Three Essays in Applied Microeconomics about Child Welfare and Child Care

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Caitlyn Keo

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Dr. Elizabeth Davis

April 2025

Acknowledgements

I would like to thank my committee members, Elizabeth (Liz) Davis, Aaron Sojourner, Joseph Ritter, and Kristine Piescher for guidance, support, and encouragement. Your careful, insightful feedback and suggestions allowed me to improve this dissertation. I am particularly grateful to my advisors, Liz and Aaron, for being great mentors. Your willingness to discuss economics and my dissertation research has been invaluable during my PhD years.

Thank you to Kristine West and Caroline Krafft for the early encouragement and for showing me the amazing things that economists do. I am grateful to my undergraduate experience doing research as a student at St. Kate's and I was glad to continue to have you as leading role models during graduate school.

Finally, I would like to acknowledge the support of my network, especially, my parents, siblings, and extended family. You have set the grounds for me to explore education at the highest level. Thank you for your enormous support and love.

Dedication

To my parents, Christina and Pisith Keo, for their love, support, and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The three independent studies in this dissertation empirically contribute to the economics of child welfare and the child care markets. The first two papers shed light on how state reforms and federal accountability can affect outcomes for children involved in child welfare. The third paper looks at the effect of regulations and the economy on child care markets and the persistence of family child care providers remaining in business. Below, I introduce each study.

Chapter 2 studies child protection investigations and how a reform that allows for alternative responses to reports of child maltreatment affects the likelihood of future maltreatment for children in the child welfare system. Traditionally, all screened-in child protection services (CPS) reports are investigated to formally determine whether the alleged child maltreatment is confirmed. This conventional investigative response (IR) to all reports has been met with concerns that investigations can be adversarial for children and families, particularly for those at the intersection of poverty and suspected child maltreatment. Therefore, many U.S. states have adopted differential response (DR) systems. In a DR, "dual track," system, the agency reserves the IR track for reports in which the child appears to be at imminent risk of subsequent maltreatment. For lower risk reports, the agency uses an alternative response (AR) track. While services to families are typically provided in the AR track, there is concern that, without an investigation, children could be more at risk for future maltreatment. With child maltreatment data from 15.7 million individual reports in the National Child Abuse and Neglect Data System from 2004 to 2019 and the staggered adoption of DR reforms over time, a difference-in-differences design assesses how DR affects subsequent maltreatment for reported children. Estimates imply that DR policies on average do not increase the likelihood of a re-report within three years. These findings suggest that on average children are as safe or safer in states with DR.

In Chapter 3, I look at whether federal reviews and accountability causes states to improve their child welfare permanency outcomes. Permanency — legal membership in a safe and stable family — is a vital objective of the child welfare system. Legislative measures such as the Adoption and Safe Families Act of 1997 clearly emphasize a preference for family reunification for children in foster care. However, despite legislation, specific initiatives to improve reunification and reduce re-entry were not made until federal oversight was introduced in the early 2000s through the Child and Family Services Review (CFSR). Following the CFSR, the U.S. Children's Bureau keeps states accountable to federal standards by advising and approving Program Improvement Plans (PIP). Given the variation in the timing of CFSR assessments, this study employs a two-way fixed effects differences-in-differences model to evaluate the impact of the two rounds of the CFSR on two key permanency outcomes: the 12-month foster care re-entry rate and the 12-month family reunification rate. The findings indicate that, all else being equal, states identifying the need for improvement in re-entries experienced substantial reductions in their foster care re-entry rate after the first round of the CFSR, without simultaneously causing significant decreases in the 12-month reunification rate.

Finally, my co-author and I in Chapter 4 look at the decline family child care in Minnesota. The number of family child care (FCC) providers has been declining nationally for over a decade, raising concerns about availability and access to child care for families at the local level. In many states, new licensing regulations have been enacted for FCC providers over the last decade and

have been viewed as burdensome and intrusive to some in the FCC community. We focus on two major changes in regulations in Minnesota: new health and safety regulations which required additional trainings in 2014, and expanded background check regulations in 2017. Using interrupted time series analysis, we find that the rate of providers exiting the child care market (i.e., closing their businesses) was not higher after these regulations were enacted. To examine factors related to FCC closings, we estimate monthly and annual survival analysis models predicting business closures. Our monthly results suggest that the 2014 health and safety regulations and the 2017 background checks were significantly associated with a decreased likelihood of exit rather than an increase. Several economic and business factors are also associated with exits.

Chapter 2. Reforming Child Protection Services: The Effect of Differential Response on Subsequent Maltreatment for Reported Children

2.1 Introduction

Traditionally, child protection services (CPS) agencies respond to all screened-in reports of child maltreatment with a formal investigation. However, many families, agencies, and practitioners expressed concerns that the traditional investigative response is inflexible and adversarial. In some cases, an alternative response may be more appropriate. Over time, many state agencies did away with the investigation-fits-all approach to handling CPS reports. Child and family needs are diverse and complex. Yet, balancing child safety and family preservation can be challenging.

Most U.S. states have reformed CPS with a differential response (DR) system, which separates child maltreatment reports into two (or more) response tracks. These tracks are generally known as an alternative response (AR) track and an investigative response (IR) track. In a DR system, criteria are established in which the traditional IR track is usually restricted to reports for children who have been severely maltreated or who, based on the initial report, appear to be of moderate to high risk of further abuse or neglect. The non-investigative AR track is established for reports that are considered of low to moderate risk. The AR track provides services to the family and by design will not result in a determination of maltreatment or with the child in foster care. Reforms for dual tracks are intended to allow CPS to engage with families in a way that moves away from a solely punitive and adversarial stance towards a more collaborative and strengths-based approach.

Critics of CPS reforms often argue that DR does not provide the same level of child safety as traditional investigative, one-track systems. To study DR policies across states over a relatively long period using microdata, I specifically look at states that adopt DR between 2004 and 2019. I leverage large amounts of data on child maltreatment reports from the National Child Abuse and Neglect Data System (NCANDS) for 2004 to 2019. To proxy subsequent maltreatment for reported children, I use re-reporting as my primary outcome of interest. A re-report is a subsequent report of maltreatment that meets statutory requirements for follow-up assessment or investigation. My key research question directly asks, how do differential response policies affect re-reporting? My empirical approach aims to understand the impact of DR policies on child safety and whether children in states with dual tracks are more likely to be further maltreated or to be re-involved with CPS compared to children in states without DR.

To date, there is limited long-term research that assesses the outcomes of DR compared to traditional child welfare models. A limited number of studies have used aggregated state-level data to explore the effect of DR on safety outcomes and found reductions in child welfare system involvement as measured by re-reporting rates and substantiation rates (Fluke et al., 2019; Johnson-Motoyama et al., 2023; Piper, 2017). For instance, Piper (2017) utilized NCANDS data and found that re-reporting rates for AR cases were lower than IR cases when fewer than a third of the accepted reports for CPS intervention were assigned to the AR pathway. Another study, which analyzed data from six states between 2004 and 2013, discovered that increased utilization of the DR pathway led to a decrease in the overall rate of substantiated re-reports (Fluke et al., 2019).

The main empirical analyses look at changes in subsequent maltreatment in states that adopt DR policies versus those that don't. I leverage the staggered adoption of DR reforms over time to estimate average treatment effects on the treated (ATT) and use difference-in-difference (DD) strategies to understand how DR policies affect re-reporting. I find that DR policies decrease the likelihood of a re-report within 3 years. Disaggregated results show that this treatment effect is statistically significant across predicted risk terciles, and children on average experience a 2-3 percentage point lower risk of a re-report within 3 years in states with DR. When the data is disaggregated by race and ethnicity, the decrease in the likelihood of a re-report is largest for Black children. When disaggregated by the reason for the report, I find that the effect of DR policies on re-reporting is not statistically significant among reports of neglect, physical abuse, sexual abuse, or emotional or psychological abuse.

I show the child and case characteristics associated with assignment to alternative response using a linear probability model (LPM) to predict placement on the AR track. Several characteristics of reports affect the probability of AR track assignment versus IR track assignment, including who makes the report. Reports made by education, child care, substitute care personnel, and medical and mental health personnel are more likely to receive AR track placement relative to reports made by social services personnel. Prior victims, children reported due to physical abuse or sexual abuse, and younger children (<5 years old) are less likely to receive an alternative response. All else equal, race and ethnicity also play a role in who receives an alternative response.

Contributing to the existing literature in public economics and child welfare, this paper provides evidence as to how DR reforms in child welfare affect safety outcomes for reported children. To date, there is limited long-term research that assesses the safety outcomes of DR systems compared to traditional child welfare systems. This lack of evidence can make it challenging to evaluate the effectiveness of the DR approach fully. This study contributes to the literature by using administrative CPS data at the child level. Earlier studies used aggregated state-level data on re-reporting rates (e.g., Johnson-Motoyama et al., 2023; Piper, 2017). The state-level findings are informative, but microdata and further analysis are of importance to understand the nuanced effects of DR policies on children of various characteristics. I also provide evidence across numerous states over a relatively long period, which improves upon state and county-specific program evaluations which are often completed in shorter durations or during pilot programs (e.g., Fuller et al., 2013; Murphy et al., 2013).

2.2 Background on Child Welfare Systems

At intake, CPS agencies[1] are responsible for all allegations of child maltreatment (abuse and neglect).[2] These allegations are called referrals, and they are often made by phone through a 24-hour hotline. If a referral meets the state and local criteria for CPS involvement it is "screened-in" and is then referred to as a report. Referrals that are screened out receive no further CPS action. This paper focuses only on reports.

[1] The child welfare system encompasses a wide range of responsibilities entrusted to local authorities whose primary goal is to ensure the safety, permanency, and well-being of children. Responding to allegations of child maltreatment is one of the foremost tasks in child welfare. In response to maltreatment allegations, the relevant jurisdiction– referring to the power, right, or authority to interpret and apply the law– differs across states. There are 39 states with state-administered CPS systems, while 9 states have county-administered systems, and 2 states have hybrid systems. The District of Columbia is also considered a state-administered system (Weigensberg et al., 2022). Throughout this study, I use the terms 'jurisdiction' and 'agency' interchangeably to discuss CPS.

[2] Federal law (42 U.S.C.A. § 5106g) broadly defines child maltreatment to include abuse and neglect. Abuse refers to harmful actions inflicted upon a child, such as physical or sexual abuse, while neglect pertains to the failure to provide a child with safe and consistent care. Following the broad federal definition of child maltreatment, narrower more concrete definitions of maltreatment which elicit CPS involvement differ at the state level and change over time (Child Welfare Information Gateway, 2014).

2.2.1 Stages of Child Welfare in a Traditional Investigative System

In a traditional CPS system, *all* reports of maltreatment are investigated (see the left panel A of Figure 2.1). Given federal law, a CPS investigation of a maltreatment report either substantiates the maltreatment allegation or determines that no maltreatment occurred, making the report unsubstantiated. Unsubstantiated reports require no further action on behalf of CPS.[3,4] A substantiated report results in two common forms of CPS intervention. One option is removing the child(ren) from their home and placing them in foster care, often while CPS provides services to the family (parents or guardians). The second option is allowing the child(ren) to remain in their home while providing services to the family. In either scenario, the overall objective is to ensure child safety and strengthen families.

