Evaluating State-Driven Changes to the Medicaid Program: Unintended, Intended, and Methodological Implications

Dissertation Advisor: Richard G. Frank

Carrie E. Fry

Evaluating State-Driven Changes to the Medicaid Program:
Unintended, Intended, and Methodological Implications

Abstract

This dissertation consists of two empirical policy papers and one methods paper. All three papers examine the effect of changes to the Medicaid program. The first paper examines the impact of Medicaid expansion of jail-based recidivism. The second paper estimates the effect of retroactive eligibility waivers in Medicaid on enrollment. The third explores the implications of choosing a study design on the estimated effects of these changes by re-analyzing three published papers.

In Chapter 1, co-authors and I estimate the impact of Medicaid expansion on recidivism. Previous research on the relationship between financial access to care and re-offense is mixed, and much of the published work is subject to selection bias, does not have a comparison group, or lacks a defined intervention. We use the variation introduced by the 2012 Supreme Court ruling in *NFIB vs. Sebelius* to derive causal estimates of this relationship using 48 months of booking and release data from six county jails. Three of the six counties are in Medicaid expansion states, and three are in non-expansion states. We conduct three case studies using a comparative interrupted time series analysis (CITS) to estimate the differential change in the probability of re-arrest and the number of arrests between the expansion and non-expansion counties. We find mixed results across these three case studies – in two case studies, we estimate declines in the probability of re-arrest of 5 and 13 percent. In the third, we estimate an increase of similar magnitude. We find a similar pattern of results with the number of arrests. To put these mixed results in context, we supplement our quantitative analysis with information from site visits and stakeholder interviews to identify mediators and moderators of the relationship between financial access to care and recidivism.

In Chapter 2, I estimate what happens to Medicaid enrollment after the implementation of a retroactive eligibility waiver in a state's Medicaid program. Retroactive eligibility provides 90 days of Medicaid coverage prior to a person's date of application, given that the beneficiary was eligible in those 90 days. In the past five years, seven states have eliminated retroactive eligibility for some portion of the Medicaid population. However, we know of no study that examines what happens to enrollment, beneficiary financial status, or health outcomes after the removal of this provision. We use 24 months of Medicaid enrollment data in four of the seven retroactive eligibility waiver states to estimate the relationship between retroactive eligibility removal and changes in enrollment. Using a difference-in-differences (DID) and geographically similar comparison states, we find no impact of retroactive eligibility on Medicaid enrollment in any of the four states. However, the confidence intervals suggest that we may be under-powered. To address the power concerns, we combine the four retroactive eligibility states and their comparators in a pooled analysis. Here, we find a 10 percent decline in Medicaid enrollment at five and six months after waiver implementation, suggesting that removing retroactive eligibility may have a 'chilling' effect on Medicaid enrollment in the months after implementation.

In Chapter 3, I explore the differences between two similar study designs – CITS and DID. Both of these designs use two time periods and a comparison group; they also use the change in the comparison group to estimate the counterfactual for the treated group without treatment. However, the use of these two study designs is disciplinary, and the respective disciplines prefer one design over the other. This is due, in part, to the lack of mathematical formalization for CITS. To understand the differences (if any), we first carefully write down the potential outcome model for two versions of each design – a general version of CITS, a linear version of CITS, DID with time fixed effects, and DID with time fixed effects and group-specific trends – and conduct a modeling exercise to estimate the counterfactuals for each. We, then, re-analyze three published studies to understand the situations where one of these designs might be preferable to the others. We find that general CITS and DID with time fixed effects and group-specific trends produce the same counterfactual and estimate the same treatment effects. The only difference between these two designs is the language used to describe them. We also find that when researchers lean into each design's respective constraints – linearity for CITS and a zero difference for DID – counterfactual and treatment effect estimation differ. Empirical researchers should provide a clear explanation of the counterfactual assumptions being made and the model specification to allow for a more transparent evaluation of the plausibility of these assumptions.

Contents

Dissertation Advisor: Richard G. Frank

Carrie E. Fry

To all the women who have forged this path before me.

May we leave it wider than we found it.

Acknowledgements

Getting a PhD is a significant accomplishment that takes hard work, grit, and determination. All too often, however, this feat is attributed solely to the person who achieved it. But opportunity is not distributed equally. There are many, many other people who, if given the same opportunities in life, could be here instead of me. Acknowledging our relative positions of power and privilege that give us these opportunities is necessary to dismantle the systems that perpetuate the opportunity gap.

These kinds of achievements also take a village. I did not and could not have done this on my own. I would like to thank and acknowledge the following people for their role in my life

I'd like to thank my advisors – Richard Frank, Laura Hatfield, and Ben Sommers. All three have been constant sources of support over the past four years and have made my PhD experience more fun. I consider myself very fortunate to have the three of them in my corner. In addition to my advisors, a number of other faculty members have been instrumental in my success as a graduate student – Mary Beth Landrum, Sara Bleich, Sherri Rose, Meredith Rosenthal, Haiden Huskamp, and Kathy Swartz. As female faculty members, you all truly lead by example.

The administrative staff in the Health Policy PhD programs are what keep the program afloat. Debbie Whitney and Colleen Yout were the first line of defense for me and countless other graduate students for concerns both personal and academic, and their impact on graduate student life is immeasurable.

The administrative staff in the Department of Health Care Policy and the Department of Health Policy and Management contributed greatly to my success, even though it was not in their job description. Aurora DeMattia, Sarah Chambers, Lauren Jett, and Laura Cowieson all made sure that I was paid on time, facilitated meetings with faculty members, and ensured that I met requirements for certain funding opportunities. Emily Crawford let me into my office countless times when I forgot my keys at home.

The cohort of 2018 – Lucy Chen, Alee Lockman, Summer Rak, Amanda Speller, Karen Smith, Jason Buxbaum, Travis Donahue, Lyndon James, Bijan Niknam, and Max Pany – in two semesters, you all taught me more about being a mentor and leader than any other experience in my life. Thank you. You are passionate and brilliant scholars. I cannot wait to see how each of you change the world.

The cohort of 2016 – my cohort. Thank you all for making a Southerner feel at home in New England, for supporting, celebrating, and championing me and each other, and for holding space for vulnerability. In particular, I'd like to thank Adrianna McIntyre for many much-needed lunches at Tasty Burger, her wonkiness, and countless invitations to social events at conferences. Caroline Kelly-Geiger for help and support through coursework and being the one constant presence at Friday morning writing group.

To my friends and program mates – Kate Lofgren, Lauren Taylor, Amanda Krieder, Alyssa Bilinski, Christine Baugh, Jamie Daw, Micah Aaron, and Rebecca Gourevitch for their wisdom, guidance, and support.

My peer mentoring group – Alex McDowell and Michael Anne Kyle. I would not have survived this without you both. Your unwavering support combined with your willingness to question my assumptions made me a better person. Our collective experiences over the course of four years is a beautiful testament to the perseverance and resilience of brilliant, passionate, empathetic, and driven women. You both inspire me every day.

To my little sister, Starr, for keeping me young at heart, making me laugh, and showing me how brilliant your generation is. The kids are alright.

To the Standard Errors softball team – thank you for getting me out of the house every Sunday, letting me air out my competitive side, being a source of lightness mixed with wonkiness, and making me a champion.

To Karen Jandorf – you have been such an important part of my life for so long. You model, inspire, and nurture justice, compassion, and empathy.

To my dad, Greg Fry, thank you for being prouder of me than you'd ever let on. To my mom, Ami Fry, who I know will answer my phone call at any time, day or night, to comfort and support me. Growing up, you embodied the strength of women. Thank you for being my biggest cheerleader. To my siblings, Britta Oakley and Austin Fry, for countless tears from laughter and sadness – I couldn't ask for better siblings.

To my dog, Banjo, who is a source of unconditional love and always knows how and when to cheer me up.

To my partner in life's journey, Joe Dunn. You embody perseverance every day by pursuing your passion regardless of the outcome. For that example, I am forever grateful.

List of Tables

List of Figures

Chapter 1:
Medicaid Expansion's Spillover to the Criminal Justice System - Evidence from Six Urban Counties

Co-authored with Richard G. Frank, PhD and Thomas G. McGuire, PhD

Dissertation Advisor: Richard G. Frank

Carrie E. Fry

Abstract

Spillovers from the Affordable Care Act's (ACA's) Medicaid expansion to other social-sector outcomes have received little attention. One spillover that may be especially salient for public policy is the impact of expanded Medicaid eligibility on jail-related outcomes. This study compares recidivism outcomes in three non-expansion counties to nearby expansion counties before and after Medicaid expansion. Using forty-eight months of arrest data from six urban county jails, we conduct comparative interrupted time series analyses to describe changes in the probability of re-arrest and the number of arrests before and after Medicaid expansion. Consistent with previous literature, we find mixed results. In two case studies, Medicaid expansion is associated with decreased rates of recidivism. In the other, we find differential increases in jail-based recidivism after Medicaid expansion. We use contextual information from site visits and stakeholder interviews to understand the factors that may mediate and moderate the relationship between Medicaid expansion and return to jail.

Introduction

Prior to the authorization of state options for Medicaid expansion, low-income adults without dependent children were ineligible for Medicaid in most states. This population overlapped with the jail-involved population significantly. Both groups were predominantly young, low-income, racial/ethnic minorities, and male. In 2014, men made up 85.3 percent[1] and young adults (18-34 years old) made up 60 percent of the jail-involved population.[2] Additionally, men of color are more likely to be jailed than their white counterparts – one in 106 white men were incarcerated in 2006 versus 1 in 36 Latino men and 1 in 7 African-American men.[2] Given this overlap, early estimates suggested that 25-30 percent of those released from jail in a given year would enroll in Medicaid in expansion states.[2]

Although federal law permits Medicaid coverage to continue during a person's incarceration, 46 states do not continue coverage.[3] The vast majority of people released from jail thus have historically found themselves in the community without immediate access to health insurance of any kind.[4] This is particularly problematic for individuals with mental illnesses as Medicaid is the largest payer for treatment services for these diagnoses in the United States and an increasingly large payer for treatment for substance use disorders. Thus, disruption in coverage reduces the likelihood of receiving timely community-based behavioral healthcare services after release from incarceration. Together, these conditions reinforce a cycle whereby jail-involved people return to the community with little support.

Jail-involved individuals have higher rates of chronic physical health conditions (e.g., asthma and diabetes), communicable diseases (e.g., HIV and Hepatitis C), mental illnesses, and substance use disorders compared to the general population.[5] Estimates from 2007-2009 report that 63 percent of sentenced jail inmates meet diagnosable criteria for drug dependence or abuse,[6] and nearly two-thirds have a diagnosable mental illness at booking or in the 12 months prior to arrest.[2] Indeed, jail-involved people have a 14 percentage point higher rate of serious mental illness (SMI) than the general population.[7]

These high risks stem in large part from the conditions under which justice-involved people live and from their limited ability to obtain appropriate care for their health needs. For instance, one-third of individuals taking prescription drugs do not have access to necessary medication while in jail, and more than half (60 percent) of inmates who require routine blood testing had no testing while in jail.[2] Additionally, most jail-involved people do not have access to healthcare services in the community, due to gaps in the healthcare safety net and/or a lack of health insurance coverage.

Providing timely access to healthcare services, particularly treatment for mental illnesses and substance use disorders, may be one way to reduce rates of re-offense. Indeed, recent evidence found that improved access to evidence-based treatment for mental and addictive illnesses can improve re-entry outcomes.[8] Other evidence, however, on the relationship between access to healthcare services and recidivism found mixed results. Additionally, many of these previous studies were retrospective and used non-experimental designs that do not account for selection into treatment, precluding reliable causal inference.

Early work found that recently released individuals with a severe mental illness (SMI) in King County, WA and Pinellas County, FL who obtained Medicaid coverage upon release from prison were 16 percent less likely to be re-arrested in the following year and spent more time in the community before re-arrest (102 vs. 93 days) compared to similar individuals who did not get Medicaid coverage.[9] In contrast, a study using administrative records from 2006-2007 found that although expedited Medicaid enrollment for individuals with SMI released from Washington State's prisons led to greater Medicaid enrollment and mental health service use, the intervention was not associated with reduced rates of recidivism at twelve[10] or thirty-six months post-release.[11]

In a more recent study using these same data in Washington, Domino and colleagues examined the relationship between access to 'timely' mental health services (defined as those received within twelve months of being released) and rates of recidivism for individuals recently released from prison.[12] The authors found that the receipt of 'timely' mental health services was associated with an increased rate of recidivism, specifically for technical violations, at twelve months post-release.[12] Interpreting these results as causal is difficult, since those more likely to access services may also be less likely to violate parole or re-offend.

The U.S. Government Accountability Office (GAO) studied re-entry programs in Florida and Michigan. In Florida, the GAO found that access to community services, including through Medicaid, upon release from prison was not associated with a decreased likelihood in re-arrest, overall. Some groups (black and older individuals) did experience a slight decrease in the likelihood of re-arrest [13]. During the study period, however, Florida's Medicaid program did not pay for substance use disorder treatment, which may have mitigated any potential effect.[13] Michigan's implemented program consisted of prison 'in-reach' sessions, health screenings, and connections to healthcare services for individuals soon-to-be-released from prison.[14] After the implementation of the program, the recidivism rate fell by 18.2 percentage points for two-year parolees and 8.4 percentage points for one-year parolees who received these services compared to the recidivism rates for these same individuals prior to the implementation of these services. However, this study lacked a control group or a well-defined intervention.[14]

The ACA and Recidivism

The ACA's coverage provisions (including Medicaid expansion) not only increased eligibility for previously ineligible populations, but they also changed the type of behavioral healthcare services that low-income adults have access to. Prior to the ACA, the public behavioral health system was funded predominantly through categorical Medicaid programs (typically excluding single childless adults from eligibility) and state and federal budgeted funding mechanisms (e.g. federal block grants). These funding mechanisms often did not require evidence-based treatment in specialty substance use disorder treatment programs, nor did they encourage adequate treatment capacity.

Since the enactment of the ACA, treatment providers in expansion states generally must meet conditions of participation in Medicaid, which require greater capacity and provision of evidence-based behavioral health care. Additionally, federal legislation and regulations attempt to standardize Medicaid benefits for behavioral healthcare services. Together, the Mental Health Parity and Addictions Equity Act of 2008 and the ACA require all Medicaid managed care plans to cover treatment services for mental illness and substance use disorder as essential health benefits, and further require that these benefits must also be covered and managed at parity with medical/surgical benefits. The ACA also requires that all enrollees from the expansion population have coverage for mental health and substance use disorder care that is at parity with medical/surgical coverage.

Despite federal legislation and regulation to standardize these benefits, the fee-for-service behavioral health services provided by Medicaid vary by state, as fee-for-service Medicaid programs are not subject to the parity provisions of MHPAEA or the ACA (with the exception of expansion adults). For instance, 43 of the 51 jurisdictions surveyed cover inpatient psychiatric hospitals stays, four states require a co-payment for these services, and 17 states place a limit on these services.[15] Coverage of residential psychiatric services is even more variable – only 23 state Medicaid programs provide this benefit. While all 50 states provide coverage for buprenorphine for the treatment of opioid use disorder, 21 require a co-payment for this service and 19 require a prior authorization.[15] Additionally, ten states (including Louisiana, in our study sample) do not cover methadone for OUD.

In addition to changes in Medicaid benefit design and covered services, the target populations of the ACA's Medicaid expansion differ from those included in previous studies examining the relationship between access to healthcare services and recidivism. Taken together, these imply that previous findings about Medicaid and recidivism may not generalize to the expansion population. In this study, we extend previous analyses to the Medicaid expansion population and provide one of the first quasi-experimental analyses of the relationship between gaining health insurance coverage and criminal justice outcomes. To do so, we compare recidivist outcomes among jail-involved individuals in three non-expansion counties to nearby expansion counties before and after Medicaid expansion.

Data and Empirical Methods

We examine the relationship between expanded Medicaid eligibility and recidivism with an intent-to-treat analysis that takes advantage of the plausibly exogenous variation provided by Supreme Court ruling in *NFIB vs. Sebelius*, which gave states the option to expand their Medicaid program. We used forty-eight continuous months of individual-level booking and release dates from six urban county jails (three in expansion states and three in nearby non-expansion states) and comparative interrupted time series regression analysis to describe the level and trends of a) the rate of re-arrest and b) the number of arrests before and after Medicaid expansion in expansion and non-expansion counties. We also estimate the impact of expanded Medicaid eligibility on rates of recidivism for the whole sample and for the largest racial/ethnic groups. Finally, in a qualitative analysis, we identify county-level re-entry or diversion programs/policies that may explain the differential relationship by county-pair (if any) between expanded Medicaid coverage and recidivism.

County Selection and Characteristics

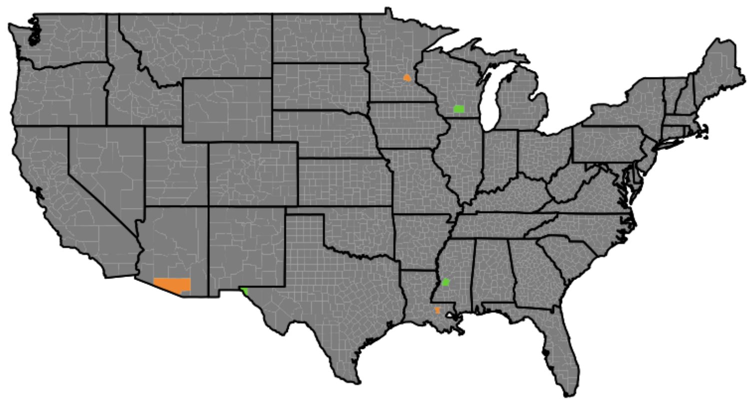
We initially selected four county pairs – Hennepin County, MN (expansion) and Dane County, WI (Midwest region); Pima County, AZ (expansion) and El Paso County, TX (Southwest region); East Baton Rouge Parish (EBR), LA (expansion) and Hinds County, MS (Southeast region); and St. Louis, MO and East St. Louis, IL. Three of these four pairs agreed to participate and provide data for our study; the St. Louis locales declined to participate (see Figure 1.1). The initial choice of county pairs was based on a number of factors. First, geographic diversity was important. Because regions of the United States have distinct cultures, practices, and attitudes toward mental illness, substance abuse, and the criminal justice system, we included regions that represented these distinctions.

County pairs were also selected because they generally had comparable poverty rates, household income, rates of jailing, approaches to pre-release coordination and eligibility determination (Table 1.1). We also wanted paired counties to be similar to one another in demographic characteristics and to be in adjacent or near-adjacent states. Additionally, it was necessary to study county jails with enough volume to provide sufficient statistical power and data systems that were capable of tracking recidivism during our study period. In addition to these initial criteria, we conducted site visits and stakeholder interviews to understand the contexts in which Medicaid expansion occurred.

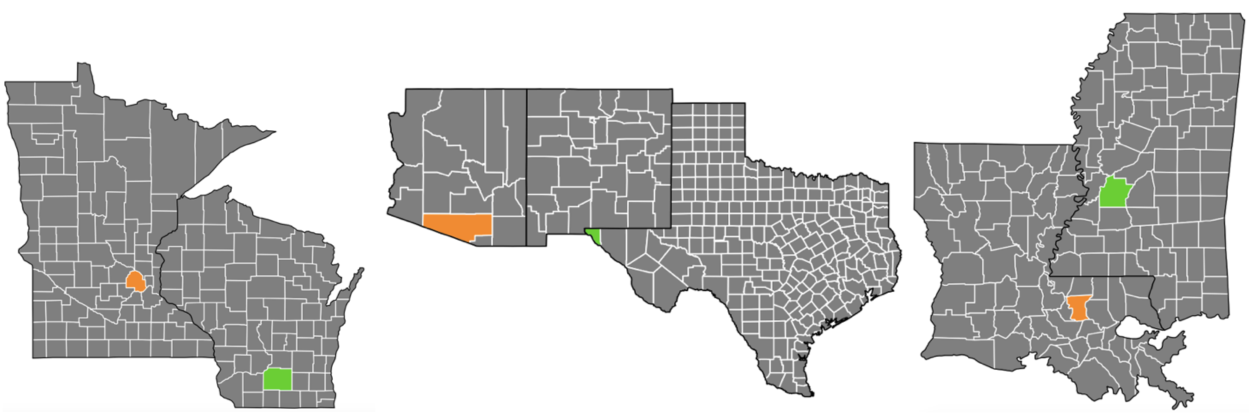
<u>Midwest</u>

In the Midwest, both study counties had similar proportions of the population under 18 years of age – 20 percent for Dane County and 22 percent for Hennepin County (Table 1.1). White, non-Hispanic individuals accounted for 85 percent of all residents in Dane and 77 percent in Hennepin.[16] More than 95 percent of the Dane County population had graduated high school at age 25 compared to 93 percent in Hennepin County.[17] The poverty rates in the two counties were identical at 11 percent, the median household incomes in 2018 were $62,865 in Dane County and $65,834 in Hennepin.[16] Both counties had population-adjusted jail rates in 2013 that were below the national average – 122 for Dane and 82 for Hennepin.[18]

Figure 1.1: Study counties
Panel A.



Panel B.
Midwest Southwest Southeast



Sources/Notes: SOURCES Authors' analysis of county-level booking data NOTES Medicaid expansion counties are in orange; non-expansion counties are in green. Counties were chosen to provide geographic diversity across the United States and based on demographic similarity. Counties also add to have adequate volume in the county jail to provide adequate power for the analyses and had to be able to track recidivist outcomes for the entire study period. Counties in the Midwest are Hennepin County, MN (expansion) and Dane County, WI (non-expansion). Counties in the Southwest are Pima County, AZ (expansion) and El Paso County, TX (non-expansion). Counties in the Southeast are East Baton Rouge Parish, LA (expansion) and Hinds County, MS (non-expansion).

In addition to similar demographic characteristics, both Wisconsin and Minnesota provide similar levels and types of Medicaid coverage for behavioral health conditions (Table A.1). Both counties have implemented programs to link the justice-involved with behavioral health care services. In Dane County, the jail has an Americorp volunteer on site three days a week to provide inmates with enrollment assistance into BadgerCare (Wisconsin's Medicaid program) prior to release. Additionally, Dane County, WI has several jail diversion initiatives, including electronic monitoring and reduced sentences for community program participation, that allows individual to remain in the community and receive community-based behavioral health services post-adjudication.

Similarly, Hennepin County, MN has a countywide Integrated Access Team that identifies, screens, and refers justice-involved individuals for treatment and assures continuity of care whether the individual is in or out of jail. Additionally, re-entry staff work with the community-based case managers to connect inmates to medication assistance in the community. In addition to re-entry programs, Hennepin County, MN has diversion programs for individuals with behavioral health needs.

In 2018 (which is after the study period in the Midwest County pair), Hennepin County opened a comprehensive social services facility that provides detoxification and mental health crisis services, employment counseling, and Medicaid eligibility assistance. All providers housed at the drop-in center accept Medicaid reimbursement for their services.

However, Dane County, WI is not a pure 'non-expansion' county. Wisconsin expanded its Medicaid program's eligibility to 100% FPL for non-disabled adults without dependents at the roughly the same time as the ACA's Medicaid expansion. Thus, Wisconsin had likely provided financial access to behavioral health services to the population at highest risk for criminal justice involvement. This may attenuate differences in outcomes between the Midwest pair.

