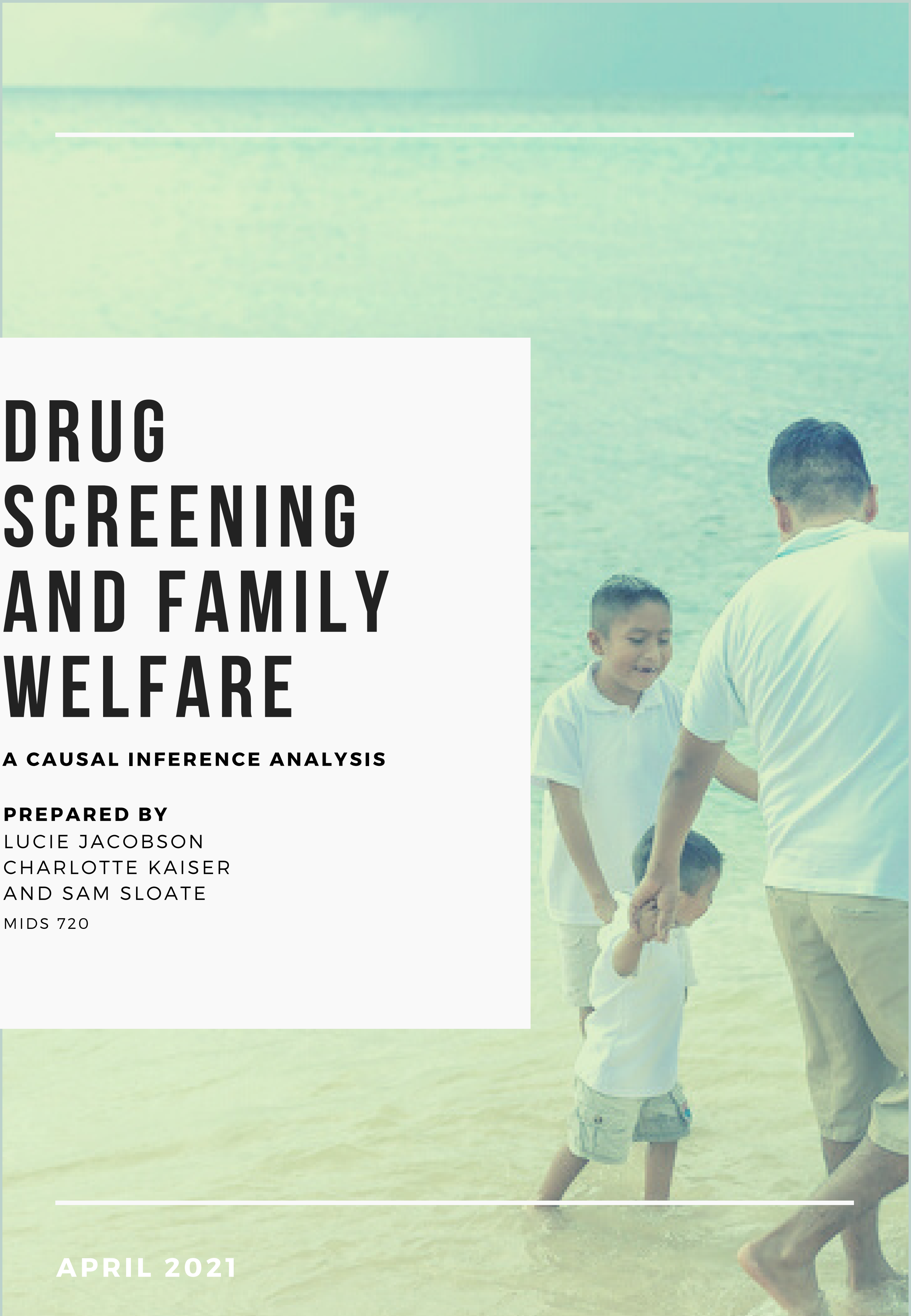


DRUG SCREENING AND FAMILY WELFARE

A CAUSAL INFERENCE ANALYSIS

PREPARED BY
LUCIE JACOBSON
CHARLOTTE KAISER
AND SAM SLOATE
MIDS 720

APRIL 2021

EXECUTIVE SUMMARY

DRUG SCREENING QUESTIONNAIRES CAUSE FAMILIES TO **AVOID APPLYING FOR ASSISTANCE**

65

FEWER FAMILIES

RECEIVE WELFARE PER 100,000 PEOPLE NATIONWIDE AFTER DRUG SCREENING LAWS WERE ENACTED

State welfare offices administer welfare (TANF) to those below certain income thresholds. Recently, many states have implemented drug screening questionnaires to determine whether applicants should undergo a drug test.

Although applicants are still entitled to TANF if they screen and test positive, this policy may be deterring people from applying for public assistance.

We find evidence of a causal link between implementation of a state drug screening questionnaire and a reduction in TANF caseload across the United States. On average, 65 families per 100,000 people drop off the TANF caseload following policy implementation, or 2,084 families total across the United States. The average caseload per state was 251 families in 2019.

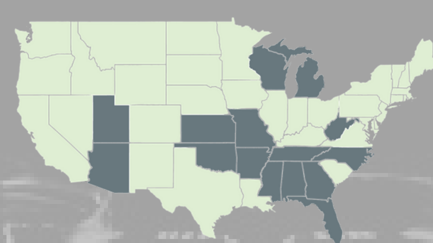
We validate this finding in an analysis of select states. Tennessee and Michigan see decreases in caseloads that exceed declining trends among neighboring states, while Utah saw no significant change in TANF caseload.

THE PROBLEM

In recent years, 15 state welfare offices have implemented drug screening questionnaires for those applying for the nation's public assistance program: Temporary Assistance for Needy Families (TANF). An average of 251 families per state received TANF in 2019, down from 677 families per state in 2001. Although lower caseloads may mean less poverty, drug screening policies may have contributed to this decline by deterring people from applying. In states with a drug screening policy, an applicant suspected of drug use is required to take a drug test before receiving welfare. A positive drug test usually triggers a referral to substance use disorder treatment program.

Despite the fact that drug tests are meant only to identify clients in need of additional support, and clients are entitled to TANF regardless of the results of their drug test, this drug screening policy may deter people from applying for welfare and thus may decrease the TANF caseload. A caseload reduction has certain negative effects. For example, children may be left without food or resources because of their parent's choice to not apply for welfare for fear of a drug test. Parents may not be referred to services they need if they do not want to fill out a drug questionnaire and potentially have to undergo a drug screening.

Caseworkers see TANF clients on a semi-regular basis as part of the program requirements, but they cannot see people who do not come in. If drug screening requirements deter people from applying for TANF, this policy may create a barrier to accessing supportive services for those who may most need services.



States with TANF Drug Screening Policies

Alabama
Arizona
Arkansas
Florida
Georgia
Kansas
Michigan
Mississippi
Missouri
North Carolina
Oklahoma
Tennessee
Utah
West Virginia
Wisconsin

THE PROBLEM

Although policymakers may want to prevent drug users from accessing state benefits, drug testing welfare applicants in practice catches very few users and is not cost-effective. Thirteen states screened more than 263,000 suspected drug users in 2019 , of which less than 1% tested positive.[1] Taxpayers paid over $200,000 for these tests, which is more than taxpayers save from the lower caseloads. Studies do not even find consistent evidence that TANF recipients are more likely to use drugs than the general population.[2] Thus, if drug screens prevent people from applying for welfare and do not have significant economic or practical benefits, they may be doing more harm than good.