2.2.2 Reforming Child Welfare to Allow for Differential Response

Over the last few decades, CPS in most U.S. states transitioned to a DR system, also known as multiple response and family assessment.[5] In 2019, 30 states had DR systems, with 25 operating at the state level and 5 at the county level (Weigensberg et al., 2022) (see Appendix Figure C1). Additionally, by 2019, 9 states and the District of Columbia had previously implemented DR,

---

[3] A report is substantiated when the CPS investigation determines that there is confirmation of reasonable cause to believe that at least one child on the report has been abused or neglected. A report is unsubstantiated when an investigation determines that no child maltreatment occurred. This determination might occur when there is insufficient evidence under state law or agency policy to support the alternative conclusion (Child Welfare Information Gateway, 2003).

[4] An important note regarding many CPS agencies is that if a referral is screened out, or a report is unsubstantiated, future re-referrals or re-reports of suspected child maltreatment cannot reference the details of the previous allegations that did not become substantiated reports. In other words, previous unsubstantiated allegations (including screened-out referrals) hold no weight in the decision of substantiation for the current report.

[5] Note, that state legislation proceeds reforms and implementation of DR. Additionally, DR is not a uniformly defined reform. Differences in DR implementation between jurisdictions can be found in Appendix D.

sometimes as a pilot program, and later discontinued DR in favor of a traditional investigation-only system.[6]

CPS agencies that are reformed for DR separate child maltreatment reports into either the traditional investigative response (IR) or alternative response (AR). Panel B of Figure 2.1 shows a standard DR example for an agency that has an AR track in addition to its IR track.[7] A DR system establishes criteria such that the IR track is usually restricted to moderate- to high-risk reports for children who have been severely maltreated or who appear to be at risk of neglect or abuse. The non-investigative AR track works with lower-risk families to address the underlying causes of their maltreatment report with referrals to services. The non-investigative AR track does not require a formal finding for the alleged maltreatment. Hence, by policy design, reports on the AR track cannot be substantiated.

Reforms focus on the underlying reason for the suspected maltreatment allegation and largely aim to address the issues that are persistent for families with low incomes. Specifically, active reform uses the AR track to focus on meeting the needs of families through services in a way that promotes family well-being and self-sufficiency (Piper et al., 2019). The AR track can be used to deliver a variety of services including, but not limited to, parenting, mental health, and substance abuse interventions.

[6] Reasons for discontinuation generally include funding or resource limitations (e.g., the Alaska DR program was discontinued in 2009 due to the loss of legislative funding allocated for the program), inconsistent implementation, leadership changes, and safety concerns following high-profile child fatalities (Alaska Citizen Review Panel, 2015; Child Welfare Information Gateway, 2014; David Scharfenberg, 2015).

[7] The initial referral screening stage is omitted. This starts with screened-in reports, as in the data.

2.3 Data and Descriptive Analysis

2.3.1 Data on Child Maltreatment Reports

In this study, I use the restricted National Child Abuse and Neglect Data System (NCANDS) Child File datasets for 2004 through 2019.[8] This administrative data is voluntarily collected from state agencies and disseminated by the National Data Archive on Child Abuse and Neglect (NDACAN) which is located at Cornell University (Children's Bureau, 2019). The data consist of the demographics of reported children (e.g., age at the time of report, race, Hispanic ethnicity), the reasons for the maltreatment report, the reporters' characteristics (e.g., social services, medical or mental health personnel, family and friends), whether the child has previously been a victim, and an indicator of placement on an IR or AR track.

In the NCANDS data, observations are at the report-child level and contain detailed information specific to each report. Any given alleged maltreatment report (identified by a report ID) may correspond to one child or several children (identified uniquely by a child ID). For most states, multiple reports for the same child are linkable within state across years.[9] This is essential for assessing subsequent maltreatment re-reporting, which is the primary outcome of interest and is discussed in detail in section 3.4.

To construct the analytic sample of maltreatment reports, I first make systematic state-year exclusions and then make child exclusions across states (see Appendix A). Specifically, in state

---

[8] These Child Files are based on submission years which are fiscal years. Throughout my analysis, I acknowledge the year (and month) of the report based on the calendar year (month) the report was received rather than the fiscal year data was submitted. This is derived from the 'report date' which is the month, day, and year that the responsible agency was notified of the suspected child maltreatment. Thus, I keep all reports that were made in the calendar years 2004 to 2019. This drops the reports that were submitted to the Children's Bureau in 2004 or later but were received by CPS as early as 2001.

[9] However, there are several breaks in child ID linkage that are documented by NDACAN. In Appendix A, I discuss how I deal with breaks in child IDs in constructing my sample.

years with linkable child IDs, I restrict the sample to children ages 0-17 with reports where ‘neglect,’ ‘physical abuse,’ ‘emotional or psychological abuse,’ or ‘sexual abuse’ were at least one of the (up to four) documented reasons for the report.[10] Among states with DR, I limit to those that begin DR between 2004 and 2019 (dropping always-treated states), that do not discontinue the use of DR before 2019, and that only make the track determination after a report is screened in.[11] After the child exclusions and the state-year exclusions, the sample is constructed with over 19 million report-child observations from 21 states (including the District of Columbia). The state-year panel is unbalanced due to differences in years in which each state reports to the national system.

2.3.2 Covariates for Local Demographic and Economic Factors

Socioeconomic factors such as income, education, employment, and social support are associated with child maltreatment. Relationships between low household income and child abuse and neglect are consistently observed (Berger et al., 2017; Berger & Waldfogel, 2004; Paxson & Waldfogel, 1999, 2003). Socioeconomic determinants of child maltreatment further include parental work status (Paxson & Waldfogel, 1999) and unemployment (Brown & De Cao, 2018). Accordingly, empirical research on child maltreatment across disciplines has explored the effects of policies that address the socioeconomic factors related to maltreatment – predominately maltreatment characterized as physical abuse and neglect (e.g., Berger et al., 2017; Ginther & Johnson-Motoyama, 2017; Raissian & Bullinger, 2017; Rittenhouse, 2023).

[10] Most states officially recognize four major types of child maltreatment: neglect, physical abuse, sexual abuse, and emotional or psychological abuse. Appendix Figure C1 shows the number of annual reports-child observations by reasons for report.
[11] NCANDS data do not provide the appropriate universe to analyze states that screen out to AR. In those states, it would be more appropriate to use data on referrals to CPS, rather than data on CPS reports.

I draw on two additional datasets for 2004 to 2019 to account for county- and state-level differences in socioeconomic factors. First, county-level covariates for demographics come from one-year samples from the American Community Survey (ACS) (Ruggles et al., 2019). ACS also provides the state-year log of the child population, the share of the population that is non-Hispanic black, non-Hispanic Asian, non-Hispanic other race, and Hispanic any race, and the share of children living below the poverty threshold. Some counties in the ACS are masked, in which case I compute the characteristics of the residual parts of the state averaged together. Then, the University of Kentucky National Welfare Data (2023) account for state-year socioeconomic factors that may affect maltreatment: log of population, log of average personal income, unemployment rate, log of the state minimum wage; log of Supplemental Nutrition Assistance Program (SNAP) benefits for a family of three, and log of Temporary Assistance for Needy Families (TANF) benefits for a family of three.

2.3.3 Differential Response Treatment Variables

Both state-level treatment and child-level DR treatment are observed and leveraged in this study. Child-level treatment indicates that a reported child is placed on the AR track, implying a state-year with DR in place. I create a binary treatment variable for report placement on the 'AR track' that is equal to one if the NCANDS report disposition is 'alternative response disposition-victim' or 'alternative response disposition-non-victim.'[12] Then, placement on the 'AR track' equals zero

[12] A report disposition is the final determination that is concluded from the CPS response to a report of child maltreatment. Dispositions categories include substantiated; indicated or reason to suspect; alternative response disposition-victim; alternative response disposition-not a victim; unsubstantiated; unsubstantiated due to intentionally false reporting; closed-no finding; other; and, unknown or missing. If there are multiple children on the report, the

if there was one of the other report dispositions since all other dispositions would have been concluded on the IR track.

State-level treatment in this study indicates that a state is reformed for differential response either at the state level or the county level. Following the methods of Johnson-Motoyama et al. (2022), if any report in that state year received a placement on the 'AR track,' the state is coded as utilizing DR that year. This state-level variable is binary and equals one in the post-DR period. The treatment variable is equal to zero for both the pre-DR period for states that eventually reform or in the states that never reform for DR since no child has received an AR placement that year. Appendix Figure C2 shows DR treatment status by year for states in my sample. Appendix Tables C1 and C2 show descriptive analyses for key variables with consideration to state and child-level treatment status.[13]

2.3.4 Outcome Measures

I focus on re-reporting as my primary outcome of interest. A re-report commands another CPS response. In the literature, it is common to proxy for subsequent maltreatment by identifying when a reported child is later re-reported within a given interval of time.[14] Various intervals of time (e.g.,

---

report disposition takes on the most severe of the dispositions. For example, a report that is "substantiated" means that at least one child on the report was found to be a victim.

[13] Specifically, within states that adopt DR policies between 2004 and 2019, Appendix Table C2.1 shows the characteristics of report-child observations separately before and after the policy. Then after the policy, it further disaggregates the IR and the AR track observations. Then, Appendix Table C2.2 shows the characteristics of report-child observations that are investigated in state-years with no DR reform (conventional investigative CPS structure) and state-years that are post-DR reform (CPS with DR reform).

[14] There are different ways to proxy for subsequent child maltreatment. Using administrative CPS data, the commonly used measures include re-referrals (i.e., a referral or report, followed by another referral), re-reporting (i.e., a report, followed by another report), or recurrence (i.e., revictimized as indicated by a substantiated report, followed by another substantiated report). Since NCANDS does not include screened-out referrals, I cannot look at re-referrals for this study. Furthermore, since reports cannot be substantiated on the AR track, I do not look at recurrence.

6 months, 12 months, 18 months, 60 months) are used in impact evaluations and empirical studies (e.g., Antle et al., 2009; Baron et al., 2023; Cuomo & Carrión, 2011; Putnam-Hornstein et al., 2021; Shipe et al., 2022).

I measure a child maltreatment re-report with an indicator for when the current report is followed by a subsequent report within 6 months, 12 months, and 3 years. The intervals are calculated based on the initial report date. To account for potential censoring for intervals of longer duration, the 6-month and 12-month re-report indicators are created for reports from 2004 to 2018, and the 3-year re-report indicator is calculated for reports from 2004 to 2016. Models using the 3-year duration window are shown in my preferred specifications because they use the most information for each case. I use the shorter windows of time in my robustness checks to test for sensitivity to different definitions of the outcome.

2.3.5 Descriptive Statistics: Trends in the Outcomes

The proportion of reported children who are re-involved with CPS via a subsequent screened-in report increases as the duration window increases (6 months, 12 months, and 3 years). Figure 2.2 shows the trends in the outcomes by year. There is a slight incline over time for the proportion of reported children who get re-reported across all three intervals. Looking at the most recent years of the data, in 2018 about 22 percent of reported children were re-reported within 6 months and 31 percent were re-reported within 12 months. In 2016 nearly 50 percent of all reported children were re-reported within 3 years.