<u>Southwest</u>

In the Southwest, we deliberately chose counties on the U.S./Mexico border with large Hispanic/Latino populations – 83 percent in El Paso and 36 percent in Pima.[16] An estimated 77 percent of El Paso's population had graduated high school by age 25 compared to 88 percent in Pima County.[17] The proportion of the population aged 18 or less in El Paso was 27 percent compared to 21 percent in Pima. Prior to expansion, El Paso had a poverty rate of 20 percent and median household income of $41,637, and Pima's were 19 percent and $46,162, respectively.[17] In 2013, El Paso had a jail rate of 324 per 100,000 population compared to Pima's rate of 289.[18]

Table 1.1: Comparison of county-level general population characteristics and pre-expansion sample characteristics

	Midwest		Southwest		Southeast	
	Hennepin N=22,146	*Dane* N=9,489	*Pima* N = 32,222	*El Paso* N = 26,576	*EBRP* N = 19,185	*Hinds* N = 7,639
General Population						
Population	1,223,149	523,643	1,010,025	835,593	446,753	242,891
% <18 years	22.2	21.0	23.0	27.9	22.8	25.0
% white	75.6	85.8	53.0	13.1	48.8	25.8
% HS diploma or equivalent	92.6	95.0	87.6	75.7	89.4	85.8
Median HH income	$65,834	$62,865	$46,162	$41,637	$49,285	$37,324
Pre-trial jail rate (per 100,000)	82	122	289	324	537	466
Jail-Involved Population						
Demographic Characteristics						
% female	22.7 (22.6 – 22.8)	24.2 (24.0 – 24.4)	26.1 (26.0 – 26.2)	27.5 (27.4 – 27.6)	24.9 (24.8 - 25.1)	17.7 (17.5 – 17.9)
% black	--	--	--	--	66.1 (66.0 – 66.3)	80.8 (80.6 – 81.0)
% Hispanic/Latino	--	--	41.7 (41.6 – 41.8)	81.6 (81.5 – 81.7)	--	--
Arrest Characteristics						
% parole violation	--	--	7.4 (7.2 – 7.6)	1.5 (1.4 – 1.6)	--	--
% misdemeanor	46.7 (46.2 – 47.3)	55.9 (55.1 – 56.8)	--	--	--	--
% felony	30.6 (30.1 – 31.1)	37.3 (36.5 – 38.1)	--	--	--	--

Sources/Notes: SOURCES General population data come from the American Community Survey, except for the jail rate. The proportion of individuals under age 18 and the proportion of individuals who are white are 2015 1-year estimates. The proportion who have a high school diploma or equivalent at by age 15 years, the median household income, and the percent living in poverty are 2015 5-year estimates. The population-standardized pre-jail incarceration rate is for 2013 and comes from the Vera Institute's trends on jail Jail-involved population comes from authors' analyses of arrest data from county jails. Sample sizes represent the pre-period number of individuals arrested in each county. The pre-expansion period for the Midwest and Southwest is from July 1, 2012 – December 31, 2013. The pre-expansion period for the Southeast is from January 1, 2015 – June 30, 2016. NOTES Values are means and (95% confidence intervals).

The two counties also have similar levels of coverage for behavioral health conditions in the Medicaid program (Table A.1). However, Texas limits the number of individual and group therapy sessions to 30 per person per year, whereas Arizona has no limit. Additionally, Texas does not provide coverage for residential psychiatric treatment, and Arizona does.

In terms of diversion and integration program, El Paso passed the Criminal Justice Mental Health Jail Diversion Collaboration Resolution in 2011. The purpose of this resolution was to engage community stakeholders in efforts to divert both pre-arrest and post-arrest individuals with behavioral health conditions, an estimated 30-35 percent of the jailed population, from the justice system to appropriate treatment.

The sheriff in Pima County also implemented initiatives to reduce criminal justice involvement among individuals with behavioral health conditions in 2011. At this time, Pima County, AZ built a Crisis Response Center and Behavioral Health Pavilion to provide integrated care to those experiencing behavioral health crises and help them avoid unnecessary incarceration. Additionally, the Pima County Sheriff's Department Mental Health Investigative Support Team coordinates responses with Pima County Behavioral Health and other law enforcement agencies when individuals with behavioral health conditions are involved in criminal justice events.

Southeast

In the Southeast, Hinds County and East Baton Rouge had large African-American populations – 47 percent in EBR and 72 percent in Hinds County prior to expansion.[16] More than 22 percent of EBR's population was under the age 18, while Hinds County had 25 percent in this age range in 2015.[16] More than 89 percent of the EBR population had graduated high school by age 25 compared to nearly 86 percent in Hinds County.[17] Both Southeast counties had pre-trial jail rates well above the national average (327 per 100,000 population) – 466 in Hinds and 537 in EBR in 2015/[18] The poverty rates were 27 percent and 19 percent in Hinds and EBR, respectively, and median household income was $37,324 in Hinds and $49,285 in EBR prior to expansion.[17]

Mississippi and Louisiana provide similar levels of coverage for behavioral health services, and neither state covers residential psychiatric treatment (Table A.1). Additionally, Louisiana does not cover methadone for the treatment of opioid use disorder, though Mississippi does.

Though Louisiana has expanded its Medicaid program under the ACA, little progress has been made in engaging local sheriffs to enroll eligible jail inmates in the state's expanded Medicaid program, even as part of re-entry planning. In addition to these challenges, efforts at implementing a program to facilitate Medicaid enrollment upon discharge have been stymied by the EBR's data system that does not maintain information by social security number but by an inmate ID unique to the jail.

Similar obstacles exist in Hinds County, MS. The county is under a Department of Justice consent decree for failing to provide adequate healthcare services in the jail system. In addition, little capacity exists in the community behavioral healthcare system to treat individuals leaving jail, and re-entry planning is severely limited. In fact, in 2015, a newly elected sheriff terminated a mental health diversion program, which demonstrated savings to the jail system in its first year of implementation.

These statistics and descriptions suggest that there are a number of similarities between the counties in each county pair, in terms of the nature of their local populations, the economic circumstances in each county, the likely pressures on the law enforcement system and efforts (or lack thereof) to integrate the criminal justice and behavioral healthcare systems. The qualitative data support our inferences within each county pair. However, we have also identified circumstances where changes to the state's Medicaid program (Dane County, WI) or criminal justice system practices (Hinds County, MS) may drive our results toward finding no changes.

Description of Data

Data were obtained from each county jail and include individual-level booking and release dates. Arrestees in each county are given unique identification numbers upon first arrest in the county and, thus, could be followed over the study period. The data spanned four years in each county – two before and two after Medicaid expansion. Additionally, each county-pair provided individual-level characteristics of arrestees, but these characteristics differed across counties and county-pair. Each county-pair provided at least the gender of the arrestee. In the Southwest and Southeast counties, the county provided the race/ethnicity of the arrestee. In the Southwest and Midwest counties, the dataset contained a measure of the severity of the crime. In the Southwest, we know whether the arrest was for a technical violation; and in the Midwest, we know whether the arrest was for a misdemeanor, felony, or some other charge.

Using these data, we created three distinct time periods over the 48 months of data – a lookback period, a pre-expansion observation period, and a post-expansion observation period. The six-month lookback period captures an individual's prior history with the criminal justice system. There were then 18 months in the pre-expansion study period, and 24 months in the post-expansion study period for each county-pair. Arizona and Minnesota expanded their Medicaid programs beginning January 1, 2014, while Louisiana expanded its Medicaid program beginning June 1, 2016. Observations are at the person-month level.

Cohort Construction and Outcome Measures

Within each county, individuals entered the study cohort when they are first arrested in the study period and are followed through the duration of the study period. If someone is arrested in the pre-period, we considered them 'at-risk' for the remainder of the study period. If someone is arrested only in the post-period, we considered them 'at-risk' for the remaining portion of the post-period but not for the pre-period. Individuals who were only arrested in the lookback period were excluded from the analysis. Arrestees who were booked into the county jail as a transfer from one prison to another (either state or federal) or as a "hold" for state/federal charges were also excluded from the sample.

Most of the literature on the relationship between Medicaid coverage or behavioral health services and recidivism relies on a single outcome – re-arrest. Prior evidence from Florida suggests that 30 percent of men and 20 percent of women are re-arrested within 18 months of being released [13]. Relying on this single measure of criminal activity may miss other ways in which access to behavioral health services may affect recidivism, such as reducing the frequency with which an individual interacts with the criminal justice. Thus, we not only focus on re-arrest rates, but we also include the number of arrests in our study.

Quasi-Experimental Design

Despite the high rate of Medicaid eligibility in the jailed population, our data do not indicate who was eligible for or enrolled in Medicaid coverage post-expansion or who received behavioral health services as a result of increased financial access to these services. Therefore, in an intent-to-treat analysis, we compare our outcomes of interest before and after Medicaid expansion between expansion and non-expansion counties for each county-pair. There is, however, strong evidence on the likely eligibility for Medicaid of the re-entry population in expansion states. The GAO estimated that for two states that expanded Medicaid (New York and Colorado), 80% to 90% on people in their prison systems were eligible for Medicaid in 2014 (US GAO 2014). Similarly, Massachusetts reports that 91% of people released from its correction system were eligible for Medicaid.

We conducted comparative interrupted time series (CITS) regression analysis. The CITS design takes advantage of an exogenous source of variation (i.e., the state's decision to expand Medicaid) between a treatment and comparison group[20-22] and allows for the estimation of both short and longer-term relationships between the outcome and exposure.

In both a difference-in-differences and CITS design, the counterfactual is constructed by assuming the change seen in the comparison group from the pre-period to the post period would be the same change seen in the treated group if not for the treatment. However, in a difference-in-differences design, the assumption is constrained so that the average change between the two groups is the same. In CITS, the counterfactual is constructed by assuming that change in linear trends in the comparison group is a good stand-in for the treated group without treatment. The only reason for deviation from these linear trends is the "interruption" in the treated group (i.e., all other reasons for deviation affect the treated and control group in the same way). Given the drivers of the outcome in this study (e.g., policing practices, criminal justice practices, and access to behavioral health services) and how they may vary between the counties, assuming linearity in the time trend of the outcomes in each of the two groups seems more reasonable than assuming that they change in the same average way over time.

Statistical Analysis

Because of heterogeneity in the criminal justice systems, health care delivery, legal dynamics (e.g. border policy) and populations across our study pairs, we chose to analyze each county pair separately. First, we computed pre-expansion means and variances for the likelihood of re-arrest and number of arrests in both the treatment and comparison counties for the full sample and for stratified samples in each county-pair. Because the policing, criminal justice, behavioral health, and health insurance systems treat individuals differently based on gender and race/ethnicity, stratification on these dimensions aims to ensure that we are making 'apples-to-apples' comparisons in our estimation strategy. Next, we computed pre-expansion monthly means of each outcome of interest to create pre-expansion trends.

If Medicaid expansion leads to reduced rate of re-arrest and number of arrests, then the composition of the pre- and post-period cohorts may differ in each observation period with individuals arrested in the post-period at higher risk, on average, compared to those arrested in the pre-period. We, thus, computed the number of individuals arrested and the number of arrests in both periods to check for compositional shifts in severity.

For the probability of re-arrest, we used a linear probability model. For the number of arrests (a count), we used ordinary least squares (OLS). We used the following general specification for each of the outcomes:

$$y_{ist} = \alpha + \beta_1\boldsymbol{X}_i + \beta_2\boldsymbol{X_i}TIME_t + \beta_3\boldsymbol{P}_{it} + \beta_4\boldsymbol{P_{it}}TIME_t + \lambda_1 TREAT_s + \lambda_2 POST_t + \lambda_3 TREND_{st} + \lambda_4$$
$$TREAT_s * TREND_{st} + \lambda_5 TREND_{st} * POST_t + \boldsymbol{\gamma} TREAT_s * POST_t + \boldsymbol{\eta} TREAT_s * POST_t * TREND_{st} +$$
$$QUARTER_t + \epsilon_{ist},$$

where: *i* is an individual in the cohort, *s* is the county, and *t* is month. $\boldsymbol{X_i}$ is a vector of time-invariant individual characteristics that are allowed to vary with the outcome over time, $\boldsymbol{P_{it}}$ is a vector of time-varying individual characteristics with effect, γ is the CITS estimator for the *level* shift in the outcome, η is the CITS estimator for the *trend* shift in the outcome and the parameter of interest, QUARTER is a vector of quarter fixed effects.

Time and individual-varying characteristics consist of charge severity – whether the arrest was a misdemeanor or felony (Midwest region) and whether the arrest was for a technical violation

(Southwest region). Time-invariant, individual-varying characteristics include the arrestee's gender, race/ethnicity (where available), and prior history with the criminal justice system. We interacted both the time-varying and time-invariant individual level covariates with the monthly time trend to adjust for the possibility that the relationship between the covariates and the outcome of interest is also time-varying. Within each county pair, we used a Bonferroni correction for multiple testing.

As noted, we stratify the CITS regression analyses in each county pair by gender and major local racial/ethnic groups. If baseline differences in racial or ethnic make-up were driving both arrest patterns and Medicaid expansion, then we would expect the causal estimates for these minority groups to differ significantly from that of the pooled sample.

Robustness Checks and Falsification Tests

A falsification test often used in CITS artificially places the empirical implementation date within a 'clean' study period (i.e., either the pre-period or the post-period). It could be that the outcome is noisy, and estimated effects might appear by chance. We varied the implementation date to three different months in the pre-period – months 7, 10, and 13 in the time series. Varying the implementation date within only the pre-period is preferable to strategies that do so within the whole study period, as the latter includes the treatment effect and may lead to detecting a spurious effect at a falsified intervention point. Our falsification strategy is also not ideal, as it cannot account for any anticipatory effects of Medicaid expansion – particularly in the Southeast county pair where other coverage provisions of the ACA had been in place for over two years prior to Medicaid expansion. Thus, we expect a weakening of the estimated impacts as the intervention date is moved away from the "true" date.

Additionally, because the pre and post-period observation windows are of different length, individuals are 'at-risk' for being arrested longer in the post-period than in the pre-period. This difference should be handled by the CITS design, since this pre and post observation period is shared by the intervention and comparison groups. Nonetheless, we checked for any effect of this pre-post discrepancy by conducting sensitivity analyses where we truncate the post-period to 18 months to balance the exposure time pre and post.

Finally, case studies raise general concerns about the validity of statistical inference. With only two clusters, estimating and accounting for intra-cluster correlation is not possible (both the within and between-group variance cannot be estimated in only two groups). Thus, typical econometric procedures for standard error adjustment (e.g., clustering, bootstrapping, permutation inference) are not feasible in this setting. We conducted the full sample CITS at the county/month year as check on statistical inference. Aggregating to the cluster/time level results in less biased standard errors even in a small number of clusters [23] but precludes individual-level covariate adjustment.

Results

In the pre-period, we found that the three Medicaid expansion counties arrested a greater number of individuals than the comparison counties (22,146 vs. 9,489 in the Midwest pair; 32,222 vs. 26,576 in the Southwest pair; and 19,185 vs. 7,639 in the Southeast pair), largely due to the fact that the expansion counties are larger in population than the non-expansion counties. Similarly, we found that there are more arrests in the pre-period in the Medicaid expansion counties compared to the non-expansion counties (33,082 vs. 13,405 in the Midwest pair; 46,569 vs. 31,966 in the Southwest pair; and 22,905 vs. 9,499 in the Southeast pair).

In the post-period, we found that a greater number of individuals were arrested in both the expansion and non-expansion counties of each pair (26,759 and 11,606 in the Midwest pair; 38,186 vs. 31,300 in the Southwest pair; and 20,371 and 9,300 in the Southeast pair) and that there were a larger number of arrests, as well (42,904 vs. 17,758 in the Midwest pair; 59,343 vs. 40,924 in the Southwest pair; and 25,311 vs. 12,166 in the Southeast pair). This suggested that the composition of the cohort may be changing from the pre-period to the post-period.

A primary concern with compositional shifts leading to fewer arrests was that this may imply that the least risky individuals (i.e., those with less impairing conditions) would be disproportionately enrolled in Medicaid and have greater financial access to behavioral health care services after expansion because they have the resources and functional capacity to do so. If individuals with fewer impairments were more likely to enroll in Medicaid and gain access to behavioral health services, our cohort would include a larger proportion of individuals more likely to have higher rates of recidivism and commit potentially more serious crimes. This scenario would lead to bias and overestimation of the effect of expansion on recidivist behavior. However, the direction of the observed change (i.e., an increase in the number of individuals arrested) implies that the marginal individual arrested in the post-period is likely not of higher risk than the marginal individual arrested in the pre-period.

Data on the types of arrests supports this interpretation. Of the arrests in the Southwest county pair, 3,933 and 3,394 were for a parole violation in the pre and post-period, respectively. Additionally, 22,951 (49.4 percent) of Midwestern arrests in the pre-period were for misdemeanors compared to 17,036 (50.0 percent) in the post-period, and 15,121 (32.5 percent) pre-period arrests and 11,191 (32.8 percent) post-period arrests were for felonies. While there appears to have been fewer post-expansion arrests, the composition of these arrests was almost identical, suggesting that the marginal individuals arrested after expansion are at roughly the same risk as individuals arrested prior to expansion.

Cohort and Arrest-Level Characteristics

In general, arrestees were much more likely to be men in each county pair (see Table 1.2). The proportion of the cohort that is female was greater in the non-expansion county (Dane, WI) in the Midwest pair (24.2 vs. 22.7%), in the non-expansion county (El Paso, TX) in the Southwest pair (27.5 vs. 26.1%), and in the expansion county (East Baton Rouge Parish, LA) in the Southeast pair (24.9 vs 17.7%). These differences are small in all except the Southeast pair. In terms of racial composition, the proportion of Hispanic/Latino arrestees in the pre-period in El Paso County was almost twice the proportion in Pima County, AZ (81.6 vs. 41.7%). Similarly, arrestees in the pre-period in Hinds County, MS were more likely to be black compared to arrestees in EBR, LA (80.8 vs. 66.1%).

In the Southwest county pair, a pre-period arrest in Pima County, AZ was nearly five times more likely to be for a parole violation compared to an arrest in El Paso County, TX (7.4 vs. 1.5% of all arrests), and pre-period arrests in Dane County, WI were more likely to be for misdemeanors (55.9 vs. 46.7%) or felonies (37.3 vs. 30.6%) compared to arrests in Hennepin County, MN. Although these baseline differences in arrestee demographics and arrest charge/severity were stable across the study period, they suggest that our comparison counties differ, at least in racial/ethnic composition.

Table 1.2: Pre-Medicaid expansion outcomes for full sample and stratified sample by gender and race/ethnicity

	Midwest		**Southwest**		**Southeast**	
	Hennepin N=22,146	*Dane* N=9,489	*Pima* N = 32,222	*El Paso* N = 26,576	*EBRP* N = 19,185	*Hinds* N = 7,639
Probability of re-arrest						
Full Sample	30.2 (29.6 - 30.8)	28.4 (27.5 - 29.3)	27.5 (27.0 - 28.0)	16.0 (15.6 - 16.5)	16.0 (15.5 - 16.5)	20.2 (19.3 - 21.1)
Male	31.9 (31.2 - 32.6)	30.0 (29.0 - 31.1)	29.0 (28.4 - 29.6)	17.0 (16.5 - 17.5)	17.3 (16.6 - 17.9)	21.6 (20.6 - 22.7)
Female	24.7 (23.5 - 25.9)	23.5 (21.8 - 25.2)	23.3 (22.4 - 24.2)	13.4 (12.6 - 14.2)	12.2 (11.3 - 13.1)	13.7 (11.8 - 15.5)
Black	--	--	--	--	17.5 (16.8 - 18.2)	20.9 (19.9 - 22.0)
Hispanic/Latino	--	--	26.1 (25.4 - 26.9)	16.0 (15.5 - 16.5)	--	--
Number of arrests						
Full Sample	1.54 (1.52 - 1.55)	1.46 (1.45 - 1.48)	1.50 (1.49 - 1.51)	1.22 (1.21 - 1.23)	1.21 (1.20 - 1.22)	1.27 (1.26 - 1.28)
Male	1.57 (1.56 - 1.59)	1.50 (1.47 - 1.52)	1.54 (1.52 - 1.56)	1.23 (1.22 - 1.24)	1.23 (1.22 - 1.24)	1.29 (1.27 - 1.31)
Female	1.42 (1.39 - 1.44)	1.36 (1.33 - 1.39)	1.40 (1.38 - 1.42)	1.18 (1.17 - 1.20)	1.15 (1.14 - 1.16)	1.17 (1.14 - 1.20)
Black	--	--	--	--	1.23 (1.22 - 1.24)	1.28 (1.26 - 1.30)
Hispanic/Latino	--	--	1.44 (1.42 - 1.45)	1.22 (1.21 - 1.23)	--	--

Sources/Notes: SOURCES Authors' analyses of arrest data from county jails. The pre-expansion period for the Midwest and Southwest is from July 1, 2012 – December 31, 2013. The pre-expansion period for the Southeast is from January 1, 2015 – June 30, 2016. NOTES Values are means and (95% confidence intervals). We only provided descriptive statistics for the predominant racial/ethnic group in each county pair, where available. No race/ethnicity data were available in the Midwest, and no arrest-level data were available for the Southeast. '-' indicates that variable was unavailable or not analyzed for the given county pair. Sample sizes are for the full sample – stratification based on gender or race/ethnicity reduces the sample size.

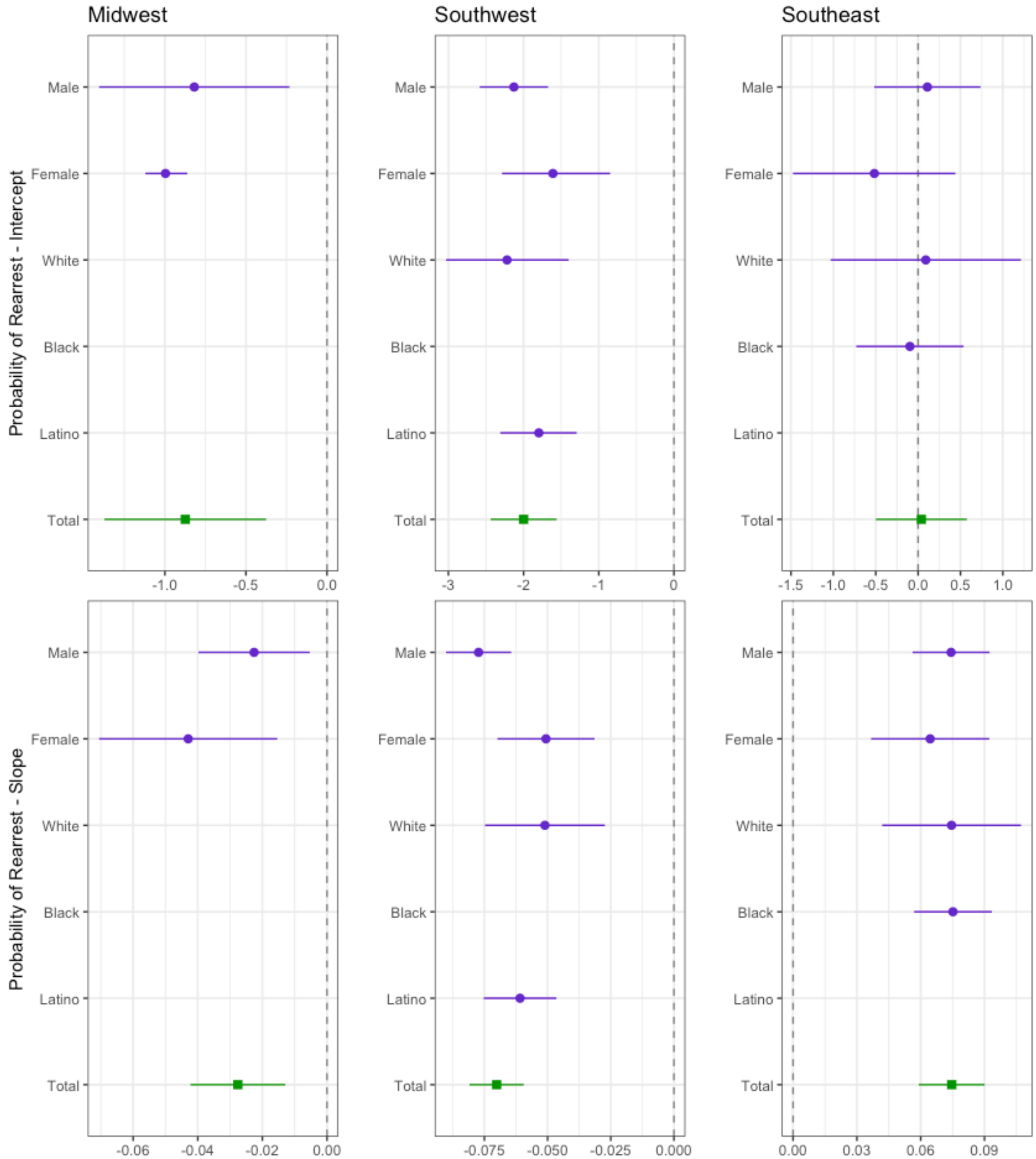
Probability of Re-Arrest

The probability of re-arrest in the pre-period was higher in Hennepin County (expansion; 28.2 percent; 95% CI: 27.6 – 28.9) compared to Dane County, WI (non-expansion; 26.6 percent; 95% CI: 25.7 – 27.6) and in Hinds County, MS (non-expansion; 20.2 percent; 95% CI: 19.3 – 21.1) compared to EBR, LA (expansion; 16.0; 95% CI: 15.5 – 16.5; Table 1.2). The greatest difference in the likelihood of re-arrest was in the Southwest county pair, where Pima County, AZ's (expansion) rate was 11.5 percentage points higher than that of El Paso County, TX (non-expansion). When stratified by gender and race/ethnicity in each county pair, the pattern of results was the same. The likelihood of re-arrest among male and female arrestees was higher in Hennepin County, MN; Pima County, AZ; and Hinds County, MS. Among black arrestees in the Southeast, the probability of re-arrest was 3.4 percentage points higher in Hinds County, MS than in EBR, LA. Additionally, the probability of re-arrest among arrested Hispanic/Latino individuals in the pre-expansion period was 8.1 percentage points higher in Pima County, AZ compared to the El Paso County, TX. While differences in the pre-period trend may be cause for concern in a difference-in-differences analysis, baseline or pre-period trend differences in the outcomes do not invalidate causal inferences for CITS, which identifies the causal effect off the change in level and trend of the outcome in the comparison group from the pre-period to the post-period.