1. Gomez, Amanda Michelle, and Josh Israel. "What 13 States Discovered after Spending Hundreds of Thousands Drug Testing the Poor." ThinkProgress, 26 Apr. 2019, archive.thinkprogress.org/states-cost-drug-screening-testing-tanf-applicants-welfare-2018-results-data-0fe9649fa0f8/.
2. "Drug Testing Welfare Recipients: Recent Proposals and Continuing Controversies." Office of the Assistant Secretary for Planning and Evaluation, United States Department of Health and Human Services., 21 Feb. 2017, aspe.hhs.gov/basic-report/drug-testing-welfare-recipients-recent-proposals-and-continuing-controversies#How.

THE QUESTION

Did the implementation of drug screening policies significantly reduce state TANF caseloads?

If drug screening polices **caused** decreases in states' welfare caseload, we should see **large decreases** post-policy implementation as compared to other states without such policies.

If drug screening polices **did not cause** decreases in states' welfare caseload, we should see **no change**, or increases, in post-policy implementation as compared to other states without such policies.

THE APPLICATION

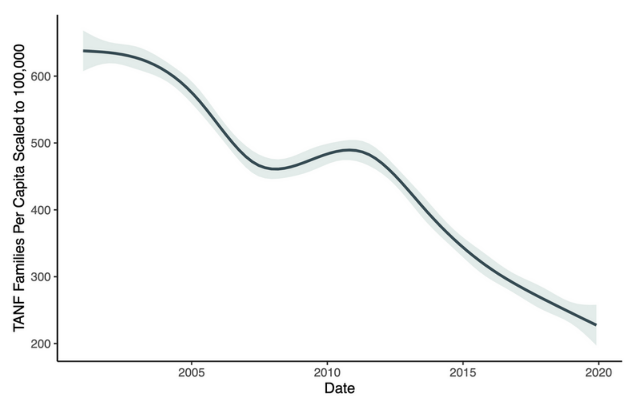
Average number of families on TANF per state:

In 2001: **677**
In 2019: **251**

TANF caseloads have been falling over time. Knowing whether drug screening laws caused part of this decline can help policymakers understand the situation and decide if and how to remedy it.

The number of families on TANF per 100,000 people has declined over time.

Generalized Additive Model Regression of TANF Families per Capita Scaled to 100,000 (2001-2019)



THE DESIGN

DATA SOURCES

The data for these analyses come from the following sources:

States with TANF Policies
National Conference of State Legislatures

TANF Caseload Data
U.S. Office of Family Assistance

Federal Poverty Level Data
IPUMS USA

State Unemployment Rate Data
Bureau of Labor Statistics

State Population Data
U.S. Census Bureau

Additional Demographic Data
IPUMS USA

DATA NOTES

Data from 2001-2019 were used for this analysis, as 2000 Census data measured poverty levels differently than 2001 and onwards. Some 2020 data were also not available for some controls and so were omitted. All states were included with the exception of the District of Columbia and Puerto Rico.

To answer the question of *whether drug screening policies significantly reduced state TANF caseloads*, we used a causal inference design.

First, we perform a state-level analysis for three states that implemented drug screening policies: Tennessee, Michigan, and Utah. States with screening policies fell broadly in three geographic regions: West, North, and Southeast. One state was selected from each region and compared to neighboring states with similar pre-policy trends in caseload. Neighboring states are likely to have comparable demographic characteristics and face similar shocks, which makes it likely that a policy state's outcomes would look similar if no policy was put in place.

To assess the drug screening policy effect in these three states, we use a difference-in-difference (DD) model. It is not enough to look at a state before and after their policy; TANF caseloads may have already been falling for external reasons. Our DD model compares select states to control states and looks for an effect above and beyond the existing trends in states that did not implement policies. The goal of a difference-in-difference analysis is to prove that outcomes for a policy state would be similar to other states if the policy state did not implement a screening policy.

THE DESIGN

WHY CAUSAL INFERENCE?

Analyses that rely on observations or correlations are not enough to determine whether or not drug screening polices caused a change in TANF caseloads. For example, a rash of tornadoes struck central and southern states in 2014, which is the same year that Mississippi implemented a drug screening policy. An observational analysis may see TANF cases decline sharply, but may not identify the tornado as the cause. By controlling for yearly shocks and comparing Mississippi to neighboring states, we may see that TANF cases did not decline *more than* comparison states. A causal analysis would indicate no policy effect, whereas a correlational study would.

Legislators need to know the true effects of their policies in order to continue, amend, or terminate them. Thus, causal research designs are necessary in policy evaluation.

FURTHER READING

For more information on causal analysis, see the web resource from Duke University's Professor Fresh here.

The difference-in-difference statistical method helps confirm that any changes in TANF caseloads came from the policy itself and not external shocks. For example, a policy change at the federal level around the same time might have reduced caseloads. However, this reduction would be seen in all states. If Michigan's caseload decreased around the time of its drug screening policy, but Indiana's caseload did also, there would be no effect of the drug screening policy above and beyond the trend in Indiana.

Additionally, we control for factors that may affect TANF caseloads, such as changes in a state's unemployment rate, which helps confirm that any effect of drug screening laws on caseloads is causal and not merely correlational.

Last, we perform a DD analysis of all states with policies as compared to all states without policies to assess the aggregate effect.

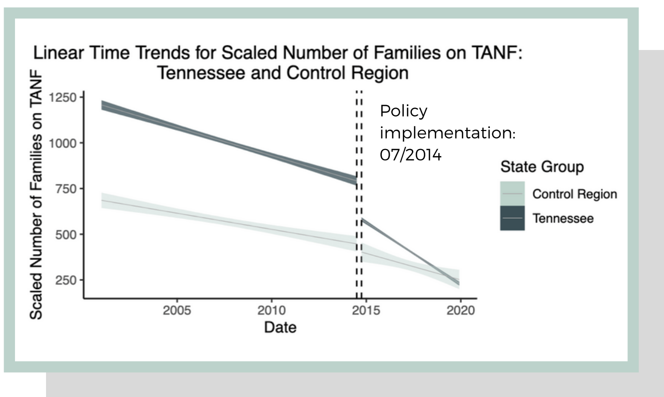
To check the robustness of our estimates, we accounted for a six month lag in behavior changes both on the all states model and the individual state models.

Further detail on methodology can be found in Appendices 1 and 2.

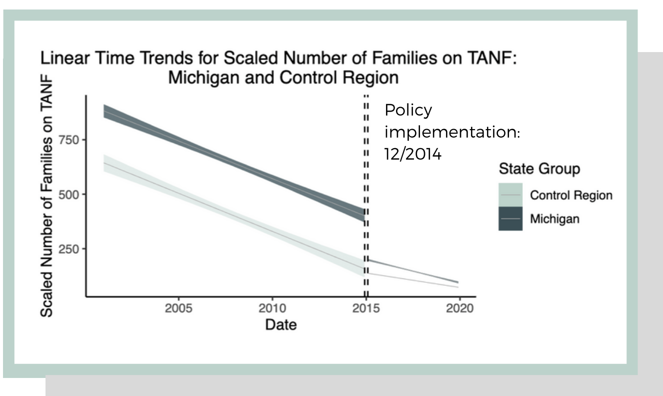
THE FINDINGS

SELECT STATE ANALYSIS

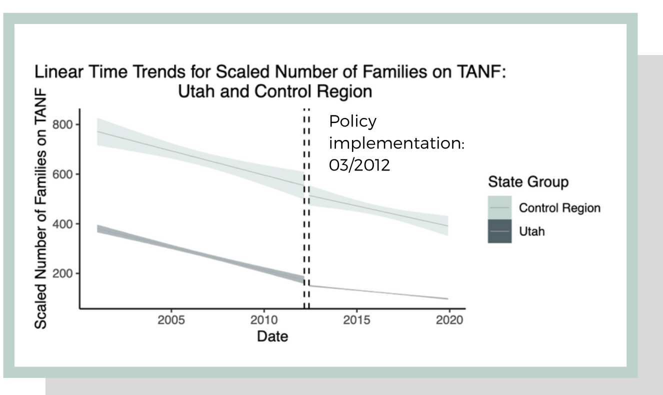
Difference-in-difference analyses of Tennessee and Michigan validate results from the national analysis, while results from Utah indicate no effect in that state (Appendices 3 and 4). Robustness checks produced similar estimates (Appendix 5).