2.3.6 Descriptive Analysis: Predictors of an Alternative Response

The DR approach generally allows CPS agencies to respond differently to maltreatment reports based on factors such as the type of maltreatment and the severity of the report. To examine what predicts an alternative response, I use a linear probability model (LPM) where the dependent variable is the binary child-level treatment, AR track equals one, in periods with DR. The coefficients on predictors of an AR track placement from my preferred specification that includes child and report characteristics (coefficients shown in Figure 2.3), local demographic and economic factors, and state and year fixed effects (coefficients not shown) describe the predictive relationship. This preferred specification is reported in column 3 of Appendix Table C2.3. Reports made by education, child care, substitute care personnel (5% level), and medical and mental health personnel (10% level) are more likely to receive an AR relative to reports made by social services personnel. Prior victims, children reported due to physical abuse or sexual abuse, and younger children (<5 years old) are less likely to receive an AR. Moreover, Black, and Asian or Pacific Islander children are less likely to receive an AR, relative to White children, as are Hispanic children relative to non-Hispanic children. These results accord with research from Texas which showed that assignment to the AR track was largely driven by child race, although I cannot also study whether it is driven by neglect and poverty as found by Choi et al. (2021).

2.4 Empirical Strategy

2.4.1 Difference-in-Difference: The Effect of Differential Response Policies on Re-reporting

How do differential response policies affect re-reporting? To quantify the impact of DR policy (treatment), I leverage the variation in timing, and I measure the average treatment effect on the treated (ATT) using differences-in-differences (DD) methods. In a staggered treatment timing design, problems arise when the magnitude of the treatment effect is correlated with the timing of treatment (Baker et al., 2022; Borusyak et al., 2023; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021). I bypass the concerns of treatment effect heterogeneity by estimating a treatment effect for each event and aggregating these treatment effects to obtain the weighted average treatment effect. Specifically, I use the estimator proposed by Callaway and Sant'Anna (2021) since I have a staggered, binary treatment.

The general form of the DD model with variation in treatment timing is expressed as:

$$Y_i = \alpha_1 DR_{s(i)t(i)} + \alpha_{s(i)} + \alpha_{t(i)} + \varepsilon_i \tag{1}$$

where Y_i denotes the re-reporting indicator outcome for report-child i. $DR_{s(i)t(i)}$ is a binary variable that indicates whether report-child i is in a state s that had the DR reform in year t. Year t refers to the year the initial report was screened in. State fixed-effects, $\alpha_{s(i)}$, account for time-invariant differences between states, such as differences in funding, child welfare practices, or child maltreatment definition. The year fixed effects, $\alpha_{t(i)}$, control for the common temporal trend in specific calendar year *t,* in which the report was screened-in by an agency. Finally, ε_i is the error term with standard properties.

In my analytic sample, I have 21 states in total: 13 control states and 8 treated states. Having relatively few treated clusters may raise potential problems that invalidate the conventional cluster robust inference. Primarily, the concern is that when there are too few clusters, the *t*-statistics tend to over-reject (Cameron et al., 2008; Cameron & Miller, 2015; Mackinnon & Webb, 2017, 2018). Therefore, I model equation (1) using wild cluster bootstrapped t-statistics to create confidence intervals (Roodman et al., 2019). The *p*-values from the wild bootstrapped *t*-statistics are clustered by state with 999 repetitions. While the coefficients and the *t*-statistics are the same as they would be in a cluster robust model, the bootstrapping of the *t*-statistics changes the *p*-values (which reveal the significance) since it uses the bootstrapped distribution rather than the normal distribution. Compared to the asymptotic normal distribution, the bootstrapped distributions lead to better approximations (Cameron et al., 2008; MacKinnon et al., 2023).

2.4.2 Heterogeneous Effects by Predicted Risk Tercile

To further understand the range of treatment effects that DR policies have on children, I assess whether there are heterogeneous effects of DR policies across predicted risk terciles. To do this, I created a re-report propensity score in periods without DR and assigned it to observations in periods without DR *and* periods with DR to hold constant the "risk" of a maltreatment re-report. In other words, in periods with no DR, meaning pre-DR and never DR state-years, I assess the predictors of a subsequent maltreatment report using a LPM where the dependent variable is a binary re-reporting outcome and I include controls for child characteristics and report characteristics, and local demographic and economic characteristics (see Appendix Table C2.4).[15] The models that pool across state and year in periods with no DR reforms, meaning they do not include state and year fixed effects, are used to create the re-report propensity scores. Then, in periods with DR, this creates a reasonable counterfactual risk measure for what re-reporting would be in the absence of the DR reform. Using the propensity score, I partition report-child observations in each state into risk terciles for analysis: low-risk, medium-risk, and high-risk.

2.5 Results

2.5.1 Difference-in-Difference: The Effect of Differential Response Policies on Re-reporting

First, I show the effect of DR reforms on re-reporting outcomes by leveraging the staggered treatment timing. To assess the pre-trends, Figure 2.4 shows the corresponding event study with 5 periods before and 5 periods after the DR reform. From the event study, there may be a slight pattern to the pre-trends that produce significant effect estimates several years before the policy (e.g., t= -5) but the pattern is also consistent with evidence of no rise in the likelihood of a re-report following the adoption of DR policies in year 0. As shown in column 1 of Table 2.1, I then estimate that DR policies on average reduce the likelihood of a re-report within 3 years using a Callaway and Sant'Anna DD approach with the never treated state-year report-child observations as the comparison group.

[15] In the absence of DR reforms, the positive and significant predictors of re-reporting and re-reporting with substantiation include being a prior victim, neglect, and psychological and emotional abuse. Physical abuse is associated with a higher likelihood of being re-reported, but not necessarily re-reported with substantiation. Older children (>5 years old) are less likely to be re-reported and less likely to be re-reported with substantiation. Relative to White children, Black, American Indian, Alaskan Native, and Asian or Pacific Islander children are less likely to be re-reported, and Black and Asian or Pacific Islander children are less likely to be re-reported with substantiation. Children on substantiated reports are less likely to be re-reported but are marginally more likely to be re-reported with substantiation in 3 years.

Since DR policies are intended to allow CPS to work with lower risk families by providing assessments and referrals to services via the AR track, we might expect to find the effect of DR policies on re-reporting to be concentrated for reports of low to medium risk. Columns 2-4 show the heterogeneity in the treatment effects across predicted risk tercile. The ATT coefficients show that DR decreases the likelihood of a re-report by 2-3 percentage points. which is a relatively large effect compared to the average likelihood of a re-report across predicted risk terciles: low risk tercile [Pr(re-report within 3 years)=0.35]; medium risk tercile [Pr(re-report within 3 years)=0.47]; and high risk tercile [Pr(re-report within 3 years)=0.60]. While the treatment effect is significant across predicted risk terciles, it is best (precisely) estimated for children who would be considered medium risk in the absence of the policy. Interestingly, the medium risk tercile is also the tercile in which the AR track is used at a higher rate in state-years with DR policies [Pr(AR|DR, middle risk tercile) =0.13]. Intuitively, the data also shows that the AR track is used less often for reported children in the high risk tercile in state-years with DR policies [Pr(AR|DR, high risk tercile) =0.09].

2.5.2 Heterogeneous Effects Across Baseline Characteristics: Race and Ethnicity

Child welfare policies and practices merit attention for many reasons including because they can reduce or contribute to racially disparate processes (Semanchin Jones, 2015). DR has been recognized as a strategy that could potentially reduce disproportionality and disparity in foster care placements (Lemon et al., 2008; Martin & Connelly, 2015). One evaluation of a pilot DR program in 10 Ohio counties in 2006 showed a decrease in foster care placements across all races, including a reduction in the number of African American children (Kaplan & Rohm, 2010). However, a decline in the number of African American foster care placements does not necessarily lead to reductions in disparities or disproportionalities. To definitively determine the overall efficacy of DR at reducing racial disproportionality and disparities in the child welfare system a stronger evidence base is needed (Fluke et al., 2010).

Table 2.2 shows the treatment effects disaggregated by the race (columns 1 through 5) and ethnicity (columns 6 and 7) of the reported child. DR policies are associated with a lower likelihood of a re-report within 3 years for Black children, which is marginally significant at the 10 percent level. However, on average, White children, American Indian and Alaskan Native children, Asian or Pacific Islander children, and children of two or more races are unaffected by the policy as measured by a re-report within 3 years. Columns 6 and 7 do not unveil a disparate impact of the policy when disaggregating by Hispanic ethnicity.

2.5.3 Heterogeneous Effects Across Baseline Characteristics: Reason for the Report

A CPS maltreatment report can list up to four reasons for the allegation. The treatment effects are disaggregated by the reason for the maltreatment report in Table 2.3. While statistically null, DR policies on average are associated with a lower likelihood of a re-report within 3 years for children who are reported to be neglected. There is no effect of DR on children who are reported for sexual abuse. Note that less than one percent of children who are reported for sexual abuse are placed on the AR track in state years with DR. Hence, a zero effect is consistent in that these children should not be affected by the policy since the AR track is not used for reports of their demeanor. Alternatively, on average, children who are reported to be emotionally or psychologically abused have an increased likelihood of a re-report within 3 years in states with DR. Children who are reported with emotional or psychological abuse as one of the reasons for the report are placed on the AR track at relatively higher rates than other reasons [Pr(AR| DR, emotional or psychological abuse) =0.16].

2.5.4 Robustness Checks

Robustness to the Choice of Estimator

Appendix Table B2.1 shows sensitivity to the choice of estimator. The ATT coefficients from my main specification are the weighted average (by group size) of all available group-time average treatment effects estimates from the unconditional Callaway and Sant'Anna model with reported children from never treated states as the comparison group. This main specification is also shown in Appendix Table B2.1 panel A. In panel B, I change the comparison group from reported children in the *never treated* states to reported children in *not-yet treated* state-years (i.e., the pre-DR state years and the never DR states). The coefficients are very similar, but the estimated ATT for the high risk tercile is no longer marginally significant.

In panel C, I include covariates in the model for child, report, and local characteristics in my Callaway and Sant'Anna model with never treated as the comparison group. For the models with covariates, I augment equation (1) with additional term $\alpha_2 X_i$, where X_i is a vector of control variables. I use the estimation proposed by Callaway and Sant'Anna (2021) that is based on doubly robust methods (Sant'Anna & Zhao, 2020). The benefit of the doubly robust DD estimator is that the modeling conditions by Sant'Anna and Zhao (2020) are less stringent than the outcome regression and the inverse probability weighting procedures (Callaway & Sant'Anna, 2021). The ATT estimate in column 1 of panel C is similar in magnitude to the ATT in column 1 of panel A.

However, the coefficients showing the heterogeneous effects across risk terciles are no longer significant and are now positive rather than negative.

The conventional two-way fixed effects (TWFE) model is a variance-weighted average of multiple 2 by 2 DD estimates which compares different treatment and control groups (Goodman-Bacon, 2021). TWFE regressions may not always estimate a convex combination of treatment effects making it challenging to give the treatment estimates a causal interpretation. The TWFE estimator could produce coefficients that have a different (i.e., opposite) sign than every unit's treatment effect (De Chaisemartin & D'haultfoeuille, 2023).

I show the ATT from conventional TWFE models, first without covariates and then with covariates, in panels D and E respectively. I continue to use wild cluster bootstrapped *t*-statistics to create confidence intervals in the TWFE specifications (Roodman et al., 2019). The *p*-values from the wild bootstrapped *t*-statistics are clustered by state with 1000 repetitions (Roodman, 2022). The TWFE models are consistent with heterogeneous treatment effects rather than homogeneous treatment effects, and the direction of the treatment effect is the opposite of the Callaway and Sant'Anna models. In panels D and E, the treatment (DR) has a positive 2-percentage point effect on the likelihood of a re-report within 3 years. This ATT is marginally significant at the 10 percent level when there are covariates included in the models as seen in panel E.