After Medicaid expansion, the short-term probability of re-arrest (i.e., the level change) declined by a statistically significant amount in the Midwest and Southwest county pair (Figure 1.2). The largest decline was in the Southwest county pair where the probability of being re-arrested declined by 2.0 percentage points (95% CI: -2.4, -1.6) in the month following expansion, which is a 7.27 percent decrease from the pre-period mean.

The estimated effect of Medicaid expansion on rates of re-arrest in Pima County relative to El Paso County was greatest among white (-2.2 percentage points in the month after expansion; 95% CI: -3.0, -1.4) and male arrestees (-2.1 percentage points; 95% CI: -2.6, -1.7). Reductions in recidivism were similar among female and Hispanic/Latino arrestees compared to the full sample.

Figure 1.2: Changes in the probability of re-arrest between Medicaid expansion and non-expansion counties



Sources/Notes: SOURCES Authors' analyses of arrest data from county jails. The study period for the Midwest and Southwest is from July 1, 2012 – December 31, 2015. The pre-expansion period for the Southeast is from January 1, 2015 – June 30, 2018. Observations are at the person-month level. NOTES Estimates are from comparative interrupted time series regressions. Regressions are linear probability models. Each full sample regression is adjusted with gender and prior contact with the criminal justice system. The Midwest pair also adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African-American and the interaction of this variable with the monthly time trend. Stratified regression analyses in each county pair adjust for these same covariates except for the variable that the sample was stratified on.

In the Midwest county pair, the probability of an individual being arrested declined by 0.87 percentage points in the month after expansion (95% CI: -1.4, -0.4 percentage points) in Hennepin County, MN relative to Dane County, WI, which amounts to a 2.9 percent decrease from the pre-period rate of re-arrest. Similar declines in the probability of re-arrest were experienced among male and female arrestees. In the Southeast county pair, Medicaid expansion was not associated with changes in the probability of re-arrest.

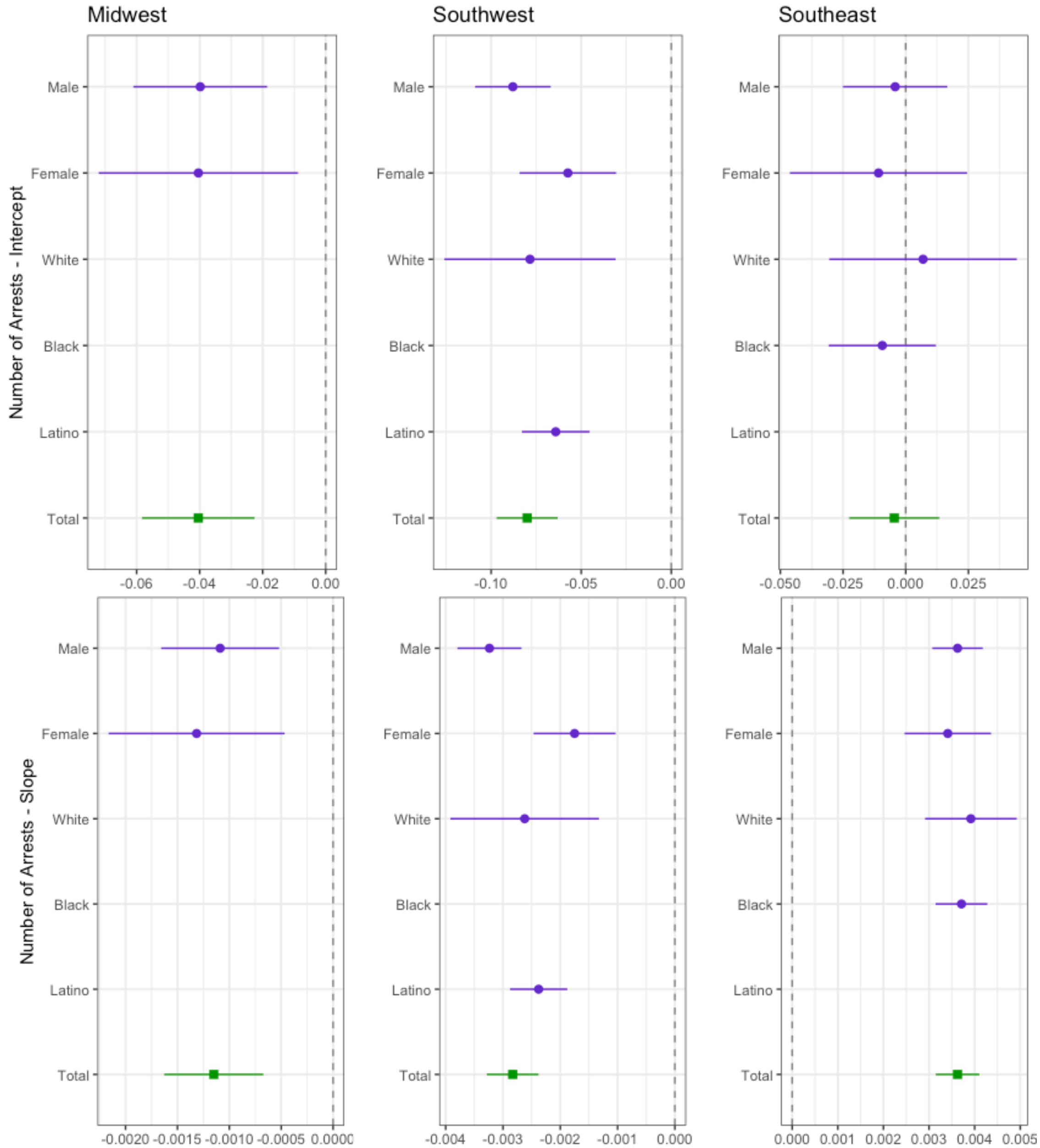
In addition to the immediate effects of Medicaid expansion on rates of re-arrest, we found sustained decreased rates of recidivism for arrestees in the Midwest (-0.03 percentage points per month; 95% CI: -0.04, -0.01) and Southwest (-0.07 percentage points per month; 95% CI: -0.06, -0.08) county pairs (Figure 1.2). The effects were similar for all subgroups examined in these two county pairs. While there was no change in the level of the outcome in the Southeast, arrestees in EBR, LA experienced a change in trend relative to their counterparts in Hinds County, MS (0.07 percentage points per month; 95% CI: 0.06, 0.09). This change in the long-term probability of re-arrest is similar for male, female, and African American arrestees.

By extrapolating the level and slope change to the end of the study period, we found that Medicaid expansion led to an average decline in the probability of re-arrest of 1.49 percentage points (an 4.92 percent decrease) in Hennepin Co., MN compared to Dane Co., WI; an average decline of 3.6 percentage points (an 13.1 percent decrease) in Pima Co., AZ compared to El Paso Co., TX. In the Southeast, a differential trend increase resulted in a 1.61 percentage point increase (a 10.1 percent increase) in the probability of re-arrest in EBR relative to Hinds Co., MS.

Number of Arrests

Like the descriptive statistics for the likelihood of re-arrest, the average number of arrests in the pre-period was higher in Hennepin County, MN than in Dane County, WI (1.54 vs. 1.46). The average number of arrests for men and women in the expansion county was also higher than that of the non-expansion county in the Midwest pair (Table 1.2). Similarly, the average number of arrests was higher in Pima County, AZ compared to El Paso County, TX (1.50 vs. 1.22) in the full sample and among all stratifications.

Figure 1.3: Changes in the number of re-arrests between Medicaid expansion and non-expansion counties



Sources/Notes: SOURCES Authors' analyses of arrest data from county jails. The pre-expansion period for the Midwest and Southwest is from July 1, 2012 – December 31, 2015. The pre-expansion period for the Southeast is from January 1, 2015 – June 30, 2018. Observations are at the person-month level NOTES Estimates are from comparative interrupted time series regressions. Regressions are Poisson regression models with an identity link. Each full sample regression is adjusted with gender and prior contact with the criminal justice system. The Midwest pair also adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African American and the interaction of this variable with the monthly time trend. Stratified regression analyses in each county pair adjust for these same covariates except for the variable that the sample was stratified on. [§] denotes that p-value is not statistically significant after Bonferroni adjustment for multiple comparisons.

Arrestees in Hinds County, MS (non-expansion) had a higher average number of arrests compared to the EBR, LA (1.27 vs. 1.21), which was also the case for male, female, and black arrestees in the Southeast pair.

Overall, Medicaid expansion resulted in a decline in the average number of arrests per person in the Midwest (-0.04; 95% CI: -0.02, -0.06) and Southwest (-0.08, 95% CI: -0.06, -0.10) expansion counties compared to the respective non-expansion counties in the month after Medicaid expansion (see Figure 1.3). In both the Midwest and the Southwest, the change in the number of arrests did not differ significantly in the stratified samples compared to the full sample estimate.

Similar to the probability of re-arrest, the change in the number of arrests in the Southeast county pair is close to zero and not statistically significant (-0.004; 95% CI: -0.02, 0.01), suggesting that there's no immediate impact of Medicaid expansion and recidivism.

The longer-term effects of Medicaid expansion on the average number of arrests per person were similar to the pattern of results seen in the analyses with the probability of re-arrest. The average number of arrests per person/month decreased more in Hennepin County, MN compared to Dane County, WI (-0.001; 95% CI: -0.001, -0.002). The average number of arrests per month decreased for the full sample, male, female, and Hispanic/Latino arrestees (roughly 0.003 arrests per month) in Pima County, AZ compared to El Paso County, The average number of arrests per month after Medicaid expansion increased for all groups in EBR, LA compared to the rate in Hinds County, MS (0.004 arrests per month; 95% CI: 0.003, 0.004).

Via linear extrapolation of the level and slope change, we found that Medicaid expansion led to an average decline of 0.1 arrests per person in Hennepin Co., MN compared to Dane Co., WI, and an average decline of 0.2 arrests in Pima Co., AZ compared to El Paso Co., TX two years after Medicaid expansion (the end of our study period). In terms of percent changes, this is a 5.8 percent decrease in the Midwest, and a 13.3 percent decrease in the Southwest. Taking into account only the trend increases in the Southeast, Medicaid expansion resulted in 0.2 more arrests per person in EBR compared to Hinds Co., MS, which is equivalent to a 12.2 percent increase.

Falsification Tests

As suggested above, falsification tests, such as these, may not be ideal when assessing whether we are isolating the causal effect of Medicaid expansion on recidivist behavior. These tests did not allow for us to specify any ramp up to Medicaid expansion, particularly in the Southeast county pair where the ACA's other coverage provisions had been in place for more than two years prior to Medicaid expansion. With these caveats, the falsification tests suggested that we were capturing the causal effect (see Tables A.2 and A.3). In all three county pairs, we did not detect any change in the intercept at the false implementation points for the probability of re-arrest or the number of arrests. However, we did detect a differential change in the slope of the line just prior to expansion for both outcomes in the Midwest and Southwest.

Limitations

There are several limitations in our study. First, our data consisted of booking data. We did not know who gains health insurance under Medicaid expansion. Thus, we conducted an intent-to-treat analysis. Additionally, we could not differentiate between individuals who are lost to follow up vs. never re-arrested. We assumed that rates of attrition are similar over time and across the counties in each county pair.

Second, from interviews and site visits, we identified changes to the behavioral health and criminal justice systems that add context to our results, but also highlighted that in some cases, changes may not be attributed entirely to Medicaid expansion. Hinds County, MS discontinued a mental health diversion program in late 2015 after the current sheriff was elected in August 2015 (expansion happened in LA in June 2016), despite evidence that the program saved the county $250,000 in the year after implementation.

Third, baseline differences in the composition of each county's arrested population and the overall arrest activities may suggest that these are not perfect comparison counties. We chose the counties based on their similarity on the county's full demographic characteristics, rather than the characteristics of the jailed population. To the extent that the characteristics of the jailed populations and differential policing/arrest activity were stable over time, our design netted out these differences.

Discussion

Overall, Medicaid expansion reduced both the probability of re-arrest and the number of arrests in two of the three county pairs. In the Midwest and Southwest, the estimated effects at two years post-expansion were consistent with estimates from other studies on the relationship between access to healthcare services and recidivism (between a 5 and 13 percent decrease). Additionally, the mixed nature of the findings (an increase in the Southeast) is also consistent with prior literature.

These estimates are similar to other initiatives to reduce recidivism. Adult drug courts reduce recidivism rates by roughly eight percent,[24] One meta-analysis of educational and vocational training programs found that these programs were associated with a 13 percentage point decrease is recidivism,[25] while another estimated these programs to reduce recidivism by seven to nine percent,[24] However, many of the studies included in these meta-analyses suffer from the same selection issues as found in previous studies on the effect of increased access to healthcare services on recidivism.

However, our estimates might be smaller than the true effect of Medicaid expansion on recidivism. We did not measure first-order effects – health insurance coverage and access to care – of Medicaid expansion in the jail-involved population nor did we measure the change in recidivism in individuals who obtained Medicaid coverage and subsequent behavioral health treatment. If we were able to conduct a treatment-on-the-treated analysis, then our estimates would scale by the proportion of individuals who enrolled in Medicaid coverage due to expansion. If we were to use the previously published proportion of jail-involved individuals that would be eligible for Medicaid expansion,[2] scaling our estimates would suggest that expansion is associated with a 16-32 percent reduction in the rate of recidivism.

In the Southeast, we failed to detect a change in the level in EBR relative to Hinds Co with the implementation of Medicaid expansion, and the change in the slope resulted in an overall increase in the probability of re-arrest and number of arrests 24 months after expansion. This could be a result of changes to behavioral health and criminal justice practices required by the federal consent decree in Hinds County, MS and the lack of integration and coordination between these two systems in East Baton Rouge Parish, LA.

Additionally, we stratified our analyses by race and gender to address within-pair heterogeneity and make comparisons across individuals who are treated more similarly by the policing, healthcare, and criminal justice systems. The estimates for these stratified groups were either the same size or larger than the full sample, which strengthens our inferences. Moreover, our qualitative analysis of the efforts occurring in each of the counties (see the Methods section) allowed us to draw out the contexts that may make Medicaid expansion more or less effective in reducing recidivist behavior. Indeed, our results mirrored the previous literature – enhanced financial access to healthcare services *contributes* to a reduction in recidivism.

Other contributing factors include coverage of evidence-based treatment for mental illnesses and substance use disorders; adequate capacity in the community's behavioral health treatment system; the provision of mental illness and substance use disorder treatment services in jail, both pre-adjudication and while incarcerated; the coordination and continuity of care across the criminal justice and behavioral healthcare systems; the implementation of jail diversion programs that keep individuals with mental illness or substance use disorder from entering the criminal justice system; and the availability of other social programs, such as supportive housing and employment, that improve the social status of individuals with mental illness or substance use disorder. Reducing rates of re-arrest, particularly for individuals with severe mental illness and/or substance use disorders, requires coordinated efforts between multiple social service systems and increased integration of those systems could be an important policy lever to increase time in the community.

Chapter 2:
Do Medicaid retroactive eligibility waivers compel beneficiaries to enroll in coverage when eligible?

Co-authored with Benjamin D. Sommers

Dissertation Advisor: Richard G. Frank

Carrie E. Fry

Abstract

The primary goal of Medicaid coverage is to provide financial protection against large health care expenses for eligible low-income populations. One way that Medicaid provides this financial protection is the retroactive eligibility provision, which provides Medicaid coverage for the 90 days prior to enrollment for those who were eligible for Medicaid during this period. In recent years, a number of states have received approval from the Centers for Medicare and Medicaid Services to waive the retroactive eligibility provision. In doing so, states have argued that the removal of retroactive eligibility will result in greater Medicaid enrollment because patients won't wait until they get sick to sign up; in turn, advocates of the waiver claim it will improve coverage continuity and health outcomes for beneficiaries. Despite these claims, there has been little research examining the impact of retroactive eligibility waivers. In this paper, we examine the relationship between retroactive eligibility waivers and monthly Medicaid enrollment for four states that implemented a retroactive eligibility waiver between 2016-2018 using a difference-in-differences analysis. We found that implementation of a retroactive eligibility waiver did not impact average Medicaid enrollment in the year after implementation. When we estimated a dynamic treatment effect for each month after implementation, we found no meaningful differences in our comparative case study analyses. When we pooled the states together to address potential power issues, we found a statistically significant 10% decrease in Medicaid enrollment at months five and six after implementation. In supplemental analysis, we obtained similar results using the synthetic control method. Future research should further examine the impact of retroactive eligibility waivers on the health and financial status of newly enrolled Medicaid beneficiaries, enrollment continuity, and the financial status of health care providers.

Introduction

Retroactive eligibility, a statutory requirement of a state's Medicaid program, provides Medicaid coverage for the 90 days prior to application, given that the beneficiary would have been eligible for Medicaid in those 90 days and that the medical expenses incurred during that time were for services covered by Medicaid. Over half of uninsured children and one quarter of uninsured adults are eligible but unenrolled in Medicaid.[26] Because these individuals may only enroll at or after a health care encounter, the goal of retroactive eligibility (retroactive eligibility) is to provide financial protection for large expenses that may have been incurred prior to enrollment. As a concrete example, imagine a parent who is severely injured in a car accident and uninsured but eligible for Medicaid coverage. The family may need time to gather application materials and apply for Medicaid coverage, but the beneficiary needs health care services immediately. Retroactive eligibility acts a stop gap measure to provide financial reimbursement of these services once the patient's application for Medicaid is approved. In addition to protecting the beneficiary, retroactive eligibility also provides compensation to the health care facility that provides the services in this situation.

For over two decades, states have received permission to limit or waive retroactive eligibility in the Medicaid program. In its push to give states more flexibility to design their Medicaid programs, the Centers for Medicare and Medicaid Services (CMS) under the Trump administration has approved eight Section 1115 waivers that eliminate retroactive eligibility. Prior to this, states received approval to eliminate retroactive eligibility for a limited population of beneficiaries. Most notably, a number of states recently received approval to waive this provision for Group VIII beneficiaries under the Affordable Care Act (ACA). Most recently, however, Iowa received permission to waive the retroactive eligibility provision for all covered populations except pregnant people and women under one year of age, the most expansive removal to date.

The removal of retroactive eligibility is largely under-studied. No state with a Section 1115 waiver has provided a formal evaluation of waiving retroactive eligibility. In published guidance, CMS provides a sample logic model and set of hypotheses for states to consider when evaluating this change to the Medicaid program.[27] In approving waivers to elimination retroactive eligibility, CMS has argued that this change will increase enrollment by people when they are healthy relative to those with access to retroactive eligibility. Additionally, CMS suggests that a retroactive eligibility waiver will incentivize early enrollment and greater enrollment continuity, since beneficiaries will have no access to coverage during lapses or gaps in enrollment.

While there is little evidence on the impact of retroactive eligibility waivers, there is considerable evidence on Medicaid's relationship with health and financial outcomes. Though not all studies find positive effects of gaining Medicaid on specific health outcomes,[28,29] a large body of evidence shows improvements across a range of health outcomes from gaining Medicaid,[29–33] including multiple studies from both before and after the ACA demonstrating reductions in mortality.[34,35]

The evidence is even stronger that Medicaid results in improved financial status for beneficiaries. Prior evidence from coverage expansions demonstrates that Medicaid is an effective way to reduce the financial burden of medical costs in low-income populations[30,36–39] and losing Medicaid coverage results in less financial security.[40] Retroactive eligibility is likely one mechanism through which Medicaid may reduce burdensome medical bills or alleviate other medical care-related financial hardships by providing coverage for expenses incurred prior to Medicaid enrollment. Thus, it is possible that a retroactive eligibility waiver results in greater financial hardship. Indeed, an informal, interim evaluation in Indiana found that 13.9% of enrollees would have previously been eligible for the program and incurred over $1,500 in medical expenses in the retroactive eligibility period.[41]

However, there is currently no literature on the impact of a retroactive eligibility waiver on enrollment. The logic model published by CMS would suggest that eliminating retroactive eligibility would result in increased enrollment after the waiver is implemented.[27] However, it is also possible that a retroactive eligibility waiver would result in decreased enrollment, as prospective beneficiaries may see less benefit or value in Medicaid coverage without retroactive eligibility or health care providers may be less likely to enroll individuals since there is no possibility of reimbursement for services previously rendered. An informal, prospective analysis in Iowa suggests the latter – the retroactive eligibility waiver could result in 3,344 fewer enrollees per month, resulting in a decline in annual Medicaid spending of $36.8M ($27.1M federal share, $9.7M state share).[41] This decline represents roughly 0.5% of the average monthly enrollment in 2018, the majority of whom be children and Medicaid expansion adults.

In this paper, we evaluate the relationship between a retroactive eligibility waiver and Medicaid enrollment to estimate one of the 'short-term' outcomes hypothesized by CMS. We do so with four comparative case studies of states that implemented a retroactive eligibility waiver after the ACA's implementation using monthly Medicaid enrollment data from CMS and quasi-experimental design.

Data and Methods

Our outcome of interest was monthly Medicaid enrollment, which is publicly-available from CMS.[42] Since September 2013, CMS has required that states provide monthly data on the number of applications and determinations for Medicaid and CHIP, as well as monthly enrollment in these programs. For each month, CMS provides provisional and updated data. For each month in the study period, we obtained the updated total number of enrollees in Medicaid in each state. We log-transformed these data for our analyses. Details of a state's retroactive eligibility waiver, including the population affected and the conditions of the waiver, were obtained from the Kaiser Family Foundation.[43] The date of implementation for each retroactive eligibility waiver was taken from the state's approved Section 1115 waiver that contained the retroactive eligibility provisions.

To adjust for potential time-varying confounders in DID regression analysis, we used data from publicly available sources as covariates. We obtained the federal medical assistance percentage (FMAP) value for each state and fiscal year combination during our study period from the Office of the Assistant Secretary for Planning and Evaluation.[44] The FMAP is the proportion of Medicaid expenditures that the federal government covers. A state's FMAP is determined by its per capita income and is statutorily set between 50% and 83%. In FY 2020, the highest FMAP rate is for Mississippi at 76.98%. The seasonally adjusted monthly unemployment rate for each state was obtained from the Bureau of Labor Statistics.[45] Using micro data from the American Community Survey, we obtained the proportion of each state's population that are of women of child-bearing age (19-44).[46] States with a higher proportion of the population in this demographic likely have higher Medicaid enrollment, as Medicaid covers just under half of the births nationwide.[47] We also obtained annual eligibility levels for Section

1913 parents (i.e., parents who were eligible for Medicaid prior to the ACA's expansions) of the Medicaid program from the Kaiser Family Foundation.[48-51]

Study Period and Comparison States

Table 2.1 lists the states we include in our study sample, as well as details about each state's retroactive eligibility waiver, the comparison states used, and the time period for each comparative case study. Our study sample consists of states that implemented an approved retroactive eligibility waiver since the ACA's Medicaid expansion went into effect on January 1, 2014, but the implementation of this waiver did not coincide with the implementation of another major Medicaid policy. For instance, Indiana's retroactive eligibility waiver for its adult expansion population was implemented on the same date as the state's Medicaid expansion. We cannot parse out the differential effect of retroactive eligibility on enrollment from the large increase due to Medicaid expansion, so we excluded Indiana from our sample.