Results indicate that after Tennessee's drug screening policy took effect in July 2014, there were **332 fewer families** enrolled in TANF per 100,000 state residents on average as compared to the control states, holding all else equal.



Results also indicate that after Michigan's drug screening policy took effect in December 2014, there were **150 fewer families** enrolled in TANF per 100,000 state residents on average as compared to the control states, holding all else equal.



Utah's March 2012 policy had **no significant effect** on TANF caseloads as compared to the control region. Thus, in some states, drug screening policies appear not to impact the number of families on the TANF caseload. Policymakers could look in to Utah's policy implementation for potential ways to keep drug screening policies while not reducing the number of families receiving TANF.

THE FINDINGS

NATIONAL ANALYSIS

DRUG SCREENING POLICIES CAUSED TANF CASELOADS TO **DROP** BY

2,085

FAMILIES IN ALL

OR, **65** FAMILIES PER 100,000 PEOPLE

An analysis of all states shows that TANF caseloads dropped by approximately 2,085 families, or 65 families per 100,000 people, as a result of implementing drug screening questionnaire policies. This result is statistically significant and shows that policies designed to have minimal negative impact on caseloads are in fact a strong factor of overall TANF caseload reductions in recent years. The estimate was the same when allowing for a six month lag after the policy implementation. (Appendix 5). The median number of families on TANF over the 2001-2019 period was 409 per 100,000, while the median number of total famlies was 14,945. Thus, the change due to screening policies is a significant decrease.

However, these data only account for the number of total people on TANF in a state at one time, not the number of new enrollments. When fewer people enroll in TANF over time, the number of cases decreases. But, the number of cases also decreases when people already on TANF leave at quicker rates than they had previously. Although we expect the number of TANF leavers to be similar in treated and control states, an important caveat to these findings is that total caseload data, not new enrollment data, were used.

THE CONCLUSION

KEY TAKEWAYS

DRUG SCREENING QUESTIONNAIRES **CAUSE FAMILIES TO AVOID APPLYING** FOR ASSISTANCE

POINT 1

Drug screens should not deter applications; applicants are still entitled to receive assistance if they test positive.

POINT 2

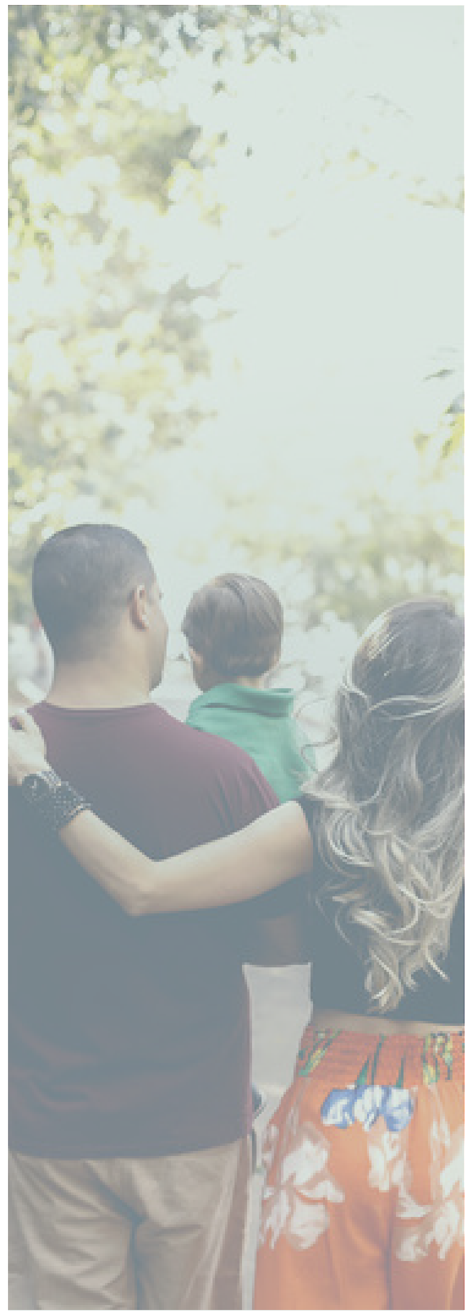
The number of families on TANF decreased across the nation after select states implemented drug screening policies.

POINT 3

Policymakers should consider if the benefits of drug screens outweigh their financial and caseload reduction repercussions.

POINT 4

Some states, like Utah, enacted screening policies without caseload reductions. These states could be valuable policy case studies for effective implementation of drug screenings.



APPENDIX 1:
Methodology & Exploratory Data Analysis

Our methodology consisted of five parts:

- **Exploratory data analysis** to understand the data;
- **Identification** of select policy states for further analysis;
- A fixed effects **difference-in-difference** analysis comparing three select states to two neighboring states each ("select states model");
- A fixed effects **difference-in-difference** analysis comparing all states with screening policies to those without ("all states model"); and
- A **robustness check** using alternate model specifications.

Exploratory data analysis is described in further detail below.

Exploratory Data Analysis

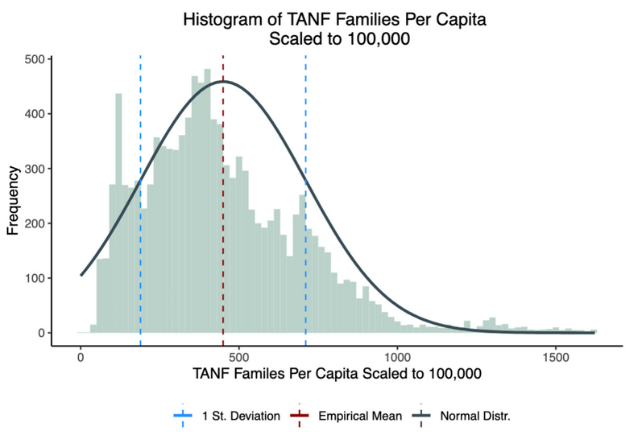
The compiled data consists of 11,628 observations and 11 variables; defintions of all variables are below. The year range within the data is 2001 to 2019, inclusive.

The number of households on TANF is scaled by state population and multiplied by a constant of 100,000 to yield the number of households on TANF per 100,000 state residents. This is the outcome variable of interest for the study.The number of families below the federal poverty line is similarly scaled by state population and multiplied by a constant of 100,000 to yield the number of households below the federal poverty line per 100,000 state residents.