This study provides an application in which the conventional TWFE methods and heterogeneity robust DID estimators offer differ conclusions about treatment effects. As mentioned by De Chaisemartin and D'haultfoeuille (2023), general conclusions surrounding the underlying reasons for the estimated differences in treatment effects require a broad set of applications. Regardless of the explanation, the estimators proposed by Callaway and Sant'Anna (2021) are preferred for the binary and staggered treatment designs since they are heterogeneity-robust estimators.

Robustness to the Outcome Definition: Length of Time for a Re-report

I check the sensitivity to the definition of my primary outcome variable: a re-report within 3 years. Specifically, I assess re-reporting within shorter duration windows. I add measures for when the current report is followed by a subsequent report within 6 months and 12 months. The intervals are calculated based on report dates. To account for potential censoring for intervals of longer duration, the 6-month and 12-month re-report indicators are created for reports from 2004 to 2018. Looking across column 1, the size of the treatment effect increases between as the duration window increases. When changing the duration window to a re-report within 6 months or 12 months, I continue to find that, on average, children are as safe or safer in states with DR policies. Appendix Table B2.2 shows that on average, the existence of DR is associated with a 1-2 percentage point reduction in the likelihood of a re-report within 6 or 12 months, but the difference is not statistically significant.

Robustness to the Outcome Definition: Re-report with Substantiation

Substantiation, meaning confirmed maltreatment, is only possible following an investigation. Another important indication of subsequent maltreatment is a re-report that gets substantiated (i.e., confirmed subsequent maltreatment). I measure re-reporting with substantiation within 6 months, 12 months, and 3 years. These outcomes also account for potential censoring: the 6-month and 12-

month re-report with substantiation indicators are created for reports from 2004 to 2018, and the 3-year re-report with substantiation indicator is calculated for reports from 2004 to 2016. Column 1 in Appendix Table B2.3 shows that, on average, there is no effect of DR on re-reporting with substantiation within any of the three bounded time intervals. When I disaggregate the data by risk tercile, I continue to find that DR policy does not on average affect the likelihood of a re-report within 6 or 12 months. However, within 3 years, I find that among low risk reported children, DR policies lead to a 1 percentage point reduction in the likelihood of a re-report within 3 years which is significant at the 5 percent level. These results reaffirm the null hypothesis that children in states with DR policies are as safe or safer as children in states with traditional investigative CPS protocol.

Robustness to the Sample

Finally, I test the sensitivity of the analysis to various changes in the sample in Appendix Table B2.4. Panel A shows my preferred specification. In panel B, I drop the two states (New York and Colorado) in my sample that have specific county-level DR systems rather than state-level DR systems. Compared to the main specification in panel A, I find that the treatment effect grows slightly for the medium risk tercile from a negative 3-percentage point treatment effect to a negative 4-percentage point treatment effect, but the statistical significance drops from a 5 percent significance level to a 10 percent significance level. Then, in panel C, I drop the reported children in small, masked counties.[16] Since there are noticeable administrative differences in CPS between large and small counties, this exclusion to the sample makes the CPS jurisdictions relatively more

[16] In NCANDS counties are masked if there are fewer than 1000 reports in the fiscal year.

similar. Compared to the main specification in panel A, I find the effect size remains the same across specifications. The statistical significance increases for medium-risk children (from a 5 percent significance level to a 1 percent significance level) and low-risk children (a 10 percent significance level to a 5 percent significance level).

2.6 Summary and Concluding Remarks

2.6.1 Summary and Implications

The investigative response to *all* reports has been met with pertinent concerns that CPS involvement, in the form of an investigation, can be adversarial for children and their caretakers, particularly because it may place undue scrutiny on families at the intersection of poverty and suspected child maltreatment. These concerns surrounding investigative protocol are a part of a larger movement in child welfare which emphasizes family-focused, strengths-based approaches to child welfare (National Quality Improvement Center, 2011). However, other concerns emerge as a result of CPS reforms for dual tracks. In this study, I test whether DR reforms keep children as safe as traditional investigative, one-track systems, by measuring differences in the likelihood of a re-report within 3 years under both policy regimes: traditional, investigative regimes and DR regimes.

While previous studies predominately look at child safety using state-level outcomes such as re-reporting rates, I look at outcomes for individual children, and unlike previous studies, my models leverage microdata to understand how reported children are affected by the adoption of this popular reform. In my main empirical analyses, I find that on average DR policies decrease the likelihood of a re-report within 3 years. Disaggregated results show that negative effects exist across risk terciles, and are driven by children who would be considered medium risk in the absence of the policy based on propensity score weighting from the periods without DR. When disaggregated by race, the decrease the likelihood of a re-report is driven by Black children. When disaggregated by the reason for the report, I find null effects of DR policies on re-reporting. The results that are significant in the aggregate analysis, as well as the disaggregated analysis, are nonnegligible relative to the average likelihood of re-reporting at each of these intervals.

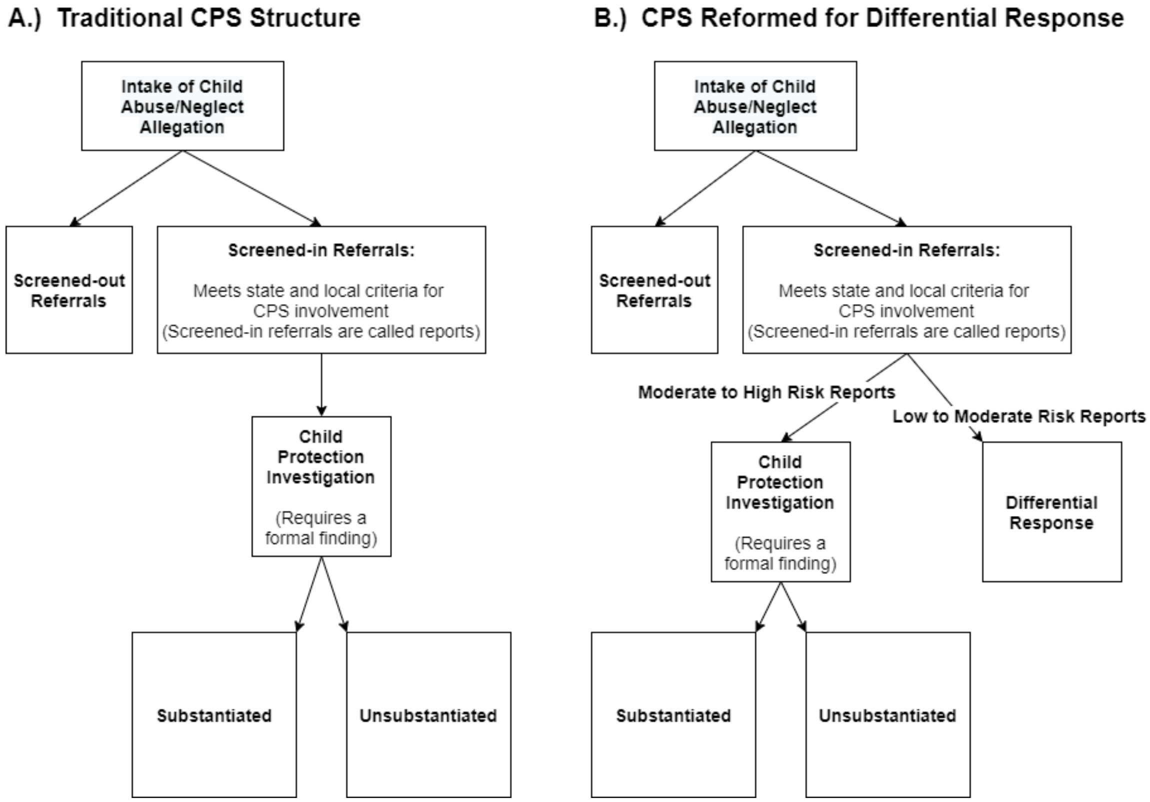
My results are consistent with the null hypothesis that reported children in states with DR policies are as safe or safer on average. Assuming that a re-report is a reliable proxy, DR does not increase the likelihood of subsequent maltreatment.

2.6.2 Limitations

To study the effect of DR on child safety in a feasible context, I made numerous specification decisions. Thus, the conclusions of this study should be understood relative to its limitations. The first limitation is related to the data. Ideally, we want to know if the child suffers maltreatment after the initial report, yet we only know if the child is re-reported. It is known that administrative CPS data underestimates child maltreatment (e.g., Hussey et al., 2006; Swahn et al., 2006), thus my outcomes (re-reporting) will miss any subsequent maltreatment that went unreported. Missed ('undercounted' or 'underreported') incidences of chronic or acute maltreatment would exist for children who did not get referred to CPS, and for children whose incidences were referred to CPS but were screened out (i.e., they did not formally become reports). Data on referrals to CPS (i.e., allegations at intake) would marginally improve upon the data-bound limitation by including the screened-out referrals. Yet, identifying incidences that are not referred to CPS would be more difficult.

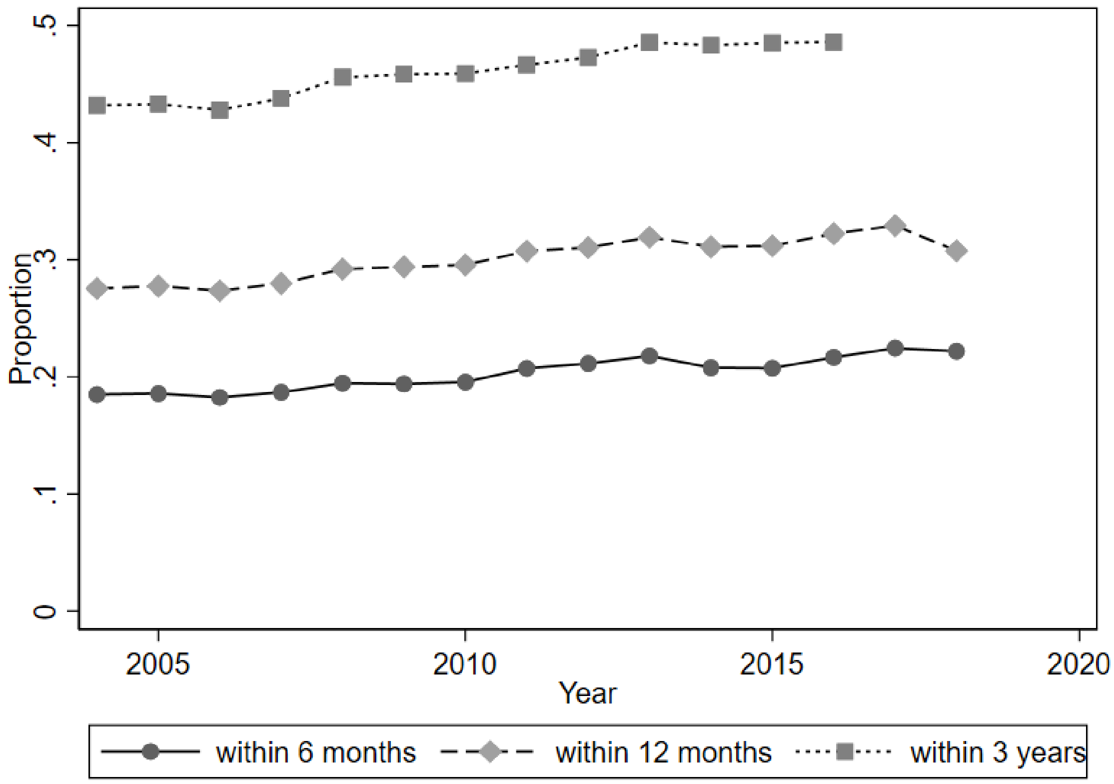
The second limitation is related to the exclusions made in the sample creation. In this study I limited my analyses to DR systems that only make the determination for AR *after a report has been screened in*. Thus, I cannot comment on DR systems that screen out to the AR track. Here, data on referrals to CPS (with response indicators for screened out AR) would drastically improve upon this limitation and make it possible to incorporate both determination timings. Furthermore, to have a consistent sample throughout this study, I excluded states with both implemented and discontinued DR over this period. I also excluded states that were always treated (e.g., Minnesota).