Table 2.1: Retroactive eligibility waiver states and study details

State	Implemented	Waiver applies to	Comparison states	Study period
AR	1/5/2017	Expansion adults	IN, KY, LA, NM, WV	1/2016 - 12/2017
FL	11/30/2018	All adults above 21	AL, GA, MS, SC, TN	11/2017 - 10/2019
IA	11/1/2017	All state plan enrollees except pregnant women, children under 1	IN, MI, MN, OH, SD	11/2016 - 10/2018
NH	11/30/2018	Expansion adults	CT, MA, NY, RI, VT	11/2017 - 6/2019

SOURCES/NOTES: **Sources** States with approved retroactive eligibility waivers were obtained from the Kaiser Family Foundation. Details on the population covered by a retroactive eligibility waiver were also obtained from KFF. Implementation dates were obtained by reading the approved CMS waivers obtained from the federal Medicaid website. **Notes** Our sample states include states with retroactive eligibility waivers that were implemented after the ACA's Medicaid expansion began on January 1, 2014 and where the implementation date was not the same as the state's Medicaid expansion date. Comparisons states were chosen based on geographic proximity to the retroactive eligibility state and Medicaid expansion status. The post-period of the New Hampshire analysis was truncated to not overlap with the implementation of New Hampshire's Medicaid work requirement.

For each comparative case study, our study period consisted of 12 months prior to implementation of a retroactive eligibility waiver and up to 12 months after the retroactive eligibility waiver was implemented. In Arkansas, we truncated the post-implementation period to 11 months because Arkansas received approval for its Medicaid work requirement in January 2018, which may have resulted in a decline in Medicaid enrollment despite the work requirement not being implemented until June 2018. We do a similar truncation in New Hampshire, which implemented its work requirement in June of 2019. In both Arkansas and New Hampshire, the retroactive eligibility provision was included in the state's work requirement waiver. Both of these waivers have been invalidated by federal courts. This reinstated retroactive eligibility in New Hampshire in July 2019 (we truncate our post period to June 2019) and in Arkansas in March 2019 (our study period ends in December 2017). New Hampshire's legislature passed a law requiring that the state move forward with new retroactive eligibility waiver in 2019.[43]

Comparison states were chosen based on geographic proximity, whether the treated state was a Medicaid expansion state, and whether the comparator implemented major Medicaid reforms during the study period. For each retroactive eligibility state, we selected five comparator states.

Difference-in-Differences Analysis

To quantify the impact of a retroactive eligibility waiver on Medicaid enrollment, we use a difference-in-differences (DID) design. Violations of the parallel trends assumption of DID may suggest that the counterfactual assumption made in DID is implausible. However, the null hypothesis that the trends are different can only be rejected – the alternative hypothesis that they are parallel cannot be confirmed.[52] One way to relax the parallel trends assumption is to incorporate group-specific linear trends. Here, the counterfactual is constructed by extrapolating the growing linear difference between the two groups in the pre-period and then assume that any level change in the comparison group represents the change in the retroactive eligibility group. The counterfactual assumption of DID and details of our model specifications are in Appendix B.

Because the states in our sample implement retroactive eligibility waivers at different times, pooling the four retroactive eligibility states and their respective comparison states results in a DID treatment effect that is the weighted average of all possible 2x2 combinations.[53] Additionally, the retroactive eligibility waivers affect in each state may produce differential treatment effects, as they affect different proportions of the Medicaid population. For instance, a larger proportion of the Medicaid population in Iowa is subject to the waiver compared to the population in Florida. We analyze each retroactive eligibility state separately in four comparative case studies as our main specification but conduct a pooled analysis as a sensitivity analysis. All regression models are adjusted using the covariates discussed above, as well as a dummy variable for whether or not the state allows for adults to be continuously enrolled in Medicaid for 12 months (vs. requiring renewal every six months). In the Arkansas case study, we also include a binary (i.e., dummy) variable for July 2016 in Louisiana, when the state expanded its Medicaid program.

DID with longer time series and clustered data results in standard errors that are too small.[54] Initial recommendations to handle the intra-cluster correlation and serial autocorrelation introduced by these data structures are to use sandwich estimators to adjust the standard errors. Yet, even clustered standard errors can be mis-estimated when the number of clusters are small. Because we have a small number of clusters in each comparative case study, we use a wild-cluster bootstrap technique to adjust the standard errors.[55,56] In a small number of clusters, this procedure results in fewer instances of Type I error than other standard error estimation procedures.[55,57]

Supplemental Analyses

In addition to the DID analysis, we used a synthetic control method to construct a synthetic comparison group for each of our four retroactive eligibility states using the same study period as used in the respective DID analysis. Rather than using a reasonable set of matched states as the comparison states, we allow for SCM to construct the synthetic comparison from all US states with data for the entire study period. As is standard with SCM analysis[58,59], we omit other states with major policy changes during the study period that could affect the outcome. For this reason, we omit the three other retroactive eligibility states for each case study. We also omit New Mexico, as it does not have published Medicaid enrollment data for all months in the study period. Details of our SCM analysis, including the covariates used, are in Appendix B.

Limitations

As with all quasi-experimental studies, our analysis has a number of limitations. First, DID and SCM rely on untestable causal assumptions. We conduct sensitivity analyses with both designs to assess the plausibility of these assumptions but cannot confirm that these assumptions are met. We conduct DID with group-specific trends to test the plausibility of the parallel trends assumption. The details of the SCM sensitivity analyses are in Appendix B.

Second, a number of states, including Florida, experienced large declines in the number of children enrolled in Medicaid in 2018.[60] This could bias our results either toward zero if the anticipated effect is a decrease in enrollment or toward a larger effect if the expected change is an increase. In Florida, using lagged outcomes as covariates in the SCM analysis could help mitigate this bias, as these changes occur in the pre-period. In Iowa, where 2018 is the post-period, we do not include any states that experienced large declines in child enrollment in Medicaid in our comparison group. The enrollment data produced by CMS does not break down Medicaid enrollment by eligibility category, which prevented us from sub-setting the enrollment data to only the beneficiaries targeted by the retroactive eligibility waiver. Thus, enrollment trends in groups not affected by the waiver, such as the decline in child enrollment in 2018, may be biasing our estimation strategy in the Florida comparative case study.

Lastly, with limited data, our analysis could be underpowered to detect the effect size that we might expect to see. The one prospective analysis in Iowa suggests that the implementation of a retroactive eligibility waiver is associated with a decline in monthly Medicaid enrollment of roughly 3400 enrollees or 0.5% of the average monthly enrollment. If proponents of retroactive eligibility waivers expect an increase in enrollment of roughly the same size, we may also be under-powered to detect the expected effect.

Results

Prior to the implementation of a retroactive eligibility waiver, the chosen DID comparison states are generally similar to the retroactive eligibility states on selected pre-period state-level demographic and policy characteristics (Table B.1). Of notable exception is the unemployment rate in Arkansas (4.0%) versus that of its comparison states (5.7%) and the FMAP in Florida, which is 8.9 percentage points lower than its comparison states.

Difference-in-Differences

Using Equation 1, we found no differential change in monthly Medicaid enrollment in the retroactive eligibility states compared to the selected comparison states (Table 2.2). The direction and magnitude of the treatment effects were mixed, though not statistically significant. Retroactive eligibility resulted in a 6.0 percent increase in enrollment in Iowa and a 5.5 percent decrease in enrollment in Florida. The effects in Arkansas and New Hampshire were smaller – we found a 2.6 percent increase in enrollment in Arkansas and a 0.4 percent decrease in Medicaid enrollment in New Hampshire after the implementation of a retroactive eligibility waiver compared to their respective comparison groups. When we adjusted for the possibility of differential pre-trends (i.e., parallel pre-trend violations) using Equation 2, we found that the treatment effect estimates did not significantly change.

Table 2.2: Change in Medicaid enrollment after retroactive eligibility waiver implementation

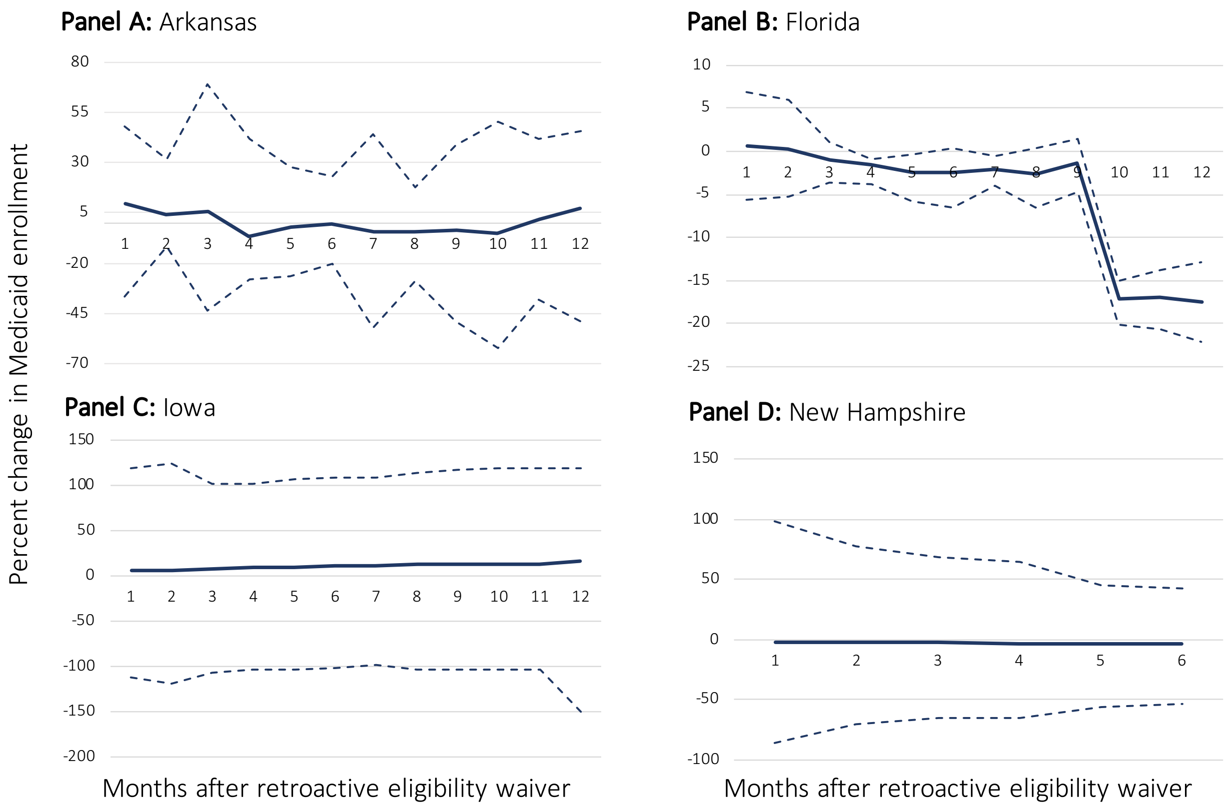
	Without group trends		**With group trends**	
	% change in enrollment	*Bootstrapped 95% CIs*	*% change in enrollment*	*Bootstrapped 95% CIs*
Arkansas (N = 144)	-0.06	-15.5, 16.9	4.0	-25.8, 46.2
Florida (N = 144)	-6.6	-26.4, 17.3	-3.8	-13.8, 7.3
Iowa (N = 144)	9.6	-82.0, 59.5	6.6	-47.5, 50.6
New Hampshire (N = 120)	-2.8	-39.8, 23.9	-1.7	-22.8, 22.7

SOURCES/NOTES: **Sources** Authors' regression analysis of Medicaid monthly enrollment data from CMS. **Notes** Difference-in-difference regression models are population-weighted, log-linear models with wild-cluster bootstrapped standard errors. The "without group trends" model reflects Equation 1, and the "with group trends" model reflects Equation 2 (See Appendix A). Models include state fixed effects and are adjusted for the proportion of the population that of child-bearing age, the state's FMAP, Section 1931 eligibility levels, whether the state allows for 12 months of continuous eligibility, and the seasonally adjusted unemployment rate. The post-period of the New Hampshire analysis was truncated to not overlap with the implementation of New Hampshire's Medicaid work requirement.

When we used Equation 3 to trace out the dynamic effects of the DID without accounting for group-specific trends, we found no change in Arkansas (Figure 2.1, Panel A). From the first to the second month, we estimated a sharp decline in the differential change in Medicaid enrollment. From month two onward, there was a steady increase in the differential change in Medicaid enrollment, but this change was not different from zero. Again, in Arkansas, we could not rule out large positive or negative effects at any time point in the post-period.

In Florida (Figure 2.1, Panel B), we estimated a slow differential decline in the first nine months after the retroactive eligibility waiver. However, this decline was not different from zero at any time point in the first nine months. In month 10, we estimated a 16.0 percent decline (95% CI: -24.2, -12.3) in Medicaid enrollment in Florida relative to the comparison states. This decline persisted in months 11 (15.7%; 95% CI: -24.0, -9.3) and 12 (15.7%; 95% CI: -26.0, -6.5).

Figure 2.1: Percent change in Medicaid enrollment each month after retroactive eligibility waiver without group-specific trends in DID estimation

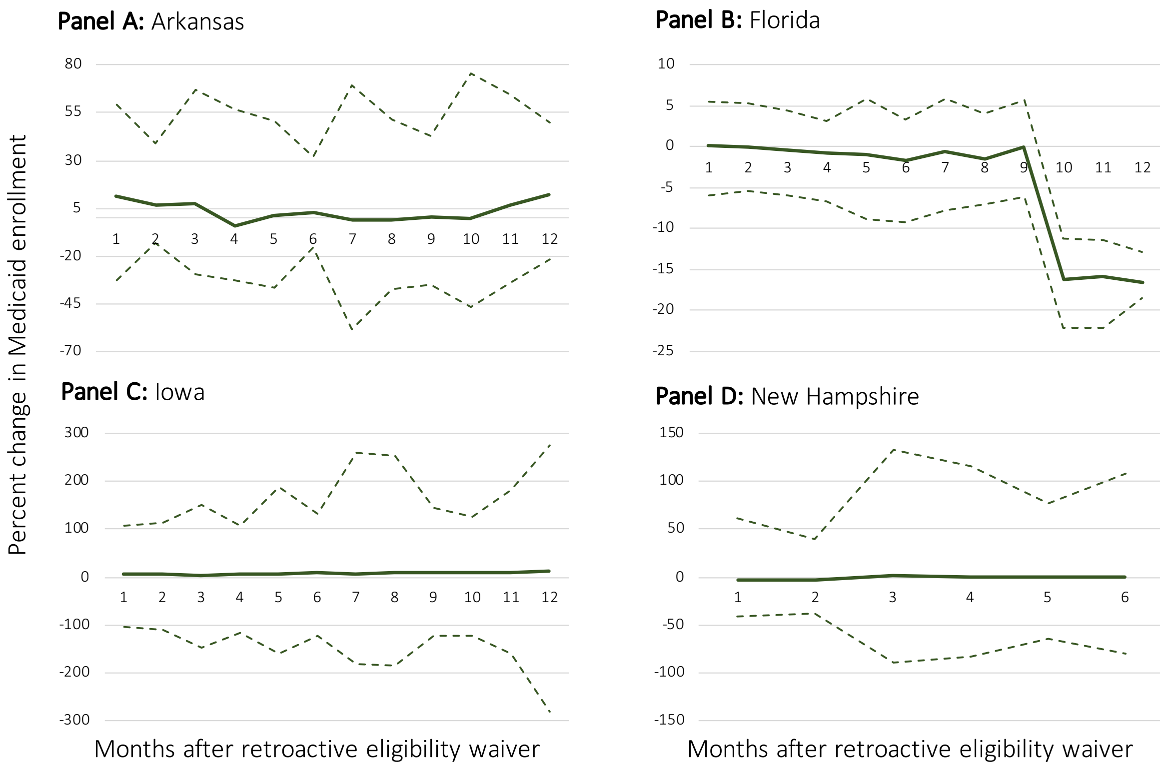


SOURCES/NOTES: **Sources** Authors' regression analysis of Medicaid monthly enrollment data from CMS. **Notes** Difference-in-difference regression models are log-linear models with wild-cluster bootstrapped standard errors and reflect Equation 3. Models include state fixed effects and are adjusted for the proportion of the population that of child-bearing age, the state's FMAP, Section 1931 eligibility levels, whether the state allows for 12 months of continuous eligibility, and the seasonally adjusted unemployment rate. Dashed lines are the 95% confidence intervals from the bootstrapped standard error procedure. The post-period of the New Hampshire analysis was truncated to not overlap with the implementation of New Hampshire's Medicaid work requirement.

In Iowa (Figure 2.1, Panel C), we estimated a gradual increase in the differential change in Medicaid enrollment relative to its comparison states from 3.2 percent to 7.6 percent over the course of the post period. However, these changes were not statistically different from zero at any post-period time point. In New Hampshire (Figure 2.1, Panel D), our estimates produced no clear pattern of differential change in Medicaid enrollment for the six months after the retroactive eligibility waiver was implemented.

When we adjusted for the possibility of differential pre-trends in our DID analysis using Equation 4, our results were qualitatively similar in Arkansas (Figure 2.2, Panel A), Iowa (Figure 2.2, Panel C), and New Hampshire (Figure 2.2, Panel D). We still cannot rule out positive or negative large effects of a retroactive eligibility waiver on Medicaid enrollment. In Florida (Figure 2.2, Panel B), our monthly estimates with group-specific trends looked similar to our estimates produced by Equation 3. However, the estimates at month 10 and month 12 were no longer statistically significant at the α = 0.05 level (-13.4; 95% CI: -27.5, 0.1 and -12.4; 95% CI: -36.3, 14.1, respectively).

Figure 2.2: Percent change in Medicaid enrollment each month after retroactive eligibility waiver with group-specific trends in DID estimation



SOURCES/NOTES: **Sources** Authors' regression analysis of Medicaid monthly enrollment data from CMS. **Notes** Difference-in-difference regression models are log-linear models with wild-cluster bootstrapped standard errors and reflect Equation 3. Models include state fixed effects and are adjusted for the proportion of the population that of child-bearing age, the state's FMAP, Section 1931 eligibility levels, whether the state allows for 12 months of continuous eligibility, and the seasonally adjusted unemployment rate. Dashed lines are the 95% confidence intervals from the bootstrapped standard error procedure. The post-period of the New Hampshire analysis was truncated to not overlap with the implementation of New Hampshire's Medicaid work requirement.

When we pooled the retroactive eligibility states together to address potential power issues in the comparative case studies, we did not find a difference in enrollment in the retroactive eligibility states relative to the comparator states in the difference-in-differences specifications with and without group trends (Table 2.3). Using the DID specification where we traced out dynamic treatment effects, we found statistically significant declines in Medicaid enrollment the fifth (10.1 percent; 95% CI: -19.5, -0.80) and sixth (10.8 percent; 95% CI: -19.1, -2.5) months after implementation of a retroactive eligibility waiver across both specifications. There were similar declines in Medicaid enrollment ten months after a retroactive eligibility waiver (12.1; 95% CI: -20.3, -3.9).

While there is no *a priori* reason to believe that retroactive eligibility waivers would take 10 months to produce declines in enrollment, it is possible that the declines in months five and six after implementation are attributable to the elimination of retroactive eligibility, provided that it takes time for knowledge of this change to diffuse to among beneficiaries and health care providers.

Table 2.3: Pooled DID analysis of retroactive eligibility and Medicaid enrollment

	Without group trends		**With group trends**	
	% change in enrollment	*Clustered 95% CIs*	*% change in enrollment*	*Clustered 95% CIs*
DID Estimate	-2.3	-5.4, 0.71	-1.7	-5.4, 2.0
Dynamic Estimates				
Month 1	-0.07	-3.2, 3.0	-0.23	-3.5, 3.0
Month 2	1.6	-1.3, 4.6	1.5	-1.9, 4.8
Month 3	0.74	-1.8, 3.2	0.54	-2.3, 3.4
Month 4	-0.44	-2.6, 1.7	-0.65	-2.9, 1.6
Month 5	-9.9	-18.3, -1.5	-10.1	-19.5, -0.80
Month 6	-10.6	-17.8, -3.4	-10.8	-19.1, -2.5
Month 7	-0.20	-2.9, 2.5	-0.48	-3.8, 2.9
Month 8	0.59	-2.5, 3.7	0.31	-2.5, 5.2
Month 9	1.7	-1.4, 4.7	1.4	-2.5, 5.2
Month 10	-11.8	-18.5, -5.0	-12.1	-20.3, -3.9
Month 11	-0.40	-2.7, 1.9	-0.75	-3.9, 2.4
Month 12	0.8	-1.6, 3.2	0.46	-3.1, 4.1

SOURCES/NOTES: **Sources** Authors' regression analysis of Medicaid monthly enrollment data from CMS. **Notes** Difference-in-difference regression models are log-linear models with clustered standard errors and reflect all four equations. Models include state fixed effects and are adjusted for the proportion of the population that of child-bearing age, the state's FMAP, Section 1931 eligibility levels, whether the state allows for 12 months of continuous eligibility, and the seasonally adjusted unemployment rate. The dynamic estimates do not include New Hampshire and its comparators in months 9-12, as the post-period was truncated to not overlap with the implementation of New Hampshire's Medicaid work requirement.

Synthetic Control Method (SCM)

In all four retroactive eligibility states, SCM produced a synthetic control group that looks nearly identical on pre-period averages of the lagged outcome and covariates (Table B.2). Because of the size of Florida's total and Medicaid-enrolled population, the state's most similar on these dimensions are expansion states. Thus, almost all states that comprise synthetic Florida are Medicaid expansion states.

The evolutions of the outcome in each case study over the pre and post retroactive eligibility period are shown in Figures B.1 – B.4. Using an 'eyeball test', the SCM seems to match the outcome across the retroactive eligibility state and synthetic control group in the pre-period fairly well. The 'eyeball test' is confirmed by the root mean squared prediction error (RMSPE) in each of the four comparative case studies. In all four of our case studies, we obtained RMSPEs under 0.5, suggesting that the fit of our SCM is quite good. Visual inspection of the post-period outcome trends between the retroactive eligibility state and synthetic control group showed no substantial deviation of the retroactive eligibility state's trend in Arkansas, Iowa, or New Hampshire (Figures B.1, B.3, and B.4, respectively). In Florida, however, we found a decrease in enrollment in the synthetic control group relative to Florida in month 11 after the retroactive eligibility waiver was implemented.

We also conducted a 2x2 difference-in-difference analysis using the outcomes from the SCM analysis. We then tested the estimated treatment effect using a z test. Using this technique, we found no statistically significant change in Medicaid enrollment using the SCM analysis, which is consistent with the results of our comparative case study DID analysis. Results and discussion of sensitivity analyses for the SCM analysis can be found in the Appendix B and Table B.4.

Discussion

Overall, we found no consistent change in Medicaid enrollment after the implementation of a retroactive eligibility waiver using two quasi-experimental methods. Our comparative case study DID analyses were likely under-powered to detect the effect size of interest, as indicated by the small treatment effects and large confidence intervals. We conducted an analysis where we pooled where we found that the implementation of a retroactive eligibility waiver is associated with a decrease in enrollment of roughly 10% in our pooled analysis at five and six months after implementation of a retroactive eligibility waiver. However, there was no reason *a priori* to think that we would find a change in enrollment at this time. Future research should investigate whether this decline was among the population subject to the retroactive eligibility waiver and thus attributable to the implementation of the waiver.

