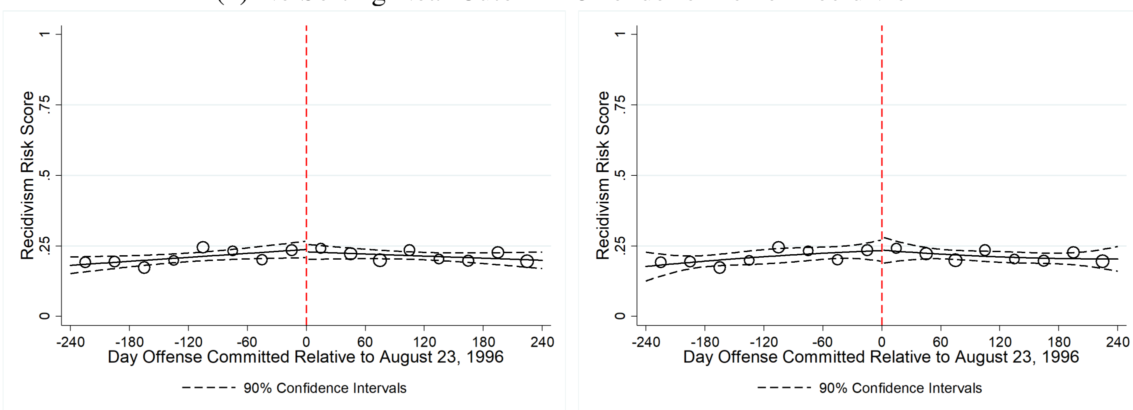
Notes: See notes from Figures A2.1a-b above. The dependent variables in these figures are offender characteristics: age at intake and race, and risk of recidivism.

Figure A2.1: Visual Evidence that RD Identifying Assumption Holds

(g) No Sorting Near Cutoff in Offender's Sex

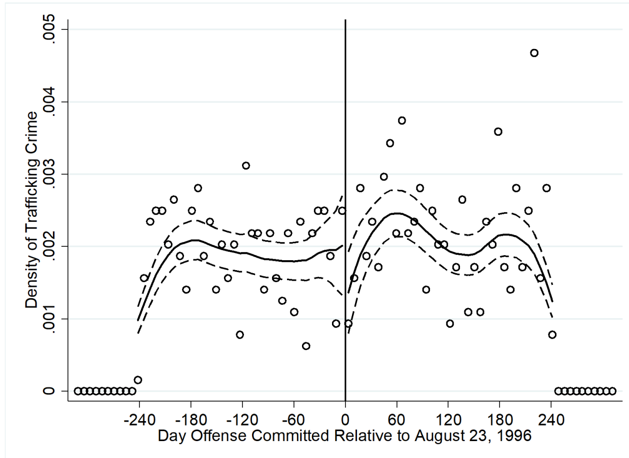


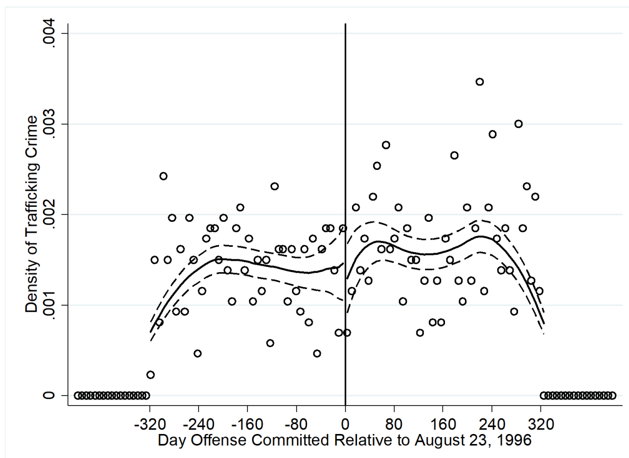
(h) No Sorting Near Cutoff in Offender's Risk of Recidivism



Notes: See notes from Figures A2.1a-b above. The dependent variables in these figures are offender characteristics: sex and risk of recidivism. See Figure 2.1 or Table A2.2 for notes about the calculation of risk of recidivism.

Figure A2.2: No Break in the Density of Drug Trafficking Crime Near August 23, 1996

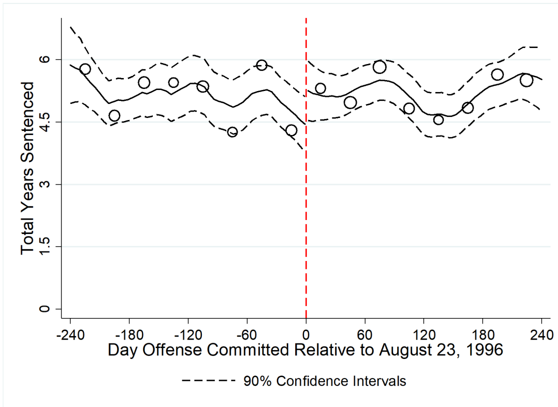




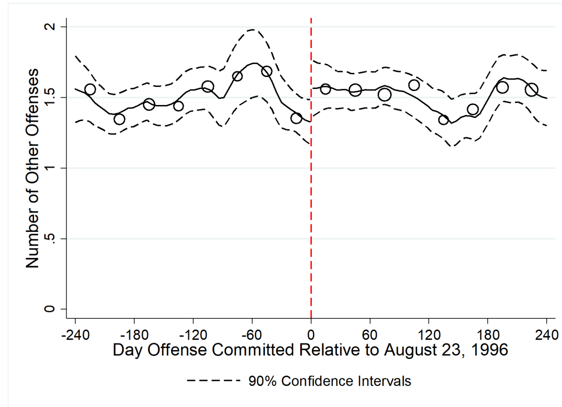
Notes: Both figures display the density of drug trafficking crime on each day in a narrow band around August 23, 1996. The figure the first row shows this for a bandwidth of 240 days before and after August 23, 1996 while the figure in the second row shows this for bandwidth of 320 days before and after August 23, 1996. Neither figure shows a statistical break in the density of drug trafficking crimes near the cutoff date—this is further evidence against endogenous sorting. I use the Stata program **DCDensity.ado** provided by Justin McCrary and Brian Kovak to conduct this test.

Figure A2.3: Non-parametric Visual Evidence that RD Identifying Assumption Holds

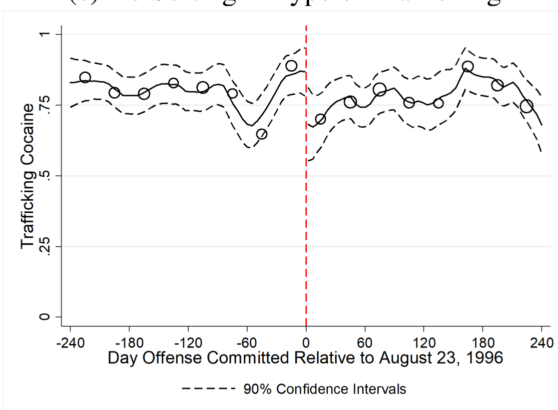
(a) No Sorting in Years Sentenced



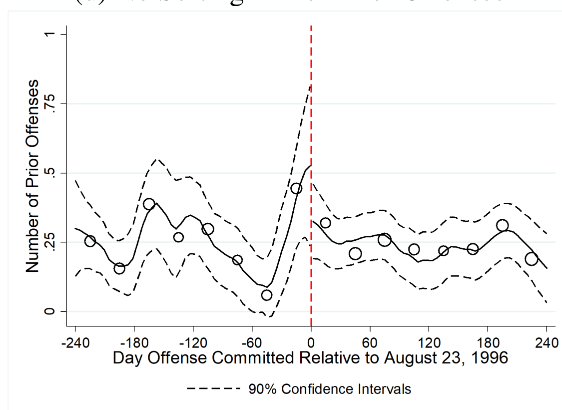
(b) No Sorting in # of Concurrent Offenses



(c) No Sorting in Type of Trafficking



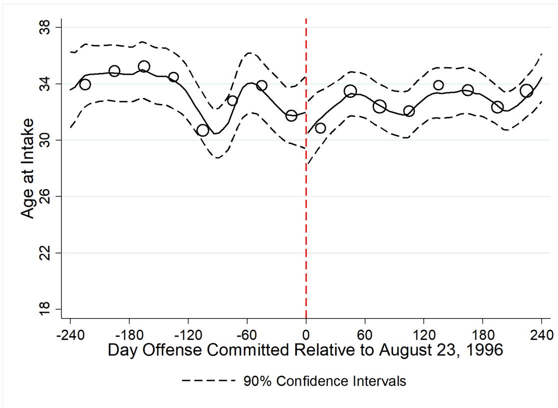
(d) No Sorting in # of Prior Offenses



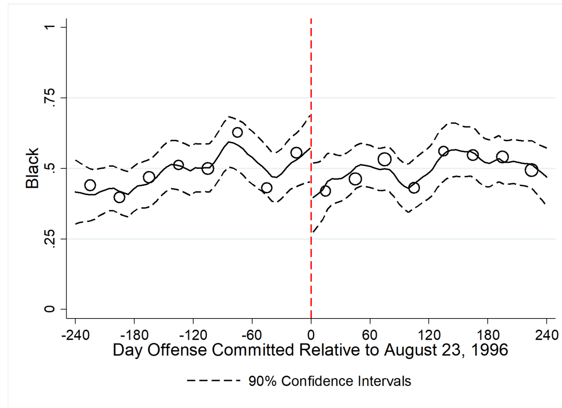
Notes: The figures above display the lines from two locally smoothed regressions, estimated separately on each side of the cutoff using the offense-level micro data. I also overlay a scatter plot of 30-day bin averages of the dependent variable weighted by the number of offenses in each 30-day bin. See Figures A2.1a-b for notes about the running variable. The dependent variables in these figures are offender characteristics: total years sentenced, number of concurrent offenses, number of prior offenses, and type of trafficking. All figures are made with Stata command **lpolyci** using the default settings.

Figure A2.3: Non-parametric Visual Evidence that RD Identifying Assumption Holds

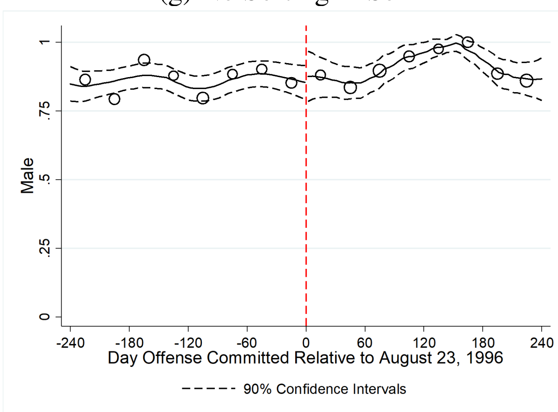
(e) No Sorting in Age at Intake



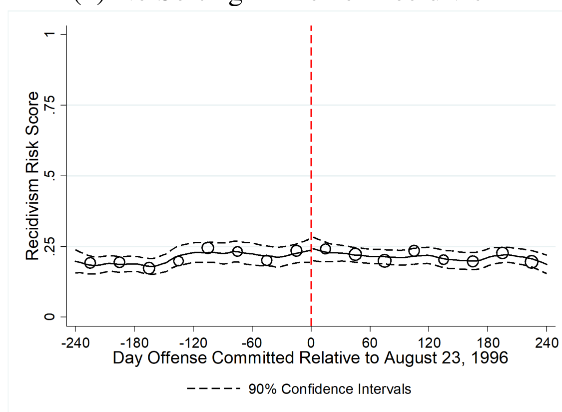
(f) No Sorting in Race



(g) No Sorting in Sex



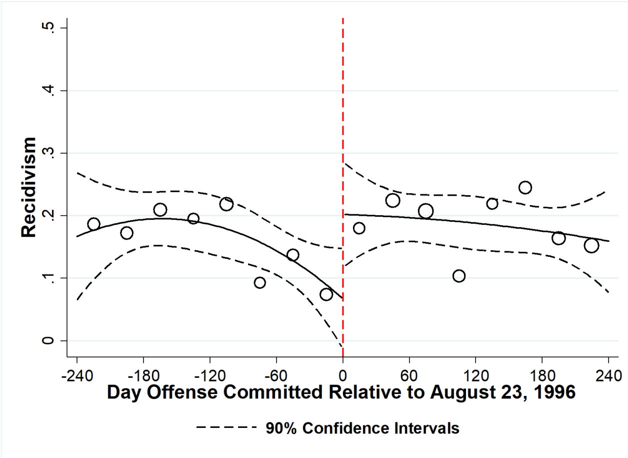
(h) No Sorting in Risk of Recidivism



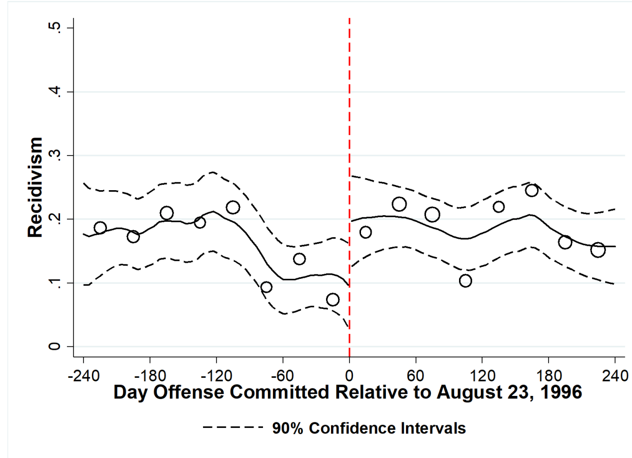
Notes: See the notes for Figures A2.3a-d. The dependent variables in these figures are offender characteristics: age at intake, race, sex, and risk of recidivism. See Figure 2.1 or Table A2.2 for notes on how risk of recidivism is calculated.

Figure A2.4: Visual Evidence of Main Result: Offenders Subject to SNAP Ban are More Likely to Recidivate

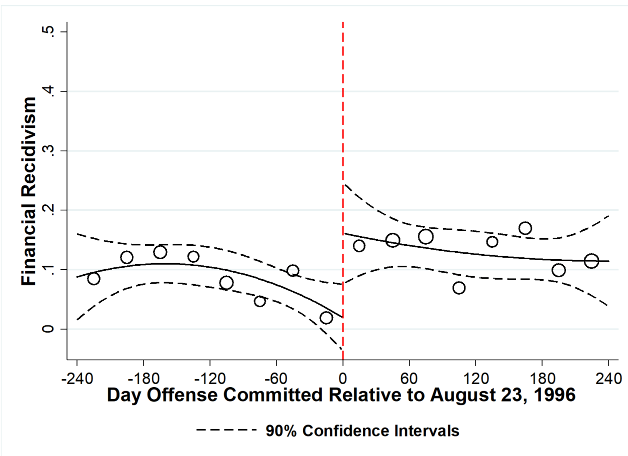
(a) Any Recidivism, Quadratic



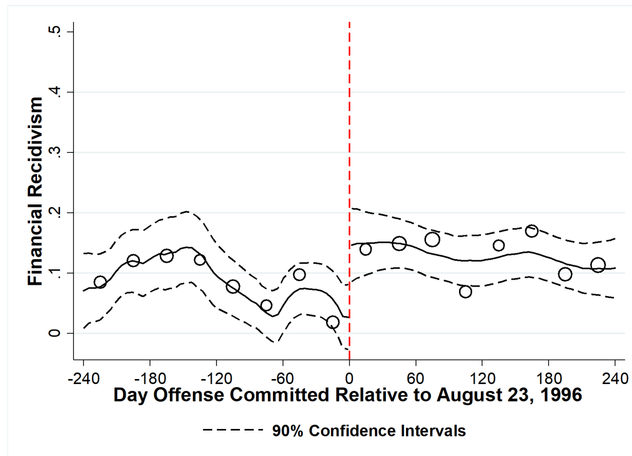
(b) Any Recidivism, Nonparametric



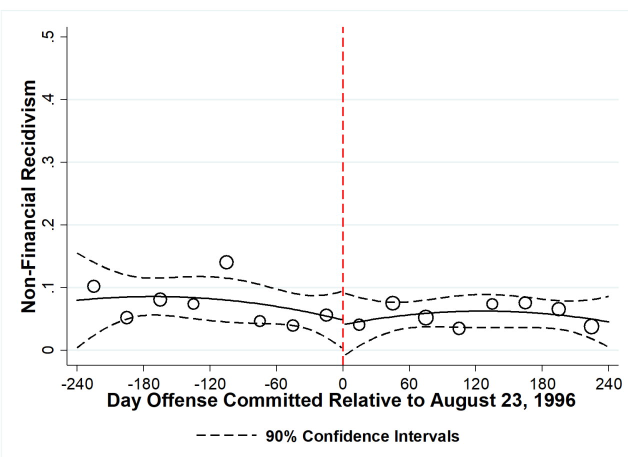
(c) Financial Recidivism, Quadratic



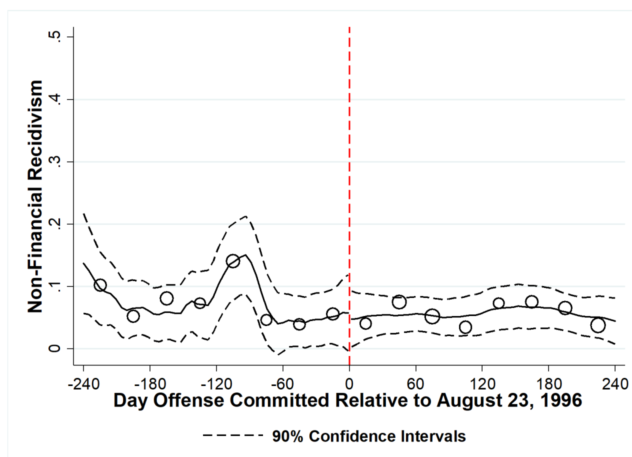
(d) Financial Recidivism, Nonparametric



(e) Non-Financial Recidivism, Quadratic



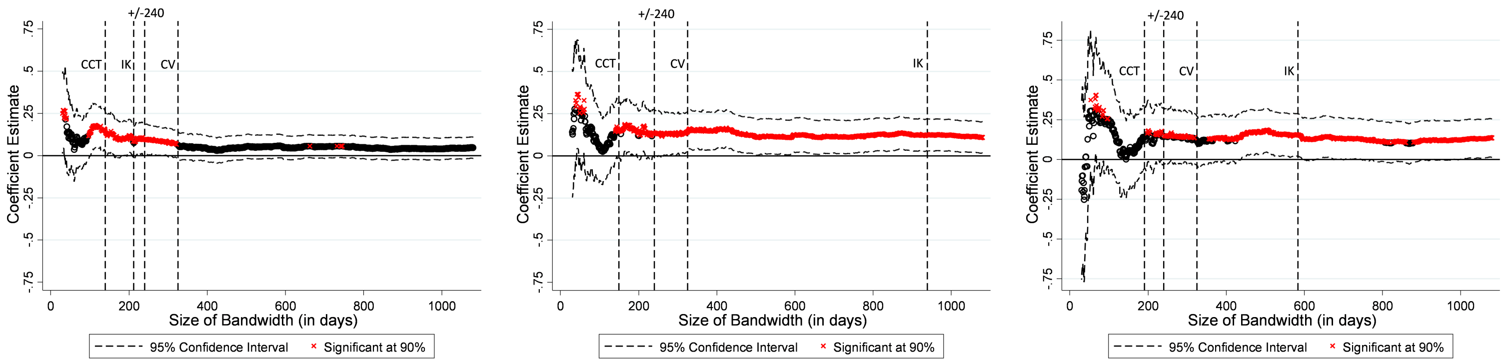
(f) Non-Financial Recidivism, Nonparametric



Notes: The figures in the first column display the lines from two local quadratic regressions, estimated separately on each side of the cutoff using the offense-level micro data. The figures in the second column display the lines from two locally smoothed regressions, estimated separately on each side of the cutoff using the offense-level micro data. I also overlay a scatter plot of 30-day bin averages of the dependent variable weighted by the number of offenses in each 30-day bin. See Figures A2.1a-b for notes about the running variable. The dependent variables in these figures are offender outcomes: recidivism, financial recidivism, and non-financial recidivism. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

Figure A2.5: Estimate of Effect over Many Bandwidths

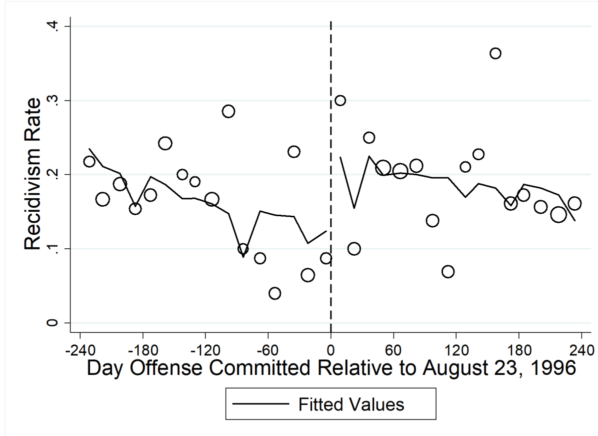
(a) Linear Polynomial (b) Quadratic Polynomial (c) Cubic Polynomial



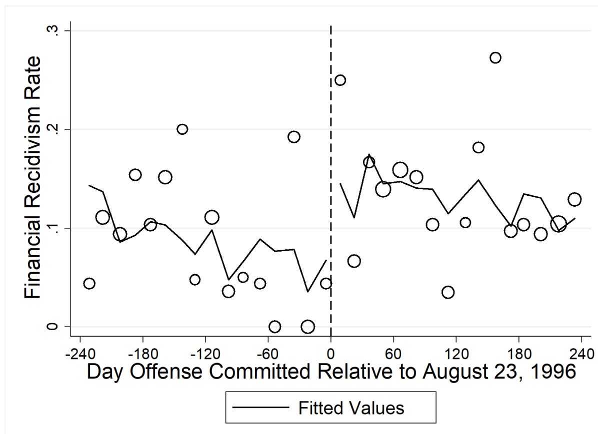
Notes: The figures above display the coefficient estimates from regressions with bandwidths ranging from ±30 days from August 23, 1996 to ±1080 days from August 23, 1996. The coefficient estimate is plotted on the y-axis and the corresponding bandwidth that yields that coefficient is plotted on the x-axis. Each figure includes four vertical lines denoting the Calonico, Cattaneo, Titiunik (CCT) optimal bandwidth, the Ludwig, Miller Cross-Validation (CV) optimal bandwidth, the Imbens, Kalyanaraman (IK) optimal bandwidth, and the consistent ±240 day bandwidth used throughout the paper. In Figure A2.5a, the regressions include a linear polynomial of the running variable. In Figure A2.5b, the regressions include a quadratic polynomial of the running variable. In Figure A2.5c, the regressions include a cubic polynomial of the running variable. 95% confidence intervals are plotted and coefficients are marked red when significant at the 90% level. Bandwidths greater than ±327 days are asymmetric since the data only includes offenses occurring after October 1, 1995.

Figure A2.6: Visual Evidence of Time-Series Result: Offenders Subject to SNAP Ban are More Likely to Recidivate

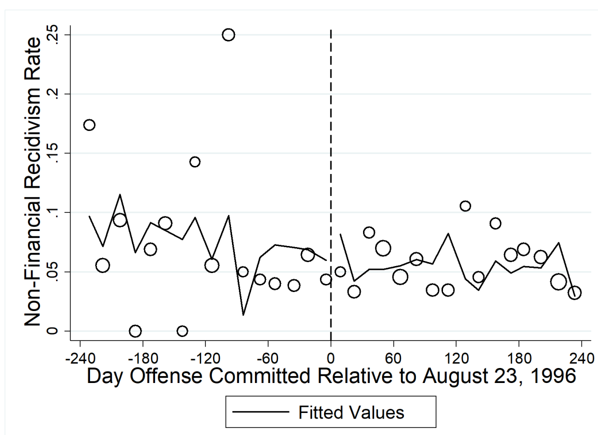
(a) Any Recidivism



(b) Financial Recidivism

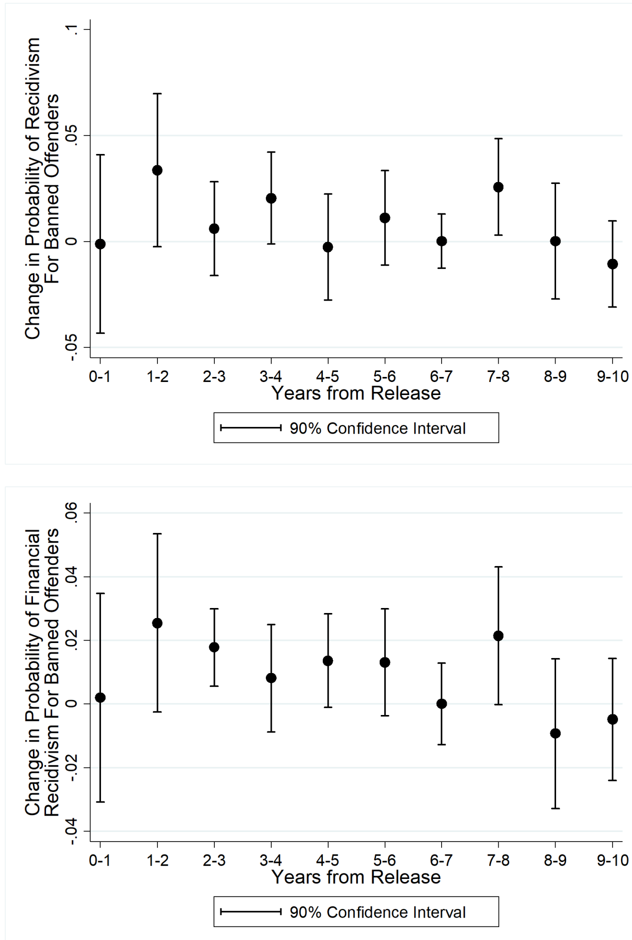


(c) Non-Financial Recidivism



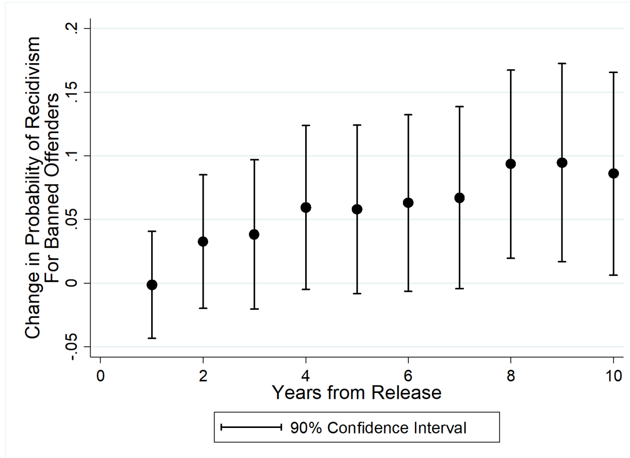
Notes: The figures above plot the lines of fitted values from time-series regressions modeling recidivism rates as an AR(1) process (number of lags chosen using the model with the highest AIC). All figures are overlaid with a scatter plot of the dependent variable averaged in 15-day bins. See Figures A2.1a-b for notes about the running variable. See Table A2.11 for notes about the time-series estimation. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

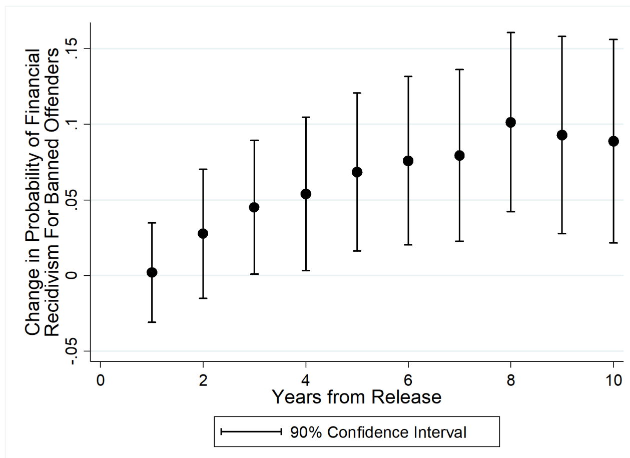
Figure A2.7: Effect of SNAP Ban on Timing of Re-incarceration



Notes: The first figure above displays the coefficient from ten separate regressions to illustrate how the SNAP ban affects timing of re-incarceration. For example, the coefficient plotted at "1-2" on the x-axis is the coefficient from a regression of whether or not the offender returns to prison within 1-2 years after release on whether or not the offender is banned from SNAP (committed a drug-trafficking offense on or after Aug 23, 1996). The second figure displays ten coefficients from similar regressions that use timing of financial recidivism as the dependent variable instead of timing of any recidivism. All regressions use a linear polynomial of the running variable, uniform kernel, and a bandwidth of ±240 days. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

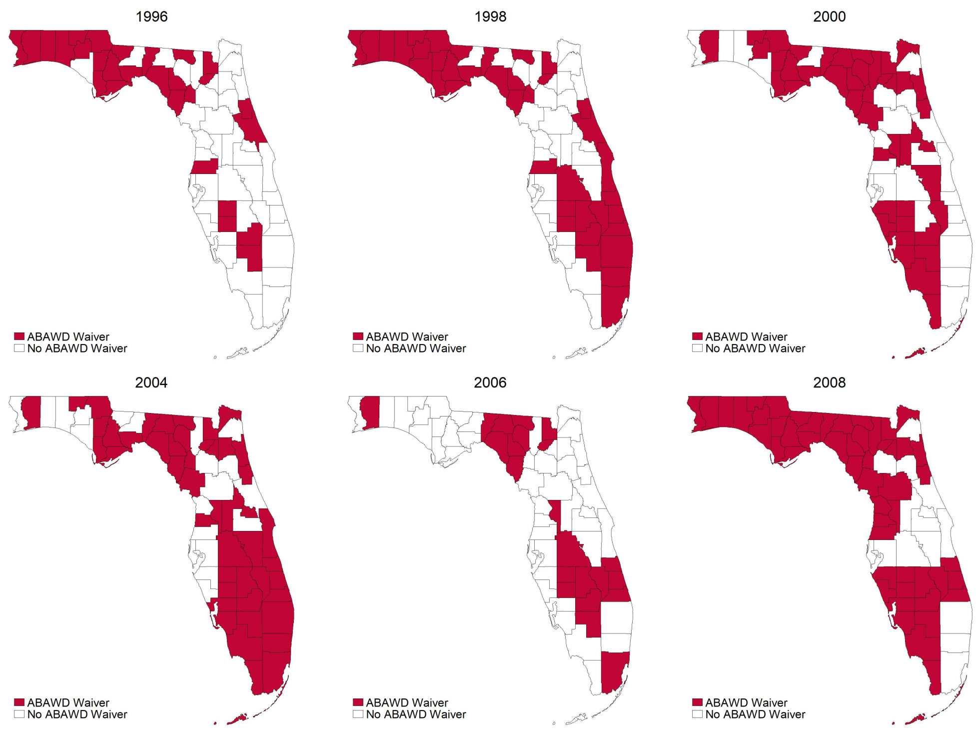
Figure A2.8: Effect of SNAP Ban on Timing of Re-incarceration, Cumulative





Notes: The first figure above displays the coefficient from ten separate regressions to illustrate how the SNAP ban affects timing of re-incarceration. For example, the coefficient plotted at "1" on the x-axis is the coefficient from a regression of whether or not the offender returns to prison within 0-1 years after release on whether or not the offender is banned from SNAP (committed a drug-trafficking offense on or after Aug 23, 1996). Similarly, the coefficient plotted at "5" is the coefficient from a regression of whether or not the offender returns to prison within 0-5 years after release on whether or not the offender is banned from SNAP. The second figure displays ten coefficients from similar regressions that use timing of financial recidivism as the dependent variable instead of timing of any recidivism. All regressions use a linear polynomial of the running variable, uniform kernel, and a bandwidth of ± 240 days. See Table A2.3 for a definition of financially and non-financially motivated crimes and the associated recidivism measures.

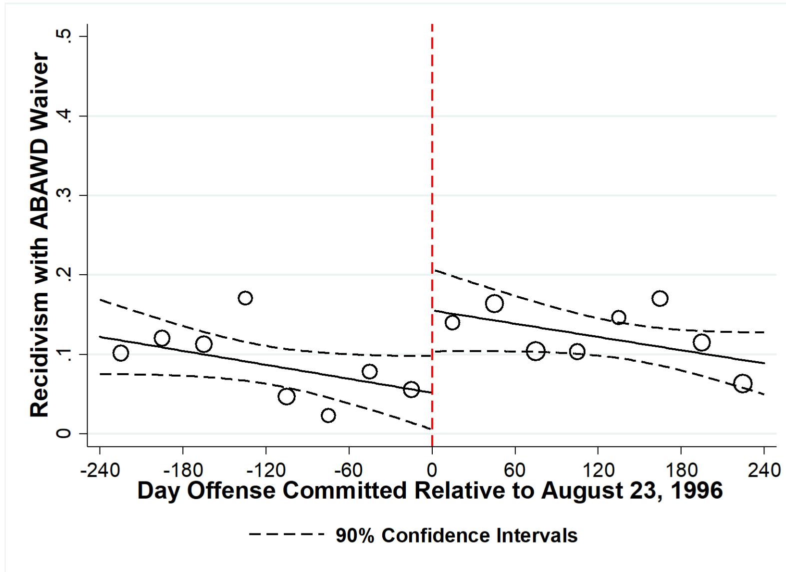
Figure A2.9: Geographic Variation in ABAWD Work Requirement Waivers, 1996-2008



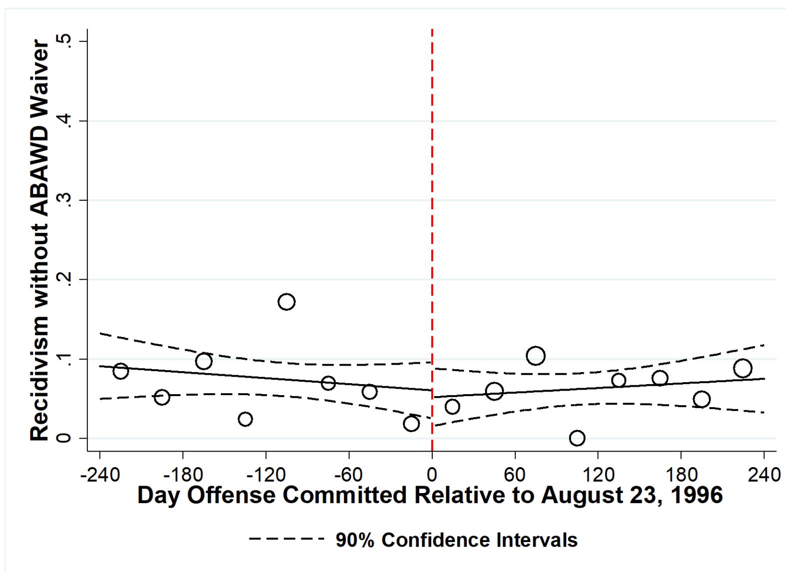
Notes: The figures above display which Florida counties have an ABAWD work requirement waiver at any point in a given year. When a county is filled in with red, it has an ABAWD work requirement waiver at some point in that year. When a county is filled in with white, it never has an ABAWD work requirement waiver in that year. I display every even-numbered year starting in 1996 and ending in 2008. I do not display years past 2008 since there is a nationwide ABAWD work requirement waiver in place from 2009-2016. Also, there is a nationwide ABAWD work requirement waiver in place from 2001-2003, so I do not display the map for 2002. An animation showing the above maps for every month-year combination from 1996-2009 is available here: `https://www.dropbox.com/s/kufg1ieiwtjm0b6/Waivers%20by%20County-Month.gif?dl=0`

Figure A2.10: Effect of SNAP Ban on Recidivism with/without ABAWD Work Waivers

(a) Recidivism in Time/Place with ABAWD Work Waiver



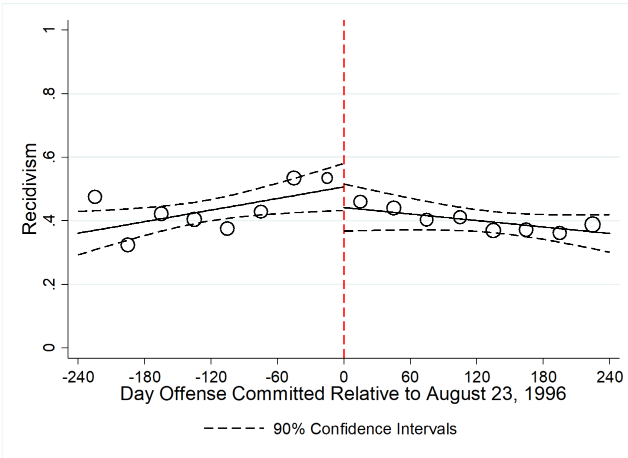
(b) Recidivism in Time/Place without ABAWD Work Waiver



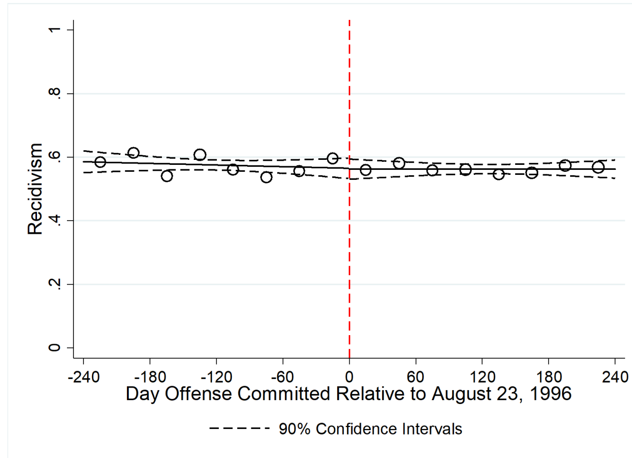
Notes: The figures above (and the following RD plots more generally) display the lines from two local linear regressions, estimated separately on each side of the cutoff using the offense-level micro data. I also overlay a scatter plot of 30-day bin averages of the dependent variable weighted by the number of offenses in each 30-day bin. See Figure A2.1a-b for notes about the running variable. The dependent variable in Figure A2.10a is whether or not the offender returns to prison for a crime committed in a time and place when an ABAWD work waiver was in effect. The dependent variable in Figure A2.10b is whether or not the offender returns to prison for a crime committed in a time and place when an ABAWD work waiver was not in effect. See Table A2.19 for more detail about this estimation and the ABAWD work requirement more generally.

Figure A2.11: Visual Evidence of Main Result: Offenders Subject to SNAP Ban are More Likely to Recidivate

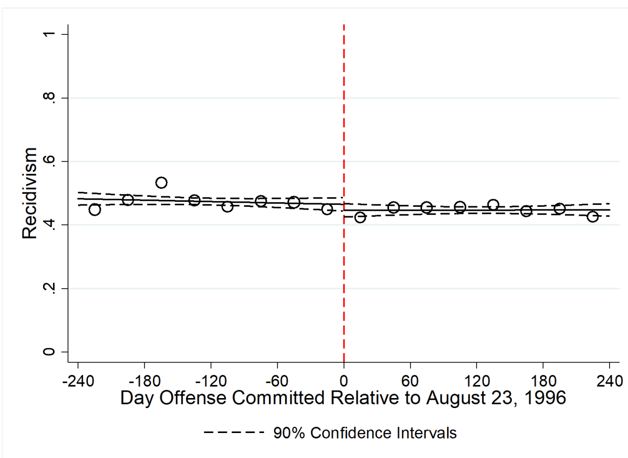
(a) DUI or Revoked License



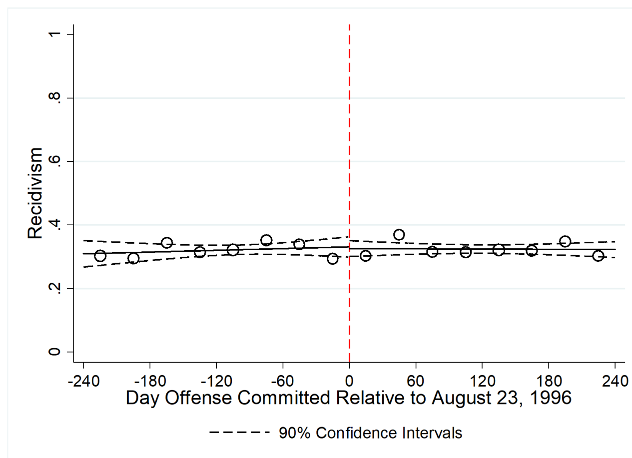
(b) Drug Possession



(c) Property Crime



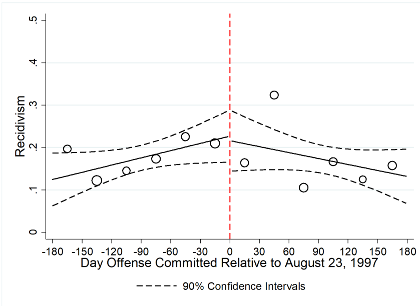
(d) Violent Crime



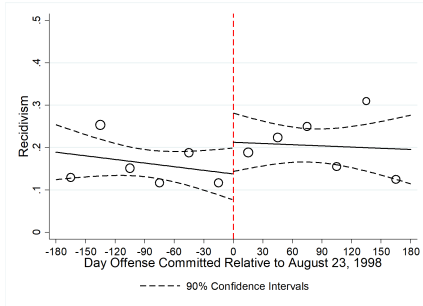
Notes: The figures above plot the lines from local linear regressions of recidivism outcomes on the running variable (days before and after August 23, 1996), estimated separately on each side of the cutoff for several different “placebo” crimes (crimes that do not lead to permanent ban from SNAP in Florida). All figures are overlaid with a scatter plot of recidivism averaged in 30-day bins. See Figures A2.1a-b for notes about the running variable. See Figure A2.4 for general notes about the creation of the RD plots for drug traffickers. These plots employ the same method but on a sample of offenders who do not commit drug trafficking but instead commit the following crimes: DUI/driving with a revoked license, drug possession, property crime, and violent crime.

Figure A2.12: Drug Traffickers in Other Years are Not More Likely to Recidivate

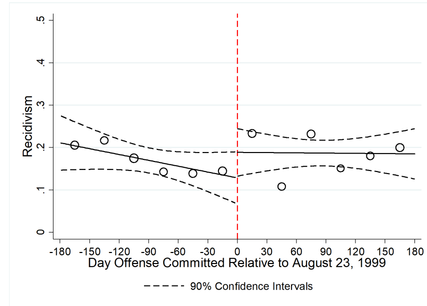
(a) 1997



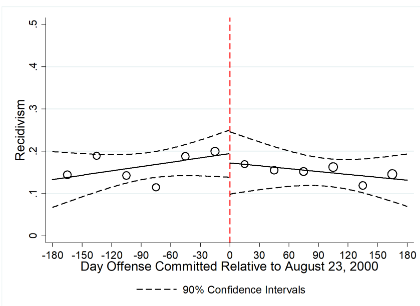
(b) 1998



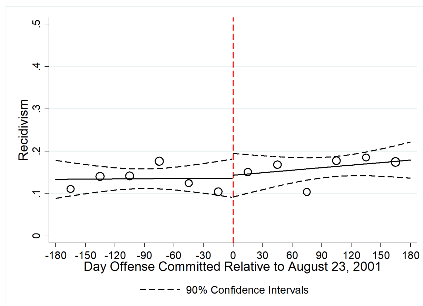
(c) 1999



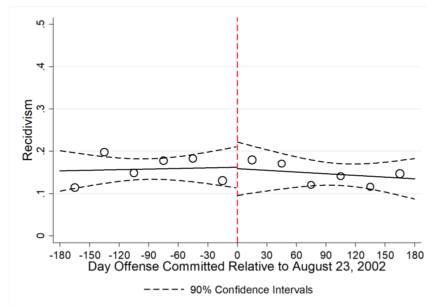
(d) 2000



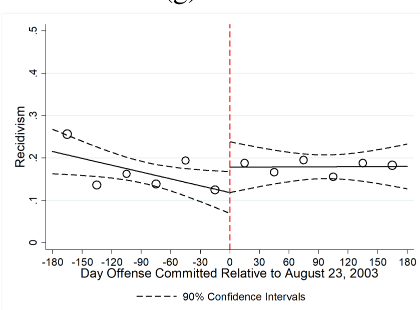
(e) 2001



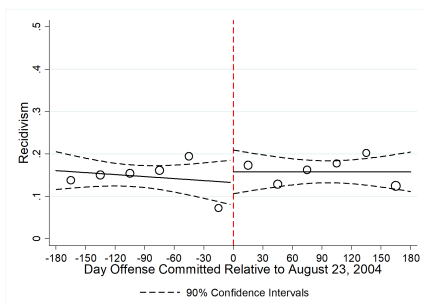
(f) 2002



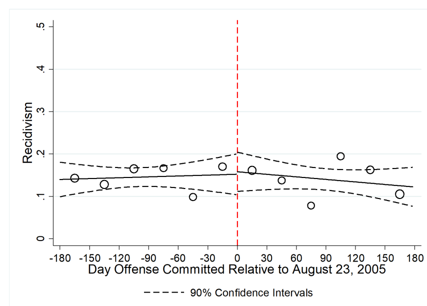
(g) 2003



(h) 2004



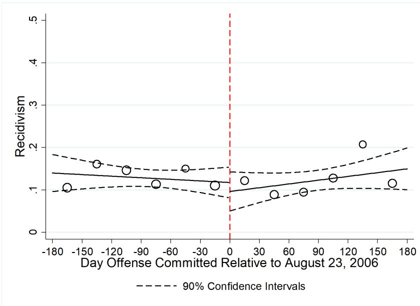
(i) 2005



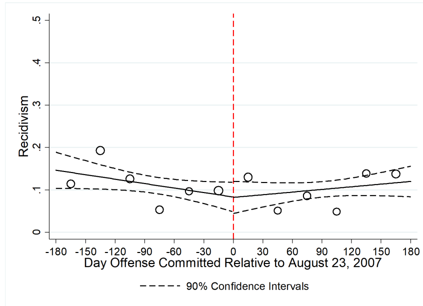
Notes: The figures above plot lines from local linear regressions of recidivism outcomes on the running variable (days before and after August 23 of a given year), estimated separately on each side of the cutoff. All figures are overlaid with a scatter plot of the recidivism averaged in 30-day bins. In these figures, the running variable is centered around placebo dates (dates that do not determine ban status). See Figure A2.4 for general notes about the creation of the RD plots for drug traffickers around August 23, 1996. These plots employ the same method but on a sample of offenders who commit drug trafficking around August 23 in the years 1997-2012.

Figure A2.12: Drug Traffickers in Other Years are Not More Likely to Recidivate

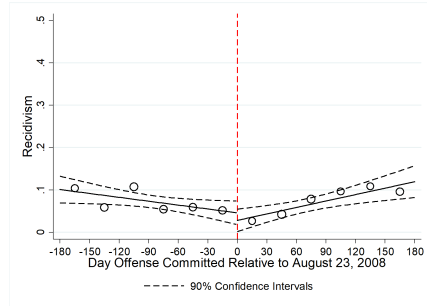
(j) 2006



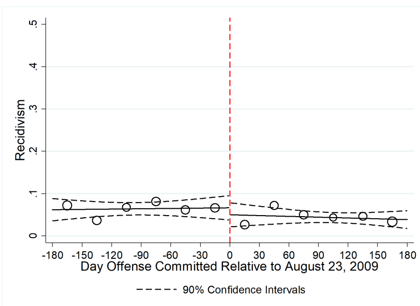
(k) 2007



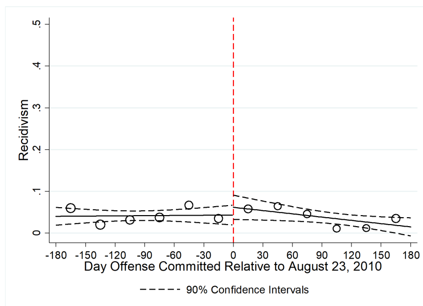
(l) 2008



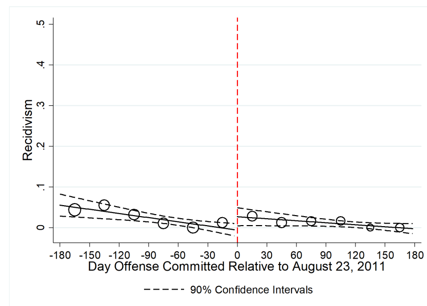
(m) 2009



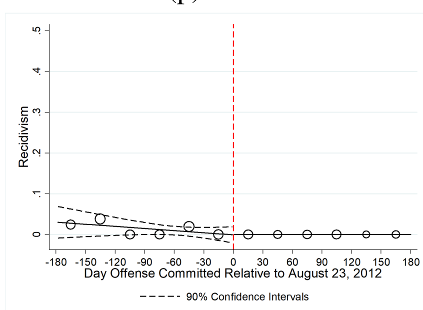
(n) 2010



(o) 2011

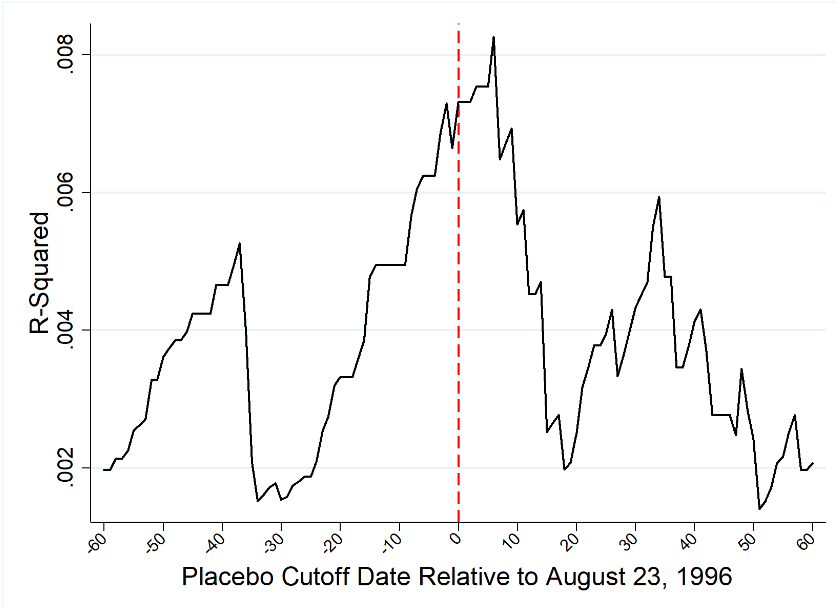


(p) 2012



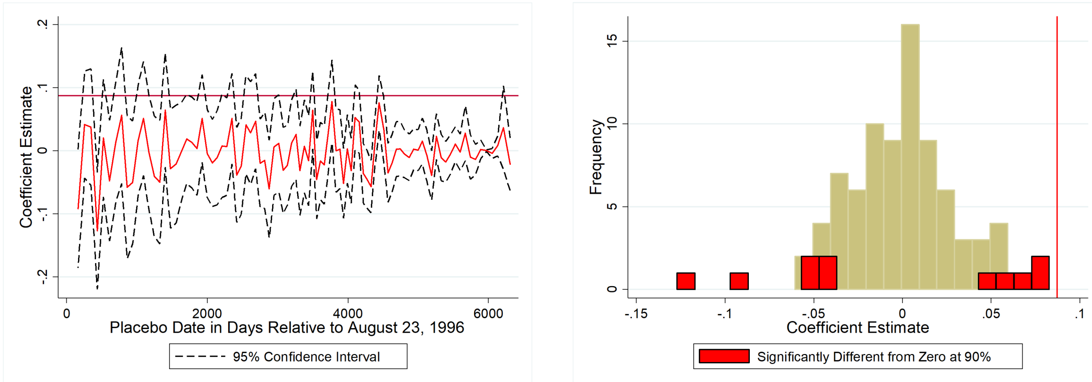
Notes: The figures above plot lines from local linear regressions of recidivism outcomes on the running variable (days before and after August 23 of a given year), estimated separately on each side of the cutoff. All figures are overlaid with a scatter plot of the recidivism averaged in 30-day bins. In these figures, the running variable is centered around placebo dates (dates that do not determine ban status).

Figure A2.13: Test for Other Significant Breaks in Bandwidth



Notes: The figure above follows Card, Mas, and Rothstein (2008) in identifying the “true” cutoff as determined by the data. To do this, I construct 120 placebo cutoffs (one for each of the 60 days before and after August 23, 1996). I then code placebo dummy variables for whether or not the offender committed their offense on or after each placebo date. Finally, I run 120 regressions of financial recidivism on each placebo dummy and plot the R-squared from each regression (no controls included). The “true” cutoff should have the highest R-squared. I detect the “true”cutoff at August 29, 1996 which is only six days from the date of the policy cutoff. The 15 days with the highest R-squared are all within nine days of August 23, 1996 and August 23, 1996 itself has the fifth highest R-squared.

Figure A2.14: Ganong-Jaeger Randomized Cutoffs Placebo Test



Notes: The figures above follow a randomization inference test outlined in Ganong & Jaeger (2015). To create these figures, I calculate the 5th-95th percentiles of the running variables—days before or after August 23, 1996. At every percentile, I construct a placebo cutoff and run 46 separate regressions of recidivism on a dummy for whether or not the offender committed the offense on or after the placebo date. From here, I plot the coefficient estimates and confidence intervals on the y-axis against the running variable on the x-axis in the first figure. In the second figure, I plot a histogram of the coefficient estimates (most are near zero) and highlight the estimates which are significant. In addition, I plot a vertical red line indicating the value of the coefficient at the true cutoff (August 23, 1996). Less than 10% of the placebo estimates are positive and significant.

2.10 Appendix B. Additional Information

Further Review of Related Literature.

A. Offender Reentry. Former offenders face a number of challenges when looking for legal work. First, many employers require employees to disclose criminal backgrounds on job applications and/or agree to criminal background checks. Pager, Western, and Sugie (2009) conduct an audit study in which they randomly assign a criminal background to some applicants. They find that applicants with criminal histories are half as likely to be called back by interviewers—this gap is even wider for black applicants. In recent years, offender advocates have encouraged cities and states to adopt laws that "ban the box" that asks applicants about criminal background. In fact, Shoag and Veuger (2016) show that after a city enacts "ban the box" legislation, employment from high-crime Census tracts increases.[1] In many cases, state occupational licensing laws only serve to exacerbate the troubles former offenders have in the legal labor market. Ex-felons are subject to more than 3,000 restrictive occupational licensing exclusions according to the American Bar Association (Council of Economics Advisors (CEA) 2016).

While the employment consequences associated with simply having a criminal background are large, incarceration and the prison experience can also negatively affect employment outcomes. For one, even if offenders are not explicitly tagged with their criminal backgrounds in the application process, many are left with large gaps in their work history as a result of their incarceration (Raphael 2011). Kroft, Lange, and Notowidigdo (2013) show that long-term unemployment in itself is penalized by potential employers. Incarcera-

[1] Agan and Starr (2018) find similar results to Pager, Western, and Sugie (2009) with a field experiment in which they sent applications to employers in New Jersey and New York City before and after "ban the box" went into effect. Employers who asked about criminal history in their sample were 62% more likely to call back applicants if they did not have a criminal record. The authors also point out the importance of statistical discrimination in this setting. Before "ban the box" went into effect, employers were 7% more likely to call back white applicants than black applicants, but this number balloons to 45% after "ban the box." It appears that "ban the box" may help offenders find work, but in doing so, it can diminish the employment prospects for young black men in general. This statistical discrimination spillover of "ban the box" policy is also explored by Doleac and Hansen (2016) who find that employment of young, low-skilled Black and Hispanic men decreases after "ban the box" takes effect in a metropolitan area.

tion may also prevent human capital accumulation, deteriorate bonds with legal job-finding networks, and/or create bonds with illegal job-finding networks (Bayer, Hjalmarrson, and Pozen 2009; Schmitt and Warner 2010). Mueller-Smith (2015) finds that an extra year of incarceration leads to a 4 percentage point drop in employment after release and a 30 percent decline in formal earnings. The stigma of a criminal background, the occupational restrictions, and the negative effects of incarceration are piled onto people who tend to have low education and low formal work experience even prior to incarceration, rendering them even less equipped to find legal work post-release (Raphael 2011).

Finding a job is not the only hurdle waiting for offenders as they transition back into their community. Once released, many offenders must navigate complicated and restrictive parole conditions that, if violated, could land them back in prison. Even more, offenders with families may return to a poverty-stricken or fractured homes—a family is 40% more likely to be in poverty when the father is incarcerated and incarceration increases probability of divorce or separation (CEA 2016). These stressors, among others, may contribute to the elevated mortality rate of offenders in the first couple of weeks after release, the majority of which is the result of drug overdoses (Schanzenbach et al. 2016)

Since offenders struggle to find legal work upon release, many reentry programs focus on increasing the employment prospects of offenders. In general, research has found mixed results on whether or not these programs are effective in curbing recidivism. Berk (2007) finds that work release does increase post-release earnings and that these earnings gains correlate with lower rates of re-incarceration but only for those offenders originally convicted of financially motivated crimes.[2] Another popular approach for helping offenders find legal work is through transitional employment programs. The National Supported Work (NSW) Demonstration, for example, provided a minimum wage job to ex-offenders for 12-18 months. Uggen (2000) finds the program decreased 3-year re-arrest rates for offenders above the age of 26 at the start of the program by about 20%. For younger

[2] Berk evaluates a work release program in Florida by comparing minimum custody inmates who participated in the program to minimum custody inmates who did not.

offenders, however, the program was ineffective.[3]

Still, other work has consistently found that offenders who face better labor market conditions upon release are less likely to recidivate. Schnepel (2018), for example, finds that the availability of "good jobs" (manufacturing and construction work) reduces recidivism for offenders released in California whereas availability of other low-wage jobs has no effect. Yang (2017b) also finds that being released in a time and place with good labor market conditions decreases probability of recidivism.

B. Financial Need and Crime. I find that offenders who are denied access to SNAP have higher rates of reincarceration. This result contributes to the literature above on prisoner reentry and recidivism, but it also adds to a long literature in economics and criminology that argues that financial motivations often underlie criminal behavior. In a seminal theoretical paper on criminal behavior, Becker (1968) points out the trade-off between participation in the legal labor market and the illegal labor market. Becker discusses how increased opportunities in the legal labor market could decrease participation in the illegal labor market. Most recent empirical investigations of the Becker model confirm this—Gould, Weinberg, and Mustard (2002) find that unemployment and wages for low-skilled men in a county are significantly related to crime in that county.[4]

Other empirical work also suggests that legal and illegal sector jobs may be substitutes. Mastrobuoni and Pinotti (2015), for example, find that recidivism (rearrest) and overall criminal activity decreases once immigrants become legal citizens, presumably because with citizenship comes many new job opportunities. The theoretical and empirical literature about legal opportunities and crime or recidivism suggests that financial need is a determinant of criminal behavior.

A nascent subset of this literature explores the effects of transfer programs on crime,

[3]Uggen evaluates the impact of the NSW by analyzing a randomized controlled trial in which some offenders were assigned to receive transitional employment while others were simply required to self-report employment and criminal information.

[4]Using Bartik-style instrumental variables, they show that higher unemployment leads to more crime and higher wages leads to less crime.

and supports the claim that financial need is a catalyst for criminal behavior. Chioda, Mello, and Soares (2015) estimate the effect of a conditional cash transfer in Brazil named Bolsa Familia. They find that as the number of children receiving the cash transfer from Bolsa Familia increases, crime decreases.[5] Similarly, Das and Mocan (2016) show that short-term employment from a public works program in India insures against negative income shocks, and as a result, decreases crime.

C. Transfer Programs and Labor Supply. In addition to the work on labor supply effects covered in the main text, Moore (2014) examines a PRWORA policy that removed drug and alcohol addictions as qualifying disabilities for DI. Moore uses this policy change in a difference-in-difference framework to determine the effect of DI on labor supply. Specifically, he compares people thrown off the DI rolls by this policy to people who had drug and alcohol addictions but were able to stay on DI for another condition. Moore finds that 22% of people removed from DI increase their labor supply to levels beyond the DI eligibility threshold. The effects of PRWORA and pre-PRWORA welfare waivers on outcomes such as labor force participation, welfare caseloads, and fertility/family structure are further documented (Blank 2002).

Hoynes and Schanzenbach (2012) also estimate labor supply effects for groups other than female-headed households. They find that the introduction of Food Stamps in a county causes a imprecisely estimated decrease in head of household annual earnings in noneldery households with low education. However, the authors find no change in hours worked and an increase in labor force participation. Focusing on female-headed households, the authors show that for those households all measures of labor supply decrease after the introduction of Food Stamps. For female-headed households, labor force participation falls by about 6 percent and this decline is even sharper for nonwhite female heads. The authors also find evidence of changes in labor supply along the intensive margin with female-headed households decreasing both hours worked and annual earnings. Their paper provides valuable

[5]The authors use the expansion of Bolsa Familia in 2008 and the demographic composition of schools to instrument for the number of children receiving funds from Bolsa Familia.

evidence about the labor supply response of female-headed households to Food Stamps, but evidence for the labor supply of males is limited, and there is no consideration of illegal labor supply.

Finally, I draw inspiration from Deshpande (2016), who estimates labor supply effects of Supplemental Security Income (SSI) child disability support. PRWORA required that children receiving SSI undergo a medical review at age 18 if their birthday occurred on or after August 23, 1996. Deshpande demonstrates that undergoing a medical review caused many kids to lose SSI benefits. Using the August 23, 1996 cutoff in a regression discontinuity design, she finds that 18-year-olds who lose SSI do increase their labor supply but not by enough to offset the loss of SSI. Her paper also uses one impactful piece of PRWORA to estimate the effect of transfers on labor supply.

Miscellaneous Details.

Throughout the paper, I focus on one specific definition of recidivism–return to prison. Recidivism has many definitions in the criminology literature. For example, recidivism can be defined as re-arrest, re-conviction, re-offense, and so on (Maltz 1984). In addition, recidivism is often defined with respect to some time frame (such as the 3-year or 5-year re-arrest rate). The definition I use in this paper is a return to a Florida prison for a new offense. I do not observe re-arrest, re-offense, or re-conviction. These events all occur more often than re-incarceration for a new offense. In Table A2.3, I show the results are robust to using 10-year, 8-year, and 5-year recidivism rates.

It's also worth noting that the crime for which an offender is convicted can feasibly differ from the crime which an offender committed. I observe the crime(s) for which the offender is convicted, which may not be the crime(s) they committed. For example, conviction crime and true offense crime may differ as a result of plea bargaining. That said, for the measure of treatment (the SNAP ban), only conviction crime and the date the offense was committed matters. In addition, the classifications financial and non-financial are broad—it is unlikely that slippage from offense crime to conviction crime will move a person from the financial to non-financial category (or vice versa).

Since the SNAP ban can be modified and repealed at the state level, offenders subject to the ban in one state could, in principle, move to another state and become eligible for SNAP. I do not find evidence that drug traffickers subject to the ban are more likely to migrate out of Florida and move away from the ban. Using the residence each offender plans to live at upon release (as reported on their release plan), I test for a change in the probability of that residence being outside of Florida. Offenders subject to the SNAP ban are not more likely to report a planned residence outside the state of Florida. Still, it is possible that offenders move to a place not listed on their release plan. In that case, the estimates in this paper will be attenuated.

While I provide numerous summary statistics on the offender population in Tables 2.1 and A2.1, I do not report the marital status of offenders because that information is not made publicly available in the OBIS database. This is potentially important for understanding how the SNAP ban affects ex-offenders. In 2013 and 2014, about 15% of Broward County jail inmates in Florida reported being married or having a significant other while the remaining 85% reported being single, divorced, separated, or widowed (ProPublica 2017). Unfortunately, to my knowledge, that is the best information available about marital status of Florida inmates.

In interpreting the main results, it is also important to consider the state's reentry policies/strategies. Florida abandoned its traditional parole system prior to 1995 and moved to a fixed sentencing system. With fixed sentencing (also known as structured sentencing or truth-in-sentencing), offenders must serve a certain percentage of their sentence (typically 80-90%). About 31% of offenders have some form of post-release supervision in Florida.

Finally, the regression used to create the risk score has a McFadden's R^2 of 0.20 and correctly predicts the recidivism outcome in 79% of drug trafficking cases within 212 days of August 23, 1996 (the IK optimal bandwidth for any recidivism). I can also calculate the risk score based on only those offenders subject to the ban and not in the ± 212 day IK bandwidth–the results do not change. I also test for heterogeneity in the effect by sentence length and by risk score. The coefficients are not statistically different from zero, but the point estimates imply that the effect of the ban on any recidivism is muted for riskier offenders and for offenders who serve longer sentences.

2.11 Appendix C. Conceptual Model of SNAP and Illegal Labor Supply

To more clearly illustrate the mechanisms described in the main body of the paper, I present a simple conceptual model. In the traditional static labor supply model with transfers, individuals choose $c =$ consumption and $l =$ leisure subject to $h =$ hours worked and $wh + y^{transfer} =$ total income to maximize utility:

$$
\begin{aligned}
\max_{c,l} u(c,l)\ s.t.\ c &= wh + y^{transfer} \\
l &= 1 - h
\end{aligned}
$$

This model is agnostic about whether h is supplied in the legal or illegal sector. For ex-offenders, this is an important distinction because they have ties to the illegal labor market, and they have difficulty finding work in the legal labor market. To highlight this distinction, I expand the model above to include$h^I =$ hours worked in the legal labor market and $h^L =$hours worked in the illegal labor market. In addition, I assume that individuals must satisfy a fixed level of consumption $\bar{c}$.

$$
\begin{aligned}
\max_{h^I,h^L} u(w^I h^I + w^L h^L + y^{transfer}, 1 - h^I - h^L)\ s.t.\ w^I h^I + w^L h^L + y^{transfer} &\geq \bar{c} \\
1 - h^I - h^L &\geq 0
\end{aligned}
$$

For simplicity, I further assume that ex-offenders face no additional cost of supplying illegal hours relative to legal hours. This implies that ex-offenders will optimally allocate all working hours to one sector. In general, I assume ex-offenders command a higher wage in the illegal labor market (w^I) than they command in the legal labor market (w^L)–this is a reduced form way of representing the difficulty of finding legal work versus illegal work for ex-offenders. When $w^I > w^L$ the maximization problem above reduces to the following:

$$\underset{h^I,h^L}{max} u(w^I h^I + y^{transfer}, 1-h^I) \; s.t. \; w^I h^I + y^{transfer} \geq \bar{c}$$
$$1-h^I \geq 0$$

Assuming that neither of the constraints binds, then differentiating the first order condition of the problem above with respect to $y^{transfer}$ and h^I yields the following comparative static[1]:

$$dh_I/dy^{transfer} < 0 \; iff \; w^I \times u_{11} - u_{21} < 0$$

Thus, for ex-offenders optimally consuming above $\bar{c}$ and working $h^I < 1$, a decrease in transfers will lead to an increase in hours worked in the illegal sector if leisure is a normal good.

For ex-offenders optimally consuming at $\bar{c}$ and working $h^I < 1$, we recover the following comparative static:

$$dh_I/dy^{transfer} < 0 \; iff \; w^I > 0$$

Notice that for these individuals, the response of h^I to a change in $y^{transfer}$ does not depend on preferences. For offenders consuming at $\bar{c}$, a decrease in transfers always leads to a increase in hours worked in the illegal sector.