Figure 2.1. Types of Child Protection Services (CPS) structure



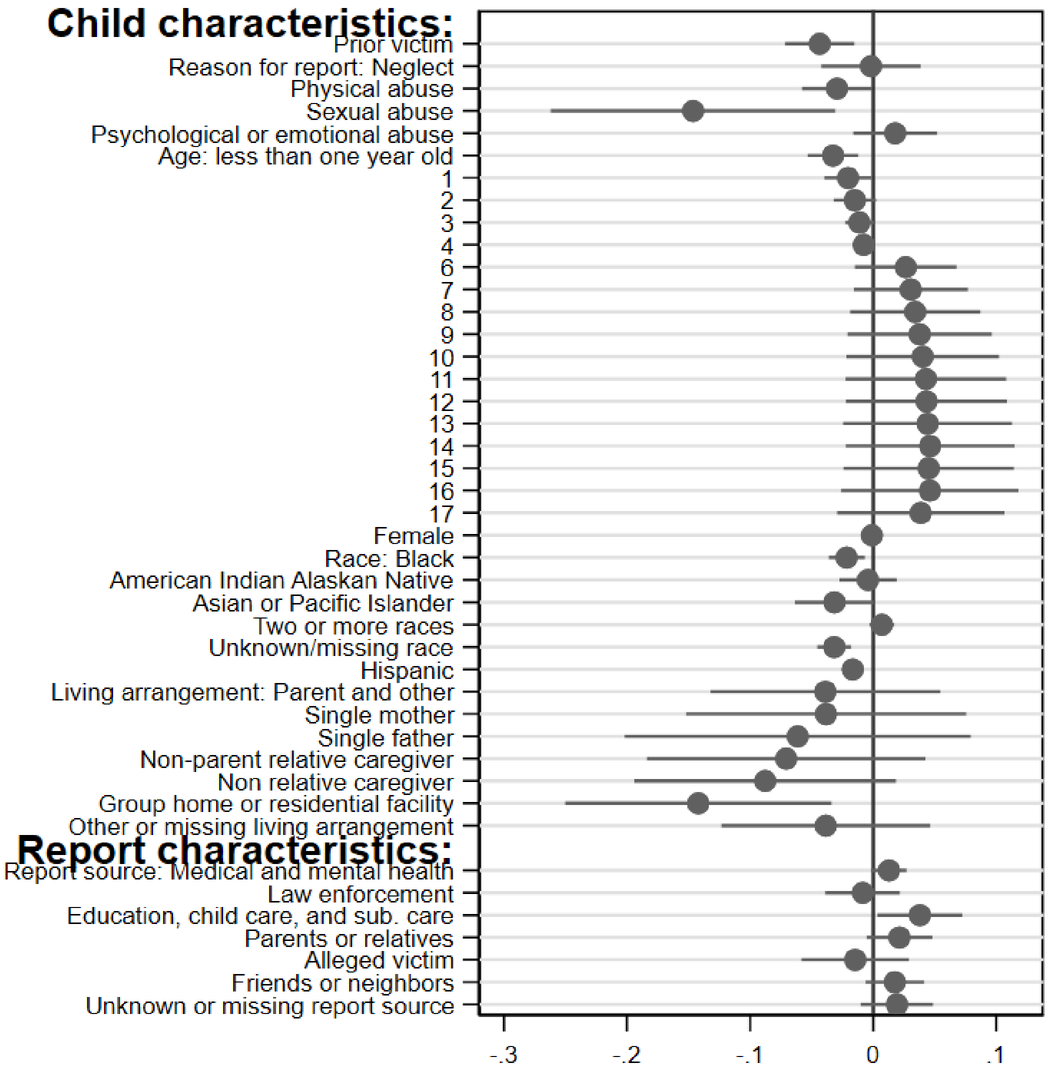
Notes. Author's creation. Panel A is a traditional one-track investigative CPS system. Panel B is a DR system that has two tracks, and *only* makes the track determination after a report is screened in.

Figure 2.2. Trends in re-reporting outcomes by year of the report



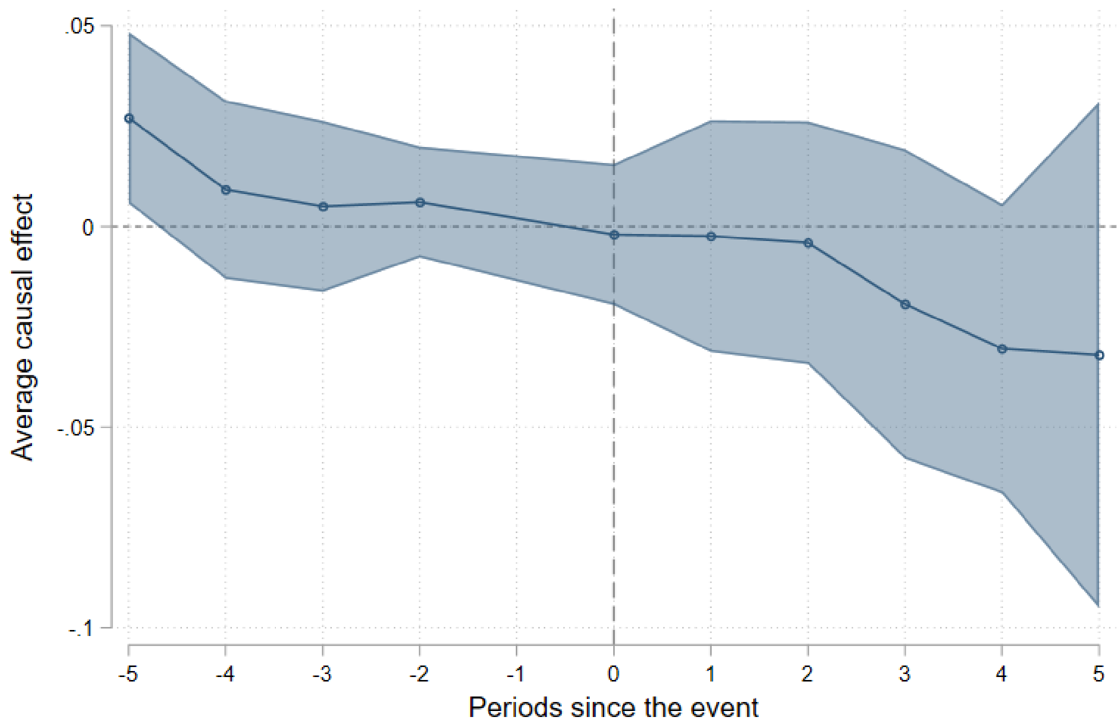
Notes. For my analytic sample of maltreatment reports, I show the proportion of report-child observations that are re-reported each year given different intervals of time.

Figure 2.3. Predictors of an alternative response in periods with differential response



Notes. In this coefficient plot, the circles indicate point estimates and bars indicate the 95 percent confidence intervals. Omitted reference groups which are not shown include: non-prior victim, no neglect, no physical abuse, no sexual abuse, no psychological or emotional abuse, age 5, race=White, non-Hispanic, living arrangement=both parents (married, unmarried, or unknown marital status), report source=social services personnel. The coefficients for months, local characteristics, and state and year fixed effects are not shown. The ordinary least squares regressions use cluster robust standard errors that are clustered by state. The full specification is shown in column 3 of Appendix Table C2.3.

Figure 2.4. Event studies estimating the effects of differential response reforms on re-reporting outcomes



Notes. Unconditional Callaway and Sant'Anna event study with never treated comparison group. Year t-1 is omitted. Bars indicate the 95 percent confidence intervals.

Table 2.1. Estimating the effect of differential response on re-reporting

	ATT	ATT by risk tercile		
		Low risk tercile	Middle risk tercile	High risk tercile
	(1)	(2)	(3)	(4)
Differential Response	-0.02	-0.02*	-0.03**	-0.03*
	[-0.059, 0.019]	[-0.038, 0.003]	[-0.059, -0.008]	[-0.056, 0.004]
Observations	15,660,574	4,202,890	4,239,644	4,263,808
Mean: outcome	0.46	0.35	0.47	0.60
Mean: AR | DR	0.11	0.11	0.13	0.09

Notes. The dependent variable is a re-report within 3 years. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the unconditional Callaway and Sant'Anna models in equation (1) with never treated as the comparison group. Risk terciles were constructed within state using weights from the periods without DR to assign a reasonable counterfactual risk measure for report-child observations in DR periods; the assigned "risk" of a re-report is based on what it would be in the absence of the DR reform. Wild bootstrapped t-statistics are used to create confidence intervals and determine significance because of the small number of treated state clusters. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table 2.2. Heterogeneous effects of differential response across baseline characteristics: race or ethnicity of the reported child

	ATT by race					ATT by ethnicity	
	White	Black	American Indian, Alaskan Native	Asian or Pacific Islander	Two or more races	Non-Hispanic	Hispanic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Differential Response	-0.02	-0.04*	-0.00	0.01	-0.02	0.01	-0.01
	[-0.052, 0.013]	[-0.079, 0.004]	[-0.044, 0.041]	[-0.023, 0.047]	[-0.058, 0.017]	[-0.038, 0.048]	[-0.025, 0.005]
Observations	8,494,825	2,863,724	115,800	94,091	581,336	94,091	581,336
Mean: outcome	0.48	0.47	0.49	0.28	0.57	0.28	0.57
Mean: AR | DR	0.13	0.10	0.09	0.08	0.16	0.08	0.16

Notes. The dependent variable is a re-report within 3 years. Race categories (1-5) are mutually exclusive. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the unconditional Callaway and Sant'Anna models in equation (1) with never treated as the comparison group. Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table 2.3. Heterogeneous effects of differential response across baseline characteristics: reason for the report

	ATT by reason for the report			
	Neglect	Physical abuse	Sexual abuse	Psychological/ emotional abuse
	(1)	(2)	(3)	(4)
Differential Response	-0.03	0.01	-0.00	0.06
	[-0.065, 0.008]	[-0.053, 0.066]	[-0.042, 0.038]	[-0.032, 0.155]
Observations	12,143,294	4,390,501	1,492,429	1,103,376
Mean: outcome	0.48	0.44	0.38	0.46
Mean: AR | DR	0.09	0.13	0.00	0.16

Notes. The dependent variable is a re-report within 3 years. Reasons for a maltreatment report are not mutually exclusive. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the unconditional Callaway and Sant'Anna models in equation (1) with never treated as the comparison group. Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent

Chapter 3. Accountability and the Role of Federal Oversight on Foster Care Permanency Outcomes

3.1 Introduction

In the early 2000s, a major federal push for accountability in the child welfare system addressed the fact that child welfare outcomes varied greatly state by state. The implementation of the Child and Family Services Reviews (CFSR) began to assess each state's performance on numerous items related to permanency and two other key areas: safety and well-being. Each item on the CFSR that is evaluated and identified as an 'area needing improvement' is addressed in the state's written Program Improvement Plan (PIP) in a process overseen by the federal Children's Bureau**.** This process works to keep state child welfare agencies accountable to federal standards.