There are 3 missing “NA” observations within the data. These pertain to Delaware with dates of 10/2015, 11/2015, and 12/2015 and are omitted from the data set. There are 5 observations where households on TANF are reported as 0: Idaho with dates of 11/2009 and 12/2009, and Missouri with dates of 1/2006, 2/2006, and 3/2006. These are identified as missing values in the raw data, and are omitted from the dataset. The observations pertaining to Washington, D.C. are identified as high-range outliers and are omitted. Subsequent to omissions, the data consists of 11,392 rows.

Variable Name	Variable Type	Definition
drug_law	binary (1=Yes/0=No)	Whether the drug policy is implemented
scale_tanf_fams	numeric	Number of TANF family applications (per 100,000 residents)
all_tanf	numeric	Number of TANF family applications (absolute)
scale_num_below_fpl	numeric	Number Below Federal Poverty Level (per 100,000 residents)
all_belowfpl	numeric	Number Below Federal Poverty Level (absolute)
unemp_rate	numeric	Unemployment Rate (in %)
state	character	U.S. State
year_cat	factor	Year

APPENDIX 1:
Methodology & Exploratory Data Analysis

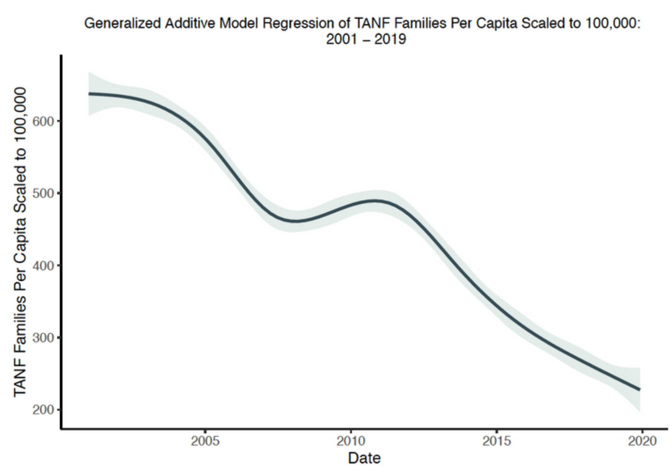


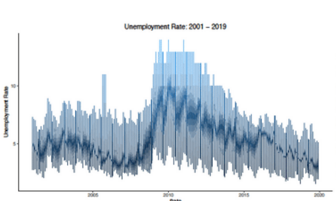
Distribution of Outcome and Control Variables of Interest:

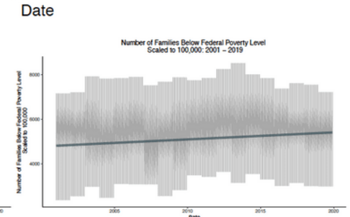
The distribution of the outcome variable of interest is presented to the left. The number of households on TANF by state is a discrete count variable, but approximate Normality is assumed under sufficient number of observations. The data is skewed right.

Two control variables are included within the data: the state unemployment rate percentage and the number of households below the federal poverty line per 100,000 state residents.

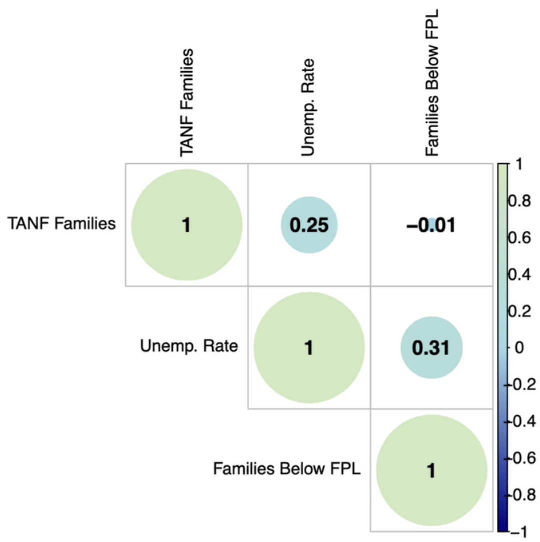
The unemployment rate across all states displays varying time trends, emphasized by a significant increase in unemployment rate under the Great Recession. The number of households below the federal poverty line per 100,000 state residents does not display a significant linear trend across the included states.







APPENDIX 1:
Methodology & Exploratory Data Analysis



The correlation heatmap indicates that no issues of multicollinearity are present within the data: a general criterion for the presence of multicollinearity is an absolute correlation coefficient greater than 0.70 among two or more feature variables. A modest positive correlation of 0.25 is observed between the number of households on TANF per 100,000 state residents and unemployment rate.

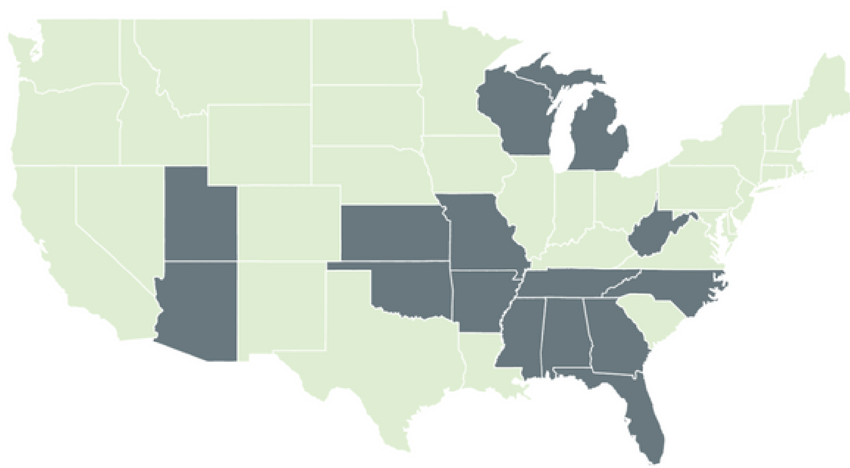
The correlation between the number of households on TANF per 100,000 state residents and the number of households below the federal poverty line per 100,000 state residents is -0.01, indicating that this variable likely does not contribute significant information for the prediction of the number of households on TANF per 100,000 state residents. This is most likely the result of the constant trend of the number of households below the federal poverty line over the considered time range.

APPENDIX 2:
IDENTIFICATION OF POLICY AND COMPARISON STATES

Identification of Policy States

15 states implemented drug screening policies for those applying to TANF. They primarily fall in three regions: the West (Utah, Arizona), the North (Michigan, Wisconsin), and the Southeast (Kansas, Oklahoma, Missouri, Arkansas, Alabama, Mississippi, Georgia, Florida, Tennessee, Kentucky, West Virginia).

State Drug Law Enaction



We selected Tennessee, Michigan, and Utah ("target states"), to perform a difference-in-difference analysis on because they were close to a number of states that did not implement screening policies. We then compared trends in the state unemployment rate, the scaled number of families below the federal poverty level, and scaled number of families on TANF before policy enactment between the target states and neighboring states. Non-screening states with similar pre-policy trends in TANF caseload, and similar unemployment and poverty trends from 2001-2019, were selected as comparison states. We selected two control states for each target state.