Finally for those ex-offenders who are optimally working at $h^I = 1$, a decrease in $y^{transfer}$ will not induce an change in h^I; while these offenders may desire to increase h^I when $y^{transfer}$ falls, they cannot because of the constraint on their total time. In a more complex model, perhaps, even these offenders could respond by increasing the severity or "riskiness" of the crimes they choose to commit.

[1]The denominator of $dh_I/dy^{transfer} = \dfrac{-(w^I \times u_{11} - u_{21})}{w^I \times (w^I \times u_{11} - u_{12}) - (w^I \times u_{21} - u_{22})}$ is negative based on the second order condition.

For drug traffickers in Florida who committed their offense prior to August 23, 1996, total income is the sum of earned income and transfer income (including SNAP). Those drug traffickers who committed their offense on or after August 23, 1996 are denied SNAP benefits. Because of this, transfer income for those committing an offense prior to the cutoff date is higher than transfer income for those committing an offense on or after the cutoff date. The comparative statics above yield a clear prediction: ex-offenders who are banned from SNAP will optimally choose to work more hours in the illegal sector (when possible) than ex-offenders who are not banned from SNAP. I empirically test whether or not offenders denied SNAP increase illegal labor supply (measured as whether or not they are re-incarcerated for a financially motivated crime), and I find evidence that suggests that they do.

This model motivates two heterogeneity tests I conduct. I began the model by assuming that $w^I > w^L$to represent the difficulty that ex-offenders have in finding legal work versus illegal work. However, finding legal work (or increasing hours in the legal labor market) is more feasible for some ex-offenders than for others. For one, ex-offenders released during good legal labor markets may enjoy higher legal wages or may have an easier time finding legal work in general. Similarly, recall that Pager, Western, and Sugie (2009) find that offenders who are black face greater discrimination in the legal labor market than offenders who are white. To capture this in the model above, I assume that offenders released in good legal labor markets and offenders who are white are more likely to face $w^L > w^I$. The SNAP ban does not affect illegal labor supply in the model above when $w^L > w^I$, and thus, it should have less of an affect for groups more likely to face $w^L > w^I$.

To test the prediction regarding offenders released in good legal labor markets, I estimate the interaction between access to SNAP and state-level unemployment rate at the time of the offender's release. Taking this the data, I find noisy but positive estimates of the effect of state-level unemployment on offenders subject to the ban. This is consistent with the model above. When the unemployment rate is high, offenders are more likely to face

$w^I > w^L$ and thus, the effect of the ban should be larger. To test the prediction regarding race of the offender, I estimate the interaction between access to SNAP and whether or not the offender is black. In testing for heterogeneity by race, I find noisy results that suggest black offenders subject to the ban are more likely to recidivate than white offenders subject to the ban. Although these estimates are not statistically different than zero, the magnitude and direction are consistent with the model above.

Finally, the model suggests that when the disparity in $y^{transfer}$ between banned and non-banned offenders is greater, we should observe that the ban has a stronger effect. I use county-by-month variation in the work requirement imposed on Able-Bodied Adults Without Dependents (ABAWDs) to test how the effect of the ban differs when benefit generosity for the non-banned offenders is higher. The work requirement stipulates that unemployed ABAWDs may only receive SNAP benefits for three months out of every three years. If the ABAWD is employed more than 20 hours per week or is enrolled in a SNAP employment and training program, then they may receive SNAP benefits for more than three months. This requirement was waived nationally from 2009-2016. In addition, the requirement is waived for Labor Surplus Areas (counties in Florida with especially high unemployment) and for counties where Florida chooses to apply a special exemption that allows states to exempt 15% of the state's caseload from the requirement (the 15% exemption) (USDA 2016b).

Using information from the Florida Department of Children and Families from 1996-2016, I create a measure for each month and county in Florida indicating whether or not the work requirement for ABAWDs is waived. I then estimate the effect of the ban on the probability an offender recidivates at a time and place where the ABAWD work requirement is waived versus the probability an offender recidivates at a time and place where the ABAWD work requirement is in effect. I find that the effect of the ban is strongest when benefit generosity for the non-banned offenders is high, which is consistent with the conceptual model above.

The static labor supply model can be extended to a dynamic setting in which offenders search for jobs over time. In the dynamic model, suppose offenders face a cost of job search that decreases with time out of prison, but that they also receive financial support from family members that decreases with time out of prison (Western et al. 2015). If the cost of the job search is highest immediately after release, then SNAP benefits may be most vital in this transition period. However, if family support is also highest immediately after release, then SNAP benefits may be more important years later when family support has waned. The model yields an ambiguous prediction about when support from SNAP is most important. In addition, once the cost of searching is incorporated, the model predicts increased recidivism among banned offenders via two channels: (1) the banned offenders are given less transfer income and thus have an incentive to increase labor supply and (2) the non-banned offenders are given more transfer income and thus have assistance that may mitigate the cost of legal job search. In this paper, I do not distinguish between these two channels. However, given that over half of all offenders (many of which have access to SNAP) are unemployed even a year after release, it does not appear that the second channel plays much of a role.

2.12 Appendix D. Cost-Benefit Analysis of the SNAP Ban

Recall, cost per offender is defined as:

$$\begin{aligned}\text{Cost per Offender} &= [(\text{Marginal Cost of Year of Incarceration}) \times (\text{Mean Years Sentenced}) \\ &\quad \times (\text{Marginal Increase in Probability of Offending due to the Ban})] \\ &\quad + [(\text{Victim Cost}) \\ &\quad \times (\text{Marginal Increase in Probability of Offending due to the Ban})]\end{aligned}$$

In columns (1)-(4) of Table D2.1, I estimate the total societal cost of the SNAP ban. To be clear, this cost estimate is intended to highlight the potential benefit of reducing recidivism by providing SNAP or other financial support post-release. A more comprehensive cost-benefit analysis of the ban is beyond the scope of this paper, as it would require estimates of the effect on legal employment and the deterrence effect of the ban for would-be first-time traffickers. In this calculation, I only include the cost of incarcerating the offenders and the cost of victimization. To start, I assume that drug traffickers who return to prison are sentenced to about 3 years, a statistic supported by the data from Florida Department of Corrections. I use an estimate of the marginal cost of incarcerating an inmate for one year from the US Department of Justice ($9,600 per year) (US DOJ 2011) and I use an estimate of victimization costs from the National Institute of Justice ($11,000) (Miller, Cohen, and Wiersema 1996). All dollar values in this section are adjusted to 2016 dollars.

In columns (5) and (6), I estimate the net cost for taxpayers. In other words, I ignore the private benefit drug traffickers get from SNAP benefits. Introducing this assumption requires an additional assumption about how long a drug trafficker would spend on SNAP if given the opportunity. The average length of time spent on SNAP is about 10 months (USDA 2011). I assume that drug traffickers would spend about the same amount of time on SNAP as the average recipient. I also assume the average SNAP benefit for men in Florida is about $150—this is consistent with the summary statistics on SNAP benefits in Table 2.2. Again, in columns (5) and (6), I treat the SNAP funds not disbursed to drug traffickers as a benefit, this is a highly conservative assumption which assumes an extra dollar of SNAP would have no effect on the welfare of a former drug trafficker. In other words, we ignore the benefit of SNAP to drug traffickers and estimate only the cost to non-banned taxpayers. In that case, the benefit per offender is defined as the following:

$$\text{Benefit per Offender} = \text{Monthly Food Benefit} \times 12 \times \text{Mean Time on SNAP}$$

Table D2.1: Cost-Benefit Analysis of SNAP Ban

	Societal Cost				Taxpayer Cost	
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Time Served for Recidivating Offenders	3 years	3 years	3 years	3 years	3 years	3 years
Marginal Cost of Incarceration	\$9,600	\$9,600	\$9,600	\$9,600	\$9,600	\$9,000
Mean Months on SNAP	-	-	-	-	12	12
Monthly SNAP Benefit	-	-	-	-	\$150	\$150
Mean Cost of Victimization	0	\$11,000	0	\$11,000	\$11,000	\$11,000
Effect of SNAP Ban	1.7 pp	1.7 pp	9.5 pp	9.5 pp	1.7 pp	9.5 pp
Net Cost per Offender	\$490	\$677	\$2,736	\$3,781	-\$1,123	\$1,981

Notes: In the exercise above, "Net Cost per Banned Offender" is equal to the cost per banned offender minus the benefit. When calculating the taxpayer cost in (5) and (6), Benefit per Offender includes $Monthly\,Food\,Benefit \times 12 \times Mean\,Time\,on\,SNAP$ since taxpayers save that amount by denying drug traffickers SNAP benefits.

I assume that the effect of the SNAP ban on recidivism is approximately 1.7 percentage points in columns (1) and (2). In other words, for every 100 drug traffickers banned from SNAP, about 2 will recidivate because of the ban. This is the lower bound of the confidence interval on the main result in Table 2.3. This assumption yields my most conservative, traditional cost-benefit estimates. In the two columns that follow, I assume the effect of the SNAP ban is 9.5 percentage points—this is the point estimate from column (1) in Table 2.3, Panel B. In columns (2) and (4) above, I assume the cost of victimization is about \$11,000 dollars on average. This cost of victimization is within the range of victimization costs for burglary, robbery, and theft provided by the National Institute of Justice (Miller, Cohen, and Wiersema 1996). The National Institute of Justice does not estimate a cost of victimization for drug crimes. Since the National Institute of Justice focuses on the material costs of crime and risk of death in these estimates, this number is an underestimate of the true costs of victimization (which also includes psychic costs, such as fear or trauma). Again, this yields conservative estimates of the net cost per offender.

In most cases, I find that the SNAP ban costs the state of Florida a substantial amount of money per offender banned. Even assuming the lower bound for the effect of the SNAP ban, I find the societal cost of the ban in Florida is about $677 per banned offender. With approximately 19,000 banned offenders, this implies the ban has cost Florida over 12 million dollars to date. Assuming the ban increases recidivism by 9.5 percentage points (the point estimate from the main results), I find the societal cost of the ban in Florida is about $3,781 per banned offender or approximately 70 million dollars to date. This estimate ignores the cost to the families of drug traffickers, all costs of crime for Florida citizens, and many other criminal justice costs (enforcement, trials, etc.). It also assumes the ban has zero deterrence effect for potential drug traffickers and no effect on the legal employment margin for those banned.

To drive the estimated net cost to zero, we must focus on the cost to taxpayers, ignoring the private benefit that drug traffickers and their families receive from the transfer. If I assume that the drug traffickers, if not banned, would spend about 1 year on SNAP and that the SNAP ban increases recidivism by about 1.7 percentage points (the lower bound estimate), then the SNAP ban has a net benefit of $1,123 per banned offender. However, if we assume the SNAP ban increases recidivism by 9.5 percentage points (the point estimate), we recover a net cost of the SNAP ban of $1,981 per banned offender.

An important question that is beyond the scope of this paper is whether SNAP is the most efficient means of post-release financial support for reducing recidivism. Hendren (2017) suggests SNAP is highly inefficient in that it has potentially large negative labor supply effects. Hendren applies the estimate from Hoynes and Schanzenbach (2012) to the marginal value of public funds formula and finds that SNAP funds have a lower marginal value than funds spent on other programs. While a large decrease in legal labor supply may make SNAP less efficient than other programs, in general, it is not clear what the implication of that result is for SNAP and offender reentry. This paper argues that the decrease in recidivism is driven by a decrease in illegal labor supply. In that way, what makes SNAP less efficient generally (large labor supply response) may make it more efficient for reentry policy if the labor supply response of offenders is primarily on the illegal labor supply margin.

2.13 Appendix E. Data Construction

I use six separate datasets. First, the "Inmate Release Offenses CPS" and "Inmate Release Offenses Prpr" data include information about current and prior offenses, respectively, of released inmates. In addition, I link this data to the "Inmate Release IncarHist" dataset that details the admit and release date for each prison spell—this allows me to accurately calculate the time between release and the next offense. I also use data on active inmates, "Inmate Active Offenses" and "Inmate Active Offenses Prpr", to determine recidivism for those offenders who were released but returned and are currently serving a sentence. Finally, demographic information (age, sex, race) comes from the "Inmate Release Root" data. All datasets are publicly available from the Florida Department of Corrections. For the purposes of this paper, offender information such as full name, exact birthdate, or Florida offender ID are not necessary. Before beginning the data construction described below, I de-identify the data by assigning a new unique ID to each offender and by stripping the data of name and exact birthdate.

To construct the sample of offenders for the recidivism analysis, I start by combining the de-identified versions of "Inmate_release_offenses_CPS", "Inmate_release_offenses_prpr", "Inmate_active_ offenses_CPS", and "Inmate_active_offenses_prpr" from the FL OBIS Access database available here: `http://www.dc.state.fl.us/pub/obis_request.html` (downloaded on April 7, 2016). The combination of these tables is the totality of information that FL provides about released inmate offense history.[1] Next, I remove duplicate observations and offenses for which the adjudication was withheld.

After that, I manually tag drug trafficking offenses. I identify drug trafficking crimes by tagging offense types that contain the string "TRAFF" but do not contain the string "STOLEN PROPERTY", "HUMAN" or "SEX." Other crime categories are identified us-

[1]Florida also provides records about which offenders are currently under community supervision. Very few drug traffickers in my sample are in this dataset and offenders committing an offense after August 23, 1996 are not more ot less likely to be under community supervision. For this reason, I do not consider community supervision as a pertinent outcome.

ing a combination of manual string matching and an official categorization of offenses provided by the Florida DOC here: `http://www.dc.state.fl.us/AppCommon/offctgy.asp`. Exact strings used to identify specific crime categories are included in the data construction code.

Next, I collapse the data by offender ID, date of adjudication, and county of conviction, keeping the minimum date at each level and the maximum sentence length. For the trafficking offenses, I keep both the minimum and maximum offense date to insure that I am accurately classifying offenders as banned or not banned. If the resulting offense date for the offender does not equal the trafficking date, I replace it with the trafficking date—this is important since trafficking date determines treatment status, so I must measure this correctly. That said, this line of code affects a small number of observations. In general, I use the minimum trafficking offense date when necessary. However, I have estimated the main results using the maximum trafficking offense date, and this distinction does not matter. After that, I collapse further to the level of offender id, date of offense, and county of conviction. Again, I keep the minimum date and the maximum sentence length. And again, for trafficking offenses, I keep both the minimum and maximum date.

Next, I bring in the "Inmate_release_incarhist" table that includes information about the exact receipt and release date from prison. Since the previous data tables do not include receipt or release date to prison, I have to match offenders based on adjudication year and receipt year. This, naturally, will lead to some mistakes but I expect it is negligible. To do this matching, I drop duplicates at the level of offense ID and receipt year. Essentially, this means I leave out offenders who enter prison twice in the same year. This is not a big portion of drug traffickers or felony offenders in general. Next, I collapse the data by offender ID and receipt date. This yields a dataset in which each observation is a unique prison stay and in which the variables indicate all of the offenses associated with a given stay. I drop all observations with offense years before 1950 or after 2016.

Using this, I can calculate amount of time after release before an offender recidivates.

I calculate "time until recidivism" as the difference between the release from prison stay[t] and the offense date for offense[t+1], if offense[t+1] exists. A small number of observations have a negative time to recidivism because offenders occasionally are arrested for crimes committed prior to stay[t] after they are released. This is not correlated with treatment. I remove these offenses and recalculate time to recidivism.[2] Since the data only includes inmates released after October 1, 1997, I exclude any observations with release dates prior to October 1, 1997 when doing recidivism analyses. I also drop all offenders with reported "race" as "Hispanic" due to special restrictions non-citizen immigrants face after committing a felony and after PRWORA's restrictions on SNAP receipt. Unfortunately, immigrant status is not available in the data. Outside of this, there are no other major data cleaning steps, only variable construction and analysis. Data and code necessary to reproduce all analyses (in the main text and in the online appendix) are available on the AEA website for this paper.

Finally, when providing the public database of released offenders, Florida includes the following disclaimer which I pass along here, "The Florida Department of Corrections updates this information regularly, to ensure that it is complete and accurate; however this information can change quickly. Therefore, the information in this file may not reflect the true current location, status, release date, or other information regarding an inmate. This database contains public record information on felony offenders sentenced to the Department of Corrections. This information only includes offenders sentenced to state prison or state supervision. Information contained herein includes current and prior offenses. Offense types include related crimes such as attempts, conspiracies and solicitations to commit crimes. Information on offenders sentenced to county jail, county probation, or any other form of supervision is not contained. The information is derived from court records provided to the Department of Corrections and is made available as a public service to in-

[2] In the code, I keep a variable that codes recidivism based on whether or not an offender has a prison stay after they are released (even if that stay is for a crime committed before stay[t]. Using this variable as the dependent variable, I get the same results. Offenders who commit drug trafficking on or after August 23, 1996 are more likely to have a stay[t+1].

terested citizens. The Department of Corrections makes no guarantee as to the accuracy or completeness of the information contained herein. Any person who believes information provided is not accurate may contact the Department of Corrections. The Florida Department of Corrections is not responsible for misinterpretation or inaccurate reporting by entities or persons utilizing this information."

Chapter 3: Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums

3.1 Introduction

Racial differences in sentencing are a persistent concern in America. In recent federal cases, black offenders face sentences that are 20 percent longer than the sentences handed down for white offenders (United States Sentencing Commission (USSC) 2017). These added years are costly for society at large and for the people incarcerated. The Bureau of Prisons (BOP) estimates the direct care cost of incarcerating a person is about \$11,000 (in 2015 dollars) per year (US Department of Justice (DOJ) 2011). Mueller-Smith (2015) estimates an additional year in prison causes a 30 percent decrease in formal earnings post-release and significant lost wages while incarcerated. Even more, those incarcerated must confront serious physical and psychological costs of prison, in addition to the more intangible cost of their lost freedom (Haney 2001; The Hamilton Project 2016; BOP 2020). Due to racial sentencing disparities, these costs are disproportionately borne by black and Hispanic offenders.[1] For policy to confront these disparities, we must

[1] In the USSC variable **newrace**, four values are recorded for the offender's "race"–(1) non-Hispanic white, (2) non-Hispanic black, (3) Hispanic, and (4) other. As such, throughout the paper, I will frequently use the term "race" in reference to Hispanic ethnicity to be consistent with this terminology used in the USSC data.

understand the root causes. One explanation for disparate sentences is that people of different races are different *upon entry* into the criminal justice system. Another explanation, however, is that *after entry* into the system, people are treated differently by race.

In this paper, I examine racial sentencing disparities and test the second explanation: that agents in the criminal justice system (police, prosecutors, judges, etc.) treat black and Hispanic defendants differently than similar white defendants.[2] To do this, I focus on federal crack-cocaine cases and the application of mandatory minimum sentences. Approximately 20 percent of all federal drug cases involve a crack-cocaine offense, and racial sentencing differences are particularly large in these cases. In 2016, black and Hispanic crack-cocaine offenders were sentenced to over 6 years, on average, compared to only 3.5 years for white crack-cocaine offenders (USSC 2017). In addition, the structure of mandatory minimum sentencing and recent changes in crack-cocaine mandatory minimums provide a unique opportunity to study discretion and racial disparities in the criminal justice system.

In federal drug trafficking cases, a mandatory minimum sentence is triggered if the drug trafficking crime involves an amount of drugs equal to or above a threshold amount. This sentencing cliff generates strong incentives for law enforcement agents. Legal rules about police sting operations and the type of evidence admissible in federal court give both police and prosecutors power to influence the amount used in sentencing. If police or prosecutors want to increase the likelihood of a harsh sentence, they can use their discretion to move the amount of drugs to the threshold amount or just above

[2] I use the term "offender" to describe someone in the final sentencing data or someone who has committed an offense (e.g. when talking about offender responses to the Fair Sentencing Act). Otherwise, I use the term "defendant."

it. This paper studies whether police or prosecutors respond to this sentencing incentive and whether their responses are racially disparate. Specifically, I test for an excess mass (or bunching) of cases at and above the mandatory minimum threshold (i.e. the use of discretion to increase the likelihood of a harsh sentence) and for differences in the excess mass by race (i.e. a racial disparity in the use of discretion).

With the Fair Sentencing Act (FSA) in 2010, the 10-year mandatory minimum threshold for crack-cocaine was increased from 50g (i.e. 50 grams) to 280g.[3] Crack-cocaine is the only drug for which the federal mandatory minimum threshold has changed since the adoption of mandatory minimums in the 1980s. The shift to 280g is especially useful since the new threshold is set at a point with zero bunching prior to 2010. All other mandatory minimum thresholds are set at somewhat natural bunching points (50g, 500g, 1000g) that do not vary over time.[4]

Using this time variation in the mandatory minimum threshold, I implement a difference-in-bunching design where I first assume the pre-2010 distribution of drug amounts is a good counterfactual for the post-2010 distribution (i.e. what the post-2010 distribution would look like with the pre-2010 thresholds) (Kleven 2016). I find the fraction of cases bunched at and above 280g increases after 2010, and that the increase is

[3]The FSA also shifted the 5-year threshold from 5g to 28g. I focus on the higher, 10-year mandatory minimum threshold for drugs in this paper. There are two reasons why I do not study bunching at 28g of crack-cocaine (the lower, 5-year mandatory minimum threshold) in detail. First, 28g is below the pre-2010 10-year mandatory minimum threshold of 50g–this yields incentives for prosecutors to shift cases that would have been charged both above 50g into the 28-50g range and cases that would have been charged below 50g into the 28-50g range. Second, estimating whether the racial disparity in bunching is conditional on underlying observed drug amount requires a range below the threshold that is not subject to strategic sentencing incentives. This is a reasonable assumption for the 60-280g range pre-2010, but would not be a reasonable assumption for the 6-28g range pre-2010 because those cases may be bunched at 50g.

[4]These amounts exhibit bunching in all drug types, even for drugs where they are not the relevant thresholds. I expect this bunching is due to a "round number" bias by police, prosecutors, offenders, etc.

much larger for black and Hispanic offenders than for white offenders.[5] I then show further evidence that, under a few additional assumptions, this disparity in bunching at 280g is **conditional** on the observed drug trafficking of offenders and is not due to a difference in underlying observed drug trafficking by race.

To be clear, this is not intended as an evaluation of the FSA, which is likely responsible for a decline in sentences after 2010 (USSC 2015a). Rather, these results imply that police or prosecutors dampened the effect of the FSA by increasing the drug amount charged for some defendants. In addition, these results do not imply that the use of discretion or a racial disparity in the use of discretion began after 2010. Instead, I take the shift to 280g as an opportunity to detect these behaviors that are otherwise difficult to detect.

I use data at multiple stages in the criminal justice process to estimate who is responsible for the bunching at 280g. First, I use drug seizure records on quantities and prices and survey data on drug use and selling to show that offenders do not respond to the relaxed sentencing rules in a way that would induce this increase in cases at 280g (or the disproportionate increase by race). Second, since the bunching occurs in federal sentencing, it is possible that more cases with drug quantities at or above 280g are sent to federal court after 2010. I examine data on state-level drug convictions from Florida, and I do not find a shifting composition of cases after 2010. Third, local and federal law enforcement can influence the drug quantity involved in an offense by choosing amounts involved in sting operations. However, the data on drug seizures made by local and federal agencies do not show increased bunching at 280g after 2010.

[5] Note, I do not find evidence of bunching just below 280g for the drug amount used in sentencing. Moreover, comparing the pre-2010 and post-2010 distributions of crack-cocaine amounts suggests that these are cases that, had they been sentenced prior to 2010, would have been recorded below 280g.

Finally, prosecutors can legally influence the drug quantity involved in an offense because, according to the USSC Guidelines, the quantity of drugs used to determine sentencing is not strictly tied to the quantity found on the offender at the time of arrest (USSC 2015b). I do find bunching at 280g after 2010 in case management data from the Executive Office of the US Attorney (EOUSA). I also find that approximately 30% of prosecutors are responsible for the rise in cases with 280g after 2010, and that there is variation in prosecutor-level bunching both within and between districts. Prosecutors who bunch cases at 280g also have a high share of cases right above 28g after 2010 (the 5-year threshold post-2010) and a high share of cases above 50g prior to 2010 (the 10-year threshold pre-2010). Also, bunching above a mandatory minimum threshold persists across districts for prosecutors who switch districts. Moreover, when a "bunching" prosecutor switches into a new district, all other attorneys in that district increase their own bunching at mandatory minimums. These results suggest that the observed bunching at sentencing is specifically due to prosecutorial discretion.

The US Supreme Court issued a 5-4 decision in *Alleyne v. United States* on June 17, 2013 that changed the evidentiary standard necessary for facts that raise a defendant's exposure to mandatory minimum sentencing (Bala 2015). Previously, prosecutors could present evidence on drug quantities to the presiding judge, and the judge would decide, based on the preponderance of evidence, whether the mandatory minimum applied. The Supreme Court ruling in *Alleyne* requires that prosecutors present this evidence to the jury, which evaluates it based on the stricter "beyond a reasonable doubt" standard. The case management data from the EOUSA show that from 2011-2013, approximately 9.1% of cases were recorded in the range of 280-290g. From 2014-2016, however, 6.8% of cases were recorded in the 280-290g range. Using a difference-in-discontinuities design, I show that the practice of bunching ballooned in the run up to *Alleyne*, and that this bunching was reined in by the Supreme Court decision (though it was not eliminated entirely). This suggests prosecutors were submitting evidence under the judicial fact-finding system that would not hold up under the scrutiny of a jury.

After documenting a racial disparity in bunching at 280g and studying the role of prosecutorial discretion in producing that disparity, I then explore whether the racial disparity could be attributed to taste-based discrimination from prosecutors. Since prosecutor tastes are unobservable, I focus on testing alternative explanations and, in doing so, I demonstrate that taste-based discrimination remains a viable explanation after accounting for several alternatives. I introduce a simple model of prosecutor objectives and discuss four potential sources of the racial disparity. First, I explore the possibility that the racial differences in bunching at 280g are driven by another a factor correlated with race. I show that racial differences in bunching exist even among observably similar offenders. For example, the increase in cases at and above 280g for black and Hispanic offenders with a college education is larger than the increase for white offenders with a college education. This is true for interactions with individual characteristics such as sex, age, criminal history, and other elements of the current offense. It is also true for interactions with district-level characteristics such as fraction of offenders who are white, pre-2010 plea rates, and pre-2010 fraction of cases declined. Race is a consistent factor in determining the amount of bunching at 280g after 2010.

Next, I test whether the disparity could be the result of racial differences in costs to the prosecutor of charging a defendant with 280g. Costs to the prosecutor are determined by defense attorneys, judges, potential juries, and other actors involved with the case. First, I show that there is no difference in type of defense attorney retained by race for federal crack-cocaine cases. Second, the increase in bunching at 280g is similar in districts with high versus low pre-2010 rates of private counsel retention. Third, I show that bunching at 280g is unrelated to judge race or political party and that, unlike prosecutors, judges with a high share of cases at 280g post-2010 are not any more likely to have cases at 28g post-2010 or at 50g pre-2010. Fourth, the increase in bunching at 280g is similar in districts with high versus low fractions of cases declined due to "weak evidence" or "lack of resources." These analyses suggest the racial disparity in bunching is not caused by racial differences in defense counsel, that bunching is not related to judges or judge characteristics, and that costs of developing a case are not a major determinant of the rise in bunching at 280g.

Finally, I consider statistical versus taste-based discrimination. I show that the racial disparity in bunching can be almost entirely explained by a measure of state-level racial animus based on Google search data developed by Stephens-Davidowitz (2014). In other words, black and Hispanic offenders convicted in states with higher levels of racial animus are more likely to be bunched at 280g than white offenders convicted in those states. In states with lower levels of racial animus, however, black, Hispanic, and white offenders are all equally likely to be bunched at 280g. The persistent racial differences even after controlling for and interacting race with observables, the within-district variation in prosecutor-level bunching, and the correlation between the racial disparity in bunching and state-level racial animus all support a model of discrimination in which the disproportionate use of discretion is a result of prosecutor tastes. Of course, a more detailed model of statistical discrimination could incorporate those facts, and I cannot reject such a model.

Taken together, these results suggest a subset of federal prosecutors use their discretion to tag some defendants with drug amounts that will trigger mandatory minimum sentences, and that they do this disproportionately for black and Hispanic defendants. Even more, the decrease in bunching after the Supreme Court tightens evidentiary standards in *Alleyne* suggests these cases are reliant on relatively weak evidence. In several additional analyses, I rule out various explanations for why the racial disparity exists, but I am unable to rule out a simple model of taste-based discrimination.

Broadly, this paper adds to an extensive literature on racial disparities and discrimination in the criminal justice system (e.g. Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Grogger and Ridgeway 2006; Antonovics and Knight 2009; Anwar, Bayer, and Hjalmarsson 2012; Rehavi and Starr 2014; Pfaff 2017; Arnold, Dobbie, and Yang 2018; West 2018; Sloan 2019). The vast majority of papers on this topic focus on racial bias from police officers and test for bias in two ways: (1) using a version of the outcome (or hit-rate) test proposed by Becker (1957) or (2) by documenting same-race versus other-race bias.

Along with recent work by Anbarci and Lee (2014) and Goncalves and Mello (2018), I implement a new test for racial bias in criminal justice that uses insights from the bunching literature.[6,7] Both Anbarci and Lee (2014) and Goncalves and Mello (2018) study the prevalence of police officers discounting speeding tickets by race. They show substantial bunching just below the point where the fine increases. Both papers argue that this is a result of officer leniency and that officers exhibit racial bias in their leniency.[8] I contribute to this literature by examining racial bias from prosecutors (a relatively under-studied group), and by showing racial differences in bunching at the point where sentences increase.

This paper also contributes new evidence to the empirical literature on prosecutorial discretion and decision-making (e.g. Glaeser, Kessler, and Piehl 2000; Bjerk 2005; Boylan 2005; Shermer and Johnson 2010; Rehavi and Starr 2014; Yang 2017; Nyhan and Rehavi 2017; on defense attorneys: Agan, Freedman, and Owens 2018; Arora 2018; Carr and McClain 2018; Sloan 2019). Bjerk (2005), for example, finds that prosecutors are more likely to charge defendants with a misdemeanor if a felony charge would invoke a "three-strikes" sentence. Sloan (2019), using random assignment of prosecutors to cases in New York County, shows that being assigned to an opposite-race prosecutor increases a defendant's likelihood of conviction, particularly in property crime cases.

The most closely related work, Rehavi and Starr (2014), finds that black offenders

[6]Note, my paper is not the first to acknowledge the existence of bunching in the amount of drugs recorded in US federal sentencing or the possibility that it could be used as a test of prosecutorial discretion and discrimination. However, this paper is the first, to my knowledge, to take advantage of the time variation in the crack-cocaine 10-year mandatory minimum threshold to isolate bunching that is solely due to the prosecutor. In addition, I examine data at multiple stages in the criminal justice process and conduct several additional empirical tests that all suggest bunching is due to prosecutorial discretion and negatively affects minority defendants. Related work in this area is discussed in more detail in Section 3.2.1

[7]Recently, economists have also studied bunching around cliffs and notches in test scores as evidence of manipulation in educational settings. See Diamond and Persson (2017) and Dee et al. (2017).

[8]Anbarci and Lee (2014) show that white officers discount more for white drivers and black officers discount more for black drivers. Goncalves and Mello (2018) demonstrate that only some officers practice this leniency and that those officers are, on average, more lenient toward white drivers than minority drivers.

receive harsher sentences than white offenders arrested for the same crime. Using linked data from US Marshals, US courts, and US federal sentencing, they show that this disparity is driven by prosecutorial discretion over initial charging decisions, in particular, the decision to bring a charge with a mandatory minimum sentence.[9] In this paper, I provide novel evidence that prosecutors are selectively harsh by race using a new source of identification–the sharp change in the crack-cocaine mandatory minimum threshold. I argue that the sudden increase in cases just meeting that threshold is indicative of discretion, and that the burden of this discretion falls disproportionately on black and Hispanic offenders. Through a series of tests, I find that prosecutors are responsible for the increase of cases at 280g. In addition, I quantify the fraction of prosecutors exercising this type of discretion, and I show that this can be mitigated by increasing evidentiary standards.

Finally, the racial disparity in bunching at 280g has meaningful implications for the racial sentencing gap. Depending on the counterfactual sentence imputed for the affected offenders, bunching at 280g can account for 2-7 percent of the racial disparity in crack-cocaine sentences. A conservative estimate suggests that being bunched at 280g adds 1-2 years to an offender's sentence. Multiple estimates suggest the cost of incarceration (combining direct care costs and the cost of lost current and future wages for the offender) is approximately \$60,000 per person per year (Donohue 2009; Mueller-Smith 2015).[10] I find 3.6% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.2% of white crack-cocaine offenders. Assuming 3.6% and 1.2% of all drug cases from 1999-2015 were subject to similar discretion by race implies total costs of 1.3 billion

[9] Rehavi and Starr (2014) do not focus on racial disparities in drug offenses due to data limitations.

[10] The majority of inmates in the Survey of Inmates in Federal Corrections (2004) report earning formal wages in the month before arrest.

dollars for black and Hispanic offenders versus 148 million dollars for white offenders. In terms of incarceration, the disparity implies 21,000 years sentenced due to this discretion for black and Hispanic offenders versus 2,500 years sentenced for white offenders.

All of the calculations above are based on the amount of discretion and the disparity detected right at and above the 10-year mandatory minimum threshold for crack-cocaine. To the extent that prosecutors exercise similar discretion to push defendants just above 5-year mandatory minimum thresholds or exercise discretion in less obvious ways (pushing defendants far beyond thresholds, for example), the cost estimates will only be higher and the effect on racial sentencing differences will only be greater.

3.2 Institutional Background and Prosecutor Objectives

3.2.1 Institutional Background

The Fair Sentencing Act, Mandatory Minimums, and Drug Quantities

Debate about federal mandatory minimum policy has overwhelmingly focused on the disparity between the threshold amounts for crack-cocaine and powder-cocaine. Prior to 2010, the threshold for the crack-cocaine 10-year mandatory minimum was 50 grams whereas the 10-year threshold amount for powder-cocaine was 5000g, a 100-to-1 disparity. In part due to the recommendations of the USSC and in part due to the political climate, the threshold amounts for crack-cocaine were increased in August 2010 by the Fair Sentencing Act. The upper threshold was changed from 50g to 280g, and offenders sentenced after the Fair Sentencing Act are subject to the new threshold.[11] In this paper,

[11] It is not clear why 280g, in particular, was chosen. One potential reason is that lawmakers wanted to set the threshold at 10 ounces (283.495g), but in keeping with the convention of setting the threshold in

I use this change from 50g to 280g to study bunching at mandatory minimum thresholds and its relation to discretion and racial disparities in the criminal justice system.

This paper is not the first to acknowledge bunching in the amount of drugs recorded in US federal sentencing.[12] Bjerk (2017) briefly discusses bunching in the distribution of drug amounts, but posits that bunching arises from negotiation downward by prosecutors and defendants.[13] A 2015 Bureau of Justice Statistics (BJS) working paper on federal sentencing disparities also investigates the idea that prosecutors could "game" the drug weight sentencing guidelines (Rhodes, Kling, Luallen, and Dyous 2015). That paper provides a cursory look at bunching above mandatory minimum thresholds for all drugs by race, but does not address the bunching that is always present at round-number amounts (50g, 100g, 500g, etc.). As such, the authors conclude prosecutorial discretion in this form does not differentially affect black and Hispanic offenders.[14]

I depart from previous work in several ways. First, I show that excess mass at the threshold comes from cases below the threshold rather than above it. I also show that the bunching is more pronounced in trial cases, which suggests that drug amounts are being moved above the cutoff and not negotiated down to it. Second, I take advantage of the time variation in the crack-cocaine 10-year mandatory minimum threshold to isolate bunching that is solely due to prosecutor choices. Finally, I examine data at multiple

grams or kilograms, chose 280g as the closest "round" number.

[12] In concurrent work, Knorre (2017) finds evidence of bunching in reported drug amounts from Russian police. Knorre does not investigate potential discriminatory behavior or the consequences of the observed bunching.

[13] Since Bjerk's paper focuses on sentencing consequences of mandatory minimums for all drug types, he does not empirically investigate the cause of the observed bunching in crack-cocaine offenses. In addition, he does not compare outcomes before and after the Fair Sentencing Act of 2010.

[14] The working paper is an extensive and excellent treatment of sentencing disparities. In that light, it is reasonable that the authors did not do a "deep dive" on this "bunching" test, which is a small piece of the broader paper.

stages in the criminal justice process and conduct several empirical tests that all suggest prosecutorial discretion negatively affects minority defendants.

Procedural Background

In Figure A3.1, I illustrate a simplified timeline from arrest to sentencing. Arrests are made by local or federal police, and after arrest, cases are handled by state or federal prosecutors. Prosecutors decide whether to try the case in court. Federal arrests typically stay in the federal system, but local arrests can be shifted to federal court or tried in both state and federal court. A case tried in federal court can end in conviction, acquittal, or dismissal. For convictions, a probation officer, partly in consultation with the prosecutor, prepares a pre-sentence report (PSR) that details facts relevant to sentencing. At sentencing, the judge considers statements from the prosecution, the defense, and the PSR to make factual determinations (e.g. the amount of drugs involved) and decide the defendant's sentence. In 2015, approximately 70% of drug arrests referred to federal prosecutors were prosecuted and 90% of those prosecuted ended in a conviction (BJS 2016). The drug quantity used in sentencing can be influenced at many of these stages. Below, I describe the legal discretion that police and prosecutors have over the drug quantity.

First, police can influence drug amounts by choosing the amount of drugs involved in "reverse sting" operations (operations in which agents will sell drugs to offenders) or by extending traditional sting operations (operations in which agents will buy drugs from offenders) until the total transacted amount is above the threshold (Honold 2014). Outside of these two levers, it is unlikely that law enforcement agents across multiple agencies could systematically manipulate drug amounts since evidentiary protocols require the precise logging and controlled storage of evidence.

Second, prosecutors can influence drug amounts because mandatory minimum sentencing is determined by the amount of drugs the offender is responsible for trafficking, which is not strictly based on the amount of drugs they are holding at the time of arrest (Honold 2014; USSC 2015b; Lynch 2016). For one, prosecutors can rely on the testimony of informants or law enforcement to establish "historical weight," the amount of drugs a defendant is responsible for outside of the actual drugs seized (Lynch 2016). In addition, mandatory minimums also apply to drug trafficking conspiracy crimes in which the total amount trafficked by the group in question can be applied to all members of the group (Lynch 2016). The USSC Guidelines (2015b) specifically state, "Types and quantities of drugs not specified in the count of conviction may be considered in determining the offense level. Where there is no drug seizure or the amount seized does not reflect the scale of the offense, the court shall approximate the quantity of the controlled substance."

Criminologist Mona Lynch has compiled compelling qualitative evidence about the reach of federal sentencing guidelines in her book *Hard Bargains*. Lynch finds that prosecutors use informants to establish "relevant" quantities, and she interviews a prosecutor about how relevant quantities can be established: "The actual heroin sales directly tied to Mr. Samuels and his son were of 1g and 4g, respectively; the rest was arrived at on the mere say-so of confidential informants. [...] She told me that she could have established enough historical weight, through those (conspirators) she had 'flipped,' to get Mr. Samuels to at least a ten-year mandatory minimum sentence, if not more."

In Section 3.5.3, I examine data from a national survey on drug use/selling, state-level convictions, local police agencies, the Drug Enforcement Administration, and the Executive Office of the US Attorney to estimate the source of the bunching at 280g. I also conduct several tests in Sections 3.5.3-3.5.5 to rule out alternative explanations related to the role of offenders, state courts, police, defense attorneys, probation officers, and judges. Ultimately, I find evidence that prosecutorial discretion leads to bunching at 280g in the case of drug trafficking.

3.2.2 Prosecutor Objectives

Prosecutors have discretion over the drug quantity charged in federal drug trafficking cases. In addition, the data suggests prosecutors exercise this discretion and that they exercise it differentially by race. In this section, I discuss the literature on prosecutor objectives from the fields of economics, criminology, and law–all of which admit self-interested and/or biased prosecutors.

Then, in light of the literature on prosecutor objectives, I discuss how sentence-maximizing prosecutors may respond to the Fair Sentencing Act. Prosecutors may desire high sentences due to career concerns, beliefs that long sentences are ideal (for retribution or future deterrence), or to wield them as tools in plea bargaining.[15] Although this conceptual discussion describes prosecutor objectives as homogenous, I ultimately find that only a subset of prosecutors behave in this way.

Related Literature

Since the 1970s, economists have produced several theoretical models of plea-bargaining based on prosecutor objective functions. This work began with the canonical economic model of the courts from Landes (1971), which assumes that prosecutors

[15] Mandatory minimums also provide certainty about sentence length. Thus, in this context, prosecutors who desire certain sentences will behave similarly to prosecutors who desire long sentences.

maximize the expected sum of sentences subject to resource constraints. Following Landes (1971), several papers emerged modeling resource-constrained prosecutors trying to achieve an ideal punishment for guilty parties and no punishment for innocent parties (Grossman and Katz 1983; Reiganum 1988; Bjerk 2007; and Baker and Mezzetti 2011).

Empirical work finds that prosecutors are, in part, career-focused (Glaeser, Kessler, and Piehl 2000; Boylan 2005). Boylan (2005) shows that for US attorneys longer sentences are associated with positive career outcomes (appointed to a federal judgeship or hired by a large private firm). In addition, recent work demonstrates partisan bias (Nyhan and Rehavi 2017) and racial bias (Rehavi and Starr 2014; Sloan 2019) in prosecutorial decisions, suggesting that prosecutors may seek harsh punishments for some offenders and lenient punishments for others.

These findings that prosecutors can be self-interested and biased are echoed and often-times preceded by insights from criminologists and legal scholars.[16] Discussions of prosecutorial discretion in law reviews frequently note that career-oriented prosecutors focus on securing lengthy sentences or high conviction rates (Bibas 2004; Simon 2007; Barkow 2009; Sklansky 2017). Stuntz (2004) argues that prosecutors lean on harsh sentences to secure guilty pleas. He even specifically notes the usefulness of sentencing guidelines (e.g. mandatory minimums) in this regard: "plea bargains outside the law's shadow depend on prosecutors' ability to make credible threats of severe post-trial sentences. Sentencing guidelines make it easy to issue those threats."

[16] Officially, the EOUSA cites Berger v. United States, 295 U.S. 78 (1935) to describe the role of the US attorney as an agent "[...] whose interest, therefore, in a criminal prosecution is not that it shall win a case, but that justice shall be done. [...] the twofold aim of which that guilt shall not escape or innocence suffer. [...] It is as much his duty to refrain from improper methods [...] as it is to use every legitimate means to bring a just one." However, the quote offers a description of the prosecutorial ideal rather than the reality. In fact, the case in Berger v. United States, is itself a case about prosecutorial misconduct.

Finally, criminologists and political scientists have also documented prosecutorial bias along race, gender, and partisan lines (Spohn, Gruhl, and Welch 1987; Mustard 2001; Farrell 2003; Ulmer, Kurlychek, and Kramer 2007; Gordon 2009; Shermer and Johnson 2010; Fischman and Schanzenbach 2012; Ulmer, Painter-Davis, and Tinik 2014; Franklin and Henry 2019; King 2019). Fischman and Schanzenbach (2012) show that sentence lengths are concentrated at mandatory minimums, that this concentration grows when judges are given more discretion over other aspects of sentencing, and that the increase in bunching at mandatory minimum sentence lengths is especially large for black and Hispanic offenders. Farrell (2003) and Ulmer, Kurlychek, and Kramer (2007) both use state court data to show that black offenders are more likely to receive a mandatory minimum penalty than white offenders, even after conditioning on several aspects of the offense. Ulmer et al. (2007) conclude, "prosecutors have great influence through charging, sentence bargaining, and, in the case examined here, the application of mandatory minimums. [...] Too often, studies of sentencing and sentencing discretion focus on judges and leave out prosecutors, crucial players in the courtroom work groups."

Prosecutor Responses to the Fair Sentencing Act

This work from economics, criminology, and law suggests that prosecutors will value crossing the mandatory minimum threshold in drug cases (for sentence length and/or sentence certainty) and that they will value it differentially by race (due to racial bias). By law, cases above the mandatory minimum threshold must receive a sentence of at least five or ten years (increased certainty), and in practice, longer sentences are handed down in cases just above the threshold (increased sentence length; see Section 3.5.2). Assuming that gathering new evidence to raise the drug quantity charged beyond the amount seized is costly and that the cost is increasing in the amount of new evidence gathered, these objectives yield predictions for how prosecutors will behave in the face of mandatory minimum thresholds and how they will behave when those thresholds change.

Prior to 2010, the mandatory minimum thresholds in federal court for crack-cocaine were 5g (for a five-year mandatory minimum sentence) and 50g (for a ten-year mandatory minimum sentence). After 2010, these thresholds shift to 28g and 280g. The shift in mandatory minimum thresholds after 2010 should lead to the following relative changes: (1) an increase in the density from 0-5g, (2) an ambiguous change in the density from 5-28g, (3) an increase in the density from 28-50g, (4) a decrease in the density from 50-280g, (5) an increase in the density from 280-290g, and (6) no change in the density above 290g. Note, for these ranges, and whenever ranges are listed, the upper bound of the range is not inclusive. See Figure A3.2 for an illustration of these changes.

These changes should occur because some cases worth bunching at 5g or 50g before 2010 will also be worth bunching at 28g or 280g after 2010 and some will no longer be worth it. Also, some cases that were not bunched at 5g or 50g before 2010 will be worth bunching at 28g or 280g after 2010. In Section 3.5.5, I introduce a simple model of prosecutor objectives to motivate a discussion about the racial disparity in bunching. I use that model in Appendix C to formally discuss why the changes described above should occur. In Section 3.5.1, I show that the empirical evidence is consistent with this simple conceptual model of prosecutor responses to the shifting thresholds.

This conceptual discussion and the empirical analysis that follows is rooted in broad ideas about prosecutor bias and prosecutors' desire for long sentences and/or certain sentences, but it also captures a specific phenomenon that has received some attention in law and criminology–federal prosecutors using sentencing guidelines and mandatory minimums to secure guilty pleas or harsh sentences (Stuntz 2004; Honold 2014; Lynch 2016). As noted in the previous section, Honold (2014) and Lynch (2016) explicitly acknowledge prosecutors exploiting legal rules about the type of evidence admissible in drug mandatory minimum cases to secure longer and more certain sentences.

In 1983, legal scholar and eventual judge Frank Easterbrook wrote, "Rules could command, for example, that all cases involving a sale of cocaine weighing more than 50 grams be prosecuted and all others not. Rules of this sort produce the arbitrary and unexpected consequences so well known to tax and welfare lawyers; it is far from clear that one can design rules to achieve a particular end. People will change their conduct to take advantage of lacunae." Since then, such rules have been implemented, but researchers have paid scant attention to the ways people have changed their conduct to take advantage of them. In this paper, I document changing conduct by prosecutors that disproportionately affects black and Hispanic defendants–behavior that has been discussed and researched qualitatively by legal scholars and criminologists but that has remained relatively unexplored empirically.

3.3 Data

To estimate the degree of bunching at the 10-year mandatory minimum threshold, I use data on federal cases that include the amount of drugs recorded at sentencing. I then bring in several other datasets from different stages in the criminal justice process to estimate who is responsible for the bunching at 280g.[17]

Figure A3.1 shows a simplified timeline from arrest to sentencing and describes how the data I use is related to each step. This timeline also acknowledges that selection into/out of the data can occur at each step. As Knox, Lowe, and Mummolo (2019) discuss, bias in selection into the dataset of interest can distort the ultimate measure of bias. My empirical approach takes any bias in selection as given, and assumes this bias does not change sharply in 2010. I show evidence to this effect: drug selling and crack-cocaine usage does not increase after 2010, drug quantities seized do not increase after 2010, and the composition of cocaine offenses in state/local convictions does not change after 2010. Penalties remain high for offenses involving less than 280g, suggesting that there is little reason for selection into federal sentencing to change pre- versus post-2010. Also, Rehavi and Starr (2012) use linked data to show that the probability a case is filed in federal court and the probability a defendant is convicted is the same for black and white defendants (conditional on arrest). Finally, as long as selection into the data is biased in favor of white defendants (i.e. police are more lenient with white defendants or prosecutors are more likely to dismiss white defendant cases), then the estimate of the racial bias in this paper will be an underestimate.

[17] I am not able to link defendants/offenders across these datasets. However, given the nature of the findings and the information available in each dataset, analyzing them independently is sufficient to show where the bunching first occurs and to rule out alternative explanations. Finally, a dataset of defendants/offenders linked from arrest to sentencing does exist, but the codebook for that data suggests it does not include a measure of drug quantity seized at arrest.

3.3.1 United States Sentencing Commission (USSC) Data

To estimate the degree of bunching at or above 280g, I use data provided by the USSC on recorded drug amounts in all federal drug cases sentenced from 1999-2015.[18] I focus on cases that involve a crack-cocaine offense since that is the only drug for which the mandatory minimum threshold changes over time. Approximately 7.8% of offenders in this sample are labeled as white, 10.6% as Hispanic, and 81.6% as black. Table 3.1 summarizes additional information about age, education, citizenship, and details about the offense, all of which are used as covariates in later analyses (see Appendix D for further details on this dataset and others).

I restrict these data to cases in which the amount of drugs is non-missing and is not recorded as a range. Approximately 20% of cases are excluded for this reason, but the fraction of missing cases for crack-cocaine does not change discontinuously at 2010, though it does increase in 2013 and 2014. Furthermore, in Appendix A, I show that including cases coded as a range only exacerbates the degree of bunching and the racial disparity in bunching. I also remove cases that are flagged for having data issues with the drug quantity variable and cases where the court does not accept or changes the findings of fact. Less than 2% of cases are excluded for these reasons.

Using the cleaned data, I plot two histograms (Figures 3.1a-b) that zoom in on the density around 280 grams for the years before and after 2010. Prior to 2010, the density around 280g is smooth. After 2010, however, 280g becomes the new mandatory mini-

[18]These amounts are derived from pre-sentence reports prepared by a probation officer and in consultation with the defendant, the defendant's counsel, and the prosecuting attorney. In the event the court rejects an amount in the pre-sentence report, the new amount is recorded in the statement of reasons report and reported in the USSC drug quantity field.

mum threshold and in that same time, the number of cases at and above 280g spikes.[19] Figures 3.1c-d display how the fraction of cases recorded as 280-290g changes over time. This shows even more clearly that the spike in cases at 280-290g coincides exactly with the policy change. These figures also highlight the racial disparity in bunching at the threshold that occurs after 2010.

3.3.2 Additional Data

In addition to data on federal sentences from the USSC, I incorporate several other datasets to understand the source of the bunching in drug trafficking cases. I describe these datasets here.

Florida State Inmate Database

These data include the year an offender is convicted, a description of the offense, and the offender's race. In Florida, drug offense descriptions typically include the name of the drug involved, and occasionally, the descriptions include a range for the amount of drugs involved (these broad ranges are: 0-28g, 28-200g, 200-400g, and 400+g). Also, Florida does not separately categorize crack versus non-crack cocaine offenses and instead describes all such drug offenses as "cocaine."[20] The fraction of all cocaine cases from 200-400g still exhibits a sharp increase in the USSC federal data, and thus, a mirrored decrease should be detectable using the broad categories in Florida. Summary statistics for these data, the NIBRS drug seizures, and the DEA drug exhibits are reported in Table A3.1.

[19] See Figure A3.3 for a plot of the histogram from 0-500g.

[20] Data from Missouri Department of Corrections indicates that, in Missouri, approximately 80% of state-level cocaine offenses are crack-cocaine offenses.

National Incident Based Reporting System] (NIBRS) Property Segment

The FBI collects data from local law enforcement agencies about crime, and many agencies report this data at the incident-level. The incident-level reports make up the data in the NIBRS property segment. These data are submitted voluntarily by agencies and thus, are not representative of national or state-level crime. For this reason, I use a balanced panel of agencies from 2000-2015. Upon receipt, the FBI checks the reports for errors and contacts agencies for corrections if necessary. The property segment of this database includes information about drug seizures and drugs involved in arrests.[21] The offender segment of this database includes information on offender race, sex, and age for all offenders involved in the incident.[22]

DEA System to Retrieve Information from Drug Evidence (STRIDE)

The STRIDE database contains information about all drug evidence from the DEA and other agencies that was submitted to DEA laboratories for analysis. I obtained the data from a Freedom of Information Act request for all records pertaining to the drug "cocaine" from 2000 to 2015. This information includes the year and month the drugs were acquired, the weight of the drugs in grams, the type of drug (cocaine, cocaine hydrochloride, cocaine base, etc.), drug potency, and the price from undercover purchases.

Executive Office of the US Attorney (EOUSA), Caseload Data

The EOUSA releases case-level data on cases (excluding certain redacted cases)

[21] See Shively (2005) and Bibel (2015) for a discussion of well-known issues with NIBRS data, such as reporting and measurement of hate crimes and sexual assault, differential coverage, and data quality. To the best of my knowledge, there are no known issues with the drug quantity field of the NIBRS property segment.

[22] For tractability, I limit the offender segment to incidents that involve 5 or fewer offenders. This covers 99% of all incidents. Also, the fraction of incidents with 5 or fewer offenders does not meaningfully change after 2010 (99.1% in 2000, 99.1% in 2005, 99.0% in 2010, and 99.3% in 2015). Finally, it is not correlated ($\rho = 0.0001$) with the probability an incident involves 280-290g of crack-cocaine.

processed by the US Attorney's office. These data are derived from information entered into the Legal Information Office Network System (LIONS) case management system. The EOUSA notes that each district may use LIONS differently, and as such, the data should not be used to make cross-district comparisons. The analyses using these data are robust to the inclusion of district fixed effects and various methods of accounting for missingness in the drug quantity data (a data quality issue that varies across districts). The EOUSA data includes a wealth of information about drug cases and other cases, including type of drug, quantity of the drug, an ID for the lead attorney on the case, and an ID for the judge on the case. Summary statistics are reported in Table A3.2.

National Survey on Drug Use and Health (NSDUH)

The NSDUH is a survey of non-institutionalized US civilians aged 13 or older that primarily asks questions about drug use and mental health. The respondents are randomly sampled based on state and age, with larger states and younger individuals oversampled. I use two questions asked from 2002-2016: (1) "have you ever, even once, used crack-cocaine?" and (2) "during the past 12 months, how many times have you sold illegal drugs?" These data provide detail about drug use and drug selling that is not based on interactions with law enforcement.

Google Trends Data on Racial Animus from Stephens-Davidowitz (2014)

To measure racial animus at the state-level, I use data introduced by Stephens-Davidowitz (2014). Stephens-Davidowitz uses Google search data from 2004-2007 (accessed via the Google Trends tool) and measures relative search volume in every US state for a specific racial slur and its plural form. Since Google searches are virtually anonymous, this measure may provide a less filtered view of racial attitudes than common survey measures. In fact, it is positively correlated with racial animus as measured by implicit association tests or questions about interracial marriage from the General Social Survey.[23] Even more, it is highly predictive of President Obama's vote share in the 2008 and 2012 US elections (Stephens-Davidowitz 2014). The construction of the measure is covered in much greater detail in Stephens-Davidowitz (2014).

Implicit Association Test (IAT) Data on Racial Animus for Lawyers

The IAT data from Project Implicit (Xu et al. 2019) contains the results of implicit association tests for racial bias for over 3 million individuals. The implicit association test for racial bias is designed to test how strongly a person links black people with the concept of "bad" and white people with the concept of "good." This is accomplished by having a person sort words into "good" and "bad" categories, sort people into "black" and "white" categories, and finally, sort both words and people into "black" and "white" categories paired with "good" or "bad" categories. The time it takes to sort into "black/good" relative to "black/bad" and "white/bad" relative to "white/good" is the basis of a person's score. See "Project Implicit" for more detail. Although recent research casts doubt on the validity of the IAT for detecting bias (Oswald et al. 2013), the data has two advantages. First, it can be aggregated to the federal district, a sub-state geography. Second, it can be calculated solely for people reporting an occupation of "Lawyers, Judges, and Related Workers."

[23]It is also correlated at the Census region level with responses to these questions from respondents with a graduate degree. This suggests it is not solely reflective of racial animus from people with low levels of education. See Figures A3.4a-i.

3.4 Methodology

This paper has four main goals. First, to quantify the bunching at 280g after 2010 and the racial disparity in bunching at 280g. Second, to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of observed evidence or a difference in the likelihood a case is bunched **conditional** on the observed evidence (i.e. a **conditional racial disparity**). Third, to estimate who causes the bunching at 280g after 2010. And fourth, to explore and test various explanations for the racial disparity in bunching, including discrimination. In this section, I detail methodology for the first three goals. I reserve the discussion of potential discrimination and related tests for Section 3.5.5.

Throughout, I use what Kleven (2016) terms the "difference-in-bunching" method. This approach estimates the degree of bunching by comparing the actual distribution to an empirical counterfactual distribution. To estimate bunching at 280g and the racial disparity in bunching, the ideal counterfactual is the post-2010 distribution with the pre-2010 thresholds. I assume the pre-2010 distribution is a good counterfactual in this sense for all parts of the drug quantity distribution. Section 3.4.1 details the estimation of bunching and the racial disparity under this assumption.

To estimate a **conditional** racial disparity in bunching at 280g, the ideal counterfactual is the post-2010 distribution with no mandatory minimum threshold (or any other incentive to increase the amount charged). I assume the pre-2010 distribution is a good counterfactual in this sense for the part of the drug quantity distribution above 50g. Section 3.4.2 outlines tests for a conditional racial disparity under this assumption.

Finally, to estimate who causes the bunching at 280g, I test for changes in drug quantity at multiple stages in the criminal justice process leading up to sentencing. Here, again, the assumption is that at each of these stages the pre-2010 distribution is what the post-2010 distribution would be if the thresholds had not changed. Thus, I use the same methods detailed in Section 3.4.1. In the Results section, I detail methodology and results for several additional analyses.

3.4.1 Bunching at 280g and Racial Disparity in Bunching

I define a case as "bunched" at 280g as any case in the narrow range 280-290g (not including 290g). I then compare the fraction of cases from 280-290g in the post-2010 distribution of drug weights to the fraction of cases from 280-290g in the pre-2010 distribution. Specifically, I estimate the following linear probability model:

$$(\text{Charged}\,280-290\text{g})_{it} = \alpha + \beta\,\text{After2010}_{it} + Z_i + g(t) + \varepsilon_{it} \tag{3.1}$$

where $(\text{Charged}\,280-290g)_{it}$ is equal to one if offender **i** in year **t** is charged with 280-290g and is equal to zero if the offender is charged with less than 280g or equal to or above 290g.[24] After2010_{it} is equal to one if the offender **i** in year **t** is sentenced in 2011-2015 and is equal to zero if the offender is sentenced in 1999-2010. β is the change in an offender's probability of being charged with an amount in the narrow 280-290g range as a result of being sentenced after the threshold amount is increased to 280g. Z_i represents case-level covariates (such as offender education, race, age, conviction state, etc), and $g(t)$ represent time trends. In most specifications, I limit the sample to 0-1000g

[24] State conviction data does not include precise drug weights. In those cases, I use the dependent variable (Convicted with 200-400g), equal to one if the offender is convicted with 200-400g and equal to zero otherwise.

to remove extreme outliers and exclude Z_i and $g(t)$, however I show that the result is robust to altering this sample range and robust to including numerous controls.

To estimate heterogeneity in bunching by race, I extend the model as follows:

$$\begin{aligned}(\text{Charged}\,280-290\text{g})_{it} = \alpha + \beta(\text{After2010} \times \text{White})_{it} \\ + \delta(\text{After2010} \times \text{BlackOrHispanic})_{it} + \text{BlackOrHispanic}_{it} + Z_i + g(t) + \varepsilon_{it}\end{aligned} \tag{3.2}$$

Now, β represents the change in a white offender's probability of being charged with 280-290g as a result of being sentenced after the threshold is increased, and δ represents the change for black and Hispanic offenders.[25]

Models (1) and (2) quantify the excess mass at 280-290g by using regression analysis on the case-level microdata and comparing the pre- and post-2010 distributions. This follows work by: Kleven et al. (2011), Behagel and Blau (2012), Sallee and Slemrod (2012), Chetty, Friedman, and Saez (2013), Dwenger et al. (2016), Goncalves and Mello (2018), and Traxler et al. (2018). This approach is also appropriate for the empirical setting. I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change in probability differs by race. In addition, some analyses in the paper preclude aggregating the data into bins because they rely on data that do not include precise drug quantities.[26]

[25]Combining black and Hispanic offenders into one category, although common in analyses of the criminal justice system, is a crude categorization. Splitting these groups into separate variables yields similar results. There is a larger increase in bunching for black offenders than white offenders and a larger increase for Hispanic offenders than white offenders. The increase in bunching is similar for black and Hispanic offenders. In a model with district-by-time effects and a limited set of offender-level controls, the increase for Hispanic offenders is slightly larger than the increase for black offenders, although the two estimates are not statistically different (p-value=0.1426). For expositional reasons, I combine these groups throughout the paper. However, it is worth noting that these groups' experience with law enforcement and with discrimination in the US, in general, is varied and complex in a way that is not accounted for in this analysis (RWJF 2018).

[26]In Appendix B, I show that the results in this paper are robust to alternative methods of quantifying bunching above the threshold. One approach, introduced by Saez (2010) and Chetty et al. (2011), constructs

To understand where the excess mass at 280-290g comes from (i.e. where the post-2010 distribution has less mass relative to the pre-2010 distribution), I estimate a series of models similar to the equation (3.1) that replace the dependent variable with different drug quantity ranges:

$$(\text{Charged X-Yg})_{it} = \alpha + \beta \text{After2010}_{it} + Z_i + g(t) + \varepsilon_{it} \tag{3.3}$$

In these models, β represents the change in an offender's probability of being charged with an amount of drugs between X and Y grams as a result of being sentenced after the threshold is increased. I estimate equation (3.3) for 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 280-290g, 290-470g, 470-600g, and 600-1000g. The prosecutor objectives discussed in Section 3.2.2 yield specific predictions about many of these ranges–an increase in the 0-5g and 280-290g ranges and a decrease in the 50-60g, 60-100g, and 100-280g ranges.[27] The missing mass analysis addresses a critical question for policy implications: how would offenders who were charged with 280-290g post-2010 have been charged pre-2010? If those offenders would have been charged below 280g, then the bunching at 280-290g post-2010 may represent an effort to increase sentence lengths for some offenders.

a high-order polynomial counterfactual density from the actual bunched density. Kleven (2016), however, notes that this standard bunching estimation is typically used in settings where there is no variation in the kink/notch, and calls this a "minimalist approach" that "may not be compelling in all contexts." Additionally, he argues "more sophisticated alternatives exist that require richer data and/or richer variation." The Fair Sentencing Act in 2010 provides richer variation in this setting. A second alternative approach takes advantage of that variation by aggregating the post-2010 distribution and the scaled pre-2010 distribution into 10g bins and comparing them directly in levels. The results in this paper are robust to both.

[27] In Appendix A, I report the analysis by race for more narrow ranges. Since the ranges involved are much wider than the previous bins, I include a time trend (centered at zero in 2011) and state fixed effects to account for broad differences in drug trafficking over time and across states. In some specifications, I also estimate the "jump" in the probability of being below or above the 280-290g range after 2010. This approach yields similar results, and it is discussed in more detail in Appendix A.

3.4.2 Racial Disparity Conditional on Observed Drug Behavior

Now, I outline the assumptions necessary to estimate whether the racial disparity in bunching at 280g is due to differences in the underlying distributions of observed evidence or a difference in the likelihood a case is bunched **conditional** on the observed evidence.

Institutional Setting

Consider a simplified criminal court setting with drug cases, prosecutor discretion over amount charged, and mandatory minimum sentences. Assume the **seized evidence** s in a case is drawn from a discrete distribution $G_r(.)^t$ that is specific to each race r and time-period t (pre- vs. post-2010). The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than s by collecting additional evidence $a-s$. Seized evidence s is a noisy measure of true drug trafficking.

I observe the amount charged a. Publicly available data from the USSC does not report the seized evidence s for each case, and true drug trafficking is unknown to the researcher and the prosecutor. The prosecutor chooses a based on a variety of factors. The first goal of the empirical analysis is to identify racial disparities in a conditional on s (i.e. a **conditional racial disparity**). The second goal (addressed in Section 3.5.5) is to model under what conditions the disparity reflects discrimination by prosecutors and to conduct empirical tests of that model.

In this section, I detail the identifying assumptions necessary to estimate the conditional racial disparity. The set-up closely follows Goncalves and Mello (2018) who use a difference-in-bunching design to estimate police officer bias in speeding tickets. For now, consider the prosecutor's objective a function of tastes (including racial biases), career concerns, the sentence that would be justified under law if true drug trafficking were observed, and costs associated with building the case.

The amount of drugs charged a maps onto a **mandatory minimum sentencing schedule** $l(a)^t$ that differs pre-2010 $t = 0$ and post-2010 $t = 1$.

$$l(a)^t = \begin{cases} 1 & \text{if } a < mm_L^t \\ 5 & \text{if } mm_L^t \leq a < mm_U^t \\ 10 & \text{if } mm_U^t \leq a \end{cases} \tag{3.4}$$

If a is below the lower threshold for time period t, the defendant is sentenced to 1 year. If a is equal to or above the lower threshold but below the upper threshold, the defendant is sentenced to 5 years. If a is equal to or above the upper threshold, the defendant is sentenced to 10 years. A mandatory minimum does not, by law, require a discontinuous increase in sentence length at the thresholds. In practice, sentences do jump at 50g pre-2010 and 280g post-2010.

Given the seized evidence s (unobserved in the data but observed by the prosecutor) in the case and the defendant's race r (observed in the data), the prosecutor charges a final amount a (observed in the data) that is equal to a mandatory minimum threshold $mm = \{5, 28, 50, 280\}$ (i.e. "bunching" at the threshold) with a **bunching probability** $Pr(a = mm|s, r)^t$ (unobserved in the data). Finally, let defendants be in one of two broad **race** categories: white $r = w$ or black/Hispanic $r = bh$.

Defining the Conditional Racial Disparity

Now, I define a racial disparity in the amount charged a conditional on s and outline key equations.

There is a **conditional racial disparity** in bunching at 280g after 2010 if $Pr(a = 280|s, bh)^1 > Pr(a = 280|s, w)^1$. In other words, a conditional racial disparity exists if a black or Hispanic defendant with amount seized s is more likely to be bunched at 280g than a white defendant with the same amount seized s.