Permanency is a vital objective of the child welfare system. Permanency means the child has legal membership in a safe and stable family. While this can reflect children transitioning from foster care to adoption or legal guardianship, this study focuses on permanency in the form of family reunification. Family reunification occurs when a child who has been in foster care (state custody) is placed back with his or her family and the parents regain custody of the child. Since foster care is often used as a temporary intervention,[17] family reunification is the most common

[17]Foster care agencies vary in how aggressive they are in removing children from their homes as they balance family preservation and child protection. There is a small body of literature that exists looking at the causal effects of foster care which leverages the decision-making tendencies by caseworkers or investigators. Early research by Doyle (2008; 2007) provides evidence from Illinois that among children on the margin of foster care placement, children placed in foster care fared worse in terms of future crime and employment than those who remained in their own home. However, new research shows the foster care system in Illinois is not representative of other states. Gross and Baron (2021) use data from Michigan to show that compared to children who were not placed in foster care, those who were placed in foster care made significant, persistent gains in safety and education outcomes when most children were

case goal for children in foster care. In 2019, there were over 664,000 children ages 0-17 in foster care across the country and 52 percent had reunification as their stated case goal (among those with a non-missing case goal). Reunification requires that the parents have made requisite changes supporting that the child will be returning to a safer home environment than he or she left.[18]

There is a growing body of literature that looks at the determinants of family reunification (Akin 2011; Cheng 2010) and foster care re-entry (Koh & Testa, 2011; Jedwab & Shaw, 2017; McDonald et al., 2006; Shaw 2006; Kimberlin et al., 2009; Wulczyn et al., 2020). However, to date, there are no studies that look at the causal effects of child welfare policies, reviews, or program reforms on family reunification and foster care re-entry outcomes in the child welfare system. This is an understudied area that is of great importance to child welfare policy considering the hundreds of thousands of children in foster care each year.

The purpose of this study is to assess whether the early rounds of the CFSR directly affected the foster care re-entry rate, and whether addressing the re-entry rate had possible unintended consequences — specifically, a lower family reunification rate. Given the variation in treatment (CFSR) timing, a two-way fixed effects differences-in-differences model is used to assess the influence of the first two rounds of the CFSR on the foster care re-entry rates and the family reunification rates. Despite the Adoption and Safe Families Act of 1997 which articulates clear

---

reunified with their birth parents about three to five years after initial placement. One explanation for this is that birth parents improved their parenting skills while their child was in foster care to become able to provide a more stable and safer environment for their child upon reunification.

[18] While little causal research looks at how parents change their behaviors with child welfare involvement, Grimon (2020) provides evidence from Allegheny county in Pennsylvania that child welfare encourages parents to seek help they may need (even when no child is removed from the home). This is particularly evident for mothers whose enrollment in housing assistance, mental health and substance abuse services increases with child welfare involvement.

preferences for family reunification for children in foster care, distinct efforts to improve reunification by reducing re-entry were not made until there was federal oversight in the early 2000s per the CFSR. I find that the first round of the CFSR significantly contributed to the national decline in the average state-year re-entry rate as states that had previously been above national standards lowered their rates following the first round of the CFSR. These improvements were sustained over time.

3.2 Background: A Push for Accountability

Federal child welfare statutes aim to create more stable and safe home environments for children and preference family reunification whenever possible. This is evident in the longstanding Adoption and Safe Families Act (ASFA) of 1997. This Act requires that states demonstrate that reasonable efforts were made to prevent the removal of children from their homes and to reunify children who been removed (Congress, 1997). It is crucial to ensure that children who are reunified with family are returning to more stable environments to avoid re-occurring abuse or neglect and subsequent foster care placements (i.e., foster care re-entry). Foster care re-entry following reunification is an important benchmark for permanency and a strong indicator that the agency's programs and practices are not effective in supporting reunification or are possibly reunifying families too early. Unsuccessful family reunification puts children at further risk of abuse and neglect and creates undesirable churn of children in and out of the foster care system. An important improvement in permanency would be for there to be a lower rate of children re-entering foster care following family reunification.

In the early 2000s, the Children's Bureau began implementing the Child and Family Services Reviews (CFSR) to gain an understanding of how each state's child welfare system performs for children and families who receive services. These reviews are a form of outcome monitoring for numerous items in three key areas: safety, permanency, and well-being. To date there have been three complete rounds of the CFSR[19] rolled out over the following years**:**

- CFSR round 1 was administered between 2001-2004
- CFSR round 2 was administered between 2007-2010
- CFSR round 3 was administered between 2015-2018

The 12-month re-entry rate is a specific permanency item that is referenced in the first two rounds of the CFSR.[20] In the first round of the CFSR, foster care re-entries needed to meet two standards: 1) the state's incidence of foster care re-entry within 12 months of discharge from a prior episode needs to be at or below the national standard of 8.6 percent, and 2) 90 percent or more of the applicable reviewed permanency cases must have been rated as a 'strength' on the item because there was not a re-entry into foster care within a 12-month period. If these two conditions were met, the state would receive the overall rating of 'strength' for foster care re-entries. Alternatively, if the 90 percent threshold is not reached for applicable cases *or* the state did not meet the national standard for re-entry rates, 8.6 percent, the state received an overall rating of 'area needing improvement' for that item.

---

[19] The Children's Bureau began to conduct a fourth round of CFSR beginning in Federal Fiscal Year 2022, which is beyond the period included in this study.

[20] In the first and second rounds of the CFSR, 'foster care re-entries' is 'Item 5' for 'Permanency Outcome 1.'

There were changes made to CFSR round 2 for numerous evaluated items. One major change for the foster care re-entries item was that states now only had to meet one standard: more than 90 percent of the applicable (randomly reviewed) permanency cases must have been rated as a 'strength' on the item because there was not a re-entry into foster care within a 12-month period. Now, if the 90 percent threshold is not reached for applicable cases, the state would receive an overall rating of 'area needing improvement' for that item. The state's incidence of foster care re-entry was not considered for the determination of whether this was an 'area needing improvement' for the second round of the CFSR.

The Children's Bureau uses the CFSR as a tool to then assist states in enhancing their capacity to achieve positive outcomes related to safety, permanency, and well-being. Following the CFSR, each item that is not in accordance with national performance standards needs to be addressed in the state's written Program Improvement Plan (PIP), which is advised and approved by the Children's Bureau and published**.** These PIP must include measurable goals for improvement, action steps, and an implementation timeline. Examples of actionable steps related to improvements in foster care re-entry rates might include the adoption or improvement of practices such as trial home visits prior to reunification or providing services to families.

While the content of each state's PIP is not measured and included in this study, the plans are presumably a possible mechanism of change for states needing to improve. This theory of change is seen in my logic model (Figure 3.1).

3.3 Data and Empirical Strategy

3.3.1 Data

This paper combines four data sources for 2000-2019 for all 50 US states. First, foster care files that state agencies were required to report to the Children's Bureau are used from the Adoption and Foster Care Analysis and Reporting System (AFCARS). These data are disseminated from the National Data Archive on Child Abuse and Neglect at Cornell University (National Data Archive on Child Abuse and Neglect, 2019). I appended the annual AFCARS child case-level files so that I have a child-case-year level panel. In addition to data on foster care re-entry and reunification, I have child-level information about age, race, and reasons for removal from their families (e.g., physical abuse, sexual abuse, neglect, parent incarceration).

My second source of information comes from publicly available PDF reports that were downloaded from the Children's Bureau website. Each state has a CFSR report corresponding to each round of review. These provide the treatment year for each round and an indicator for whether the state identified foster care re-entries were a 'strength' or an 'area needing improvement' that round. Further, the reports provide the state's 12-month foster care re-entry rate and the percentage of the applicable permanency cases were rated as a 'strength' because there was not a re-entry into foster care within a 12-month period.

Two additional datasets provide annual state demographic characteristics and economic conditions. I construct state-year characteristics on demographics and unemployment rates from the IPUMS American Community Survey (Ruggles et al., 2019). I calculate the natural log of average household income, adjust for inflation, and put it into 2018 dollars. Additionally, I calculate the percent of households that are below the poverty threshold.[21] Annual unemployment rates by state are calculated as the number of unemployed individuals aged 25-64 divided by the number of individuals aged 25-64 in the labor force.

3.3.2 Permanency Outcomes

Then, in alignment with the federal priorities articulated in the CFSR (described in section 2), I focus my analysis on state-level measures of permanency. The state-level 12-month re-entry rate (in percent) is calculated via equation (1). Using the child-case-year observations, I can identify children who re-enter foster care within 12 months of discharge based on the child's current removal date and the date of their last discharge.[22] I aggregate the number of children re-entering within 12 months for the numerator of equation (1). Note, this is aggregated given the calendar year they were discharged, which is not necessarily the calendar year they re-entered. For the denominator, I aggregate the number of children reunified with family to the state-year level given the calendar year a child was discharged. Given that I do not know how many children discharged in 2019 will re-enter within one year (i.e., 2019 observations are right-censored), the re-entry rate is calculated for 2000-2018.

$$\text{Re} - \text{entry rate} = \frac{\#\ \text{re} - \text{entering within 12 mo.}}{\#\ \text{reunified this year}} \times 100 \tag{1}$$

[21] The official "poverty line" depends on total family income, the size of the family, the number of people in the family who are children, and the age of the householder (under/over age 65) (*Poverty* variable in IPUMS).
[22] Each child-year observation has the date of the previous foster care discharge if the current foster care placement is a re-entry (i.e., the current placement is the 2+ time in foster care).

Given that re-entry is conditional on reunification in this study, I also directly assess reunification as a secondary outcome of interest. In equation (2), I identify the children in foster care who are reunified with family within 12 months of entering foster care. For each state-year, I calculate the reunification rates (in percent) corresponding to the calendar year children entered foster care, which might not be the calendar year they went home if reunified. While reunification may occur beyond the 12-month threshold, as clearly seen in Figure Figure , this standard timeframe is referenced in the CFSR for achieving permanency in a timely manner.

$$\text{Reunification rate} = \frac{\#\ \text{reunified within 12 mo.}}{\#\ \text{entering FC this year}} \times 100 \tag{2}$$

States greatly vary in the rate at which they reunify children back with birth families and states that have higher 12-month reunification rates consistently have higher 12-month re-entry rates. In fact, the state-year 12-month reunification rates are positively correlated with 12-month re-entry rates. This is shown in Figure 3.2 where the correlation coefficient 0.261 is statistically significant at the 1 percent level. One way to address re-entry rates could be to reunify fewer families in a timely (e.g., 12-month) manner. Therefore, throughout this study, I assess whether changes to 12-month re-entry rates coincided with changes to 12-month reunification rates.