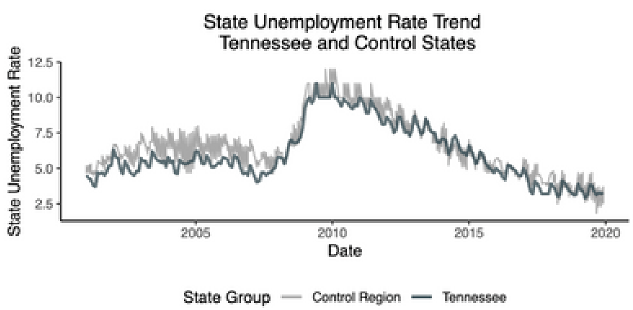
APPENDIX 2:

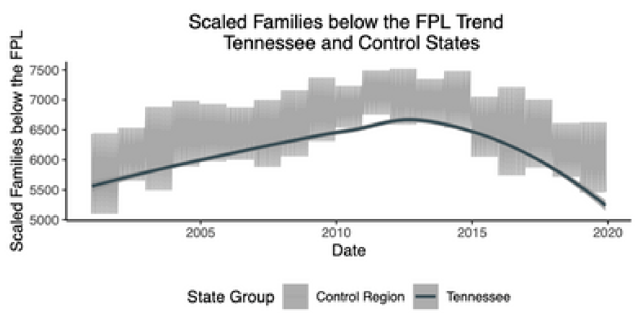
IDENTIFICATION OF POLICY AND COMPARISON STATES

Case: Tennessee

Policy Date: 07/2014

Control Region: Kentucky, South Carolina



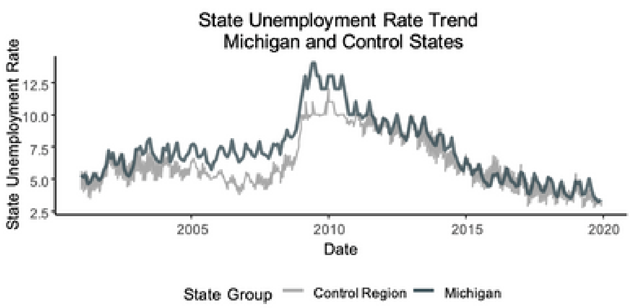


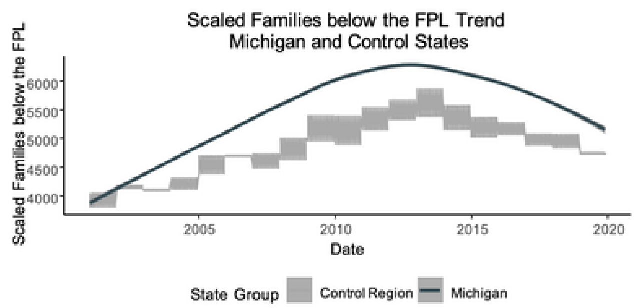
(Left): State unemployment rate levels and trends are comparable across Tennessee and control states of Kentucky and South Carolina. A naive Welch Two Sample t-Test has a test statistic of -3.5411 and an associated p-value of 0.0004378, indicating that there is sufficient evidence at any standard significance level to conclude that the means of the two samples are different. However, a Granger Causality Test has an F-test statistic of 5.0563 and an associated p-value of 0.007126, indicating that there is sufficient evidence at the $\alpha = 0.01$ level to conclude that the lagged unemployment rate within the control states provides information for the unemployment rate within Tennessee; that is, the time series are statistically comparable.

(Right): The number of households below the federal poverty line per 100,000 state residents is consistently slightly higher within control states than within Tennessee. A naive Welch Two Sample t-Test has a test statistic of -9.3551 and an associated p-value of < 2.2e–16, indicating that there is sufficient evidence at any standard significance level to conclude that the means of the two samples are different. A Granger Causality Test has an F-test statistic of 3.5399 and an associated p-value of 0.06121, indicating that there is evidence at the $\alpha = 0.10$ significance level to conclude that the lagged number of households below the federal poverty line within the control states provides information for the number of households below the federal poverty line within Tennessee. Although these results are not overly robust, the treatment and control parallels are considered sufficient for variable matching within the context of the study as the time trends and the overall counts are relatively comparable.

APPENDIX 2:
IDENTIFICATION OF POLICY AND COMPARISON STATES

Case: Michigan
Policy Date: 12/2014
Control Region: Indiana, Illinois





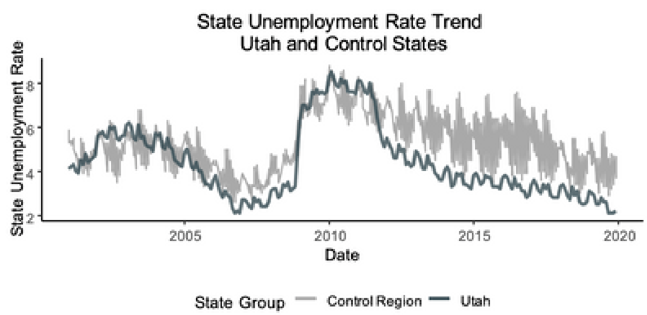
(Left): State unemployment rate levels and trends are relatively comparable across Michigan and control states of Indiana and Illinois, with deviation occurring within the years 2005 to 2010. A naive Welch Two Sample t-Test has a test statistic of 5.0126 and an associated p-value of 8.063−7, indicating that there is sufficient evidence at any standard significance level to conclude that the means of the two samples are different. A Granger Causality Test, however, has an F-test statistic of 13.615 and an associated p-value of 2.654−6, indicating that there is sufficient evidence at any standard significance level to conclude that the lagged unemployment rate within the control states provides information for the unemployment rate within Michigan.

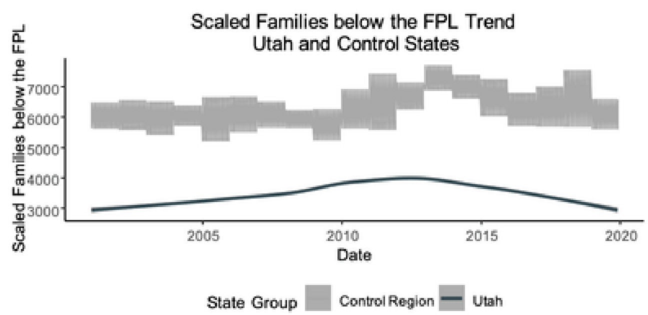
(Right): The number of households below the federal poverty line per 100,000 state residents is consistently slightly higher within Michigan than within control states. A naive Welch Two Sample t-Test has a test statistic of 11.64 and an associated p-value of < 2.2−16, indicating that there is sufficient evidence at any standard level of significance to conclude that the means of the two samples are different. A Granger Causality Test has an F-test statistic of 3.0203 and an associated p-value of 0.0836, indicating that there is evidence at the $\alpha = 0.10$ significance level to conclude that the lagged number of households below the federal poverty line within the control states provides information for the number of households below the federal poverty line within Michigan. Although these results are not overly robust, the treatment and control parallels are considered sufficient for variable matching within the context of the study as the time trends and the overall counts are relatively comparable.

APPENDIX 2:

IDENTIFICATION OF POLICY AND COMPARISON STATES

<u>Case: Utah</u>
Policy Date: 03/2012
Control Region: New Mexico, Montana





(Left): State unemployment rate levels and trends are relatively comparable across Utah and control states of New Mexico and Montana, with increased variability occurring after the year 2010. A naive Welch Two Sample t-Test has a test statistic of -7.1137 and an associated p-value of 5.454−12, indicating that there is sufficient evidence at any standard level of significance to conclude that the means of the two samples are different. A Granger Causality Test has an F-test statistic of 2.9399 and an associated p-value of 0.08812, indicating that there is sufficient evidence at the $\alpha = 0.10$ significance level to conclude that the lagged unemployment rate within the control states provides information for the unemployment rate within Utah.