I observe the final amount charged, which can be written for the following ranges as:

$$Pr(a = j|r)^t = \begin{cases} \begin{cases} (a) \quad Pr(s = 50|r)^0 + \sum\limits_{k<50} Pr(s = k|r)^0 \times Pr(a = 50|s = k, r)^0 & \text{if } j = 50 \\ (b) \quad Pr(s = j|r)^0 & \text{if } 50 < j \end{cases} & \text{if } t = 0 \\ \begin{cases} (c) \quad Pr(s = j|r)^1 \times (1 - Pr(a = 280|s, r)^1) & \text{if } 50¡\ j < 280 \\ (d) \quad Pr(s = 280|r)^1 + \sum\limits_{k<280} Pr(s = k|r)^1 \times Pr(a = 280|s = k, r)^1 & \text{if } j = 280 \\ (e) \quad Pr(s = j|r)^1 & \text{if } 280 < j \end{cases} & \text{if } t = 1 \end{cases} \tag{3.5}$$

Equations **(3.5a)** and **(3.5b)** express the probability a case is charged with a given amount a prior to 2010. First, the probability a defendant is charged with an amount a equal to 50g is equal to the probability the seized evidence s is 50g plus the likelihood that a case with s under 50g gets moved up to 50g **(eqn. 3.5a)**. Second, since there is no sentencing benefit of charging an amount above 50g, the probability a case is charged above 50g **(eqn 3.5b)** is equal to the probability s is equal to that amount.

Equations **(3.5c)-(3.5e)** express the probability a case is charged with a given amount

a after 2010. The probability a case is charged with an amount below 280g and above 50g **(eqn. 3.5c)** is equal to the probability that s is equal to that amount and that the case does not get moved up to 280g given the amount s. The probability a case is charged with 280g **(eqn. 3.5d)** is equal to the probability s is 280g plus the likelihood that a case with s under 280g gets moved up to 280g. As in **(eqn. 3.5b)**, the probability a case is charged above 280g **(eqn 3.5e)** is equal to the probability that s is equal to that amount. Throughout, I assume that prosecutors don't suppress evidence, i.e. $a \geq s$.[28]

Difference-in-Bunching Estimator and the Conditional Racial Disparity

To estimate whether $Pr(a = 280|s, r)^1$ differs for black/Hispanic vs. white defendants, I compare the distribution of amounts charged after 2010 to the distribution of amounts charged prior to 2010.

Under the assumption that $Pr(s = k|r)^0 = Pr(s = k|r)^1$–i.e., the probability a case with a defendant of race r has seized evidence $s = k$ does not change pre- vs. post-2010–the **difference-in-bunching coefficients (eqn. 3.2)** $\delta - \beta$ yields the following:

$$\begin{aligned}\delta - \beta = [\sum_{k<280} Pr(s = k|bh) \times Pr(a = 280|s = k, bh)^1] \\ - [\sum_{k<280} Pr(s = k|w) \times Pr(a = 280|s = k, w)^1]\end{aligned} \tag{3.6}$$

$\delta > 0$ and $\beta > 0$ imply that prosecutors increase a in response to the Fair Sentencing

[28] In reality, it is possible for prosecutors to reduce the drug amount charged or choose not to pursue a drug charge entirely. Introducing this possibility means the disparity in bunching could be due to: (1) a difference in underlying observed drugs, (2) a conditional disparity in bunching, or (3) a conditional disparity in suppressing. The empirical evidence I show is consistent with (2) and (3), both of which are disparities conditional on underlying observed drugs. For that reason, I focus on the simpler case.

Act, and $\delta - \beta > 0$ implies that they increase a more for black and Hispanic defendants. This alone is of interest–it shows that prosecutors use their discretion to increase sentences in response to the FSA and that the burden of this falls on minority defendants. However, $\delta - \beta > 0$ could be driven by different underlying distributions of seized evidence s (i.e. different $Pr(s = k|r)$) or by disparate treatment conditional on s (i.e. different $Pr(a = 280|s, r)^1$–a conditional racial disparity).

The goal of this section is to outline how to test whether $\delta - \beta > 0$ is due to a conditional racial disparity. I detail two tests. For the first test, $\delta - \beta$ can be rewritten as follows:

$$\delta - \beta = \overbrace{[\sum_{k \leq 50} Pr(s = k|bh) \times Pr(a = 280|s = k, bh)^1 - \sum_{k \leq 50} Pr(s = k|w) \times Pr(a = 280|s = k, w)^1]}^{H}$$
$$+ \overbrace{[\sum_{50 < k < 280} Pr(s = k|bh) \times Pr(a = 280|s = k, bh)^1 - \sum_{50 < k < 280} Pr(s = k|w) \times Pr(a = 280|s = k, w)^1]}^{I} \quad (3.7)$$

First, I test whether the H term can explain $\delta - \beta > 0$. I observe $Pr(a = 50|r)^0$ and $Pr(a = 50|r)^1$. Equation **(3.5)** implies that:

$$Pr(a = 50|bh)^1 - Pr(a = 50|bh)^0 = -[Pr(s = 50|bh) \times Pr(a = 280|s = 50, bh)^1]$$
$$- [\sum_{k < 50} Pr(s = 50|bh) \times Pr(a = 50|s = k, bh)^0] \quad (3.8)$$

Under the assumption that $Pr(a=50|s,r)^0 \geq Pr(a=280|s,r)^1$ for all $s \leq 50$, equation **(3.8)** is greater than the $\sum_{k\leq 50} Pr(s=k|bh) \times Pr(a=280|s=k,bh)^1$ term from equation **(3.7)**. Thus, if the sum of equation **(3.8)** and $\delta - \beta$ is greater than zero, then the term H cannot explain $\delta - \beta > 0$. In other words, the shift **from** 50g for black and Hispanic offenders is an upper bound for the movement **to** 280g that can be explained by amounts seized at 50g or below. If this shift is not enough to explain the racial disparity in bunching at 280g, then the racial disparity must be due to term I.

Second, I test whether racial differences in $\sum_{50<k<280} Pr(s=k|r)$ from term I can explain $\delta - \beta > 0$. From equation **(3.5b)**, $Pr(a=k|r)^0 = Pr(s=k|r)^0 \forall 280 > k > 50$. Thus, I can test if $\sum_{50<k<280} Pr(s=k|w)^0 = \sum_{50<k<280} Pr(s=k|bh)^0$ by testing if $\sum_{50<k<280} Pr(a=k|w)^0 = \sum_{50<k<280} Pr(a=k|bh)^0$. In other words, if the distributions of pre-2010 charged amounts from 50-280g are approximately equal by race, then the racial disparity in bunching must be due to a racial disparity in the probability a case is bunched at 280g **conditional** on the seized evidence.

Now, I turn to the second test for a conditional racial disparity. The assumptions above also imply:

$$Pr(a=50<k<280|r)^1 = Pr(s=k|r)^1 \times (1 - Pr(a=280|s=k,r)^1) \tag{3.9}$$

$$Pr(a=50<k<280|r)^0 = Pr(s=k|r)^0 \tag{3.10}$$

The difference between equation **(3.9)** and **(3.10)** by race can be estimated as follows:

$$(\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010}\times\text{BlackOrHispanic})_{it} + \gamma\text{After2010}_{it} + \lambda\text{BlackOrHispanic}_i + \varepsilon_{it} \tag{3.11}$$

The coefficient $\delta^X = Pr(a = 280|w,s)^1 - Pr(a = 280|bh,s)^1$. Then, $\delta^X < 0$–i.e., black and Hispanic defendants are more likely to be shifted away from a given amount X after 2010–implies that there is a racial disparity in amount charged a conditional on the underlying evidence seized s.

3.5 Results

3.5.1 Main Results

Primary Bunching Estimates and Robustness

Using final sentencing data from the USSC, I estimate the effect of being sentenced after 2010 on whether an offender is sentenced for a drug amount between 280-290g. Column 1 of Table 3.2 indicates that offenders sentenced after the threshold increases to 280g are more likely to be charged with amounts just above 280g. An offender sentenced after 2010 is 3.5 percentage points more likely to be charged with a drug amount between 280-290g. Column 2 shows that this increase in bunching is driven by black and Hispanic offenders, who are approximately three times as likely to be charged with 280-290g after 2010 compared to white offenders. Figures 3.1a-d display graphical evidence of bunching at 280-290g and the racial disparity in that bunching.[29]

This result is robust to various sample restrictions (e.g. limiting to post-2006 years);

[29]Figures A3.5-A3.7 and B3.1-B3.4 present alternative ways to visualize this phenomenon. In particular, Figure A3.6 shows that the total number of cases at 280-290g increases after 2010.

the inclusion of state fixed effects, time trends, state-specific time trends, and offender-level controls (e.g. education, criminal history, age, etc.); clustering standard errors at the state-level; the use of Logit/Probit/Poisson models instead of a linear probability model; wider definitions of the bunching range (e.g. 280-380g); and the inclusion of cases with weights coded as range. See Tables A3.3-A3.7 for these results. I also conduct a simple bounding exercise in Table A3.8 that accounts for potential substitution into other drug types or selection into the case's drug weight being coded as a range. Table A3.9 presents a difference-in-differences analysis of bunching using other drug types for which the mandatory minimum threshold did not change. These additional tests confirm the main results. Offenders sentenced after 2010 are more likely to be charged with 280-290g, and this increase is disproportionately large for black and Hispanic offenders.

Source of the Excess Mass at 280g

To understand the reason for this bunching at 280g, I analyze other parts of the drug quantity distribution. If the excess mass in 280-290g after 2010 comes from above 290g, bunching may be the result of negotiation between prosecutors and defendants (Bjerk 2017). However, if the excess mass comes from below 280g, it is possible that prosecutors are shading amounts upward to exceed the threshold and secure longer and/or more certain sentences.[30]

In Table 3.3, I show the change in the probability of being recorded in several

[30] To be clear, it is impossible to say with certainty that the "missing mass" in the distribution is where cases in the "excess mass" would be recorded had they been sentenced prior to 2010. This is true for nearly all bunching analyses (panel bunching designs that follow the same unit over time are more convincing in this respect). As is typical in bunching analyses, I assume that the missing mass is indicative of where the "excess" cases would be located in the counterfactual. This is not guaranteed by the research design. Instead, this is another piece of suggestive evidence that the bunching is a result of cases being shifted in a way that is consistent with a simple conceptual model of prosecutor behavior and the empirical evidence of no offender response.

different ranges: 0-5g, 5-28g, 28-50g, 50-60g, 60-100g, 100-280g, 290-470g, 470-600g, and 600-1000g. Table 3.3 shows that the probability a case is recorded in those ranges matches the conceptual discussion in Section 3.2.2.[31] In Figures A3.7a-i, I plot the share of cases over time in each of these ranges. I estimate the regressions in Table 3.3 by race in Table A3.10a. The results are similar but noisier since it requires cutting the already narrow ranges by race. Table A3.10b and Figures A3.7j-k shows results by race using broader ranges: 0-280g and 290-1000g. In Table A3.10c, I re-estimate Table 3.3 including only years from 2007-2015, and I find similar results.

Summing the coefficients in columns 4-6 of Table 3.3 implies that the change in probability from 50g-280g can account for 87% of the increase in the 280-290g bin. Is it possible that some offenders charged with 280-290g post-2010 would have been charged below 50g prior to 2010? A fixed cost of evidence-gathering could explain this behavior. For example, if an offender is arrested with 10g of physical evidence prior to 2010, it may not be worthwhile to collect evidence to push them from a 5-year sentence to a 10-year sentence. After 2010, however, that same offender would face a 1-year sentence without additional evidence-gathering. Once prosecutors pay the fixed cost to gather evidence, it may then be worthwhile to gather enough evidence to reach the 10-year sentence.

Finally, I examine the degree of bunching in the subset of cases that go to trial.

[31] Although it is not clear from these analyses, there is excess mass at 50g (the pre-2010 threshold) even after the threshold changes in 2010. This persistent excess mass at 50g is likely due to round-number bias from offenders, police, or prosecutors. The powder cocaine distribution, which never has a mandatory minimum threshold at 50g, exhibits similar excess mass at 50g. For crack-cocaine, the fraction of cases from 50-60g is about 1.5 times the fraction of cases from 40-50g. For powder cocaine, that ratio is similar–the fraction of cases from 50-60g is about 1.7 times the fraction of cases from 40-50g. While conventional bunching estimation would address the presence of round-number bias by accounting for it in the estimation of the polynomial counterfactual, the difference-in-bunching method accomodates round-number bunching directly because that bunching will be present in both the counterfactual (pre-2010) and actual (post-2010) distributions (Best et al. 2018).

If the bunching is a result of lenient prosecutors rounding down, we should expect less bunching in trial cases where incentives for leniency are muted. However, the degree of bunching and the racial disparity in bunching is only heightened in trial cases (see Column 3 of Table 3.2). In fact, the only cases with 280-290g that go to trial are those of black and Hispanic offenders. As before, the increased bunching is accompanied by a falling share of cases below 280g ($\beta = -0.109$ and $SE = 0.022$) and a small, rising share of cases above 290g ($\beta = 0.034$ and $SE = 0.019$).[32] This is further evidence that the observed bunching is a result of shading up rather than negotiating down. In Section 3.5.3, I show additional evidence from prosecutor case management data that cases bunched at 280g would likely be recorded below 280g in the absence of strategic prosecutor behavior around the mandatory minimum threshold.

Estimating the Conditional Racial Disparity in Bunching at 280g

The results above indicate that there is a racial disparity in bunching at 280g. However, those results alone are not enough to understand why there is a racial disparity in bunching. It could be that there are different underlying distributions of observed drug behavior by race. For example, suppose black and Hispanic defendants are more likely to be arrested with 200g and white defendants are more likely to be arrested with 100g. If defendants with 200g are more likely to be moved to 280g, then a racial disparity will emerge. On the other hand, suppose that among defendants with 200g, black and Hispanic defendants are more likely to be moved to 280g–this would imply there is a disparity in bunching conditional on observed drug amount.

Section 3.4.2 outlines the assumptions and empirical tests necessary to estimate the

[32]See Table A3.10d for missing mass results using trial cases only.

conditional racial disparity in this setting. I conduct both tests outlined in that section, and both tests suggest that the disparity in bunching is driven by a conditional racial disparity rather than racial differences in the underlying distribution of observed drug amount.

The first test relies on decomposing the potential bunching at 280g. For the first part of that test, I estimate the racial difference in the shift away from the 50-60g range. Table 3.4 reports this result. Black and Hispanic offenders are less likely to be charged with 50-60g after 2010. However, the decrease in the 50-60g range is not large enough to explain the racial disparity in bunching at 280g. Adding the decrease from 50-60g for black and Hispanic offenders in column (1) to the increase to 280-290g for black and Hispanic offenders in column (2) yields a new bunching coefficient of 0.0293. The new coefficient is still about three times larger than the coefficient for white offenders, and it is statistically different from the coefficient for white offenders at the one percent level (p-value = 0.003).

For the second part of the first test, I test whether the distributions of charged amounts from 60-280g are equal by race prior to 2010. Figure 3.2a plots the distributions by race, and they are very similar. A Kolmogorov-Smirnov test of equality fails to reject the null that the distributions are equal (p-value = 0.788). Alternative evidence from drug seizure records confirms black and white offenders are seized with similar amounts (see Table 3.6a and Figure A3.8a-b). Since the racial disparity in bunching at 280g cannot be accounted for by racial differences in movement from 50g or by racial differences in the distribution from 60-280g, this implies the disparity is a conditional racial disparity.

The second test for a conditional racial disparity in bunching relies on estimating racial differences in movement away from other narrow ranges. Figure 3.2b plots the coefficients from equation **(3.11)** divided by the share of cases in each range to show a percent difference by race. There is a noisy decrease from 160-280g, but at several amounts, the coefficient is significantly different from zero or marginally significant. This implies that at those amounts, black and Hispanic offenders are more likely to be bunched at 280g than white offenders. Again, this implies there is a conditional racial disparity in bunching at 280g.

3.5.2 Sentencing Consequences

In order to understand the policy implications of this bunching, I estimate the sentencing consequences of crossing the mandatory minimum threshold. Since mandatory minimum sentencing only gives guidelines about minimum sentencing, it is possible that being above the amount has no effect on actual sentencing.[33] I investigate this by estimating the following:

$$Sentence_i = \alpha + \beta_1 Above280_i + \beta_2 Amount_i + \beta_3 (Above280 \times Amount)_i + \varepsilon_i \tag{3.12}$$

where $Sentence_i$ is the sentence handed down for offender **i**, $Above280_i$ is equal to one if the offender is recorded with 280g or more of crack-cocaine and zero otherwise, and $Amount_i$ is equal to the offender's recorded drug quantity centered at 280g. For the main results, I focus on cases sentenced after 2010. In Table A3.11, I estimate similar regressions using the pre-2010 data. As long as the offenders who are bunched above the threshold are not negatively selected from the population just below the threshold, then

[33]In other words, judges could choose to treat defendants with 270g the same as defendants with 280g and apply the mandatory minimum sentence of 10 years to both.

β_1 will provide a conservative estimate of the sentencing penalty associated with crossing the mandatory minimum threshold after 2010. The bunching above 280g suggests this assumption may be violated. As such, I also estimate (12) for states with low levels of bunching above 280g.

I find that bunching at 280g does have sentencing consequences. Offenders recorded with 270-280g after 2010 have a mean sentence of 9.6 years whereas offenders recorded with 280-290g after 2010 have a mean sentence of 11.2 years. Figure 3.3a plots sentencing outcomes by drug weight from 230-330g and the linear fit on each side of the 280g threshold for cases sentenced after 2010. The discontinuity (β_1) is the sentencing penalty from crossing the mandatory minimum threshold. Figure 3.3b shows that there is no discontinuity in predicted sentence, where sentence is predicted from a model using pre-2010 cases and several offender characteristics. Figure 3.3c plots actual sentence for the subset of cases sentenced in states that have low levels of bunching. Even in states where there is little manipulation around the threshold, there is a sentencing penalty of about 2 years.[34,35] See Figure A3.9 for robustness to bandwidths from 10g to 250g.

This estimate assumes that an offender bunched at 280g would be charged with an amount just below 280g in the absence of the 280g threshold. However, the results in Section 3.5.1 suggest that offenders bunched at 280g come from throughout the distribution below 280g. The average sentence after 2010 for offenders in the 50-280g range is 7.9 years. Using that value as the counterfactual sentence implies a sentencing consequence

[34]This is possible because although offenders are negatively selected (in terms of sentence) on some characteristics, like race, they are positively selected on others, like criminal history score.

[35]These estimates indicate that there is a sentencing penalty for crossing the mandatory minimum threshold (both before and after 2010), not that sentences were longer after 2010 or that the sentencing penalty of triggering the mandatory minimum was higher after 2010.

of 3.3 years.

3.5.3 Potential Mechanisms

The four mechanisms I evaluate are: (1) offender responses to the FSA, (2) a shifting composition of cases between state and federal court, (3) law enforcement discretion, and (4) prosecutorial discretion. For these analyses, I present visual evidence as well as a formal analysis of the microdata showing the main bunching results for each mechanism in Table 3.5 and Tables 3.6a-b. Ultimately, I find bunching at 280g in prosecutor case management files from the EOUSA but not at an earlier stage. This implies that prosecutors are responsible for the excess mass at 280g in final sentences. In Section 3.5.3, I discuss several additional empirical tests that also suggest prosecutors are responsible for the bunching of cases at and above 280g.

Offender Behavior

If black and Hispanic offenders respond differently than white offenders to the Fair Sentencing Act, a racial disparity in bunching at 280g may reflect prosecutors' reactions to those different responses rather than racial discrimination. In Table 3.6a, I show that black and Hispanic offenders are not arrested with more drugs following the Fair Sentencing Act, but instead, are holding slightly smaller amounts when arrested after 2010 (after controlling for state fixed effects, sex, and age).[36] In Table 3.6b, I show that black and Hispanic respondents to the NSDUH are not more likely to report having ever used crack, selling drugs in the past 12 months, or having used crack and selling drugs after

[36]Likewise, I find no evidence of a response in the DEA STRIDE data on drug amounts or drug prices (see Figure A3.10).

2010. This implies that the racial disparity in bunching cannot be attributed to differential responses in drug-carrying by race.[37]

Shifting of Cases Between State and Federal Courts

Drug Convictions in Florida Courts

The USSC data covers the universe of federal drug cases, but it is possible that the type of cases prosecuted in federal court versus state court changes after 2010. Cases can be prosecuted federally for many reasons (see Appendix D for a discussion of the reasons a case can enter federal court). State and local authorities could send more of their high weight, 280g cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010.

To test this possibility, I use state-level data on cocaine offense convictions from Florida.[38] Florida classifies drug offenses using broad ranges: 0-28g, 28-200g, 200-400g, and 400+g. The USSC data show a sharp 3.6 percentage point increase in cases with 200-400g convicted in a Florida federal district after 2010 (see Table 3.5, column 7 and Figure A3.11a). If the bunching in federal cases is due to state and local authorities sending more 280g cases to federal prosecutors, then there should be a mirrored decrease in the fraction of state-level cases in Florida with 200-400g. Even more, the decrease should be especially pronounced for black and Hispanic offenders.

I do not find a decrease in state convictions for 200-400g in general or by race. Figure 3.4a plots the share of all cocaine cases in Florida that are for offenses with 200-

[37]Other papers also find that offenders do not respond or respond only modestly to a change in punishments/sanctions. For example, Lee and McCrary (2017) finds that offenders do not discontinuously decrease offending at age 18, despite a discontinuous increase in the probability of a harsh sentence at that age.

[38]In Appendix A, I show similar results for North Carolina. I do not include NC in the main analysis because many of its drug convictions do not include any information about drug type involved.

400g of cocaine by race. Columns 1 and 2 of Table 3.5 confirm this. The probability a state-level drug conviction is in the 200-400g range in Florida does not meaningfully change after 2010. This implies that shifting from state and local courts to federal courts cannot explain the sharp rise of cases at 280g in federal sentencing.[39] In Table A3.12 and Figure A3.11b-c, I show these results are robust to alternative sample restrictions and to using similar data from North Carolina.

Bunching by Law Enforcement Agency Sending Case to EOUSA

The EOUSA prosecutor case management files (which I analyze in more detail below) include a field that indicates the law enforcement agency that sends the case to the EOUSA. If the bunching at 280g is caused by a shift from state courts to federal courts, then bunching should only be present in cases with state law enforcement involved. In Figure A3.11d, I plot the fraction of cases with 280-290g over time by the type of agency involved. I find that bunching at 280g is present in cases with state law enforcement involvement and in cases that are sent from Federal agencies (see Table A3.13 for a formal test). This is further evidence that the bunching at 280g after 2010 is not the result of state to federal case shifting.

[39] Since there are many more cases convicted at the state-level versus federal-level, it is possible that a minor, undetectable shift in Florida would be detectable at the Federal-level. This is not the case for the 200-400g range. First, the state-federal disparity in number of cases is due to states prosecuting more minor possession cases than the federal courts. There are 150 crack or powder cocaine cases in the 200-400g range convicted in federal court districts located in Florida after 2010. There are only 200 cases in this range convicted in Florida state courts after 2010. Re-coding 150 of the 200 Florida cases as if they were not in the 200-400g range does yield a detectable effect. Similarly, re-coding 150 cases not in the 200-400g range as if they were in the 200-400g range also yields a detectable effect. This simple simulation implies that a shift of cases from Florida to the federal system would be detectable in the state data. A related concern is that the large number of cases in urban counties may mask shifting in rural counties. I split the analysis by counties with greater than 5000 cocaine convictions from 2000-2015 and counties with less than 5000 cocaine convictions from 2000-2015. I do not find substantial shifting for either group. For small counties (those with less than 5000 cocaine convictions), I find a decrease in cases with 200-400g of about 0.1 percentage points. For large counties (those with more than 5000 cocaine convictions), I find no change in cases with 200-400g (less than 0.02 percentage points).

Law Enforcement Discretion

NIBRS, Local Law Enforcement Drug Seizures

Using a balanced panel of agencies in the NIBRS data on drug crime, I examine the distribution of drug seizure quantities. If local law enforcement is the source of bunching, I should observe an increase in bunching at 280-290g after 2010. Figure 3.4b plots the fraction of drug seizures with 280-290g over time and does not show an increase in drug seizures with 280-290g after 2010, in general or by race. These results are also shown in Columns 3 and 4 of Table 3.5.[40] In addition, only 5 incidents total are reported with 280-290g in the NIBRS after 2010. This suggests that discretion in local law enforcement and drug sting tactics cannot explain the bunching in drug amounts after 2010.

DEA STRIDE, Federal Law Enforcement Drug Seizures

I also test for bunching in drug quantities from the DEA's STRIDE database.[41] This data includes exhibits sent to DEA laboratories from both federal and local law enforcement agencies. Figure 3.4c plots the share of cocaine exhibits with weights from 280-290g from 2000-2015. There is no increase in exhibits with 280-290g after 2010. Again, Table 3.5 also shows this result. In fact, there are less than 20 total cocaine exhibits in the DEA data with 280-290g after 2010. This further suggests that local and federal law enforcement are not responsible for the observed bunching at 280g after 2010.

Prosecutorial Discretion

Bunching in Prosecutor Case Management Files

[40] This result is robust to using only states that have full coverage by 2012 (i.e. states in which all agencies are participating in NIBRS) and 90-100% coverage from at least 2008-2015 (DOJ 2012). See Table A3.14 and Figure A3.12.

[41] The analysis in this section uses unvalidated DEA data, and I claim authorship and responsibility for all inferences and conclusions that I draw from this information.

The EOUSA provides case-level data extracted from their internal case management system. Using this data, I test for bunching in the quantity of drugs recorded in the case management system. Figure 3.4d shows that there is a sharp increase in the fraction of cases recorded with 280-290g after 2010.[42] Since I find no evidence of bunching in data from earlier stages, this suggests that the bunching occurs once the case is in the hands of the prosecutor.

Table 3.5 indicates that the fraction of cases in 280-290g increases by 7.7 percentage points after 2010. This is twice the increase I find in the final sentencing data. This difference is likely driven by missing values in the EOUSA files. Re-coding each missing value as though it were not in the 280-290g range (i.e. equal to zero) yields an increase of about 3.5 percentage points after 2010, which is consistent with estimates from the sentencing data. The main results below are robust to missing value re-coding.[43]

I also examine bunching at 280g for cases received by the EOUSA before the Fair Sentencing Act is signed into law. These cases are less likely to be influenced by offender or police responses to the FSA. For cases that are received by the EOUSA 60 days before the FSA but sentenced after the FSA, 2.7% are bunched at 280-290g. For cases that are received by the EOUSA 60 days before the FSA and sentenced before the FSA, 0.4% are bunched at 280-290g. The timing of bunching in these cases further suggests the increase in bunching at 280g is due to prosecutor decisions.[44]

The EOUSA data do not contain a field for race of the defendant. I can impute race for cases from the EOUSA data that contain a sentence month and year (not all cases

[42] See Figure A3.13a for a plot of bunching at 280g by the month the case is received.

[43] See Table A3.15 and Figure A3.13b-c for the main bunching results after re-coding the 280-290g dummy variable as equal to zero when the drug weight is missing.

[44] Figure A3.13d plots bunching by year sentenced for cases received before the FSA.

received are sentenced) by using the racial composition of sentencing in each year-month from the USSC sentencing data. As before, I find an increase in 280-290g cases after 2010 and a particularly large increase in months with more black and Hispanic offenders sentenced (see Table A3.15). In Table A3.15, I also show that the disproportionate bunching for black and Hispanic offenders (using imputed race) is robust to including prosecutor fixed effects.

Prosecutor-level Bunching Estimates

To further explore bunching by prosecutors, I use the ID of the lead attorney on each case and test for heterogeneity in bunching by attorney. Since each attorney only has a small number of cases and since I do not know the specific circumstances of each case, I cannot pinpoint "bad behavior" from any individual attorney. However, by estimating bunching separately for each attorney, I can calculate the fraction of prosecutors responsible for the observed bunching. Also, I can compare the distribution of cases for bunching and non-bunching attorneys to further understand where the excess mass at 280-290g is coming from (i.e. where there is relatively less mass in the bunching attorney distribution compared to the non-bunching attorney distribution).

Prior to 2010, approximately 0.4% of all cases with a drug quantity less than 1000g were recorded as having 280-290g. I use this statistic as a benchmark to detect attorneys who bunch after 2010.[45] For each attorney, I calculate the percentage of their cases with 280-290g of drugs after 2010. I classify an attorney as a "bunching" attorney if their bunching is greater than or equal to 0.4%. For this analysis, I limit the sample to attorneys

[45] I can also use the district-level pre-2010 average to account for district fixed effects in cases at 280-290g. Even more, I can use each attorney's pre-2010 behavior as their own benchmark to detect bunching post-2010. Both approaches yield similar results.

with 10 or more cases after 2010. Results are similar when using lead attorneys with 5 or more drug cases after 2010 or with 15 or more drug cases after 2010.

The majority of these attorneys exhibit no bunching.[46] In other words, their fraction of cases with 280-290g post-2010 is at or below the pre-2010 average. Approximately 30.4% of prosecutors, however, do have a higher than normal percentage of cases with 280-290g after 2010. Drawing 50 samples (stratified on lead attorney ID and with replacement) from the data and re-calculating the fraction of bunching attorneys in each sample yields a standard error of 0.024. This implies a 90% confidence interval on the estimate of about 26.4-34.3%. Over 50% of these attorneys have two or more cases at 280-290g and over 25% have three or more cases at 280-290g.[47] The fraction of bunching attorneys is also significantly different at the one percent level from the fraction calculated by randomly re-assigning cases to prosecutors (see Figure A3.15).

In Figure A3.16, I map the number of bunching attorneys in each state (using attorneys with 5 or more drug cases post-2010 to increase the set of states that have eligible attorneys).[48] The attorney-level bunching cannot be accounted for by district fixed effects. The within-district standard deviation in the 280-290g bunching metric is 0.13, the between-district standard deviation is similar at 0.12, and district fixed effects only explain about 6% of the variance in the attorney-level bunching metric.

[46] Figure A3.14 plots a histogram of the resulting measure for the 128 attorneys who served as lead attorney on at least 10 drug cases after 2010.

[47] While this statistic is only calculated for the 128 attorneys with 10 or more drug cases post-2010, this ratio of non-bunching to bunching attorneys holds for the entire data. In fact, those bunching attorneys with 10+ cases post-2010 do not even account for half of the total observed bunching. Removing the bunching attorneys with 10+ cases post-2010 decreases the bunching estimate from 0.078 to 0.054. In other words, the majority of bunching at 280g is accounted for by prosecutors with fewer than 10 cases post-2010.

[48] The number of bunching attorneys in a state is positively correlated with racial animus in that state (see Table A3.16).

Further Evidence on Source of Excess Mass at 280g

In Table 3.7, I estimate the likelihood a case is charged below 280g, with 280-290g, or above 290g for the bunching versus the non-bunching attorneys. This echoes the approach that Goncalves and Mello (2018) use to formally estimate bunching in speeding tickets in Florida.[49] For this analysis, I use two definitions of a bunching attorney: (1) attorneys who have an above-average share of cases with 280-290g post-2010 and (2) attorneys who have an above-average share of cases with 50-60g pre-2010. Definition (2) provides a classification of bunching attorneys that is not mechanically related to the fraction of cases in the 280-290g range.[50]

The key idea is that the non-bunching attorneys provide a counterfactual density since they are not responding to the mandatory minimum thresholds in the same way as the bunching attorneys. Comparing these two groups, I see that non-bunching attorneys (in both definitions) have more cases below 280g post-2010 than bunching attorneys and a similar number of cases above 290g post-2010. This provides further evidence, from different data and a different source of variation, that those attorneys who bunch at mandatory minimum thresholds are shading up the reported quantity of crack-cocaine.

Additional Evidence on Prosecutor-level Bunching

Next, I identify attorneys who switch from one federal district to another federal district, and, using the two definitions above, I test whether bunching is persistent across districts. Definition (2) is important for this analysis because there are few attorneys

[49] They compare lenient police officers to non-lenient police officers.

[50] In Appendix A, I show that the results using definition (1) are robust to categorizing the prosecutor for defendant **i** as a bunching or non-bunching attorney leaving out defendant **i** from the determination, and that all results are robust to bootstrapping the standard errors to adjust for error in the bunching classification. I also show that these results are robust to using attorneys with 15+ cases or 5+ cases. See Tables A3.17-A3.19.

who switch districts and have a sufficient number of cases post-2010 in both districts. Table A3.20 shows these results. I find that an attorney who bunches at the 10-year mandatory minimum threshold in their first district is more likely to bunch at the 10-year threshold in their second district than an attorney who does not bunch at the 10-year threshold in their first district. In other words, bunching at the 10-year mandatory minimum threshold is a behavior that persists across districts, suggesting that bunching is related to a characteristic of the prosecutor and not another actor in the district (e.g. police, judge, or defense attorney).[51]

In Figure A3.17, I examine how other prosecutors in a district change their bunching behavior when a bunching prosecutor enters. I find that that when a bunching attorney switches into a new district, all other attorneys in that district begin bunching more. To conduct this test, I classify bunching attorneys using data from 1994-1999 and definition (2). I then identify the districts that those attorneys move into, and I study the attorneys in that district after the first bunching attorney moves in post-1999. This means earlier years are over-represented. I show that bunching increases in a district once a bunching attorney enters, but that it does not decrease once the bunching attorney leaves. This is suggestive evidence that the increase in bunching is not related to a temporary shift, such as competition among attorneys, but that it may be related to something more permanent, such as learning about techniques or developing new beliefs/norms. Figure A3.18 shows that bunching at the 10-year mandatory minimum increased by 60% from 1988-90 to 2010, which is consistent with the practice of bunching being learned over time. The

[51] Recall, Table A3.13 shows that the increase in bunching at 280-290g is similar for most police agencies sending cases. This also suggests that the variation in bunching at the prosecutor level is due to prosecutor choices and not choices made by investigators.

figure notes in Appendix A contain a more detailed discussion of these results.

Finally, in Table A3.21, I show that attorneys who bunch at 280-290g post-2010 also have more cases bunched at 28-29g (the five-year mandatory minimum) post-2010 and more cases bunched at 50-60g pre-2010 (the pre-2010 ten-year mandatory minimum). Likewise, attorneys who bunch at 50-60g pre-2010 also have more cases bunched at 28-29g post-2010 and 280-290g post-2010. One concern about the estimation of prosecutor-level bunching is that the variation across prosecutors could be due to noise alone, especially since I only require prosecutors to have 10 or more cases after 2010. These results that show prosecutor-level bunching is persistent across time, across districts, and across mandatory minimum thresholds provide strong evidence that the prosecutor-level bunching metric does contain a signal of prosecutor type.

While it may be surprising that prosecutors could induce this bunching, recall that this ability is explicitly written into federal sentencing guidelines. One tool prosecutors can use to increase the weight used at sentencing is tying the defendant to a larger drug conspiracy. Cases with 280-290g after 2010 are more likely to have a lead charge of "drug conspiracy" than cases with 290g-1000g (see Table A3.22). Prior to June 2013, the evidence about relevant quantities did not need to satisfy the "beyond a reasonable doubt" evidentiary standard, because the "principles and limits of sentencing accountability under this guideline are not always the same as the principles and limits of criminal liability" (USSC, 2015). A Supreme Court decision in June 2013 changed the evidentiary standard, and I evaluate that change below.

3.5.4 The Impact of *Alleyne v. United States*

On January 14, 2013, the Supreme Court began hearing arguments in the case *Alleyne v. United States*. The petitioner, Allen Alleyne, argued that facts that increase the mandatory minimum sentence for a defendant are "elements" of the alleged crime and should be evaluated by a jury. In a 5-4 decision on June 17, 2013, the Court ruled in favor of Alleyne and issued a decision that changed the evidentiary standard for evidence related to mandatory minimum sentencing enhancements (Bala 2015).

Prior to this decision, evidence on drug quantities was presented to the judge during the "sentencing phase" of a trial. The presiding judge would then decide, based on the legal standard of "a preponderance of evidence," whether the mandatory minimum sentence applied. The Supreme Court decision required that evidence that would raise the minimum sentence for a defendant be presented to the jury and evaluated based on the stricter legal standard of "beyond a reasonable doubt." I estimate how prosecutors reacted to this decision by comparing the change in bunching around June 17, 2013 to the change around June 17th in other years after 2010. If prosecutors are inflating drug amounts to levels that could not be supported at trial, then there will be a decrease in bunching for cases received after the Supreme Court decision.

Using the EOUSA case management data, I estimate the discontinuity in the prevalence of bunching for cases received around June 17, 2013 relative to the discontinuity for cases received around June 17 in all years after 2010 excluding 2013:

$$
\begin{aligned}
(\text{Recorded}\,280-290g)_{it} = \alpha_0 &+ \beta_1 \text{AfterJune17}_{it} + \beta_2 \text{DaysFrom}_{it} + \beta_3 (\text{After}\times\text{DaysFrom})_{it} \\
&+ \delta_1 (\text{AfterJune17} \times \text{Year2013})_{it} + \delta_2 (\text{DaysFrom}\times\text{Year2013})_{it} \\
&+ \delta_3 (\text{After}\times\text{DaysFrom}\times\text{Year2013})_{it} + D_{it} + \varepsilon_{it} \qquad (3.13)
\end{aligned}
$$

where After_{it} is equal to one if case **i** is received after June 17th of year **t** but before January 1st of year **t+1** and is equal to zero if case **i** is received before June 17th of year **t** but after January 1st of year **t**. DaysFrom_{it} is the number of days from June 17th that case **i** is received, and Year2013_{it} is equal to one if case **i** is received in 2013 and is equal to zero if it is received in 2011-2012 or 2014-2016.[52] D_{it} represents day-of-week fixed effects. The coefficient β_1 is the average discontinuity in the fraction of cases with 280-290g after June 17 from 2011-2016. The coefficient δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.[53,54]

Column 2 of Table 3.8 shows this result using a bandwidth of 130 days (the Imbens-Kalyanaraman optimal bandwidth for 2013) before and after June 17th in each year. The coefficient in the first row indicates that, on average, there is approximately no change in bunching after each June 17th from 2011-2016.[55] The next coefficient, labeled "After June 17, 2013", shows the change in bunching that is specific to June 17, 2013. I find

[52] I do not include 2017 in these analyses since the data do not include the full year.

[53] In response to *Alleyne*, Attorney General Eric Holder released a memo in August 2013 instructing US attorneys to decline to charge quantities necessary to trigger the mandatory minimum in cases with low-level and non-violent offenders who have little criminal history. The decrease in bunching could be a result of this memo and not the Supreme Court decision. To address that concern, I narrow the bandwidth of the RD design to 60 days before/after June 17th. Even then, I find a discontinuous decrease in bunching (although the standard errors are much larger). Also, using updated EOUSA data, I find that there is no change in bunching after May 12, 2017, the day Attorney General Jeff Sessions rescinded the August 2013 Holder memo.

[54] I do not conduct the traditional RD identifying assumption tests in this section. For one, the EOUSA data contain very few case-level covariates. Even more, the resulting discontinuity, whether it arises from prosecutors rushing to try cases before the Supreme Court decision or solely from prosecutors changing their behavior immediately after the decision, reveals that prosecutors were submitting evidence to judges that they believed would not hold up if submitted to a jury. That said, the density of cases is displayed in Figure A3.19a.

[55] The coefficient on AfterJune17 for 2013 is at least twice as large as the next largest all other years from 1999-2016 (when estimating the non-2013 years separately instead of pooling). See Figure A3.19b.

that bunching changes discontinuously only after June 17, 2013. In fact, the fraction of cases recorded with 280-290g drops by about 15 percentage points after the ruling in Alleyne. This is also the case for the 120-day and 60-day bandwidth, although as I narrow the bandwidth, I lose precision.[56] Table A3.23 shows that the decrease in bunching after Alleyne is robust to imputing missing values as zero. Figure A3.20 shows robustness to additional bandwidth choices and choice of polynomial.

Figure 3.5 illustrates the large discontinuity in the fraction of cases with 280-290g around June 17, 2013. Although it does not eliminate it entirely, it is clear that Alleyne at least somewhat reined in the practice of bunching. This suggests that prosecutors were using discretion to build cases on evidence that was unlikely to pass "beyond a reasonable doubt" scrutiny from juries.

3.5.5 Discrimination and Alternative Explanations

Now, I introduce a simple model of prosecutor objectives to discuss potential explanations for the racial disparity in bunching at 280g and to motivate empirical tests of those explanations.

Model of Prosecutor Objectives

First, I detail the prosecutor's decision problem, which determines the probability $Pr(a = mm|s,r)^t$ that a case with a given amount seized s and defendant race r is charged with an amount a that is equal to the mandatory minimum threshold $mm =$

[56] I do not find a decrease in the fraction of cases recorded with 280-290g after the announcement that the Supreme Court would hear the case (in October 2012) or after the oral arguments (in January 2013). Unlike some Supreme Court cases, the ultimate ruling in June 2013 was not clear from the outset. At the time, the New York Times referred to the case as a "murky area of sentencing law" on which the Supreme Court had issued "contradictory rulings." For this reason, the announcement and the arguments alone would not provide sufficient evidence of whether the law would ultimately change.

$\{5, 28, 50, 280\}$. Although I do not estimate any of the parameters in the following model directly, I use it to illustrate channels through which $Pr(a = mm|s, r)^t$ may differ by race and to discuss suggestive empirical tests of those various channels.[57]

The prosecutor for the case chooses the **amount (in grams) of drugs charged** a, and can charge amounts higher than **seized evidence** s by collecting additional evidence $a - s$. Seized evidence s is a noisy measure of **true drug trafficking** d, which is unobservable to the prosecutor. For a given case, prosecutor i chooses the amount of drugs charged a to solve the following problem:

$$\max_a \pi(l(a)^t) - \gamma(r,x) \times c_g(a-s) - c_d(|l(a)^t - (l^*(s,r,x) + \phi_i(r,x))|) \tag{3.14}$$

The function $\pi(.)$ represents the **career benefits** a prosecutor gets from securing a longer sentence. There are also costs to the prosecutor associated with increasing a, such as the **cost of gathering the additional evidence** $c_g(a-s)$ to build the case. This cost $c_g(a-s)$ is increasing in $a-s$.[58] This cost is determined by other actors the prosecutor must face in the process of working a case. Judges, defense attorneys, juries, witnesses, or other actors in the criminal justice system who are racially biased may present fewer obstacles to entering the additional evidence $a-s$ for cases involving black and Hispanic defendants. Also, if defendants of one race procure better defense counsel, that counsel may make it more difficult for the prosecutor to use additional evidence $a-s$. These **cost**

[57] Note, I write down a static model below, but it can incorporate reputational benefits or reputational costs associated with bunching. The data are not amenable to testing dynamics at the prosecutor-level. I focus on the static problem because it has clear connections to empirical tests I can conduct.

[58] Again, I assume that prosecutors don't suppress evidence and thus, $a \geq s$.

differences by race (and other defendant characteristics) are captured in $\gamma(.)$.

The prosecutor also faces a **psychic cost of deviating** $c_d(.)$ from **the sentence that would be justified by law if true drug trafficking were observed** $l^*(d)$. Since true drug trafficking d is unobservable, prosecutors form an expectation of d by solving a signal extraction problem given the seized evidence s, defendant race r, and other characteristics x. This yields $l^*(s,r,x)$.[59]

Finally, a **prosecutor specific taste parameter** $\phi_i(r,x)$ is added to the sentence $l^*(s,r,x)$, reflecting the prosecutor's animus for defendants based on race r or other characteristics x. Assume that only ϕ_i varies at the prosecutor level.

Writing down the prosecutor's objective function makes explicit the various channels that could cause a conditional racial disparity in the probability a defendant is bunched at 280g. First, the disparity could be due to taste-based racial discrimination: $\phi_i(bh,x) > \phi_i(w,x)$. Second, it could be due to statistical discrimination: $l^*(s,bh,x) > l^*(s,w,x)$. Third, it could be due to racial differences in the cost (to the prosecutor) of building a case: $\gamma(bh,x) < \gamma(w,x)$. All three of the channels could also be related to other characteristics x that are correlated with race r rather than race itself.

Empirical Tests of Discrimination and Other Explanation

Other Offender Characteristics

First, I test the explanation that the racial disparity in bunching at 280g is driven by a characteristic correlated with race. To do this, I estimate how bunching differs by various observable offender characteristics. Specifically, I estimate equation (3.2) fully interacted with binary variables for the following offender characteristics: college edu-

[59] I model the signal extraction problem in Appendix C.

cation or more, male, above the median age for offenders, offense involves a weapon, above the median criminal history score, above the median number of other current offenses, and convicted in a state with an above median fraction of black or Hispanic cases pre-2010.

This partially addresses concerns that white and black and Hispanic offender's are different on a wide range of other characteristics and that race may be a proxy for those characteristics. By estimating bunching by race and education, for example, I can compare black offenders with a college education to white offenders with a college education. If the racial disparity still exists within education categories, then this suggests that the racial disparity is driven by attitudes about race. In Table 3.9, I show that the racial disparity in bunching exists even within all of these observably similar groups.

The observable characteristics from the USSC data are only a subset of what the prosecutor observes about a defendant. One concern is that black and Hispanic drug offenders may be more likely to operate in drug organizations or gangs, and that prosecutors may charge offenders from gangs with higher amounts for various reasons. The 2004 Survey of Inmates in Federal Correctional Facilities (SIFCF) indicates that black and Hispanic federal drug offenders are **less** likely to be a member of a drug organization than white federal drug offenders. Also, they are less likely to report income from illegal activities prior to arrest.[60] Also, although the amount charged is endogenous to the presence

[60]The SIFCF is a nationally representative survey of inmates in federal prisons. Over 3,000 inmates from 39 federal prisons were interviewed for the 2004 survey. The interviews were conducted by the US Census Bureau on behalf of the Bureau of Justice Statistics. At the beginning of the interview, inmates are told their answers are confidential and that their responses cannot be released to the prison or to anyone else in a way that would identify them. These data contain information on whether the offender was involved in a drug organization/gang. Although the statistics are based on self-reports, it does not appear black and Hispanic offenders report differently than white offenders on other sensitive questions, such as whether police used force during their arrest or whether they have had thoughts of revenge.

of a conspiracy charge, there is a racial disparity in bunching for offenders charged with conspiracy and for offenders not charged with conspiracy (see Table 3.9, column 8). As in the SIFCF data, white offenders are also more likely to face a conspiracy charge. This further suggests that differences in gang participation by race do not explain the racial disparity in bunching at 280g.

Costs to the Prosecutor of Bunching at 280g

In this section, I test the explanation that the racial disparity is due to racial differences in the costs to the prosecutor of bunching a case at 280g.

First, I test whether racial difference in defense counsel could explain the racial disparity in bunching. The data do not include the offender's type of defense counsel in all years. This information is available for 1999-2002, but in those years, black, Hispanic, and white crack-cocaine offenders are equally likely to be represented by private counsel.[61] The 2004 Survey of Inmates in Federal Correctional Facilities also indicates that private counsel retention is the same by race. Using data from the 1999-2002 USSC files, I construct each district's private counsel retention rate and tag districts as below or above median private counsel retention. I find that bunching and the racial disparity in bunching is similar in places with low and high private counsel retention (see Table A3.24).

Next, I consider whether the racial disparity in bunching can be attributed to judge bias. I am able to match approximately half of the cases in the EOUSA files to a judge race and political party. For these cases, I do not find any evidence that judge race or political party influences the probability a case is bunched at 280g (see Table A3.25).[62]

[61] 21.0% of white offenders, 22.7% of black offenders, and 21.7% of Hispanic offenders retain private counsel from 1999-2002.

[62] I have also examined heterogeneity in bunching by race of the head US attorney in the district and the racial composition of prosecutors, judges, defenders, and probation officers in the district. I do not find

Also, unlike prosecutors, judges with a high share of cases at 280g post-2010 are not any more likely to have cases at 28g post-2010 or at 50g pre-2010 (see Table A3.26).

I also test whether district-level differences in costs of gathering evidence are related to bunching at 280g. I find that the increase in bunching at 280g is similar in districts with a low and high fractions of cases declined due to "weak evidence" or "lack of resources" (see Table A3.23).[63] This suggests that costs of developing evidence are not related to the rise in bunching at 280g.

Taste-based vs. Statistical Discrimination

Lastly, I consider taste-based vs. statistical discrimination. These two explanations are difficult to disentangle. A simple model of statistical discrimination would imply that prosecutors within the same district should be equally likely to bunch cases at 280g and that, after accounting for other offender characteristics, the racial disparity in bunching should decrease. I find that there is variation in the level of bunching across prosecutors within districts, and that the racial disparity exists within observably similar defendant groups. While these results could be reconciled by a more detailed model of statistical discrimination, they suggest that the simple model outlined above does not explain the racial disparity.

One potential explanation of these results is that some prosecutors have biased tastes against black and Hispanic drug offenders and believe they should be punished more harshly than white drug offenders. To explore the taste-based discrimination mechanism, I use a state-level measure of racial animus constructed by Stephens-Davidowitz (2014)

robust results on these margins.

[63]The EOUSA files contain information about why a case is declined for about 60% of its cases.

based on intensity of Google searches including racial slurs in each state. I match this measure to the USSC Sentencing data using the state of the federal district in which the offender is convicted. I take this measure of racial animus as a potentially valid measure of prosecutor tastes for several reasons: about half of government lawyers work in the same state they were born in (author's calculation from 2000 and 2010 publicly available Census samples), assistant US attorneys must reside in the district they serve in, and assistant US attorneys have a choice over where to apply.[64]

Again, I estimate equation (3.2) fully interacted with a dummy variable for high racial animus states that is equal to one if the state where the offender is convicted is above the median on a measure of racial animus from Stephens-Davidowitz (2014) and equal to zero if it is below the median. If racial animus is correlated with some state-level preference for harsh sentencing, then I should find an effect for both white and black and Hispanic offenders. However, if the effect is driven by racist beliefs about black and Hispanic offenders, then it should only be present for those groups.

I find that in states with a higher level of racial animus, bunching at 280-290g is more prevalent specifically for black and Hispanic offenders.[65,66] These results are in

[64]Recall that Alleyne v. US made the jury more important in mandatory minimum cases after 2013. This change led to stricter evidentiary standards for mandatory minimum cases (beyond a reasonable doubt versus preponderance of evidence). However, if juries are, on average, more racially biased than judges, then the effect of Alleyne v. US may be buffered by the increased racial bias of juries. I find that the fraction of cases at 280-290g in low racial animus states (below median) fell by 40% from 2011-2012 to 2014-2017. In high racial animus states (above median), the fraction of cases at 280-290g fell by 20%. This is suggestive evidence that Alleyne was, in fact, less effective in states with high racial animus. However, in all states, the increase in evidentiary standards led to a net decrease in cases at 280-290g.

[65]The racial animus measure was developed to measure animus against black people. I assume that this is correlated with animus for Hispanic people, so I focus on the pooled results. However, the estimates are similar if I exclude black offenders or Hispanic offenders.

[66]Specifically, I split states by above/below the median racial animus. States above the median racial animus measure are: AL, AR, CT, DE, FL, GA, IL, IN, KY, LA, MD, MI, MO, MS, NC, NJ, NV, NY, OH, OK, PA, RI, SC, TN, and WV. States below the median racial animus measure are: AK, AZ, CA, CO, HI, IA, ID, KS, MA, ME, MN, MT, ND, NE, NH, NM, OR, SD, TX, UT, VA, VT, WA, WI, and WY.

Tables 3.9-3.10. Column 8 of Table 3.9 shows that in states with high levels of racial animus, black and Hispanic offenders are substantially more likely to be charged with an amount at or slightly above the mandatory minimum threshold.

Table 3.10 explores the robustness of this result. Columns 1-4 introduce individual and district-level controls interacted with the after 2010 by race dummy variables, and the relationship between animus and bunching is unchanged. Columns 5 and 6 estimate the relationship between bunching and the continuous measure of state-level animus from Google Trends. The coefficient in column 5 is not statistically significant (p-value = 0.2), but the magnitude is much larger than the coefficient for white offenders. Also, based on that coefficient, white and black and Hispanic offenders at low-levels of animus are not statistically different from each other, but they are statistically different at higher levels of animus. Column 6 re-estimates column 5 after eliminating outliers in the animus measure (states with animus below the 1st percentile or above the 99th percentile).

In column 7 of Table 3.10, I introduce a district-level of racial animus by aggregating implicit association test scores for people reporting an occupation of "lawyers, judges, and related workers." Since many states contain multiple federal districts, I include state fixed effects interacted with after 2010 by race dummy variables. The estimate, then, is identified from within state variation in the IAT animus measure. I find the average IAT score of lawyers in a federal district is correlated with higher bunching for black and Hispanic offenders (p-value = 0.14).

3.6 Conclusion

For federal drug crimes, a sharp increase in sentencing is triggered when the offense involves at or above a certain amount of drugs. In this paper, I show that there is substantial bunching at and above that point where the mandatory minimum sentence increases, and that bunching is disproportionately larger for black and Hispanic offenders. I use the pre-2010 distribution of drug weights, when the threshold is at 50g instead of 280g, to show that the racial disparity in bunching at 280g post-2010 is conditional on observed drug amounts.

Since the bunching only appears in prosecutor case management data and the final sentencing data but not in data on state-level convictions or drug seizures, it is likely a result of prosecutorial discretion. Several additional tests confirm this. In fact, just 20-30% of attorneys account for 100% of the bunching observed in the case management data. In addition, bunching becomes less prevalent among prosecutors following a Supreme Court decision that requires stricter evidentiary standards for drug quantity evidence. This, in addition to numerous other tests discussed above, suggests that prosecutors are shading drug amounts upward to induce longer sentences.

Why do some prosecutors bunch black and Hispanic defendants at 280g more often than white defendants? The racial disparity cannot be explained by observable individual characteristics or district characteristics. Black and Hispanic crack-cocaine defendants are just as likely to retain private counsel as white defendants. Also, bunching at 280g is unrelated to judge race, political party, and the judge's share of cases at other mandatory minimum thresholds. Since only a subset of prosecutors practice bunching and there is variation across prosecutors within federal districts, a simple model of statistical discrimination does not apply either. This suggests the disparity may be the result of taste-based discrimination. In fact, I find the racial disparity in bunching at 280g is largest in federal districts in states with higher levels of racial animus.

Finally, the bunching in drug weights and the racial disparity in bunching has meaningful implications for the racial sentencing gap. Depending on the counterfactual sentence imputed for the affected offenders, bunching at 280g can account for 2-7 percent of the racial disparity in crack-cocaine sentences. A highly conservative estimate suggests that being bunched at 280g adds 1-2 years to an offender's sentence. Multiple estimates suggest the cost of incarceration (combining direct care costs and the cost of lost current and future wages for the offender) is approximately $60,000 per person per year (Donohue 2009; Mueller-Smith 2015). I find 3.6% of black and Hispanic crack-cocaine offenders are bunched at 280g after 2010 versus 1.2% of white crack-cocaine offenders. Assuming 3.6% and 1.2% of all drug cases from 1999-2015 were subject to similar discretion by race implies total costs of 1.3 billion dollars for black and Hispanic offenders versus 148 million dollars for white offenders. In terms of incarceration, the disparity implies 21,000 years sentenced due to this discretion for black and Hispanic offenders versus 2,500 years sentenced for white offenders.

3.7 Tables and Figures

Table 3.1: Summary Statistics for USSC Sentencing Data

	1999-2010	2011-2015
Black or Hispanic	0.921	0.939
	(0.270)	(0.239)
Age (in years)	31.187	34.166
	(8.517)	(8.748)
Male	0.915	0.916
	(0.279)	(0.277)
College or more	0.126	0.148
	(0.332)	(0.355)
High school or more	0.509	0.598
	(0.500)	(0.490)
Not US citizen	0.046	0.033
	(0.209)	(0.178)
Weapon involved	0.262	0.296
	(0.440)	(0.456)
Number of other current offenses	1.606	1.720
	(1.427)	(1.735)
Criminal history points	5.713	6.512
	(5.474)	(5.586)
Drug weight (in grams)	102.530	116.968
	(156.957)	(169.892)
Sentence (in years)	9.294	7.807
	(7.057)	(5.833)
Observations	47,439	9,445

Notes: The table above describes defendants found in the USSC sentencing data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. The statistics above are derived from the cleaned USSC data in which the following cases are removed: cases with missing drug weight values (including those cases with weights coded as a range), cases with reported problems in the drug weight variables, cases where judges change or do not accept the findings of fact for drug weights, and cases at and above 1000g.

Table 3.2: Effect of Changing Mandatory Minimum Threshold on Bunching at 280-290g

	Pr(280-290g Crack-Cocaine Recorded)		
	(1)	(2)	(3)
After 2010	0.0347***		0.0754***
	(0.00204)		(0.0132)
After 2010 x White		0.0125**	
		(0.0053)	
After 2010 x Black or Hispanic		0.0360***	
		(0.0021)	
Constant	0.0051***	0.0032***	0.00333***
	(0.0003)	(0.0010)	(0.00118)
P-value: W (White) = BH (Black or Hispanic)	-	0.0000	-
Trial Cases Only	No	No	Yes
Observations	56,884	52,745	2,823

Notes: Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 3.1 for notes on sample selection. The row "P-value: W (White) = BH (Black or Hispanic)" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In the remaining tables, I abbreviate the label to "P-value: W= BH." Specifications with the race and after 2010 interactions also include a dummy variable equal to one for black and Hispanic offenders and equal to zero for white offenders. Coefficients are estimated from the following regression for Column 1:

$$(1) \quad (\text{Charged}\,280-290g)_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \varepsilon_{it}$$

and the following regression for Column 2:

$$(2) \quad (\text{Charged}\,280-290g)_{it} = \alpha_0 + \beta_1(\text{After2010} \times White)_{it} + \beta_2(\text{After2010} \times \text{BlackOrHispanic})_{it} + \text{BlackOrHispanic}_{it} + \varepsilon_{it}$$

Column 3 re-estimates equation (1) excluding cases that end in a plea deal (i.e. trial cases only). I do not re-estimate equation (2) on the trial-only sample because there are zero white offenders with 280-290g in trial cases after 2010.