Another possible reason for detecting no consistent relationship is because there was none. As CMS points out in its evaluation logic model, one of the moderating factors to the relationship between a retroactive eligibility waiver an increased enrollment is beneficiary knowledge and understanding.[27] Previous literature suggests that Medicaid beneficiaries (like most privately-insured Americans) do not fully understand their benefit package or know when the benefit package changes.[61]

Conditional on beneficiaries understanding retroactive eligibility and knowing that they can no longer receive retroactive coverage, other moderators may inhibit enrollment. Over the course of the Trump administration, states have implemented barriers, including retroactive eligibility waivers, to slow Medicaid enrollment. Together, it is likely that these collective efforts have had a negative effect on Medicaid enrollment over the past several years. This does not invalidate our study design, however, as retroactive eligibility waivers were implemented at separate and distinct times from these other efforts.

Future research should explore the effect of retroactive eligibility waivers on other outcomes, as it is possible that other outcomes are affected by this change even in enrollment does not change. As suggested by CMS' logic model, the impact of retroactive eligibility waivers on enrollment continuity, health outcomes, and the financial status of beneficiaries and hospitals are of particular interest. Understanding how retroactive eligibility provisions affects these outcomes is important to know whether CMS should continue to approve these kinds of waivers, as other at least two other states (Nebraska and Utah) have a retroactive eligibility waiver pending with CMS.[43]

Chapter 3:
Do Methodological Birds of a Feather Flock Together?

Co-authored with Laura A. Hatfield, PhD

Abstract

The use of quasi-experimental methods for health services and health policy research has proliferated over the last two decades, as researchers focus on program evaluation and studies of causal inference. Two popular such methods, difference-in-differences (DID) and comparative interrupted time series (CITS), compare observations before and after intervention in treated and comparison groups. Both methods compare the change in the treated group relative to the change in the comparison group and rely on strong, untestable counterfactual assumptions. However, the methodological literature on CITS lacks the mathematical formality for DID, which can obscure comparisons of the two approaches. In this paper, we use the potential outcomes framework to formalize the estimands for two versions of CITS — a general version described by Bloom (2005) and a linear version often used in health services research. We then compare these outcome models to two of their DID counterparts — DID with time fixed effects and DID with time fixed effects and group trends. We show that these two designs begin to diverge in counterfactual construction and treatment effect estimation when one leans into each design's respective constraints. For CITS, this constraint is linearity. In DID, the constraint is a constant difference between the two groups. In the constrained situation, the choice between these two designs matter, and researchers should consider a number of factors (e.g., the data-generating mechanism, the plausibility of linearity, the presence of diverging pre-trends) when deciding which of these two designs to use. We also demonstrate that the most general versions of CITS and DID (general CITS and DID with time fixed effects and group trends), when estimated with linear regression, produce the same counterfactuals and estimate nearly identical treatment effects. The only difference between these two designs is the language used to describe them. We suggest that empiricists carefully write down the outcome model and counterfactuals they are assuming to allow for a more transparent evaluation of the plausibility of the assumptions being made, regardless of the language being used.

Introduction

Observational methods in social science research have proliferated over the last two decades, as more studies seek to make causal inferences, and observational designs have been demonstrated to reliably estimate experimental results.[21,62–65] Among the designs used in health policy and health services research, difference-in-differences (DID) is particularly popular thanks to the method's visual and conceptual simplicity. For instance, over 300 peer-reviewed publications have evaluated the Affordable Care Act's Medicaid expansion, and the majority use DID.[66] DID compares the change in outcome before and after treatment in a treated group to the change in outcome in a group that does not receive treatment (the comparison group). Advances in DID methodology have been developed in the econometrics and statistics literature in parallel to the method's growing use.[23,53,67–72]

A comparative interrupted time series (CITS; also known as interrupted time series with a control or controlled interrupted time series) is superficially similar. It uses treated and comparison groups to quantify changes before and after an intervention and relies on untestable counterfactual assumptions. The methods are so similar that critiques of observational methods have lumped DID and CITS together, asserting these methods are the same.[73] Despite these similarities, proponents of DID and CITS each strongly prefer their respective method. While this is partly disciplinary (DID is preferred in economics, while CITS is preferred in health services, education policy, and clinical epidemiology research), the lack of mathematical formalization of CITS may contribute to the confusion.

To further add to the confusion, there are different formulations of both DID and CITS. In fact, education and health services researchers usually define CITS differently. The education policy literature defines CITS more generally than the health services literature. Throughout the remainder of this paper, we will discuss two of the most-often implemented versions of DID — DID with time fixed effects ("FE DID") and DID with time fixed effects and group-specific pre-trends ("FE DID with group trends") — and CITS — the generalized version of CITS ("general CITS") used in education research and the fully linear version of CITS ("linear CITS") used in health services research.

In some situations, only one method is feasible. FE DID requires only two observation points (one pre- and one post-intervention), FE DID with group trends requires at least five observation points (at least four in the pre-period and one in the post-period. General CITS also requires at least five observation points (four in pre and one in the post-period), while linear CITS requires at least eight (four pre- and four-post intervention). Both versions DID can estimate only one treatment effect (i.e., a level shift) or can estimate an effect at each post-period time point. Both versions of CITS usually estimate more than one treatment effect. General CITS estimates as many treatment effects as there are post-period time points, which could be a single time point or many. Linear CITS estimates two treatment effects — a level shift and a trend shift. Many researchers who prefer linear CITS cite the ability to estimate both immediate and sustained effects as an advantage of the method. In DID, this is often accomplished with time fixed effects interacted with treatment, which may be preferred because it is less parametric than linear CITS. However, linear CITS gives the magnitude and direction of the growing treatment effect, while DID with time-varying treatment effects or general CITS cannot. With both general CITS and DID with time-varying treatment effects, researchers must choose some way to summarize the treatment effect for decisionmakers.

Critics of DID state that the method is overly simplistic, restrictive, and inflexible. In the canonical specification of DID with two time points per group, this study design precludes adjustment for changes in the rate of the outcome prior to intervention. However, empirical DID often includes more than two time points to observe changes in the outcome prior to intervention. The counterfactual assumption of DID states that the average change in outcome from pre- to post-intervention in the two groups would have been the same *if not for treatment*. In practice, most researchers typically impose the most restrictive version of this assumption — the parallel evolution of outcomes at all pre- and post-period time points, which is often referred to as the “parallel pre-trends” assumption. However, alternative modeling specifications relax this.[68,69] Other specifications, such as modeling group-specific trends, are often used to address instances when the adjusted pre-period trends are not parallel between the two groups.[74]

Critics of CITS argue that the method makes unnecessary parametric assumptions. But these parametric assumptions offer a way out of DID's parallel pre-trend conundrum, since both versions of CITS explicitly model different outcomes trends in the two groups in the pre-period. The counterfactual assumption of general CITS is that the difference in the post period from the pre-period linear extrapolation would have been the same for the treatment and comparison group *if not for treatment*. Similarly, the counterfactual assumption of linear CITS is that the change in level and trend in the comparison group from the pre-period extrapolation would have been the same in the treated group *if not for treatment*. The tradeoff for modeling a less restrictive relationship between the two groups is a parametric assumption of linearity assumption. Deviations from linearity in the pre-period, like those shown in Baicker & Svoronos,[75] may lead to a biased estimation of the treatment effect. Additionally, those authors demonstrate that irregularities in linearity near the edges of each period may have increased influence on treatment effect estimation.

DID and CITS differ in a number of ways, including functional form, counterfactual extrapolation, and treatment effect estimation. All four versions we consider add the change seen in the comparison group to the treatment group's pre-period outcomes to construct the counterfactual. FE DID does not require any parametric specification of the outcome. FE DID with group trends assumes that the growing difference between the two groups is linear but does not specify a parametric form for the outcome. Both versions of CITS assume that the outcome trend is linear; general CITS assumes linearity in the pre-period trend only, while linear CITS assumes linearity in both the pre- and post-period trends.

All four designs extrapolate the pre-period outcomes to the post-period to construct the counterfactual in the post-period. FE DID assumes and extrapolates no trend difference between the comparison and treatment group. FE DID with group trends assumes that the observed trend difference in the pre-period would continue to the post period. General CITS also assumes that the observed pre-period trend difference would continue into the post period. The counterfactual is constructed in FE with group trends by extrapolating the linear difference between the two groups and then adding the comparison group's level change to the pre-period level of the treated group. Counterfactual construction is similar in the general version of CITS. The distance between the linear extrapolation of the comparison group's pre-period outcome and each observed post-period outcome is added to the treated group's linear extrapolation. Like general CITS, linear CITS extrapolates both groups' pre-period trend. To construct the counterfactual, the difference between the comparison group's extrapolated intercept and trend and its observed post-period intercept and trend is added to the extrapolated intercept and trend in the treated group.

In DID without time fixed effects, a treatment effect can be estimated using a 2x2 table. With FE DID and FE DID with group trends, treatment effects can be estimated using linear regression and other non-parametric approaches. General and linear CITS can only be estimated via linear regression. Other outcome models (particularly those on the multiplicative scale) make it difficult to separate the treatment effect estimates of general and linear CITS. With both versions of FE DID, we can estimate an average treatment effect, or we can estimate a treatment effect at each post-period time point (or an ATT_t). General CITS also estimates a treatment effect at each post-period time point. Linear CITS estimates a change in the level of the outcome and a change in the trend of the outcome, which can be used to construct a treatment effect at each post-period time point. We summarize the functional form, extrapolation, and treatment effect estimation of each version of these two designs in Table 3.1.

Table 3.1. Comparison of Features of CITS and DID

Design	**# obs.**	**Functional Form**	**Extrapolation**	**Linearity**	**Treatment effect; estimation**
CITS					
General	≥ 5	Linear, additive	Linear	Pre-period trend	$ATT_{(t)}$; Linear regression
Linear	≥ 8	Linear, additive	Linear	Pre & post-period trend	Level, trend[1]; Linear regression
DID					
FE with group trends	≥ 5	Linear[2], additive	Linear	Differential growth between groups	$ATT_{(t)}$; Linear regression/non-parametric approaches
FE	≥ 2	Additive	Zero difference	None	$ATT_{(t)}$; Linear regression/non-parametric approaches

[1] Can be used to construct an $ATT_{(t)}$

[2] Linearity is assumed for the evolution of the difference between the treated and control, rather than for the outcome trend itself.

Despite these critiques and differences, there have been no papers to our knowledge that compare these two methods to determine instances (if any) in which one design is more appropriate. The paper proceeds in the following way. First, we identify and define the untreated potential outcomes for all four versions of these two study designs: the general and linear formulation of CITS, FE DID, and FE DID with group trends. Readers who wish to skip the mathematical formalization can proceed to the three empirical examples illustrating the issues involved in choosing between CITS and DID and their respective formulations.

As we will demonstrate, CITS and DID, in their most general forms, produce the same counterfactuals and estimate the similar treatment effects. The only meaningful difference between general CITS and FE DID with group trends is the words used to describe the study design. Differences in counterfactuals and estimated treatment effects between CITS and DID arise when researchers lean into each design’s respective constraints. When linearity is imposed in both the pre- and post-period (as in linear CITS), treatment effects may be biased when non-linearities are present in the outcome trend. When zero differential growth is imposed on the counterfactual assumption (as in the parallel pre-trend assumption of FE DID), differential trends may also result in bias in the estimated treatment effects.

We conclude with guidance on when and how one might want to use the more restrictive versions of each study design and improving the understanding of the models being implemented in health services research. Specifically, we advise health policy and health services researchers who use two-group, two-time period quasi-experimental designs to provide a careful description write down the outcome model and counterfactual assumptions, regardless of the nomenclature used to describe the study design.

Comparison of Study Designs’ Potential Outcomes

Below we define the untreated and treated potential outcomes for four study designs — a general version of CITS, a linear version of CITS, DID with time fixed effects, and DID with time fixed effects and group-specific trends.

Additionally, we illustrate how the counterfactual is constructed in each study design. We suppose the true data-generating model is, $y \sim (t/2)^2$, in the comparison group and $y \sim (t/3)^2$ in the treated group. This is non-linear, but the differential growth between the two groups is linear. The data-generating model does not change from the pre-period to the post-period. We fit a series of models and show how each would predict the treated group's untreated potential outcomes in the post-period (i.e., the counterfactual outcomes for the treated group). We plot both the true outcomes (the dots) and the model extrapolations (the lines) for both the comparison group (pink) and treated group (blue) in Figure 3.1.

Let Y^0 be the untreated potential outcomes; t indicate time with t_0 being the time of the intervention (with $t < t_0$ indicating the pre-period and $t \geq t_0$ indicating the post period); β be the parameters governing the pre-period outcomes, and $\breve{\beta}$ be parameters governing post-period outcomes. Superscripts indicate group membership — 0 for the comparison group and 1 for the treated group.

General formulation of CITS

This formulation is found in Bloom and Riccio's 2005 analysis of the Jobs-Plus program.[76] In this formulation of CITS, separate lines are fit through the pre-period outcomes of the comparison and treated groups. These lines are extrapolated into the post-period for both groups. The counterfactual is constructed by first measuring the distance from the extrapolated line to the observed post-period outcomes in the comparison group at each post-period time point, $t \geq t_0$. Then, these distances are added to the extrapolated line for the comparison group at each time point.

The comparison group's untreated potential outcomes are,

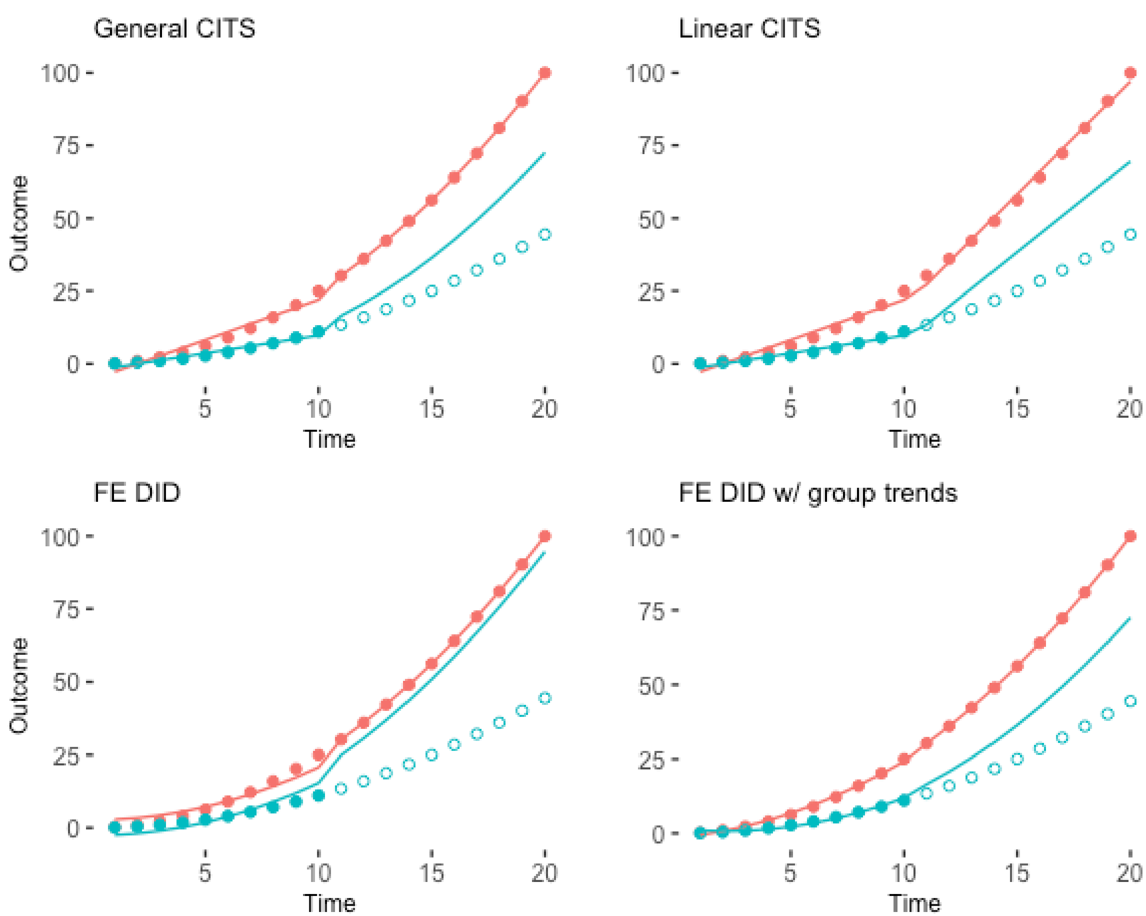
$$E[Y^0] = \beta_0^0 + \beta_1^0 t + \sum_{k=t_0}^{\mathrm{T}} \breve{\beta}_k^0 I_{k=t},$$

and the treated group's untreated potential outcomes are,

$$E[Y^0] = \beta_0^0 + \beta_0^1 t + \sum_{k=t_0}^{\mathrm{T}} \breve{\beta}_k^0 I_{k=t} + \beta_0^1 + \beta_1^1 t.$$

The difference between the models are β_0^1 and β_1^1, which represent the differential level and trend of the treated group's outcomes, respectively. In this version of CITS, we construct a counterfactual outcome at each time point, $t \geq t_0$, using the $\breve{\beta}_k^0$ parameters. In Figure 3.1, the general version of CITS (upper left panel) fits the true data well in the pre-period and produces the counterfactual outcome that we would expect to see, given the comparison group's post-period outcomes. A linear approximation in the pre-period is reasonable for both groups, and this model allows for flexible modeling of the post-period outcomes.

Figure 3.1: Comparison of counterfactual scenarios in non-linear models



NOTES: The untreated group's true outcomes (pink dots) are generated from the model, $y \sim \left(\frac{t}{2}\right)^2$, and the treated group's true outcomes in the pre-period (blue dots) are generated from the model, $y \sim \left(\frac{t}{3}\right)^2$. The treated group's true, unobservable post-period outcomes are the empty dots. Lines are the predicted values of the outcomes using linear regression for each respective study design.

Linear Formulation of CITS

Like the more general formulation of CITS, the fully linear version of CITS fits a different line through each group in the pre-period and extrapolates those lines into the post-period. Instead of measuring the distance from each point in the post-period to the extrapolated line, the linear version of CITS fits another line through the comparison group's post-period outcomes. The counterfactual is constructed by adding the differences in level and trend from the comparison group to the linear extrapolation of the treated group.

The comparison group's untreated potential outcomes are,

$$E[Y^0] = \beta_0^0 + \beta_1^0 t + \breve{\beta}_0^0 + \breve{\beta}_1^0\, t,$$

and the treated group's untreated potential outcomes are,

$$E[Y^0] = \beta_0^0 + \beta_1^0 t + \breve{\beta}_0^0 + \breve{\beta}_1^0\, t + \beta_0^1 + \beta_1^1 t.$$

Like the general version of CITS, β_0^0 and β_1^0 are the pre-period level and trend for the comparison group in the linear version of CITS. $\breve{\beta}_0^0$ and $\breve{\beta}_0^1$ are the comparison group's level and trend in the post period ($t \geq t_0$). The differential parameters in the treated potential outcomes are β_0^1 and β_1^1, which are the differential level and trend for the treated group in the pre-period.

Because the outcomes are reasonably linear *within* each group and study period and because the linear version of CITS estimates a different line in each group and study period, the linear version of CITS seems to fit the outcomes fairly well and estimates a reasonable counterfactual for the treated group (upper right panel of Figure 3.1). The lines do not go directly through the observed outcomes but are reasonable linear approximations.

DID with time fixed effects

DID with time fixed effects (FE DID) assumes a constant difference between the treated and comparison groups. FE DID measures the average level change from the pre-period to the post-period in the comparison group. To construct the counterfactual, FE DID assumes that the same change would have been seen in the treated group *absent treatment*.

In the formulation of the potential outcomes for FE DID, we use γ_t to represent the time fixed effects, which are estimated throughout the pre- and post-period. The untreated potential outcomes in the comparison group are,

$$E[Y^0] = \beta_0^0 + \breve{\beta}_0^0 + \gamma_t,$$

and the untreated potential outcomes in the treated group are,

$$E[Y^0] = \beta_0^0 + \breve{\beta}_0^0 + \beta_0^1 + \gamma_t$$

β_0^0 and $\breve{\beta}_0^0$ are the levels of the comparison group's untreated potential outcomes in the pre- and post-period, respectively. β_0^1 also represents the differential level of the treated group in the pre-period. In the lower left panel of Figure 2, we can see the "parallel pre-trends" or constant difference assumption of DID. Because this, the FE DID fails to capture the growing difference between the two groups.

DID with time fixed effects and group-specific trends

Unlike FE DID above, FE DID with group-specific trends adds linear time trends for the pre-period outcomes in the treated and comparison groups, allowing these two groups to have diverging trends. These lines are fit after the outcomes are effectively de-meaned via time fixed effects. This assumes that the average change in the comparison group would occur in the treated group *absent treatment*. But unlike FE DID, the differential trend is extrapolated prior to the addition of the average change to the treated group's pre-period level. Thus, the counterfactual is constructed by accounting for the differential trends across the two groups. The untreated potential outcomes in the comparison group are,

$$E[Y^0] = \beta_0^0 + \beta_1^0 t + \breve{\beta}_0^0 + \gamma_t,$$

and the untreated potential outcomes in the treated group are,

$$E[Y^0] = \beta_0^0 + \beta_1^0 t + \breve{\beta}_0^0 + \beta_0^1 + \beta_1^1 t + \gamma_t.$$

Like FE DID, this allows for a level difference in the pre-period (β_0^1). Also, $\breve{\beta}_0^0$ is the change in the average outcome for the comparison group in the post-period. Unlike the FE DID version above, this allows linear trends in the pre-period in the comparison (β_1^0) and treated (β_1^1) groups. Because the true data-generating model has a differential trend across the two groups, accounting for this differential trend improves the model fit in the pre-period (Figure 3.1, bottom right panel). Additionally, this model produces a counterfactual that takes into account the growing differential between the two groups in a way that the FE DID model cannot.

This counterfactual construction is identical to that of general CITS because the fixed effects are not being extrapolated into the post-period. The extrapolation occurring in FE DID with group trends is a differential linear time trend, which is the same extrapolation in general CITS. The difference is the assumption made about the pre-period outcome. In the general version of CITS, the outcome's trend is assumed to be linear, but FE DID with group trends assumes the difference between the groups is growing linearly. In the simple linear model cases we consider here, these two sets of assumptions turn out to be identical.

Empirical Examples

The choice between CITS and DID is one about the assumptions of the data-generating model. Of the four models we discuss, three of them require some assumption of linearity and one of them (FE DID) assumes a constant difference between the two groups. In the following three empirical examples, we discuss the decisions that the researchers might make to choose between DID and CITS, reanalyze the data from each paper, and compare our findings to the original study results. In each study re-analysis, we first create an event study plot by fitting covariate-adjusted linear regression models with a dummy variable for each time point pre- and post-intervention relative to the intervention time point. Then, we fit four different models to each dataset 1) FE DID, 2) FE DID with group trends, 3) general CITS, and 4) linear CITS. In both DID formulations, we estimate time-varying treatment effects to compare the dynamics of the outcome to those produced by both versions of CITS.