(Right): The number of households below the federal poverty line per 100,000 state residents is consistently higher within the control states than within Utah. A naive Welch Two Sample t-Test has a test statistic of -69.819 and an associated p-value of < 2.2−16, indicating that there is sufficient evidence to conclude at any standard significance level that the means of the two samples are different. A Granger Causality Test has an F-test statistic of 0.8603 and an associated p-value of 0.3547, indicating that there is not sufficient evidence at any standard significance level to conclude that the lagged number of households below the federal poverty line within the control states provides information for the number of households below the federal poverty line within Utah. While the general treatment and control parallels are again considered sufficient for variable matching within the context of the study, more advanced matching methods may certainly be explored in further analysis.

APPENDIX 3:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: ALL STATES

All States

The difference-in-difference model for all states was specified as

$$TANF = \beta_0 + \beta_1 DrugLaw_{it} + \beta_2 Poverty_{it} + \beta_3 Unemployment_{it} + \beta_3 State_i + \beta_4 Year_t + \varepsilon_{it}$$

where *TANF* is the number of families per 100,000 people receiving TANF assistance per state-year, *DrugLaw* is an indicator variable that takes a value of 1 if a drug screening policy existed in a given state (i) in a given year (t) as 0 otherwise, *Poverty* is the per capita number of people below the federal poverty level per state *i* and year *t*, *Unemployment* is the unemployment rate in state *i* and year *t*, *State* is a fixed effect indicator for each of the 50 states, and *Year* is a fixed effect indicator for each year in the 2001-2019 time range. The coefficient on *DrugLaw* is the coefficient of interest.

We use the above model to conduct a fixed effects linear regression with state-level clustered standard errors (SEs). For the initial model with all states ("complete" model), the number of TANF applications per 100,000 state residents was regressed on the treatment indicator, i.e., whether a state has implemented the drug screening policy or not. Control variables that measure the state-specific unemployment rate and the number of state residents below the federal poverty level per 100,000 state residents are also included. This basic regression analysis excludes a pre/post-indicator (and treatment-time interaction term) as the policy was implemented at different points in time for the states included. The basic regression includes year- and state-fixed effects to account for time invariant state specific effects and external events or shocks, as well as state-level clustered standard errors. Estimates are seen in Table 1 below; state- and year- fixed effect coefficients omitted for brevity. Estimates for total number of families are seen in Table 1(A).

Table 1. Complete model

	Complete
Drug Law (1=Yes/0=No)	-65.2272^{***}
	(4.2369)
Number Below Federal Poverty Level (per 100,000)	-0.0058^{*}
	(0.0035)
Unemployment Rate (in %)	15.2914^{***}
	(1.0950)

Table 1 (A). Complete model with absolute numbers

Drug Law (1=Yes/0=No)	$-2{,}085.0080^{***}$
	(607.4156)
Number Below Federal Poverty Level (absolute)	-0.0120^{***}
	(0.0033)
Unemployment Rate (in %)	$2{,}949.1330^{***}$
	(152.7838)

When regressing the complete model, the effect of the drug policy on TANF applications is negative and statistically significant, with a p-value below 0.01. Screening policy implementation is associated with on average 65 TANF household applications less per 100,000 state residents, or 2,085 families in all. The estimates of both control variables are also statistically significant.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

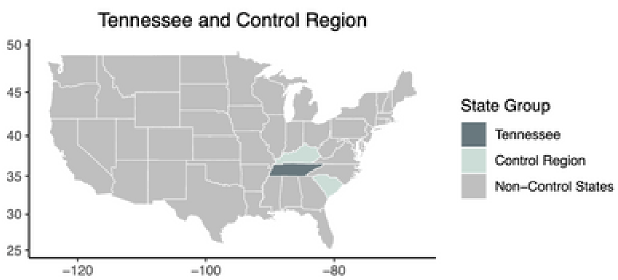
Select States

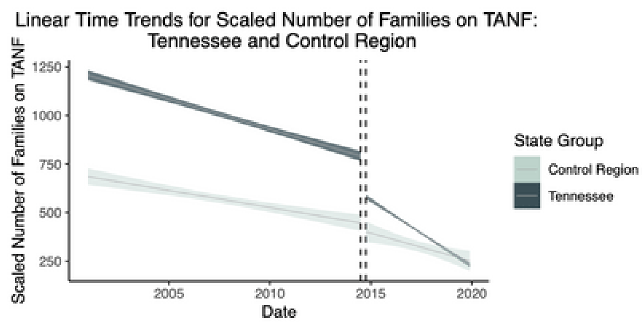
Following the all states complete model, we look at our selectively chosen target states and their control states. The control states provide a counterfactual for what the target states' caseloads may have been if screening policies had not been implemented. Because control states have similar pre-intervention trends, are in the same region, and have similar trends in poverty and unemployment, we believe our difference-in-difference analysis is a valid causal estimate.

The model is the same as in the all states section above, using the target and control states as opposed to all states. The figures on the following pages show each difference-in-difference plot; we are interested in the gap on either side of the policy implementation year. We also show regression tables, the residuals versus the fitted value, and the Normal Q-Q plot for each state.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

Case: Tennessee
Policy Date: 07/2014
Control Region: Kentucky, South Carolina





Tennessee and the selected control states maintain parallel trends pertaining to the number of families on TANF prior to the date of policy enactment in Tennessee. Subsequent to the date of policy enactment, the trend pertaining to families on TANF in Tennessee exhibits an increased negative slope while a stable trend is maintained within the selected control states.

Table 2 (A). Policy Regression for Tennessee.

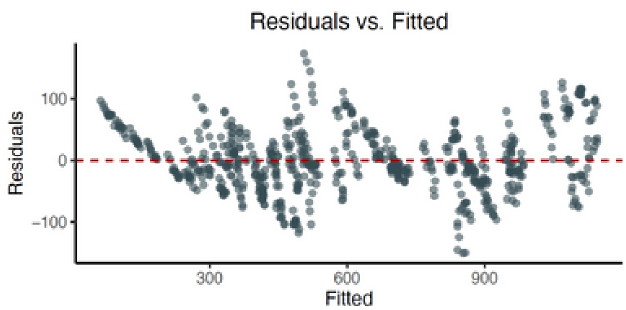
	Complete
Drug Law (1 = Yes/0 = No)	-332.1027***
	(9.2148)
Number Below Federal Poverty Level (per 100,000)	0.0117
	(0.0162)
Unemployment Rate (in %)	-15.0213***
	(3.0013)
N	684
R^2	0.9708
Adjusted R^2	0.9698

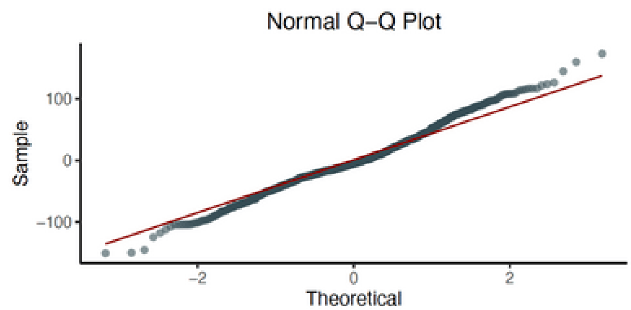
Standard errors given in parentheses.
***p<0.01; **p<0.05; *p<0.1.
TANF rate scaled to 100,000 inhabitants.