*** p<0.01, ** p<0.05, * p<0.1

Table 3.3: "Missing Mass" in the Distribution of Drug Amounts, Comparing Pre- and Post-2010 Distributions

Panel A. Analysis of Changes in the 0-100g Range.					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0172***	-0.0711***	0.0358***	-0.0061**	-0.0089**
	(0.0038)	(0.0048)	(0.0039)	(0.0028)	(0.0036)
Constant	0.1139***	0.2920***	0.1099***	0.0714***	0.1232***
	(0.0015)	(0.0021)	(0.0014)	(0.0012)	(0.0015)
Predicted Change from Conceptual Model	Increase	Decrease	Ambiguous	Decrease	Decrease
Observations	56,884	56,884	56,884	56,884	56,884
Panel B. Analysis of Changes in the 100-1000g Range.					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(6)	(7)	(8)	(9)	(10)
After 2010 (Actual Change)	-0.0152***	0.0347***	0.0055**	0.0019	0.0062***
	(0.0043)	(0.0020)	(0.0024)	(0.0017)	(0.0020)
Constant	0.1929***	0.0051***	0.0439***	0.0214***	0.0263***
	(0.0018)	(0.0003)	(0.0009)	(0.0007)	(0.0007)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	56,884	56,884	56,884	56,884	56,884

Notes: Robust standard errors estimated jointly by seemingly unrelated regression in parentheses. The estimates in this table are based on the USSC data. See Table 3.1 for notes on sample selection. The predicted change from the conceptual model of prosecutor behavior in Section 3.2.2 is displayed in the row labeled "predicted change from conceptual model." Coefficients are estimated from the following regression for each range:

$$(3) \quad (\text{Charged X-Yg})_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \varepsilon_{it}$$

Tables A3.9f-g display versions of this table with race interactions. Tables A3.9a-e display versions of this table with time trend interactions.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Racial Difference in Shifting from 50g Compared to Shifting to 280g

	Pr(50-60g)	Pr(280-290g)
	(1)	(2)
After 2010 x Black or Hispanic	-0.0066**	0.0360***
	(0.0029)	(0.0021)
After 2010 x White	-0.0006	0.0125***
	(0.0111)	(0.0053)
Constant	0.0653***	0.0032*
	(0.0042)	(0.0010)
P-value: BH = W	0.6000	0.0000
Observations	52,745	52,745

Notes: Robust standard errors standard errors estimated jointly by seemingly unrelated regression in parentheses. The estimates in this table are based on the USSC data. See Table 3.1 for notes about sample selection. Coefficients are estimated from the following regression for each range:

$$(4) \quad (\text{Charged X-Yg})_{it} = \alpha_0 + \beta_1 \text{After2010}_{it} + \varepsilon_{it}$$

Adding the coefficient in column (1) for black and Hispanic offenders to the coefficient in column (2) for black and Hispanic offenders yields a new coefficient of 0.0293. This coefficient is still larger than the coefficient in column (2) for white offenders and the two are statistically different at the one percent level (p-value = 0.0084).
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Bunching Analysis for Potential Mechanisms

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures					
	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)	(4)	(5)
After 2010	0.00005		-0.0002***		-0.0006***
	(0.0005)		(.0001)		(0.0002)
After 2010 x White		0.0004		-0.0001	
		(0.0011)		(0.0001)	
After 2010 x Black or Hispanic		0.0002		-0.0003***	
		(0.0005)		(0.0001)	
Constant	0.0051***	0.0085***	0.0004***	0.0002***	0.0010***
	(0.0003)	(0.0005)	(0.00005)	(0.0001)	(0.0001)
Data Analyzed	FL Convictions	FL Convictions	Drug Seizures, NIBRS	Drug Seizures, NIBRS	Drug Evidence, DEA STRIDE
Drugs Included	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine	Cocaine, all types
P-value: W = BH	-	0.8148	-	0.2382	-
Observations	214,573	214,573	203,700	191,774	100,306
Panel B. Analysis of Bunching in Prosecutor Case Files and Final Sentencing					
	Pr(280-290g)	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)
	(6)	(7)	(8)	(9)	(10)
After 2010	0.0783***	0.0408***		0.0347***	
	(0.00561)	(0.0126)		(0.00204)	
After 2010 x White			0.0031		0.0125**
			(0.0292)		(0.0053)
After 2010 x Black or Hispanic			0.0447***		0.0360***
			(0.0130)		(0.0021)
Constant	0.0039***	0.1096***	0.1242***	0.0051***	0.0032***
	(0.0004)	(0.0072)	(0.0156)	(0.0003)	(0.0010)
Data Analyzed	EOUSA Case Management System	USSC Sentencing, FL only	USSC Sentencing, FL only	USSC Sentencing	USSC Sentencing
Drugs Included	Crack-cocaine	Cocaine, all types	Cocaine, all types	Crack-cocaine	Crack-cocaine
P-value: W = BH	-	-	0.1566	-	0.0000
Observations	19,363	6,856	6,856	56,884	52,745

Notes: Robust standard errors in parentheses. When possible, the specifications above use a sample of offenses with drug amounts between 0 grams and 1000 grams. Analyses of state-level drug convictions do not make this restriction since the state reports broad drug weight categories instead of specific amounts. When broad categories (e.g. 200-400g) are analyzed, a linear trend in year is included. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In Panel A: columns 1-2 show an analysis of reported drug amounts for state-level drug convictions in Florida, columns 3-4 show an analysis of weights for seized drugs reported to the FBI through the National Incident Based Reporting System, and column 5 shows an analysis of weights for drugs sent to DEA laboratories. In Panel B: column 6 shows an analysis of weights recorded in case management files from the Executive Office of the US Attorney, columns 7-8 show an analysis of weights from USSC sentencing data for federal convictions in FL using broad drug categories and all types of cocaine, and columns 9-10 show the main bunching results from Table 3.2 for all federal crack-cocaine convictions in the USSC sentencing data. Coefficients in columns 1, 3, 5, 6-7, and 9 are estimated from the regression in equation (1) of Table 3.2, with a linear time trend included for columns 1 and 7 (the broad drug categories). Coefficients in columns 2, 4, 8, and 10 are estimated from the regression in equation (2) of Table 3.2, with a linear time trend included for columns 2 and 8.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6a: Offender Drug-Holding Behavior by Race, After Fair Sentencing Act in 2010

	Weight	Pr(280-290g)	Weight	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-280g)	Pr(270-280g)	Pr(280-290g)	Pr(>290g)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010 x White			0.0768	0.0342***	-0.0298***	0.0000	-0.0058***	-0.0000	-0.0000	0.0015**
			(0.6040)	(0.0041)	(0.0037)	(0.0017)	(0.0012)	(0.0000)	(0.0001)	(0.0007)
After 2010 x Black			-2.9470***	0.0531***	-0.0264***	-0.0077***	-0.0171***	-0.0001***	-0.0002**	-0.0016***
			(0.2774)	(0.0029)	(0.0026)	(0.0011)	(0.0010)	(0.0001)	(0.0001)	(0.0004)
Black	1.716***	0.0001	2.4062***	-0.0951***	0.0707***	0.0101***	0.0131***	0.0001***	0.0001	0.0009**
	(0.265)	(0.0001)	(0.2867)	(0.0026)	(0.0024)	(0.0010)	(0.0009)	(0.0001)	(0.0001)	(0.0004)
Constant	10.266***	0.0003**	9.8706***	0.7280***	0.2031***	0.0345***	0.0303***	0.0001	0.0003**	0.0038***
	(0.436)	(0.0001)	(0.4458)	(0.0041)	(0.0037)	(0.0016)	(0.0015)	(0.0001)	(0.0001)	(0.0006)
Observations	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677	191,677
P-value: W = B	-	-	0.0000	0.0002	0.4433	0.0001	0.0000	0.0282	0.2444	0.0002

Notes: Robust standard errors estimated jointly by seemingly unrelated regression in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-3 show the relationship between race of offender and drug weight seized, in general. Column 4 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 5-11 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W= B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." Coefficients in column 1 are estimated from the following regression:

$$(5) \quad Weight_i = \alpha_0 + \beta_1 Black_i + X_i + Z_s + \varepsilon_i$$

where $Weight_i$ is the weight of the drugs seized, $Black_i$ is an indicator of whether the offender is recorded as black or white, X_i includes offender age and sex, and Z_s is a vector of state fixed effects. The coefficients in column 2 are estimated from the same specification with a dummy variable for the 280-290g range as the dependent variable. Coefficients in column 3 are estimated from the following regression:

$$(6) \quad Weight_{it} = \alpha_0 + \beta_1 (\text{Black} \times \text{After2010})_{it} + \beta_2 (\text{White} \times \text{After2010})_{it} + X_i + Z_s + \varepsilon_{it}$$

The coefficients in columns 4-10 are estimated from the same specification with dummy variables for the range of interest as the dependent variable.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6b: Drug Use and Drug Selling After the Fair Sentencing Act

	Ever Use Crack	Sold Drugs in Past Year	Use Crack & Sold Drugs
	(1)	(2)	(3)
After 2010 x White	0.0019**	-0.0009**	-0.0007***
	(0.0009)	(0.0005)	(0.0002)
After 2010 x Black or Hispanic	-0.0053***	-0.0031***	-0.0010***
	(0.0015)	(0.0009)	(0.0003)
Black or Hispanic	0.0033***	0.0039***	-0.0009***
	(0.0012)	(0.0007)	(0.0003)
Constant	0.0342***	0.0145***	0.0037***
	(0.0005)	(0.0007)	(0.0001)
Observations	763,335	762,322	762,054
P-value: W = BH	0.0000	0.0257	0.3350

Notes: Robust standard errors in parentheses. This analysis uses data from the National Survey on Drug Use and Health. Column 1 shows that the fraction of respondents answering "yes" to the question, "have you ever, even once, used crack-cocaine?" does not increase after 2010. Column 2 shows that the fraction of respondents answering a number greater than zero to the question, "how many times have you sold illegal drugs in the past 12 months?" does not increase after 2010. Column 3 shows that the fraction of people answering yes to both of these questions does not increase after 2010. All specifications use year-specific sampling weights. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Coefficients in are estimated from the following regression:

$$\text{Outcome}_{it} = \alpha_0 + \beta_1(\text{BlackOrHispanic} \times \text{After2010})_{it} + \beta_2(\text{White} \times \text{After2010})_{it} + \text{BlackOrHispanic}_i + \varepsilon_{it} \quad (7)$$

*** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.1794***	0.2170***	-0.0376
	(0.0629)	(0.0393)	(0.0461)
Constant	0.9184***	-	0.0816*
	(0.0435)	-	(0.0435)
Observations	989	989	989

Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0254)	(0.0172)	(0.0168)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes: Standard errors clustered at the prosecutor level and estimated jointly by seemingly unrelated regression in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(8) \quad (\text{Charged X-Yg})_i = \alpha_0 + \beta_1 \text{AttyBunchesAt280}g_i + \varepsilon_i$$

where AttyBunchesAt280g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases (for columns 1-3) and to prosecutors with 10+ cases post-2010. Note, column (2) is a mechanical relationship, hence the missing standard error. Table A3.22 shows that this result is robust to using leave-out-means to classify bunching attorneys. Coefficients in panel B are estimated from the following regression for each range:

$$(9) \quad (\text{Charged X-Yg})_i = \alpha_0 + \beta_1 \text{AttyBunchesAt50}g_i + \varepsilon_i$$

where AttyBunchesAt50g is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases (for columns 4-6) and to prosecutors with 10+ cases pre-2010. *** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Change in Bunching by Prosecutors after *Alleyne v. United States* Decision

	Pr(Case Recorded with 280-290g)			
	(1)	(2)	(3)	(4)
After June 17th, 2011-2016	0.0070	-0.0049	0.0041	-0.0206
	(0.0260)	(0.0284)	(0.0295)	(0.0406)
After June 17th, 2013	-0.1740**	-0.1518*	-0.1433	-0.1289
	(0.0813)	(0.0920)	(0.0935)	(0.1246)
Constant	0.1620	0.1626	0.1576	0.2093
	(0.1520)	(0.1519)	(0.1520)	(0.1776)
Bandwidth	±150 days	±130 days	±120 days	±60 days
Observations	1,937	1,672	1,513	754

Notes: Standard errors clustered at the date the case is received in parentheses. The estimates in this table are based on the EOUSA data. The coefficients above are estimated from the following regression discontinuity style model:

$$(10) \quad (\text{Recorded}\,280-290g)_{it} = \alpha_0 + \beta_1 \text{AfterJune17}_{it} + \beta_2 \text{DaysFrom}_{it} + \beta_3 (\text{AfterJune17} \times \text{DaysFrom})_{it} + \delta_1 (\text{AfterJune17} \times \text{Year2013})_{it} + \delta_2 (\text{DaysFrom} \times \text{Year2013})_{it} + \delta_3 (\text{AfterJune17} \times \text{DaysFrom} \times \text{Year2013})_{it} + D_{it} + \varepsilon_{it}$$

where AfterJune17 is a dummy variable equal to one for cases received after June 17th in each year, DaysFrom, the running variable, is the date the case was received centered at zero on June 17th, and Year2013 is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. The ±130 day bandwidth is selected from the Imbens-Kalyanaraman optimal bandwidth procedure for the year 2013. Figure 3.4 shows graphical evidence of the discontinuity in bunching around June 17, 2013. Figure A3.21 shows further robustness checks. *** p<0.01, ** p<0.05, * p<0.1

Table 3.9: Degree of Bunching Post-2010 by Race and Offender Characteristics

	Pr(280-290g)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After '10 x White (W)	0.0171**	0.0065	0.0143	0.0129**	0.0160**	0.0103*	0.0149*	-0.0018**	0.0085
	(0.0068)	(0.0063)	(0.0087)	(0.0062)	(0.0071)	(0.0058)	(0.0076)	(0.0009)	(0.0095)
After '10 x Black or Hispanic (BH)	0.0363***	0.0235***	0.0424***	0.0303***	0.0452***	0.0306***	0.0471***	0.0088***	0.0156***
	(0.0023)	(0.0072)	(0.0037)	(0.0024)	(0.0036)	(0.0025)	(0.0173)	(0.0015)	(0.0040)
After '10 x W x Char.	-0.0207***	0.0109	-0.0024	-0.0015	-0.0095	0.0089	-0.0074	0.0283***	0.0067
	(0.0072)	(0.0100)	(0.0109)	(0.0120)	(0.0107)	(0.0135)	(0.0110)	(0.0109)	(0.0123)
After '10 x BH x Char.	-0.0042	0.0131*	-0.0102**	0.0191***	-0.0163***	0.0157***	-0.0188	0.0686***	0.0250**
	(0.0061)	(0.0076)	(0.0046)	(0.0051)	(0.0044)	(0.0047)	(0.0183)	(0.0052)	(0.0118)
Constant	0.0032***	0.0022	0.0031**	0.0033***	0.0013*	0.0036***	0.0031**	0.0018**	0.0052**
	(0.0011)	(0.0015)	(0.0014)	(0.0011)	(0.0008)	(0.0012)	(0.0014)	(0.0009)	(0.0020)
Characteristic	College	Male	Above Med. Age	Weapon	Above Med. Crim. Hist. Points	Above Med. # of Counts	State Above Med. % of BH Cases	Conspiracy Charge	State Above Med. Racial Animus
P-value: W = BH	0.0074	0.0764	0.0031	0.0085	0.0002	0.0012	0.0885	0.0000	0.5114
P-value: W+Char. = BH+Char.	0.0000	0.0177	0.0043	0.0007	0.0078	0.0352	0.0183	0.0000	0.0440
Observations	52,389	49,049	52,712	52,233	52,725	52,742	52,692	52,745	51,679

Notes: Robust standard errors in parentheses for columns 1-6. Standard errors clustered at the state level are in parentheses for columns 7 and 9. "Characteristic" or "Char." represents a dummy variable that is an offender or case characteristic. The specific offender characteristic of interest is noted in the "Characteristic" row. For example, when the "Characteristic" is "College", then "Characteristic" is equal to one if the offender's educational attainment is college or more and is equal to zero if the offender's educational attainment is less than college. See Table 3.1 for notes on sample selection. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "P-value: W+Char. = BH+Char." reports the p-value from a test of the null hypothesis that the combined coefficients on "(After 2010 x White)+(After 2010 x White x Characteristic)" is equal to the combined coefficients on "(After 2010 x Black or Hispanic)+(After 2010 x Black or Hispanic x Characteristic)." Male is equal to one if the offender is male and equal to zero if not. Above median age is equal to one if the offender is above the median age for offenders and equal to zero if not. Weapon is equal to one if the offense involves a weapon and equal to zero if not. Above median crim. hist. points is equal to one if the offender has a criminal history score above the median criminal history score for offenders and equal to zero if not. Above the median # of other counts is equal to one if the offender has above the median number of other criminal counts for offenders and equal to zero if not. Column 7 examines differences in bunching for offenders convicted in states with above/below the median fraction of black and Hispanic cases. Column 8 tests for differences in bunching for offenders with a "drug conspiracy" charge versus those without. The final column examines differences in bunching for offenders convicted in states with above/below the median level of racial animus. The coefficients in columns 1-9 are estimated from the following regression:

$$
\begin{aligned}
(11) \quad (280-290g)_{it} = \alpha_0 + \beta_1(\text{After2010}\times\text{W})_{it} + \beta_2(\text{After2010}\times\text{BH})_{it} + \beta_3(\text{After2010}\times\text{W}\times\text{Characteristic}^H)_{it} \\
+ \beta_4(\text{After2010}\times\text{BH}\times\text{Characteristic}^H)_{it} + \beta_5\text{Characteristic}^H_{it} + \beta_6\text{BH}_{it} + \beta_5(\text{Characteristic}^H\times\text{BH})_{it} + \varepsilon_{it}
\end{aligned}
$$

*** p<0.01, ** p<0.05, * p<0.1

Table 3.10: Robustness Tests for Relationship between Racial Animus and the Racial Disparity Bunching at 280g

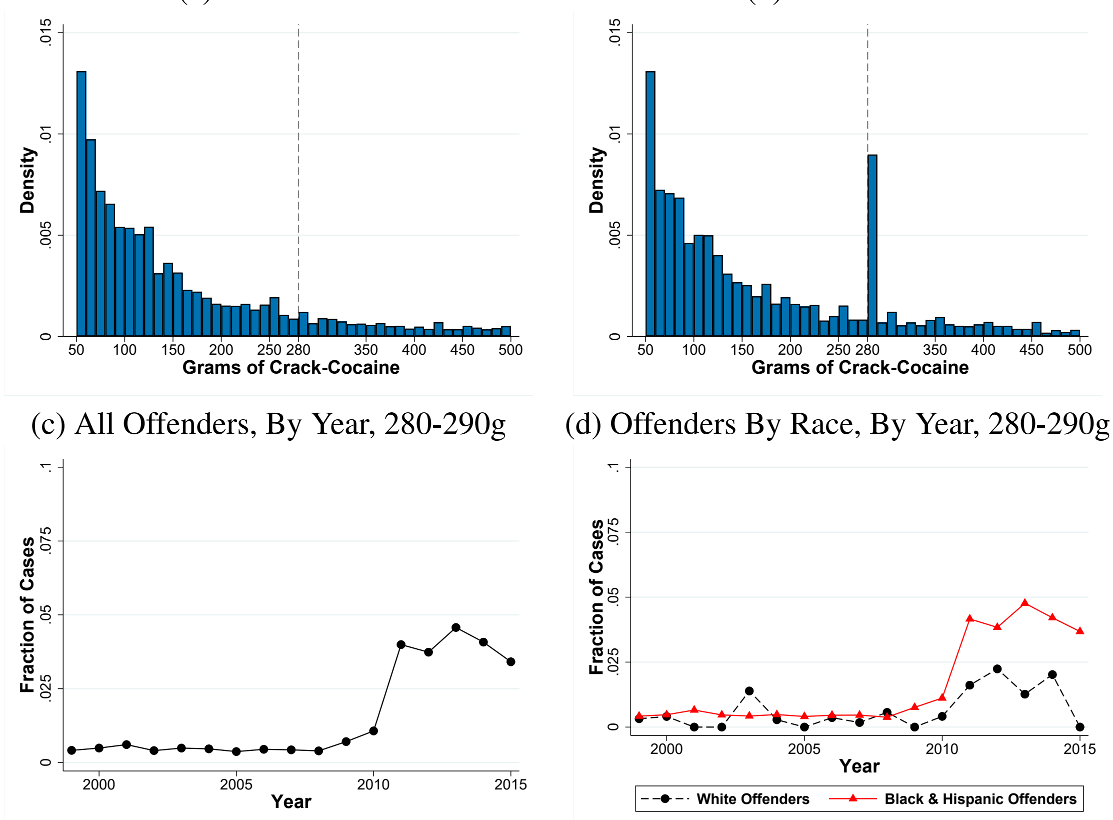
	Pr(280-290g)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After '10 x W x Above Med. Animus	0.0067	-0.0033	0.0063	-0.0047			
	(0.0123)	(0.0228)	(0.0128)	(0.0245)			
After '10 x BH x Above Med. Animus	0.0250**	0.0267**	0.0269**	0.0279**			
	(0.0118)	(0.0108)	(0.0124)	(0.0108)			
After '10 x W x Continuous Animus					0.0001	0.0008	
					(0.0004)	(0.0008)	
After '10 x BH x Continuous Animus					0.0007	0.0015***	
					(0.0005)	(0.0004)	
After '10 x IAT-Lawyers							-0.0075
							(0.0095)
After '10 x BH x IAT-Lawyers							0.0155
							(0.0105)
Constant	0.0052**	-0.0334	0.0040*	-0.0282	0.0099**	0.0111	0.0037
	(0.0020)	(0.0295)	(0.0023)	(0.0300)	(0.0044)	(0.0081)	(0.0059)
Other Controls Included	None	Offender Controls	District Economic Controls	Offender + District Controls	None	None	State x After 2010 x Race FEs
Sample Restrictions	None	None	None	None	None	Outliers Removed	None
Observations	51,679	51,679	47,692	47,692	51,679	49,188	51,679

Notes: Standard errors clustered at the state level in parentheses for columns 1-6. Standard errors clustered at the district level are in parentheses for column 7. See Table 3.1 for notes on sample selection. The first four columns examine differences in bunching for offenders convicted in states with above/below the median level of racial animus. Column 1 reports this result with no additional controls; column 2 introduces individual controls (college, male, age, criminal history, and state caseload) interacted with the after 2010 by race dummy variables; column 3 introduces district controls for economic characteristics (median household income in 2016, non-white share of population in 2010, population density in 2010, fraction with college in 2010, poor share in 2010, log of wage growth for high school graduates, black-white and Hispanic-white differences in incarceration and income conditional on parent income rank at the 25th percentile, job density in 2013, and annual job growth from 2004-2013) interacted with the after 2010 by race dummy variables; column 4 combines all controls from columns 2-3. Column 5 examines the relationship between animus and bunching using the continuous measure of animus from Google Trends, the p-value is less than 0.2 and the coefficient is several times larger than the coefficient for white offenders. Column 6 re-runs column 5 with outlier states (states with animus above the 99th percentile or below the 1st percentile) removed. Column 7 introduces a district level measure of animus, the implicit association test scores for lawyers (and other legal-service workers) aggregated to the district level. Since the measure is at the district level, I include state fixed effects interacted with the after 2010 by race dummy variables. The estimate is identified from within-state variation in the IAT-animus measure, and the p-value on the estimate is 0.14. The IAT measure is scaled to the median difference between the minimum and maximum score in states, meaning a one unit increase is approximately equivalent to moving from the minimum score in a state to the maximum score.
*** p<0.01, ** p<0.05, * p<0.1

Figure 3.1: Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010

(a) 1999-2010 (b) 2011-2015

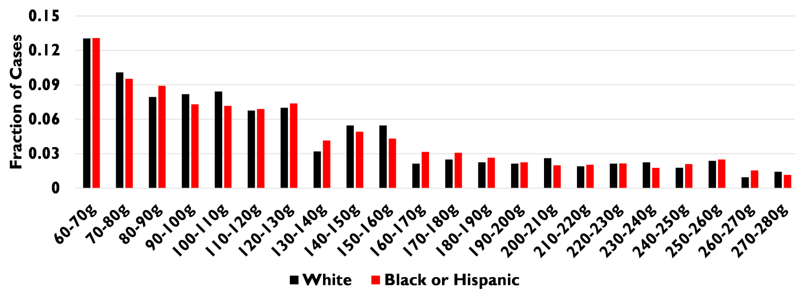
(c) All Offenders, By Year, 280-290g (d) Offenders By Race, By Year, 280-290g



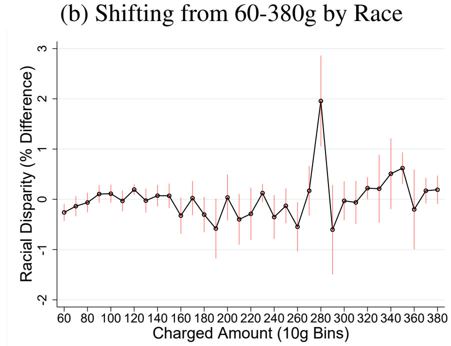
Notes: Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at 50 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). Panels (c) and (d) display the fraction of crack-cocaine cases with 280-290g by year, in general and by race. The denominator in panel (c) is all crack-cocaine cases under 1000g. The denominators in panel (d) are all crack-cocaine cases under 1000g, by race. Histograms showing the full density from 0-500g are in Figures A3.3a-b. Figures 1c-d with confidence intervals are in Figures A3.3c-d.

Figure 3.2: Testing for Conditional Racial Disparity in Bunching

(a) Distribution of Pre-2010 Charged Amount by Race, 60-280g



(b) Shifting from 60-380g by Race

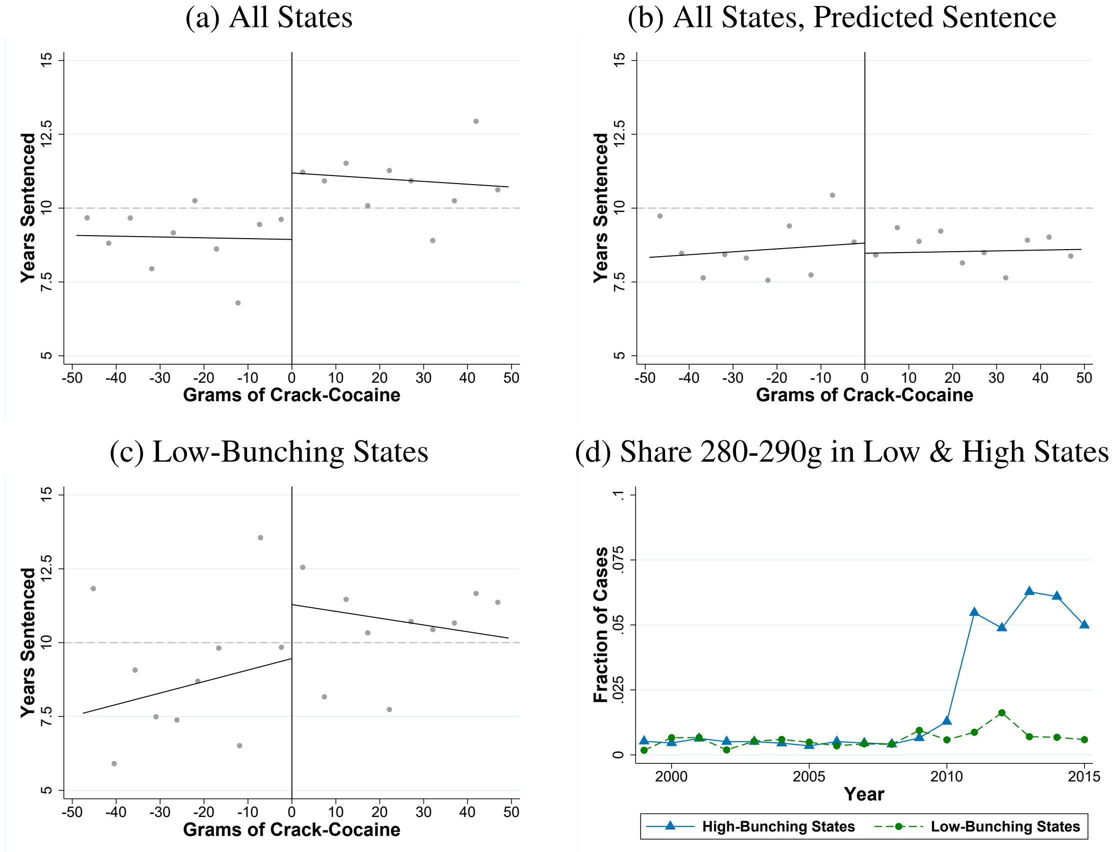


Notes: Panel (a) plots the distribution of charged amounts pre-2010 from 60-280g. A Kolmogorov-Smirnov test of the equality of the distributions by race fails to reject the null that the distributions are equal (p-value=0.788). Panel (b) plots the coefficient δ^X for each 10g bin starting at X divided by the share of cases in each 10g bin.

$$\begin{aligned} (12) \quad (\text{Charged X-Yg})_{it} &= \alpha + \delta^X(\text{After2010} \times \text{BlackOrHispanic})_{it} \\ &\quad + \gamma \text{After2010}_{it} + \lambda \text{BlackOrHispanic}_i + \varepsilon_{it} \end{aligned}$$

The plot shows these estimates for amounts from 0-380g, at higher amounts the estimates are more noisy. Figure A3.8 shows the estimates up to 1000g.

Figure 3.3: Sentencing Consequences of Crossing the Mandatory Minimum Threshold



Notes: Figure 3.3a plots the average sentence (within each 5g bin) from 230-330g for cases sentenced after 2010. A linear fit is estimated on each side of the 280g threshold. The estimated sentencing discontinuity is about 2.25 years ($se = 0.85$). Figure 3.3b is the same plot but using predicted sentence from a model of sentencing and offender characteristics using pre-2010 data. There is no discontinuity in this figure, suggesting that offenders bunched at 280g are not negatively selected on characteristics that would increase sentence length in the absence of the threshold. Figure 3.3c is the same plot but limited to the subset of states that have low-levels of bunching. The estimated discontinuity is about 2.00 years ($se = 1.73$). Figure 3.3d plots the share of cases with 280-290g by year for low- and high-bunching states. The coefficients described above are estimated from the regression:

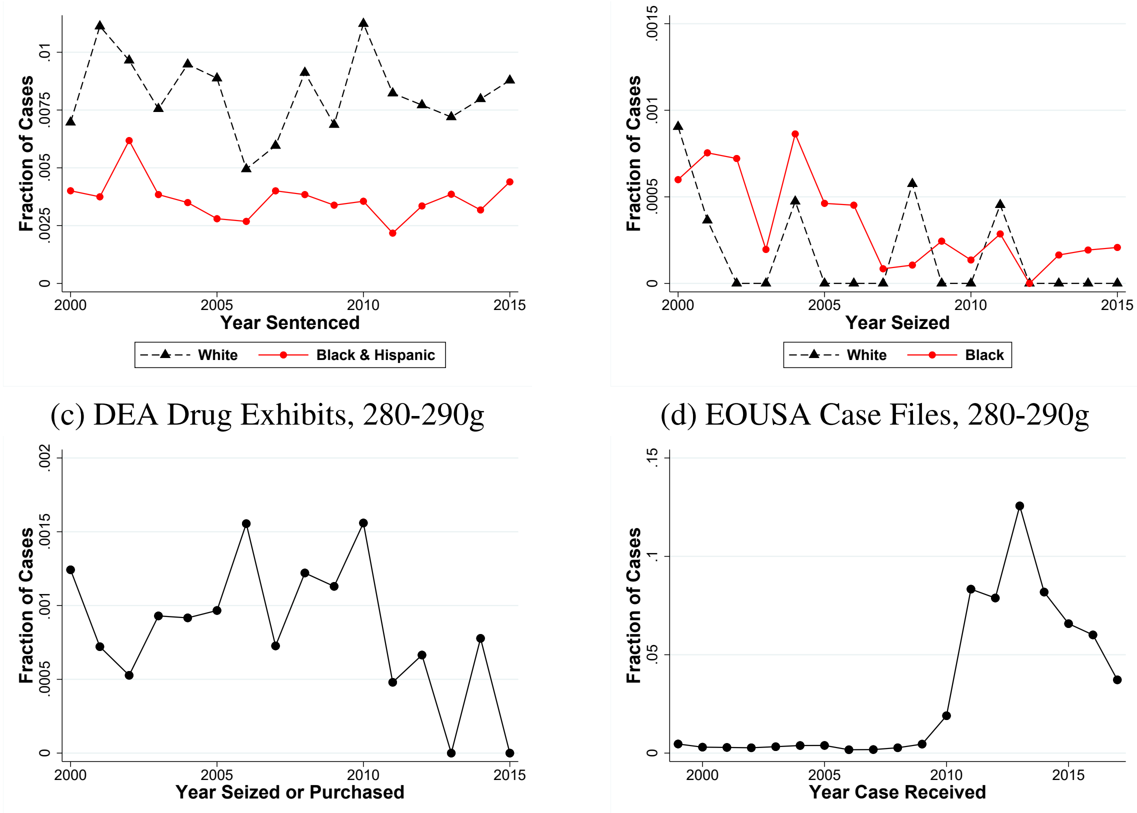
$$(13) \quad \text{Sentence}_i = \alpha_0 + \beta_1 \text{Amount}_i + \delta_1 \text{Above280}_i + \phi_1 (\text{Amount} \times \text{Above280}_i) + \varepsilon_i$$

δ_1 is the estimated discontinuity (reported in the preceding notes) in sentencing due to crossing the mandatory minimum threshold.

Figure 3.4: Changing Fraction of Cases at Various Stages of Criminal Justice System

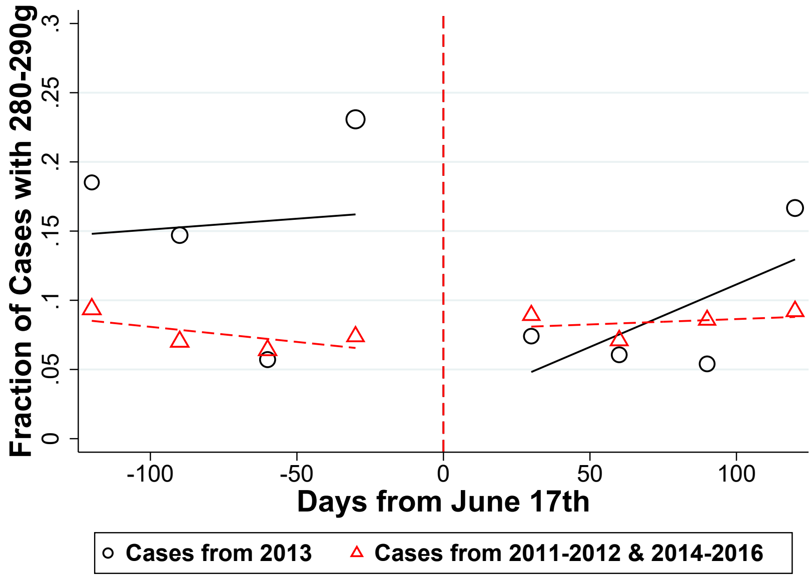
(a) Florida Convictions, By Race, 200-400g (b) NIBRS Drug Seizures, By Race, 280-290g

(c) DEA Drug Exhibits, 280-290g (d) EOUSA Case Files, 280-290g



Notes: Please note the different y-axis scales, particularly in the case of panels (b) and (c). Panel (a) plots the fraction of cocaine offenses that have a range from 200-400g in FL state prison from 2000-2015, by race. The denominators are all cocaine offenses in FL, by race. Panel (b) plots the fraction of crack-cocaine drug seizures made by local police departments and recorded as 280-290g from 2000-2015, by race. Panel (c) plots the fraction of cocaine drug exhibits sent to DEA laboratories and recorded as 280-290g from 2000-2015 (the DEA data does not include race). The denominator is all cocaine exhibits in the DEA STRIDE data. Results are similar if limited to "cocaine hydrochloride" or "cocaine base." Panel (d) plots the fraction of crack-cocaine cases recorded as 280-290g in the EOUSA caseload data (the EOUSA data does not include race). The denominator is all crack-cocaine cases in the EOUSA data with non-missing drug quantities. The EOUSA data contains many more missing values than the USSC data. Imputing missing drug weights as zero does not fundamentally change the results.

Figure 3.5: Change in Bunching by Prosecutors after *Alleyne v. United States* Decision



Notes: Panel (a) plots the fraction of cases with 280-290g in each 30-day bin for 120 days before and 120 days after June 17th. The black circles show the fraction of cases in each bin for 2013 and the red triangles show the average fraction of cases in each bin for 2011-2012 and 2014-2016. The solid black line shows a linear fit on each side of the June 17, 2013 and the dashed red line shows a linear fit on each side of June 17 for all other years. The scatter plot symbols are weighted by the total number of cases in each bin. The estimated discontinuity is $\delta = -0.1433$ and $se = 0.0935$ and is estimated from the following regression:

$$
\begin{aligned}
(14) \quad (280-290g)_{it} = {} & \alpha_0 + \beta_1 \text{AfterJune17}_{it} + \beta_2 \text{DaysFrom}_{it} + \beta_3 (\text{After}\times\text{DaysFrom})_{it} \\
& + \delta_1 (\text{AfterJune17}\times\text{Year2013})_{it} + \delta_2 (\text{DaysFrom}\times\text{Year2013})_{it} \\
& + \delta_3 (\text{After}\times\text{DaysFrom}\times\text{Year2013})_{it} + D_{it} + \varepsilon_{it}
\end{aligned}
$$

where $After_{it}$ is equal to one if case **i** is received after June 17th of year **t** but before January 1st of year **t+1** and is equal to zero if case **i** is received before June 17th of year **t** but after January 1st of year **t**. $DaysFrom_{it}$ is the number of days from June 17th that case **i** is received, and $Year2013_{it}$ is equal to one if case **i** is received in 2013 and is equal to zero if it is received in 2011-2012 or 2014-2016. D_{it} represents day-of-week fixed effects. The coefficient β_1 is the average discontinuity in the fraction of cases with 280-290g after June 17 from 2011-2016. The coefficient δ_1 is the discontinuity that is specific to June 17, 2013–the date of the *Alleyne* decision.

3.8 Appendix A. Additional Tables and Figures

Table A3.1: Summary Statistics for FL, NIBRS, and DEA Records

	Pre-2010	Post-2010	Observations
Panel A. Cocaine Felony Convictions in FL			
200-400g	0.00474	0.00432	214,573
	(0.0687)	(0.0656)	
28-200g	0.0405	0.0473	214,573
	0.197	(0.212)	
Missing drug weight	0.945	0.936	214,573
	(0.228)	(0.245)	
Black or Hispanic	0.771	0.789	214,573
	(0.420)	(0.408)	
Panel B. NIBRS Drug Seizures, Balanced Panel			
Weight (g)	10.33	7.76	203,700
	(46.19)	(44.87)	
280-290g	0.000360	0.000141	203,700
	(0.0190)	(0.0119)	
Black	0.737	0.746	191,774
	(0.440)	(0.435)	
Male	0.837	0.834	192,721
	(0.370)	(0.372)	
Panel C. DEA Drug Seizures			
Weight (g)	78.28	67.28	100,306
	(188.83)	(176.54)	
280-290g	0.00102	0.000428	100,306
	(0.0319)	(0.0207)	
Seized (vs. Purchased)	0.529	0.544	100,302
	(0.499)	(0.498)	
Price per gram (median)	42.02	47.62	37,820

Notes: The table above describes offenders found in the FL inmate database, the NIBRS drug seizure records, and the DEA drug exhibit data pre- and post-2010 (the DEA data actually describes the drugs themselves, not the offenders). The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable. The statistics above are derived from the cleaned data in which the following cases are removed for NIBRS and DEA: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded. 280-290g is a dummy variable equal to one when the weight is from 280-290g and zero when it is from 0-280g and 290-1000g, and missing when it is missing. The 200-400g and 28-200g variables follow the same logic. Missing drug weight is equal to one when the drug weight is missing. "Seized (vs. Purchased)" is equal to one if the DEA obtained the drug exhibit from a seizure versus an undercover purchase. The median price per gram is reported after removing outliers above the 95th percentile and below the 5th percentile.

Table A3.2: Summary Statistics for EOUSA Prosecutor Case Files

	Pre-2010	Post-2010	Observations
Weight (g)	72.500	97.966	19,363
	(135.219)	(162.538)	
280-290g	0.004	0.082	19,363
	(0.062)	(0.274)	
280-290g, Missing = 0	0.002	0.026	49,342
	(0.040)	(0.158)	
50-60g	0.210	0.082	19,363
	(0.408)	(0.274)	
50-60g, Missing = 0	0.086	0.026	49,342
	(0.280)	(0.158)	
Missing drug weight	0.593	0.686	49,342
	(0.491)	(0.464)	
Only Federal Law Enforcement Involved	0.642	0.647	48,501
	(0.479)	(0.478)	
Any Federal Law Enforcement Involved	0.737	0.713	48,501
	(0.440)	(0.452)	
Lead Charge = Conspiracy	0.212	0.217	46,335
	(0.409)	(0.412)	

Notes: The table above describes defendants found in the EOUSA prosecutor case management data pre- and post-2010. The mean value of each variable is reported with standard deviations in parentheses. Observation counts are displayed separately for each variable since some fields in this data are missing much more often than others. The statistics above are derived from the cleaned data in which the following cases are removed: cases with drug weights above 1000g. Weight is the weight of the drugs in grams recorded in the case management system. 280-290g is a dummy variable equal to one when the weight is from 280-290g, zero when it is from 0-280g and 290-1000g, and missing when it is missing.. "280-290g, Missing=0" is a dummy variable equal to "280-290g" but coded equal to zero when the weight field is missing. The 50-60g variables follow the same logic. Missing drug weight is equal to one when the drug weight is missing. "Only Federal Law Enforcement" is equal to one when the agency recorded as sending the case is strictly federal (i.e. DEA, FBI, or ATF) and equal to zero otherwise. "Any Federal" is equal to one if the agency sending the case has any federal involvement (i.e. "Joint DEA and state/local task force") and equal to zero otherwise. "Lead Charge = Conspiracy" is equal to one when the lead charge for the case is a drug conspiracy charge.

Table A3.3: Result Robust to Other Drug Weight Sample Restrictions

	Pr(280-290g Crack-Cocaine)				
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0119**	0.0115**	0.0115**	0.0844***	0.0258**
	(0.0050)	(0.0049)	(0.0049)	(0.0131)	(0.0116)
After 2010 x Black or Hispanic	0.0345***	0.0329***	0.0328***	0.1186***	0.0718***
	(0.0021)	(0.0020)	(0.0020)	(0.0040)	(0.0042)
Constant	0.0031***	0.0030***	0.0030***	0.0034***	0.0088***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0027)
P-value: W = BH	0.0000	0.0000	0.0001	0.0127	0.0002
Sample Restriction	0-2500g	0-25000g	No Restriction	0-1000g	50-1000g
Includes Weights Coded as a Range	No	No	No	Yes	No
Observations	55,729	58,116	58,645	59,677	24,905

Notes: Robust standard errors in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Columns 1-3 include outliers to varying extents. Column 4 reports results when the sample includes quantities coded as a range (in this analysis, the lower bound of the range is used). Column 5 excludes drug weights below 50g (i.e. excluding weights close to the 5-year mandatory minimum pre- and post-2010).
*** p<0.01, ** p<0.05, * p<0.1

Table A3.4: Result Robust to Various Sample Restrictions

	Pr(280-290g Crack-Cocaine)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0314***		0.0336***		0.0304***	
	(0.0021)		(0.0021)		(0.0022)	
After 2010 x White		0.0125**		0.0128**		0.0128**
		(0.0053)		(0.0054)		(0.0054)
After 2010 x Black or Hispanic		0.0327***		0.0348***		0.0317***
		(0.0022)		(0.0022)		(0.0023)
Constant	0.0053***	0.0032***	0.0062***	0.0030**	0.0063***	0.0030**
	(0.0004)	(0.0010)	(0.0006)	(0.0015)	(0.0006)	(0.0015)
P-value: W = BH	-	0.0004	-	0.0002	-	0.0013
Hispanic Offenders Excluded	Yes	Yes	No	No	Yes	Yes
Post-2006 Data Only	No	No	Yes	Yes	Yes	Yes
Observations	47,763	47,763	25,893	25,846	23,241	23,241

Notes: Robust standard errors in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Post-2006 Data Only" is equal to "Yes" when the data is limited to cases brought to court from 2007-2015 (after the *Booker v. United States* Supreme Court case that made sentencing guidelines optional, excluding mandatory minimum guidelines). The row "Hispanic Offenders Excluded" is equal to "Yes" when Hispanic offenders are removed from the sample.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3.5: Result Robust to Other Categorizations of Bunching

	Pr(280-300g)	Pr(280-320g)	Pr(280-380g)
	(1)	(2)	(3)
After 2010 x White	0.0154**	0.0146**	0.0137*
	(0.0061)	(0.0067)	(0.0083)
After 2010 x Black or Hispanic	0.0360***	0.0367***	0.0394***
	(0.0022)	(0.0025)	(0.0029)
Constant	0.0055***	0.0099***	0.0230***
	(0.0013)	(0.0017)	(0.0026)
P-value: W = BH	0.0016	0.0019	0.0033
Observations	52,745	52,745	52,745

Notes: Robust standard errors in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." Each column corresponds to a different definition of what it means for a case to be "bunched" above the mandatory minimum threshold. For the main results, I define a result as "bunched" if it is in the narrow range of 280-290g. In columns 1-3, I use alternative ranges: 280-300g, 280-320g, and 280-380g.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3.6: Result Robust to Controls and Alternative Std. Errors

	Pr(280-290g Crack-Cocaine)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010	0.0347***		0.0348***		0.0345***		0.0327***		0.0322***	
	(0.0082)		(0.0081)		(0.0079)		(0.0068)		(0.0066)	
After 2010 x White		0.0125**		0.0130**		0.0136**		0.0118*		0.0138**
		(0.0058)		(0.0059)		(0.0062)		(0.0060)		(0.0066)
After 2010 x Black or Hispanic		0.0360***		0.0363***		0.0358***		0.0340***		0.0333***
		(0.0086)		(0.0086)		(0.0084)		(0.0073)		(0.0071)
Constant	0.0051***	0.0032***	0.0085***	0.0064**	0.0088**	0.0085**	0.0078*	0.0074*	0.0082**	0.0075**
	(0.0005)	(0.0010)	(0.0029)	(0.0031)	(0.0035)	(0.0037)	(0.0043)	(0.0044)	(0.0031)	(0.0033)
P-value: W = BH	-	0.0181	-	0.0184	-	0.0282	-	0.0286	-	0.0695
Offender Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year Trend	No	No	No	No	No	No	Yes	Yes	Yes	Yes
State-specific Trends	No	No	No	No	No	No	No	No	Yes	Yes
Observations	56,826	52,692	51,813	51,746	51,813	51,746	51,804	51,737	51,804	51,737

Notes: Standard errors clustered at the state-level in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Offender Controls" indicates if the following offender-level controls are included: criminal history points, age, citizenship, number of current offense counts, whether a weapon was involved, and education. The rows "State Fixed Effects" and "Year Trend" indicate if the specification includes state fixed effects or a year trend as controls. The row "State-specific Trends" indicates if the specification includes state-specific linear trends. In all cases, there is a sharp increase in the fraction of cases with 280-290g after 2010 and a racial disparity in that increase by race.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.7: Result Robust to Probit, Logit, and Poisson Models

	Probit			Logit			Poisson			OLS		
	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g	280-290g		280-380g
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After 2010 x W	0.5747***	0.5606***	0.2046*	1.6031***	1.4119***	0.4804*	1.1208***	1.1615***	0.1102	0.0125**	0.0252**	0.0137*
	(0.1651)	(0.1840)	(0.1085)	(0.4518)	(0.4546)	(0.2498)	(0.4042)	(0.2758)	(0.5745)	(0.0053)	(0.0113)	(0.0083)
After 2010 x BH	0.8159***	0.9008***	0.3851***	2.0784***	2.0895***	0.8400***	2.1129***	2.1042***	0.8604	0.0360***	0.0710***	0.0394***
	(0.0337)	(0.0374)	(0.0235)	(0.0869)	(0.0878)	(0.0500)	(0.3645)	(0.2726)	(0.6351)	(0.0021)	(0.0042)	(0.0029)
Constant	-2.7258***	-2.3912***	-1.9948***	-5.7392***	-4.7715***	-3.7476***	3.5423***	2.6237***	3.6109***	0.0032***	0.0084***	0.0230***
	(0.0994)	(0.1102)	(0.0470)	(0.3020)	(0.3028)	(0.1138)	(0.3624)	(0.2202)	(0.3624)	(0.0010)	(0.0025)	(0.0026)
P-value: W = BH	0.1524	0.0701	0.1041	0.3015	0.1433	0.1580	0.0157	0.0007	0.3286	0.0000	0.0001	0.0033
Sample	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g	0-1000g	50-1000g	0-1000g
Observations	52,745	25,647	52,745	52,745	25,647	52,745	400	380	400	52,745	25,647	52,745

Notes: Robust standard errors in parentheses. The row “P-value: W = BH” reports the p-value from a test of the null hypothesis that the coefficient on “After 2010 x W” is equal to the coefficient on “After 2010 x BH,” where “W” is the “White” dummy variable and “BH” is the “Black or Hispanic” dummy variables (abbreviated for table space). In general, columns 1-3 estimate probit models, columns 4-6 estimate logit models, columns 7-9 estimate Poisson models (on binned data), and columns 10-12 estimate OLS (or linear probability) models. Columns 1, 4, 7, and 10 estimate the change in bunching at 280-290g after 2010 for all cases from 0-1000g. Columns 2, 5, 8, and 11 limit the sample to cases from 50-1000g (following column 5 of Table A3.3). Columns 3, 6, 9, and 12 extend the “bunching” definition to 280-380g (following column 3 of Table A3.5).

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

Table A3.8: Result Robust to Concerns about Selection Into/Out of Missing and Selection Into/Out of Other Drugs

	Pr(280-290g)			
	(1)	(2)	(3)	(4)
After 2010 x White	0.0583***	0.0242***	0.0005	0.0727***
	(0.0087)	(0.0059)	(0.0003)	(0.0032)
After 2010 x Black or Hispanic	0.0833***	0.0441***	0.0093***	0.2030***
	(0.0027)	(0.0021)	(0.0006)	(0.0031)
Constant	0.0033***	0.0024***	0.0004***	0.8680***
	(0.0009)	(0.0007)	(0.0001)	(0.0021)
P-value: W = BH	0.0063	0.0016	0.0000	0.0000
Drugs included	Crack-cocaine	Crack-cocaine	All	All
Dependent variable recoded to	Lower value of weight range	Upper value of weight range	Non-crack cases = 0	Non-crack cases = 1
Selection issue addressed	Into/out of missing weight	Into/out of missing weight	Into/out of other drugs	Into/out of other drugs
Observations	67,040	65,003	149,428	149,428

Notes: Robust standard errors in parentheses. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "Drugs included" indicates the type of drugs included in the analysis. In columns 1 and 2, I focus on the crack-cocaine sample to analyze how including missing exact weights (i.e. weights recorded as ranges) affects the results. In columns 3 and 4, I focus on the sample of all drugs to analyze how movement of cases into or out of other drug types affects the results. The row "Dependent variable recoded to" indicates how the dependent variable is recoded in each analysis. In column 1, the dependent variable is recoded as 1 if the lower bound of the weight range is between 280-290g and recoded as 0 otherwise. In column 2, it is recoded as 1 if the upper bound of the range is between 280-290g and recoded as 0 otherwise. Results are also robust to recoding all missings as (In 280-290)=0 or recoding all missings as (In 280-290)=1. In column 3, the dependent variable is recoded as 0 if the case is not a crack-cocaine case, and in column 4, it is recoded as 1 if the case is not a crack-cocaine case. Finally, the row "Selection issue addressed" indicates the type of selection issue being investigated in each column. In all columns, I find that the probability of being in the 280-290g range for crack-cocaine increases after 2010 and increases disproportionately for black and Hispanic offenders, regardless of selection into missing exact weights or other drug types.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.9: Difference-in-Difference Bunching Identification

	Pr(280-290g)				Pr(50-60g)	
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0011*	-0.0002				
	(0.0006)	(0.0011)				
After 2010 x Crack-cocaine	0.0336***	0.0127**				
	(0.0021)	(0.0054)				
After 2010 x Crack-cocaine x Black or Hispanic		0.0217***				
		(0.0059)				
Crack-cocaine	-0.0020***	-0.0042***	-0.0036**	0.0088	0.0151***	0.0210*
	(0.0005)	(0.0011)	(0.0016)	(0.0058)	(0.0053)	(0.0122)
Crack-cocaine x Black or Hispanic			0.0020	0.0229***	0.0108*	-0.0021
			(0.0017)	(0.0063)	(0.0057)	(0.0127)
Constant	0.0072***	0.0074***	0.0068***	0.0070***	0.0502***	0.0438***
	(0.0003)	(0.0006)	(0.0012)	(0.0026)	(0.0033)	(0.0065)
Drugs Included	All	All	Crack & Powder	Crack & Powder	Crack & Powder	Crack & Powder
Years Included	1999-2015	1999-2015	1999-2010	2011-2015	1999-2010	2011-2015
Observations	149,428	149,428	65,475	17,307	65,475	17,307

Notes: Robust standard errors in parentheses. Columns 1-2 compare crack-cocaine cases to all other drug cases. Specifically, they estimate the change in the probability a case is recorded with 280-290g after 2010 both for crack-cocaine and for other drugs. Column 1 does this in general and column 2 does this by race. This amounts to a difference-in-difference (pre- vs. post-2010 and crack vs. non-crack) estimation of the bunching (as opposed to the pre- vs. post-2010 difference that is the focus of the paper). Columns 3-6 apply this same design to estimate the probability of being recorded with 280-290g and 50-60g before and after 2010. These columns compare crack to powder cocaine alone since powder cocaine is a drug that never has a 50g mandatory minimum threshold.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.10a: Missing Mass in the Distribution of Drug Amounts by Race

Panel A. Analysis of Changes in the 0-100g Range					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	-0.0030	-0.1162***	0.0326**	-0.0006	0.0189
	(0.0179)	(0.0188)	(0.0149)	(0.0111)	(0.0143)
After 2010 x Black or Hispanic	0.0222***	-0.0696***	0.0341***	-0.0066**	-0.0100***
	(0.0039)	(0.0050)	(0.0041)	(0.0029)	(0.0038)
Constant	0.1971***	0.3242***	0.0968***	0.0653***	0.0965***
	(0.0068)	(0.0080)	(0.0050)	(0.0042)	(0.0050)
P-value: W = BH	0.1669	0.0164	0.9216	0.6000	0.0503
Observations	52,745	52,745	52,745	52,745	52,745
Panel B. Analysis of Changes in the 100-1000g Range					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(1)	(2)	(3)	(4)	(5)
After 2010 x White	0.0028	0.0125**	0.0137	0.0099	0.0294***
	(0.0162)	(0.0053)	(0.0096)	(0.0070)	(0.0090)
After 2010 x Black or Hispanic	-0.0165***	0.0360***	0.0044*	0.0016	0.0044**
	(0.0045)	(0.0021)	(0.0025)	(0.0018)	(0.0020)
Constant	0.1493***	0.0032***	0.0353***	0.0163***	0.0160***
	(0.0061)	(0.0010)	(0.0032)	(0.0022)	(0.0021)
P-value: W = BH	0.2503	0.0000	0.3470	0.2539	0.0066
Observations	52,745	52,745	52,745	52,745	52,745

Notes: Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic."
*** p<0.01, ** p<0.05, * p<0.1

Table A3.10b: Missing Mass in the Distribution of Drug Amounts by Race, with Various Time Trend Controls and State FEs

	Pr(< 280g)	Pr(280-290g)	Pr(> 290g)
	(1)	(2)	(3)
Panel A. No Interaction with Time Trend			
After 2010 x White	-0.0685***	0.0120**	0.0566***
	(0.0151)	(0.0055)	(0.0143)
After 2010 x Black or Hispanic	-0.0602***	0.0343***	0.0259***
	(0.0051)	(0.0023)	(0.0047)
Constant	0.9372***	0.0059***	0.0569***
	(0.0053)	(0.0013)	(0.0052)
P-value: W = BH	0.5840	0.0001	0.0330
Observations	52,678	52,678	52,678
Panel B. Interaction with Linear Time Trend			
After 2010 x White	-0.0403*	0.0164**	0.0240
	(0.0229)	(0.0083)	(0.0218)
After 2010 x Black or Hispanic	-0.0601***	0.0345***	0.0256***
	(0.0064)	(0.0033)	(0.0057)
Constant	0.9078***	0.0043**	0.0880***
	(0.0100)	(0.0020)	(0.0098)
P-value: W = BH	0.4063	0.0418	0.9418
Observations	52,678	52,678	52,678
Panel C. Interaction with Quadratic Time Trends			
After 2010 x White	0.0031	0.0133	-0.0164
	(0.0303)	(0.0099)	(0.0291)
After 2010 x Black or Hispanic	-0.0256***	0.0301***	-0.0045
	(0.0085)	(0.0040)	(0.0078)
Constant	0.8789***	0.0038	0.1173***
	(0.0192)	(0.0040)	(0.0188)
P-value: W = BH	0.3614	0.1150	0.6933
Observations	52,678	52,678	52,678

Notes: Robust standard errors in parentheses. The estimates in this table are based on the USSC data. See Table 3.1 for notes about sample selection. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The general model I estimate is:

$$(Charged\,X - Y g)_{it} = \alpha_0 + \beta_1(After2010 \times W)_{it} + \beta_2(After2010 \times BH)_{it} + \delta_1(After2010 \times W \times Trend)_{it}$$
$$+ \delta_2(After2010 \times BH \times Trend)_{it} + \gamma_1 BH + \phi_1(BH \times Trend) + Z_i + g(t)_t + \varepsilon_{it}$$

$Trend$ takes on the value of zero (i.e. no trend interaction), a linear trend, or a quadratic trend. $g(t)_t$ is a linear trend when no trend interactions are used and when the linear trend interaction is used. $g(t)_t$ is a quadratic trend when the quadratic trend interactions are used. Figures A3.7j-k show the total share of cases below 280g and above 280g over time, by race. For these shares, there are considerable trends over time, especially for white offenders. To quantify the break in those trends after 2010, I estimate case-level regressions that interact the dummy variable for after 2010 with a linear time trend centered at zero in 2011. Panel (a) shows the estimates without accounting for these time trends, and as a result, column 3 indicates that white offenders are more likely to be charged with amounts greater than 290g after 2010, relative to black and Hispanic offenders. This is true, but it is due to a substantial rise in cases above 290g for white offenders that begins in 2005. Panels (b) and (c) account for this by estimating the break in the trend after 2010. Both panels indicate that white, black, and Hispanic offenders have similar (and small) trend breaks in their share of cases above 290g. Likewise, both panels show bunching at 280-290g, a racial disparity in bunching, and evidence that the excess mass at 280-290g is drawn from cases that would have been charged below 280g prior to 2010. All specifications include state fixed-effects (Z_i).
*** p<0.01, ** p<0.05, * p<0.1

Table A3.10c: Missing Mass in the Distribution of Drug Amounts, Post-2007 Only

Panel A. Analysis of Changes in the 0-100g Range.					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0246***	-0.0710***	0.0323***	-0.0098***	-0.0120***
	(0.0042)	(0.0055)	(0.0044)	(0.0033)	(0.0042)
Constant	0.1065***	0.2920***	0.1134***	0.0751***	0.1263***
	(0.0024)	(0.0035)	(0.0025)	(0.0021)	(0.0026)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	25,893	25,893	25,893	25,893	25,893
Panel B. Analysis of Changes in the 100-1000g Range					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0108**	0.0336***	0.0050*	0.0026	0.0056**
	(0.0050)	(0.0021)	(0.0027)	(0.0019)	(0.0022)
Constant	0.1886***	0.0062***	0.0443***	0.0207***	0.0269***
	(0.0031)	(0.0006)	(0.0016)	(0.0011)	(0.0013)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	25,893	25,893	25,893	25,893	25,893

Notes: Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and sentenced from 2007-2015. The predicted change from the conceptual model of prosecutor behavior in Section 3.2.2 is displayed in the row labeled "predicted change from conceptual model."
*** p<0.01, ** p<0.05, * p<0.1

Table A3.10d: Missing Mass in the Distribution of Drug Amounts, Trial Cases Only

Panel A. Analysis of Changes in the 0-100g Range					
	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-60g)	Pr(60-100g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	0.0294	-0.0530**	0.0248	-0.0120	-0.0591***
	(0.0181)	(0.0216)	(0.0171)	(0.0137)	(0.0126)
Constant	0.1104***	0.2592***	0.0984***	0.0831***	0.1112***
	(0.0064)	(0.0089)	(0.0061)	(0.0056)	(0.0064)
Predicted Change from Conceptual Model	Increase	Decrease	Increase	Decrease	Decrease
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.030	0.020	0.006	0.008	0.007
Panel B. Analysis of Changes in the 100-1000g Range					
	Pr(100-280g)	Pr(280-290g)	Pr(290-470g)	Pr(470-600g)	Pr(600-1000g)
	(1)	(2)	(3)	(4)	(5)
After 2010 (Actual Change)	-0.0392**	0.0749***	0.0030	0.0217*	0.0085
	(0.0199)	(0.0131)	(0.0124)	(0.0111)	(0.0111)
Constant	0.2050***	0.0033***	0.0562***	0.0281***	0.0389***
	(0.0082)	(0.0012)	(0.0047)	(0.0034)	(0.0039)
Predicted Change from Conceptual Model	Decrease	Increase	No Change	No Change	No Change
Observations	2,841	2,841	2,841	2,841	2,841
R-squared	0.012	0.022	0.007	0.006	0.010

Notes: Robust standard errors in parentheses. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams and cases that end in a jury trial. The predicted change from the conceptual model of prosecutor behavior in Section 3.2.2 is displayed in the row labeled "predicted change from conceptual model."

*** p<0.01, ** p<0.05, * p<0.1

Table A3.11: Sentencing Consequences of Being Above the Threshold Amount

	Years Sentenced					
	(1)	(2)	(3)	(4)	(5)	(6)
Above 280g	-0.580**	0.0621			0.00410	-0.0576
	(0.289)	(0.691)			(0.294)	(0.461)
Above 280g x After 2010	2.332***	2.181**			0.971*	2.836***
	(0.508)	(1.102)			(0.535)	(0.842)
Above 50g			0.755***	0.955***	1.469***	2.101***
			(0.128)	(0.158)	(0.180)	(0.227)
Above 50g x After 2010			-1.387***	-1.063***	-1.298***	-2.058***
			(0.270)	(0.357)	(0.451)	(0.445)
Constant	12.93***	11.48***	9.664***	9.540***	13.12***	14.08***
	(0.170)	(0.565)	(0.114)	(0.116)	(3.298)	(3.709)
Bandwidth	±250g	±50g	±250g	±50g	±250g	±250g
Includes Life & <1 Month	No	No	No	No	No	Yes
Observations	29,767	2,800	49,154	14,713	29,064	31,134
R-squared	0.037	0.015	0.070	0.035	0.038	0.031

Notes: Robust standard errors in parentheses. The estimates in this table are based on the USSC data. The coefficients in columns 1-2 are estimated from the following regression discontinuity style model:

$$Sentence_{it} = \alpha + \beta_1 Above280_{it} + \beta_2 Amount_{it} + \beta_3 (Above280 \times Amount)_{it} + \delta_1 (Above280 \times After2010)_{it}$$
$$+ \delta_2 (Amount \times After2010)_{it} + \delta_3 (Above280 \times Amount \times After2010)_{it} + g(t)_t + \varepsilon_{it}$$

where $Amount_{it}$, the running variable, is the amount of drugs centered at the 280g mandatory minimum, $After2010_{it}$ is a dummy variable equal to one if the case is sentenced after 2010, and $Above280_{it}$ is a dummy variable equal to one if the case involves 280g or more of crack-cocaine. Columns 3-4 estimate equation (4) around the 50g threshold instead of the 280g threshold. Columns 5-6 estimate the sentencing penalty around the 50g threshold and the 280g threshold simultaneously. In addition, all specifications above include a time trend to capture the gradual decline in sentences over time. Column 6 includes life sentences (coded as 70 years) and sentences less than 1 month (coded as 0 years). I do not find significant differences in these sentencing discontinuities by race. I include the R-squared in this table because the dependent variable is continuous. Figures 3.3a-d show graphical evidence of the sentencing penalty. Figure A3.9 shows that the estimate of the sentencing penalty from model (5) is robust to many different bandwidths from 10g to 250g.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.12: Bunching Analysis for Potential Mechanisms, Alternative Results

Panel A. Analysis of Bunching in State Convictions and in Drug Seizures				
	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)	(4)
After 2010	0.00358		0.0185	
	(0.00873)		(0.0444)	
After 2010 x White		0.0068		-0.0008
		(0.0116)		(0.0554)
After 2010 x Black or Hispanic		0.0017		0.0192
		(0.0095)		(0.0488)
Constant	0.103***	0.1018***	0.2132***	0.1615***
	(0.00616)	(0.0068)	(0.0297)	(0.0379)
Data Analyzed	FL Convictions	FL Convictions	NC Convictions	NC Convictions
Drugs Included	Cocaine, all, Weight Only	Cocaine, all, Weight Only	Cocaine, all	Cocaine, all
P-value: W = BH	-	0.6484	-	0.2382
Observations	12,194	12,194	843	843

Panel B. Analysis of Bunching in Drug Seizures and Final Sentencing				
	Pr(280-290g)	Pr(200-400g)	Pr(200-400g)	Pr(280-290g)
	(6)	(7)	(8)	(9)
After 2010	-0.000186**		0.0332**	
	(8.67e-05)		(0.0162)	
After 2010 x White		0.0002		0.0038
		(0.0002)		(0.0513)
After 2010 x Black or Hispanic		-0.0003***		0.0346**
		(0.0001)		(0.0164)
Constant	0.000422***	0.0003***	0.143***	0.1558***
	(4.94e-05)	(0.0001)	(0.0120)	(0.0219)
Data Analyzed	NIBRS, Full Coverage States	NIBRS, Full Coverage States	USSC Sentencing, NC only	USSC Sentencing, NC only
Drugs Included	Crack-cocaine	Crack-cocaine	Cocaine, all	Crack-cocaine
P-value: W = BH	-	0.0830	-	0.5469
Observations	219,515	219,515	4,376	4,376

Notes: Robust standard errors in parentheses. When possible, the specifications above use a sample of offenses with drug amounts between 0 grams and 1000 grams. Analyses of state-level drug convictions do not make this restriction since the state reports broad drug weight categories instead of specific amounts. When broad categories (200-400g) are analyzed, a linear trend in year is included. The row "P-value: W= BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." In Panel A: columns 1-2 show an analysis of reported drug amounts for state-level drug convictions in Florida that restricts to cases where some weight range is listed in the offense description, columns 3-4 show an analysis of state-level drug convictions in North Carolina (a state where only some offenses specify the type of drug involved). Columns 5-6 show an analysis of weights for seized drugs reported to the FBI through the National Incident Based Reporting System (limiting to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015), Finally, columns 7-8 show an analysis of weights from USSC sentencing data for federal convictions in NC using broad drug categories and all types of cocaine.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.13: Variation in Bunching at 280-290g By Type of Agency Sending the Case

	280-290g	280-290g	280-290g	Weight (g)
	(1)	(2)	(3)	(4)
After 2010	0.0826***	0.0760***	0.0989***	26.09***
	(0.0180)	(0.0191)	(0.0129)	(5.659)
After 2010 × Any Federal	-0.00889			
	(0.0190)			
After 2010 × Only Federal		-0.00263		
		(0.0202)		
After 2010 × FBI			0.0160	52.99***
			(0.0198)	(11.29)
After 2010 × ATF			-0.0732***	-15.03**
			(0.0143)	(6.953)
After 2010 × State/local			-0.0229	-7.648
			(0.0231)	(11.45)
After 2010 × DEA & State/local			-0.0133	-3.980
			(0.0383)	(19.46)
After 2010 × Joint state/local			0.0148	7.345
			(0.0507)	(25.60)
After 2010 × ATF & State/local			-0.00860	-9.386
			(0.0388)	(13.18)
After 2010 × FBI & State/local			-0.0619	-17.32
			(0.0386)	(22.44)
Constant	0.00342***	0.00360***	0.00481***	77.73***
	(0.00121)	(0.00136)	(0.000876)	(1.523)
Observations	17,042	15,016	17,042	17,042

Notes: Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 interacts the after 2010 dummy variable with a dummy variable equal to one when the agency recorded as sending the case involves a federal agency (i.e. DEA, ATF, FBI). This includes agencies recorded as a federal agency joint with a state/local task force. Column 2 interacts the after 2010 variable with a variable equal to one when the agency sending the case is strictly federal (i.e. not including any involvement from state/local authorities). Column 2 does not include "joint" investigations in the sample. Column 3 provides more detail by interacting the after 2010 dummy variable with dummy variables for the top agencies (with the DEA as the reference category). Most agencies have similar levels of bunching at 280-290g post-2010. Two agencies have considerably lower levels, but as column 4 shows, those agencies are involved with lower drug weight cases, in general.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.14: Offender Drug-Holding Behavior by Race, After Fair Sentencing Act in 2010, Full Coverage States

	Weight	Pr(280-290g)	Weight	Pr(0-5g)	Pr(5-28g)	Pr(28-50g)	Pr(50-280g)	Pr(270-280g)	Pr(280-290g)	Pr(>290g)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After 2010 x White			-0.6018	0.0302***	-0.0210***	-0.0033**	-0.0058***	-0.0001	0.0001	-0.0002
			(0.5999)	(0.0041)	(0.0037)	(0.0016)	(0.0013)	(0.0000)	(0.0002)	(0.0007)
After 2010 x Black			-2.8015***	0.0403***	-0.0172***	-0.0064***	-0.0143***	-0.0001***	-0.0002**	-0.0020***
			(0.2504)	(0.0027)	(0.0025)	(0.0011)	(0.0009)	(0.0000)	(0.0001)	(0.0003)
Black	2.503***	9.21e-05	3.0414***	-0.1125***	0.0825***	0.0137***	0.0148***	0.0001	0.0002	0.0013***
	(0.260)	(0.000102)	(0.2885)	(0.0025)	(0.0023)	(0.0010)	(0.0009)	(0.0001)	(0.0001)	(0.0004)
Constant	10.01***	0.000454***	9.7586***	0.7503***	0.1856***	0.0310***	0.0284***	0.0002**	0.0004***	0.0043***
	(0.426)	(0.000152)	(0.4417)	(0.0040)	(0.0036)	(0.0016)	(0.0014)	(0.0001)	(0.0002)	(0.0006)
Observations	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043	207,043
P-value: W = B	-	-	0.0007	0.0408	0.3969	0.1075	0.0000	0.3308	0.1266	0.0205

Notes: Robust standard errors in parentheses. This analysis uses the weights of seized drugs reported to the FBI through the National Incident Based Reporting System. Ethnicity is not consistently recorded in NIBRS over this time period. As such, I refer to offenders as black or white, omitting the Hispanic label used in previous analyses. Columns 1-3 show the relationship between race of offender and drug weight seized, in general. Column 4 shows how the weight of an offender's seized drugs changes by race after 2010. Columns 5-11 show how the probability an offender's seized drugs are in a certain bin changes by race after 2010. All specifications include state fixed effects and controls for age and sex. The row "P-value: W= B" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black." The sample is limited to states that have full coverage from 2012-2015 and have at least 90% coverage from 2008-2015.

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

Table A3.15: Relationship between Bunching in EOUSA and Imputed Defendant Race

	280-290g, Missing = 0	280-290g	280-290g, Missing = 0	280-290g	280-290g, Missing = 0
	(1)	(2)	(3)	(4)	(5)
After 2010	0.0241***	-0.0318	-0.0153**	-0.00536	-0.00511
	(0.00180)	(0.0196)	(0.00654)	(0.0229)	(0.00826)
After 2010 × % Black or Hispanic (for Cases Sentenced in District-Month)		0.123***	0.0457***	0.0793***	0.0303***
		(0.0295)	(0.01000)	(0.0282)	(0.00984)
Constant	0.00159***	-0.00193	-0.00111	-0.00202	-0.000842
	(0.000195)	(0.00319)	(0.00130)	(0.00633)	(0.00259)
Prosecutor FEs	NO	NO	NO	YES	YES
Observations	49,342	13,384	32,751	13,384	32,751

Notes: Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. Column 1 displays the main bunching result using a dependent variable that is equal to one when the drug weight in the case is between 280-290g and is equal to zero if it is not in that range. Importantly, "280-290g, Missing=0" is also coded as zero if the drug weight field is missing. This is especially relevant for cross-district analyses because weight missingness varies substantially across districts. Coefficients are estimated from the following regression for column 1:

$$(Charged\,280-290g, Missing=0)_{it} = \alpha_0 + \beta_1 After2010_{it} + \varepsilon_{it}$$

Columns 2-5 interact the after 2010 dummy variable with a probabilistic estimate of defendant race (race is not available in the EOUSA files). To impute defendant race, I match EOUSA information about sentence year-month to USSC information about the racial composition of sentences in each sentence year-month. I code "% Black or Hispanic" equal to the fraction of offenders sentenced in a year-month who are black or Hispanic. In columns 4-5, I include prosecutor fixed effects. Specifications with the race and after 2010 interactions also include a variable equal to % black and Hispanic offenders in the district-month. The number of observations falls because not all cases that enter EOUSA end in a sentence. Coefficients are estimated from the following regression for columns 2 and 3 (with only the dependent variable changing):

$$\begin{aligned}(Charged\,280-290g)_{it} = \alpha_0 + \beta_1 (After2010)_{it} + \\ \beta_2 (After2010 \times \%BlackOrHispanic)_{it} + \%BlackOrHispanic_{it} + \varepsilon_{it}\end{aligned}$$

*** p<0.01, ** p<0.05, * p<0.1

Table A3.16: Relationship between Bunching in EOUSA and State-level Racial Animus

	280-290g (1)	280-290g, Missing = 0 (2)	# of Attys in State who Bunch at 280g (3)
After 2010	0.0756***	0.0163***	-
	(0.0123)	(0.00287)	-
Above Med. Racial Animus	-0.00187	-0.000390	1.737**
	(0.00122)	(0.000447)	(0.690)
After '10 × Above Med. Racial Animus	0.00150	0.0106***	-
	(0.0138)	(0.00365)	-
Constant	0.00520***	0.00182***	-
	(0.00111)	(0.000388)	-
Observations	19,241	49,051	51

Notes: Robust standard errors in parentheses. The estimates in this table are based on the EOUSA data. See Table A3.15 for a discussion of the "280-290, Missing=0" dependent variable. Columns 1 and 2 interact the after 2010 dummy variable with a dummy variable equal to one when the state where the case is received is above the median level of racial animus and equal to zero if it is below the median level. Coefficients are estimated from the following regression for columns 1 and 2 (with only the dependent variable changing):

$$(Charged\,280-290g)_{it} = \alpha_0 + \beta_1(After2010)_{it} + \beta_2(After2010 \times AboveMedRA)_{it} + AboveMedRA_{it} + \varepsilon_{it}$$

Since racial animus is a measure that varies across districts, column 2 results are particularly noteworthy (using the "missing included" version of 280-290g accounts for some of the cross-district variation in drug weight reporting). Finally, column 3 estimates a state-level regression of the number of bunching attorneys in the state (defined as an attorney whose fraction of cases at 280-290g post-2010 is above the average fraction at 280-290g pre-2010) on the above median racial animus dummy variable.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.17: Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors

	Atty. with 5+ Cases			Atty. with 15+ Cases		
Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)	(4)	(5)	(6)
Atty. Bunches at 280-290g Post-2010	-0.2193***	0.2421***	-0.0228	-0.1143	0.1882***	-0.0739
(15+ cases post-2010)	(0.0459)	(0.0339)	(0.0272)	(0.0806)	(0.0447)	(0.0640)
Constant	0.9309***	-	0.0691***	0.8855***	-	0.1145*
	(0.0242)	-	(0.0242)	(0.0617)	-	(0.0617)
Observations	1,647	1,647	1,647	699	699	699
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010						
	Below 280g	280-290g	Above 290g	Below 280g	280-290g	Above 290g
	(7)	(8)	(9)	(10)	(11)	(12)
Atty. Bunches at 50-60g Pre-2010	-0.0665***	0.0467***	0.0198	-0.0863***	0.0611***	0.0252
(15+ cases pre-2010)	(0.0245)	(0.0169)	(0.0151)	(0.0263)	(0.0167)	(0.0178)
Constant	0.9258***	0.0335***	0.0407***	0.9466***	0.0153	0.0382***
	(0.0168)	(0.0111)	(0.0115)	(0.0172)	(0.0096)	(0.0139)
Observations	1,278	1,278	1,278	956	956	956

Notes: Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Y\,g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \varepsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases post-2010 in columns 1-3 and with 15+ cases post-2010 in columns 4-6. Note, column (2) is a mechanical relationship, hence the missing standard error. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Y\,g)_i = \alpha_0 + \beta_1 AttyBunchesAt50g_i + \varepsilon_i$$

where $AttyBunchesAt50g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases and to prosecutors with 5+ cases pre-2010 in columns 7-9 and with 15+ cases pre-2010 in columns 10-12.