Medicaid expansion's spillover to the criminal justice system

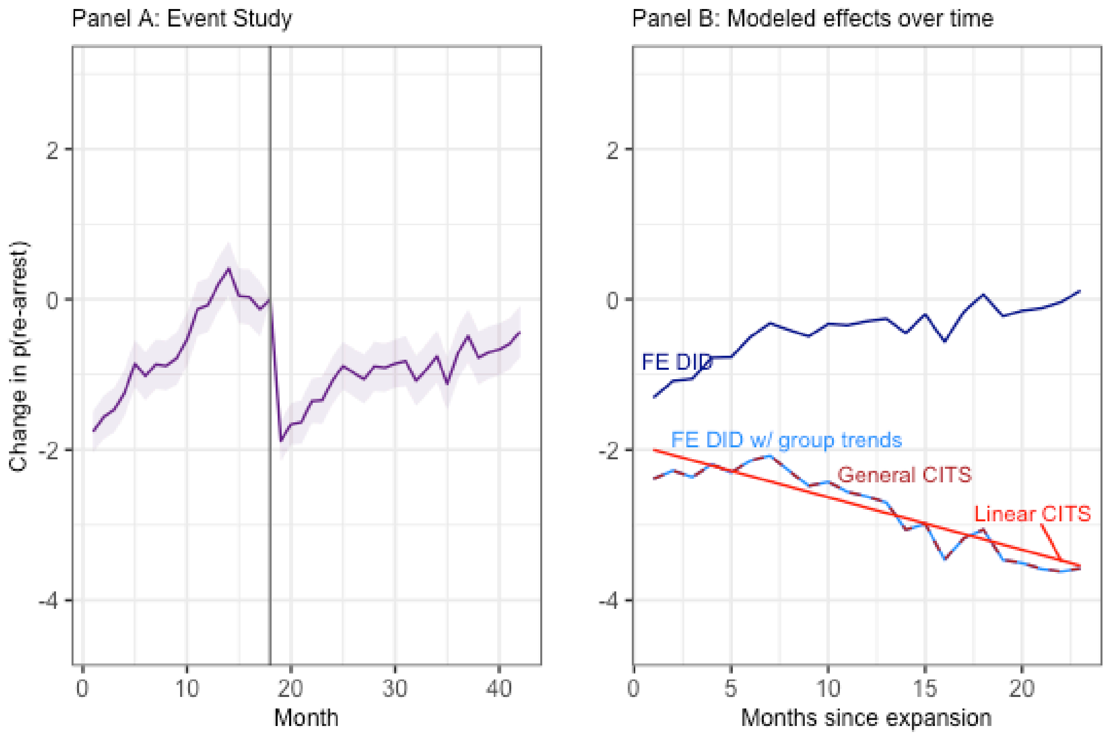
Fry, McGuire, and Frank[77] conduct three case studies to estimate the impact of the ACA's Medicaid expansion on rates of return to county jails. Each case study compares the change in recidivist outcomes for a county where Medicaid was expanded under the ACA to the change in outcomes for a county where Medicaid was not expanded. For brevity, we will re-analyze one of the three case studies (the Southwest counties) with only one outcome (the probability of re-arrest).

The authors of the original paper briefly discuss the rationale for choosing a CITS design over DID, "Given the drivers of the outcome in this study (e.g., policing practices, criminal justice practices, and access to behavioral health services) and how they may vary between the counties, assuming linearity in the evolution of the outcomes in each of the two groups seems more reasonable than assuming that they evolve in the same average way over time" (page 16). The authors also have qualitative information that suggests differential demographics, access to behavioral health resources, policing practices, and coordination between the behavioral health and criminal justice systems in the study counties. Thus, assuming that outcomes evolve in the same way between the treatment and comparison groups may not be reasonable.

The difference in the pre-period trends is large and approximately linear (Figure 3.2, Panel A). After Medicaid expansion, there is an immediate decrease in the differential probability of re-arrest and a flattening of the differential trend. From the event study plot in Panel A, there is suggestive evidence that we could expect to see negative estimates for the change in the intercept and slope. In this re-analysis, diverging pre-period outcomes suggest that FE DID is not an appropriate study design. Accounting for this divergence via CITS or FE DID with group trends may be a more appropriate way to estimate the relationship between Medicaid expansion and return to jail.

Indeed, when we estimate a linear CITS, we find that Medicaid expansion reduces the probability of any re-arrest by -2.00 (95% CI: 1.62, 2.34) percentage points in the first month after expansion (i.e., the CITS level estimate), and the decline grows by 0.07 (95% CI: 0.06, 0.08) percentage points each month (i.e., the CITS trend estimate).

Figure 3.2: Comparison of Estimates Across Study Designs - Fry et al, 2019



SOURCES/NOTES: **Sources** Authors' re-analysis of booking and release data from Fry et al., 2020. **Notes** Panel A is an event study plot, where the adjusted differential in outcome between the treated group and comparison group relative to the time of intervention is estimated for each time period before and after Medicaid expansion. Panel B provides time-varying estimates for each month after Medicaid for FE DID, FE DID with group trends, general CITS, and linear CITS. Covariate adjustment is the same for each model presented in both Panels A and B and is exactly the same as the covariate adjustment used in Fry et al., 2020.

Because FE DID with group trends extrapolates the group-specific linear trends into the post period but does not extrapolate the time fixed effects, the estimates for General CITS and FE with group-specific trends are indistinguishable from one another and hover around the CITS linear extrapolation (Figure 3.2, Panel B). The largest differences between the linear CITS extrapolation and the general CITS/FE DID with group trends estimates are at seven and 16 months after expansion. However, these estimates are not statistically different from the linear CITS estimation. This suggests that accounting for the diverging pre-period trends in DID will result in estimates similar to both the linear and general CITS estimates over the post-period, if the outcome's evolution is approximately linear.

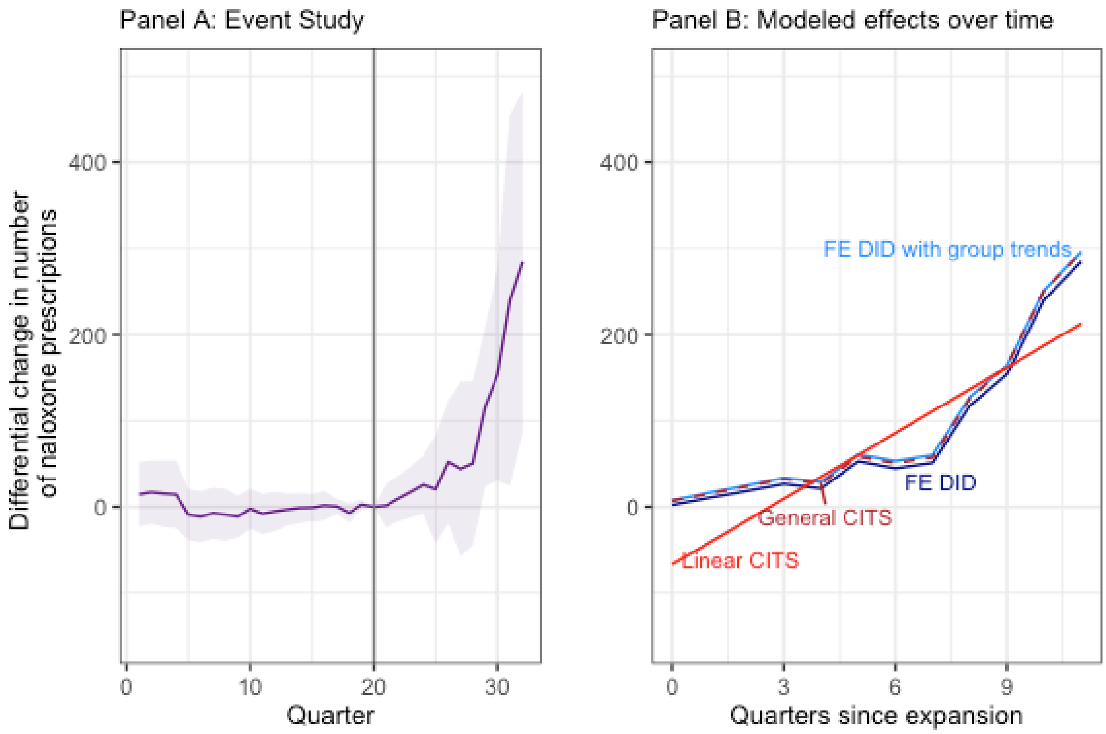
Using FE DID to estimate this relationship, we find that expansion resulted in a 0.45 (95% CI: 0.39, 0.52) percentage point decline in the probability of re-arrest. The DID FE estimate is 2.32 percentage points (or 84%) smaller than the linear CITS estimate at the midpoint of the post-period. Because FE DID does not account for the diverging pre-trends, the treatment effect appears to grow in a positive direction rather than a negative one. Adjusting for pre-period trends by one of the other three designs increases the treatment effect estimate by 54.8% in the beginning, by 111.1% in the middle, and by 330.0% at the end of the post-period.

Because the drivers of recidivism, such as policing and criminal justice practices and the availability of behavioral health treatment, are different across the counties, assuming that the change would be constant (as in FE DID) may less reasonable than assuming some form of linearity (as in the other three designs). Because the assumption of linearity in the outcome trend is reasonable in both study periods, linear CITS produces estimates similar to general CITS and FE DID with group trends. The most crucial study design choice here appears to be accounting for diverging pre-trends either via FE DID with group trends or either type of CITS.

Medicaid expansion and naloxone prescriptions

Frank and Fry[78] compare the change in total naloxone prescriptions before and after Medicaid expansion in expansion and non-expansion states. While the authors do not provide any explicit rationale for choosing a DID over a CITS, they do write, “the number of naloxone prescriptions paid for by Medicaid was essentially identical in expansion states compared to non-expansion states,” which suggests that there was no pre-period level difference in the outcome. However, DID can accommodate differences in the level of the pre-period between the treated and untreated groups, but it cannot accommodate differences in the pre-period evolution of the outcome. The adjusted event study plot in Figure 3.3 (Panel A) lends support to both the level and evolution assumptions, as the differential trend line is not significantly different from zero.

Figure 3.3: Comparison of Estimates Across Study Designs - Frank & Fry, 2019



SOURCES/NOTES: **Sources** Authors' re-analysis of Medicaid covered naloxone prescriptions from Frank & Fry, 2019. **Notes** Panel A is an event study plot, where the adjusted differential in outcome between the treated group and comparison group relative to the time of intervention is estimated for each time period before and after Medicaid expansion. Panel B provides time-varying estimates for each month after Medicaid for FE DID, FE DID with group trends, general CITS, and linear CITS. Covariate adjustment is the same for each model presented in both Panels A and B and is exactly the same as the covariate adjustment used in Frank & Fry, 2019.

Figure 3.3, Panel A also shows that the outcomes diverge non-linearly after expansion. Modeling the post-period outcome trend as linear, as in linear CITS, may result in biased treatment effects. Indeed, the CITS level estimate suggests that Medicaid expansion resulted in a *decrease* of 66.8 (95% CI: 7.6, 1259) naloxone prescriptions in the quarter after expansion, with a growing increase of 25.4 (95% CI: 7.1, 43.7) prescriptions in each quarter thereafter. If a researcher were to interpret the level change of linear CITS, the resulting treatment effect would result in a reverse policy conclusion than the linear CITS extrapolation at 18 months (six quarters) after expansion. This suggests that researchers using linear CITS should extrapolate the level and trend estimates to some reasonable point in the post-period to determine the policy effect. While the linear CITS extrapolation at 6 quarters post-expansion is similar to the estimates of the other three designs, the non-linearities in the post-period mean that the linear CITS estimates are significantly different from both FE DID designs and general

CITS at the beginning and end of the post-expansion period (Figure 3.3, Panel B). This results in an underestimate of 113% and 28% at the first and last post-period time point, respectively.

Unlike linear CITS, general CITS assumes linearity in the pre-period outcome trend but flexibly models the post-period outcomes. In Frank & Fry (2019), the pre-period outcome trend appears fairly linear, which results in general CITS estimates that are identical (within computational error) to those produced by FE DID with group trends (Figure 3.3, Panel B). The FE DID estimates are also very similar to those of general CITS and FE DID with group trends. As recommended by Bilinksi and Hatfield[52], one way to assess the plausibility of the "parallel pre-trends" assumption of DID is to compare the estimates of FE DID and FE DID with group trends. If these specifications produce estimates significantly different from one another, the "parallel pre-trends" assumption may not hold. Here, this assumption seems like a reasonable one, as the estimates do not differ across these model specifications.

Drivers of the number of naloxone prescriptions include the prevalence of opioid use disorder and the presence other laws increasing access to naloxone, such as the ability of doctors to prescribe naloxone to a friend or family member. While the pre-period outcome trends in this re-analysis are approximately parallel, there is anecdotal evidence that some states expanded Medicaid, in part, because of the prevalence of opioid use disorder. If this were the case, then assuming a constant pre-period difference in the outcome and its drivers (as in FE DID) may be unrealistic. Visual inspection of the event study plot, however, suggests that a linearity assumption (particularly in the post-period) is also not a reasonable assumption. FE DID with group trends or general CITS gives us the best of both worlds of linear CITS and FE DID. This design allows for us to model differential trends in the outcome and its drivers and flexibly model the outcome trend in both study periods.

Reformulation of OxyContin and the incidence of Hepatitis C

Powell, Alpert, and Pacula[79] explore the relationship between the 2009 reformulation of OxyContin to an abuse-deterrent form and changes in the incidence of acute Hepatitis C (HCV) infections. The reformulation of OxyContin was a national policy implemented in all U.S. states. To estimate a treatment effect of the reformulation (which happened nationwide), the paper uses a state's rate of

OxyContin abuse or misuse as the continuous treatment variable and an event study. The use of the continuous exposure variable assumes that the relationship between initial OxyContin abuse or misuse and the HCV rate is linear. The counterfactual assumptions of an event study are not clearly defined, but the method identifies a differential change relative to the time of implementation at every *pre and post-period* time point. This is different than the identifying (i.e., counterfactual) assumption of FE DID or FE DID with group-specific trends.

Despite these differences, the paper invokes a counterfactual assumption that is similar to that of DID, "The testable assumption is that OxyContin misuse rates were not predictive of hepatitis C infection trends before the reformulation. Studying the effect of a policy exposure both before and after the intervention in an event study is recommended when using difference-in-differences designs to study health policy" (page 289). While counterfactual assumptions are *not* testable, this language is suggestive of the DID parallel pre-trends assumption and the procedures often used to "test" this assumption. The study concludes that the reformulation of OxyContin did not result in an increase in the incidence of acute HCV cases in the year after formulation but did in each subsequent year of the study period (see Exhibit 4 of the original paper; the largest increase occurred four years after reformulation).

In our re-analysis, we conduct the event study (as we did in the first two re-analyses) to visually assess the respective assumptions of CITS and DID, but we do not make any causal inferences from the event study. Rather, we use FE DID, FE DID with group trends, general CITS, and linear CITS to make causal inferences. Additionally, we dichotomize the exposure or treatment variable at the median value of average pre-period rates of OxyContin misuse and abuse (as the manuscript does in Exhibits 2 and 3) for the event study and all four study designs because DID and CITS typically use a binary treatment variable.

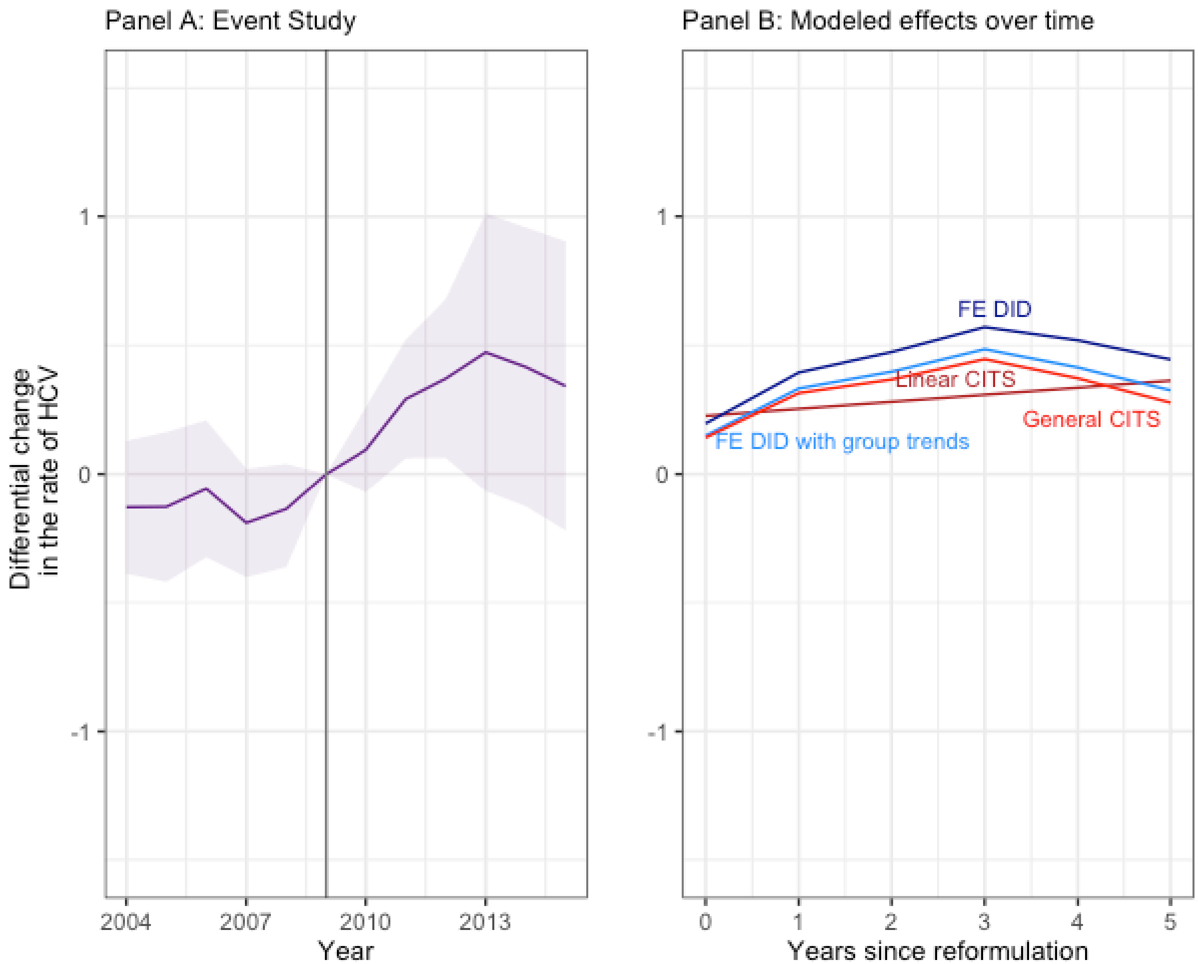
The pre-period trend appears relatively linear (Figure 3.4, Panel A), and there do not appear to be any meaningful differential trends across the comparison and treatment groups. In the post-period, the differential trend appears relatively linear and increases in the high-exposure states compared to the low-exposure states. This increase is different from zero in the second and third years after reformulation. When we use a dichotomous exposure variable, our results go in the same direction as the original paper but differ in magnitude and statistical significance.

Using a FE DID, we find that the reformulation of OxyContin leads to a statistically significant increase in the rate of acute HCV infections in years 2 through 5 after reformulation, consistent with the original findings. When we account for the possibility of diverging pre-trends with general CITS and FE DID with group trends, our estimates get slightly smaller and lose statistical significance in years 4 and 5 after reformulation. As with the reanalysis of Fry et al. (2020) and Frank and Fry (2019), the estimates from the FE DID with group trends and the general CITS are identical (Figure 3.4, Panel B). The decrease in estimates by modeling group-specific pre-period trends suggests that there is likely some differential trend in the outcome.

Linear CITS, like general CITS and FE DID with group trends, accounts for the slight difference in pre-period trend and results in smaller estimates than the FE DID. Although the linear CITS estimates look fairly similar to those produced by the other three study designs, modeling the post-period trend as linear misses the slight increase in years 2 and 3 after reformulation that the more flexible modeling of FE DID and general CITS. The result is that neither the level nor trend estimates in linear CITS are statistically significant.

Prior to the reformulation of OxyContin, there were differential trends in the outcome, perhaps driven by differential rates of risky behaviors (e.g., intravenous drug use or unsafe sexual practices) and access to services for the detection and treatment of HCV. In the absence of OxyContin reformulation and the presence of these drivers of acute HCV incidence, it is unclear that the change in acute HCV incidence would be equivalent in “high” and “low” exposure states.

Figure 3.4: Comparison of Estimates Across Study Designs - Powell, et al., 2019



SOURCES/NOTES: **Sources** Authors' re-analysis of acute HCV incidence from Powell et al., 2019. **Notes** Panel A is an event study plot, where the adjusted differential in outcome between the treated group and comparison group relative to the time of intervention is estimated for each time period before and after Medicaid expansion. Panel B provides time-varying estimates for each month after Medicaid for FE DID, FE DID with group trends, general CITS, and linear CITS. Covariate adjustment is the same for each model presented in both Panels A and B and is exactly the same as the covariate adjustment used in Powell et al., 2019. Unlike Powell et al., 2019 who use a continuous treatment variable, we use a binary exposure variable with the cutoff for "high exposure" being the median OxyContin misuse rate in the pre-period.

In the reanalysis, general and linear CITS account for the divergent pre-period outcomes like FE DID with group trends. However, again, the linearity assumption of CITS may lead us to under-estimate the treatment effect over time. Like the reanalysis in Frank and Fry (2019), the assumption of linearity in the pre-period might be a reasonable one. However, FE DID with group trends allows for modeling flexibility in both the pre- and post-period like FE DID while accounting for differential pre-trends like linear and general CITS.

Conclusion

Through mathematical formalization of the models and careful examination of the counterfactual assumptions, we have highlighted the differences between two versions of CITS and DID — a general version of CITS frequently used in education policy research, linear CITS used most often in health services research, DID with time fixed effects, and DID with time fixed effects and group-specific trends. The counterfactual is constructed similarly in all four designs — by assuming that the change in the comparison group can stand in for the change in the treated group absent treatment. In their most general forms, CITS and DID produce the same counterfactuals and estimate the same treatment effects. The only thing extrapolated in both general CITS and FE DID with group trends is the differential linear trend. Here, the difference between these two designs is disciplinary and the words used to describe each design.

When we lean into each design's respective constraints (linearity for CITS and a constant difference for DID), the counterfactuals begin to differ from one another. In these more constrained situations, the decision about which design to use should rely on the researcher's belief about the data-generating model. If one believes the evolution of the outcome and its confounders is linear, then linear CITS may be preferable. If one believes that the difference in the outcome's evolution between the two groups is zero, FE DID may be preferable. Content expertise should be used to assess which constraint (linearity vs. parallel trends) is most plausible in a given data setting/question to inform the choice between the two designs. Regardless of the design choice, researchers should provide model specifications and counterfactual assumptions in all empirical work. This allows for a more transparent evaluation of the plausibility of these assumptions regardless of the language being used.

While we believe that the decision between CITS and DID should be made via careful consideration of the counterfactual assumption and the evolution of the outcome trend (especially linear CITS and FE DID), researchers may use suggestive evidence from adjusted event study plots to assist them in their decision. Evidence of adjusted, differential pre-period trends in an event study plot may suggest that a DID may not be appropriate, as seen in the re-analysis of Fry, et al. (2020) and Powell, et al. (2019). However, where non-linear outcome trends are present, imposing the linearity assumption of CITS may not fully capture the treatment effect over time as seen in the re-analysis of Frank & Fry (2019) and Powell, et al. (2019). While FE DID may seem like the most flexible model, it comes with what some may consider an overly restrictive counterfactual assumption. While general CITS and FE DID with group trends are the same study design, choosing between linear CITS and FE DID is a matter of trading off seemingly small differences in the counterfactual and parametric form, but these different constraints matter in modeling the outcome, constructing the counterfactual, and estimating a treatment effect.

Bibliography

1. Zeng Z. Characteristics of confined inmates in local jails, 2000 and 2005-2016. In: *Jail Inmates in 2016, NCJ 251210*. Washington, D.C.: Bureau of Justice Statistics; 2018.

2. Regenstein M, Rosenbaum S. What the Affordable Care Act Means for People with Jail Stays. *Health Affairs*. 2014;33(3):448-454. doi:10.1377/hlthaff.2013.1119

3. Morrissey JP, Dalton KM, Steadman HJ, Cuddeback GS, Haynes D, Cuellar AE. Assessing gaps between policy and practice in Medicaid of jail detainees with severe mental illness. *Psychiatric Services*. 2006;57(6):803-808.