The coefficient in Table 2(A) of -332.1 families per 100,000 people is significant at the 1% level.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

Case: Tennessee
Policy Date: 07/2014
Control Region: Kentucky, South Carolina



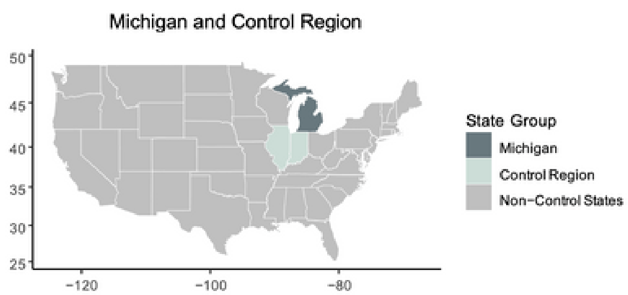


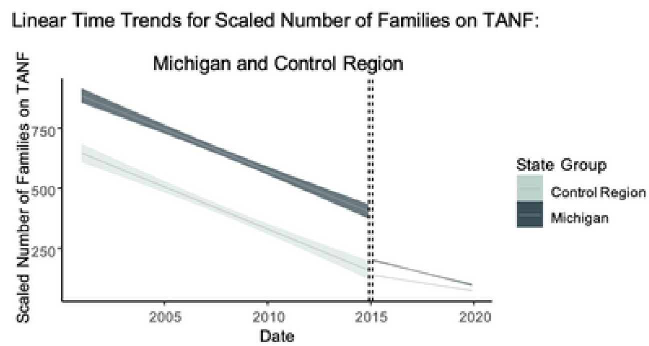
(Left): To determine if the model exhibits constant variability of residuals, a residuals versus fitted plot is generated. In the plot, the fitted values of the model are plotted on the x axis, and the residuals of the model are plotted on the y axis. The fitted values generally form a horizontal band around the residual = 0 line, indicating overall constant variability of residuals. However, non-random trends are visible within the plot that indicate a linear model may not be fully appropriate for the data, and the residuals are large in absolute magnitude.

(Right): To determine if the model has nearly normal residuals, a normal probability plot is generated. In the plot, the data are plotted by residuals generated from a theoretical normal distribution. The plot for the data follows a general linear trend, except in the tail areas of the distribution. The first points at the beginning of the range of the data fall below the line, while the last points at the end of the range of the data fall above the line. This indicates that the data exhibits thin tails. Further analysis may augment the data to explain more variability of the response variable.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

Case: Michigan
Policy Date: 12/2014
Control Region: Indiana, Illinois





Michigan and the selected control states maintain parallel trends pertaining to the number of families on TANF prior to the date of policy enactment in Michigan. Subsequent to the date of policy enactment, the trend pertaining to families on TANF in Michigan exhibits a significant downwards shift with a similar slope. A relatively stable trend is maintained within the selected control states.

Table 2 (B). Policy Regression for Michigan.

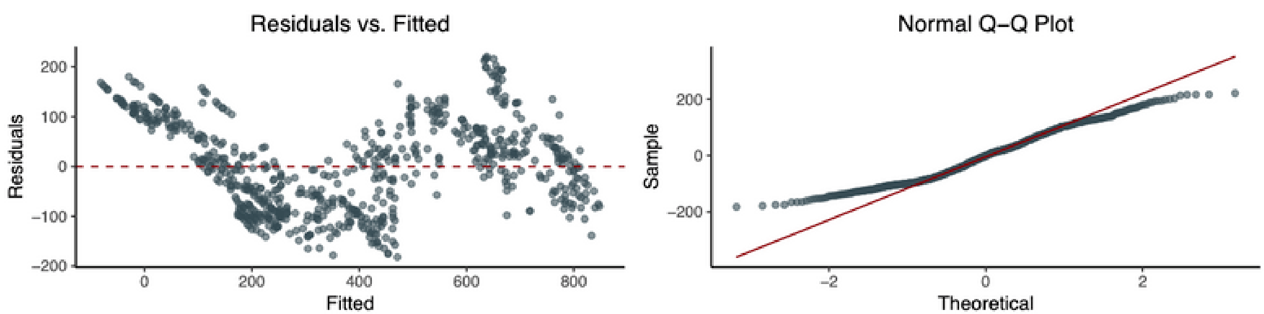
	Complete
Drug Law (1 = Yes/0 = No)	-150.8445***
	(17.4633)
Number Below Federal Poverty Level (per 100,000)	-0.0689***
	(0.0239)
Unemployment Rate (in %)	42.9279***
	(5.1240)
N	684
R^2	0.8834
Adjusted R^2	0.8793

Standard errors given in parentheses.
***p<0.01; **p<0.05; *p<0.1.
TANF rate scaled to 100,000 inhabitants.

The coefficient in Table 2(B) of -150.8 families per 100,000 people is significant at the 1% level.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

<u>*Case: Michigan*</u>
Policy Date: 12/2014
Control Region: Indiana, Illinois

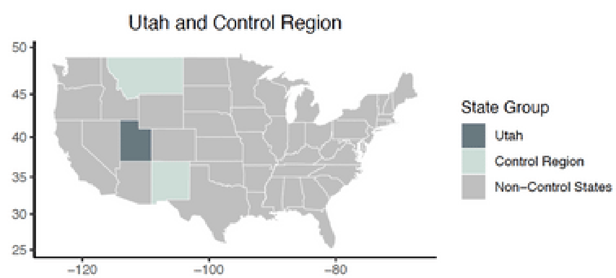


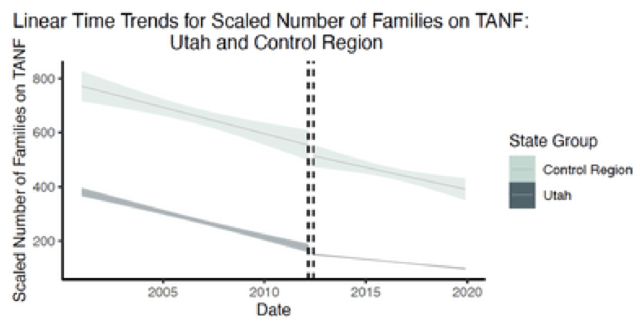
(Left): To determine if the model exhibits constant variability of residuals, a residuals versus fitted plot is generated. In the plot, the fitted values of the model are plotted on the x axis, and the residuals of the model are plotted on the y axis. The fitted values generally form a horizontal band around the residual = 0 line in the second half of the range of fitted values, but fail to do so in the first half of the range of fitted values. Furthermore, a non-random oscillating trend is visible within the plot that indicates a linear model may not be fully appropriate for the data, and the residuals are large in absolute magnitude.

(Right): To determine if the model has nearly normal residuals, a normal probability plot is generated. In the plot, the data are plotted by residuals generated from a theoretical normal distribution. The plot for the data fails to follow a linear trend in the tail areas of the distribution, significantly deviating from the theoretical line. The first points at the beginning of the range of the data fall above the line, while the last points at the end of the range of the data fall below the line. This indicates that the data exhibits thin tails. Further analysis may augment the data to explain more variability of the response variable.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

<u>Case: Utah</u>
Policy Date: 03/2012
Control Region: Montana, New Mexico





Utah and the selected control states maintain parallel trends pertaining to the number of families on TANF prior to the date of policy enactment in Utah. Subsequent to the date of policy enactment, the respective trends pertaining to families on TANF in Utah and the control states remain fairly consistent; the absolute magnitude of the regression slope pertaining to Utah appears to decrease slightly.