*** p<0.01, ** p<0.05, * p<0.1

Table A3.18: Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, Leave-One-Out Classification

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.114*	0.149***	-0.0354
(Leaving out current case in calculation)	(0.0659)	(0.0435)	(0.0463)
Constant	0.891***	0.0272***	0.0816*
	(0.0432)	(0.00765)	(0.0436)
Observations	971	971	971
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Pct. of Cases Bunched at 280-290g	-0.505***	0.527***	-0.0227
(Leaving out current case in calculation)	(0.116)	(0.0717)	(0.0976)
Constant	0.891***	0.0380***	0.0708**
	(0.0346)	(0.00791)	(0.0349)
Observations	971	971	971

Notes: Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \varepsilon_i$$

where *AttyBunchesAt*280*g* is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). **The classification for each bunching attorney is based on all cases excluding the current observation (i.e. a leave-one-out procedure).** Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\,X - Y g)_i = \alpha_0 + \beta_1 PctBunching280g_i + \varepsilon_i$$

where *PctBunchingAt*280*g* is equal to the prosecutor's fraction of cases at 280-290g post-2010 (excluding the current observation) minus the average fraction of cases at 280-290g pre-2010. These regressions are restricted to post-2010 cases and to prosecutors with 10+ cases post-2010.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.19: Missing Mass in the Distribution of Drug Amounts, Comparing "Bunching" and "Non-Bunching" Prosecutors, with Bootstrapped SEs

Panel A. Bunching at 280g Post-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(1)	(2)	(3)
Atty. Bunches at 280-290g Post-2010	-0.1794***	0.2170***	-0.0376
	(0.0659)	(0.0371)	(0.0510)
Constant	0.9184***	-	0.0816*
	(0.0435)	-	(0.0435)
Observations	989	989	989
Panel B. Bunching at 50g Pre-2010 and Distribution of Cases Post-2010			
	Below 280g	280-290g	Above 290g
	(4)	(5)	(6)
Atty. Bunches at 50-60g Pre-2010	-0.0785***	0.0575***	0.0211
	(0.0299)	(0.0177)	(0.0180)
Constant	0.9359***	0.0233**	0.0408***
	(0.0170)	(0.0105)	(0.0133)
Observations	1,135	1,135	1,135

Notes: Standard errors are calculated from 25 replications of a bootstrapping procedure that samples cases (with replacement) clustered at the prosecutor-level and calculated the bunching dummy variables within each sample. The standard errors for the constant terms are not calculated in this way; robust errors clustered at the prosecutor-level are used. The estimates in this table are based on the EOUSA data. Coefficients in panel A are estimated from the following regression for each range:

$$(Charged\, X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt280g_i + \varepsilon_i$$

where $AttyBunchesAt280g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 280g definition (i.e. the fraction of their cases that are from 280-290g is above the average fraction of 280-290g cases pre-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 280-290g is at or below the average fraction of 280-290g cases pre-2010). These regressions are restricted to post-2010 cases (for columns 1-3) and to prosecutors with 10+ cases post-2010. Note, column (2) is a mechanical relationship, hence the missing standard error. Coefficients in panel B are estimated from the following regression for each range:

$$(Charged\, X - Y g)_i = \alpha_0 + \beta_1 AttyBunchesAt50g_i + \varepsilon_i$$

where $AttyBunchesAt50g$ is equal to one if the prosecutor is classified as a "bunching" prosecutor under the 50g definition (i.e. the fraction of their cases that are from 50-60g is above the average fraction of 50-60g cases post-2010) and is equal to zero if the prosecutor is not classified as a bunching prosecutor (i.e. the fraction of their cases that are from 50-60g is at or below the average fraction of 50-60g cases post-2010). These regressions are restricted to post-2010 cases (for columns 5-8) and to prosecutors with 10+ cases pre-2010.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.20: Persistence of Attorney-level Bunching Across Districts, from Analysis of Movers

	Pr(Atty. Bunches at 10-Year Mandatory Minimum in 2nd District)			
	(1)	(2)	(3)	(4)
Atty. Bunches at 10-Year MM in 1st District	0.184*	0.162**	0.263**	0.154*
	(0.0936)	(0.0816)	(0.108)	(0.0829)
Constant	0.500***	0.432***	0.462***	0.440***
	(0.0700)	(0.0580)	(0.0809)	(0.0577)
Bunching classification	280-290g, National	280-290g, Missing=0, National	280-290g, District	280-290g, Missing=0, District
Observations	109	148	79	144

Notes: Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. For this analysis, I identify the attorneys who switch districts at some point in their career (using their initials recorded in the EOUSA case management system). I then identify the set of those attorneys who bunch at a 10-year mandatory minimum in their first district. I also limit the sample to attorneys who have at least 5+ cases in their first district and 5+ cases in their second district (this maintains the 10+ restriction but spreads it evenly across districts). Since I am analyzing movers, it is almost always the case that the cases in their first district are pre-2010 cases, meaning that the bunching classification is determined based on bunching at 50-60g. Finally, I regress an indicator equal to one if the attorney bunches at the 10-year threshold in their second district on whether they bunched at the 10-year threshold in their first district. I do this for four methods of classifying bunching attorneys. Columns 1 and 2 are detailed in Table A3.15. Columns 3 and 4 mirror those two approaches but define the "baseline" bunching at the district-level. For example, an attorney **i** in district **A** is defined as bunching at 50-60g in column 3 if their fraction of cases at 50-60g pre-2010 is above the fraction of cases at 50-60g in district **A** post-2010. In all cases, I find that an attorney who bunches above the mandatory minimum threshold in their first district is more likely to do so in their second district than an attorney who does not bunch above the mandatory minimum threshold in their first district.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.21: Relationship between Various Bunching Ranges, Attorneys

	28-29g (1)	28-29g (2)	50-60g (3)	280-290g (4)	280-290g (5)	280-290g (6)
Atty. Bunches at 280-290g Post-2010	0.144**	0.140**	0.182***			
	(0.0625)	(0.0590)	(0.0664)			
Atty. Bunches at 28-29g Post-2010				0.155***	0.0876**	
				(0.0544)	(0.0340)	
Atty. Bunches at 50-60g Pre-2010						0.0575***
						(0.0172)
Constant	0.131***	0.120***	0.155***	0.0826***	0.0479***	0.0233**
	(0.0241)	(0.0232)	(0.0288)	(0.0271)	(0.0149)	(0.0105)
Sample Years	2011-2017	2011-2017	2000-2010	2011-2017	2011-2017	2011-2017
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	843	910	1,976	483	840	1,135

Notes: Standard errors clustered at the prosecutor level in parentheses. The estimates in this table are based on the EOUSA data. Columns 1-3 estimate the likelihood an attorney who bunches at 280-290g (i.e. who has a fraction of cases at 280-290g post-2010 that is above the average fraction of 280-290g cases pre-2010) also bunches at 28-29g post-2010, 28-29g post-2010, and 50-60g pre-2010, respectively. Column 1 limits the sample to cases with below 280g to avoid a mechanical relationship. Column 2 does this by excluding only the 280-290g range from the sample. Both approaches yield similar results. Column 3, since the dependent variable is based on pre-2010 data, uses the full range of cases (0-1000g). Columns 4-6 estimate the likelihood an attorney who bunches at 28-29g post-2010 or 50-60g pre-2010 also bunches at 280-290g post-2010. As before, columns 4 and 5 exclude the 28-29g range to avoid a mechanical relationship. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of attorneys who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). In all cases, an attorney who bunches at one mandatory minimum threshold is more likely to bunch at a separate mandatory minimum threshold.

*** p<0.01, ** p<0.05, * p<0.1

Table A3.22: Bunching at 280-290g and Drug Conspiracy Charges

	Pr(Lead Charge = Conspiracy)		
	(1)	(2)	(3)
Case recorded at 280-290g	0.396***	0.307***	0.249***
	(0.0326)	(0.0329)	(0.0361)
Constant	0.166***	0.255***	0.314***
	(0.00279)	(0.00487)	(0.0156)
Sample restriction	0-1000g	50-1000g	280-1000g
Observations	18,062	8,236	1,116

Notes: Robust standard errors are in parentheses. The estimates in this table are based on the EOUSA data. The dependent variable is an indicator equal to one if the lead charge on the case is a drug conspiracy charge. Drug conspiracy charges are a tool that prosecutors can use to increase the weight involved in the offense because the total weight of the conspiracy is applied to each offender deemed involved in the conspiracy. The independent variable is whether the case involves 280-290g. Cases with 280-290g are substantially more likely to carry a lead conspiracy charge. This is true even when limiting to cases with 280-1000g only (see column 3).

*** p<0.01, ** p<0.05, * p<0.1

Table A3.23: Effect of Alleyne v. US, Accounting for Missing Values

	Pr(Case's Drug Weight is Missing)	Pr(Case is Charged with 280-290g, Missing = 0)
	(1)	(2)
After June 17th, 2011-2016	-0.0211	0.00438
	(0.0309)	(0.00869)
After June 17th, 2013	-0.0219	-0.0389*
	(0.0702)	(0.0223)
Constant	0.834***	0.0243
	(0.0690)	(0.0269)
Bandwidth	±150 days	±150 days
Observations	6,182	6,182

Notes: Standard errors clustered at the date the case is received in parentheses. The estimates in this table are based on the EOUSA data. The coefficients above are estimated from the following regression discontinuity style model:

$$
\begin{aligned}
Y_{it} = \alpha_0 &+ \beta_1 AfterJune17_{it} + \beta_2 DaysFrom_{it} + \beta_3 (AfterJune17 \times DaysFrom)_{it} \\
&+ \delta_1 (AfterJune17 \times Year2013)_{it} + \delta_2 (DaysFrom \times Year2013)_{it} \\
&+ \delta_3 (AfterJune17 \times DaysFrom \times Year2013)_{it} + D_{it} + \varepsilon_{it}
\end{aligned}
$$

where $AfterJune17$ is a dummy variable equal to one for cases received after June 17th in each year, *DaysFrom*, the running variable, is the date the case was received centered at zero on June 17th, and *Year*2013 is equal to one for cases received in 2013 (the year *Alleyne* is decided). In addition, all specifications above include day-of-week fixed effects, D_{it}, for the day the case is received. In column 1, Y_{it} is equal to one if the observation has a missing drug weight and equal to zero otherwise. There is little effect of Alleyne on the likelihood an observation has missing drug weight. In column 2, Y_{it} is equal to one if the drug weight is equal to 280-290g or if the drug weight is missing and equal to zero otherwise. There is still a decrease in bunching after Alleyne when accounting for missing values.
*** p<0.01, ** p<0.05, * p<0.1

Table A3.24: Degree of Bunching Post-2010 by Race and District-level Caseload Characteristics

	Pr(280-290g)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After '10 x White (W)	0.0172**	0.0183*	0.0161**	0.0197*	0.0131	0.0219***	0.0113*	0.0137	0.0128
	(0.0082)	(0.0100)	(0.0080)	(0.0102)	(0.0088)	(0.0083)	(0.0061)	(0.0089)	(0.0083)
After '10 x Black or Hispanic (BH)	0.0424***	0.0477***	0.0344***	0.0536***	0.0302***	0.0388***	0.0407***	0.0368***	0.0379***
	(0.0094)	(0.0035)	(0.0028)	(0.0035)	(0.0027)	(0.0029)	(0.0035)	(0.0030)	(0.0033)
After '10 x W x Char.	-0.0147	-0.0088	-0.0055	-0.0122	0.0008	-0.0191*	0.0007	-0.0043	-0.0028
	(0.0098)	(0.0116)	(0.0107)	(0.0115)	(0.0108)	(0.0103)	(0.0104)	(0.0107)	(0.0104)
After '10 x BH x Char.	-0.0187	-0.0222***	0.0027	-0.0363***	0.0124***	-0.0072*	-0.0077*	-0.0011	-0.0031
	(0.0114)	(0.0043)	(0.0043)	(0.0042)	(0.0044)	(0.0042)	(0.0045)	(0.0044)	(0.0044)
Constant	0.0024*	0.0054***	0.0054***	0.0053***	0.0053***	0.0053***	0.0045***	0.0046***	0.0046***
	(0.0014)	(0.0010)	(0.0011)	(0.0010)	(0.0011)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Characteristic	District-by-Year Above Med. # of Cases per Attorney	District Above Med. % of Guilty Cases	District Above Med. % of Declined Cases	District Above Med. % of Plea Cases	District Above Med. % of Cases Dismissed for 'Weak Evidence'	District Above Med. % of Cases Dismissed for 'Resources'	District Above Med. % Of Cases with Retained Counsel (based on '99-'02)	District Above Med. % Of Cases with Appointed Counsel	District Above Med. % Of Cases with Public Defender Counsel
P-value: W = BH	0.0246	0.0057	0.0297	0.0017	0.0609	0.0536	0.0000	0.0139	0.0049
P-value: W+Char. = BH+Char.	0.0000	0.0113	0.0007	0.0872	0.0001	0.0000	0.0191	0.0001	0.0003
Observations	52,731	52,745	52,745	52,745	52,745	52,745	49,851	49,851	49,851

Notes: Robust standard errors in parentheses. "Characteristic" or "Char." represents a dummy variable that is an district or district-by-year characteristic. The specific characteristic of interest is noted in the "Characteristic" row. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. The row "P-value: W = BH" reports the p-value from a test of the null hypothesis that the coefficient on "After 2010 x White" is equal to the coefficient on "After 2010 x Black or Hispanic." The row "P-value: W+Char. = BH+Char." reports the p-value from a test of the null hypothesis that the combined coefficients on "(After 2010 x White)+(After 2010 x White x Characteristic)" is equal to the combined coefficients on "(After 2010 x Black or Hispanic)+(After 2010 x Black or Hispanic x Characteristic)." Column 1 interacts the after 2010 by race dummy variables with a district-by-year dummy variable indicating if the district received above the median number of cases (per attorney) in the year. Column 2 studies districts above/below the median for percent of cases that end in a guilty verdict, column 3 studies districts above/below the median for percent of cases declined, and column 4 studies districts above/below the median for percent of cases that end in plea deals. Columns 5 and 6 study districts above/below the median for percent of cases declined due to "weak evidence" or "lack of resources" (as coded in the EOUSA case files, codes not present for all cases). Columns 7-9 use the USSC data from 1999-2002 on type of defense counsel to examine heterogeneity by type of defense counsel used in the district. Places with different rates of retained, appointed, or public defender defense counsel from 1999-2002 nevertheless have similar bunching at 280g post-2010.

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

Table A3.25: Relationship between Bunching at 280g and Judge Characteristics

	Pr(280-290g)	Pr(280-290g)	Pr(280-290g)
	(1)	(2)	(3)
After 2010	0.0928***	0.0891***	0.1042***
	(0.0093)	(0.0209)	(0.0151)
After 2010 $\times$White Judge		0.0045	
		(0.0233)	
After 2010 $\times$Republican Judge			-0.0197
			(0.0191)
Constant	0.0040***	0.0059**	0.0049***
	(0.0007)	(0.0024)	(0.0014)
Observations	8,359	8,359	8,359

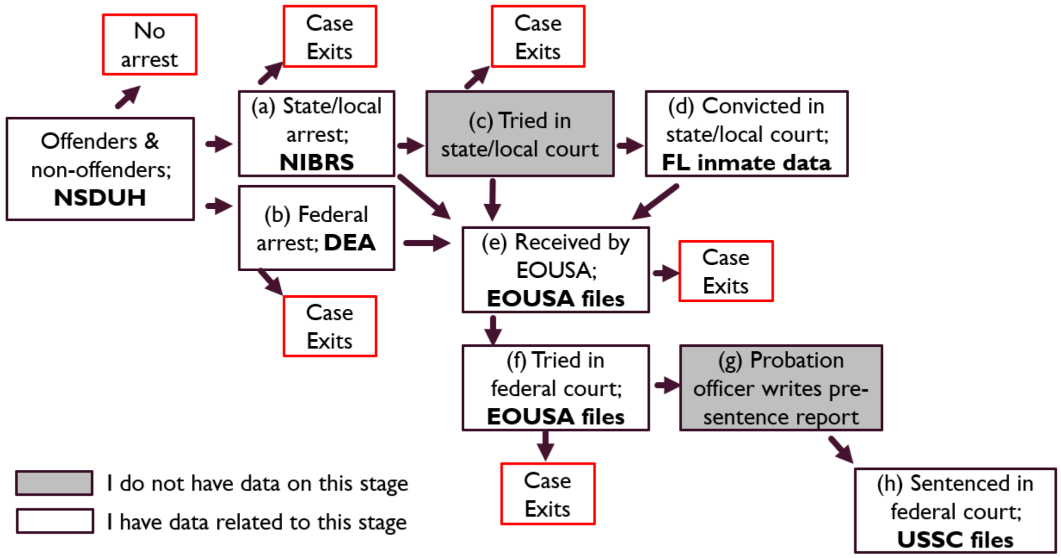
Notes: Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. I can match judge race and political party to approximately half of the cases in the EOUSA data. For data on judge characteristics, I use the file provided by Cohen and Yang (2019). I estimate whether bunching at 280g is related to judge race or judge political party. Column (1) shows that the level of bunching is similar for cases where I can match judge characteristics. Column (2) shows that judge race does not affect bunching at 280g. Column (3) shows that judge political party does not affect bunching at 280g.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3.26: Relationship between Various Bunching Ranges, Judges

	28-29g	28-29g	50-60g	280-290g	280-290g	280-290g
	(1)	(2)	(3)	(4)	(5)	(6)
Judge Bunches at 280-290g Post-2010	-0.0129	-0.00857	0.0557			
	(0.0305)	(0.0286)	(0.0412)			
Judge Bunches at 28-29g Post-2010				-0.00207	-0.0144	
				(0.0523)	(0.0329)	
Judge Bunches at 50-60g Pre-2010						0.0175
						(0.0215)
Constant	0.155***	0.143***	0.199***	0.168***	0.108***	0.0723***
	(0.0195)	(0.0185)	(0.0243)	(0.0390)	(0.0250)	(0.0180)
Sample Restriction	0-280g	0-280g, 290-1000g	0-1000g	29-1000g	0-28g, 29-1000g	0-1000g
Observations	769	827	2,710	469	789	1,270

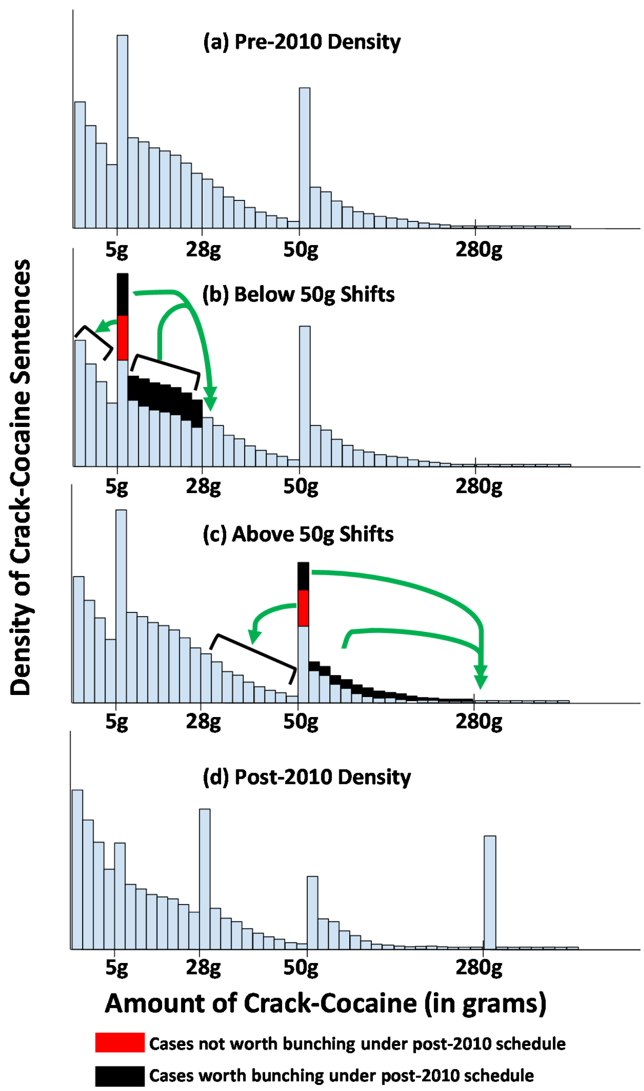
Notes: Standard errors clustered at the judge level in parentheses. The estimates in this table are based on the EOUSA data. See Table A3.21 for a discussion of the dependent and independent variables in column 1-6. The major difference is that these regressions examine judges classified as "bunching" at a given range. This is possible because the EOUSA files contain a judge ID for many cases. I use that judge ID to calculate the fraction of cases at 280-290g post-2010, 28-29g post-2010, and 50-60g pre-2010 for each judge. 28-29g is relevant post-2010 because 28g is the threshold for the 5-year mandatory minimum after 2010. 50-60g is relevant pre-2010 because 50g is the threshold for the 10-year mandatory minimum prior to 2010. All regressions in this table use the sample of judges who have 10+ cases (post-2010 for columns 1-5; pre-2010 for column 6). Judges who bunch at one mandatory minimum threshold are not more likely to bunch at other mandatory minimum thresholds.
*** p<0.01, ** p<0.05, * p<0.1

Figure A3.1: Graphical Illustration of Timeline from Arrest to Sentencing



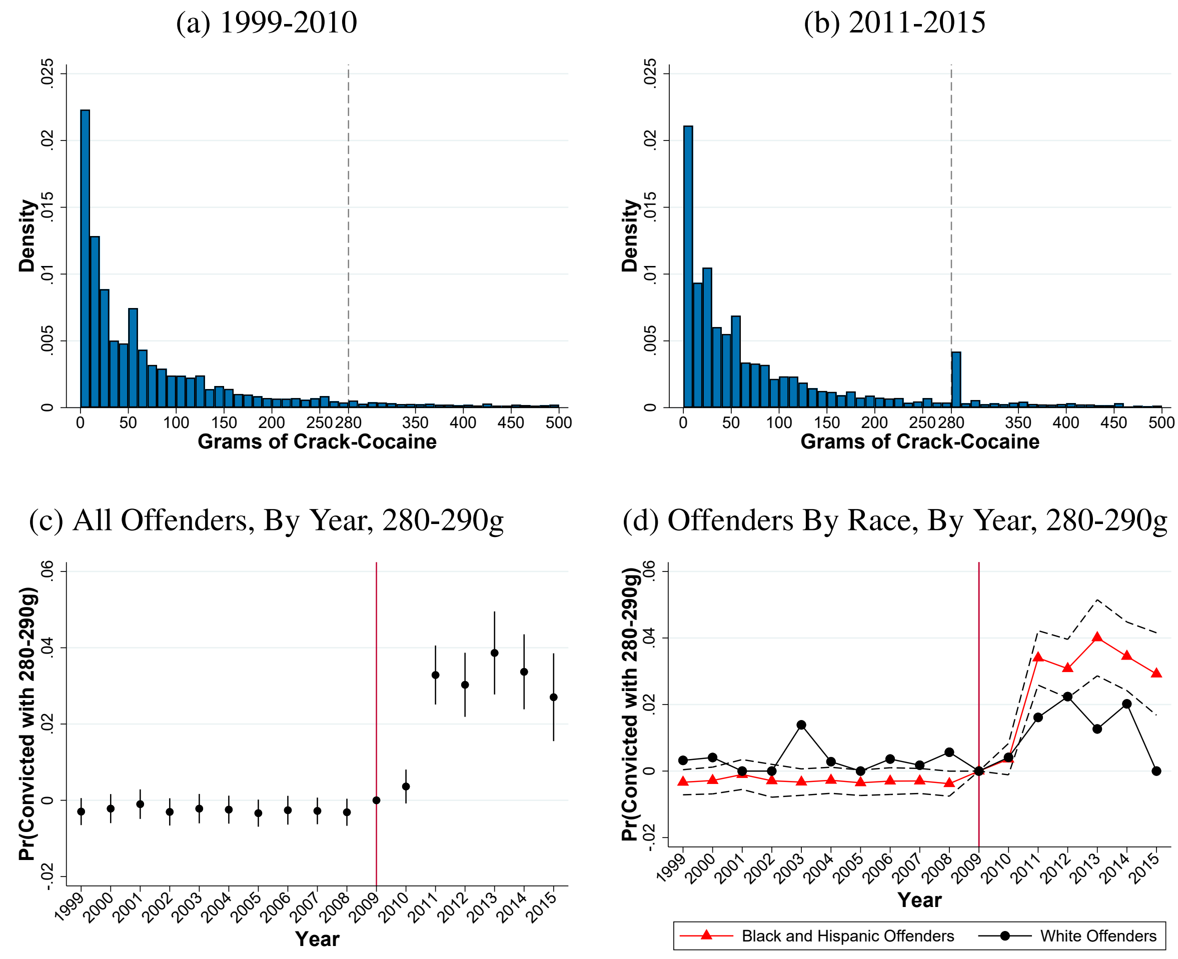
Notes: The figure above details the timeline from arrest to sentencing. Before arrest, the eventual arrestees come from the set of all people, some of whom are innocent and some of whom are guilty. Some individuals from this group are arrested by state/local police or federal police. Of those arrested by state/local police, their case can be dismissed, tried in state/local court, or passed on to federal authorities. Case tried in state/local court can leave the system if they are found not guilty, dismissed, etc., they can be convicted, or they can be sent to federal authorities. In fact, even convicted cases can be sent to federal authorities. Individuals arrested by federal police are typically referred to the EOUSA directly. Once a case is received by the EOUSA, it can leave the system via a dismissal, declination, etc., or it can be taken to federal court. For cases convicted in federal court, a probation officer prepares a pre-sentence report, and ultimately, the offender is sentenced. I have obtained data at nearly all of these steps. The two steps for which I lack data are in the middle of steps where bunching does not change, which suggests that nothing changes in the middle step.

Figure A3.2: Graphical Illustration of Conceptual Model, Prosecutor Responses to the FSA



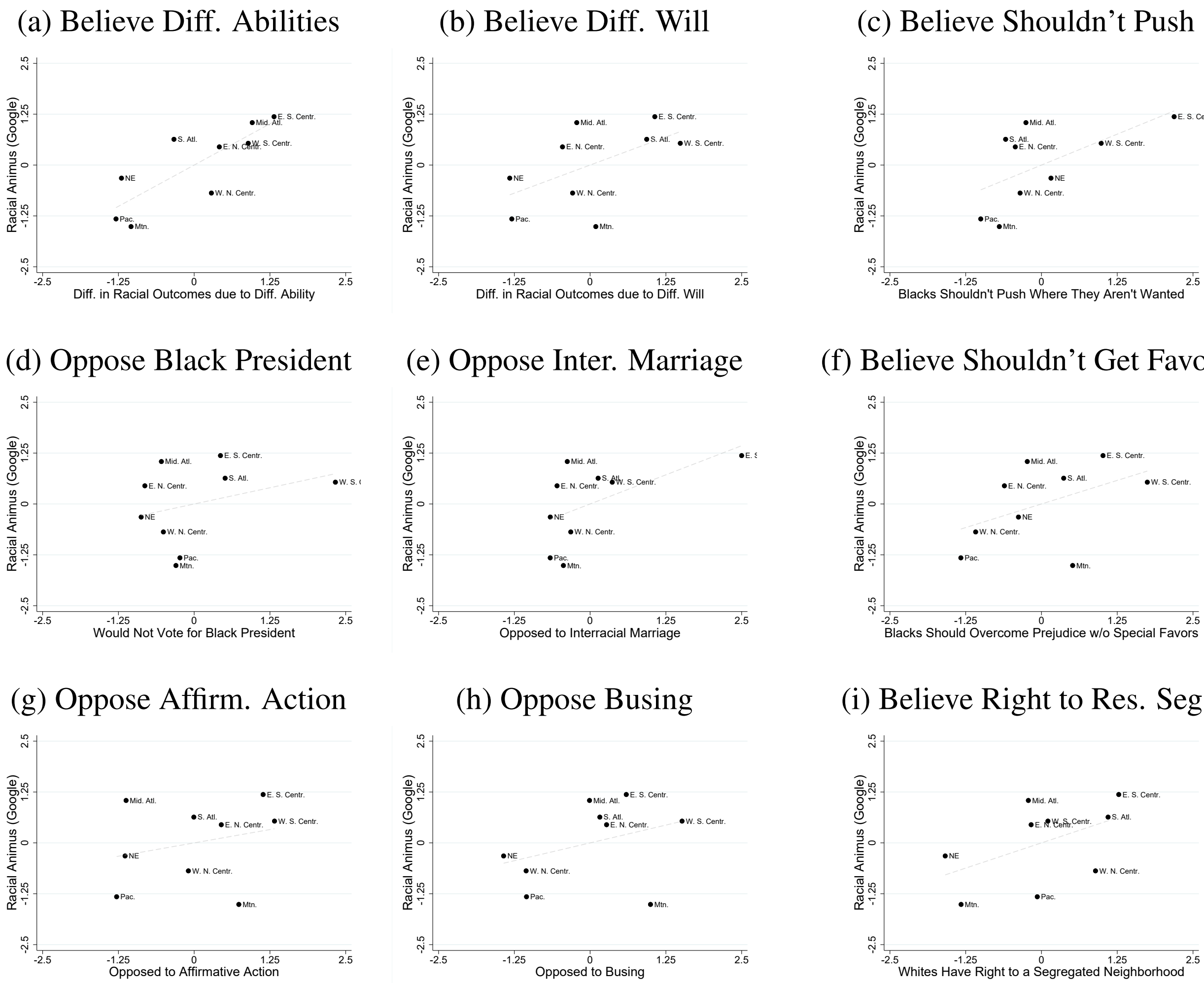
Notes: Panel (a) displays a hypothetical pre-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and prosecutor discretion. Panel (b) shows how the 0-5g, 5-28g, and 28-50g ranges will change after 2010. Some cases bunched at 5g will not be worth bunching at 28g (depicted in red), and they will shift into the 0-5g range. Some cases bunched at 5g and some cases from 5-28g will be worth bunching at 28g (depicted in black), and they will shift into the 28-50g range. Panel (c) illustrates a similar phenomena for the 50-280g range–some cases will shift down into the 28-50g range and some will shift up to the 280-290g range. Panel (d) shows the hypothetical post-2010 distribution of weights, with bunching at 5g and 50g due to round-number bias and bunching at 28g and 280g due to prosecutor discretion.

Figure A3.3: Changing Distribution of Drug Amounts Around 280g Pre- and Post-2010, USSC



Notes: Panels (a) and (b) plot the distribution of drug amounts recorded in federal crack-cocaine sentences starting at 0 grams and ending at 500 grams for 1999-2010 (when the mandatory minimum threshold was 50g) and 2011-2015 (when it was 280g). In panel (c), I estimate the main bunching coefficient by year (relative to 2010) and plot the coefficients with 90% confidence intervals. Panel (d) plots the coefficients and confidence interval for black and Hispanic offenders and the coefficients for white offenders (I do not include confidence intervals for white offenders because their estimates by year are extremely noisy).

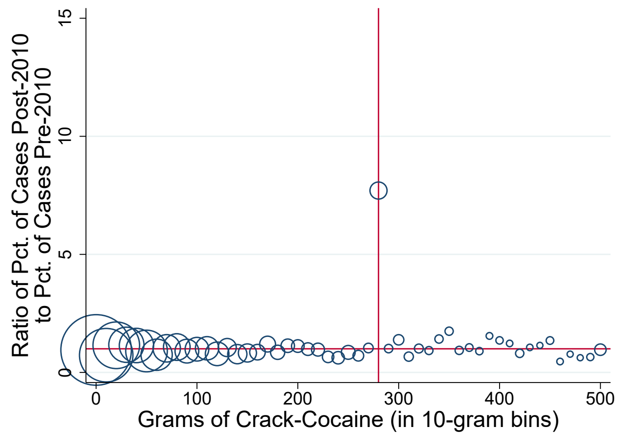
Figure A3.4: Relationship between Google Trends Racial Animus Measure and GSS Responses from Highly Educated Respondents on Attitudes about Race



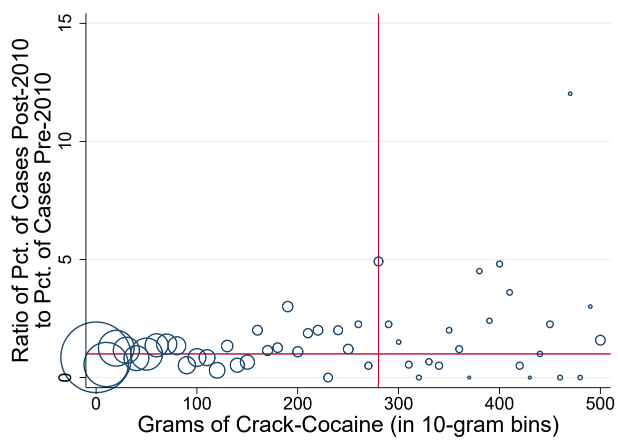
Notes: The figures above plot the relationship between the Google Trends racial animus measure (standardized and centered at zero) and various measures of attitudes about race from the General Social Survey (GSS) from 1972-2018 (not all questions are present in all years; also standardized and centered at zero). For the GSS measures, I limit the sample to respondents with a graduate degree or higher to test if the Google Trends racial animus measure is correlated with racial attitudes of highly educated people. The public sample of the GSS only includes region identifiers. I aggregate the Google Trends measure to the region level by taking the mean across all states in the region. The regions are: Northeast, West North Central, Pacific, Mountain, East North Central, Mid Atlantic, South Atlantic, West South Central, and East South Central. The GSS questions are: Do you believe... (a) racial differences in outcomes are due to different abilities by race (available 1977-2018), (b) racial differences in outcomes are due to different will by race (1977-2018), (c) black shouldn't push where they aren't wanted (1972-2002), (f) blacks should overcome prejudice without special favors (1994-2018), and (i) whites have a right to a segregated neighborhood (1972-1996)? And are you opposed to... (a) voting for a black president (1972-2010), (b) interracial marriage (1972-2002), (c) affirmative action (1994-2018), (d) desegregation busing (1972-1996)?

Figure A3.5: Bunching Ratio from 0-500g, USSC

(a) Bunching Ratio from 0-500g, Black and Hispanic Offenders



(b) Bunching Ratio from 0-500g, White Offenders

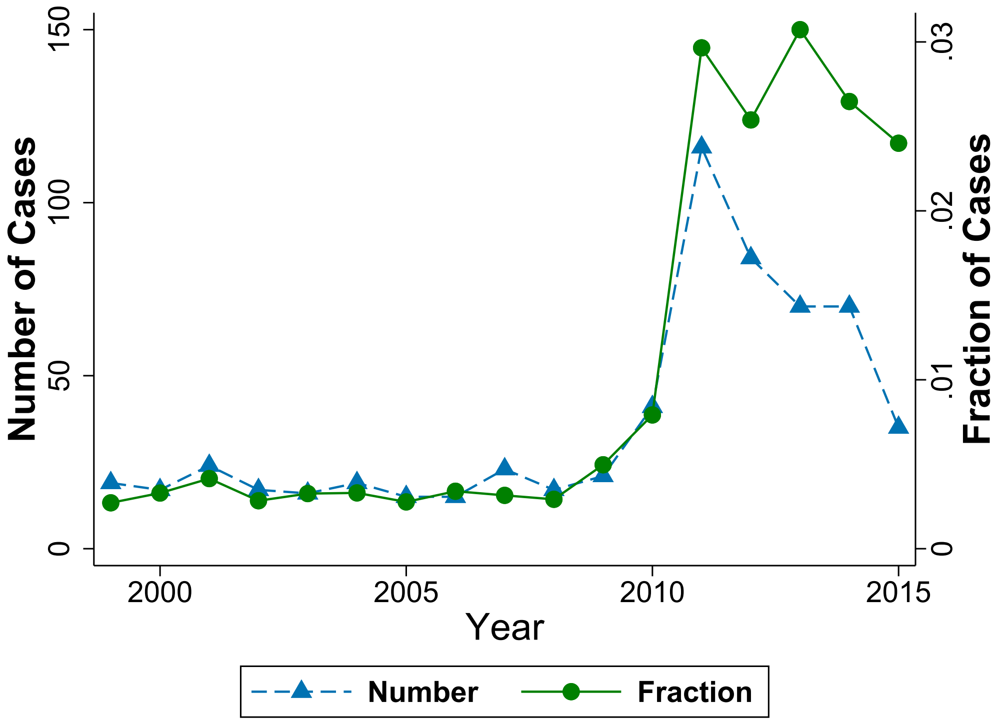


Notes: The figure above plots the bunching ratio for each 10-gram bin from 0-500 grams by race. The bunching ratio for each bin *b* is defined as follows:

$$\text{Bunching Ratio}_b = \frac{\text{\% of cases in b post-2010}}{\text{\% of cases in b pre-2010}}$$

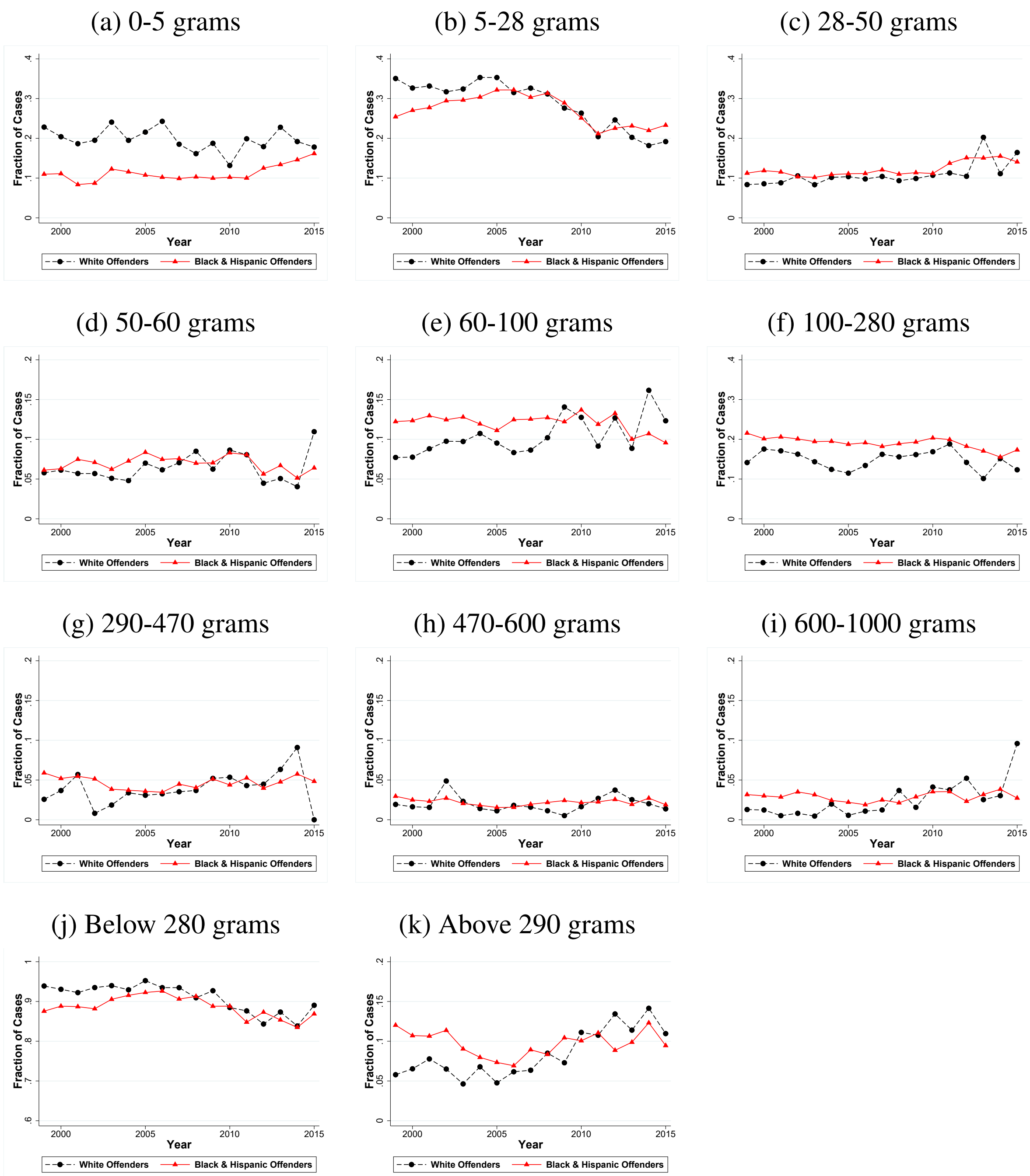
If the distributions are the same pre- and post-2010, the bunching ratio will equal 1 (marked by the horizontal red line). If the ratio is above 1, there is a higher degree of bunching in bin b post-2010. If the ratio is below 1, there is a lower degree of bunching post-2010. The size of the marker for each bin b is weighted by the total number of cases in the bin pre- and post-2010 (relative to rest of the group included in the plot, not relative to the full sample).

Figure A3.6: Number and Share of Offenses with 280-290g Over Time, USSC



Notes: The figure above plots the total number of offenses with 280-290g over time and the share (or fraction) of cases with 280-290g over time.

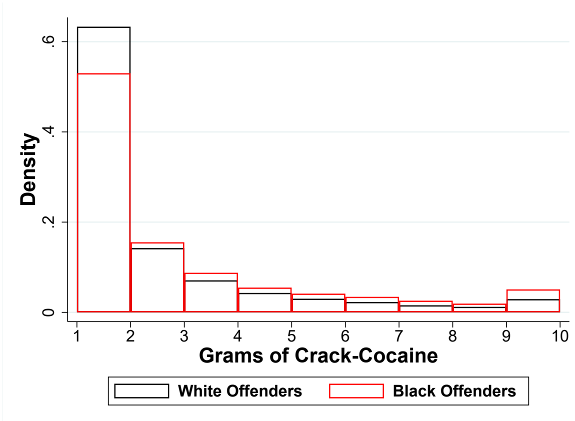
Figure A3.7: Changing Distribution of Drug Weights Over Time, By Race, USSC



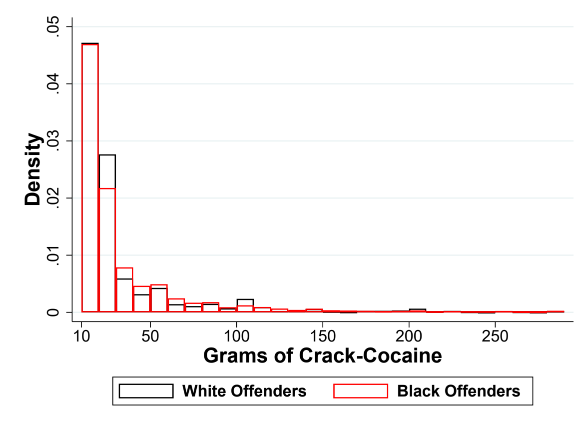
Notes: The figures above plot the share of cases in the specified range by year for white and black and Hispanic offenders. For example, panel (a) plots the share of cases with 0-5g (not including 5g) in each year from 1999-2015. Panel (b) plots the share of cases with 5-28g in each year from 1999-2015, and so on.

Figure A3.8: Alternative Figures for Conditional Racial Disparity Tests

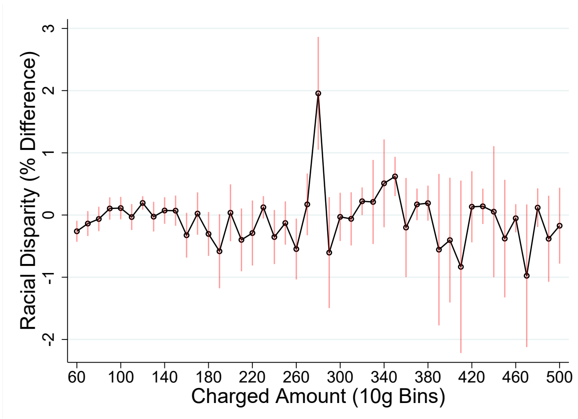
(a) Drug Seizures by Race, 0-10g, NIBRS



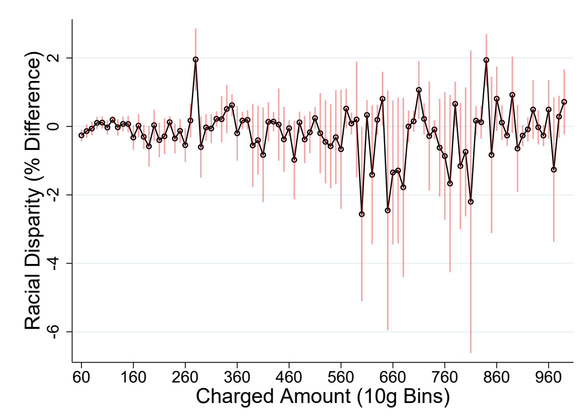
(a) Drug Seizures by Race, 10-280g, NIBRS



(c) Shifting from 60-500g by Race, USSC



(d) Shifting from 60-1000g by Race, USSC

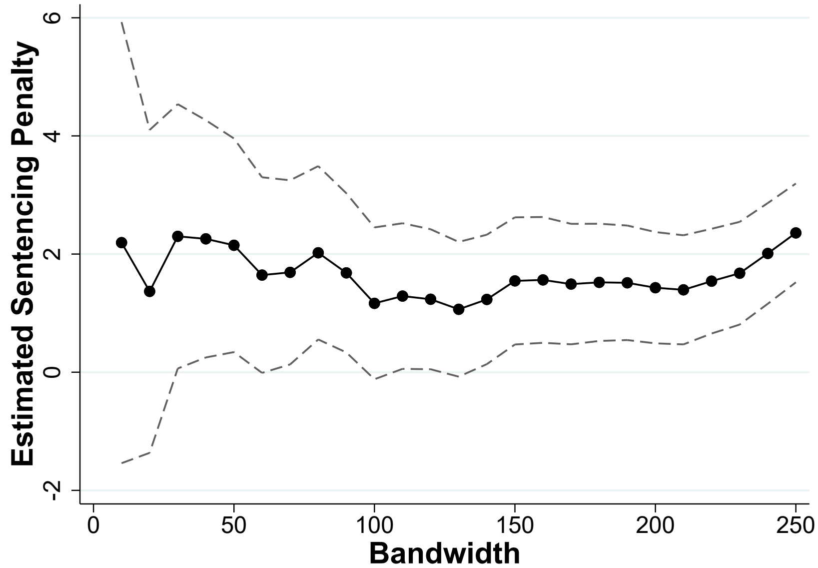


Notes: The figure in panel (a) plots the histograms of crack-cocaine amounts seized for white offenders and for black and Hispanic offenders from 0-10g. The white offenders are slightly over-represented at 1g, but otherwise, the distributions are very similar. The figure in panel (b) plots the histograms by race from 10-280g. White offenders are slightly over-represented at 20-30g, but otherwise, the distributions are very similar. These figures use the balanced sample of agencies (i.e. agencies that are present in all 16 years) in NIBRS. Panels (c) and (d) plot the coefficient δ^X for each 10g bin starting at X divided by the share of cases in that 10g bin (to calculate a percent difference).

$$\text{(12)} \quad (\text{Charged X-Yg})_{it} = \alpha + \delta^X(\text{After2010} \times \text{BlackOrHispanic})_{it} + \gamma \text{After2010}_{it} + \lambda \text{BlackOrHispanic}_i + \varepsilon_{it}$$

Since estimates are noisier at higher amounts, panel (c) shows the estimates for amounts from 0-500g alone and panel (d) shows the estimates for amounts from 0-1000g.

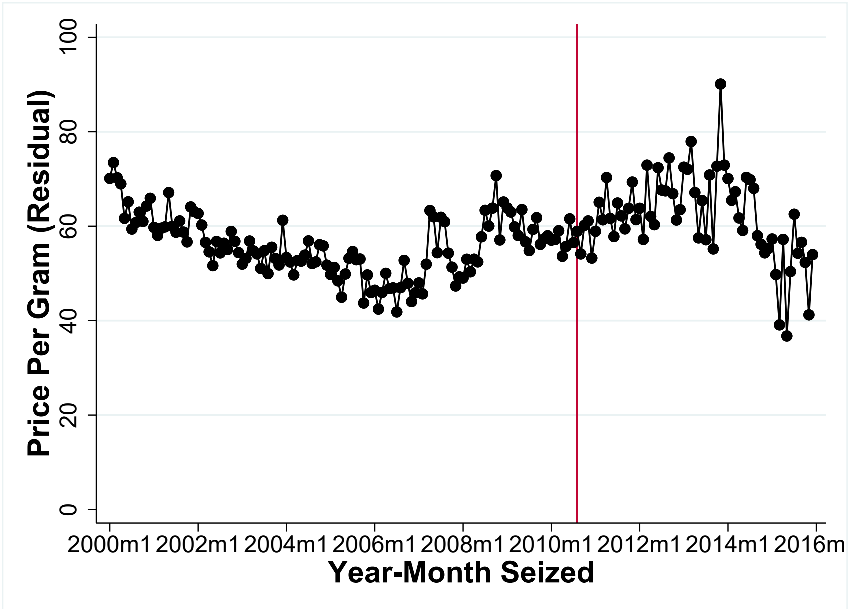
Figure A3.9: Sentencing Discontinuity Robust to Multiple Bandwidths



Notes: The figure above plots the sentencing penalty of crossing the 280g mandatory minimum threshold after 2010, as estimated using the RD difference-in-difference model specified in equation (5) of the main text. The dashed lines are 90% confidence intervals. Estimates using a quadratic in polynomial are similar in magnitude but slightly noisier. The bandwidths used in the figure above range from 10g to 250g, in 10g intervals.

Figure A3.10: Drug Prices Before and After the Fair Sentencing Act,

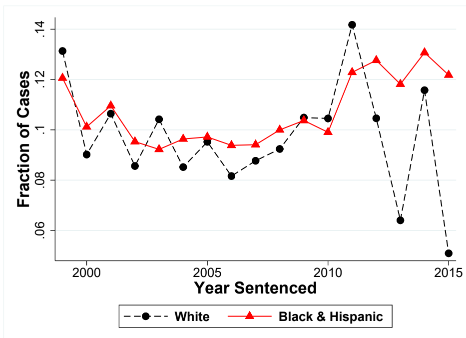
DEA



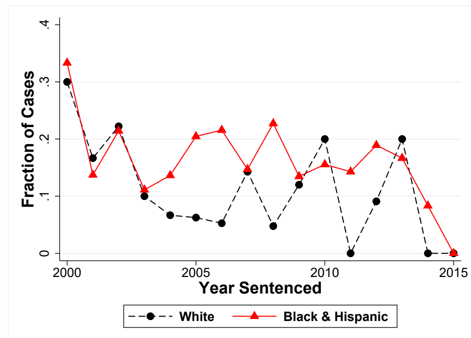
Notes: Panel (a) plots the drug price per gram (conditional on state, drug potency, type of drug, month seized, and a linear trend in year) against the year-month the drugs were seized. Outliers above the 95th percentile ($200 per gram) and below the 5th percentile ($20 per gram) are excluded. The price is smooth and increasing through the date the Fair Sentencing Act was implemented. In other words, there is no clear price response in the illegal drug market, at least in the short run. I formally estimate the discontinuity around the date the bill was signed using a bandwidth of +/- 24 months and various polynomials (linear, quadratic, cubic). The estimated discontinuity is never statistically different from zero, and it ranges from -5.5 to 2.1. Panel (b) plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015.

Figure A3.11: Alternative Figures Testing for Shifting from State/Local Authorities to Federal Court

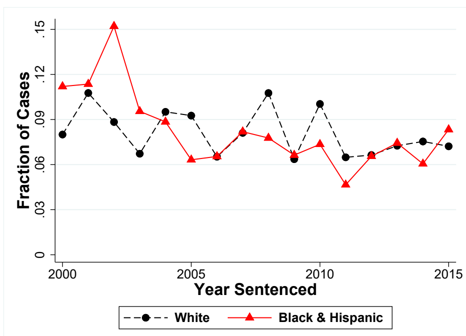
(a) Fraction of Cocaine Cases 200-400g, USSC



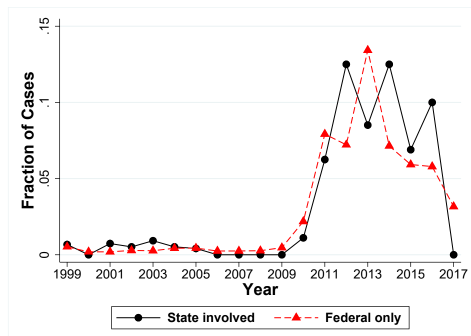
(b) Fraction of Cocaine Cases 200-400g, NC



(c) Fraction of Cocaine Cases 200-400g, FL, Alternative Sample

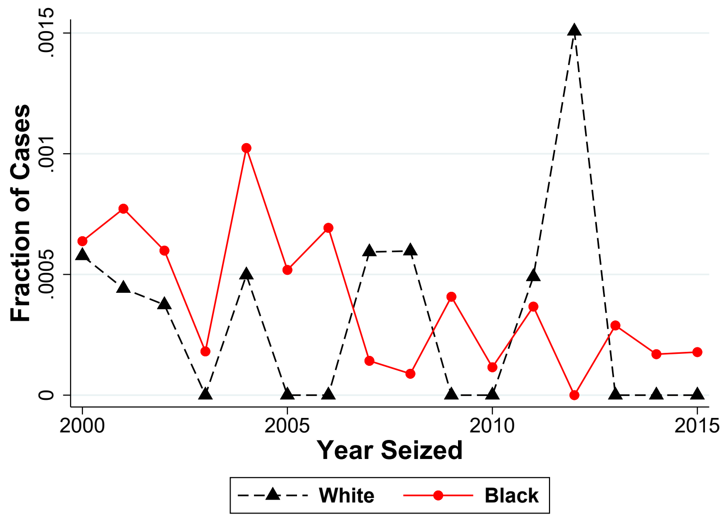


(d) Fraction of Crack-Cocaine Cases in 280-290g, by Type of Source Agency, EOUSA



Notes: The figure in panel (a) plots the fraction of cocaine offenses with 200-400g in the USSC federal sentencing data, by race. The figure in panel (b) plots the fraction of cocaine offenses that have a range from 200-400g in NC state prison from 2000-2015, by race. Many of drug convictions in NC do not include type of drug in the offense description, the figure above is limited to those offenses that specifically list 'cocaine' in the offense description. The figure in panel (c) plots the fraction of cocaine offenses with 200-400g in FL state prison by race, limiting to those offenses that list a weight range in the offense description (the figure in the main text includes all cocaine offenses and codes (Convicted 200-400g)=0 if there is not weight listed in the offense description). The figure in panel (d) plots the share of cases sent to EOUSA attorneys from sources that involve state agencies (red dashed line with triangle markers) and the share of cases sent to EOUSA attorneys from strictly Federal sources (black solid line with circle markers). This figure is limited to the top agencies sending cases and excludes joint investigations (e.g. FBI + state/local task force). The top agencies are: DEA, FBI, ATF, and state/local.

Figure A3.12: Fraction of Crack-Cocaine Seizures from 280-290g, Full Coverage States, NIBRS

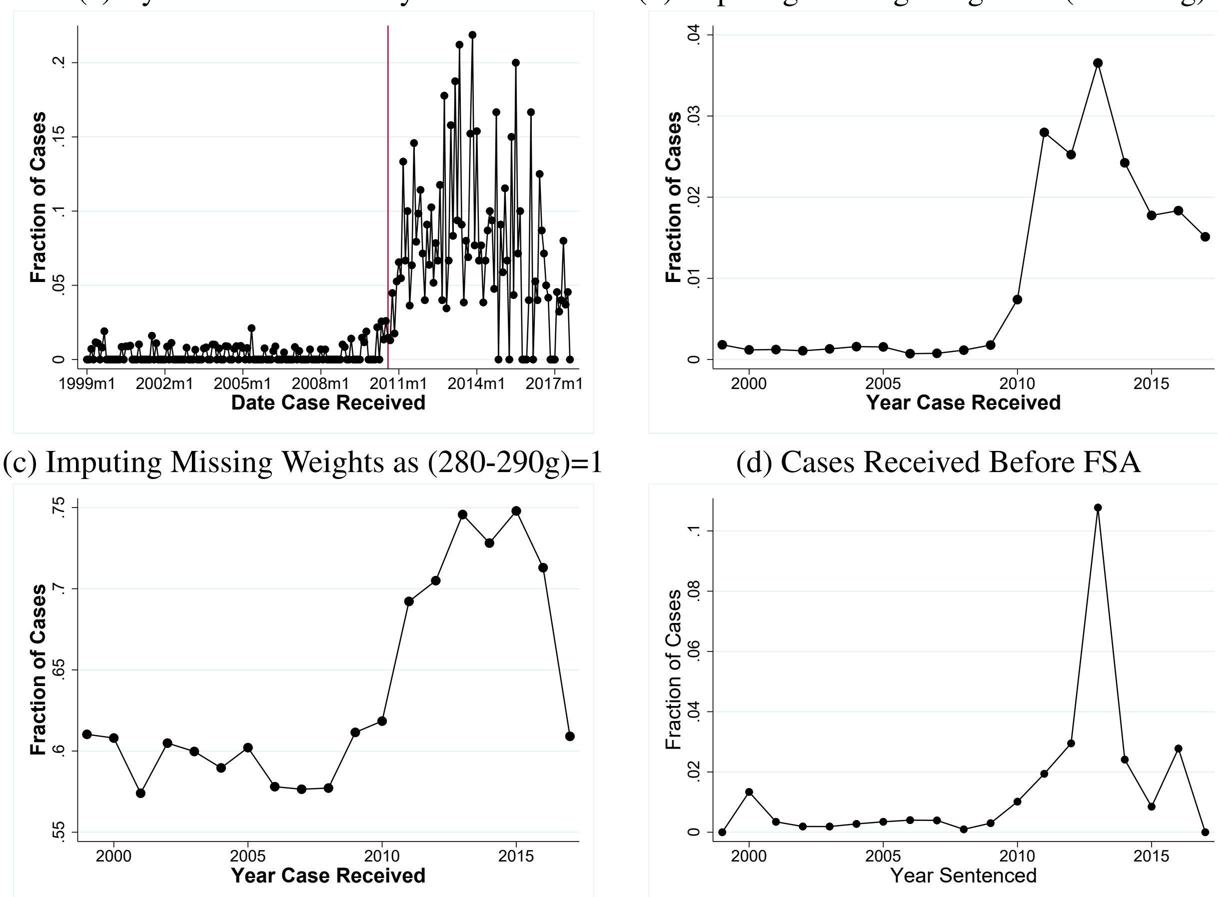


Notes: The figure above plots the fraction of crack-cocaine seizures with 280-290g by race. The sample is limited to states with full coverage (i.e. all agencies in the state participating) starting in 2012 and with 90% coverage or more from at least 2008-2015.

Figure A3.13: Fraction of Cases with 280-290g Over Time, EOUSA

(a) By Month Received by EOUSA (b) Imputing Missing Weights as (280-290g)=0

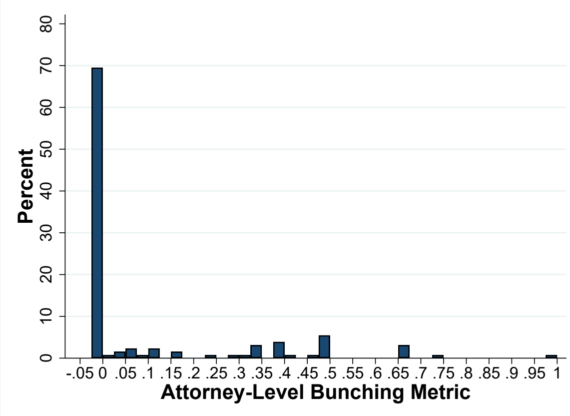
(c) Imputing Missing Weights as (280-290g)=1 (d) Cases Received Before FSA



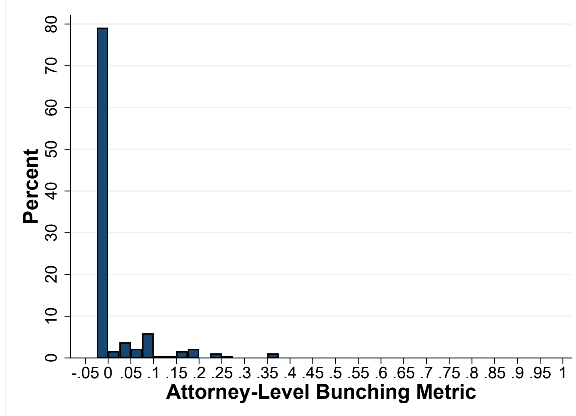
Notes: Panel (a) plots the fraction of cases with 280-290g (excluding cases with missing drug weights) by the month the case was received. The vertical red line indicates the date the Fair Sentencing Act was passed. In panel (b), I re-code the 280-290g dummy variable equal to zero if the drug weight is missing (typically, I leave the dummy variable missing if the drug weight is missing). In panel (c), I do the opposite, coding the 280-290g dummy variable equal to one if the drug weight is missing. In both cases, there is a sharp increase in the fraction of cases at 280-290g after 2010. Since panel (b) more accurately matches the statistics from the USSC final sentencing data, I use that imputed value for various robustness tests. Panel (d) plots the fraction of cases with 280-290g in each year for cases that are received by the EOUSA prior to the signing of the Fair Sentencing Act.

Figure A3.14: Histograms of Attorney-level Bunching Metric at 280-290g, EOUSA

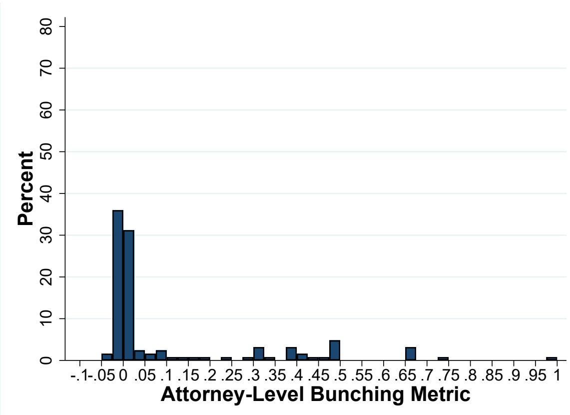
(a) 280-290g, National



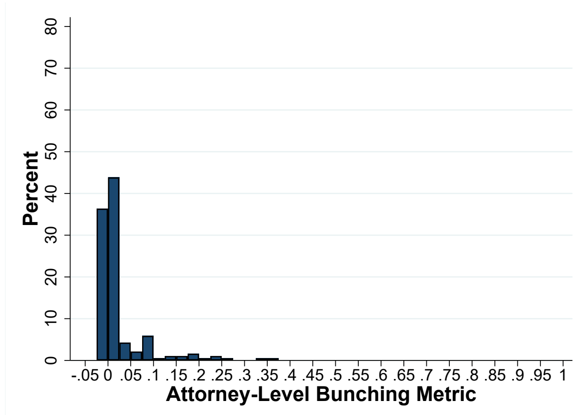
(b) 280-290g, Missing = 0, National



(c) 280-290g, District



(d) 280-290g, Missing = 0, District

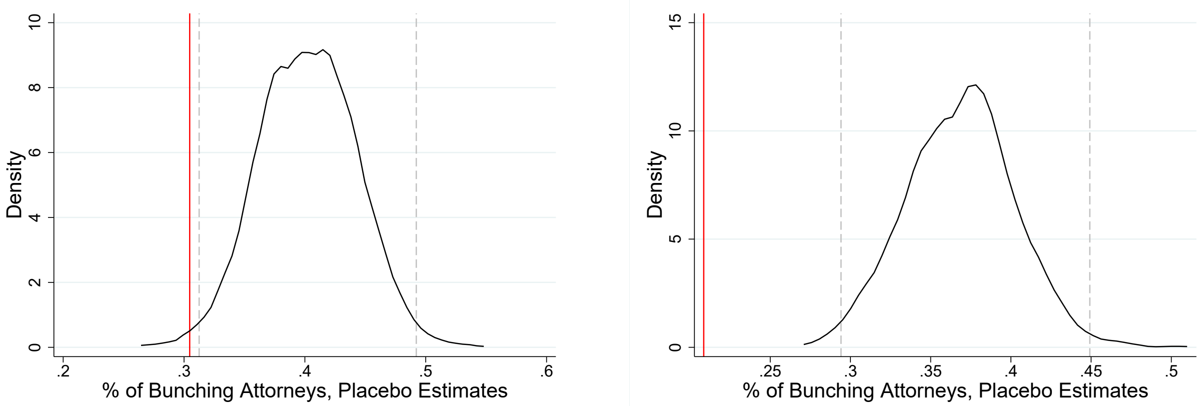


Notes: The figures above plot histograms of attorney-level bunching metrics, which are calculated as the difference between each attorney's fraction of cases with 280-290g post-2010 and the average fraction of cases with 280-290g at "baseline." In the national case (panels (a) and (b)), the baseline is the average fraction of cases with 280-290g prior to 2010. In the district case (panels (c) and (d)), the baseline for an attorney in district **A** is the average fraction of cases with 280-290g prior to 2010 in district **A**. Panels (b) and (d) include cases where the drug weight field is missing by coding the 280-290g dummy variable equal to zero when the drug weight is missing. I define an attorney as a "bunching attorney" if their bunching metric is above zero, thus the exact fraction of bunching attorneys for each panel is as follows: (a) 30.5%, (b) 20.9%, (c) 31.2%, and (d) 20.9%. These figures are limited to attorneys with 10+ cases post-2010. Limiting to 15+ cases delivers similar results. Limiting to 5+ cases decreases the fraction of bunching attorneys to: (a) 21.2%, (b) 14.2%, (c) 21.4%, and (d) 14.2%. Even imputing missing weight cases as though they **are** 280-290g cases (the highly unrealistic result in Figure A3.13c) implies that only 70% of attorneys bunch at 280-290g.