3.3.3 Foster Care Re-entries as a 'Strength' or an 'Area Needing Improvement'

As discussed in section 2, the first two rounds of the CFSR evaluated whether foster care re-entries were a 'strength' or an 'area needing improvement' according to different criteria. The criterion for CFSR round 1 was two-dimensional, such that the CFSR item evaluating foster care re-entries need to meet two standards: 1) the state's incidence of foster care re-entry within 12 months of discharge from a prior episode is needs to be below the national standard of 8.6 percent, and 2) 90 percent or more of the applicable permanency cases must have been rated as a 'strength.' Since these criteria are both observable and continuous, Figure 3.3 shows two-way scatterplots for the state foster care re-entry rate (percent as reported in the CFSR reports) and the CFSR cases rated a 'strength' (percent as reported in the CFSR reports) for each CFSR round. The dashed lines in the scatterplots represent the respective 8.6 percent threshold and 90 percent threshold. The scatterplot for CFSR round 1 shows the states that were a 'strength' are plotted in green (bottom right quadrant of the plot) and states that had foster care re-entries as an 'area needing improvement' are plotted in orange.

Whereas the criterion for CFSR round 1 was two-dimensional, the major change for CFSR round 2 was that the foster care re-entries item now only had to meet one standard: 90 percent of the applicable permanency cases must have been rated as a 'strength' on the item because there were no re-entries into foster care within a 12-month period. In Figure 3.3, the scatterplot for CFSR round 2 shows the states that had foster care re-entries as a 'strength' plotted in green (right-side of the plot) and states in which re-entries are an 'area needing improvement' are plotted in orange (left-side of the plot). As seen in the scatterplot, each state's incidence of foster care re-entry was not considered for the determination of whether re-entries being a ‘strength’ or an ‘area needing improvement’ for the CFSR round 2.

3.3.4 Descriptive Statistics of Key Variables

State-year summary statistics of key variables are shown in Table 3.1. On average, 10.5 percent of children re-enter foster care within one year of discharge and 20.5 percent of children are reunified with family within one year of entering foster care. During each round of the CFSR, based on assessment, the state can identify re-entries as a ‘strength’ or an ‘area needing improvement.’ In Table 3.1, during CFSR round 1, over half (54 percent) of all states found re-entries identified as an ‘area needing improvement.’ According to the results of the CFSR, large improvements in terms of foster care re-entries were made before the next round. In CFSR round 2, the percentage of states that identified that re-entries were an ‘area needing improvement’ dropped to 24 percent.

Figure 3.4 illustrates how these two outcome variables change between 2000 and 2018. The average re-entry rate declined from 18.5 percent on average in 2000 to 7.9 percent on average in 2018 and reunification declined from 22.3 percent on average in 2000 to 15.9 percent 2018. Additionally, over this period the 95 percent confidence intervals for the re-entry rates decrease greatly over the first decade (2000 to 2010) as there is less variation in re-entry rates across states. The confidence intervals for reunification rates on the other hand remain relatively large and consistent over time.

Research on known risk factors and the determinants of re-entry and reunification informs the current study. Child factors such as age, race, and disability are known factors that influence re-entry into foster care (Koh & Testa, 2011; Shaw, 2006; Jedwab & Shaw, 2017; McDonald et al., 2006; Kimberlin et al., 2009). Hazards of re-entry are higher for teenagers and infants (Parolini et al., 2018; Kimberlin et al., 2009), African American children relative to white children (Kimberlin et al., 2009) and children with mental health concerns relative to children without such a history (Parolini et al. 2018). These risk factors inform the compositional characteristics of each state's foster care caseload that are included as covariates in analyses.

Turning to the caseload characteristics in Table 3.1, among discharged children, the state-year average length of time that children spent in foster care was 19.9 months. Roughly a third (35.2 percent) of children in the average state caseload are nonwhite (i.e., Black or African America, American Indian or Alaskan Native, Asian, Hawaiian or Pacific Islander, or Hispanic), 47.9 percent are female, and 22.9 percent have a clinical disability.

The caseload characteristics show six reasons for removing the child include: physical abuse, sexual abuse, neglect, alcoholic parent, parent's drug abuse, and parents are in jail. Average state trends in the reasons for removal over time are shown in appendix Figure A3.3. These reasons for removal are not mutually exclusive and each case can have up to three reported reasons. For example, a child could have been removed from their household and placed in foster care due to both neglect and the drug abuse of their parent(s). Therefore, the summation of the removal reasons within a state-year does not sum to 100. Over half of children in the average state-year caseload (56 percent) were removed from their homes due to neglect. The average state caseload has 24.6 percent of its children removed from home due to having a drug abusive parent and 8.5 percent removed due to physical abuse.

Economic conditions are particularly important to control for because there has been consistently observed correlations between low household income and child abuse and neglect

(Ginther & Johnson-Motoyama, 2017; Paxson & Waldfogel, 2003; Paxson & Waldfogel, 1999; Berger et al., 2017; Berger & Waldfogel, 2004; Raissian & Bullinger, 2017). In addition to caseload characteristics, my analyses control for time-varying economic conditions. In Table 3.1 the state-year natural log of household income ranges from 10.66 to 11.40 with an average of 11.03. The average state-year has 15 percent of households below the poverty line. Some evidence shows that states with higher incomes have lower rates of child abuse and neglect (Paxson & Waldfogel, 1999) and increases in income, even through cash-assistance programs like TANF, are known to decrease child maltreatment and foster care placements (Ginther & Johnson-Motoyama, 2017). Exogenous increases in income are associated with reductions in child neglect and CPS involvement, particularly for children of low-income single-mothers (Berger et al., 2017). Research also shown that the socioeconomic determinants of child maltreatment include parental work status (Paxson & Waldfogel, 1999) and unemployment (Brown & Cao De, 2018). In Table 3.1, the state-year average unemployment rate is 0.05.

Since the caseload characteristics and economic conditions are time-varying and Table 3.1 only shows the overall sample mean, appendix Table A3.2 shows the mean and standard deviations for states in the year that states are reviewed for each round. Changes in the caseload characteristics over this period could affect the outcomes of interest.

3.4 Empirical Strategy

3.4.1 Two-way Fixed Effects Differences-in-Differences Models

The first set of analyses in this study take advantage of the variation in treatment (CFSR) timing. Each state was reviewed (i.e., treated) in one year during each round of the CFSR. The staggered treatment of each state within each round of the CFSR is an important dimension for the study design. Each state's treatment years are shown in appendix Table A3.1. The variation in the timing of CFSR rounds produces 5-7 years between rounds 1 and 2. There is variation such that states treated in the first year of round 1 (i.e., 2001) are not necessarily treated in the first year of round 2 (i.e., 2007), and states treated in the second year of round 1 are not necessarily treated in the second year of rounds 2, and so on.

Identifying the causal impact of the CFSR is tricky because all states are treated at some point during round 1 and round 2. However, states are only reviewed for one year in each round and each round spans four years. I leverage this treatment timing variation using a two-way fixed effects differences-in-differences (DD) model to assess the relationship between the CFSR and permanency outcomes: reunification and re-entry. This research design uses panel of state-year data to compare the outcome for the states that received that round of the CFSR to control states that have not yet received that round of the CFSR to quantify the impact of the treatment.[23]

The general form of the differences-in-differences model with variation in treatment timing model is expressed as follows in equation (3):

$$Y_{st} = \sum_{i=1}^{2} \beta_i R_{ist} + \gamma X_{st} + \alpha_s + \eta_t + \varepsilon_{st} \tag{3}$$

[23] Since not all states are treated in the same year this methodology allows me to analyze earlier cohorts relative to later cohorts. Thus, states treated in the first year of round *i* are compared to all states not yet treated. Then in the second year of round *i*, there are two cohorts of treated states compared to the two cohorts of untreated states. Similarly, in the third year of round *i*, there are three treated cohorts compared to one untreated cohort. The states treated in the fourth and last year of round *i* have been a control group for all earlier treated cohorts, yet do not have a control group themselves. For example, the coefficient on the β_i for $i = 1$, is the average effect of the first CFSR round relative to those who haven't received it yet. When states undergo CFSR rounds 2 and 3, similar comparisons are made such that the last year of treated states have no control group.

where, Y_{st} denotes, separately, the 12-month reunification rate (percent) and 12-month foster care re-entry rate in each state *s,* at year *t.* R_{ist} is the treatment variable indicating the CFSR review status (0/1) of each round $i \in (1, 2)$ of the CFSRs Following a in state *s,* at year *t.* review, the indicator is 0 if the state has not yet received that round of the CFSR and stays on as 1 if it has. The X_{st} vector of covariates varies at the state-year level and includes caseload characteristics (e.g., percent of the foster care caseload that is female), the average log household income (adjusted for inflation), percent in poverty, and the unemployment rate (percent. State fixed-effects, α_s, account for time-invariant differences between states, such as differences in funding, child welfare practices, or differences in definitions of child maltreatment (i.e., neglect and abuse). The year fixed effects (indicators), η_t, control for the common temporal trend in specific calendar year *t,* which relates to the year of which the child was discharged from foster care and reunified with family (which is not necessarily the year they re-entered). Finally, ε_{st} is the error term with standard properties.

Equation (3) allows to me measure the average treatment effect on the treated (ATT) following each round of the CFSR. However, my logic model specifically predicts that the CFSR reduces re-entry rates because states not in accordance with federal standards are identified and held accountable by the Children's Bureau via an approved Program Improvement Plan. This logic hypothesizes that the reviews would not drive reductions in re-entry for all states on average, but instead would specifically drive reductions for states in which re-entries are identified as an 'area needing improvement' following each round of the CFSR rather than a 'strength.' Therefore, I fully anticipate that the ATT is different for each group of states in which re-entries are a 'strength' or an 'area needing improvement.'

In equation (4), I adapt equation (3) by interacting the CFSR treatment variable R_{ist} with A_{ist}, where A_{ist} is an indicator variable (0/1) for re-entries being identified as an 'area needing improvement' in each round $i \in (1, 2)$ of the CFSRs in state *s,* at year *t*:

$$Y_{st} = \sum_{i=1}^{2} \beta_i \, R_{ist} + \sum_{i=1}^{2} \delta_i \, (R_{ist} \times A_{ist}) + \gamma X_{st} + \alpha_s + \eta_t + \varepsilon_{st} \qquad (4)$$

In equation (4), the β_i coefficients tell us the average effect of each treatment round i for states in which re-entries that round were a 'strength' and *not* an 'area needing improvement' ($A_{ist} = 0$). Then, the δ_i coefficients tell the differential effect of each treatment round *i* for states in which re-entries are an 'area needing improvement' ($A_{ist} = 1$). Since there can only be a formal 'area needing improvement' when there is a review year, there is not a standalone indicator for A_{ist}. All other terms are defined the same for equation (4) as they were for equation (3).

Since I have state-level analysis and relatively small N, I model equations (3) and (4) using wild cluster bootstrapped t-statistics to create confidence intervals (Roodman et al., 2019). The p-values from the wild bootstrapped t-statistics are clustered by state with 1000 repetitions (Roodman, 2022). While the coefficients and the t-statistics are the same as they would be in a cluster robust model, the bootstrapping of the t-statistics changes the p-values (which reveal the significance) since it uses the bootstrapped distribution rather than the normal distribution. Compared to the asymptotic normal distribution, the bootstrapped distributions lead to better approximations of my outcomes (Cameron et al., 2008; MacKinnon et al., 2022).