4. Wang EA, White MC, Jamison R, Goldenson J, Estes M, Tulsky JP. Discharge Planning and Continuity of Health Care: Findings from the San Francisco County Jail. *American Journal of Public Health*. 2008;98(12):2182-2184. doi:10.2105/AJPH.2007.119669

5. Marks JS, Turner N. The Critical Link Between Health Care and Jails. *Health Affairs*. 2014;33(3):443-447. doi:10.1377/hlthaff.2013.1350

6. Bronson J, Stroop J, Zimmer S, Berzofsky M. *Drug Use, Dependence, and Abuse among State Prisoners and Jail Inmates, 2007-2009*. Washington, D.C.: Bureau of Justice Statistics; 2017. https://www.bjs.gov/content/pub/pdf/dudaspji0709.pdf.

7. Steadman HJ, Osher FC, Robbins PC, Case B, Samuels S. Prevalence of Serious Mental Illness Among Jail Inmates. *Psychiatric Services*. 2009;60(6):761-765. doi:10.1176/ps.2009.60.6.761

8. Patel K, Boutwell A, Brockmann BW, Rich JD. Integrating Correctional and Community Health Care For Formerly Incarcerated People Who Are Eligible For Medicaid. *Health Affairs*. 2014;33(3):468-473. doi:10.1377/hlthaff.2013.1164

9. Morrissey JP, Cuddeback GS, Cuellar AE, Steadman HJ. The role of Medicaid enrollment and outpatient service use in jail recidivism among persons with severe mental illness. *Psychiatr Serv*. 2007;58(6):794-801. doi:10.1176/ps.2007.58.6.794

10. Morrissey JP, Domino ME, Cuddeback GS. Expedited Medicaid Enrollment, Mental Health Service Use, and Criminal Recidivism Among Released Prisoners with Severe Mental Illness. *Psychiatr Serv*. 2016;67(8):842-849. doi:10.1176/appi.ps.201500305

11. Grabert BK, Gertner AK, Domino ME, Cuddeback GS, Morrissey JP. Expedited Medicaid Enrollment, Service Use, and Recidivism at 36 Months Among Released Prisoners With Severe Mental Illness. *Psychiatric Services*. 2017;68(10):1079-1082. doi:10.1176/appi.ps.201600482

12. Domino ME, Gertner A, Grabert B, Cuddeback GS, Childers T, Morrissey JP. Do timely mental health services reduce re-incarceration among prison releasees with severe mental illness? *Health Services Research*. March 2019. doi:10.1111/1475-6773.13128

13. Costopoulos JS, Plewinski AM, Monaghan PL, Edkins VA. The impact of US Government assistance on recidivism: Government assistance and recidivism. *Criminal Behaviour and Mental Health*. 2017;27(4):303-311. doi:10.1002/cbm.1997

14. Sartorius PJ, Woodbury V. *Michigan Pathways Project Links Ex-Prisoners to Medical Services, Contributing to a Decline in Recidivism*. Rockville, MD: Agency for Healthcare Quality and Research; 2014. https://innovations.ahrq.gov/profiles/michigan-pathways-project-links-ex-prisoners-medical-services-contributing-decline.

15. Kaiser Family Foundation. *Medicaid Behavioral Health Services Database Notes and Methods*. San Francisco, CA: Kaiser Family Foundation; 2018. http://files.kff.org/attachment/Survey-2018-Medicaid-Behavioral-Health-Services-Database-Notes-and-Methods.

16. U.S. Census Bureau. American Community Survey, 1-year estimates. 2015.

17. U.S. Census Bureau. American Community Survey, 5-year estimates. 2015 2011.

18. Vera Institute of Justice. *Pretrial Jail Incarceration Rate*. Washington, D.C.; 2018.

19. US Government Accountability Office. *Medicaid: Information on Inmate Eligibility and Federal Costs for Allowable Services*. Washington, D.C.; 2014. https://www.gao.gov/assets/670/665552.pdf.

20. Lopez Bernal J, Cummins S, Gasparrini A. The use of controls in interrupted time series studies of public health interventions. *Int J Epidemiol*. 2018;47(6):2082-2093. doi:https://doi.org/10.1093/ije/dyy135

21. St.Clair T, Cook TD, Hallberg K. Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time Series Design by Comparison with a Randomized Experiment. *American Journal of Evaluation*. 2014;35(3):311-327. doi:10.1177/1098214014527337

22. Shadish WR, Cook TD, Campbell DT. Quasi-Experimentation: Interrupted Time Series Designs. In: *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. 2nd ed. Cengage; 2002.

23. Rokicki S, Cohen J, Fink G, Salomon JA, Landrum MB. Inference with difference-in-differences with a small number of groups: A review, simulation study, and empirical application using SHARE data. *Medical Care*. 2018;56:97-105. doi:10.1097/MLR.0000000000000830

24. Aos S, Miller M, Drake E. *Evidence-Based Public Policy Options to Reduce Future Prison Construction, Criminal Justice Costs, and Crime Rates*. Olympia, WA: Washington Institute for Public Policy; 2006.

25. Davis LM, Bozick R, Steele JL, Saunders J, Miles JNV. *Evaluating the Effectiveness of Correctional Education: A Meta-Analysis of Programs That Provide Education to Incarcerated Adults*. Santa Monica, CA: RAND; 2013. https://www.bja.gov/Publications/RAND_Correctional-Education-Meta-Analysis.pdf.

26. Haley J, Kenney GM, Wang R, Pan C, Lynch V, Buettgens M. *Improvements in Uninsurance and Medicaid/CHIP Participation among Children and Parents Stalled in 2017*. Washington, D.C.: Urban Institute; 2019. https://www.urban.org/sites/default/files/publication/100214/improvements_in_uninsurance_and_medicaid_chip_participation_among_children_and_parents_stalled_in_2017_1.pdf.

27. Centers for Medicaid and Medicare Services. *Appendix to Eligibility & Coverage Evaluation Guidance: Retroactive Eligibility Waivers*. https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/ce-evaluation-design-guidance-retro-eligibility-appendix.pdf.

28. Courtemanche C, Marton J, Ukert B, Yelowitz A, Zapata D. Early Impacts of the Affordable Care Act on Health Insurance Coverage in Medicaid Expansion and Non-Expansion States: Impacts of the Affordable Care Act. *J Pol Anal Manage*. 2017;36(1):178-210. doi:10.1002/pam.21961

29. Baicker K, Taubman SL, Allen HL, et al. The Oregon Experiment — Effects of Medicaid on Clinical Outcomes. *New England Journal of Medicine*. 2013;368(18):1713-1722. doi:10.1056/NEJMsa1212321

30. Sommers BD, Maylone B, Blendon RJ, Orav EJ, Epstein AM. Three-Year Impacts of the Affordable Care Act: Improved Medical Care And Health Among Low-Income Adults. *Health Affairs*. 2017;36(6):1119-1128. doi:10.1377/hlthaff.2017.0293

31. Graves JA, Hatfield LA, Blot W, Keating NL, McWilliams JM. Medicaid expansion slowed rates of health decline for low-income adults in Southern states. *Health Aff (Millwood)*. 2020;39(1):67-76. doi:10.1377/hlthaff.2019.00929

32. Simon K, Soni A, Cawley J. The Impact of Health Insurance on Preventive Care and Health Behaviors: Evidence from the First Two Years of the ACA Medicaid Expansions. *Journal of Policy Analysis and Management*. 2017;36(2):390-417. doi:10.1002/pam.21972

33. Wherry LR, Miller S. Early coverage, access, utilization, and health effects associated with the Affordable Care Act Medicaid expansions: A quasi-experimental study. *Annals of Internal Medicine*. 2016;164(12):795. doi:10.7326/M15-2234

34. Sommers BD, Baicker K, Epstein AM. Mortality and access to care among adults after state Medicaid expansions. *N Engl J Med*. 2012;367(11):1025-1034. doi:10.1056/NEJMsa1202099

35. Miller S, Altekruse S, Johnson N, Wherry L. *Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data*. Cambridge, MA: National Bureau of Economic Research; 2019:w26081. doi:10.3386/w26081

36. Sommers BD, Oellerich D. The poverty-reducing effect of Medicaid. *J Health Econ*. 2013;32(5):816-832. doi:10.1016/j.jhealeco.2013.06.005

37. Sommers BD, Blendon RJ, Orav EJ, Epstein AM. Changes in Utilization and Health Among Low-Income Adults After Medicaid Expansion or Expanded Private Insurance. *JAMA Internal Medicine*. 2016;176(10):1501. doi:10.1001/jamainternmed.2016.4419

38. Finkelstein A, Taubman S, Wright B, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly Journal of Economics*. 2012;127(3):1057-1106. doi:10.1093/qje/qjs020

39. Sommers BD, Gawande AA, Baicker K. Health Insurance Coverage and Health. *N Engl J Med*. 2017;377(20):2000-2001. doi:10.1056/NEJMc1711805

40. Argys LM, Friedson AI, Pitts MM, Tello-Trillo DS. *Losing Public Health Insurance: TennCare Disenrollment and Personal Financial Distress*. Atlanta, GA: Federal Reserve Bank of Atlanta; 2017.

41. Musumeci M, Rudowitz R. *Medicaid Retroactive Coverage Waivers: Implications for Beneficiaries, Providres, and States*. Washington, D.C.: Kaiser Family Foundation; 2017. https://www.kff.org/medicaid/issue-brief/medicaid-retroactive-coverage-waivers-implications-for-beneficiaries-providers-and-states/.

42. The Centers for Medicare and Medicaid Services. Monthly Medicaid & CHIP Application, Eligibility Determination, and Enrollment Reports & Data. medicaid.gov. https://www.medicaid.gov/medicaid/national-medicaid-chip-program-information/medicaid-chip-enrollment-data/monthly-medicaid-chip-application-eligibility-determination-and-enrollment-reports-data/index.html. Published 2020.

43. Kaiser Family Foundation. Medicaid Waiver Tracker: Approved and Pending Section 1115 Waivers by State. https://www.kff.org/medicaid/issue-brief/medicaid-waiver-tracker-approved-and-pending-section-1115-waivers-by-state/#Table3. Published April 7, 2020.

44. Kaiser Family Foundation. *Federal Medical Assistance Percentage (FMAP) for Medicaid and Multiplier*. Washington, D.C.: Kaiser Family Foundation; 2020. https://www.kff.org/medicaid/state-indicator/federal-matching-rate-and-multiplier/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

45. Bureau of Labor Statistics. *Local Area Unemployment Statistics*. Washington, D.C.: Bureau of Labor Statistics, Department of Labor. https://www.bls.gov/lau/#tables.

46. Ruggles S, Flood S, Goeken R, et al. *IPUMS USA: Version 10.0 [Dataset]*. Minneapolis, MN: IPUMS; 2020.

47. Kaiser Family Foundation. *Medicaid's Role for Women*. Washington, D.C.: Kaiser Family Foundation; 2019. http://files.kff.org/attachment/Fact-Sheet-Medicaids-Role-for-Women.

48. Brooks T, Miskell S, Artiga S, Cornachione E, Gates A. *Medicaid and CHIP Eligibility, Enrollment, Renewal, and Cost Sharing Policies as of January 2016: Findings from a 50-State Survey*. Washington, D.C.: Kaiser Family Foundation; 2016. https://www.kff.org/medicaid/report/medicaid-and-chip-eligibility-enrollment-renewal-and-cost-sharing-policies-as-of-january-2016-findings-from-a-50-state-survey/.

49. Brooks T, Wagnerman K, Artiga S, Cornachione E, Ubri P. *Medicaid and CHIP Eligibility, Enrollment, Renewal, and Cost Sharing Policies as of January 2017: Findings from a 50-State Survey*. Washington, D.C.: Kaiser Family Foundation; 2017. https://www.kff.org/medicaid/report/medicaid-and-chip-eligibility-enrollment-renewal-and-cost-sharing-policies-as-of-january-2017-findings-from-a-50-state-survey/.

50. Brooks T, Wagnerman K, Artiga S, Cornachione E. *Medicaid and CHIP Eligibility, Enrollment, Renewal, and Cost Sharing Policies as of January 2018: Findings from a 50-State Survey*. Washington, D.C.: Kaiser Family Foundation; 2018. https://www.kff.org/medicaid/report/medicaid-and-chip-eligibility-enrollment-renewal-and-cost-sharing-policies-as-of-january-2018-findings-from-a-50-state-survey/.

51. Brooks T, Roygardner L, Artiga S. *Medicaid and CHIP Eligibility, Enrollment, Renewal, and Cost Sharing Policies as of January 2019: Findings from a 50-State Survey*. Washington, D.C.: Kaiser Family Foundation; 2019. https://www.kff.org/medicaid/report/medicaid-and-chip-eligibility-enrollment-and-cost-sharing-policies-as-of-january-2019-findings-from-a-50-state-survey/.

52. Bilinski A, Hatfield LA. Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. *arXiv:180503273 [stat]*. January 2020. http://arxiv.org/abs/1805.03273. Accessed April 2, 2020.

53. Goodman-Bacon A. *Difference-in-Differences with Variation in Treatment Timing*. National Bureau of Economic Research; 2018. https://www.nber.org/papers/w25018.

54. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*. 2004;119:249-275. doi:10.1162/003355304772839588

55. Cameron AC, Gelbach JB, Miller DL. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*. 2008;90(3):414-427. doi:10.1162/rest.90.3.414

56. Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*. 2015;50(2):317-372. doi:10.3368/jhr.50.2.317

57. Rokicki S, Cohen J, Fink G, Salomon JA, Landrum MB. Inference with difference-in-differences with a small number of groups: A review, simulation study, and empirical application using SHARE data. *Medical Care*. 2018;56:97-105. doi:10.1097/MLR.0000000000000830

58. Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*. 2010;105:493-505. doi:10.1198/jasa.2009.ap08746

59. McClelland R, Gault S. *The Synthetic Control Method as a Tool to Understand State Policy*. Washington, D.C.: Urban Institute; 2017. https://www.urban.org/sites/default/files/publication/89246/the_synthetic_control_method_as_a_tool_0.pdf.

60. Brooks T, Park E, Roygardner L. *Medicaid and CHIP Enrollment Decline Suggests the Child Uninsured Rate May Rise Again*. Washington, D.C.: Georgetown University Health Policy Institute Center for Children and Families; 2019. https://ccf.georgetown.edu/wp-content/uploads/2019/06/Enrollment-Decline.pdf.

61. Sommers BD, Fry CE, Blendon RJ, Epstein AM. New Approaches in Medicaid: Work Requirements, Health Savings Accounts, and Health Care Access. *Health Affairs*. 2018;37(7):1099-1108. doi:10.1377/hlthaff.2018.0331

62. LaLonde RJ. Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *The American Economic Review*. 1986;76(4):604-620.

63. Cook TD, Shadish WR, Wong VC. Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *J Pol Anal Manage*. 2008;27(4):724-750. doi:10.1002/pam.20375

64. Fretheim A, Zhang F, Ross-Degnan D, et al. A reanalysis of cluster randomized trials showed interrupted time-series studies were valuable in health system evaluation. *Journal of Clinical Epidemiology*. 2015;68(3):324-333. doi:10.1016/j.jclinepi.2014.10.003

65. Ferraro PJ, Miranda JJ. Panel Data Designs and Estimators as Substitutes for Randomized Controlled Trials in the Evaluation of Public Programs. *Journal of the Association of Environmental and Resource Economists*. 2017;4(1):281-317. doi:10.1086/689868

66. Antonisse L, Garfield R, Rudowitz R, Guth M. *The Effects of Medicaid Expansion under the ACA: Updated Findings from a Literature Review*. Washington, D.C.: Kaiser Family Foundation; 2019. https://www.kff.org/medicaid/issue-brief/the-effects-of-medicaid-expansion-under-the-aca-updated-findings-from-a-literature-review-august-2019/.

67. Hartman E, Hidalgo FD. An equivalence approach to balance and placebo tests. *American Journal of Political Science*. 2018. doi:10.1111/ajps.12387

68. Mora R, Reggio I. Alternative diff-in-diffs estimators with several pretreatment periods. *Econometric Reviews*. 2019;38(5):465-486. doi:10.1080/07474938.2017.1348683

69. Athey S, Imbens GW. Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*. 2006;74(2):431-497. doi:10.1111/j.1468-0262.2006.00668.x

70. Daw JR, Hatfield LA. Matching and regression-to-the-mean in difference-in-differences analysis. *Health Services Research*. 2018;53(6):4138-4156. doi:10.1111/1475-6773.12993

71. Conley TG, Taber CR. Inference with "difference in differences" with a small number of policy changes. *The Review of Economics and Statistics*. 2011;93:113-125. doi:10.1162/REST_a_00049

72. Chabé-Ferret S. *Should We Combine Difference in Differences with Conditioning on Pre-Treatment Outcomes?* Toulouse School of Economics; 2017. https://www.tse-fr.eu/publications/should-we-combine-difference-differences-conditioning-pre-treatment-outcomes.

73. Wing C, Simon K, Bello-Gomez RA. Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*. 2018;39(1):453-469. doi:10.1146/annurev-publhealth-040617-013507

74. Angrist JD, Pischke J-S. *Mostly Harmless Econometrics: An Empiricist's Compansion*. Princeton, NJ: Princeton University Press; 2009.

75. Baicker K, Svoronos T. Testing the Validity of the Single Interrupted Time Series Design. *NBER Working Paper No 26080*. July 2019. https://www.nber.org/papers/w26080.pdf.

76. Bloom HS, Riccio JA. Using Place-Based Random Assignment and Comparative Interrupted Time-Series Analysis to Evaluate the Jobs-Plus Employment Program for Public Housing Residents. *The ANNALS of the American Academy of Political and Social Science*. 2005;599(1):19-51.

77. Fry CE, McGuire T, Frank RG. Medicaid Expansion's Spillover to the Criminal Justice System: Evidence from Six Urban Counties. *Working Paper*. 2020.

78. Frank RG, Fry CE. The impact of expanded Medicaid eligibility on access to naloxone. *Addiction*. 2019;114(9):1567-1574. doi:10.1111/add.14634

79. Powell D, Alpert A, Pacula RL. A Transitioning Epidemic: How the Opioid Crisis Is Driving The Rise In Hepatitis C. *Health Affairs*. 2019;38(2):287-294. doi:10.1377/hlthaff.2018.05232

Appendix A

Table A.1: Sample state differences in Medicaid coverage of behavioral health services

<table>
<tr><th></th><th colspan="2">Inpatient</th><th>Residential</th><th colspan="4">Outpatient</th></tr>
<tr><th></th><th>Psychiatric</th><th>Detoxification</th><th>Psychiatric</th><th>Individual therapy</th><th>Group therapy</th><th>Buprenorphine</th><th>Methadone</th></tr>
<tr><td colspan="8">Midwest</td></tr>
<tr><td>MN</td><td colspan="2">Yes</td><td>Yes</td><td colspan="2">Yes</td><td colspan="2">Yes</td></tr>
<tr><td>WI</td><td colspan="2">$3/day; $75/stay</td><td>No</td><td colspan="2">$0.50 – $3 copay</td><td colspan="2">Copay of up to $12/month</td></tr>
<tr><td colspan="8">Southwest</td></tr>
<tr><td>AZ</td><td rowspan="2">15-day max</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td>TX</td><td>Requires prior authorization</td><td>No</td><td colspan="2">30/year</td><td>Yes</td><td>Only at OTP</td></tr>
<tr><td colspan="8">Southeast</td></tr>
<tr><td>LA</td><td colspan="2">Yes</td><td>No</td><td colspan="2">Yes</td><td>$0.50 - $3 copay</td><td>No</td></tr>
<tr><td>MS</td><td colspan="2">Requires prior authorization</td><td>No</td><td>36/year</td><td>24/year</td><td colspan="2">$3 copay & prior authorization</td></tr>
</table>

SOURCES/NOTES All data comes from the Kaiser Family Foundation Medicaid Behavioral Health Services Database. Available at: https://www.kff.org/data-collection/medicaid-behavioral-health-services-database/.

Table A.2: Falsification test of relationship between Medicaid expansion and the probability of re-arrest

	Change in intercept		Change in slope	
	Coefficient	*95% CI*	*Coefficient*	*95% CI*
Midwest				
Month 7	-0.71	-2.28, 0.85	-0.07	-0.13, -0.01
Month 10	-0.88	-2.59, 0.82	-0.09	-0.14, -0.00
Month 13	-0.73	-3.70, 2.23	-0.09	-0.18, -0.00
Month 18	-0.87	-0.38, -1.12	-0.03	-0.04, -0.01
Southwest				
Month 7	-0.91	-1.93, 0.10,	-0.07	-0.11, -0.03
Month 10	-0.35	-1.55, 0.84	-0.08	-0.12, -0.04
Month 13	-0.97	-3.09, 1.15	-0.21	-0.27, -0.15
Month 18	-2.00	-2.44, -1.56	-0.07	-0.08, -0.06
Southeast				
Month 7	-0.28	-1.18, 1.13	-0.02	-0.08, 0.03
Month 10	-0.03	-1.69, 1.63	-0.03	-0.09, 0.02
Month 13	-0.34	-0.70, 0.66	0.08	-0.02, 0.17
Month 18	0.04	-0.50, 0.58	0.07	0.06, 0.09

SOURCES/NOTES: **Sources** Authors' analyses of arrest data from county jails. Observations are at the person-month level **Notes** Estimates are from comparative interrupted time series regressions. Regressions for the likelihood of re-arrest are linear probability models. Regressions for the number of arrests are Poisson regression models. Each full sample regression is adjusted with gender and prior contact with the criminal justice system (in the pre-period) and an interaction between these variables and the running monthly counter to account for a time-varying relationship between the outcome and the covariates. The Midwest pair also adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African American and the interaction of this variable with the monthly time trend. § denotes that p-value is not statistically significant after Bonferroni adjustment for multiple comparisons.

Table A.3: Falsification test of relationship between Medicaid expansion and the number of arrests

	Change in intercept		Change in slope	
	Coefficient	*95% CI*	*Coefficient*	*95% CI*
Midwest				
Month 7	0.003	-0.05, 0.06	-0.0001	-0.002, 0.002
Month 10	-0.004	-0.06, 0.14	-0.001	-0.002, 0.001
Month 13	0.02	-0.08, 0.06	-0.0001	-0.004, 0.0003
Month 18	-0.04	-0.06, -0.02	-0.001	-0.002, -0.001
Southwest				
Month 7	0.03	-0.02, 0.08	0.003	0.005, 0.001
Month 10	0.04	-0.01, 0.08	0.002	0.0005, 0.004
Month 13	0.00	-0.07, 0.07	-0.004	-0.006, -0.002
Month 18	-0.08	-0.10, -0.06	-0.003	-0.003, -0.002
Southeast				
Month 7	0.01	-0.04, 0.07	0.0002	-0.002, 0.002
Month 10	-0.01	0.01, 0.05	-0.002	-0.004, -0.0004
Month 13	-0.07	-0.06, 0.04	0.003	0.001, 0.005
Month 18	-0.004	-0.02, 0.01	0.004	0.003, 0.004

SOURCES/NOTES: **Sources** Authors' analyses of arrest data from county jails. Observations are at the person-month level **Notes** Estimates are from comparative interrupted time series regressions. Regressions for the likelihood of re-arrest are linear probability models. Regressions for the number of arrests are Poisson regression models. Each full sample regression is adjusted with gender and prior contact with the criminal justice system (in both the pre- and post-period) and an interaction between these variables and the running monthly counter to account for a time-varying relationship between the outcome and the covariates. The Midwest pair also adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African American and the interaction of this variable with the monthly time trend. [§] denotes that p-value is not statistically significant after Bonferroni adjustment for multiple comparisons.

Table A.4: Comparison of estimates with full post-period (24 months) and truncated post-period (18 months)

	Probability of Re-arrest		Number of Arrests	
	Change in Level	*Change in Slope*	*Change in Level*	*Change in Slope*
Midwest				
24 months post	-0.87	-0.03	-0.04	-0.001
18 months post	-0.96	-0.01	-0.04	-0.001
Southwest				
24 months post	-2.00	-0.07	-0.08	-0.003
18 months post	-1.84	-0.07	-0.07	-0.003
Southeast				
24 months post	0.04	0.07	-0.005	0.004
18 months post	0.17	0.05	0.14	0.004

Sources/Notes: SOURCES Authors' analyses of arrest data from county jails. Observations are at the person-month level NOTES Estimates are from comparative interrupted time series regressions. Regressions for the likelihood of re-arrest are linear probability models. Regressions for the number of arrests are Poisson regression models. Each full sample regression is adjusted with gender and prior contact with the criminal justice system (in both the pre- and post-period) and an interaction between these variables and the running monthly counter to account for a time-varying relationship between the outcome and the covariates. The Midwest pair also adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African American and the interaction of this variable with the monthly time trend.