Figure A3.15: Histograms of Randomized Attorney-level Bunching Metric at 280-290g, EOUSA

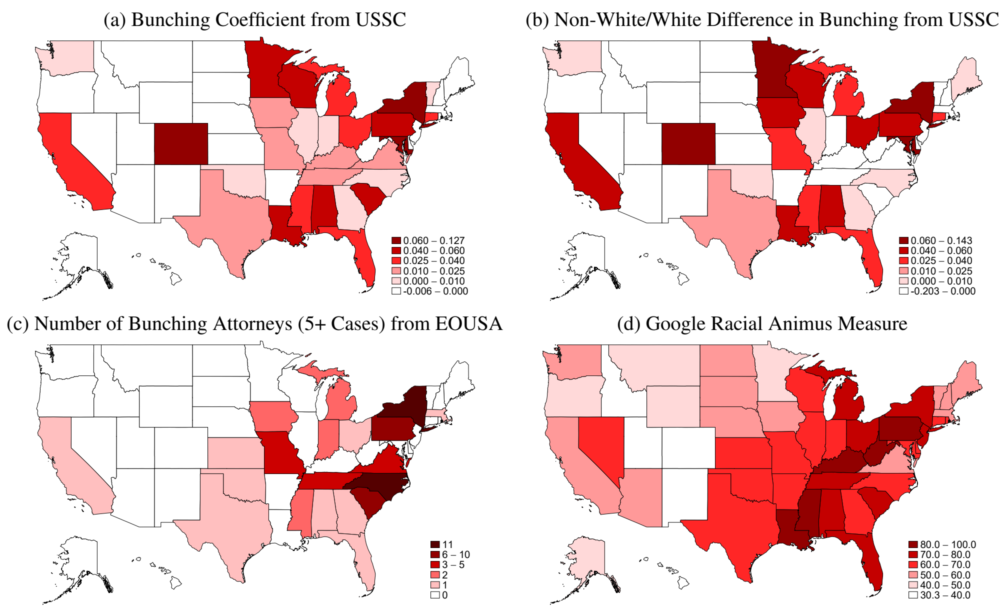
(a) 280-290g, National

(b) 280-290g, Missing = 0, National



Notes: I randomly re-assign all cases in the sample of attorneys with 10 or more cases after 2010, maintaining the same overall fraction of 280-290g cases in each year. After doing this random re-assignment, I calculate the number of bunching attorneys. I do this 1,000 times and plot the placebo estimates from the non-missing data in panel (a) and from the data with missing values imputed in panel (b). The gray dashed lines indicate the 1st and 99th percentiles of the placebo distribution and the red line indicates the fraction of bunching attorneys from the true data.

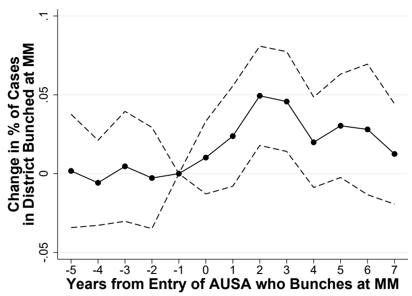
Figure A3.16: Map of State-level Bunching and State-level Racial Disparity in Bunching



Notes: Panel (a) plots the state-level bunching estimate for all states with a sufficient number of cases. Panel (b) plots the difference between the state-level bunching estimate for white offenders and the state-level bunching estimate for black and Hispanic offenders for all states with a sufficient number of cases. Panel (c) plots the number of prosecutors who bunch in each state (among those prosecutors with 5+ drug cases after 2010). Panel (d) plots the racial animus index derived from Google search volume for a racial slur and introduced by Stephens-Davidowitz (2014). For Panels (a) and (b) there are several states that do not have enough cases to estimate bunching or racial disparities in bunching at 280-290g (these states are: AZ, DE, HI, ID, MT, ND, NH, NJ, NM, NV, OR, RI, SD, UT, WY). I pool all of these states in one regression and apply the resulting coefficient.

Figure A3.17: Additional Evidence of Prosecutorial Discretion in Bunching, Alleyne Results and Movers Results, EOUSA

(a) Effect of Entry of a Bunching AUSA

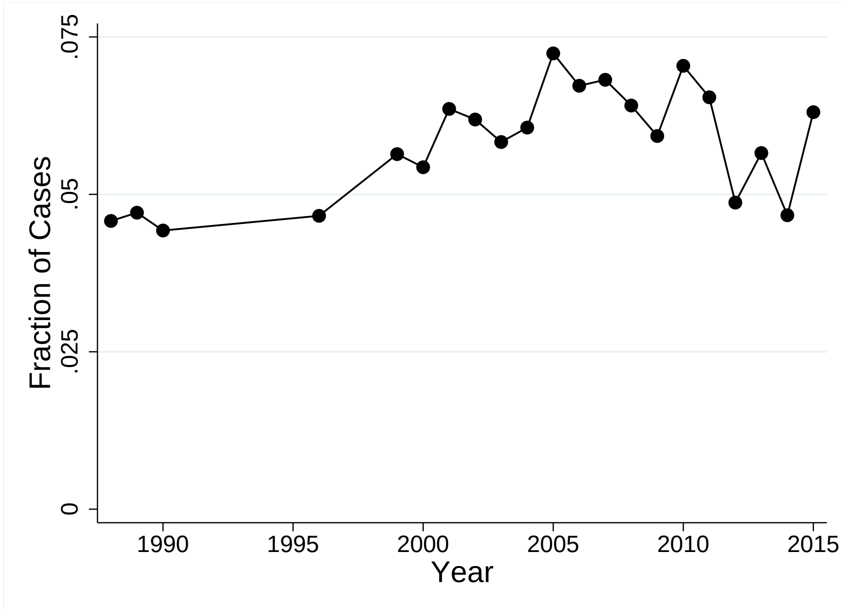


(b) Effect of Entry of a Bunching AUSA, Low-Bunching Districts

(c) Effect of Exit of a Bunching AUSA

Notes: Panels (a) and (b) plot the change in the percent of cases that are bunched at the mandatory minimum (MM) threshold (50g pre-2010 and 280g post-2010) after a "bunching" prosecutor enters a district. For these figures, I identify prosecutors who switch districts, who bunch at the mandatory minimum threshold in their first district, and who have 5 or more cases in their first district. I then identify the districts that they switch into and analyze the fraction of cases bunched at the mandatory minimum for all other prosecutors in that district. Panel (a) shows that prior to entry of a bunching prosecutor, district-level bunching does not change year-to-year, but that immediately after the bunching prosecutor enters, all other prosecutors in that district increase their fraction of cases bunched at the threshold. Panel (b) shows that this increase is driven by districts that have low-levels of bunching (below the median for all districts) prior to the entry of the bunching prosecutor. Panel (c) plots the bunching activity for the districts from which these prosecutors are leaving. This analysis is limited to the first bunching attorney from panels (a) and (b) that leaves the district. There is not a decrease in the prevalence of bunching after bunching prosecutors exit a district. This suggests bunching at the mandatory minimum threshold is not related to a temporary behavior shift, such as increased competition among attorneys, but that it may be related to something more permanent, such as learning about techniques or developing beliefs/norms. The dashed lines in panels (a)-(c) are 90% confidence intervals. Since these figures rely on prosecutors who move from one district to another and require reasonably long pre- and post-periods, I use data from 1994-2016 and identify the first moving attorney for post-1999 years only (insuring a 5-year pre-period for every district). In practice, this means the figures above are largely based on bunching at 50-60g (the pre-2010 mandatory minimum). Restricting to post-2010 moves does not yield a large enough sample of movers with sufficient cases to classify them as bunching versus non-bunching.

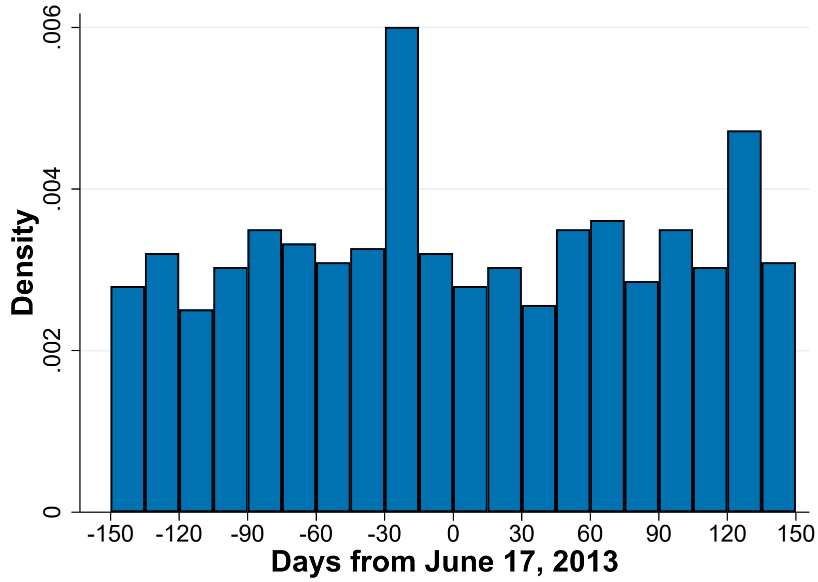
Figure A3.18: Fraction of Cases in 50-60g by Year, from USSC Sentencing Data



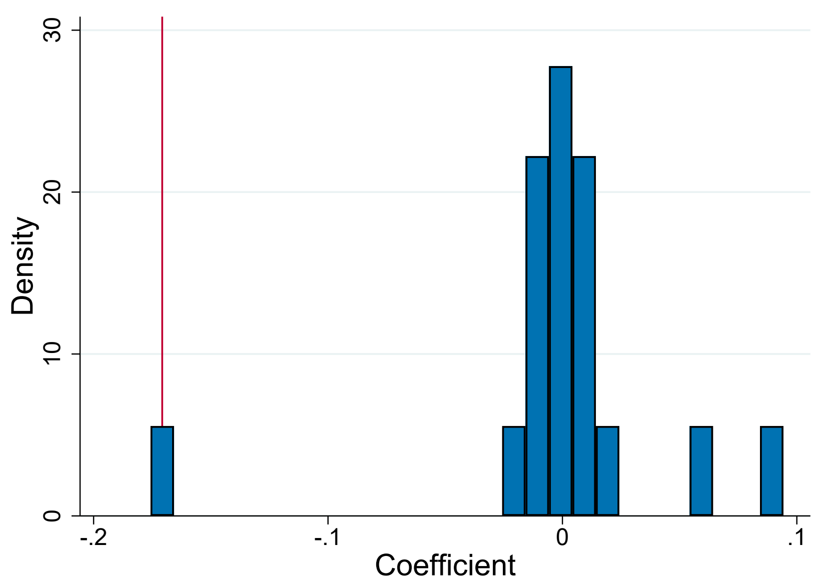
Notes: The figure above plots the fraction of all cocaine (powder and crack) cases with 50-60g by year. The sample is limited to cases with drug weights from 0-1000g. All cocaine cases are used because earlier years (1988-1990) do not distinguish between types of cocaine. This figure indicates that cases bunched above the pre-2010 10-year mandatory minimum threshold increased by about 60% from 1988-90 to 2010. Over this same time period, the average weight of cases from 0-1000g decreased. This suggests that the practice of bunching cases at the mandatory minimum was potentially learned over time, which is consistent with the evidence on movers and the spread of bunching in Figure A3.17.

Figure A3.19: Tests of Validity for Alleyne v. US Result, EOUSA

(a) Density of Cases Received Around June 17, 2013 (Date of Decision in Alleyne)

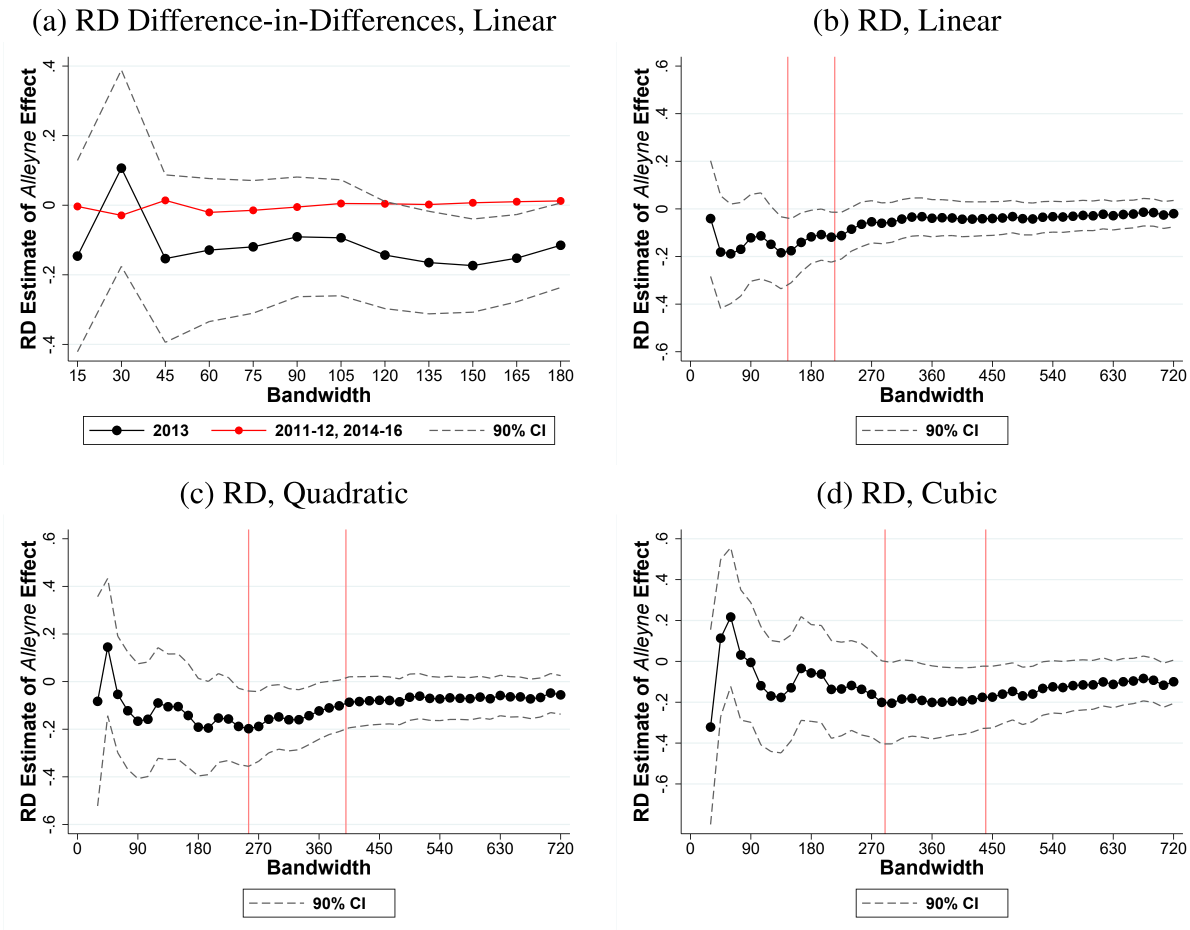


(b) Estimate of Discontinuity Around June 17 in All Years 1999-2016



Notes: Panel (a) plots the density of cases around the June 17, 2013 (centered at zero) and grouped into 15-day bins. June 17, 2013 is the day Alleyne v. US was decided. Outside of the large number of cases from -30 to -15 days before Alleyne was decided, the density is relatively smooth through that date. Panel (b) plots a histogram of the estimated discontinuity around June 17 in all years from 1999-2016. The estimates are centered at zero and the coefficient in June 2013 (marked by the red line) is twice as large as the next largest estimate of any sign and over 4 times larger than the next largest negative estimate.

Figure A3.20: Robustness of Alleyne v. US Result to Choice of Bandwidth and Polynomial, EOUSA



Notes: The figures above display estimates for the effect of Alleyne v. US (a case that strengthened evidentiary requirements) on the prevalence of bunching at 280-290g. Each panel displays estimates across many different bandwidth choices (i.e. the number of days before and after June 17 included in the regression) and different polynomial choices (i.e. the polynomial of the running variable, number of days from June 17, included in the regression) are shown across panels. Panel (a) displays coefficient estimates from the RD difference-in-differences regression for bandwidths from 15-180. Since the difference-in-difference estimates use multiple years, bandwidths above 160 days are asymmetric. The black line in panel (a) displays the estimates from 2013, the red line displays the estimates from all other years after 2010 (when nothing in particular happened around June 17). Panels (b)-(d) estimate a typical RD regression (i.e. not using variation around June 17 in other years). This allows me to extend the bandwidth to 2 years before and after Alleyne v. US. In these panels, the first red line denotes the CER-optimal bandwidth and the second red line denotes the MSE-optimal bandwidth (Cattaneo et al. 2018). In panel (b), for example, the estimate approaches zero at larger bandwidths–this is to be expected. As we get further from the cutoff, the a linear polynomial becomes an increasingly bad fit. In all three panels, the optimal bandwidths yield estimates that are statistically different from zero (or marginally statistically significant).

3.9 Appendix B. Alternative Methods of Estimating Bunching

Comparing Aggregated Pre- and Post-2010 Densities

Most papers using the "difference-in-bunching" approach can be fit into one of two categories. In one, authors estimate bunching using the conventional polynomial method (see the second section below for a detailed description) separately for groups where the threshold applies and for groups where the threshold does not apply, using the latter as a placebo test (Best et al. 2015; Fack and Landais 2016; Gelber, Jones, and Sacks 2017; Zaresani 2017; Chen et al. 2018). In the other, authors directly compare the group where the threshold applies to the group where the threshold does not apply. Even within the direct comparison category, strategies differ. Several papers compare the distributions by aggregating the data into bins and calculating the difference in levels between the actual and the counterfactual distributions (Brown 2013; Best et al. 2018; Best and Kleven 2018; Cengiz, Dube, Lindner, and Zipperer 2018). Others compare the distributions using regression analysis on the microdata (Kleven et al. 2011; Behaghel and Blau 2012; Sallee and Slemrod 2012; Chetty, Friedman, and Saez 2013; Dwenger et al. 2016; Goncalves and Mello 2018; and Traxler et al. 2018). These papers frequently estimate the difference in the probability an observation is in a given bin between the actual and the counterfactual setting.

In this paper, I employ both direct comparison methods (aggregate/binned analysis and microdata analysis). I am primarily interested in estimating the change in the probability a case is charged with 280-290g after 2010 and whether that change in probability differs by race. In addition, some analyses in the paper preclude aggregating the data into bins because they rely on data that do not include precise drug quantities. For these reasons, I follow the papers that use regression analysis on microdata to compare the pre- and post-2010 crack-cocaine distributions.

To show robustness to the other "difference-in-bunching"/direct comparison method, I aggregate the cases into 10g bins pre- and post-2010. Following Best et al. (2018), I estimate 90% confidence intervals with a bootstrap procedure that samples cases with replacement from the microdata before aggregating to the 10g bin level.[1] I compare the binned distributions to estimate the net change in bins below 280g, at 280-290g, and above 290g.

Aggregate bunching analyses yield very similar results. Figure B3.1 below plots the counterfactual scaled pre-2010 density and the actual post-2010 density. The spike at 280g in the post-2010 density is the bunching that is detected in Table 3.2. After 2010, there is a 3.5 percentage point increase in cases with 280-290g. I also show the densities by race. The bunching at 280g in the post-2010 density is larger for black and Hispanic offenders. After 2010, the rise in cases with 280-290g is about 2 percentage points higher for black and Hispanic offenders than for white offenders.

In Figure B3.2a, I plot the difference between the post-2010 and the scaled pre-2010 densities for each 10g bin and add confidence intervals by using 50 bootstrapped samples from the microdata. In addition, I also display a table of the statistical results for the binned missing mass analysis in Figure B3.2b. When this difference is below zero, it means the bin contains relatively fewer cases after 2010 and when the difference is above zero, it means the bin contains more cases after 2010.

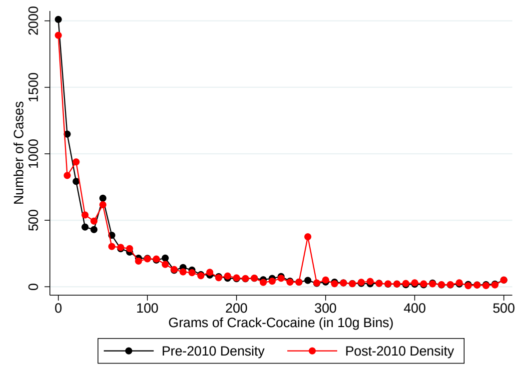
The figure shows an increase of about 340 cases in the 280-290g bin post-2010, a net increase in cases above 280g, and a net decrease below 280g. Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Again, this point estimate is noisy. In fact, summing the upper confidence interval for all bins below 280g implies a net increase in that section of the

[1] I draw 50 random samples from the microdata and do the binned analysis on each sample. The final number of cases for each bin is calculated as the mean of the number of cases across all 50 samples, and the final standard error is calculated as the mean of the standard error across all 50 samples.

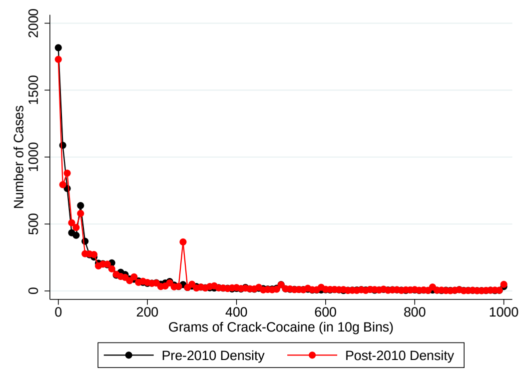
distribution. The key takeaway is that changes in the distribution below 280g can account for the excess mass at 280g, whereas changes in the distribution above 280g cannot. In other words, an offender charged with 280-290g post-2010 would likely have been charged with less than 280g had they been sentenced prior to 2010. Table B3.1 displays the results from similar binned analyses using the NIBRS data, DEA data, and EOUSA data.

Figure B3.1: Scaled Pre-2010 Distribution of Recorded Weights vs. Post-2010 Distribution

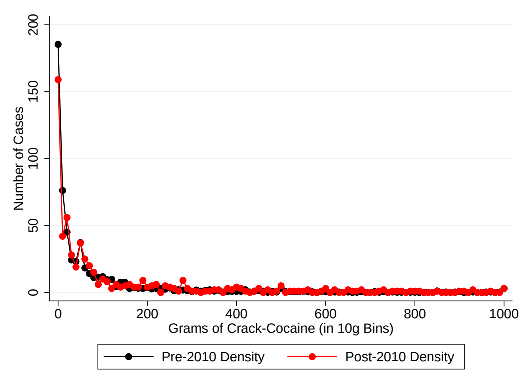
(a) All Offenders



(b) White Offenders



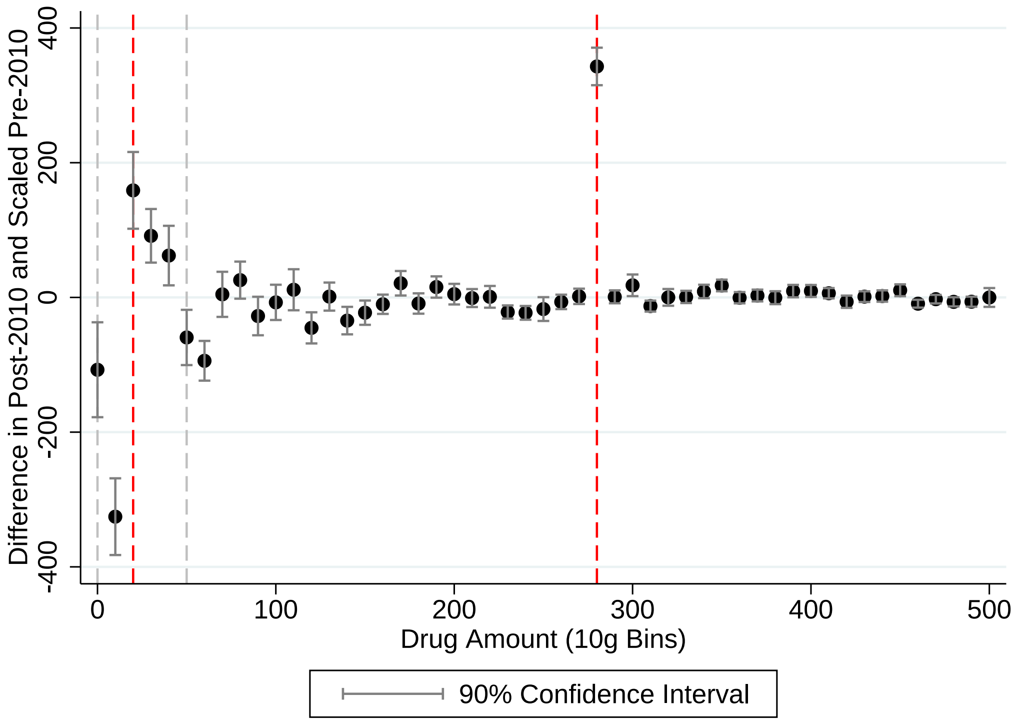
(c) Black and Hispanic Offenders



Notes: Figure B3.1a plots the scaled density of drug quantities pre-2010 (in black) and the actual density of drug quantities post-2010 (in red) for all offenders. The amounts are aggregated into 10-gram bins and limited to drug quantities under 1000g. Figures B3.1b and B3.1c do the same but restrict the sample to white offenders or black and Hispanic offenders, respectively.

Figure B3.2: Post-2010 Density Minus Scaled Pre-2010 Density

(a) Difference between Post-2010 and Pre-2010 Distribution of Drug Amounts



(b) Fraction of Bunching Accounted for by Different Ranges

Range	Net Difference	90% CI	% Bunching at 280g
0-20g	-435.56	(-558.17, -312.94)	128.67%
20-50g	293.49	(158.10, 428.88)	-86.70%
50-60g	-52.63	(-101.67, -3.59)	15.55%
60-100g	-65.00	(-184.58, 54.59)	19.20%
100-280g	-122.43	(-414.25, 169.39)	36.17%
0-280g	-382.13	(-1100.58, 336.32)	112.89%
290-500g	43.44	(-146.74, 233.62)	-12.83%

Notes: The figure above plots the difference between the post-2010 density and the scaled density of drug quantities in pre-2010 for each 10-gram bin. Confidence intervals are calculated by bootstrapping as discussed in the text. The red dashed lines correspond to the post-2010 mandatory minimum bins (28g and 280g) and the gray dashed lines correspond to the pre-2010 mandatory minimum bins (5g and 50g). Summing the changes in bins above 280g, I find a net increase in that section of the distribution after 2010. The point estimate on the net change is noisy, but even summing the lower bound of the confidence interval for all bins above 280g can only account for about 46% of the increase in the 280-290g bin. On the other hand, the net change below 280g can account for 120% of the increase in the 280-290g bin. Even the changes from 50-280g can account for 85% of the increase in the 280-290g bin. Panel B displays statistical results for relevant drug amount ranges.

Table B3.1: All Bunching Results using Aggregated/Binned Comparison with Bootstrapped SEs

	Pr(280-290g Crack-Cocaine Recorded)					
	(1)	(2)	(3)	(4)	(5)	(6)
After 2010	0.0347***		-0.0002***		-0.0006***	0.0771***
	(0.0020)		(0.0001)		(0.0002)	(0.0054)
After 2010 x White		0.0126**		-0.00002		
		(0.0062)		(0.0001)		
After 2010 x Black/Hispanic		0.0359***		-0.0003***		
		(0.0023)		(0.0001)		
Constant	-0.0003***	-0.0001***	0.000002	0.0000002***	0.000006***	-0.0008***
	(0.00002)	(0.0001)	(0.0000008)	(0.000001)	(0.000002)	(0.0001)
Data	USSC, Final Sentencing	USSC, Final Sentencing	NIBRS, Drug Seizures	NIBRS, Drug Seizures	DEA, Drug Seizures	EOUSA, Prosecutor Files
Bins	100	100	100	100	100	100
Observations	57,101	52,940	203,700	203,700	100,306	24,493

Notes. Bootstrapped standard errors in parentheses. Standard errors are calculated from the standard deviation in estimates derived from 50 replications where in each replication cases are sampled with replacement before aggregating to the 10g bin level. All specifications above use the sample of offenses with drug amounts between 0 grams and 1000 grams. Specifications with the white/non-white and after 2010 interactions also include a dummy variable equal to one for black and Hispanic offenders. Columns 1-2 show the main bunching result for the final sentencing data. Columns 3-5 show no increase in bunching for drug seizure amounts. Column 6 shows an increase in bunching in prosecutor case management files.
*** p<0.01, ** p<0.05, * p<0.1

Comparing an Estimated Counterfactual and Post-2010 Densities

Many bunching papers, for lack of variation in the threshold of interest, estimate bunching by constructing the counterfactual density from the actual bunched density. To do this, one typically aggregates the data into bins and estimates a regression of the count in each bin on a high-order polynomial of the bin's value and dummy variables for bins in the bunched "window." The estimates from that regression (not including the bunching dummy variables) can be used to predict a smooth distribution of bin counts. Authors then compare that smooth density to the actual density to calculate the degree of bunching in the actual density. My main results are also robust to this method.

To start, I collapse the data on drug quantities for all cases after 2010 to 10 gram bins. I then run a regression of the count of cases on a seventh order polynomial of the bin values and dummy variables for the bins 0-10g, 270-280g, and 280-290g. Then, using the coefficients from the seventh order polynomial and the dummy variable for the bin 0-10g, I calculate a smooth counterfactual distribution. For graphical purposes, I re-scale that smooth distribution to have the same total number of cases as the true distribution. Next, I calculate the percent of all cases that are in the 280-290g bin in the true distribution, the percent of all cases that are in the 280-290g bin in the counterfactual distribution, and the difference between those two percentages. Finally, I run a regression of the difference between the true and counterfactual distributions on a dummy variable equal to one for the 280-290g bin and equal to zero otherwise (bootstrapped standard errors are calculated by re-sampling the residuals from the polynomial estimation with 200 replications). I carry out a similar procedure to estimate the difference in bunching between white and black and Hispanic offenders (the major difference being that I estimate the counterfactual distributions separately for white and black and Hispanic offenders and that the final regression includes an interaction between the 280-290g bin dummy and a dummy for black and Hispanic offenders).

First, I construct the counterfactual density by aggregating the data to 10-gram bins, summing the number of cases in each bin. With this aggregated data, I estimate a regression of the bin counts on a seventh-order polynomial of the bin values, dummies for the 270g and 280g bins, and a dummy for the 0g bin.

$$Count_b = \alpha_0 + \sum_{i=1}^{7} \beta_i (Amount_b)^i + \gamma_1 Bin270_b + \gamma_2 Bin280_b + \delta_1 Bin0_b + \varepsilon_b \tag{1}$$

where $Count_b$ is the total number of cases in bin b, $Amount_b$ is the value of bin b, and $Bin[X]_b$ is a dummy variable indicating if the bin's value equals X. I use the parameter estimates from (8) (excluding γ_1 and γ_2) to predict a smooth density of bin counts. Furthermore, I adjust the predicted counts to force the smooth density to have the same number of cases as the actual density. I plot the counterfactual density and the actual post-2010 density below in Figures B3.3 and B3.4.

Using the predicted counts from the counterfactual density and the actual counts post-2010, I construct the percent of cases in each bin for each density. I then calculate the difference in these percentages and run the following regression, bootstrapping the standard errors from 200 replications:

$$(\text{\% in Post2010 - \% in Predicted})_b = \alpha + \beta Bin280_b + \varepsilon_b$$

The resulting $\beta = 0.0352$ and $SE_\beta = 0.0169$.

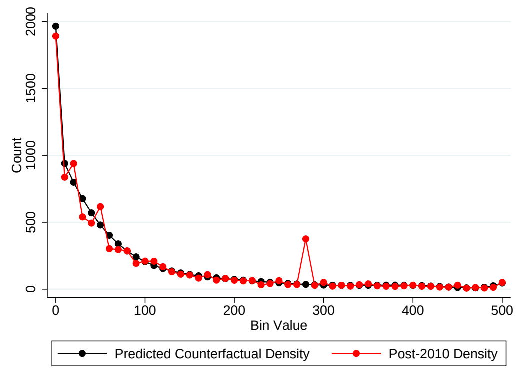
Next, I estimate:

$$(\text{\% in Post2010 - \% in Counterfactual})_{br} = \alpha + \beta Bin280_b + \gamma NonWhite_r + \delta Bin280_b \times NonWhite_r + \varepsilon_b$$

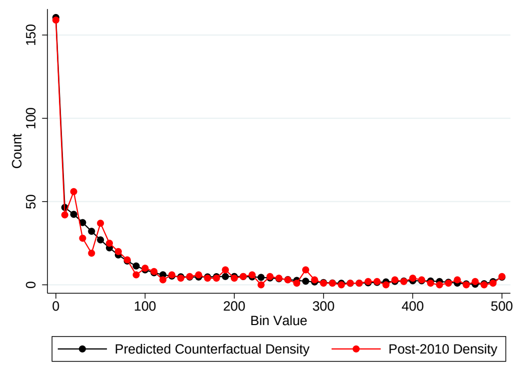
Using the Saez (2010) and Chetty et al. (2011) method, I estimate $\delta = 0.0237$ and $SE_\delta = 0.0119$. Using the difference-in-bunching method, I estimate $\delta = 0.0216$ and $SE_\delta = 0.0109$. In all analyses, I detect substantial bunching after 2010 and disproportionate bunching after 2010 for black and Hispanic offenders.

Figure B3.3: Predicted Counterfactual Density and Post-2010 Density

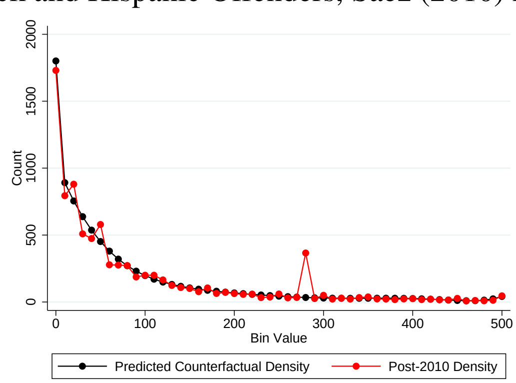
(a) All Offenders



(b) White Offenders, Saez (2010) Method



(c) Black and Hispanic Offenders, Saez (2010) Method



Notes: In panel (a), I plot a predicted counterfactual density of drug quantities (in black) and the actual density of drug quantities post-2010 (in red). In panels (b) and (c), I plot predicted counterfactual densities of drug quantities (in black) and the actual densities of drug quantities post-2010 (in red) by race. The amounts are aggregated into 10-gram bins and limited to drug quantities under 500g.

3.10 Appendix C. Supplementary Materials for Prosecutor Model

Prosecutor Responses to Changing Mandatory Minimum Thresholds

The model from Section 3.5.5 also has implications about how the optimal choice in period $t = 0$ relates to the optimal choice in period $t = 1$. I outline this in Section 3.2.2, and provide additional detail in this Appendix section.

Assuming that there are no fixed costs to building a case and that there are no changes in the objective function other than the change in the sentencing schedule, then a prosecutor who chooses not to bunch a case at a mandatory minimum threshold for a sentence X in one period would not bunch the same case at a higher mandatory minimum threshold for a sentence $Y \leq X$ in another period. In other words, a prosecutor not taking on the costs of bunching for a given gain would not take on even greater costs for the same or lesser gain.

For example, when a prosecutor chooses $a^{0*} = s < 5$, this implies that their utility from choosing s is higher than their utility of choosing 5g or 50g: $u(s)^0 > u(5)^0$ and $u(s)^0 > u(50)^0$. Since $a^{1*} = 28$ yields the same benefits as $a^{0*} = 5$ but requires greater costs, then $u(5)^0 > u(28)^0$. These two statements (and the assumptions above) imply that $u(s)^1 > u(28)^1$, which means that the prosecutor should also choose $a^{1*} = s < 5$. The same revealed preference argument can be made for why $u(s)^1 > u(280)^1$. Table C3.1 shows these possible rational choices of a^{1*} for a given a^{0*} and ranges of s.

Table C3.1: Relationship between a^{0*} and a^{1*} for relevant ranges of seized evidence

	(1)	(2)	(3)	(4)	(5)
	$s < 5$	$28 > s \geq 5$	$50 > s \geq 28$	$280 > s \geq 50$	$s \geq 280$
$a^{0*} = s$	$a^{1*} = s$	$a^{1*} = \{s, 28\}$	$a^{1*} = s$	$a^{1*} = \{s, 280\}$	$a^{1*} = s$
$a^{0*} = 5$	$a^{1*} = \{s, 28\}$	–	–	–	–
$a^{0*} = 50$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 28, 280\}$	$a^{1*} = \{s, 280\}$	–	–

Ultimately, this means that there will be an increase in the share of cases with $a < 5$ post-2010 (increases from cases previously bunched at 5g and 50g); an ambiguous change in the share of cases with $28 > a \geq 5$ (increases from cases previously bunched at 50g and decreases from cases previously bunched at 5g), an increase in the share of cases with $50 > a \geq 28$ (increases from cases previously bunched at 5g and cases previously bunched at 50g); a decrease in the share of cases with $280 > a \geq 50$ (decreases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$), and an increase in the share of cases with $a \geq 280$ (increases from cases previously bunched at 50g and cases previously left with $a = s \geq 50$). See Figure A3.2 for a graphical representation of this.

Prosecutors' Signal Extraction Problem

The racial disparity in bunching at 280g after 2010 could be due to statistical discrimination. Recall that seized evidence s is a noisy measure of true drug trafficking d. Suppose that, on average, black and Hispanic defendants have higher true drug trafficking amounts:

$$d_r \sim N(\bar{d}_r, \sigma_d^2)$$
$$\bar{d}_{bh} > \bar{d}_w$$

Since s is a noisy measure of true drug trafficking d, we can write s as follows:

$$s = d + \nu,\ \nu \sim N(\mu, \sigma_\nu^2)$$

This implies that $E(d|s,r,x) = \bar{d}_r \times (1-\alpha) + (s-\mu) \times \alpha$ where $\alpha = \sigma_\nu^2 / (\sigma_\nu^2 + \sigma_d^2)$. Since $\bar{d}_{bh} > \bar{d}_w$, $E(d|s,bh,x) > E(d|s,w,x)$. Since the prosecutor does not observe d, they instead use $l^*(E(d|s,r,x))$. I denote this as $l^*(s,r,x)$, and the setting described here implies that $l^*(s,bh,x) > l^*(s,w,x)$. In other words, the prosecutor's expectation over true drug trafficking d "justifies" a higher sentence for black and Hispanic offenders. This decreases their cost of choosing $a > s$ because the associated mandatory minimum sentence will be less of a deviation from that sentence l*. Prosecutors may also use another defendant characteristic x_1 to solve the signal extraction problem (as detailed above) and arrive at $l^*(s, r, x = x_1) > l^*(s, r, x \neq x_1)$.

3.11 Appendix D. Data Appendix

United States Sentencing Commission (USSC) Federal Sentencing Data

These data contain the universe of federal sentences from 1999-2015. The data were obtained from the ICPSR "Monitoring of Federal Criminal Sentences" series here: `https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/83`.

The data itself is compiled from several court documents: (1) the Judgment and Conviction Order (JC), (2) the Pre-sentence Report (PSR), and (3) the Statement of Reasons (SOR). The PSR is prepared by the probation officer in consultation with the prosecutor and the defense. It is a detailed report on the offender and their offenses intended to aid the judge in making the factual determinations that affect sentencing. The SOR is a form filled out by the judge that details their findings and whether/why they differ from the PSR. The JC is the final ruling in the case that outlines the adjudication and the sentence. Key variables from the data are described below:

Crack cocaine offense. Whether or not the case involves a crack cocaine offense is derived from the raw variables DRUGTYP{X} provided by USSC. These variables contain the types of drugs involved in the offense. This information is taken from the Judgment and Conviction Order (JC), if present. If it is not included in the JC, the information is taken from the Pre-sentencing Report (PSR) prepared by the probation officer assigned to the case. According to USSC, if the information in these documents conflicts, the JC takes precedent.

Drug quantity. The amount of drugs involved in the case is derived from the variables WGT{X} provided by the USSC. These variables contain the gram amount for drug {X} corresponding to DRUGTYP{X}. I use the weight corresponding to the drug type crack cocaine for each case. The values for WGT{X} are converted from variables DRGAM{X} and UNIT{X}. Information on drug amount and drug unit is taken from the Statement of Reasons (SOR), if present. If not present in the SOR, the information is taken from the

PSR. According to the USSC, if the information in these documents conflicts, the SOR takes precedent.

Offender race. I code offender race based on the USSC variables NEWRACE, which categorizes offenders as non-Hispanic white, non-Hispanic black, or Hispanic. The variable NEWRACE is a combination of raw variables MONRACE and HISPORIG. The information for these variables is taken from the PSR. In fact, the USSC notes that offender race is self-reported to the probation officer.

Other offender characteristics (e.g. education). These are also derived primarily from the PSR.

Year. The year used for analyses is derived from the variable AMENDYR, which represents the year of the guideline manual used for sentencing guidelines calculations. This information is taken from the PSR.

District. The district used for analyses is derived from the variables DISTRICT, which represents the federal district the offender is sentenced in. This information is taken from the JC, if available, and from the PSR, if not. If both documents are available, and the information conflicts, the JC takes precedent.

FL State Inmate Database

These data contain all inmates who have been released from a FL state prison since October 1997. The data were obtained here: `http://www.dc.state.fl.us/pub/obis_request.html`. Key variables from the data are outlined below:

Offense/drug quantity. The offense field indicates all of the inmate's known offenses in FL. For drug offenses, the field contains the drug name. In FL, powder-cocaine and crack-cocaine cases are both recorded as "cocaine." For many of the drug offenses, the field contains a label indicating if the offense was with 0-28g of cocaine, 28-200g, 200-400g, or 400+g.

Offender race. Offender race is included as part of the "basic inmate information" file. There is no information on how race is determined. I expect it is similar to the federal court data, in which race is self-reported. In the FL data, the race field includes labels for “black”, “Hispanic”, and “white” inmates.

In robustness tests, I use similar data from North Carolina. It also contains an offense string that provides information about drug type and quantity. However, the string does not always specific the type of drug. These data cover cases that are handled at the state/local level as opposed to federal court (those cases included in the USSC data). This is important because state and local authorities could send more of their high weight, 280g cases to federal court after 2010. Similarly, federal prosecutors could pull more of these types of cases from state and local courts after 2010. A case can enter the federal system for procedural reasons: drugs are trafficked across state lines or the arrest is made by federal agents. However, cases can also be prosecuted federally for more arbitrary reasons. Wright (2006) notes that sorting into federal versus state is often determined by law enforcement agents involved with the case and/or the prosecuting attorneys, but it is never the official purview of judges or defense attorneys.[1] Why might local law enforcement or attorneys wish to pass a case on to the federal courts? For one, local authorities may not have the time or resources to properly pursue a case. Also, Wright suggests that federal sentencing is typically harsher than state sentencing, and that this gap could motivate jurisdiction decisions.

NIBRS Property Segment

These data contain information on drug quantity and drug type for drugs seized by NIBRS-participating police departments. The data were obtained here:
icpsr.umich.edu/icpsrweb/NACJD/series/128. Key variables from the data are outlined below:

[1]**Wright, Ronald.** 2006. “Federal or State? Sorting as a Sentencing Choice.” *Criminal Justice* 21 (2): 16-21.

Drug quantity. The drug quantity field is populated when there is a drug seizure by the department. It is equal to the total quantity of drugs seized.

Offender race. The race field for NIBRS does not include an indicator for whether the offender is Hispanic. An ethnicity field is available only in later years, so I focus on white versus black offenders in this data. There is no information on how race of the offender is determined. I expect it is similar to other criminal justice data, in which race is self-reported.

For the primary analyses of the NIBRS data, I limit the sample to a balanced panel of agencies. For robustness checks, I limit to stats that have had full agency coverage in NIBRS since 2012 and over 90% coverage since 1998.

DEA STRIDE Database

These data contain information on drug quantity, drug type, and purity for seizures and undercover purchases sent to DEA labs for analysis. The data also indicate whether the drugs were obtained via seizure or undercover purchase. For drugs that were purchased, the data contains their price. The data were obtained from a FOIA request for all records related to cocaine from January 1, 1999 to December 31, 2015. Key variables from the data are:

Drug quantity. This field indicates the weight of the drug evidence received by the lab.

Drug type. This field indicates type of drug. The DEA does not use street names to refer to drugs in this data, meaning no drugs are referred to as crack-cocaine. For the main analyses, I use all drug types containing the word "cocaine," but results are similar if I focus on the "cocaine base" drug type.

Purity. This field indicates the chemical purity of the drug evidence received by the lab.

Acquisition. This field indicates whether the drug was acquired via seizure or undercover purchase.

Price. This field is populated if the drugs were acquired via undercover purchase. Price indicates the price paid for the drugs. In one robustness analysis, I plot the time series of price by month. To do this, I adjust the raw price field (described here) based on the purity of the drug, calculating a "price per pure gram."

EOUSA Case Management Files

These data contain information on cases handled by the EOUSA from the EOUSA's internal case management system: Legal Information Office Network System (LIONS). The data were obtained here:
`https://www.justice.gov/usao/resources/foia-library/national-caseload-data`.
Key variables from the data are:

Drug quantity/type. This field comes from the "controlled substances" screen of the LIONS software. According to the LIONS user manual, the controlled substances data "tracks information on controlled substances; includes type and quantity of all substances in a case." The manual instructs users to do the following: "Enter the actual quantity of the controlled substance seized. Fractions must be converted to one or two decimal places." The software itself, however, simply has a field for "quantity" to be entered with no instruction. In general, the drug weights recorded in the EOUSA data are much larger than the drug seizure weights reported by the DEA or NIBRS. In fact, drug quantities decrease in the DEA and NIBRS after 2010 but increase in the EOUSA. Also, the fraction of 280-290g cases at the district/month level in the EOUSA data is highly correlated with the fraction of 280-290g cases at the district/month level in the USSC data. These validation tests suggest the data entered into LIONS is indicative of total drugs involved/charged in the offense and not raw amount seized alone.

Staff ID/Assignment. The EOUSA data also contains an ID variable for the lead attorney assigned to the case. This ID is tied to the district. In other words, two attorneys can have the same numeric ID as long as they are in different districts. Also, this ID will not follow an attorney from one district to another.

Initials. Since the EOUSA numeric ID for lead attorney is not constant across districts, I use a field for the attorney's "initials" to follow attorneys who switch districts. The initials field is "initials of the staff member authorized to use the LIONS application." In most cases, the field contains 3 or more letters, making it likely that if I see the same initial in two different districts it is the same attorney. In practice, this initials-based ID appears to accurately identify attorneys who switch districts. First, attorneys who move from one district to another continue to bunch at 280g in the new district. Second, when an attorney moves into a new district, other attorneys in that district start to bunch more at 280g. Third, attorneys who I identify as "moved" are often disconnected from their old district in the data and connected to their new district. If the initials-based ID were totally random, we should not expect to see these three patterns.

Date received. The date the criminal case was received by the US Attorney's Office.

Sentence date. For cases that are sentenced, the EOUSA also notes the data of sentencing.

Judge ID. For cases that are brought to a judge, the EOUSA data contains an identifier for the judge involved and that identifier can be linked to a table of judge names. For robustness analyses, I examine the effect of judge race and political party on bunching at 280g. I obtain data on judge characteristics from Crystal Yang's paper on resource constraints and judicial vacancies:

`https://test.openicpsr.org/openicpsr/project/114590/version/V1/view?path=/openicpsr/114590/fcr:versions/V1/Data_2015_0150/Public-Use-Data&type=folder`

Chapter 4: The Long-run Economic Effects of School Desegregation

4.1 Introduction

In Brown v. Board of Education, the Supreme Court ruled de jure segregation of schools unconstitutional because "separate educational facilities are inherently unequal." That decision in 1954 ultimately set off a wave of desegregation plans over the next 30 years, many of which were court-ordered due to resistance from local school districts.[1] In terms of integrating schools, these plans were successful–by 1988, about 44% of black children were attending majority white schools. In the early 1990s, however, the Supreme Court issued three decisions which led to the dissolution of many court-ordered desegregation plans. Currently, only 23% of black children are attending majority white schools, a level of segregation not seen in the United States since 1968 (Orfield et al. 2014).[2]

Numerous studies document beneficial short term effects of school desegregation (e.g. Guryan 2004; Reber 2010; Johnson 2015; Bergman 2018). Recent work finds harmful short-run effects of re-segregation (Lutz 2011; Billings, Deming, and Rockoff 2014;

[1] As detailed below, court-ordered desegregation primarily followed the Civil Rights Act in 1964 and two additional cases–Green v. Kent in 1968 and Swann v. Charlotte-Mecklenburg in 1971. Brown v. Board and Brown II laid foundation for these later decisions, but did not, themselves, induce wide-scale school integration.

[2] Caetano and Maheshri (2017) find that demographic shocks explain only 60% of this change. Lutz (2011) provides causal evidence that the dissolution of a desegregation court order for a district increases segregation in that district.

Cook 2016). However, there is little evidence about the long term effects of either on final educational attainment, earnings, or neighborhood quality in adulthood (aside from Johnson 2015). Furthermore, the existing literature has primarily produced estimates of the **net** effect of desegregation by studying district-level changes induced by court orders (Guryan 2004; Reber 2010; Lutz 2011; Johnson 2015). Within-cohort evaluations of desegregation are particularly scant, making it difficult to understand the mechanisms through which school desegregation has positive effects. That is, are the positive effects due to changing peers, changing resources, or something else?

In this paper, I use within-cohort variation in busing assignment from a unique desegregation plan in Jefferson County, KY to estimate the long-run economic effect of busing. In 1975, the primarily white Jefferson County Public Schools (JCPS, "County" schools) district was ordered to integrate with the primarily black Louisville City Schools (LCS, "City" schools) district. To fix language, I will refer to the merged district as JCPS, "the merged district", or "the district", and I will refer to schools in the merged district that were in LCS prior to the merger as "former City" schools and schools in the merged district that were in JCPS prior to the merger as "former County" schools. The merged district is the union of the former County and the former City schools.

To achieve the target level of integration in each school, the merged district followed a busing plan designed by the federal district court judge. White students from minority-black, former County schools were taken by bus from their home school to a former City school, and black students from minority-white, former City schools were taken by bus from their home school to a former County school. Not all students were assigned busing in the same grades. For example, some white students were bused in 5th grade and 10th grade while others were bused in 3rd and 8th. Likewise, some black students were bused in 2nd grade through 9th grade while others were bused in 4th grade through 12th grade. The grades in which a student was assigned busing were based on the first initial of the student's last name and their race.

This conditionally random assignment procedure creates a series of natural experiments allowing me to study the causal effect of busing assignment among students within the same graduating cohort. I start by estimating the intent-to-treat effect of busing assignment relative to no busing assignment. This is possible because black students in 10th grade in 1975, for example, who have an assignment of busing in 2nd through 9th grade are not bused because they have completed those grades when busing begins. On the other hand, black students in 10th grade in 1975 who have an assignment of busing in 4th through 12th grade are bused because they have not completed 10th through 12th grade when busing begins. Furthermore, the busing assignment scheme induces random variation in the number of years assigned to busing and variation in the age at which a student is first assigned busing.[3]

To measure the long-run outcomes of students affected by this busing plan, I link confidential data from the Social Security Administration's Numident file on place of birth with confidential data from the 2000 Decennial Census and a special extract from the 2000 Decennial containing each individual's "alphabet group." I analyze characteristics of the individual's neighborhood in adulthood to study outcomes for all Census respondents as

[3]The intensive margin variation in busing assignment is, however, correlated with the age at which a student is first assigned busing, the year in which a student is first assigned busing, whether the busing assignment is disrupted (i.e. the student is assigned busing in one grade, assigned home school in a later grade, and then assigned busing again in yet a later grade), and other busing plan components.

opposed to the random sample surveyed for questions related to income.[4]

Using the linked data and variation in busing assignment induced by first letter of last name, I find that black students assigned busing to former County schools (formerly majority-white schools) live in higher quality neighborhoods as adults than black students assigned to remain in former City schools (formerly majority-black schools). This intent-to-treat effect increases in the number of years the student is assigned busing, and it is strongest for students assigned busing at earlier ages. By comparison, I find small to zero intent-to-treat effects of busing assignment on long-run neighborhood outcomes for white students. I find qualitatively similar effects on individual earnings, educational attainment, and employment using the random sample of respondents in the long-form Census.

These differences in outcomes emerge despite the fact that former City and County schools are both equally integrated after 1975. Historical data from the Office of Civil Rights show that former City schools were approximately 25.9% black from 1976-1982 while former County schools were approximately 23.5% black.[5] Despite the roughly equal integration, anecdotal accounts and empirical evidence suggest the schools were not equal with respect to staffing, facilities, neighborhood environment, other resources, and short-run student outcomes (e.g. dropout rates). Specifically, former City schools were worse along all these margins even after the district merger. Bused and not bused students were exposed to similar racial integration, but ultimately, attended different schools with

[4]In a recent paper about the long-run economic effects of Food Stamps, Bailey, Hoynes, Rossin-Slater, and Walker (2019) also analyze the characteristics of respondents' neighborhoods to leverage the full short-form Census data.

[5]Historical data from the Office of Civil Rights were collected, digitized, and provided by Ben Denckla and Sarah Reber here: https://web.archive.org/web/20150109135107/http://l1.ccpr.ucla.edu/OCR/ocr.htm

different resources. This suggests the long-run effects of busing assignment in this setting are due to improved school resources and not simply an effect of peer race.[6]

I then explore the treatment effect of busing take-up, moving beyond the intent-to-treat estimation discussed above. Estimating the effect of busing take-up versus remaining in the home school, however, is more complicated. It is possible that students do not comply with their busing assignment and instead drop out of school, move to another district, or transfer to a private school. In the publicly available 5% sample of the 1980 Decennial Census, 66.3% of white children aged 6-17 in Jefferson County are attending a public school. Over 93.3% of black children in Jefferson County are attending a public school in 1980. To produce a student-level measure of compliance, I use a novel dataset of archival yearbook and commencement program records from nearly twenty high schools in Jefferson County pre- and post-desegregation. I use this measure and another measure of compliance to scale the intent-to-treat estimates and estimate how large the effects might be after accounting for non-compliance. In future work, I will link this student-level measure of compliance directly to the 2000 Decennial Census and estimate the treatment effect of busing using an instrumental variables (IV) approach.

If non-compliance occurs equally for students assigned to busing and not assigned to busing, the IV regression will recover a local average treatment effect for those students who take up busing due to their assignment. However, the exclusion restriction is violated

[6] Although peer racial composition is similar in former County and former City schools, peer compositions may differ along other characteristics. Section 4.4.2 presents evidence that white students assigned busing to former City schools are more likely to leave the school district than white students assigned to remain in former County schools. If, for example, high income white students leave regardless of assignment but middle income white students leave only when assigned busing to a former City schools, then peer income composition will differ in former City and County schools. I discuss this in more detail in Section 4.4.4.

if assignment to busing affects non-compliance because in that case, busing assignment will affect long-run outcomes through a channel (e.g. drop-out or private school) other than actual busing. I use the yearbook and commencement program records to investigate this empirically. I find that compliance does not differ by busing assignment for black students, but that white students assigned to busing are less likely to comply with their assignment than white students assigned to remain in their home school. In Section 4.4.2, I discuss the implications of this and how it biases the estimate of the local average treatment effect for white students.

The results discussed above are based on within-cohort comparisons for cohorts in which all students are exposed to integration. Even so, I estimate similar effects to prior studies that focus on the net effect of desegregation. Johnson (2015) conducts a comprehensive study of the long-run net effect of desegregation. Using nationwide variation in desegregation court orders and data from the Panel Study of Income Dynamics, he finds that each year of exposure to school desegregation increases adult wage by 3.6 (se = 0.019) percent for black students. I estimate that assignment to busing increases the average income of a black student's neighborhood in adulthood by 3.4 (se = 0.016) percent. The effect I estimate in a setting where racial composition of schools is held constant is similar in magnitude and is not statistically different from the effect Johnson (2015) estimates in a setting where both school resources and racial composition vary. The contexts and outcomes are admittedly different, but this is suggestive evidence that the net effect of desegregation is driven by changes in school resources as opposed to changes in the racial composition of schools. In future work, I will test this directly in Jefferson County by comparing students graduating before versus after desegregation for those assigned versus not assigned busing.

This paper contributes to the literature on the effects of desegregation and re-segregation on student outcomes in two ways. First, I estimate long-run effects of school desegregation. Only one other paper estimates long-run effects and does so with a focus on district-level changes (Johnson 2015). Second, I estimate within-cohort effects of a school desegregation plan. Since district-level changes in segregation yield both dramatic changes in school resources and in racial integration, it is not possible to determine which change (or how much each change) affects outcomes. In Jefferson County, every student is exposed to racial integration after 1975, but there is still within-cohort variation in busing/school assignment. I isolate school resource effects by using this within-cohort variation in a setting where racial integration is held roughly constant. Other work using within-cohort variation in assignment focuses on re-segregating districts (Billings, Deming, and Rockoff 2014) or cross-district assignment lotteries (Angrist and Lang 2004; Cook 2016; Bergman 2018).

Ultimately, these results have important implications for the re-segregation of schools in the U.S.[7] I show that desegregation had long-run positive effects on economic outcomes of black students with no evidence of a strong negative effect for white students. Specifically, I find that busing leads black students to live in better neighborhoods many years

[7]These results also have implications for school desegregation in Jefferson County, KY. Jefferson County is one of the largest school districts in the U.S., and it has been the focus of multiple efforts to dismantle its current approach to desegregating schools, which still relies on busing students away from their neighborhood schools. For example, the Supreme Court ruled in 2007 that districts could not use race as the sole determinant for student assignment to schools (Meredith v. Jefferson County Board of Education; Parents Involved in Community Schools v. Seattle School District No. 1). And in 2017, Kentucky's state legislature took up a bill that would allow students across the state to attend the school nearest to their home (Emma Brown, "GOP bill could dismantle one of nation's most robust school desegregation efforts," The Washington Post, March 4, 2017, accessed July 15, 2019).

later. In addition, these results shed light on the current labor market situation of black and white individuals and potential intergenerational effects of desegregation. Recent research finds that neighborhood quality has long-run effects on intergenerational mobility, suggesting that desegregation may be an important channel for improving mobility of black children.

This paper also provides suggestive evidence that the gains from desegregation are primarily due to school resource effects with peer effects playing a smaller role. While this highlights the importance of equalizing school resources for black and white students, it also suggests that merging the funding of two disparate districts is not, by itself, a sufficient remedy to educational inequalities. Even after the districts merged, the former County schools produced better long-run outcomes for black students. This is likely the result of several lasting differences in the schools discussed in Section 4.2.3.

4.2 Institutional Details

4.2.1 Brief History of School Desegregation in the US

The Supreme Court ruled de jure segregation of public schools unconstitutional in 1954 (Brown v. Board of Education), and the Court handed enforcement of desegregation to district courts in 1955 (Brown II). Despite this enormous shift in policy, little changed in practice. School districts adopted "freedom of choice" plans and allowed voluntary transfers that technically complied with the law but limited its effectiveness (Cascio et al. 2008). In addition, existing residential segregation and white migration or "white flight" to suburban districts also diluted the impact of Brown and Brown II.

The 1964 Civil Rights Act (CRA) and the 1965 Elementary and Secondary Education Act (ESEA) made federal funding conditional on compliance with Brown, and as a result, school districts, especially those at risk of losing large grants, desegregated "just enough" to meet federal guidelines (Cascio et al. 2010). The CRA changed the legal environment in other ways, making it possible for the U.S. Attorney General to bring suits for plaintiffs in segregated local school districts (Johnson 2015). In 1968, the Supreme Court ruled in Green v. County School Board of New Kent County that the county's "freedom of choice" plan did not eliminate the "dual system" of separate black and white schools, and mandated the district adopt a new plan that would achieve actual integration. Finally, Swann v. Charlotte-Mecklenburg Board of Education in 1971 established that mandatory busing plans were a constitutional solution to desegregate districts that were segregated as a result of residential segregation. Of the 108 court-ordered desegregation plans documented by Welch and Light (1987), 106 were ordered after 1964, 101 were ordered after 1968, and 57 were ordered after 1971. Guryan (2004) and Reber (2005) show that these court orders increased integration, even in the presence of white flight.

The CRA, ESEA, and two critical Supreme Court decisions accelerated the process of school desegregation in the US. This process was somewhat stifled by a 1974 decision in Milliken v. Bradley that clarified schools could not be forced to desegregate across district lines unless it could be shown that the district lines were drawn with racist intentions. The Court's decision in this case meant white migration out of a desegregating district would prevent full integration because the district they migrated to could not be forced to integrate with the district they migrated from. In the early 1990s, three Supreme Court cases effectively ended court-ordered desegregation in the US. Lutz (2011) shows that when a district is released from their court order, school segregation and black dropout rates increase.

4.2.2 Busing and Desegregation in Jefferson County, KY

Like many cities in the U.S., Louisville, KY (and Jefferson County, KY) has a long history of residential and school segregation. The city charter in 1828 established public schools for white children, and in 1870, a charter established separate public schools for black children (JCPS 2019). In 1941, the Louisville City Schools (LCS) district had 57 white schools and 19 black schools. Shortly after the decision in Brown, LCS desegregated by re-drawing school attendance zones and allowing open enrollment in the high schools, subject to capacity constraints. However, the district gave students attending majority other-race schools under this plan the option to transfer to a majority same-race school. Teachers were integrated three years later in 1959. The transfer option and white migration to the Jefferson County Public Schools (JCPS) district curtailed full integration (K'Meyer 2013).

In 1972, the Kentucky Civil Liberties Union (KCLU), the local branch of National Association for the Advancement of Colored People (NAACP), and the Kentucky Commission on Human Rights (KCHR) filed a lawsuit asking the court to merge LCS, JCPS, and the small district of Anchorage to achieve de facto integration (K'Meyer 2013). Judge James Gordon initially rejected this proposal, but the Sixth Circuit Court of Appeals overturned his decision in 1973. Milliken v. Bradley put this plan in jeopardy by rejecting cross-district busing in Detroit. However, the Sixth Circuit ultimately decided the case of

LCS and JCPS qualified as an exception under Milliken, and in 1975, the districts merged (K'Meyer 2013). Judge James Gordon was tasked with enforcing the desegregation order, and due to his apparent hesitation, the Sixth Circuit in July 1975 suddenly ordered he develop a plan for the school year beginning in September 1975.

Judge Gordon's desegregation plan was unique among desegregation plans in the U.S. As was typical, it required each school consist of a certain percentage of black students (12 to 35 percent in this case). To achieve this, the plan adopted a traditional mandatory assignment that required the busing of black students from City schools that were formerly majority-black to County schools that were formerly majority-white and vice versa for white students. Students not assigned to busing would remain at their home school and have white or black students bused in. The busing assignment scheme, however, was not traditional (i.e. it was not based on zoning or grade restructuring).[8]

Under Judge Gordon's plan, students were quasi-randomly assigned to busing in a given grade based on the first initial of their last name and their race. For example, white students with the initials "A", "B", "F", or "Q" were assigned busing in 11th and 12th grade whereas white students with the initials "C", "P", "R", or "X" were assigned busing in 3rd and 8th grade. Black students were assigned busing in many more grades, but otherwise, were subject to the same initial-based assignment scheme. For example, black students with the initials "A", "B", "F", or "Q" were assigned busing in 2nd, 3rd, and 7th-12th grade whereas black students with the initials "C", "P", "R", or "X" were

[8] Welch and Light (1987) identifies grade restructuring as the primary method of desegregation for districts that desegregate by pairing formerly black and formerly white schools (the method used in Jefferson County, KY). They give an example of two schools that are grades K-6 but are racially segregated. A busing plan that desegregates these schools will typically convert one school to grades 1-3 and one school to grades 4-6 (leaving kindergarten unaffected). Under this type of plan, all students in a given cohort and school are treated with the same busing assignment.

assigned busing in 2nd-9th grade. Figure 4.1 shows this assignment plan as displayed in a July 1975 issue of The Courier-Journal.[9]

This assignment procedure generates within-cohort variation on the extensive margin (whether a student is assigned to busing), the intensive margin (the number of years a student is assigned to busing), and age of intervention (how early in childhood they are assigned busing). Note that different children are affected by the extensive versus intensive margin variation. I study the effect of these margins on long-run economic outcomes for black and white students.

This assignment plan was used for ten years, with only a minor change for white students in 1982. In 1985, the district shifted from initial-based assignment to a zoning system for junior high schools and high schools (K'Meyer 2013).[10] In 1991, the school district fully eliminated the initial-based system, moving elementary schools to a zoning system (K'Meyer 2013). In this paper, I focus on students in graduating cohorts from 1990 or earlier, meaning they are exposed to the pre-1985 system and aged 28 or older by the time of the 2000 Decennial Census.

4.2.3 Persistent Differences in Former County and Former City Schools

Prior to the merger in 1975, per pupil spending was slightly higher in the County than in the City, with JCPS spending approximately 10% more per student than LCS in 1972 (Census of Local Government Finances). The goal of the City-County merger was racial integration and equalization of school finances/resources. Anecdotally, however,

[9]Figure A4.1c shows the potential variation in busing assignment induced by this plan for a student attending the merged district as of 1975-76.

[10]See Figures A4.1a and A4.1b for the 1982 and 1985 changes as documented in The Courier-Journal.

the former City schools and former County schools remained different post-1975. For one, former County schools had long been the beneficiary of higher facilities spending. Spending on construction was 94% higher in JCPS than in LCS in 1972 (Census of Local Government Finances). Since facilities improvements, and other resources like textbooks, are a stock, former City and County schools could not immediately equalize on that margin. In fact, the district indicates in the 1975-76 School Superintendents Survey that no major capital spending occurred in that school year.[11]

Anecdotally, former City schools also had less involvement from Parent Teacher Associations (PTAs) than former County schools. An interviewee from Tracy K'Meyer's 2013 book on busing in Jefferson County states, "PTA was hard to come by. They didn't want to do anything, not in the city schools. The white kids were bused two years, so the parents weren't going to do anything in these black schools. They're just going to put in their time and then they'll go and work at their home schools."

After the merger, teachers were also assigned to schools in an effort to desegregate faculty. Despite this, archival yearbooks show that staffing was not equal in former City and County schools and that differences in staffing persisted into the 1980s. Ballard High School (a former County school) had nearly 20% more teachers with masters degrees in 1980 than Central High School (a former City school). This is consistent with Jackson (2009), suggesting that teacher labor supply responses resulted in lower teacher quality in former City schools.

I use historical school-level data from the Office of Civil Rights (OCR) from 1976,

[11] Cellini, Ferrara, & Rothstein (2010) and Goodman, Hurwitz, Park, & Smith (2018) both find school facilities are an important dimension of school quality.

1978, 1980, and 1982 to compare school characteristics and student outcomes in former City and County schools post-integration. Specifically, I estimate the following equation:

$$Y_{st} = \alpha + \beta FormerCity_s + X_s + Z_t + \varepsilon_{st} \tag{4.1}$$

where the dependent variable is a characteristic of school **s** in year **t** or a student outcome at school **s** in year **t** and $FormerCity_s$ is a dummy variable equal to one if school **s** was a City school prior to 1975 and equal to zero if school **s** was a County school prior to 1975. X_s is a set of fixed effects for the grades that are offered at school **s** (grades 1, 7, and/or 12) and Z_t is a set of year fixed effects.

First, despite the differences outlined above, former City schools and County schools have roughly equal racial composition of students post-1975. Column 1 of Table 4.1a shows that the percentage of black students is only about 2.3 percentage points higher at former City schools. Columns 2-4 show that gender composition and gender composition within race are also similar at former City and County schools. Column 5 uses a measure of classroom-level racial composition from the Office of Civil Rights surveys in 1976 and 1980. Specifically, the dependent variable is the standard deviation in the percent black in each classroom (of the 18 randomly surveyed classrooms from each school). Column 6 is the fraction of those classrooms in each school that had a particularly skewed racial composition (i.e. classroom percent black below 15% or above 35%). These results indicate that classrooms were also equally integrated at former City and former County schools.