Table 2 (C). Policy Regression for Utah.

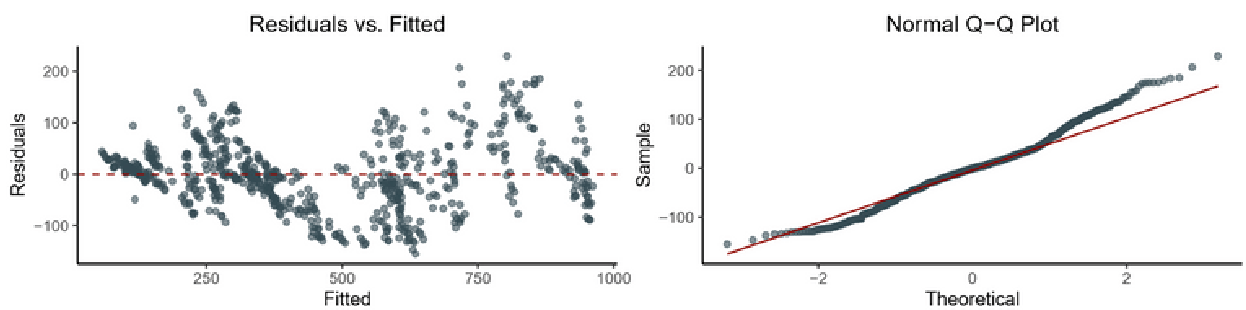
	Complete
Drug Law (1 = Yes/0 = No)	17.5039
	(13.6967)
Number Below Federal Poverty Level (per 100,000)	0.0079
	(0.0123)
Unemployment Rate (in %)	-23.0794***
	(3.9391)
N	684
R^2	0.9375
Adjusted R^2	0.9353

Standard errors given in parentheses.
***p<0.01; **p<0.05; *p<0.1.
TANF rate scaled to 100,000 inhabitants.

The coefficient in Table 2(C) of 17.5 families per 100,000 people is not significant.

APPENDIX 4:
DIFFERENCE-IN-DIFFERENCE ANALYSIS: SELECT STATES

Case: Utah
Policy Date: 03/2012
Control Region: Montana, New Mexico



(Left): To determine if the model exhibits constant variability of residuals, a residuals versus fitted plot is generated. In the plot, the fitted values of the model are plotted on the x axis, and the residuals of the model are plotted on the y axis. The fitted values generally form a horizontal band around the residual = 0 line, indicating overall constant variability of residuals. Many residuals are large in absolute magnitude.

(Right): To determine if the model has nearly normal residuals, a normal probability plot is generated. In the plot, the data are plotted by residuals generated from a theoretical normal distribution. The plot for the data follows a general linear trend, with deviation occurring at the end of the range of the data as the points curve upwards. This may indicate right skew within the data. Further analysis may augment the data to explain more variability of the response variable.

APPENDIX 5:
ROBUSTNESS CHECKS

All States

To check the robustness of our nationwide regression model, we perform a lagged fixed effects regression with state-level clustered standard errors. In doing so, we drop the first six months after the policy implementation to account for gradual, as opposed to instantaneous, behavior changes. In the all states model in Table 3 below, state- and year- fixed effect coefficients are omitted for brevity.

Table 3. All States Model Regression with Six Month Lag

	Complete	With Lag
Drug Law (1=Yes/0=No)	-65.2272***	-65.1837***
	(4.2369)	(4.3856)
Number Below Federal Poverty Level (per 100,000)	-0.0058*	-0.0048
	(0.0035)	(0.0035)
Unemployment Rate (in %)	15.2914***	15.3475***
	(1.0950)	(1.1005)

The effect of drug screening policies with lag is very similar in magnitude to the effect we find in the complete model. Thus, we can conclude that the treatment effect is indeed robust and stable in the long-term, not only the result of an immediate short-term response effect to the policy implementation.

APPENDIX 5:
ROBUSTNESS CHECKS

Select States

We perform the same analysis with lag on each of the selected states and their control states, displayed in Tables 4(A) to 4(C) below.

Table 4 (A). Tennessee Model Regression with Robustness Checks.

	Complete	With Lag
Drug Law (1 = Yes/0 = No)	-332.1027***	-355.9017***
	(9.2148)	(9.4152)
Number Below Federal Poverty Level (per 100,000)	0.0117	-0.0076
	(0.0162)	(0.0159)
Unemployment Rate (in %)	-15.0213***	-15.1476***
	(3.0013)	(2.9533)
N	684	666
R^2	0.9708	0.9731
Adjusted R^2	0.9698	0.9721

Standard errors given in parentheses.

***p<0.01; **p<0.05; *p<0.1.

TANF rate scaled to 100,000 inhabitants.

The effect of drug screening policies with lag in Tennessee is very similar in magnitude to the effect we find in the complete model for Tennessee. Accounting for lag actually increases the effect's magnitude, likely because the TANF agencies may need time to create drug screening infrastructure and TANF applicants may need time to hear about the policy and adjust their actions accordingly.

APPENDIX 5:
ROBUSTNESS CHECKS

Select States

Table 4 (B). Michigan Model Regression with Robustness Checks.

	Complete	With Lag
Drug Law (1 = Yes/0 = No)	-150.8445***	-157.7844***
	(17.4633)	(18.0334)
Number Below Federal Poverty Level (per 100,000)	-0.0689***	-0.0660***
	(0.0239)	(0.0240)
Unemployment Rate (in %)	42.9279***	44.1282***
	(5.1240)	(5.1898)
N	684	666
R^2	0.8834	0.8853
Adjusted R^2	0.8793	0.8812

Standard errors given in parentheses.
***p<0.01; **p<0.05; *p<0.1.
TANF rate scaled to 100,000 inhabitants.

The effect of drug screening policies with lag in Michigan is also very similar in magnitude to the effect we find in the complete model for Michigan. Accounting for lag increases the magnitude of the effect; however, the increase in coefficient magnitude from -150.8 to -157.8 represents only a 5% change, which is minimal.

APPENDIX 5:
ROBUSTNESS CHECKS

Select States

Table 4 (C). Utah Model Regression with Robustness Checks.

	Complete	With Lag
Drug Law (1 = Yes/0 = No)	17.5039	25.5028*
	(13.6967)	(13.9666)
Number Below Federal Poverty Level (per 100,000)	0.0079	0.0151
	(0.0123)	(0.0123)
Unemployment Rate (in %)	-23.0794***	-23.5592***
	(3.9391)	(3.9464)
N	684	666
R^2	0.9375	0.9381
Adjusted R^2	0.9353	0.9359

Standard errors given in parentheses.

***p<0.01; **p<0.05; *p<0.1.

TANF rate scaled to 100,000 inhabitants.

The effect of drug screening policies with lag in Utah becomes significant only at the 10% level. In the complete model for Utah, the effect of drug screening laws was positive and insignificant. Accounting for lag shows drug screening laws increase the TANF caseload by approximately 25 families. This result is surprising and may warrant further analysis.

One explanation for the caseload increase after the policy implementation may be that one of the control states, Montana, is further away than other control states used in the analyses. Although Montana and Utah had similar trends for unemployment rate and number of people below the federal poverty level, they were not perfect matches. However, given the available data, no better matches existed between Utah and neighboring states. There may be unobserved demographic differences between the two states that make them fairly different. If so, Montana would not be the best control state for Utah, and may have different underlying trends in TANF caseloads. Nevertheless, further investigation is warranted in future studies.