Table A.5: Comparison of individual-level and county-level CITS standard errors for estimates of the change in the probability of re-arrests and the number of arrests

	Probability of Re-arrest		**Number of Arrests**	
	Change in Intercept	*Change in Slope*	*Change in Intercept*	*Change in Slope*
Midwest				
Arrestee-level	0.11	0.007	0.038	0.002
County-level	0.57*	0.03*	0.006	0.0003
Southwest				
Arrestee-level	0.09	0.005	0.03	0.001
County-level	0.69	0.03	0.02	0.001
Southeast				
Arrestee-level	0.13	0.008	0.51	0.003
County-level	0.44	0.02	0.01	0.001

SOURCES/NOTES: **Sources** Authors' analyses of arrest data from county jails. **Notes** Estimates are from comparative interrupted time series regressions. Regressions for the likelihood of re-arrest are linear probability models. Regressions for the number of arrests are ordinary least squares regression models. The Midwest pair adjusts for whether the arrest was a felony or misdemeanor and the interaction of this variable with the monthly counter. The Southwest county pair also adjusts for whether the arrest was for a parole violation and for whether the arrestee was Hispanic/Latino plus the interactions of these two variables with the monthly counter. Regressions using the Southeast county pair also adjust for whether the arrestee was African American and the interaction of this variable with the monthly time trend. *denotes that p-value becomes non-significant at the county-level analysis compared to the arrestee-level analysis.

Appendix B

Table B.1: Comparison of outcome and state characteristics for retroactive eligibility state and comparators prior to retroactive eligibility implementation

	AR	Comparators	FL	Comparators	IA	Comparators	NH	Comparators
Medicaid enrollment (log-transformed)	13.7 (0.05)	13.8 (0.36)	15.2 (0.01)	13.8 (0.40)	13.3 (0.04)	13.8 (1.17)	12.1 (0.01)	13.6 (1.28)
Unemployment rate	4.0 (0.16)	5.7 (0.80)	3.7 (0.23)	4.0 (0.50)	3.2 (0.19)	4.0 (0.74)	2.6 (0.09)	3.7 (0.67)
% of population that is child-bearing age	29.6 (--)	30.0 (0.41)	29.6 (0.03)	30.6 (0.32)	29.1 (0.08)	29.7 (0.78)	30.6 (0.20)	31.2 (0.36)
FMAP	69.9 (0.14)	68.3 (3.48)	61.7 (0.27)	70.6 (3.36)	56.9 (0.50)	59.8 (6.40)	50.0 (--)	51.0 (0.01)
Section 1931 Parent FPL	16% (--)	25% (10%)	33% (--)	49% (30%)	51% (--)	70% (41%)	55% (0.7%)	105% (37%)
N	12	60	12	60	12	60	12	60

SOURCES/NOTES: **Sources** Monthly Medicaid enrollment is from CMS' updated enrollment files and includes all Medicaid and CHIP enrollment. The seasonally adjusted monthly unemployment rate is from the Bureau of Labor Statistics. The proportion of the population that is of child-bearing age is the authors' analysis of micro data from the American Community Survey made publicly available via IPUMS. The FMAP (Federal Medical Assistance Percentages) is reported for each fiscal year and was obtained from the Office of the Assistant Secretary for Planning and Evaluation at the Department of Health and Human Services. The FPL eligibility for Section 1931 parents was obtained from the Kaiser Family Foundation's annual survey reports on the Medicaid program.

Details of difference-in-differences analysis

DID is a quasi-experimental design that assumes that the change in enrollment seen in the comparison group from the pre-period to the post-period would have been the change seen in the retroactive eligibility state *if not for the implementation of the retroactive eligibility waiver*. The strictest form of the DID counterfactual assumption assumes that the difference between the untreated outcomes in the treated and comparison group are constant at each time point in the pre and post period – this is often referred to as the "parallel trends" assumption.

We estimate four log-linear DID regression models in each comparative case study – one with and one without group-specific linear trends. If treatment effect estimates differ significantly between these two specifications, then it is likely that the parallel trends assumption is violated. Typically, in DID, researchers estimate an average treatment effect over the entire post period. Sometimes, researchers may estimate a treatment effect at each post-period time point to trace out the dynamics of the intervention on the outcome. We estimate an average treatment effect, as well as time-varying treatment effects for both specifications. Explicitly, we specify our regression models in the following way:

Equation 1: Average treatment effect without group-specific trends

$$y_{it} = \beta_0 + \beta_1 RE_i + \beta_2 post_t + \beta_3 (RE_i * post_t) + \boldsymbol{\gamma X_{it}} + \boldsymbol{\lambda}_i + \epsilon_{it},$$

where:
y_{it} is the log-transformed Medicaid enrollment in state i in month t,
β_0 is the pre-period average in the comparison group,
β_1 is the differential pre-period average in the retroactive eligibility state,
β_2 is the post-period average in the comparison group,
β_3 is the coefficient of interest; the differential post-period average in the retroactive eligibility state,
γ is a vector of coefficients for time-varying state-level covariates,
λ are state fixed effects, and
ϵ_{it} is the residual error term.

Equation 2: Average treatment effect with group-specific trends

$$y_{it} = \beta_0 + \beta_1 RE_i + \beta_2 post_t + \beta_3 (RE_i * post_t) + \beta_4 * trend_t + \beta_5 (trend_t * RE_i) + \boldsymbol{\gamma X_{it}} + \boldsymbol{\lambda}_i + \epsilon_{it},$$

where the differential components from Equation 1 are:
β_4, which is the linear trend in the comparison group in the pre-period,
β_5, which is the differential linear trend in the retroactive eligibility state in the pre-period.

Equation 3: Dynamic treatment effects without group-specific trends

$y_{it} = \beta_0 + \beta_1 RE_i + \beta_2 post_t + \sum_{k=t_0}^{T} \beta_k I_{k=t} + \boldsymbol{\gamma X_{it}} + \boldsymbol{\lambda_i} + \epsilon_{it},$

where the differential component from Equation 1 is:
$\sum_{k=t_0}^{T} \beta_k I_{k=t}$, which is the treatment effect at post-period time, *k*.

Equation 4: Dynamic treatment effects with group-specific trends

$y_{it} = \beta_0 + \beta_1 RE_i + \beta_2 post_t + \sum_{k=t_0}^{T} \beta_k I_{k=t} + \beta_4 * trend_t + \beta_5 (trend_t * RE_i) + \boldsymbol{\gamma X_{it}} + \boldsymbol{\lambda_i} + \epsilon_{it},$

where the differential components from Equation 3 are:
β_4, which is the linear trend in the comparison group in the pre-period,
β_5, which is the differential linear trend in the retroactive eligibility state in the pre-period.

Synthetic Control Method

In addition to the difference-in-differences analysis above, we also estimate the relationship using the synthetic control method (SCM).[29,30] SCM was developed for comparative case study analysis, such as this, where there is only one treated unit. In SCM, a data-driven algorithmic process uses covariates related to the outcome and treatment to assign weights to a set of potential comparison units with the goal of creating a composite or synthetic comparison group that is nearly identical to the treated group on pre-period observables. Additionally, researchers often use lagged pre-period outcomes to create a synthetic control group's trend that mimics the evolution of the outcome in the treated group. As with DID, the trend seen in the synthetic comparison group represents an approximation for the treated group's outcomes *in the absence of treatment*.[30,31] One way to assess the fit of SCM is to "eye-ball" the pre-period trends, but this method is crude. Another way is to assess the root mean squared prediction error (RMSPE), which we want to minimize.

Deviations from the comparison group's trend in the post-period represent the causal effects of the retroactive eligibility waiver on Medicaid enrollment. The differences in outcome between the treated group and synthetic comparison group can either be averaged for the whole post period or assessed at each post-period time point.

There are a number of potential issues with SCM. One is that the outcomes of the treated state are assumed to be within the 'convex hull' of the outcomes of the untreated states. This means that there is exists a weighted combination of untreated states that will produce the pre-treatment outcomes of the treated state exactly. However, this assumption can be relaxed if the treated state comprises the any of the synthetic controls for the untreated states. We, thus, constructed a synthetic control group for every member of a state's donor pool. For each retroactive eligibility state, the treated state appears in the donor pool for at least one of its donor states (Table B.4).

Table B.2: Covariate balance between retroactive eligibility state and synthetic control

	AR	Synthetic AR	FL	Synthetic FL	IA	Synthetic IA	NH	Synthetic NH
Lagged outcome	13.7	13.7	15.2	15.2	13.4	13.4	12.1	12.1
Unemployment rate	4.0	4.0	3.69	4.01	3.23	3.32	2.61	2.69
FMAP	69.9	62.7	61.7	57.8	56.9	56.9	50.0	52.4
Section 1913 FPL	16.0%	43.1%	33.0%	55.6%	51.2%	56.0%	55.3%	53.5%
Medicaid expansion	1	0.98	0	1.00	1	1	1	1
12-month eligibility	1	084	1	0.817	1	0.196	0	0.23
% of child-bearing age	29.6	29.9	29.6	30.2	29.1	29.6	30.6	30.4
Total population	2.31M	3.55M	17.1M	19.0M	1.11M	1.00M	1.11M	1.00M

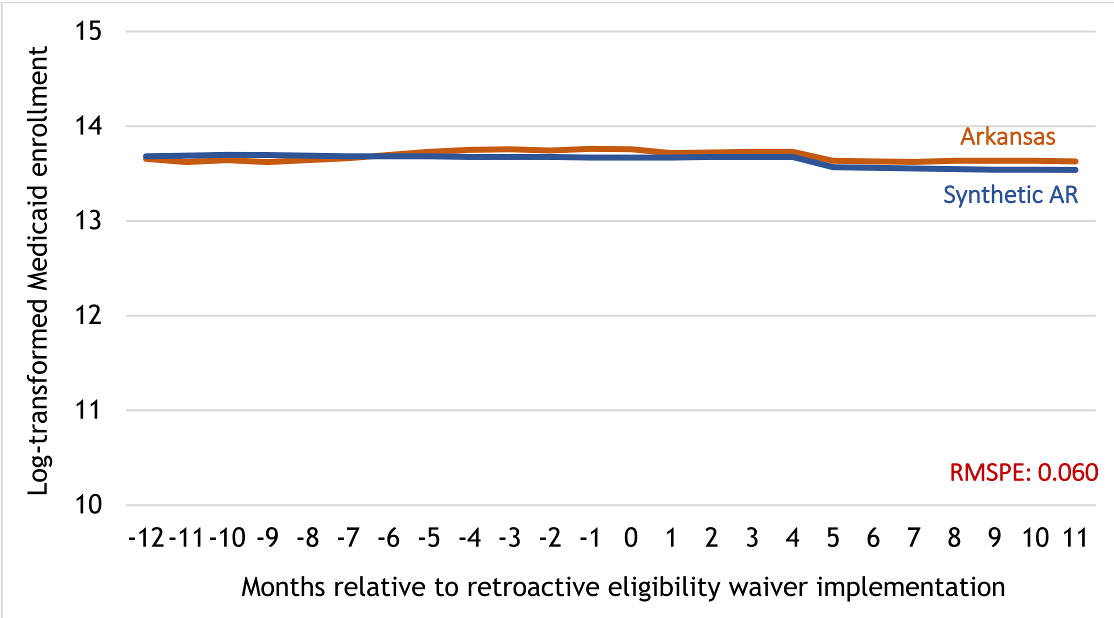
SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, the total population is from the Bureau of Labor Statistics, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** Covariate balance is obtained via synthetic control method using the covariates in this table. All covariates are averaged for the entire 12-month post period for each retroactive eligibility state.

Table B.3: Donor states and their respective weights for SCM analysis

Arkansas		Florida		Iowa		New Hampshire	
Donor State	*Weight*	*Donor State*	*Weight*	*Donor State*	*Weight*	*Donor State*	*Weight*
Colorado	0.254	California	0.331	Colorado	0.051	Massachusetts	0.092
Idaho	0.018	Indiana	0.184	Idaho	0.018	North Dakota	0.23
Indiana	0.158	Michigan	0.196	North Dakota	0.198	Vermont	0.678
Oregon	0.343	New York	0.008	Oregon	0.289		
Utah	0.227	Texas	0.282	Wisconsin	0.445		
Sum	**1.00**	**Sum**	**1.00**	**Sum**	**1.00**	**Sum**	**1.00**

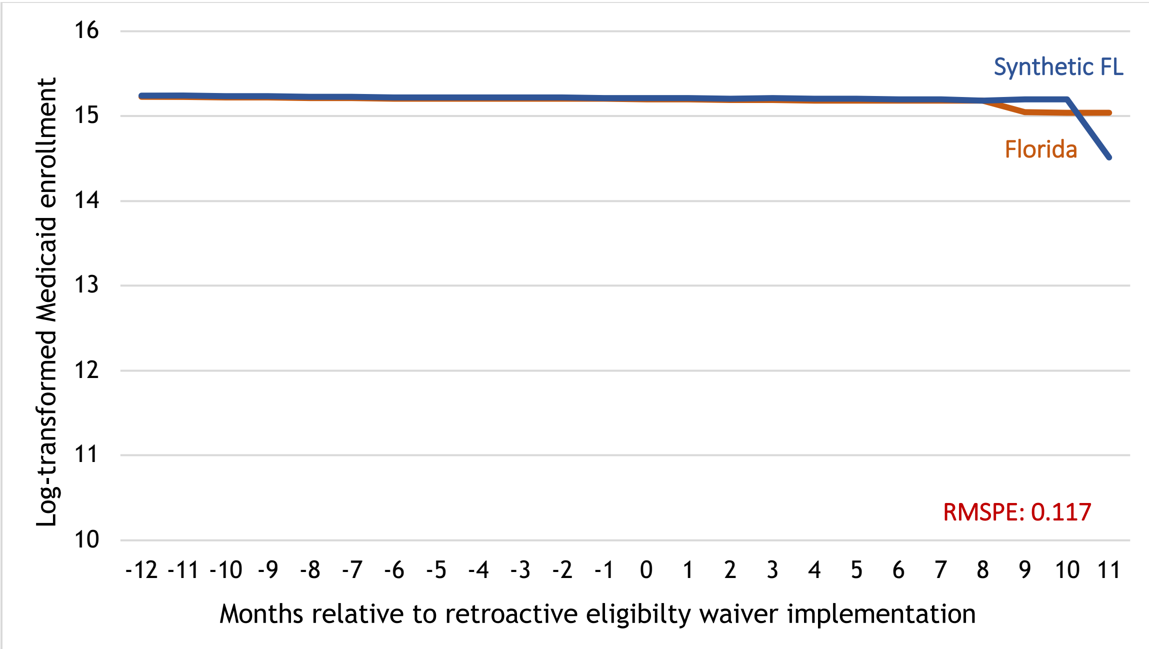
SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. **Notes** Donor pool and weights were obtained via synthetic control method using one-month lagged outcomes, the seasonally adjusted unemployment rate, the state's total population, the fiscal year FMAP, eligibility thresholds for Section 1913 parents, state expansion status, whether states allow for 12 months continuous enrollment, whether states allow for automatic re-enrollment, and the proportion of the population of child-bearing age.

Figure B.1: SCM in Arkansas



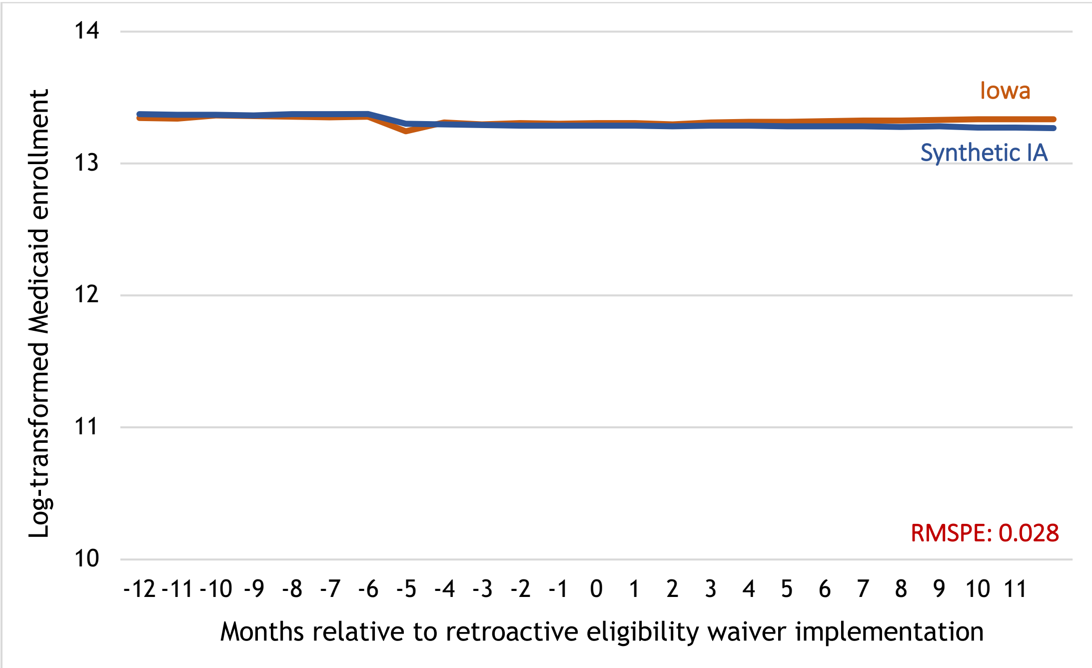
SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, total population, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** Outcome for Arkansas (retroactive eligibility state) and synthetic Arkansas shown here for 12 months before and 12 months after implementation. In calendar time, this is from January 2016 – December 2018 with implementation of retroactive eligibility waiver occurring in January 2017. Potential donor pool consists of all states except those that implemented a retroactive eligibility waiver during this time period (Iowa, New Hampshire, and Florida).

Figure B.2: SCM in Florida



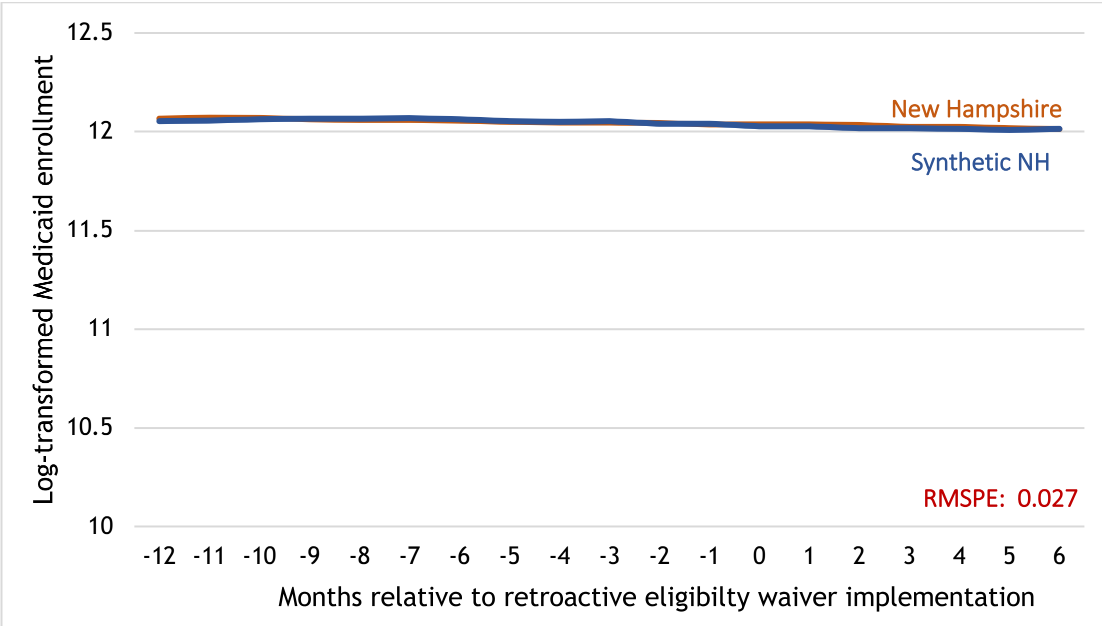
SOURCES/NOTE: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, the total population, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** Outcome for Florida (retroactive eligibility state) and synthetic Florida shown here for 12 months before and 12 months after implementation. In calendar time, this is from November 2017 – October 2019 with implementation of retroactive eligibility waiver occurring in November 2018. Potential donor pool consists of all states except those that implemented a retroactive eligibility waiver during this time period (Iowa, New Hampshire, and Arkansas) and New Mexico which did not have outcome data for September 2019.

Figure B.3: SCM in Iowa



SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, the total population, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** Outcome for Iowa (retroactive eligibility state) and synthetic Iowa shown here for 12 months before and 12 months after implementation. In calendar time, this is from November 2016 – October 2018 with implementation of retroactive eligibility waiver occurring in November 2017. Potential donor pool consists of all states except those that implemented a retroactive eligibility waiver during this time period (Arkansas, New Hampshire, and Florida).

Figure B.4: SCM in New Hampshire



SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, the total population, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** Outcome for New Hampshire (retroactive eligibility state) and synthetic New Hampshire shown here for 12 months before and 6 months after implementation. In calendar time, this is from November 2017 – June 2019 with implementation of retroactive eligibility waiver occurring in November 2018. The post period was truncated because New Hampshire's work requirement was implemented in July 2019, and Medicaid work requirements may result in disenrollment or less enrollment. Potential donor pool consists of all states except those that implemented a retroactive eligibility waiver during this time period (Iowa, Arkansas, and Florida).

Table B.4: SCM Convex hull analysis

	Donor List
Arkansas	
Colorado	**Arkansas**, Hawaii, Massachusetts, New York
Idaho	Kansas, North Dakota, South Dakota, Utah
Indiana	Arizona, Colorado, Kentucky, Minnesota, New York, Texas, Wisconsin
Oregon	**Arkansas**, Colorado, DC, Kentucky, Louisiana, New Jersey, South Carolina, West Virginia
Utah	**Arkansas**, Idaho, North Dakota, South Dakota, Texas
Florida	
California	New York
Indiana	Colorado, **Florida**, Kentucky, Michigan, Missouri
Michigan	**Florida**, Indiana, Kentucky, New York, North Carolina, Ohio
New York	California, Illinois, Massachusetts, Ohio
Texas	California, **Florida**, Vermont, Wyoming
Iowa	
Colorado	Hawaii, Indiana, **Iowa**, Massachusetts, New York
Idaho	Kansas, Maine, Mississippi, North Dakota, South Dakota, Utah
North Dakota	South Dakota, Wyoming
Oregon	Alabama, Colorado, Illinois, Indiana, **Iowa**, Nevada, Tennessee
Wisconsin	California, Hawaii, **Iowa**, Minnesota, Tennessee
New Hampshire	
Massachusetts	DC, Minnesota, New York
North Dakota	**New Hampshire**, South Dakota, Wyoming
Vermont	DC, Idaho, **New Hampshire**, North Dakota

SOURCES/NOTES: **Sources** Authors' analyses of Medicaid monthly enrollment data from CMS. The seasonally adjusted unemployment rate is from the Bureau of Labor Statistics, the total population, the fiscal year FMAP is from the Office of the Assistant Secretary of Planning and Evaluation, eligibility limits for Section 1931 parents is from the Kaiser Family Foundation, data on whether states allow 12 months of continuous eligibility are from Kaiser Family Foundation, and the proportion of the population that are of child-bearing age is derived from the American Community Survey micro data obtained from IPUMS. **Notes** For each donor state in the retroactive eligibility SCM analysis, we conducted the same SCM method and then obtained the donor states for the retroactive eligibility donor states. If the retroactive eligibility state appears in the donor list of at least one list for its donor states, then the retroactive eligibility state likely sits within the convex hull of the potential pool of donors.