Table 4.1b shows how former City and County schools differ in terms of student outcomes. These results are intended to be an indication of school quality, but admittedly, these outcomes are a function of many inputs, including student quality. Columns 1 and 7 show that former City schools have higher dropout rates and a higher rate of students referred to the courts for disciplinary action than former County schools. Former City schools did not have higher suspension rates; in fact, the coefficient suggests suspension rates were lower in these schools. Columns 2-3, 5-6, and 8-9 show the relevant outcomes (dropout, suspensions, court referral) by race. Former City schools perform especially poorly for black students in terms of dropouts and court referrals, yet they are also worse for white students. Since students are quasi-randomly assigned to former City and County schools, these effects can be attributed to the school as opposed to the student body, absent any major differences in compliance between students assigned to former City versus County schools.

Table 4.1c explores a few measures of school resources. Column 1 shows that former City schools are less likely to have a "Gifted and Talented" program for students. Column 2 indicates that there are no major differences in terms of whether these schools offer additional honors courses or other enrichment courses. Column 3 compares the student-teacher ratio at former City and former County schools. I use the number of classrooms as a proxy for the number of teachers in 1978 and 1980 because in those years, teacher data is not available and classroom data is. There are no statistically significant differences, but the coefficient implies the ratio is slightly higher in former City schools. Column 4 finds similar results using total number of teachers as the dependent variable and controlling for total number of students.

Finally, the neighborhoods where these schools are located also differ markedly. I show this in Table 4.1d. Columns 1-6 use publicly available tract-level data from the

1980 Decennial Census (obtained from NHGIS). Former city schools are located in tracts with lower rates of high school completion, higher poverty, lower employment, and lower median household income (see Columns 1-4). In addition, they are also located in tracts in which the buildings are less likely to have air conditioning and are more likely to be heated using a room heater as opposed to a central heating system (see Columns 5-6).

Column 7 uses data from the CDC's 2001-2005 prediction of daily, tract-level PM2.5 pollution. Unfortunately, pollutant data is not available at the tract level in earlier years. Nevertheless, these results show that former City schools are located in areas with higher predicted PM2.5 pollution, on average, from 2001-2005. Columns 8-10 use data from the Louisville Metropolitan Police Department on zip code level crime in 2004. Again, this the earliest year in which crime data is available at sub-county geographies. These results show that former City schools are located in zip codes with higher violent, property, and drugs/other crime.[12] Recent work finds that schools' neighborhood environments are an important input in the educational production function (e.g. Ebenstein, Lavy, and Roth 2016; Heissel, Persico, and Simon 2019).

Racial composition of schools is held roughly constant post-1975. School resources, broadly defined, likely remain different at former City and County schools. Data from several different sources and qualitative interviews suggest that former County schools had better facilities, more investment by Parent-Teacher Associations, more program offerings (like the Gifted and Talented program), higher quality teachers, were located in better neighborhoods, and ultimately, had better short-run student outcomes. Be-

[12] I limit the data to crimes occurring outside of summer months and in the hours from 6am-5pm to reflect the level of crime students would be potentially exposed to near school.

cause of this, it is reasonable to attribute any long-run differences in student outcomes to the effect of school resource differences and not an effect of school racial composition. This interpretation becomes more complicated in the presence of disenrollment responses and non-compliance. I discuss this in Section 4.3.2 and Section 4.4.2.

4.3 Data and Methodology

4.3.1 Data

2000 Decennial Census and the Numident

To estimate the effect of busing assignment on long-run outcomes, I link confidential data from the Social Security Administration's Numident file on place of birth with confidential data from the 2000 Decennial Census and a special extract containing each individual's alphabet group (the busing assignment group that they are in based on the first initial of their last name in the 2000 Census).[13]

The 2000 Decennial Census can be broken into two groups: short-form respondents and long-form respondents. The short-form data contain information on age, race, sex, household structure, and residence for almost all individuals in the U.S. in 2000. The long-form data contain more detailed information on income, educational attainment, and employment for a random sample of approximately 1 in 6 households. To take advantage

[13]The extract file is a subset of the 2000 Decennial Census that contains a unique identifier for each individual, binary variables indicating whether their last name begins with the letters: 'A,B,F,Q', 'G,H,L', 'C,P,R,X', 'M,O,T,U,V,Y', 'D,E,N,W,Z', and 'I,J,K,S' (based on the busing assignment schemes from 1975-1984), and binary variables indicating whether their last name begins with the letters: 'I,J,K,S,W,M', 'I,J,K,S,B,W', 'A,B,F,Q,H,C,O,U,V,Y,N,Z,X,E,L,R', 'G,H,L,C,P,D', 'T,D,P,G', and 'M,T,V,R,Z,X,F,A,O,U,Y,E,Q,N' (based on the busing assignment scheme from 1985-1990). I do not observe the individual's name or even the first initial of their last name. I can only access the unique identifier and these alphabet group indicator variables.

of the full sample from the short-form, I use the long-form with sampling weights to construct the following tract-level characteristics for every person in the short-form: average income, fraction of individuals with a high school degree, fraction of individuals with a bachelors degree, and fraction of individuals working in the last year or last week. Since women often change their last name at marriage, making matching problematic, I limit the main sample to men aged 28 and above, and as such, I calculate the tract-level statistics for men aged 28-55. Results are robust to various methods of calculating neighborhood characteristics.

Since I attach these tract-level characteristics to each individual in the short-form sample based on their reported tract, these tract-level characteristics are individual-level outcomes–they represent the quality of the neighborhood where the individual lives. I use the neighborhood quality results as the main results in this paper because they are estimated on the full short-form sample, maximizing statistical power. Results using individual income, education, and employment responses from the long-form sample are, in general, qualitatively similar. Table 4.2 shows summary statistics on men aged 28-55 and living in Jefferson County, KY from the publicly available sample of the 2000 Decennial Census.[14,15]

I use place of birth from the Numident as a proxy for childhood school district. This leads to some mismeasurement that will attenuate the results. Note, this measurement error differs from the issue of migration as an endogenous response to desegregation or busing assignment. For one, some students leave Jefferson County before 1975. Second,

[14] I produce summary statistics from publicly available samples to minimize disclosure risk.

[15] For comparison, Table A4.3 shows these statistics for women aged 28-55 and living in Jefferson County, KY.

some of those students in Jefferson County at school age will be attending private school prior to 1975. Third, some white students attending public school in Jefferson County will be attending LCS prior to 1975, and some black students attending public school in Jefferson County will be attending JCPS prior to 1975. In all of these cases, these students will not actually receive a busing assignment, but by treating county of birth as childhood school district, I will still code them as receiving an assignment. Table 4.3 presents migration and school attendance statistics by race for school-aged children from the publicly available 5% sample of the 1980 Decennial Census and from district-level enrollment counts (see also Figures 4.2a-4.2c). In Section 4.4.2, I use those statistics to adjust the intent-to-treat estimates for measurement error, and the results are roughly the same.

Finally, I use year of birth, month of birth, and school entry rules (from Bedard and Dhuey 2007) to define each individual's graduating cohort.[16] For the main analysis, I focus on individuals in graduating cohorts from 1965-1990.[17] Those students in graduating cohorts from 1965-1974 are not exposed to the desegregation program, allowing me to include race by alphabet group controls. Students in graduating cohorts from 1975-1990 are exposed to the pre-1985 system and are at least 28 years old in the 2000 Decennial Census. To summarize, the final sample includes men born in Jefferson County, KY who are in graduating cohorts from 1965-1990.

Archival Yearbook Records

[16]Bedard and Dhuey (2007) collect detailed information on school entry rules to estimate the effect of these rules on adult earnings.

[17]Note, I use the term "graduating cohort" to refer to the year the individual would have graduated from high school if they completed school with no grade retention. I do not require individuals in the sample to complete high school or to complete without grade retention.

I supplement the data above with data on student-level enrollment collected from archival high school yearbooks and commencement programs from Jefferson County, KY pre- and post-desegregation. To my knowledge, this is one of the first economics papers to use yearbooks as a source of student-level data.[18]

Student-level enrollment data have many benefits in this setting. First, I use the enrollment data to improve the measurement of who is exposed to the desegregation plan. In future work, I will link the data to birth and marriage indices obtained from the Kentucky Department of Libraries and Archives (KDLA) to measure busing assignment for eventually married women. Finally, the enrollment data allow me to observe actual take-up rather than busing assignment alone. By observing take-up, I can also evaluate how much take-up differs by race and for students assigned versus not assigned busing. Table 4.4a displays basic statistics about the yearbook data.

4.3.2 Methodology

Intent-to-Treat Estimates, Extensive and Intensive Margin Effects

For the main results in this paper, I estimate the intent-to-treat effect of assignment to busing by race. The identifying assumption is that, conditional on graduating cohort and race, busing assignment is exogenous to later-in-life outcomes. Since busing assignment is determined based on first initial of last name conditional on graduating cohort and race, this assumption is likely satisfied. One remaining concern is that students with certain initials perform better than students with other initials. For example, if students

[18] I began collecting yearbook data under this project, but have continued it for a subsequent project that I am working on jointly with E. Kaplan and J. Spenkuch investigating the effect of busing assignment in Jefferson County on long-run political attitudes of white students.

with initials closer to the beginning of the alphabet perform better than students with initials closer to the end, then the "A", "B", "F", "Q" group may be inherently different from the "C", "P", "R", "X" group. To account for this, I include students graduating before the desegregation plan and control for race by alphabet group fixed effects. Specifically, I estimate the following equation:

$$Outcome_{i,2000} = \alpha + \beta(BusAssign \times White)_{it} + \delta(BusAssign \times Black)_{it} + RY_{it} + RG_i + \varepsilon_{it} \tag{4.2}$$

where $Outcome_{i,2000}$ is the outcome variable measured in the year 2000 for individual **i**. $BusAssign_{it}$ is a dummy variable equal to one if individual **i** from graduating cohort **t** is assigned busing and equal to zero if not. $White_i$ is a dummy variable indicating individual **i** is white and $Black_i$ is a dummy variable indicating individual **i** is black. RY_{it} is a set of race by graduating cohort fixed effects, and RG_i is a set of race by alphabet group fixed effects. The outcome variables of interest for the short-form sample are: tract-level average income, fraction of people with a high school degree in the tract, fraction of people with a bachelors degree in the track, and fraction of people working last year or last week in the tract.[19,20] Again, I estimate (2) for men born in Jefferson County, KY and in graduating cohorts from 1965-1990. I find qualitatively similar results when estimating equation (4.2) for women who are single as of 2000.

Equation (4.2) estimates the extensive margin effect of busing assignment (not necessarily busing take-up). The assignment procedure yields extensive margin variation

[19] Again, these tract-level characteristics are based on men aged 28-55, but results are robust to using other samples.

[20] For the long-form sample, the outcomes of interest are: above/below median earnings, high school completion, bachelors degree completion, worked last year, and worked last week. I dichotimize earnings when using the smaller sample to increase statistical power.

when some students in a graduating cohort are completely past their assignment grades and some students are not. For black students, there is only extensive margin variation in the first few years after 1975 and it is present only for students in high school. The busing assignment plan also yields substantial intensive margin variation. I estimate the effect of each additional year an individual is assigned busing with the following model:

$$Outcome_{i,2000} = \alpha + \beta(YearsAssign \times White)_{it} + \delta(YearsAssign \times Black)_{it} + RY_{it} + RG_i + \varepsilon_{it} \tag{4.3}$$

I also estimate a model including both the assignment dummy variable and the years of assignment for black students. This yields the effect of an additional year of assignment conditional on assignment. I do not do this for white students because there is not meaningful variation in both assignment and years of assignment. The number of years a student is assigned busing is correlated with the age at which they are first bused and the year in which they are first bused. Also, students assigned busing for more years are more likely to have a disruption in their busing schedule, meaning that they are assigned busing for some years, assigned to remain at their home school for some later years, and then assigned busing again for even later years. When interpreting the results from equation (4.3), it is important to remember these possible patterns and the fact that the students affected by the extensive margin variation are much older than the students affected by the intensive margin variation.

Intent-to-Treat Estimates, Early vs. Late Childhood Effects

Finally, the assignment plan yields variation in the age at which an individual is first assigned busing. Prior work has found that neighborhood interventions occurring in early childhood are especially effective (Chetty and Hendren 2018). To test for age-of-assignment effects, I estimate the following equation:

$$\begin{aligned} Outcome_{i,2000} = \alpha &+ \beta_1 (BusAssign \times White)_{it} + \delta_1 (BusAssign \times Black)_{it} \\ &+ \beta_2 (GradeFirstAssign \times White)_{it} + \delta_2 (GradeFirstAssign \times Black)_{it} + RY_{it} + RG_i + \varepsilon_{it} \end{aligned} \tag{4.4}$$

where $GradeFirstAssign_{it}$ is a linear term equal to the grade in which an individual is first assigned busing (1st-12th grade) and equal to zero if individual **i** is not assigned busing. Conditional on race by graduating cohort fixed effects, the first grade in which an individual is assigned busing is perfectly collinear with the first year in which they are assigned busing and highly correlated with the number of years they are assigned busing. I compare the results from equation (4.4) to the results from the following equation:

$$\begin{aligned} Outcome_{i,2000} = \alpha &+ \beta_1 (BusAssign \times White)_{it} + \delta_1 (BusAssign \times Black)_{it} \\ &+ \beta_2 (GradeFirstAssign \times White)_{it} + \delta_2 (GradeFirstAssign \times Black)_{it} \\ &+ (R \times YearFirstAssign)_{it} + (R \times YearsAssign)_{it} + (R \times g(t))_{it} + RG_i + \varepsilon_{it} \end{aligned} \tag{4.5}$$

where $(R \times YearFirstAssign)_{it}$ is the interaction between the race dummy variables and a linear term in the first year an individual is assigned busing (centered at zero in 1974 and coded as zero for years earlier). $(R \times YearsAssign)_{it}$ is the interaction between the race dummy variables and a linear term in the number of years assigned. Finally, $(R \times g(t))_{it}$ is the interaction between the race dummy variables and a linear term in graduating cohort (coded as zero in 1965). I conduct several additional robustness tests that are described in Section 4.4.1.

LATE and the Exclusion Restriction

The prior section detailed the methodology for estimating intent-to-treat effects of busing assignment. In this section, I outline the methodology and assumptions necessary to estimate the effect of busing take-up. To estimate the local average treatment effect of busing take-up versus remaining in the home school, I employ a novel dataset of student enrollment records collected from archival yearbooks and commencement programs. As part of the busing plan, former City schools were paired with former County schools and students were bused within those pairs. For example, black students assigned busing from Central High School were assigned to one of seven County high schools that Central was paired with. Similarly, white students assigned busing from Ballard High School were assigned to Central High School.

In theory, I should be able to match a student from a pre-desegregation yearbook to a post-desegregation yearbook of the school (or schools) they would have been assigned to in a given year. However, school zone boundaries were partially re-drawn as part of the desegregation plan. For example, a black student in the Central High School zone may be re-drawn into the Louisville Male High School zone. If this occurs, I will not be able to find them in Central High School when they aren't assigned busing, and I will not be able to find them in one of Central's paired schools when they are assigned busing. As such, I measure take-up as whether I can match the student from a pre-desegregation yearbook to any post-desegregation yearbook of the system (former City or former County) that they were assigned to in a given year. In other words, if a white student in 10th grade in a 1974-75 County yearbook should be assigned busing in 1975-76, then take-up is counted if I match them to a 1975-76 yearbook for any former City school. This measurement is not perfect and is a lower bound of take-up because not all schools have yearbooks available in a given year. I also calculate adjusted measures of take-up to account for the fact that some yearbooks were not available for data collection and for the fact that year-to-year matches are low even in pre-desegregation years. Table 4.4b presents statistics on these measures of take-up.

I use the individual-level estimates of take-up from the yearbook records and a district-level measure of take-up from Cunningham, Husk, and Johnson (1978) to scale the intent-to-treat estimates. In future work, I plan to match the individual-level enrollment data to the 2000 Decennial Census and estimate the effect of busing by instrumenting for it with the student's busing assignment. If the assignment to busing versus not busing differentially affects enrollment in JCPS and if disenrollment affects outcomes differently than remaining in the home school, then the exclusion restriction is violated. I start with the assumption that assignment to busing does not influence enrollment in the district. Then, I explore the resulting biases from differential enrollment by assignment status. I estimate the following equations using two-stage least squares:

$$Bused_{it} = \alpha + \beta(BusAssign \times White)_{it} + \delta(BusAssign \times Black)_{it} + RY_{it} + RG_i + \varepsilon_{it} \tag{4.6}$$
$$Outcome_{i,2000} = \gamma + \phi(\widehat{Bused} \times White)_{it} + \lambda(\widehat{Bused} \times Black)_{it} + RY_{it} + RG_i + \omega_{it} \tag{4.7}$$

where $Bused_{it}$ is the dummy variable for busing take-up and $BusAssign_{it}$ is the dummy variable for busing assignment. Assuming busing assignment does not influence enrollment in the district, this yields the local average treatment effect for students who take up busing (due to their assignment) relative to those who remain in their home school (and have new students bused in).

If busing assignment does affect enrollment decisions, then the exclusion restriction is likely violated. To make the problem more clear, consider this context in the always-takers/never-takers/compliers/defiers framework (Angrist, Imbens, Rubin 1996). First, there are no always-takers or defiers in this setting. In other words, a student not assigned busing cannot choose to take up busing.[21] Students are only compliers (i.e. they take up busing when assigned it and they do not take up busing when not assigned it) or they are never-takers (i.e. they do not take up busing, regardless of assignment). Assuming that no one leaves the district only when they are not assigned busing and/or no one stays in the district only when they are assigned busing, then the never-takers can be divided into two groups: the always-leavers and the sometimes-leavers. Always-leavers disenroll from the district regardless of busing assignment. Sometimes-leavers disenroll from the district only when they are assigned busing. The presence of sometimes-leavers results in a violation of the exclusion restriction.[22]

How prevalent are sometimes-leavers in this setting? First, consider the problem for black students. Private school enrollment is low for school-aged black children in Jefferson County, KY. In addition, there is no evidence that desegregation plans increase private school enrollment or migration for black students. For these reasons, the relevant disenrollment margin is dropout. Although desegregation plans decrease black dropout rates (Guryan 2004), it is possible that assignment to busing decreases dropout more or less than assignment to remain in the home school. I do not find any evidence of this.

[21]This implies there are no always-takers because a student can't choose to take up busing regardless of assignment. It also implies no defiers because a student can't choose to take up busing only when they aren't assigned.

[22]This is likely an issue in many instrumental variable designs where the treatment is something individuals would like to avoid and there is a special action that can be taken to avoid treatment.

Student-level enrollment from yearbook data suggest black students assigned versus not assigned busing are equally likely to remain enrolled in JCPS. In this case, equation (4.7) yields the local average treatment effect of busing compared to remaining in the home school.

Now, consider the problem for white students. Private school enrollment is high for school-aged white children in Jefferson County, KY. Reber (2005) finds compelling evidence of white flight in response to desegregation plans. This suggests that movement to another district, enrollment in private school, and dropout are all relevant margins of disenrollment. In addition, I do find evidence that busing assignment affects disenrollment for white students. White students assigned to busing are less likely to be enrolled in JCPS in the next year than white students not assigned to busing. In this case, equation (4.7) yields biased estimates for white students because there is selection into disenrollment (i.e. some white students are "sometimes-leavers").

The bias is a function of the effect the instrument has on the outcome, independent of its effect on busing take up, and the strength of the first stage (Conley, Hansen, and Rossi 2012). In future work, I will conduct a bounding exercise on the effect for white students. Altonji, Elder, and Taber (2005) find that Catholic school attendance increases the probability of high school graduation by 0.05 to 0.08 and the probability of college enrollment by 0.02 to 0.15. I will, in future work, use those estimates, estimates of the returns to schooling, and measures of white flight from Cunningham et al. (1978) and the yearbook records to bound the effect of busing for white students.

Finally, these enrollment responses can affect the interpretation of the results in a peers versus resources framework. Although the former City school and former County schools are equally integrated, it is possible that the white students who leave when assigned to busing are different from the white students who leave when assigned to remain in their home school and have new peers bused in. If this is the case, then those black students who are assigned to remain in their home school are exposed to a different peer group than black students assigned to busing. I present some statistics comparing white students in former City and County schools and discuss this further in Section 4.4.4.

Decomposing the Net Effect

The methodology above estimates the effect of busing assignment or busing take-up for individuals within a cohort. Prior work finds that desegregation plans have a large positive net effect for black students. In other words, black students in a cohort exposed to desegregation have higher educational attainment and earnings than black students in a cohort graduating before desegregation occurs (Johnson 2015). In future work, I will estimate how much of this cross-cohort difference in Jefferson County can be explained by the within-cohort assignment differences. First, I will estimate:

$$\begin{aligned} Outcome_{i,2000} = \alpha + \beta(BusAssign \times White \times Exposed)_{it} + \delta(BusAssign \times Black \times Exposed)_{it} \\ + \eta(White \times Exposed)_{it} + \zeta(Black \times Exposed)_{it} + (R \times g(t))_{it} + RG_i + \varepsilon_{it} \end{aligned} \tag{4.8}$$

where $Exposed_{it}$ is equal to one if individual **i** is graduating in a year **t** after the desegregation plan in Jefferson County is implemented. η and ζ, then, identify the effect of desegregation exposure for students not assigned to busing, and β and δ identify the additional effect of exposure for students assigned to busing. Since I am estimating cross-cohort differences in one city, I cannot control for race by year fixed effects. Instead, I control for race interacted with a linear trend in the graduating cohort. In future work, I will also re-estimate equation (4.9) with other large counties included in the sample to control for cohort effects:

$$\begin{aligned} Outcome_{i,2000} = \alpha + \beta(BusAssign \times White \times Exposed)_{it} + \delta(BusAssign \times Black \times Exposed)_{it} \\ + \eta(White \times Exposed)_{it} + \zeta(Black \times Exposed)_{it} + RY_{it} + RG_i + \varepsilon_{it} \end{aligned} \tag{4.9}$$

4.4 Results

4.4.1 Long-run Effects of Busing Assignment

Extensive and Intensive Margin Effects

Using Numident data on place of birth and 2000 Decennial Census data on first initial of last name, I estimate the effect of busing assignment by race on long-run outcomes for men born in Jefferson County and in graduating cohorts from 1965-1990. Tables 4.5-4.7 display results for the effect of assignment on the quality of the tract where the student lives in adulthood. Tract-level characteristics are individual-level outcomes defined for the full short-form sample and are the following: tract-level average income, fraction of respondents in the tract with a high school degree, fraction with a bachelors degree, fraction who worked last year, and fraction who worked last week. These outcomes represent the quality of the neighborhood where the individual lives as an adult.

Column 1 of Table 4.5 shows that by adulthood, black students assigned to busing live in neighborhoods with average incomes approximately 3.4% higher than black students not assigned to busing (i.e. assigned to remain in their home school and have students bused in). Column 2 shows the average income result in levels (as opposed to logs in column 1). Columns 3 and 4 show that busing assignment causes black students to live in neighborhoods with more high school graduates ($\beta = 0.0055$, $se = 0.0036$) and in neighborhoods with more college graduates ($\beta = 0.0173$, $se = 0.0058$). Busing assignment does not lead black students to live in neighborhoods with higher employment (see columns 5 and 6).

The effects for white students (also in Table 4.5) are small and in no specification, statistically different from zero. For example, the coefficient in column 1 suggests that busing assignment leads white students to live in neighborhoods with average income about 0.15% lower than white students not assigned to busing. The 90% confidence interval for the estimate in column 1 is -0.58% to 0.29%. All of the estimated results in Table 4.5 are attenuated because some individuals born in Jefferson County have migrated out of the county before the desegregation plan is implemented, some are enrolled in private school in Jefferson County before the plan is implemented, and some are attending a school where they are in the minority-race. In Section 4.4.2, I scale the estimates in Table 4.5 to account for this attenuation.

Tables 4.6 and 4.7 show the effect of intensive margin increases in busing assignment. In Table 4.6, the busing assignment dummy variable is replaced with a linear term for years assigned busing. The results show that black students assigned more years of busing live in better neighborhoods as adults. As before, the effects for white students are small and are not statistically different from zero. Table 4.7 estimates the extensive margin and intensive margin effects simultaneously for black students. I only do this for black students because white students are only bused zero to two years whereas black students are bused zero to nine years. Table 4.7 suggests that the positive effects of busing are due to increases on the intensive margin and are not simply due to an extensive margin shock of busing assignment.

Early vs. Late Childhood Effects

I test for age-of-assignment effects by estimating how the first grade in which a student is assigned busing affects outcomes. Table 4.8 shows these results. Columns 1 and 3 show that, for black students, being first assigned busing in an earlier grade does lead to better outcomes (higher tract-level average income, higher fraction of college graduates in the tract) than first assignment in a later grade. Again, the estimates for white students are near zero.

The estimates in columns 1 and 3 are from a model that includes race by graduating cohort fixed effects because that is the level of randomization. However, when race by graduating cohort fixed effects are included, then the first grade in which a student is assigned busing is perfectly collinear with the first year in which they are assigned busing and highly correlated with their number of years assigned busing. I address this by estimating models that replace the race by cohort fixed effects with a race by cohort linear trend and include a linear term in first year assigned busing and a linear term in number of years assigned busing. These results, in columns 2 and 4, are similar to the results with race by cohort fixed effects, suggesting the results in columns 1 and 3 are due to an age of intervention effect rather than a year of intervention effect.

It is unlikely that these results are driven by differential compliance or measurement error by age. First, measurement error from defining the sample based on individuals born in Jefferson County should not differ much by age. From the 1980 Census, 86% of black children aged 6-8 and born in KY are still living in KY, and 81% of black children aged 15-17 and born in KY are still living in KY. This age profile (shown in Figure 4.2c) of the measurement error cannot explain the coefficient on grade first assigned busing. Similarly, enrollment in public school (shown in Figure 4.2b) does not appear to differ much by age. From ages 7-15, approximately 94-98% of black children in Jefferson County are attending a public school. Enrollment is lower for 6 year olds as well as 16 and 17 year olds, but the results are robust to dropping individuals first bused at these ages. These patterns also hold for white school-aged children from the 1980 Census.

Robustness Tests

I evaluate the robustness of the main extensive margin results by conducting placebo tests using similarly sized counties that are also under a court order for desegregation after 1968 (Welch and Light 1987). Specifically, I choose approximately 50 counties that are at least half the size of Jefferson County, KY. Then, I estimate the main extensive margin regression on each of those cities. Since the initial-based busing assignment plan is only used in Jefferson County, KY, then the "assigned busing" dummy variable should be near zero for all other cities. I conduct a similar placebo test in which I randomly assign each student to a new alphabet group. The main results are robust to these placebo tests.

Finally, the results are also robust to excluding the race-by-alphabet group fixed effects or including race-by-alphabet group fixed effects interacted with a linear trend in graduating cohort. This provides support for the identifying assumption that busing assignment (based on alphabet group) is exogenous to later-in-life economic outcomes.

4.4.2 Accounting for Enrollment Responses

Instrumental Variables Results

In this paper, the student-level enrollment data from yearbooks and the data from the U.S. Census Bureau are separated. In future work, I will link these two datasets. In this paper, I scale the estimates in Table 4.5 to account for the non-compliance and measurement error statistics presented in Table 4.3, Tables 4.4a-4.4b, and Figures 4.2a-4.2c. Table 4.9 shows these results.

In column 1 of Table 4.9, I display the main result from column 1 of Table 4.5. In column 2, I adjust for measurement error in busing assignment that arises from the fact that some students born in Jefferson County, KY are not living in Jefferson County, KY prior to the district merger. I quantify this error by measuring the fraction of students born in Kentucky who are still living in Kentucky as of 1980, by race. In column 3, I further adjust for measurement error that arises from the fact that some students living in Jefferson County, KY are already attending private schools prior to the district merger. I measure this for white students using the data from Cunningham, Husk, and Johnson (1978) on the fraction of school-aged birth cohort attending public schools in 1974.[23] For black students, I use the data from 1980 on the fraction of black school-aged residents attending public schools, assuming, since non-compliance is low for black students, that the 1980 levels are a good proxy for pre-1975 levels. Finally, in column 4, I further adjust for measurement error that arises from the fact that some students are already attending a majority other-race school prior to the district merger. I measure this directly using

[23]This is a good proxy–in 1977, 65% of the school-aged birth cohort are attending public schools and in 1980, 66% of actual residents are attending public schools.

aggregated Office of Civil Rights data available from the National Archives and Records Administration.

After adjusting for measurement error, the coefficient for black students increases to 0.0495, implying that black students assigned busing live in neighborhoods with average incomes that are about 5% higher than black students not assigned busing. On the other hand, the coefficient for white students remains small at -0.0028 even after these adjustments. In this table, I also (crudely) adjust the lower bound and the upper bound of the confidence intervals by the same measurement error statistics. This exercise yields a fairly precise adjusted confidence interval for white students of -0.0113 to 0.0056. Of course, this adjustment does not account for any error in the adjustment itself.

Columns 1-4 adjust estimates based on measurement error that arises from coding busing assignment based on place of birth. In columns 5-7 of Table 4.9, I adjust estimates to account for non-compliance. In future work, I will estimate the local average treatment effect directly. In this paper, this adjustment is intended to give a sense of how big the effect of busing take-up may be. I use three measures of non-compliance to scale the intent-to-treat estimates.

Students can exit the district by moving to another public school district, transferring to a private school, or dropping out of school. In 1980, five years after desegregation, 66.3% of white children and 93.3% of black children in Jefferson County are attending a public school.[24] The high public school enrollment of black children suggests there is not a substantial enrollment response to desegregation. In fact, dropout rates in LCS in 1974

[24] 29.4% of white children and 3.1% of black children are attending a private school. 4.3% of white children and 3.6% of black children are not enrolled in a public or private school. These statistics are calculated for individuals living in Jefferson County, KY in 1980 and aged 6-17.

are approximately 9.3% and fall to 4.0% in former LCS schools in 1976. Guryan (2004) also finds that desegregation decreases black dropout rates.

To measure white flight in response to the desegregation order, Cunningham, Husk, and Johnson (1978) calculate white public school enrollment as a share of the white school-aged birth cohort in Jefferson County, KY. In the 1974-75 school year, approximately 77.8% of the white school-aged birth cohort is attending public schools in Jefferson County. In the 1975-76 school year, this drops to 74.4%, and in 1976-1977, it drops further to 66.2%. Assuming the decline in public school enrollment is entirely due to desegregation, this implies approximately 15,000 or 15% of white students left the district because of the merger.[25] Approximately one third of this decline can be explained by rising private school enrollment, leaving the remaining amount to be explained by dropout or movement out of the district (Cunningham et al. 1978). Reber (2005) also finds white enrollment in a district decreases by about 10% within 2-3 years of desegregation. Finally, anecdotal accounts suggest parents believed that removing their children would lead to a policy reversal. Another interviewee from K'Meyer (2013) shares, "My parents thought if enough people stuck together and held their kids out of school then they would have no choice but to see, well, this isn't working and there's not enough jails to hold everybody who's not sending their kids to school."

I use the Cunningham, Husk, and Johnson (1978) estimate that 15% of white students left the district due to busing as one measure of non-compliance. When using this measure, I assume that 0% of black students left the district due to busing, based on the

[25]Total white school-aged births sums to 135,000 in the 1976-77 school year. Applying the 77.8% public school enrollment of 1974-75 to the 1976-77 birth cohort implies 105,000 white students should be enrolled. Only 90,000 white students are actually enrolled in 1976-77.

discussion of black public school enrollment above and the results from Guryan (2004). I also use two student-level measures calculated using the archival yearbook records. The first measure is the raw match rate from the pre-1975 yearbooks to the post-1975 yearbooks, by race. The second measure accounts for the fact that the data does not include every yearbook, and thus, some students cannot be matched to a post-1975 record.

Column 5 of Table 4.9 adjusts the estimate from Column 4 using the raw yearbook measure of compliance. This yields the largest coefficient estimate for white students, and it suggests that white students who are bused live in neighborhoods with average incomes that are 0.7% lower than white students who remain in former County schools. This estimate is an order of magnitude lower than the estimate for black students in this column. Since only some yearbooks were available for data collection, this column represents a lower-bound on compliance. I use an adjusted compliance measure in column 6, which suggests that white students who are bused live in neighborhoods with average incomes that are only 0.4% lower than white students who remain in former County schools. Finally, in column 7, I use the Cunningham, Husk, and Johnson (1978) measure of compliance and find similar results to column 6. Ultimately, these results suggest that even after accounting for measurement error and non-compliance, the effect of busing take-up for white students remains relatively small. This exercise, however, assumes that compliance is equal for groups assigned versus not assigned to busing. This is not the case for white students, and I discuss the implications of this in the following section.

The Exclusion Restriction

The IV results are only unbiased if the exclusion restriction is satisfied. In other words, it must be the case that busing assignment affects outcomes only through its effect on busing take-up. If busing assignment induces disenrollment and disenrollment affects outcomes, the exclusion restriction is violated. Note, however, the intent-to-treat effects in Tables 4.5-4.8 remain unbiased in this case. Table 4.4b shows that disenrollment does not differ by busing assignment for black students but it does differ for white students. In future work, I will consider the three margins of disenrollment (dropout, private school enrollment, and migration out of district) and existing estimates from the literature to put a bound on the local average treatment effect for white students.

4.4.3 Decomposing the Net Effect

Prior research on school desegregation has estimated large benefits for black students exposed to court-ordered desegregation. Many educational inputs change dramatically when school districts desegregate, and comparing students in districts before and after desegregation yields an estimate of the "net effect" of all of these changes. In this paper, I focus on within-cohort comparisons. However, in future work, I will also estimate the net effect in Jefferson County, KY and how much of that net effect is explained by within-cohort busing assignments.

4.4.4 Peers vs. Resources and Alternative Explanations

The results in Section 4.4.1 suggest that the gains from busing assignment are due to improved school resources for black students. This is because City and County schools are equally integrated after 1975, but in terms school quality, they are not equal. As described in Section 4.4.3, I will, in future work, test whether the net benefits of desegregation in Jefferson County, KY accrue primarily to students assigned busing. Comparing the magnitude of the estimates in Table 4.5 to existing estimates of the long-run net effect from Johnson (2015) does suggest that peer effects play a smaller role.

The results discussed in Section 4.4.2, however, present a challenge to this interpretation. I find that white students assigned to busing are less likely to comply with that assignment than white students assigned to remain in their home school (and have black students bused in). This differential compliance could result in a different composition of white students in former City versus former County schools. For example, consider a scenario in which all high-income families leave the district, regardless of assignment, but that middle-income families leave only when their child is assigned busing. This would mean that white students bused to former City schools would be from low-income families while white students remaining in former County schools would be from low- and middle-income families. These different peer compositions could explain the results for black students.[26]

To evaluate this possibility, I explore whether white students who are bused are observably different from white students who are not bused. In general, I find that white students who are bused are similar to white students who are not bused in terms of gender and standardized test scores measured one year after busing.[27] Columns 2-4 of Table 4.1a

[26] Another possibility is that all extremely intolerant families leave the district, regardless of assignment, but that moderately intolerant families leave only when their child is assigned busing. In this case, white students bused to former City schools would be particularly tolerant. However, this type of sorting should lead to better outcomes for black students in the former City schools because they would be exposed to peers that are, on average, more tolerant of racial integration.

[27] Test scores measured one year after busing could be affected by busing, potentially making them a poor measure of whether the bused/not bused groups differed before busing. First, there are no reports of test score by bused/not bused for pre-desegregation years. Second, the test scores in this report are measured after only one year of busing. Third, if busing did have a negative effect on white students, then the failure to find differences in the post-busing test scores would suggest that white students who were bused had higher pre-busing test scores than those not bused. In that case, we should not expect negative peer effects

shows that female students, in general and by race, are equally likely to take up busing. In other words, families did not disproportionately disenroll girls or boys in response to their busing assignment. In addition, former County schools with different characteristics have similar compliance rates.[28] In terms of test scores, a report on 2nd grade test scores in Jefferson County, KY (Natkin 1980) finds that white students who are bused score only 1.0 points lower than white students who remain in former County schools. This suggests that differential compliance based on busing assignment did not result in different peer compositions along these margins.[29]

Finally, there is reason to believe that, even if there were some peer quality differences, they may be muted in this setting. Carrell, Sacerdote, and West (2013) shows that students sort into sub-groups when exogenously assigned to a larger group of peers. And recent work using AddHealth data on peer groups finds that school desegregation does not increase interactions with other-race peers (Mele 2019). I present suggestive evidence that this occurs in the context of school desegregation in Jefferson County by using two measures of peer interaction in schools. First, I use historical data from the Office of Civil Rights on the racial composition of randomly selected classrooms at each school. Second, I collect data on racial composition of extracurricular activities and clubs from archival yearbooks.

---

for black students remaining in former City schools.

[28]Eastern High School has a 35 minute drive time to Central High School and Waggener has a 25 minute drive time, but they both have compliance rates of around 58%. At Eastern and Seneca, the student body is over 6% black pre-1975, but those schools have similar compliance rates for white students (~56-60%) as Waggener and Westport, where the student body is less than 1% black pre-1975. Jeffersontown has the highest student-teacher ratio at 25.66 and Atherton has the lowest at 23.06, but both schools have compliance rates around 40%.

[29]Data from the OCR on the percent of students in school on free or reduced price lunch also shows that, outside of a few elementary schools, former City and former County schools have an equal fraction of their students on free or reduced price lunch. I do not include these as main results because the sum of students on full price, reduced price, or free price lunch is often much lower than the sum of total students.

Figure 4.3 is a histogram showing the fraction of black students in each classroom and the fraction of black students in each school. Note, the fraction of black students at the classroom level is more disperse. In fact, almost half of all classrooms have a racial composition that is outside the 5th or 95th percentiles of the school-level distribution of percent black. In other words, half of all classrooms have a percentage of black students that is either below 15% or above 35% (the 5th and 95th percentiles). Even more, over one-third of all classrooms have compositions outside the 1st or 99th percentiles (i.e. a percentage of black students below 12.5% or above 42%). Finally, approximately 40% of all classrooms have a racial composition that is more than 50% different than their school's racial composition. These results suggest that even after segregation, black and white students had limited interaction at the classroom level.

This point is further highlighted in Figure 4.4, a histogram showing the fraction of black students in each club for the 1974-75 school year (pre-desegregation) and the same fraction for each club in the 1975-76 school year (post-desegregation). This data was collected from high school yearbooks for one City and three County high schools. Although extracurriculars become slightly more integrated after desegregation, they remain disproportionately segregated. This is suggestive evidence that black students and white students remained relatively segregated in terms of peer interaction even after the district merger. Since the evidence on peer differences suggests white students who are bused are similar to white students remaining in former County schools, and the evidence on peer interaction suggests that students were sorted into racially segregated sub-groups, it is unlikely that the main results are driven by peer effects.

4.5 Conclusion

In this paper, I study a unique busing assignment plan in Jefferson County, KY to estimate long-run economic effects of school desegregation. In 1975, the district (under a court order) assigned students to be bused to new schools (versus stay at their home school and have new students bused in) based on their race and the first initial of their last name. I find black students assigned busing to former County schools (formerly majority-white schools) live in better neighborhoods at adulthood than black students assigned to remain in former City schools (formerly majority-black schools). Black students in former County schools realize these gains despite continued segregation at the classroom level, limited interaction with other-race peers via extracurricular activities, and the fact that the former City and County schools are merged into one district after 1975.

This effect for black students is increasing in the total number of years a student is assigned busing and is larger for students assigned busing in earlier grades. On the other hand, busing assignment has small to zero effect on white students. Since former City and former County schools had similar racial compositions after desegregation, these results suggest school desegregation in this setting improved outcomes for black students by giving them access to better schools.

4.6 Tables and Figures

Table 4.1a: Differences in Student Composition between Former City and County Schools, Post-1975

	Percent Black	Percent Female	Percent Black Female	Percent White Female	Std. Dev. in Classroom Percent Black	Percent of Classrooms with Skewed Percent Black
	(1)	(2)	(3)	(4)	(5)	(6)
Former City School	0.0232	0.00292	0.00349	-0.00113	0.00839	0.0228
	(0.0176)	(0.00475)	(0.0113)	(0.00682)	(0.0130)	(0.0310)
Constant	0.251***	0.479***	0.477***	0.465***	0.164***	0.540***
	(0.0302)	(0.00723)	(0.0145)	(0.0195)	(0.0163)	(0.0514)
Years Included	'76, '78, '80, '82	'76, '78, '80, '82	'76	'76	'76, '80	'76, '80
Observations	382	382	102	102	191	191
R^2	0.152	0.045	0.038	0.071	0.098	0.084

Notes: Table 4.1a is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. Columns 1-4 display basic school demographics and indicate that racial composition, gender composition, and gender composition within race were roughly equal in former City and former County schools. Columns 5-6 use a measure of classroom-level racial composition. The OCR survey instructed schools to randomly select 18 classrooms in their school and provide data on the racial composition at the classroom level. I calculate the standard deviation in percent black at the classroom level for each school (column 5) and the percentage of classrooms that are below 15% black or above 35% black (the 5th and 95th percentiles of percent black at the school level). These results indicate that former City schools are also equally integrated at the classroom-level. All specifications include controls for grades offered at the school (1st grade, 7th grade, and 12th grade) and for year fixed effects (when applicable). Standard errors are clustered at the school-level.
* p < 0.1; ** p < 0.05; *** p < 0.01

Table 4.1b: Differences in Student Outcomes between Former City and County Schools, Post-1975

	Dropouts	Black Dropouts	White Dropouts	Suspensions	Black Suspensions	White Suspensions	Court Referrals	Black Court Referrals	White Court Referrals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Former City School	0.0100***	0.0179*	0.00862***	-0.00888	-0.0176	-0.00745	0.000602*	0.00166	0.000216
	(0.00339)	(0.00961)	(0.00322)	(0.00646)	(0.0118)	(0.00527)	(0.000338)	(0.00106)	(0.000178)
Constant	0.0123	0.0359	0.00712	0.0399**	0.0598**	0.0308**	-0.000142	-0.000902	-0.000128
	(0.00850)	(0.0239)	(0.00716)	(0.0154)	(0.0285)	(0.0119)	(0.000336)	(0.00117)	(0.000159)
Years Included	'76	'76	'76	'76, '78, '80, '82	'76, '78, '80, '82	'76, '78, '80, '82	'76	'76	'76
Observations	75	75	75	217	217	217	74	74	74
R^2	0.782	0.628	0.706	0.663	0.640	0.628	0.289	0.349	0.096

Notes: Table 4.1b is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. Columns 1-3 display dropout rates, in general and by race. Columns 4-6 display suspension rates, in general and by race. Columns 7-9 display court referral rates (i.e. the student is referred to court or juvenile authority for disciplinary action), in general and by race. The data reported in the OCR survey is based on the prior school year, but the measure of total students used in the denominator is from the current school year. For columns 4-6, I can interpolate the population between gap years and the results do not change. The OCR documentation also notes that data on suspensions was often prone to error because some schools reported the total number of suspensions and some reported the total number of unique students suspended. All specifications include controls for grades offered at the school (1st grade, 7th grade, and 12th grade) and for year fixed effects (when applicable). Standard errors are clustered at the school-level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.1c: Differences in Resources between Former City and County Schools, Post-1975

	Has Gifted & Talented Program	Has Additional Honors Courses	Student-Teacher Ratio	Total Teachers
	(1)	(2)	(3)	(4)
Former City School	-0.128**	0.0327	0.427	-2.422
	(0.0553)	(0.101)	(0.806)	(1.845)
Constant	-0.0432	0.577***	18.74***	16.65***
	(0.117)	(0.209)	(1.664)	(3.440)
Years Included	’76, ’78, ’80, ’82	’76	’76, ’78, ’80	’76, ’78, ’80
Observations	382	102	246	246
R^2	0.129	0.102	0.037	0.886

Notes: Table 4.1c is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. Column 1 displays the likelihood the school has a Gifted and Talented program–it is considerably lower in former City schools. Column 2 looks at the probability the school offers additional honors courses, and it does not differ by former City or County school. Column 3 displays the student-teacher ratio. This is constructed using data on total number of teachers for 1976 and total number of classrooms in 1978 and 1980. Column 4 displays the level instead of the ratio, including a control for total number of students in the school. All specifications include controls for grades offered at the school (1st grade, 7th grade, and 12th grade) and for year fixed effects (when applicable). Columns 1-2 and 4 include controls for the total number of students in the school. Standard errors are clustered at the school-level.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.1d: Differences in School Locations between Former City and County Schools, Post-1975

	Tract HS Completion Rate, 1980	Tract Below Poverty, 1980	Tract Employment Rate, 1980	Tract Median Household Income, 1980	Percent of Buildings in Tract with No AC, 1980	Percent of Buildings in Tract with Room Heater, 1980	Tract Predicted PM2.5 Levels, 2001-2005 Average	Zip Code Violent Crime per Capita, 2004	Zip Code Property Crime per Capita, 2004	Zip Code Drugs/Other Crime per Capita, 2004
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Former City School	-0.154***	0.168***	-0.0801***	-9,057***	0.302***	0.157***	0.590***	0.00633***	0.0273*	0.0325***
	(0.0410)	(0.0406)	(0.0203)	(1,371)	(0.0468)	(0.0453)	(0.0518)	(0.00132)	(0.0155)	(0.00969)
Constant	0.705***	0.0889***	0.912***	22,053***	0.0995**	0.0155	15.97***	0.00316***	0.00746	0.00750
	(0.0575)	(0.0337)	(0.0246)	(1,709)	(0.0441)	(0.0320)	(0.113)	(0.00111)	(0.00910)	(0.00658)
Observations	99	99	99	99	99	99	99	99	99	99
R^2	0.153	0.385	0.316	0.547	0.309	0.256	0.386	0.415	0.135	0.335

Notes: Table 4.1d is derived from historical Office of Civil Rights data, publicly available Decennial Census data from 1980 (from NHGIS), recent data on predicted pollution from the CDC, and recent data on zip code level crime from the Louisville Metropolitan Police Department (LMPD). In the OCR data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. I then geocode each school to a Census tract and match it to various tract-level or zip-level outcomes. Columns 1-6 display tract-level characteristics from the 1980 Decennial Census. Former City schools are located in neighborhoods with lower high school completion, higher poverty, lower employment rates, and lower median household income. Also, former City schools are located in areas where buildings are less likely to have air conditioning and are more likely to be heated by a room heater. Column 7 uses tract-level data on predicted PM2.5 pollution from the CDC (estimated using EPA data). I collapse this data to the average from 2001-2005. Former City schools are located in areas that have higher pollution levels as of 2001-2005 (earlier data not available). Finally, columns 8-10 use crime data from the LMPD from 2004 (earlier data not available). I collapse this to the zip code level, keeping only crimes that occur during school months and from the hours from 6am-5pm. I calculate crime rates using zip code level population. Former City schools are located in areas that have higher crime rates as of 2004. All specifications include controls for grades offered at the school (1st grade, 7th grade, and 12th grade). Standard errors are clustered at the school-level.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.2: Summary Statistics on Men Aged 28-55 in Jefferson County, KY from the Public Sample of the 2000 Decennial Census

	White Men	Black Men
	(1)	(2)
Income	48735.8	25931.1
	(53685.6)	(30515.1)
High School Degree	0.908	0.851
	(0.288)	(0.356)
Bachelors Degree	0.321	0.133
	(0.467)	(0.340)
Worked Last Year	0.922	0.781
	(0.269)	(0.414)
Born in KY	0.705	0.767
	(0.456)	(0.423)
Institutionalized	0.00386	0.0198
	(0.0620)	(0.139)
Observations	5,785	1,010

Notes: Table 4.2 is derived from the publicly available sample of the 2000 Decennial Census. I limit the sample to men aged 28-55 and living in Jefferson County, KY as of 2000. I report averages on income, educational attainment, employment, state of birth, and institutionalization by race. Standard deviations are in parentheses.

Table 4.3: Summary Statistics from the Public Sample of the 1980 Decennial Census, Aggregate Office of Civil Rights Data, and Cunningham et al. (1978)

	White, 6-17 y.o.	Black, 6-17 y.o.
Panel A. Pr(Living in KY | Born in KY), 1980		
Living in KY	0.8165	0.8319
	(0.3871)	(0.3741)
Sample	Born in KY	Born in KY
Observations	17,136	1,463
Source Data	1980 Census	1980 Census
Panel B. Pr(In Public School | In Jefferson County), 1980		
Enrolled, Public	0.6709	0.9408
	(0.4700)	(0.2361)
Sample	Living in Jefferson County	Living in Jefferson County
Observations	2,647	676
Source Data	1980 Census	1980 Census
Panel C. Pr(In Public School | In Jefferson County), 1974-75		
Enrolled, Public	0.7784	–
	–	–
Sample	School-aged birth cohort	–
Observations	–	–
Source Data	Cunningham et al. (1978)	–
Panel D. Pr(In Majority Same-Race School | In Public), 1974-75		
In Majority Same-Race	0.8045	0.8382
	–	–
Sample	In Public School in Jefferson County	In Public School in Jefferson County
Observations	–	–
Source Data	NARA/OCR Aggregates	NARA/OCR Aggregates

Notes: Panel A is derived from the publicly available sample of the 1980 Decennial Census. I limit the sample to boys aged 6-17 who were born in KY, and I report the probability that they are still living in KY, by race. Panel B is also derived from the publicly available sample of the 1980 Decennial Census. I limit the sample to boys aged 6-17 who are living in Jefferson County, KY, and I report the probability that they are enrolled in public school, by race. Panel C is derived from Cunningham et al. (1978)–see notes for Figure 4.2a. Panel D is derived from aggregated OCR data provided by NARA. I calculate the fraction of white students in a majority white school prior to desegregation (column 1) fraction of black students in a majority-black school prior to desegregation (column 2). I adjust the coefficients from Table 4.5 by these numbers in Table 4.9.

Table 4.4a: Summary Statistics for Student Characteristics from Yearbooks

Panel A.	City Schools, post-1975	County Schools, post-1975
Black	0.172	0.153
	(0.377)	(0.360)
Male	0.498	0.490
	(0.500)	(0.500)
Observations	1,240	13,355
Panel B.	**Black Students, 1975**	**White Students, 1975**
'A', 'B', 'F', 'Q'	0.191	0.168
'G', 'H', 'L'	0.154	0.173
'C', 'P', 'R', 'X'	0.163	0.174
'M', 'O', 'T', 'U', 'V', 'Y'	0.155	0.154
'D', 'E', 'N', 'W', 'Z'	0.162	0.159
'I', 'J', 'K', 'S'	0.174	0.171
Observations	1,090	20,887
Panel C.	**City Schools, 1972-74**	**County Schools, 1972-74**
Continually Enrolled –i.e. Matched to a Yearbook	0.586	0.649
	(0.493)	(0.477)
Observations	752	2,867

Notes: All statistics above are derived from newly collected data from archival school yearbooks. Panel A limits the sample to schools post-desegregation and reports statistics on student race and gender. This confirms the findings in Table 4.1a. Former City and former County schools are equally integrated. Panel B limits to students in the year prior to desegregation and shows the fraction of students in each alphabet group. Panel C uses data on Central High School and Ballard High School from 1972-74 to calculate year-to-year match rates in the years prior to desegregation. These columns show that even in years prior to desegregation, the match rates are low. I adjust the post-desegregation match rates by these numbers in Table 4.9.

Table 4.4b: Regression Results for Student-Level Enrollment and Busing Take-up from Yearbooks

	Bused	Bused	Continually Enrolled (Matched to a Yearbook)
	(1)	(2)	(3)
Assigned Busing x White	0.934***	0.489***	-0.143***
	(0.0125)	(0.0182)	(0.0195)
Assigned Busing x Black	0.976***	0.455***	-0.0389
	(0.0134)	(0.0307)	(0.0373)
Black	0.0122*	0.00568*	-0.201***
	(0.00700)	(0.00327)	(0.0228)
Constant	0	0	0.667***
	(7.13e-11)	(3.36e-10)	(0.00700)
Sample	Matched Students	Matched+ Unmatched	Matched+ Unmatched
Observations	3,756	6,090	6,090
R-squared	0.930	0.433	0.027

Notes: All statistics above are derived from newly collected data from archival school yearbooks. I limit the sample to schools in cluster 1 (discussed in text) because we have yearbooks for almost every school in that sample. I then attempt to match everyone in a pre-desegregation yearbook to an expected post-desegregation yearbook (former City or former County). I adjust the match rates in columns 2-3 since some yearbooks are missing even when limiting to the cluster with the most complete set. Column 1 limits the sample to students who I match to a post-desegregation yearbook. This column indicates that among students who remain enrolled in the merged district, they almost always comply with their busing assignment. Column 2 runs the same regression as Column 1 but includes students who disenroll (i.e. students for whom I do not find a yearbook match). This columns indicates that approximately half of students assigned busing comply with their busing assignment. Finally, Column 3 regresses whether I find a yearbook match at all on whether the student was assigned busing. This column indicates that white students assigned busing are more likely to leave the district, but this is not the case for black students. Robust standard errors are in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.5: Extensive Margin Effect of Busing Assignment on Neighborhood Characteristics in Adulthood

	Log(Avg. Income)	Avg. Income	% with HS	% with BA	% Worked Last Year	% Worked Last Week
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned Busing x Black	0.0336**	1,890***	0.00551	0.0173***	0.00187	-0.00307
	(0.0156)	(565)	(0.00360)	(0.00580)	(0.00458)	(0.00527)
Assigned Busing x White	-0.00145	-168	0.000160	-0.00128	0.000283	0.0000260
	(0.00264)	(171)	(0.000782)	(0.00168)	(0.000419)	(0.000545)
P-value, AB x B = AB x W	0.0273	0.000561	0.148	0.00227	0.730	0.560
Observations	149,000	149,000	149,000	149,000	149,000	149,000
R^2	0.0863	0.0457	0.0344	0.0257	0.131	0.161

Notes: Standard errors are clustered at the level of variation, race by grade cohort by alphabet group. The sample in all specifications is men born in Jefferson County, KY (based on the Numident) and graduating in years 1965-1990 (based on year of birth, month of birth, and school entry rules) who respond to the short-form Census in 2000. Assigned Busing x Black (AB x B) is a dummy variable equal to one if the respondent reports a non-white race and if the respondent is coded as "assigned busing" based on grade cohort, race, and alphabet group. It is equal to zero if the respondent reports race as "white" or is coded as "not assigned busing" based on grade cohort, race, and alphabet group. Assigned Busing x White (AB x W) is defined similarly. All specifications include grade cohort fixed effects interacted with race fixed effects and alphabet group fixed effects interacted with race fixed effects. The dependent variables are continuous tract-level averages derived from the 2000 Decennial long-form data using sample weights. Log(avg. income) is the natural log of the average income in the tract, avg. income is the level of the average income in the tract, % with HS is the fraction of respondents in the tract with a high school diploma or more, % with BA is the fraction of respondents in the tract with a bachelors degree or more, % worked last year is the fraction of respondents in the tract who worked last year, and % worked last week is the fraction of respondents in the tract who worked last week. These averages are based on men aged 28-55 as of the 2000 Census. For tracts with a small number of respondents, the county-level average is used. I attach these tract-level outcomes to short-form respondents to use the largest sample possible. Essentially, this measures whether busing leads people to live in "better" neighborhoods later in life (i.e. neighborhoods with higher average income, higher education levels, higher levels of employment in the prior year and prior week). The regressions themselves are not weighted.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.6: Intensive Margin Effect of Busing Assignment on Neighborhood Characteristics in Adulthood

	Log(Avg. Income)	Avg. Income	% with HS	% with BA	% Worked Last Year	% Worked Last Week
	(1)	(2)	(3)	(4)	(5)	(6)
Years Assigned x Black	0.00848**	290*	0.00311***	0.00335**	0.00150*	0.000949
	(0.00393)	(167)	(0.000917)	(0.00162)	(0.000885)	(0.00117)
Years Assigned x White	-0.000601	-79.9	0.000136	-0.000323	0.000324	0.000210
	(0.00133)	(87.0)	(0.000413)	(0.000872)	(0.000212)	(0.000295)
P-value, YA x B = YA x W	0.0294	0.0507	0.00336	0.0462	0.198	0.541
Observations	149,000	149,000	149,000	149,000	149,000	149,000
R^2	0.0863	0.0457	0.0344	0.0257	0.131	0.161

Notes: Standard errors are clustered at the level of variation, race by grade cohort by alphabet group. The sample in all specifications is men born in Jefferson County, KY (based on the Numident) and graduating in years 1965-1990 (based on year of birth, month of birth, and school entry rules) who respond to the short-form Census in 2000. Years Assigned x Black (YA x B) is a continuous variable equal to the number of years the respondents is coded as "assigned busing" based on grade cohort, race, and alphabet group if the respondent reports as non-white race and equal to zero if the respondent reports race as "white." Years Assigned x White (YA x W) is defined similarly. All specifications include grade cohort fixed effects interacted with race fixed effects and alphabet group fixed effects interacted with race fixed effects. The dependent variables are continuous tract-level averages derived from the 2000 Decennial long-form data using sample weights. Log(avg. income) is the natural log of the average income in the tract, avg. income is the level of the average income in the tract, % with HS is the fraction of respondents in the tract with a high school diploma or more, % with BA is the fraction of respondents in the tract with a bachelors degree or more, % worked last year is the fraction of respondents in the tract who worked last year, and % worked last week is the fraction of respondents in the tract who worked last week. These averages are based on men aged 28-55 as of the 2000 Census. For tracts with a small number of respondents, the county-level average is used. I attach these tract-level outcomes to short-form respondents to use the largest sample possible. Essentially, this measures whether busing leads people to live in "better" neighborhoods later in life (i.e. neighborhoods with higher average income, higher education levels, higher levels of employment in the prior year and prior week). The regressions themselves are not weighted.

* p < 0.1; ** p < 0.05; *** p < 0.01

Table 4.7: Extensive and Intensive Margin Effect of Busing Assignment on Neighborhood Characteristics in Adulthood, Black Students

	Log(Avg. Income)	Avg. Income	% with HS	% with BA	% Worked Last Year	% Worked Last Week
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned Busing x Black	0.0143	1,310*	-0.00238	0.0102	-0.00202	-0.00603
	(0.0180)	(700)	(0.00410)	(0.00690)	(0.00520)	(0.00603)
Years Assigned x Black	0.00787*	234	0.00321***	0.00292*	0.00158*	0.00121
	(0.00422)	(180)	(0.000993)	(0.00173)	(0.000946)	(0.00126)
Observations	149,000	149,000	149,000	149,000	149,000	149,000
R^2	0.0863	0.0457	0.0344	0.0257	0.131	0.161

Notes: Standard errors are clustered at the level of variation, race by grade cohort by alphabet group. The sample in all specifications is men born in Jefferson County, KY (based on the Numident) and graduating in years 1965-1990 (based on year of birth, month of birth, and school entry rules) who respond to the short-form Census in 2000. See Table 4.5 notes for a definition of Assigned Busing x Black. See Table 4.6 notes for a definition of Years Assigned x Black. All specifications include grade cohort fixed effects interacted with race fixed effects and alphabet group fixed effects interacted with race fixed effects. Also, the specifications above include an Assigned Busing x White dummy variable that is not reported because white students do not have meaningful variation in number of years assigned busing (they are only bused for 0-2 years). The dependent variables are continuous tract-level averages derived from the 2000 Decennial long-form data using sample weights. Log(avg. income) is the natural log of the average income in the tract, avg. income is the level of the average income in the tract, % with HS is the fraction of respondents in the tract with a high school diploma or more, % with BA is the fraction of respondents in the tract with a bachelors degree or more, % worked last year is the fraction of respondents in the tract who worked last year, and % worked last week is the fraction of respondents in the tract who worked last week. These averages are based on men aged 28-55 as of the 2000 Census. For tracts with a small number of respondents, the county-level average is used. I attach these tract-level outcomes to short-form respondents to use the largest sample possible. Essentially, this measures whether busing leads people to live in "better" neighborhoods later in life (i.e. neighborhoods with higher average income, higher education levels, higher levels of employment in the prior year and prior week). The regressions themselves are not weighted.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.8: Grade-of-Assignment Effects on Neighborhood Characteristics in Adulthood

	Log(Avg. Income)	Log(Avg. Income)	% with BA	% with BA
	(1)	(2)	(3)	(4)
Grade First Assigned x Black	-0.00746	-0.0111***	-0.00331*	-0.00406***
	(0.00547)	(0.00282)	(0.00195)	(0.00117)
Grade First Assigned x White	0.000207	-0.000519	0.0000712	-0.00163***
	(0.000705)	(0.000579)	(0.000420)	(0.000386)
Assigned Busing x Black	0.109*	0.112***	0.0508**	0.0418**
	(0.0579)	(0.0399)	(0.0206)	(0.0165)
Assigned Busing x White	-0.00294	0.00505	-0.00179	0.0101*
	(0.00601)	(0.00816)	(0.00361)	(0.00563)
Year Fixed Effects	YES	–	YES	–
Year Trend + Controls	–	YES	–	YES
Observations	149,000	149,000	149,000	149,000
R^2	0.0863	0.0862	0.0257	0.0255

Notes: Standard errors are clustered at the level of variation, race by grade cohort by alphabet group. The sample in all specifications is men born in Jefferson County, KY (based on the Numident) and graduating in years 1965-1990 (based on year of birth, month of birth, and school entry rules) who respond to the short-form Census in 2000. Grade First Assigned x Black is a continuous variable equal to the first grade in which a respondent is coded as assigned busing (1st grade through 11th grade) based on grade cohort, race, and alphabet group if the respondent reports a non-white race and equal to zero if the respondent reports race as “white.” Grade First Assigned x White is defined similarly. See Table 4.5 notes for definitions of Assigned Busing x Black and Assigned Busing x White. All specifications alphabet group fixed effects interacted with race fixed effects. The columns with year fixed effects include grade cohort fixed effects interacted with race fixed effects. The columns with a year trend and controls include linear trends in grade cohort, number of years assigned busing, and year first assigned busing interacted with race fixed effects. The dependent variables are continuous tract-level averages derived from the 2000 Decennial long-form data using sample weights. Log(avg. income) is the natural log of the average income in the tract, and % with BA is the fraction of respondents in the tract with a bachelors degree or more. These averages are based on men aged 28-55 as of the 2000 Census. For tracts with a small number of respondents, the county-level average is used. I attach these tract-level outcomes to short-form respondents to use the largest sample possible. Essentially, this measures whether busing leads people to live in “better” neighborhoods later in life (i.e. neighborhoods with higher average income, higher education levels, higher levels of employment in the prior year and prior week). The regressions themselves are not weighted.

* p < 0.1; ** p < 0.05; *** p < 0.01

Table 4.9: Effects Adjusted to Account for Measurement Error and Non-compliance

	Log(Avg. Income), Measurement Error Adjustments				Log(Avg. Income), Take-up Adjustments		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Assigned Busing x Black							
Coefficient Estimate	0.0336	0.0404	0.0429	0.0512	0.1126	0.0660	0.0512
Lower Bound of Confidence Interval	0.0080	0.0096	0.0102	0.0122	0.0269	0.0157	0.0122
Upper Bound of Confidence Interval	0.0592	0.0711	0.0756	0.0902	0.1983	0.1162	0.0902
Assigned Busing x White							
Coefficient Estimate	-0.0015	-0.0018	-0.0023	-0.0028	-0.0062	-0.0037	-0.0033
Lower Bound of Confidence Interval	-0.0058	-0.0071	-0.0091	-0.0113	-0.0248	-0.0146	-0.0133
Upper Bound of Confidence Interval	0.0029	0.0035	0.0045	0.0056	0.0124	0.0073	0.0066
Adjustment Factor	Nothing	Pr(In KY | Born in KY)	Pr(In Public School, pre-1975 | In Jeff. Co.)	Pr(In JCPS/LCS | White/Black & In JCPS or LCS)	Pr(Take-up, Yearbooks)	Pr(Take-up, Yearbooks w/ Adjustment)	Pr(Take-up, Cunningham et al. White Flight)
Black Students	–	0.8352	0.9689	0.8382	0.4250	0.7083	–
White Students	–	0.8149	0.7780	0.8045	0.4894	0.7529	0.85
Observations	149,000	149,000	149,000	149,000	149,000	149,000	149,000

Notes: See Table 4.5 for general notes about how the coefficients in column 1 are estimated. In column 2, I adjust the estimate in column 1 to account for measurement error induced by students migrating out of Jefferson County, KY prior to the desegregation order. The statistic used for this adjustment is from Table 4.3, Panel A. In column 3, I further adjust the estimate in column 2 to account for measurement error induced by students attending private school in Jefferson County, KY prior to the desegregation order. The statistic used for this adjustment is from Table 4.3, Panels B & C. In column 4, I further adjust the estimate in column 3 to account for measurement error induced by students attending majority-other race schools prior to the desegregation order. The statistic used for this adjustment is from Table 4.3, Panel D. Column 5 adjusts the estimate in column 4 by a lower-bound measure of busing take-up (from Table 4.4b, column 2). Column 6 uses a scaled measure of take-up that accounts for the fact that yearbook-to-yearbook match rates are low even in pre-desegregation years (from Table 4.4a, Panel C). Finally, column 6 uses a measure of take-up derived from Cunningham et al. (1978). Note, the adjustments in columns 5-7 are not cumulative (the adjustments in columns 1-4 are). Note, this table also includes a crude adjustment of the lower and upper bounds of the confidence intervals for these coefficients. This is not ideal and does not account for error in the adjustment factors. Nevertheless, it gives some sense of how these intervals change.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Figure 4.1: Busing Assignment Plan in Jefferson County, KY

How to tell when your child will be bused . . . unless

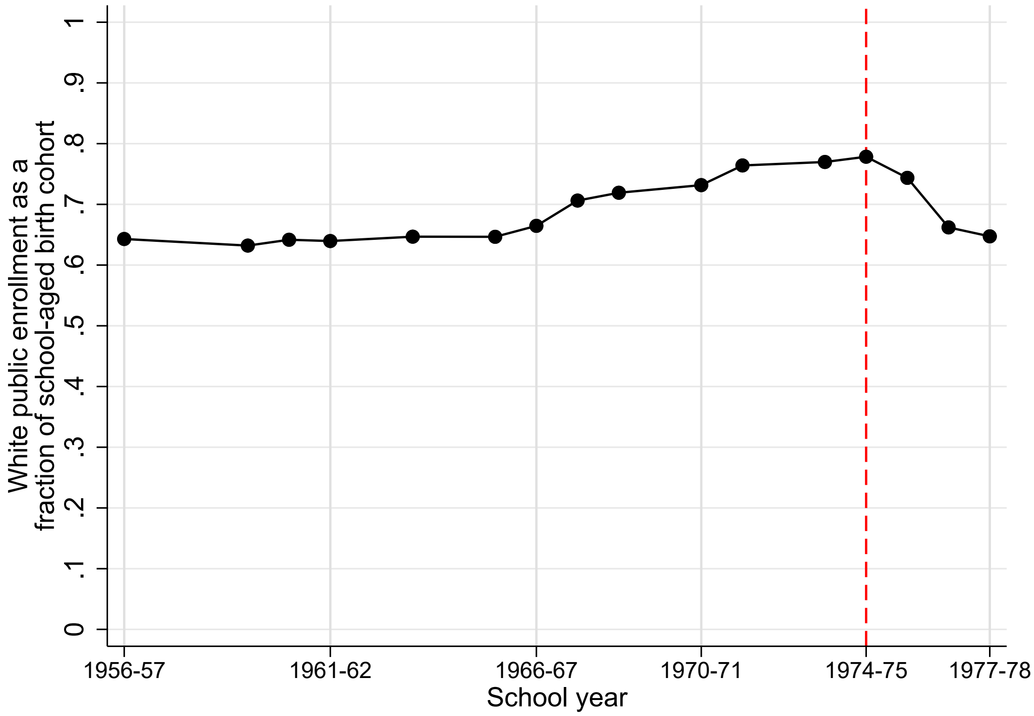
If child's last name begins with letters:	White child will be bused in grades:	Black child will be bused in grades:
A, B, F, Q	11, 12	2, 3, 7, 8, 9, 10, 11, 12
G, H, L	2, 7	2, 3, 7, 8, 9, 10, 11, 12
C, P, R, X	3, 8	2, 3, 4, 5, 6, 7, 8, 9
M, O, T, U, V, Y	4, 9	2, 3, 4, 5, 6, 10, 11, 12
D, E, N, W, Z	5, 10	4, 5, 6, 7, 8, 9, 10, 11, 12
I, J, K, S	6	4, 5, 6, 7, 8, 9, 10, 11, 12

Exempted students:
- Kindergarten students
- First graders during the fall quarter *
- Students who will be seniors this year, however, in subsequent years seniors will participate in the plan
- Students in special schools, primarily for the emotionally or physically handicapped
- Students attending schools exempted under the plan

* In grade one no child will be bused during the fall quarter; after that, entire classes will be bused with their teachers on a schedule to be determined later.

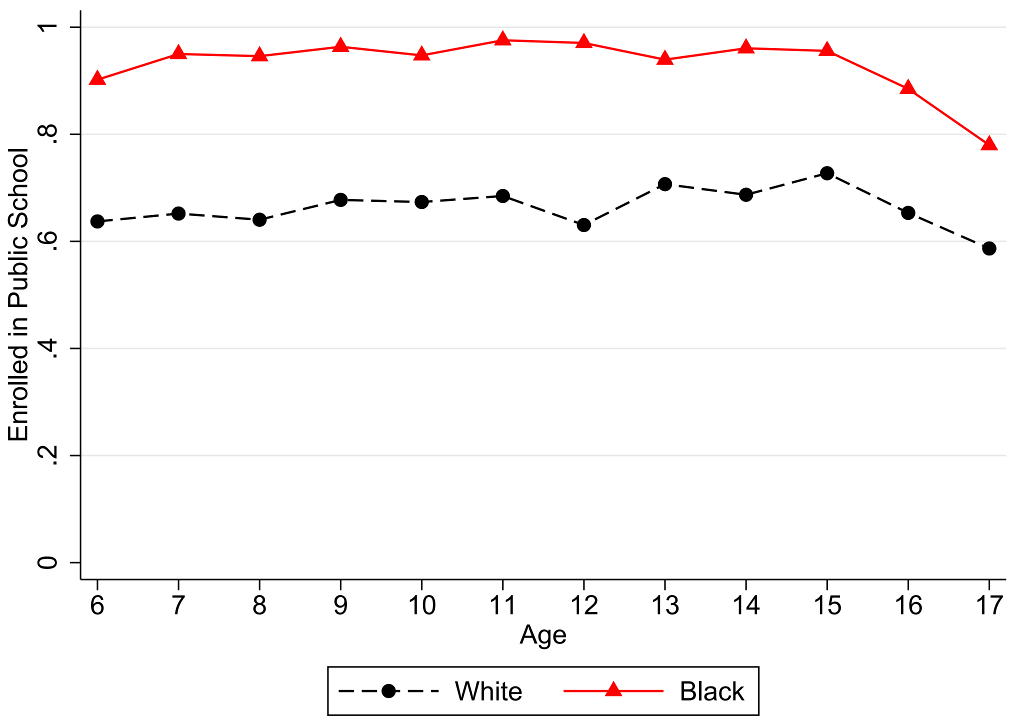
Notes: The plan depicted above was printed in the July 31, 1975 issue of The Courier-Journal. This plan, known as 'the alphabet plan', was unchanged from 1975-1982. A minor change was made for white students in 1982. In 1985, the district adopted a zoning system for middle and high school students, abandoning the alphabet plan for those students. In 1991, the district moved to a zoning system for elementary school students. The 1982 plan and the 1985 plan are displayed in Figures A4.1a and A4.1b. Figure A4.1c depicts the potential variation in busing assignment induced by this plan for a student in school by the 1975-76 school year.

Figure 4.2a: White Public School Enrollment as a Share of White School-aged Birth Cohorts



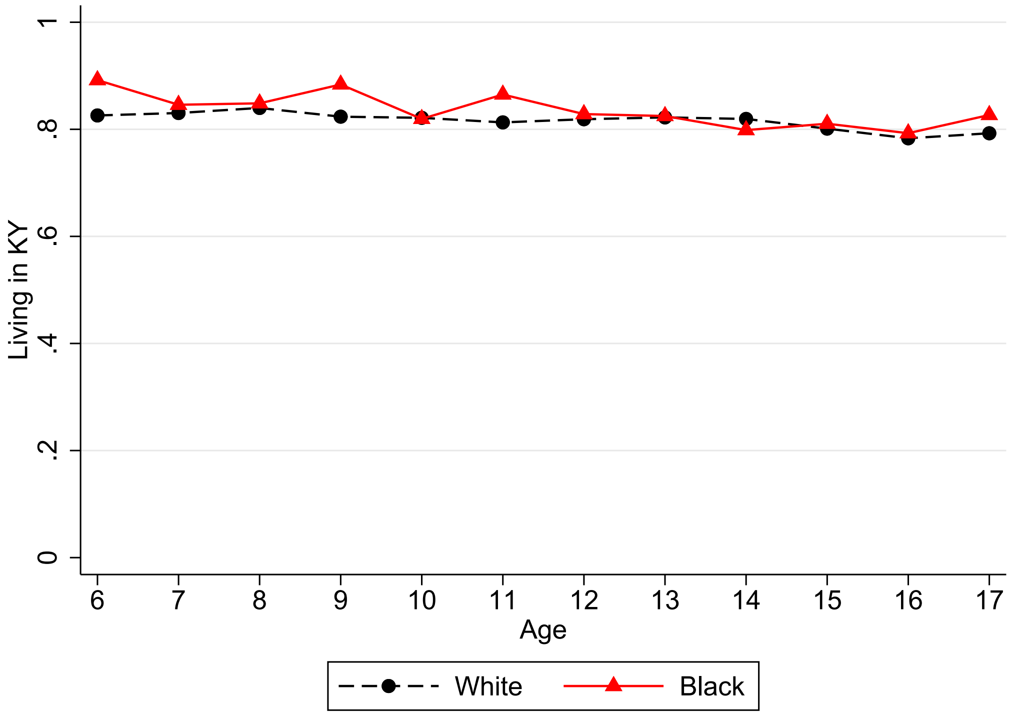
Notes: Figure 4.2a is derived from Cunningham et al. (1978). The authors measure white public school enrollment in Jefferson County, KY for several years from 1956 to 1977. They also calculate the school-aged birth cohort, which is the number of people who should be school aged in Jefferson County, KY based on birth records alone. In their paper, they argue that white flight is not as stark as it seems. Part of the decline in the level of white enrollment is due to a decline in the school-aged birth cohort. I digitized their figure, calculated the share directly, and I plot that share here. White public school enrollment falls after the desegregation order in 1975.

Figure 4.2b: Public School Enrollment in 1980, by Race and Age



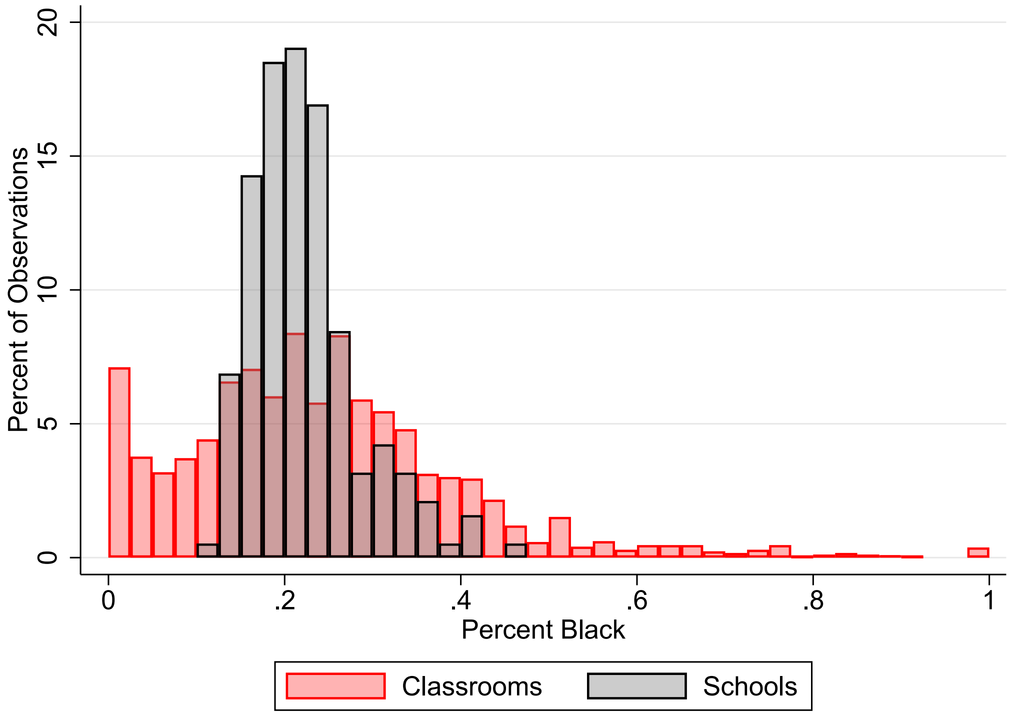
Notes: Figure 4.2b is calculated from publicly available Decennial Census data from 1980. In this figure, I plot the fraction of children enrolled in public school in Jefferson County, KY in 1980 for ages 6-17 and by race. Enrollment in public school is fairly constant from ages 6-15 but falls at ages 16 and 17. This is due to a rise in non-enrollment. Figures A4.3a and A4.3b show private school enrollment and non-enrollment by age and race.

Figure 4.2c: Fraction Born in KY Who Are Still Living in KY in 1980, by Race and Age



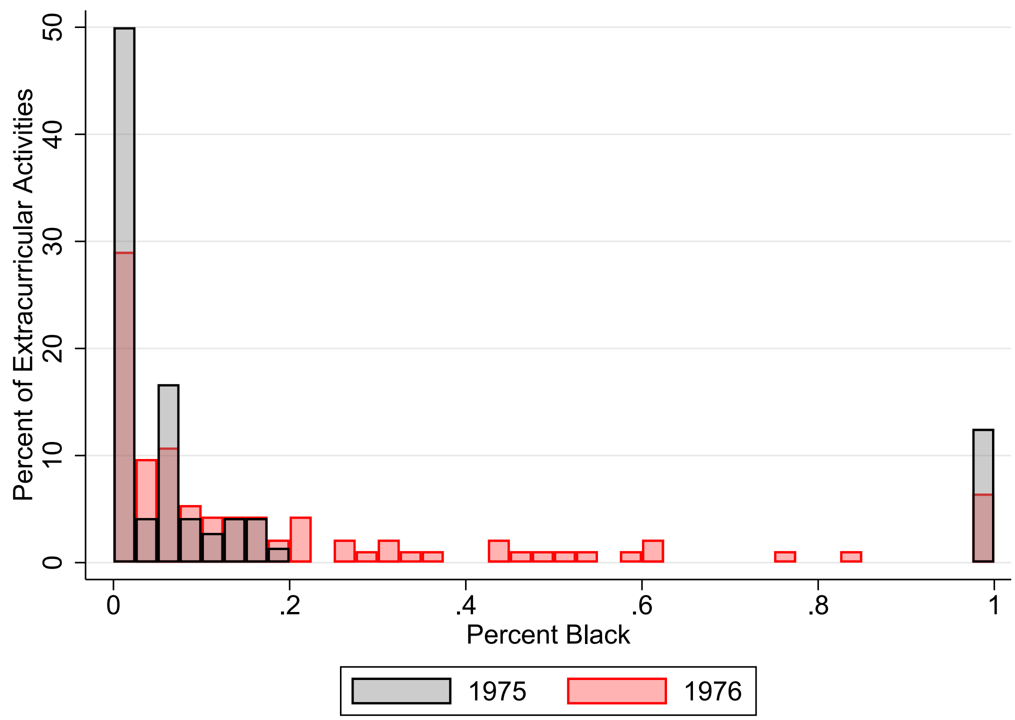
Notes: Figure 4.2c is calculated from publicly available Decennial Census data from 1980. In this figure, I plot the fraction of children born in KY who are still living in KY as of 1980 for ages 6-17 and by race. The probability of migrating out of the state is constant from ages 6-17. I use this as a proxy for the migration rate out of Jefferson County, KY to approximate the attenuation bias in the coefficients from Table 4.5 (see Table 4.3, Table 4.5, and Table 4.9).

Figure 4.3: Classroom Racial Composition versus School Racial Composition



Notes: Figure 4.3 is calculated from historical Office of Civil Rights data. As part of the OCR survey, schools were instructed to randomly select 18 classrooms in their school and provide information about the racial composition of the classroom. I have coded this information from the files for 1976 and 1980. The figure above plots a histogram of the percent black in each school over a histogram of the percent black in each classroom. The percent black at the classroom-level is more disperse–almost half of all classrooms are below 15% black or above 35% black (the 5th nd 95th percentiles of percent black at the school-level).

Figure 4.4: Racial Composition of Extracurricular Activities, Pre- and Post-Desegregation



Notes: Figure 4.4 is calculated using newly collected data from archival yearbooks in Jefferson County, KY. A research assistant collected data on the racial composition of extracurricular activities from one city high school (Central High School) and three county high schools (Ballard High School, Eastern High School, and Atherton High School). I plot the percent black at the club-level for 1975, the year before desegregation, and for 1976, the year after desegregation. Although clubs become slightly more integrated in 1976, they are far more segregated than the student bodies themselves.

4.7 Appendix A. Additional Tables and Figure

Table A4.1a: Differences in Racial Composition between Former City and Former County Schools, post-1975

	Percent Black								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Former City School	0.0265	0.0306*	0.0174	0.0184	0.00598	0.0356	0.0254	0.00772	
	(0.0189)	(0.0170)	(0.0174)	(0.0189)	(0.0197)	(0.0216)	(0.0205)	(0.0279)	
Former City x 1976									0.0160
									(0.0180)
Former City x 1978									0.00835
									(0.0221)
Former City x 1980									0.0357
									(0.0217)
Former City x 1982									0.0354
									(0.0219)
Constant	0.271***	0.227***	0.267***	0.258***	0.277***	0.335***	0.321***	0.203***	0.254***
	(0.0320)	(0.0250)	(0.0318)	(0.0324)	(0.0327)	(0.0381)	(0.0379)	(9.31e-10)	(0.0298)
Years Included	’78, ’80, ’82	’76, ’80’, ’82	’76, ’78, ’82	’76, ’78, ’80	’76, ’78, ’80, ’82	’76, ’78, ’80, ’82	’76, ’78, ’80, ’82	’76	’76, ’78, ’80, ’82
Excluding Schools	-	-	-	-	Grade=1	Grade=7	Grade=12	Grade=1, Grade=7	-
Observations	280	282	293	291	130	300	312	18	382
R-squared	0.105	0.167	0.158	0.162	0.239	0.196	0.185	0.066	0.156

Notes: Table A4.1a is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. Columns 1-8 examine racial composition at former City (vs. former County) schools, making various retrictions outline above. Column 9 estimates the relationship by year.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A4.1b: Differences in Free/Reduced Lunch between Former City and Former County Schools, post-1975

	Percent Free/Reduced Price Lunch		
	(1)	(2)	(3)
Former City School	0.0345	-0.00567	0.123***
	(0.0260)	(0.0231)	(0.0376)
Constant	0.310***	0.331***	0.265***
	(0.0327)	(0.0293)	(0.0460)
Observations	95	37	100
R-squared	0.358	0.670	0.389

Notes: Table A4.1b is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. The table above examines the percent of students with free or reduced price lunch at former City (vs. former County) schools. In column 1, I remove 5 elementary schools with outlier rates of free or reduced price lunch (above 50%). Column 2 removes all elementary schools from the sample. Column 3 shows the regression with no restrictions. One caveat with this data is that the total number of children receiving lunch (full price + reduced price + free) is often much lower than the total number of children at the school.

Table A4.1c: Differences in Racial Composition between Former City and Former County Schools, post-1975

	Gifted and Talented Program	Total Teachers	Student-Teacher Ratio	Suspension Rate
	(1)	(2)	(3)	(4)
Former City School x 1976	-0.0916**	-2.965**	0.495	0.00312
	(0.0406)	(1.202)	(0.811)	(0.0112)
Former City School x 1978	-0.125*	-7.398	0.160	0.00922
	(0.0639)	(5.101)	(2.438)	(0.00959)
Former City School x 1980	-0.121*	0.626	-0.566	-0.0299**
	(0.0725)	(2.807)	(1.763)	(0.0119)
Former City School x 1982	-0.176*			-0.0224***
	(0.106)			(0.00820)
Constant	-0.0551	17.49***	16.73***	0.0386**
	(0.117)	(3.674)	(3.028)	(0.0166)
Observations	382	242	242	382
R-squared	0.130	0.859	0.049	0.675

Notes: Table A4.1c is derived from historical Office of Civil Rights data. In this data, I classify schools as former City schools if they were in the Louisville City Schools district and not in the Jefferson County Schools district from 1968-1974. In the 1976-82 data, I remove schools with low student populations (less than 200) and with an abnormally high percentage of black students (above 50%). On inspection, these are primarily non-traditional schools, such as vocational schools. The table above examines the presence of gifted and talented programs, student-teacher ratio, total teachers, and suspension rates over time.

Table A4.2: Distribution of Surnames by First Initial, from Data on Top 100 Surnames

Alphabet Group	Percentage of Surnames
'A', 'B', 'F', 'Q'	0.165
'G', 'H', 'L'	0.171
'C', 'P', 'R', 'X'	0.171
'M', 'O', 'T', 'U', 'V', 'Y'	0.160
'D', 'E', 'N', 'W', 'Z'	0.157
'I', 'J', 'K', 'S'	0.175

Notes: The statistics above are calculated using publicly available data on the top 1,000 surnames in the United States. Surnames are removed if the percentage of individuals who are black or white is less than 75 percent.

Table A4.3: Summary Statistics on Women Aged 28-55 in Jefferson County, KY from the Public Sample of the 2000 Decennial Census

	White Women	Black Women
	(1)	(2)
Income	25448.8	22043.4
	(26890.1)	(25672.2)
High School Degree	0.924	0.886
	(0.265)	(0.318)
Bachelors Degree	0.314	0.128
	(0.464)	(0.334)
Worked Last Year	0.826	0.801
	(0.379)	(0.399)
Born in KY	0.699	0.766
	(0.459)	(0.423)
Institutionalized	0.00186	0.00379
	(0.0430)	(0.0615)
Observations	6,139	1,296

Notes: Table A4.3 is derived from the publicly available sample of the 2000 Decennial Census. I limit the sample to women aged 28-55 and living in Jefferson County, KY as of 2000. I report averages on income, educational attainment, employment, state of birth, and institutionalization by race. Standard deviations are in parentheses.

Figure A4.1a: Busing Assignment Plan Change in Jefferson County, KY–1982

First letter of last name	This year	1983-84 school yr	1984-85 school yr
I, J, K, S	1,6	1	1
G, H, L	2,3,7	2,3,7	2,3
C, P, R, X	8	6,8	6,7,8
M, O, T, U, V, Y	4,5,9	4,5,9	4,5
D, E, N, W, Z	10	10	9,10
A, B, F, Q	11,12	11,12	11,12

Notes: The plan depicted above was printed in an issue of The Courier-Journal. The original alphabet plan is displayed in Figure 4.1a. This plan details a change that was made for white students in 1982. In 1985, the district adopted a zoning system for middle and high school students, abandoning the alphabet plan for those students. In 1991, the district moved to a zoning system for elementary school students.

Figure A4.1b: Busing Assignment Plan in Jefferson County, KY–1985

Alphabet groupings

Within each cluster, black and white elementary school students will be bused according to the first letters of their last names. Here are the alphabet groups:

Blacks

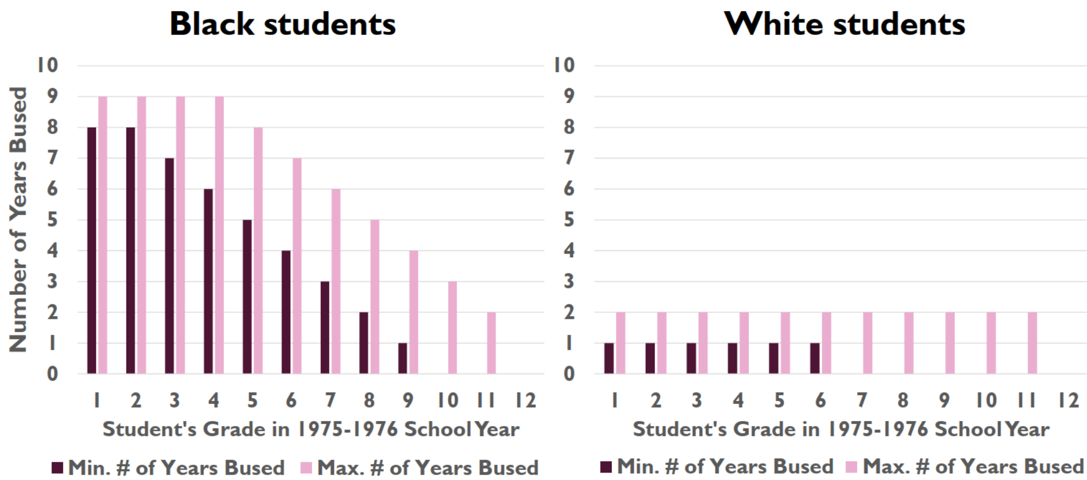
Last name begins with	Grades Bused
I, J, K, S, W, M	4, 5
A, B, F, Q, H, C, O, U, V, Y, N, Z, X, E, L, R	1, 2, 3
T, D, P, G	1, 4, 5

Whites

Last name begins with	Grades Bused
I, J, K, S, B, W	1
G, H, L, C, P, D	2, 3
M, T, V, R, Z, X, F, A, O, U, Y, E, Q, N	4, 5

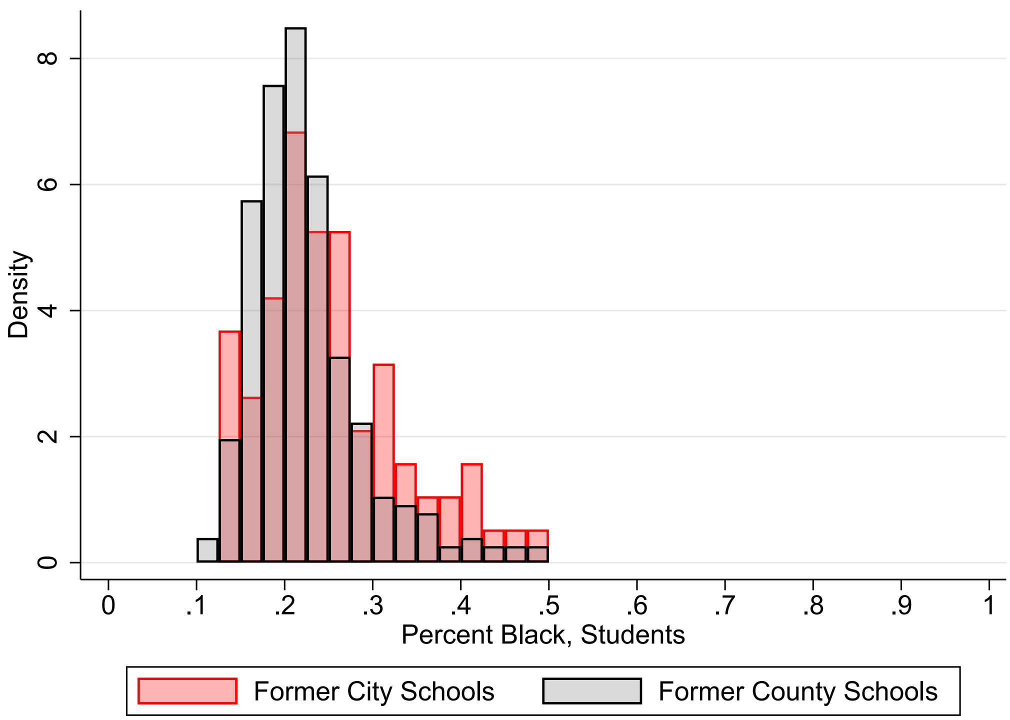
Notes: The plan depicted above was printed in an issue of The Courier-Journal. The original alphabet plan is displayed in Figure 4.1a. In 1985, the district adopted a zoning system for middle and high school students, abandoning the alphabet plan. The plan for elementary school students in 1985 is displayed above. In 1991, the district moved to a zoning system for elementary school students.

Figure A4.1c: Potential Variation in Busing Assignment for Student in 1975-76



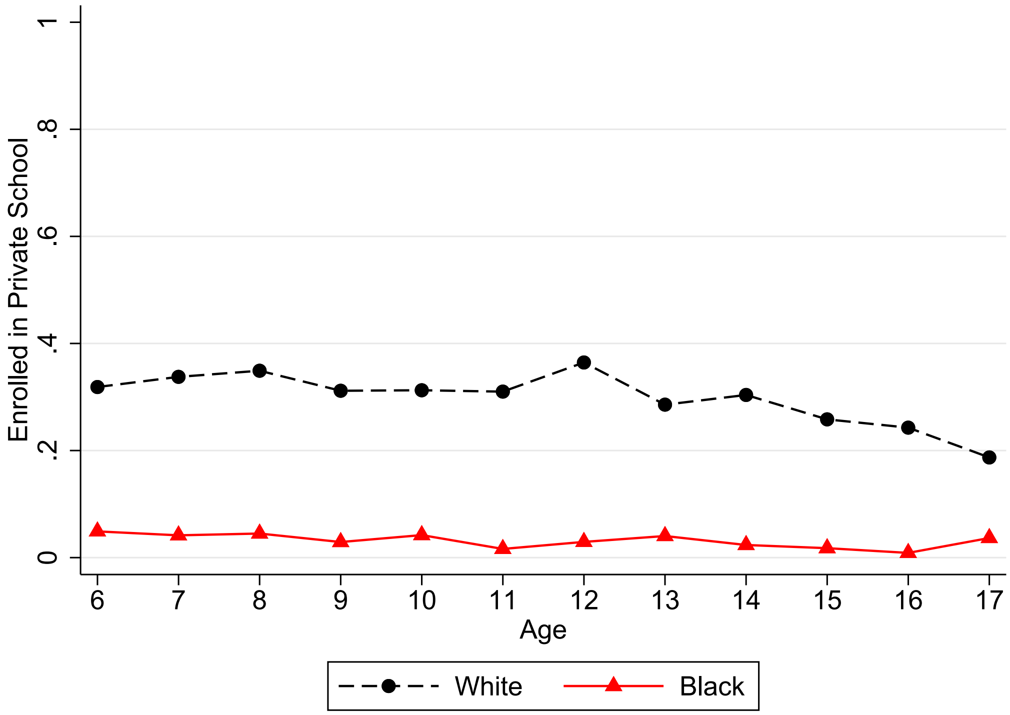
Notes: The figure above displays the potential variation in busing assignment for students attending the merged district in 1975-76. The mininum number of years assigned and the maximum number of years assigned by race are based on the plan displayed in Figure 4.1a.

Figure A4.2: Histogram of Percent Black in School, 1976-1982



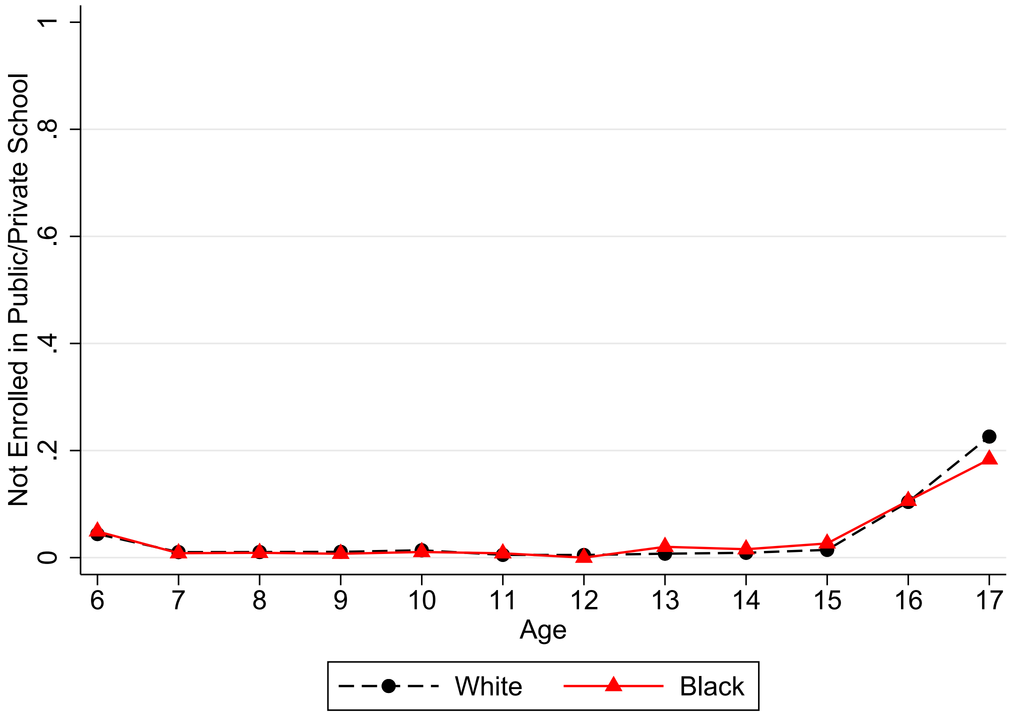
Notes: Figure A4.2 is calculated from historical Office of Civil Rights data. The figure above plots a histogram of the percent black in each school for former City schools (red) and former County schools (gray).

Figure A4.3a: Private School Enrollment in 1980, by Race and Age



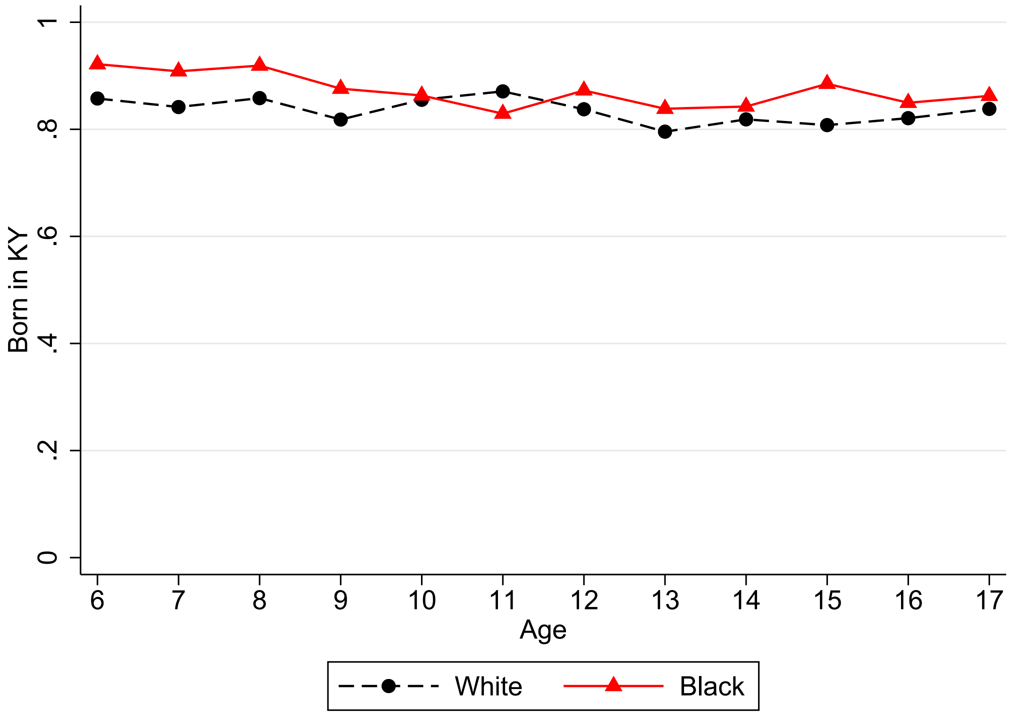
Notes: Figure A4.3a is calculated from publicly available Decennial Census data from 1980. In this figure, I plot the fraction of children enrolled in private school in Jefferson County, KY in 1980 for ages 6-17 and by race. Enrollment in private school is fairly constant from ages 6-15 but falls at ages 16 and 17. This is due to a rise in non-enrollment.

Figure A4.3b: No School Enrollment in 1980, by Race and Age



Notes: Figure A4.3b is calculated from publicly available Decennial Census data from 1980. In this figure, I plot the fraction of children not enrolled in school in Jefferson County, KY in 1980 for ages 6-17 and by race. Non-enrollment is fairly constant from ages 6-15 but increases at ages 16 and 17.

Figure A4.3c: Fraction Living in KY Who Are Born in KY in 1980, by Race and Age



Notes: Figure A4.3c is calculated from publicly available Decennial Census data from 1980. In this figure, I plot the fraction of children living in KY as of 1980 who were born in KY for ages 6-17 and by race.

Bibliography

[1] Agan, Amanda, Matthew Freedman, and Emily Owens. 2018. "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense." NBER working paper.

[2] Agan, Amanda, and Sonja B. Starr. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *The Quarterly Journal of Economics* 133 (1): 191-235.

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

[4] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–184.

[5] Almond, Douglas, Hilary Hoynes, and Diane Whitmore Schanzenbach. 2010. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *Review of Economics and Statistics* 93 (2): 387–403.

[6] Anbarci, Nejat, and Jungmin Lee. 2014. "Detecting Racial Bias in Speed Discounting: Evidence from Speeding Tickets in Boston." *International Review of Law and Economics* 38: 11–24.

[7] Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–455.

[8] Angrist, Joshua D., and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review* 94 (5): 1613–1634.

[9] Antonovics, Kate L., and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91 (1): 163–177.

[10] Anwar, Shamena, and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1): 127–151.

[11] Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics* 127 (2): 1017–1055.

[12] Arnold, David, Will Dobbie, and Crystal S. Yang. 2017. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics*.

[13] Arora, Ashna. 2018. "Too Tough on Crime? The Impact of Prosecutor Politics on Incarceration." Working paper.

[14] Baker, Scott, and Claudio Mezzetti. 2011. "Prosecutorial Resources, Plea Bargaining, and the Decision to Go to Trial." *The Journal of Law, Economics, and Organization* 17 (1): 149-167.

[15] Bailey, Martha J., Hilary Hoynes, Maya Rossin-Slater, and Reed Walker. 2019. "Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program." Working paper.

[16] Bala, Nila. 2015. "Judicial Fact-Finding in the Wake of *Alleyne*." *New York University Review of Law & Social Change* 39 (1): 1–44.

[17] Barkow, Rachel E. 2008. "Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law." *Stanford Law Rev*iew 61: 869.

[18] Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections." *The Quarterly Journal of Economics* 124 (1): 105–47.

[19] Bedard, Kelly, and Elizabeth Dhuey. 2007. "Is September Better than January? The Effect of Minimum School Entry Age Laws on Adult Earnings." Unpublished manuscript.

[20] Becker, Gary S. 1957. *The Economics of Discrimination: An Economic View of Racial Discrimination*. University of Chicago.

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

[22] Behaghel, Luc, and David M Blau. 2012. "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age." *American Economic Journal: Economic Policy* 4 (4): 41–67.

[23] Berger v. United States, 295 U.S. 78. 1935.

[24] Bergman, Peter. 2018. "The Risks and Benefits of School Integration for Participating Students: Evidence from a Randomized Desegregation Program." Working paper.

[25] Berk, Jillian. 2007. "Does work release work?." Unpublished manuscript. Providence, RI: Brown University.

[26] Berk, Richard A., and David Rauma. 1983. "Capitalizing on Nonrandom Assignment to Treatments: A Regression-Discontinuity Evaluation of a Crime-Control Program." *Journal of the American Statistical Association* 78 (381): 21-27.

[27] Berk, Richard A., Kenneth J. Lenihan, and Peter H. Rossi. 1980. "Crime and Poverty: Some Experimental Evidence From Ex-Offenders." *American Sociological Review* 45 (5): 766–786.

[28] Best, Michael, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6): 1311–55.

[29] Best, Michael, and Henrik Jacobsen Kleven. 2018. "Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK." Working paper.

[30] Best, Michael, James Cloyne, Ethan Ilzetzki, and Henrik Jacobsen Kleven. 2018. "Interest Rates, Debt and Intertemporal Allocation: Evidence from Notched Mortgage Contracts in the United Kingdom." Working paper.

[31] Bibas, Stephanos. 2004. "Plea Bargaining Outside the Shadow of Trial." *Harvard Law Review* 2463–2547.

[32] Bibel, Daniel. 2015. "Considerations and Cautions Regarding NIBRS Data: A View From the Field." *Justice Research and Policy* 16 (2): 185-194.

[33] Billings, Stephen B., David J. Deming, and Jonah Rockoff. 2014. “School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg.” *The Quarterly Journal of Economics* 129 (1): 435–476.

[34] Bjerk, David. 2005. “Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing.” *The Journal of Law and Economics* 48 (2): 591–625.

[35] Bjerk, David. 2007. “Guilt Shall Not Escape or Innocence Suffer? The Limits of Plea Bargaining When Defendant Guilt Is Uncertain.” *American Law and Economics Review* 9 (2): 305–329.

[36] Bjerk, David. 2017. “Mandatory Minimums and the Sentencing of Federal Drug Crimes.” *Journal of Legal Studies* 46 (1): 93-128

[37] Blank, Rebecca M. 2002. “Evaluating Welfare Reform in the United States.” *Journal of Economic Literature* 40 (4): 1105–66.

[38] Boylan, Richard T. 2005. “What Do Prosecutors Maximize? Evidence from Careers of U.S. Attorneys.” *American Law and Economics Review* 7 (2): 379-402.

[39] Brown, Kristine M. 2013. “The Link between Pensions and Retirement Timing: Lessons from California Teachers.” *Journal of Public Economics* 98 (February): 1–14.

[40] Bureau of Justice Statistics (BJS). 2018. “2018 Update on Prisoner Recidivism: A 9-Year Follow-up Period (2005-2014).” United States Department of Justice. `https://www.bjs.gov/content/pub/pdf/18upr9yfup0514.pdf` (accessed April 2020).

[41] Bureau of Justice Statistics (BJS). 2019a. “Justice Expenditure And Employment Extracts, 2016 - Preliminary.” United States Department of Justice. `https://www.bjs.gov/index.cfm?ty=pbdetail&iid=6728` (accessed April 2020).

[42] Bureau of Justice Statistics (BJS). 2019b. “Prisoners in 2017.” United States Department of Justice. `https://www.bjs.gov/content/pub/pdf/p17.pdf` (accessed April 2020).

[43] Bureau of Labor Statistics (BLS). 1996-2016. “Local Area Unemployment Statistics: Florida, Seasonally Adjusted - LASST120000000000003.”

United States Department of Labor. https://data.bls.gov/timeseries/LASST120000000000003 (accessed November 2016).

[44] Bureau of Prisons (BOP). 2020. "Prison Safety." https://www.bop.gov/about/statistics/statistics_prison_safety.jsp (accessed April 2020).

[45] Bushway, Shawn D., and Gary Sweeten. 2007. "Abolish Lifetime Bans for Ex-felons." *Criminology & Public Policy* 6 (4): 697–706.

[46] Caetano, Gregorio, and Vikram Maheshri. 2017. "Explaining Recent Trends in US School Segregation: 1988-2014." Working paper.

[47] Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust non-parametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6): 2295–2326.

[48] Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *The Quarterly Journal of Economics* 123 (1): 177–218.

[49] Card, David, David Lee, Zhuan Pei, and Andrea Weber. 2014. "Local Polynomial Order in Regression Discontinuity Designs." Brandeis University Department of Economics and International Business School Working Paper 81.

[50] Carr, Jillian, and Analisa Packham. 2017. "SNAP Benefits and Crime: Evidence from Changing Disbursement Schedules." Working paper.

[51] Carr, Jillian, and Vijetha Koppa. 2017. "The Effect of Housing Vouchers on Crime: Evidence from a Lottery." Working paper.

[52] Carr, Jillian, and William McClain. 2018. "Evidence On Court System Bias From Strategic Judge Assignment." Working paper.

[53] Carrell, Scott, Bruce Sacerdote, and James West. 2013. "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica* 81 (3): 855–882.

[54] Cascio, Elizabeth, Nora Gordon, Ethan Lewis, and Sarah Reber. 2008. "From Brown to Busing." *Journal of Urban Economics* 64 (2): 296–325.

[55] Cascio, Elizabeth, Nora Gordon, Ethan Lewis, and Sarah Reber. 2010. "Paying for Progress: Conditional Grants and the Desegregation of Southern Schools." *The Quarterly Journal of Economics* 125 (1): 445–482.

[56] Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics* 125 (1): 215–261.

[57] Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2018. "The Effect of Minimum Wages on Low-Wage Jobs: Evidence from the United States Using a Bunching Estimator." Working paper.

[58] Chen, Zhao, Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Yi Xu. 2018. "Notching R&D Investment with Corporate Income Tax Cuts in China." Working paper.

[59] Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and economic opportunity in the United States: An intergenerational perspective." *The Quarterly Journal of Economics* 135 (2): 711-783.

[60] Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2): 749–804.

[61] Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7): 2683–2721.

[62] Chetty, Raj, and Nathaniel Hendren. 2018. "The Effects of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

[63] Chioda, Laura, Joo MP De Mello, and Rodrigo R. Soares. 2016. "Spillovers from Conditional Cash Transfer Programs: Bolsa Famlia and Crime in Urban Brazil." *Economics of Education Review* 54: 306-320.

[64] Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics* 94 (1): 260–272.

[65] Cook, Jason. 2016. "Segregation, Student Achievement, and Postsecondary Attainment: Evidence from the Introduction of Race-Blind Magnet School Lotteries." Working paper.

[66] Council of Economic Advisors (CEA). 2016. "Economic Perspectives on Incarceration and the Criminal Justice System." Executive Office of the President of the United States. `https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/-CEA%2BCriminal%2BJustice%2BReport.pdf`

[67] Cunningham, George K., William L. Husk, and James A. Johnson. 1978. "The Impact of Court-Ordered Desegregation on Student Enrollment and Residential Patterns (White Flight)." *Journal of Education* 160 (2): 36–45.

[68] Das, Satadru, and Naci Mocan. 2016. "Analyzing the Impact of the World's Largest Public Works Project on Crime." NBER Working Paper Series No. 22499.

[69] Dee, Thomas S., Will Dobbie, Brian A. Jacob, and Jonah Rockoff. 2017. "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations." Forthcoming at *American Economic Journal: Applied Economics*.

[70] Deshpande, Manasi. 2016. "Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls." *American Economic Review* 106 (11): 3300–3330.

[71] Diamond, Rebecca, and Petra Persson. 2017. "The Long-Term Consequences of Teacher Discretion in Grading of High-Stakes Tests." Working paper.

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

[73] Doleac, Jennifer L., and Benjamin Hansen. 2016. "Does 'Ban the Box' Help or Hurt Low-Skilled Workers? Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden." NBER Working Paper Series No. 22469.

[74] Donohue, John J., III. 2009. "Assessing the Relative Benefits of Incarceration: Overall Changes and the Benefits on the Margin." *Do Prisons make Us Safer? The Benefits and costs of the Prison Boom*, edited by Steven Raphael and Michael A. Stoll, 269–342. New York: Russell Sage Foundation

[75] Drug Enforcement Administration (DEA). 2015. "National Heroin Threat Assessment Summary." https://www.dea.gov/divisions/hq/2015/hq052215_National_Heroin_Threat-_Assessment_Summary.pdf.

[76] Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy* 8 (3): 203–32.

[77] Easterbrook, Frank H. 1983. "Criminal Procedure as a Market System." *The Journal of Legal Studies* 12 (2): 289–332.

[78] Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution." *American Economic Journal: Applied Economics* 8 (4): 36–65.

[79] Fack, Gabrielle, and Camille Landais. 2016. "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France." *Journal of Public Economics* 133 (January): 23–40.

[80] Farrell, Jill. 2003. "Mandatory Minimum Firearm Penalties: A Source of Sentencing Disparity?" *Justice Research and Policy* 5 (1): 95-115.

[81] Fischman, Joshua B., and Max M. Schanzenbach. "Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums." *Journal of Empirical Legal Studies* 9 (4): 729-764.

[82] Florida Department of Children and Families (FL DCF). 1996-2016. "ABAWD Waivers." Freedom of Information Act Request (received December 2016).

[83] Florida Department of Corrections (FL DOC). 2018. "The Offender Based Information System (OBIS) Database." http://www.dc.state.fl.us/pub/obis_request.html (accessed April 2016).

[84] Florida State Legislature, Statute 893.135. "Trafficking[...]" http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&-Search_String&URL=0800-0899/0893/Sections/0893.135.html

[85] Food and Nutrition Services (FNS). 2017. "SNAP State Activity Report, Fiscal Year 2016." U.S. Department of Agriculture. https://fns-prod.azureedge.net/sites/default/files/snap/FY16-State-Activity-Report.pdf.

[86] Foley, C. Fritz. 2010. "Welfare Payments and Crime." *Review of Economics and Statistics* 93 (1): 97–112.

[87] Franklin, Travis W., and Tri Keah S. Henry. 2019. "Racial Disparities in Federal Sentencing Outcomes: Clarifying the Role of Criminal History." *Crime & Delinquency*.

[88] Ganong, Peter, and Simon Jäger. 2017. "A Permutation Test for the Regression Kink Design." *Journal of the American Statistical Association*, forthcoming.

[89] Gassman-Pines, Anna, and Laura Bellows. 2015. "Food Instability and Academic Achievement: A Quasi-Experiment Using SNAP Benefit Timing." Working paper.

[90] Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2013. "Estimating Earnings Adjustment Frictions: Method and Evidence from the Social Security Earnings Test." NBER working paper.

[91] Gelman, Andrew, and Guido Imbens. 2018. "Why Higher-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics*.

[92] Gilna, Derek. 2016. "Report Calls for End of Welfare and Food Stamp Restrictions for Felony Drug Offenders." *Prison Legal News*. `https://bit.ly/3dkuu9Y`.

[93] Glaeser, Edward L., Daniel P. Kessler, and Anne Morrison Piehl. 2000. "What Do Prosecutors Maximize? An Analysis of Federalization of Drug Crimes." *American Law and Economics Review* 2 (2): 259-290.

[94] Goncalves, Felipe, and Steven Mello. 2018. "A Few Bad Apples? Racial Bias in Policing." Working paper.

[95] Goodman, Joshua, Michael Hurwitz, Jisung Park, and Jonathan Smith. 2018. "Heat and Learning." *American Economic Journal: Economic Policy*. Forthcoming.

[96] Gordon, Sanford C. 2009. "Assessing Partisan Bias in Federal Public Corruption Prosecutions." *American Political Science Review* 103 (04): 534–54.

[97] Gould, Eric D., Bruce A. Weinberg, and David B. Mustard. 2002. "Crime Rates and Local Labor Market Opportunities in the United States: 1979–1997." *Review of Economics and Statistics* 84 (1): 45–61.

[98] Government Accountability Office (GAO). 2005. "Drug Offenders: Various Factors May Limit the Impacts of Federal Laws That Provide for Denial of Selected Benefits." Report to Congressional Requesters. `http://www.gao.gov/assets/250/247940.pdf`.

[99] Gregory, Christian A., and Partha Deb. 2015. "Does SNAP Improve Your Health?" *Food Policy* 50 (January): 11–19.

[100] Grogger, Jeffrey, and Greg Ridgeway. 2006. "Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness." *Journal of the American Statistical Association* 101 (475): 878– 887.

[101] Grossman, Gene M., and Michael L. Katz. 1983. "Plea Bargaining and Social Welfare." *American Economic Review* 73 (4): 749-757.

[102] Guryan, Jonathan. 2004. "Desegregation and Black Dropout Rates." *American Economic Review* 94 (4): 919–943.

[103] Haney, Craig. 2001. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." National Policy Conference.

[104] Hansen, Benjamin. 2015. "Punishment and Deterrence: Evidence from Drunk Driving." *American Economic Review* 105 (4): 1581–1617.

[105] Haskins, Ron. 2006. "Interview: Welfare reform, 10 years later." *The Brookings Institution*, August 24. `http://www.brookings.edu/research/interviews/2006/08/24welfare-haskins`.

[106] Heissel, Jennifer, Claudia Persico, and David Simon. 2019. "Does Pollution Drive Achievement? The Effect of Traffic Pollution on Academic Performance." National Bureau of Economic Research. Working paper.

[107] Hendren, Nathaniel. 2017. "Efficient Welfare Weights." National Bureau of Economic Research. Working paper.

[108] Honold, Dan. 2014. "Note: Quantity, Role, and Culpability in the Federal Sentencing Guidelines." *Harvard Journal on Legislation* 51: 389-314.

[109] Hoynes, Hilary, and Diane Whitmore Schanzenbach. 2012. "Work Incentives and the Food Stamp Program." *Journal of Public Economics* 96 (1–2): 151–162.

[110] Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903–934.

[111] Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933–959.

[112] Imbens, Guido, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615–35.

[113] Immigration and Nationality Act (INA). 8 USCS 1101. `https://www.law.cornell.edu/uscode/text/8/1101`.

[114] Jackson, Kirabo. 2009. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." *Journal of Labor Economics* 27 (2): 213–256.

[115] Jacob, Brian A., and Jens Ludwig. 2012. "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery." *American Economic Review* 102 (1): 272–304.

[116] Jacob, Brian A., Max Kapustin, and Jens Ludwig. 2015. "The Impact of Housing Assistance on Child Outcomes: Evidence from a Randomized Housing Lottery." *The Quarterly Journal of Economics* 130 (1): 465–506.

[117] JCPS. 2019. "History of JCPS." `https://www.jefferson.kyschools.us/about/jcps-history`.

[118] Johnson, Rucker C. 2015. "Long-Run Impacts of School Desegregation & School Quality on Adult Attainments." National Bureau of Economic Research. Working paper.

[119] K'Meyer, Tracy E. 2013. "From Brown to Meredith: The Long Struggle for School Desegregation in Louisville, Kentucky, 1954-2007." University of North Carolina Press.

[120] Kaeble, Danielle, and Mary Cowhig. 2018. "Correctional Populations in the United States, 2016." Bureau of Justice Statistics. `https://www.bjs.gov/content/pub/pdf/cpus16.pdf`.

[121] Kaplan, Ethan, Jorg Spenkuch, and Cody Tuttle. 2019. “The Long-run Political Consequences of Racial Integration.” Working paper.

[122] King, Ryan D., and Michael T. Light. 2019. “Have Racial and Ethnic Disparities in Sentencing Declined?” *Crime and Justice*.

[123] Kleven, Henrik Jacobsen. 2016. “Bunching.” *Annual Review of Economics* 8: 435–464.

[124] Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez. “Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark.” 2011. *Econometrica* 79 (3): 651–692.

[125] Knorre, Alexey. 2017. “How to detect fraud in drug crime records? An analysis of 585,000 crime records in Russia.” Working paper.

[126] Knowles, John, Nicola Persico, and Petra Todd. 2001. “Racial Bias in Motor Vehicle Searches: Theory and Evidence.” *Journal of Political Economy* 109 (1): 203–229.

[127] Knox, Dean, Will Lowe, and Jonathan Mummolo. 2019. “The Bias Is Built In: How Administrative Records Mask Racially Biased Policing.” Working paper.

[128] Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo. 2013. “Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment.” *The Quarterly Journal of Economics* 128 (3): 1123–1167.

[129] Kuziemko, Ilyana. 2013. “How Should Inmates Be Released from Prison? An Assessment of Parole Versus Fixed-Sentence Regimes.” *The Quarterly Journal of Economics* 128 (1): 371-424.

[130] Landes, William M. 1971. “An Economic Analysis of the Courts.” *Journal of Law and Economics* 14 (1): 61-107.

[131] Lee, David S., and Justin McCrary. 2017. “The Deterrence Effect of Prison: Dynamic Theory and Evidence.” *Advances in Econometrics* 38.

[132] Levitt, Steven D., and Sudhir Alladi Venkatesh. 2000. “An Economic Analysis of a Drug-Selling Gang’s Finances.” *The Quarterly Journal of Economics* 115 (3): 755–789.

[133] Lofstrom, Magnus and Steven Raphael. 2016. Crime, the Criminal Justice System, and Socioeconomic Inequality. *Journal of Economic Perspectives* 30(2): 103-26.

[134] Looney, Adam, and Nicholas Turner. 2018. "Work and opportunity before and after incarceration." Washington, DC: Brookings Institution.

[135] Luallen, Jeremy, Jared Edgerton, and Deirdre Rabideau. 2017. "A Quasi-Experimental Evaluation of the Impact of Public Assistance on Prisoner Recidivism." *Journal of Quantitative Criminology* May: 1-33.

[136] Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start improve children's life chances? Evidence from a regression discontinuity design." *The Quarterly Journal of Economics* 122 (1): 159-208.

[137] Lutz, Byron. 2011. "The End of Court-Ordered Desegregation." *American Economic Journal: Economic Policy* 3 (2): 130–168.

[138] Lynch, Mona. *Hard Bargains: The Coercive Power of Drug Laws in Federal Court.* New York: Russell Sage Foundation, 2016.

[139] Mabli, James, and Jim Ohls. 2015. "Supplemental Nutrition Assistance Program Participation Is Associated with an Increase in Household Food Security in a National Evaluation." *The Journal of Nutrition* 145 (2): 344-351.

[140] Mallar, Charles D., and Craig V. D. Thornton. 1978. "Transitional Aid for Released Prisoners: Evidence from the Life Experiment." *The Journal of Human Resources* 13 (2): 208–36.

[141] Maltz, Michael D. 1984. *Recidivism*. Orlando, Florida: Academic Press, Inc.

[142] Mastrobuoni, Giovanni, and Paolo Pinotti. 2015. "Legal Status and the Criminal Activity of Immigrants." *American Economic Journal: Applied Economics* 7 (2): 175–206.

[143] Mathematica Policy Research (MPR). 1996-2014. "SNAP Quality Control Data, Public Use Files." `https://host76.mathematica-mpr.com/fns/`. (accessed June 2017).

[144] McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714.

[145] Mele, Angelo. 2019. "Does school desegregation promote diverse interactions? An equilibrium model of segregation within schools." *American Economic Journal: Economic Policy*. Forthcoming.

[146] Miller, Ted, Mark Cohen, and Brian Wiersema. 1996. "Victim Costs and Consequences: A New Look." Final Summary Report to the National Institute of Justice.

[147] Moore, Timothy J. 2015. "The Employment Effects of Terminating Disability Benefits." *Journal of Public Economics* 124 (April): 30–43.

[148] Mueller-Smith, Michael. 2015. "The Criminal and Labor Market Impacts of Incarceration." University of Michigan Working Paper.

[149] Mueller-Smith, Michael, and Kevin Schnepel. 2017. "Diversion in the Criminal Justice System: Regression Discontinuity Evidence on Court Deferrals." Working paper.

[150] Munyo, Ignacio, and Martín A. Rossi. 2015. "First-Day Criminal Recidivism." *Journal of Public Economics* 124 (April): 81–90.

[151] Mustard, David B. 2001. "Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts." *Journal of Law and Economics* 44: 285–314.

[152] Natkin, Gerald L. 1980. "The Effects of Busing on Second Grade Students Achievement Scores." Technical Report.

[153] Nyhan, Brendan, and M. Marit Rehavi. 2017. "Tipping the Scales? Testing for Political Influence on Public Corruption Prosecutions." Working paper.

[154] Orfield, Gary, Genevieve Siegel-Hawley, and John Kucsera. 2014. "Sorting Out Deepening Confusion on Segregation Trends." The Civil Rights Project.

[155] Oswald, Frederick, Gregory Mitchell, Hart Blanton, and Philip Tetlock. 2013. "Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies." *Journal of Personality and Social Psychology* 105 (2): 171-192.

[156] Pager, Devah, Bruce Western, and Naomi Sugie. 2009. "Sequencing Disadvantage: Barriers to Employment Facing Young Black and White Men with Criminal

Records." *The Annals of the American Academy of Political and Social Science* 623 (1): 195–213.

[157] Panel Study of Income Dynamics (PSID), public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. `https://simba.isr.umich.edu/default.aspx` (accessed November 2016).

[158] Pfaff, John. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform*. New York: Basic Books, 2017.

[159] Phillips, Ryan. 2016. "Georgia may soon lift ban on food stamps for drug felons." *Athens Banner-Herald*, April 26. `http://www.onlineathens.com/article/20160426/NEWS/304269967`.

[160] ProPublica. 2017. "COMPAS Recidivism Risk Score Data." `https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis` (accessed March 2017).

[161] Raphael, Steven. 2011. "Incarceration and Prisoner Reentry in the United States." *The Annals of the American Academy of Political and Social Science* 635 (1): 192–215.

[162] Reber, Sarah J. 2005. "Court-Ordered Desegregation: Successes and Failures in Integration Since Brown vs. Board of Education." *Journal of Human Resources* 40 (3): 559-590.

[163] Reber, Sarah J. 2010. "School Desegregation and Educational Attainment for Blacks." *Journal of Human Resources* 45 (4): 893–914.

[164] Rehavi, M. Marit, and Sonja B. Starr. 2012. "Racial Disparity in Federal Criminal Sentences and Its Sentencing Consequences." Univ. of Michigan Law School Program in Law and Economics Working Paper Series. Working paper no. 12-002.

[165] Rehavi, M. Marit, and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6):1320–54.

[166] Reinganum, Jennifer F. 1988. "Plea Bargaining and Prosecutorial Discretion." *American Economic Review* 78 (4): 713-728.

[167] Rhodes, William, Ryan Kling, Jeremy Luallen, and Christina Dyous. 2015. "Federal Sentencing Disparity: 2005-2012." Bureau of Justice Statistics Working Paper Series.

[168] Robert Wood Johnson Foundation (RWJF). 2018. "Discrimination in America: Final Summary."

[169] Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. "Integrated Public Use Microdata Series: Version 6.0 [dataset]." Minneapolis: University of Minnesota (accessed November 2016).

[170] Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

[171] Sallee, James M., and Joel Slemrod. 2012. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96 (11): 981-999.

[172] Schanzenbach, Diane Whitmore, Ryan Nunn, Lauren Bauer, Audrey Breitwieser, Megan Mumford, and Greg Nantz. 2016. "Twelve Facts about Incarceration and Prisoner Reentry." Economic Facts, The Hamilton Project, The Brookings Institution, Washington, DC. `https://www.brookings.edu/research/twelve-facts-about-incarceration-and-prisoner-reentry/`.

[173] Schmitt, John, and Kris Warner. 2010. "Ex-Offenders and the Labor Market." Center for Economic and Policy Research. `http://cepr.net/documents/publications/ex-offenders-2010-11.pdf`.

[174] Schnepel, Kevin T. 2018. "Good Jobs and Recidivism." *The Economic Journal* 128: 447-469.

[175] Schwartz, J., C. Spix, G. Touloumi, L. Bachrov, T. Barumamdzadeh, A. le Tertre, T. Piekarksi, et al. 1996. "Methodological Issues in Studies of Air Pollution and Daily Counts of Deaths or Hospital Admissions." *Journal of Epidemiology and Community Health* (1979-) 50: S3–11.

[176] Shermer, Lauren O'Neill, and Brian D. Johnson. 2009. "Criminal Prosecutions: Examining Prosecutorial Discretion and Charge Reductions in U.S. Federal District Courts." *Justice Quarterly* 1-37.

[177] Shively, Michael. 2005. "Study of Literature and Legislation on Hate Crime in America." National Institute of Justice.

[178] Shoag, Daniel, and Stan Veuger. 2016. "Banning the Box: The Labor Market Consequences of Bans on Criminal Record Screening in Employment Applications." Working paper.

[179] Short, Kathleen. 2015. "The Supplemental Poverty Measure: 2014." U.S. Census Bureau, Current Population Reports. `https://www.census.gov/content/dam/Census/library/publications/2015/demo/-p60-254.pdf`.

[180] Simon, Jonathan. 2007. *Governing Through Crime: How the War on Crime and American Democracy Created a Culture of Fear*. Oxford: Oxford Univ. Press

[181] Sklansky, David Alan. 2018. "The Problems with Prosecutors." *Annual Review of Criminology* 1: 451–469.

[182] Sloan, CarlyWill. 2019. "Racial Bias by Prosecutors: Evidence from Random Assignment." Working paper.

[183] Spohn, Cassia, John Gruhl, and Susan Welch. 1987. "The Impact of the Ethnicity and Gender of Defendants on the Decision to Reject or Dismiss Felony Charges." *Criminology* 25 (1): 175–192.

[184] Stephens-Davidowitz, Seth. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data." *Journal of Public Economics* 118: 26-40.

[185] Stuntz, William J. 2004. "Plea Bargaining and Criminal Law's Disappearing Shadow." *Harvard Law Review* 117 (8): 2548.

[186] Sugie, Naomi. 2012. "Punishment and Welfare: Paternal Incarceration and Families' Receipt of Public Assistance." *Social Forces* 90: 1403-1427.

[187] The Hamilton Project. 2016. "Twelve Facts about Incarceration and Prisoner Reentry." Economic Facts, The Hamilton Project, The Brookings Institution, Washington, DC

[188] Traxler, Christian, Franz G. Westermaier, and Ansgar Wohlschlegel. 2018. "Bunching on the Autobahn? Speeding Responses to a 'Notched' Penalty Scheme." *Journal of Public Economics* 157: 78–94.

[189] Uggen, Christopher. 2000. "Work as a Turning Point in the Life Course of Criminals: A Duration Model of Age, Employment, and Recidivism." *American Sociological Review* 65 (4): 529–546.

[190] Ulmer, Jeffrey T., Megan C. Kurlychek, and John H. Kramer. 2007. "Prosecutorial Discretion and the Imposition of Mandatory Minimum Sentences." *Journal of Research in Crime and Delinquency* 44 (4): 427-458.

[191] Ulmer, Jeffrey, Noah Painter-Davis, and Leigh Tinik. 2014. "Disproportional Imprisonment of Black and Hispanic Males: Sentencing Discretion, Processing Outcomes, and Policy Structures." *Justice Quarterly*.

[192] U.S. Congress. *Congressional Record*. 104th Cong., 2nd sess., 1996. Vol. 142. `https://www.gpo.gov/fdsys/pkg/CREC-1996-07-23/content-detail.html`.

[193] U.S. Department of Agriculture (USDA). 2011. "Dynamics of Supplemental Nutrition Assistance Program in the Mid-2000s." `https://www.fns.usda.gov/snap/dynamics-supplemental-nutrition-assistance-program-participation//-mid-2000s`.

[194] U.S. Department of Agriculture (USDA). 2016a. "Annual Report to Congress SNAP Employment and Training (E&T) Pilot Projects Authorized by the Agricultural Act of 2014." `https://www.fns.usda.gov/sites/default/files/snap/SNAP-E-and-T-2016-report.pdf`.

[195] U.S. Department of Agriculture (USDA). 2016b. "Supplemental Nutrition Assistance Program (SNAP): Able-Bodied Adults Without Dependents (ABAWDs)." `http://www.fns.usda.gov/snap/able-bodied-adults-without-dependents-abawds`.

[196] U.S. Department of Health and Human Services (HHS). 2015. "Characteristics and Financial Circumstances of TANF Recipients, Fiscal Year 2013." `https://www.acf.hhs.gov/sites/default/files/ofa/tanf_characteristics_fy2013.pdf`.

[197] U.S. Department of Justice (DOJ). 2011. "The Department of Justice's International Prisoner Transfer Program." https://oig.justice.gov/reports/2011/e1202.pdf.

[198] U.S. Department of Justice (DOJ). 2012. "NIBRS Participation by State."

[199] U.S. Sentencing Commission (USSC). 2015a. "Report to the Congress: Impact of the Fair Sentencing Act of 2010."

[200] U.S. Sentencing Commission (USSC). 2015b. "United States Sentencing Commission (USSC) Guidelines Manual."

[201] U.S. Sentencing Commission (USSC). 2017. "Demographic Differences in Sentencing: An Update to the 2012 Booker Report."

[202] Webb, Dan K., and Scott F. Turow. 1982. "The Prosecutor's Function in Sentencing." *Loyola U. Chicago Law Journal* 13 (4): 641-668.

[203] Welch, Finish, and Audrey Light. 1987. "New Evidence on School Desegregation." U.S. Commission on Civil Rights Clearinghouse (Washington, DC) Publication 92.

[204] West, Jeremy. 2018. "Racial Bias in Police Investigations." Working paper.

[205] Western, Bruce, Anthony A. Braga, Jaclyn Davis, and Catherine Sirois. 2015. "Stress and Hardship after Prison." *American Journal of Sociology* 120 (5): 1512–1547.

[206] Wolkomir, Elizabeth. 2018. "How SNAP Can Better Serve the Formerly Incarcerated." Center on Budget and Policy Priorities. https://www.cbpp.org/research/food-assistance/how-snap-can-better-serve-the-formerly-incarcerated.

[207] Wright, Ronald. 2006. "Federal or State? Sorting as a Sentencing Choice." *Criminal Justice* 21 (2): 16-21.

[208] Xu, Frank, Brian Nosek, Anthony Greenwald, Kate Ratliff, Yoav Bar-Anan, Emily Umanksy, Mahzarin Banaji, Nicole Lofaro, Colin Smith, Jordan Axt. 2019. "Project Implicit Demo Website Datasets."

[209] Yang, Crystal S. 2016. "Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies." *American Economic Journal: Economic Policy* 8 (4): 289–332.

[210] Yang, Crystal S. 2017a. "Does Public Assistance Reduce Recidivism?" *American Economic Review: Papers and Proceedings* 107 (5): 551–555.

[211] Yang, Crystal S. 2017b. "Local Labor Markets and Criminal Recidivism." *Journal of Public Economics* 147 (March): 16–29.

[212] Zaresani, Arezou. 2016. "Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program." University of Calgary working paper.