3.3 Results

3.5.1 Reunification and Re-entry Rates

In Table 3.2, I present the main results from my two-way fixed effects difference-in-difference models (equations 3 and 4) in which the 12-month re-entry rates (panel A) and reunification rates (panel B) are the dependent variables (in percent). Each specification includes the foster care caseload controls and state-year economic controls shown in Table 3.2 (full results are in the appendix). In panel A, column 1, I use the model specified in equation (3) and find that the CFSR rounds 1 and 2 had no significant effect on the re-entry rate. In column 2, I use the model specified in equation (4) to interact each round of the CFSR with whether the state identified during that round that re-entries are an area that needs an improvement. This allows me to test for differential effects of the CFSR since we anticipate that reductions in re-entry primarily occur for states in which the CFSR identified that re-entries needed to be improved to become aligned with federal standards.

Column 2 shows that states that identified that re-entries need improvement in the first round of the CFSR reduced foster care re-entries by 2.46 percentage points on average (calculated as 3.4 0– 5.86), a difference that is significant at the 1 percent level. Comparatively, the coefficient for states that were already aligned with federal standards was positive and significant; re-entries increased by 3.40 percentage points in these states. In round 2 of the CFSR, states that identified that re-entries needed improvement experienced nearly no significant change in re-entry rate (calculated as 0.10 – 1.04).

Reasonably we might anticipate lagged effects if states are not able to respond to the review outcomes immediately in the next year. Columns 3 and 4 in panel A, lag the effects of the CFSR by one year. In column 4, states that identified that re-entries need improvement in round 1 of the CFSR had an average 3.25 percentage point reduction in the foster care re-entry rate (calculated as 3.42– 6.67).

If a state is aggressive in reunifying families, it may have a higher rate of re-entry even if it has a program of support. In this case, the state's response to the review might be to reunify lower rates of families. Therefore, it is possible that the decline in re-entries is a direct effect of a decline in reunification. In panel B, I assess the impact of the CFSR on the 12-month reunification rate. In columns 1 and 2, the coefficients on each round of the CFSR are not statistically significant. In columns 3 and 4, which lagged the effects by one year, I still find that all coefficients are not significant.

While there are two CFSR rounds analyzed in this study, the initial round of the CFSR resulted in the largest reductions in the re-entry rate. Impressively, results in column 2 regarding both re-entry (panel A) and reunification (panel B) following CFSR round 1 imply that states were reunifying at the same (statistically similar) rate following the first CFSR round yet had a significantly lower rate of children re-entering. This is a positive sign that improvements to the re-entries did not come at the cost of lower rates of reunification. These trends in the average 12-month reunification and foster care re-entry rates (percent) are reflected in Figure 3. relative to the timing of CFSR round 1 (year = 0). Pre-trends are shown up to 4 years prior to treatment and the post-trends show up until 4 years after the first round. CFSR round 2 does not occur until 5-7 years after the first round. States in the left panel with re-entries as a strength had lower average re-entry and reunification rates than states in the right panel in which re-entries are an area needing improvement. For states in which re-entries are an area needing improvement, there has been a general decline in re-entry rates both before and after the treatment year 0.[24]

3.5.2 Foster Care Re-entry Rates by Age-groups

The average annual re-entry rate for children 0-17 masks large differences in re-entry rates by age-group (Figure 3.). I define age-groups as: under 1 year old (infant), 1-2 years old (toddler), 3-5 years old (preschool), 6-11 years old (elementary), 12-13 years old (middle school) and 14-17 years old (high school). This follows the age groups defined by Wulczyn et al. (2020) who look at the determinants of re-entry for these groups. In a relatively monotonic fashion, the re-entry rate generally increases on average (and never decreases) with increasing age-groups in which infants under 1 years old have the lowest 12-month re-entry rates and high schoolers, ages 14 -17, have the highest 12-month re-entry rates. The rates for 3-5-year-olds and 6-11-year-olds are relatively identical.

Given the heterogeneity in re-entry rates illustrated in Figure 3., we might expect heterogeneity in which some age-groups experience larger effects than others. To formally test this, I estimate equation (4) separately by age group. The results are shown in Columns 1 in Table 3.3 is the same as panel A, column 4 in Table 3.2. Indeed, it appears that the 12-month re-entry rate for reunified children differs by age-group. The effect on states that identified re-

[24] Appendix Figure A4 shows the 12-month re-entry rates by treatment cohort years (e.g. cohort of states treated in 2001, vs cohort of states treated in 2002, etc.) separately for states in which re-entries were a 'strength' or an 'area needing improvement' to visually assess whether states that are treated later (e.g. the cohort treated in 2004) appear to have trends that allude to them making changes in anticipation of the upcoming reviewing. There are some downward pre-trends for the 2002 and 2003 cohorts among states in which states are an 'area needing improvement.' Other cohorts do not show changing pre-trends in the three years before round 1.

entries need improvement, as shown by the interaction term on *CFSR Round 1 × re-entries need improvement*, is substantial and negative across all groups (summing the coefficients on CFSR Round 1 and the interaction term) and statistically significant for children ages 3–5, 6–11, 12–13 and 14–17. In states in which re-entries were identified as an area needing improvements in round 1, re-entry rates decreased by 1.4 percentage points for children ages 3–5 (calculated as 2.5 – 3.9 in column 4), 1.1 percentage points for children ages 6–11 (calculated as 3.3 – 4.4 in column 5), 4.3 percentage points for children ages 12–13 (calculated as 3.2 – 7.5 in column 6), and 2.9 percentage points for children ages 14–17 (calculated as 4.3 – 7.2 in column 7). The disaggregated age-groups for children over the age of one in columns 3–7, reveal consistently positive and significant effects on the term *CFSR round 1* which is the effect on states that did not need to improve re-entries because re-entries were a strength.

The coefficients on the interaction terms *CFSR round 2 × re-entries need improvement* across columns 2–7 are all insignificant. The coefficients related to *CFSR round 2* across columns 2-7 are insignificant, except for being positive and significant on *CFSR round 2* for 3–5-year-olds (column 4) and 12–13-year-olds (column 6).

3.5.3 Robustness

Heterogeneous treatment effects may influence the estimates from my two-way fixed effect difference-in-difference models. In canonical DD, a regression contains two time periods, "pre" and "post", and two groups, "treatment" and "control" and thus is a function of pre/post and treat/control means. However, several recent papers point out that the two-way fixed effects estimator (year and state fixed effects in my case) is a weighted average of treatment effects, in which some of the weights can be negative if the treatment effect is not constant across groups and over time (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfoeuille, 2020; Goodman-Bacon, 2018).

In this study, since the treatment (i.e., CFSR review status) turns on at different times for each state, the regression DD coefficient is a weighted average of canonical "2x2" DDs (Goodman-Bacon, 2018). Every state is part of the control group in some 2x2 DD terms. The weights depend on the sample and come from two things: the size of each subgroup (what share of states that are in the treatment and control group for a given pair, and what share of time periods are used in a given 2x2 subsample), and the variance of treatment (how close to the beginning/end of the subsample window does treatment turn on). I address the possibility of heterogeneous treatment effects and follow Goodman-Bacon (2018) to decompose the DD estimate from my preferred specification (column 2) in Table 3.2. A scatterplot of the 2x2 DDs against the weights is shown in Appendix Figure A3.6 (Goodman-Bacon et al., 2019). I determine from this Goodman-Bacon analysis that the DD estimator is reliable and consistent.

3.6 Summary and Concluding Remarks

3.6.1 Summary and Policy Implications

Pressures from the Children's Bureau provoked states to improve their outcomes. The results from my differences-in-differences model with variation in treatment timing suggest that the federal push for accountability in the child welfare system and the implementation of the CFSR is associated with substantial re-entry reductions for states that did not meet federal standards. For states needing improvements, these reductions in re-entry following CFSR round 1 were large relative to the mean re-entry rate. Among these states needing to improve, these results are driven by reductions in re-entry rates for children ages 3-17 in CFSR round 1. This suggests that the overall introduction of the CFSR is associated with a reduction in foster care re-entries, but not necessarily each review round. Furthermore, improvements in re-entry have been relatively sustained. The PIP (mechanism of change) which corresponds to the implementation of new or improved practices for reunified families may also inform other in-home, prevention-focused efforts and policy initiatives. However, this does not allow me to rule out that the additional rounds did not contribute to the sustained improvements.

Re-entry in this study is conditional on having been reunified and states could respond to high re-entry rates (and other areas needing improvement) by reunifying fewer children. This might be a way of addressing the re-entry outcome on the CFSR, that could have negative consequences for children. Throughout this study, I assess whether changes to 12-month re-entry rates coincided with changes to 12-month reunification rates. Impressively, the reduction in the re-entry rate following the first round occurred without simultaneously significant reductions in the 12-month reunification rate. This is an example of how states addressed one issue (re-entry) without it coming at the expense of another (reunification).

The results suggest that state legislative priorities benefit from federal oversight. As discussed throughout, all states with foster care re-entries as an 'area needing improvement' needed to address the issue in a PIP that was approved by the Children's Bureau. The proposed theory of change in my logic model (Figure 3.1) hypothesizes that the PIP (mechanism of change) potentially corresponds to the implementation of new or improved practices. While the current analysis essentially assumes that PIPs have the same, averaged anticipated effects, states might implement a variety of new or improved programs, and some may be more effective than others. Therefore, it would be of great interest to further analyze the possible mechanisms behind the changes in re-entry rates. The PIP which strengthens reunified families and reduces re-entry might also inform other effective in-home, prevention-focused efforts to avoid first-time placements of children in foster care. Reducing the overall use of foster care is a renewed federal priority as seen from the Family First Prevention Services Act that took effect in October 2019.

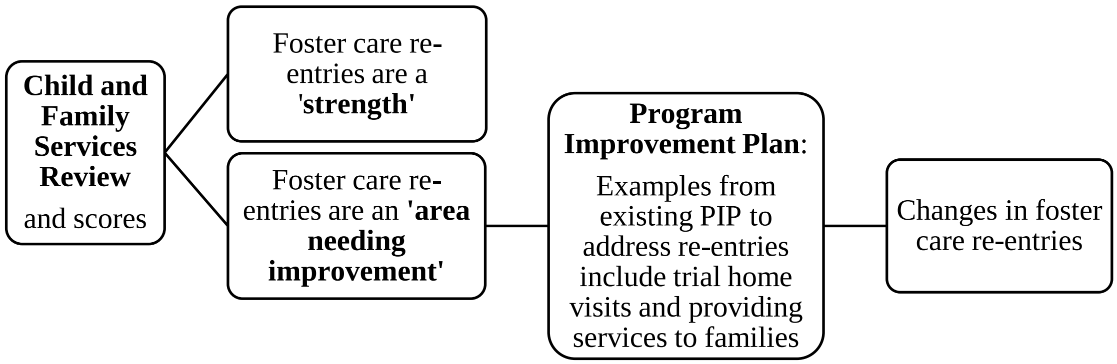
Identifying the content in the PIP would help policy makers know what works best. This is particularly important because qualitative research has concluded that child welfare agencies frequently institute new programs or practices without an evidence base of its effectiveness (Horwitz et al., 2014). The development of a better-informed social network between local and state colleagues, as well as the development of inter-state networks, may increase the use of evidence-based practices.

3.6.2 Limitations

One common concern with the difference-in-difference approach is that the treatment would be endogenous to unobserved state-level shocks. If a state systematically is reviewed after experiencing negative persistent shocks to permanency, the approach would underestimate the benefit of the CFSR. Another possible limitation is that if a state is reviewed after one or two bad years, the re-entry rates would likely revert to the mean, and the model would spuriously capture the impact even if the CFSR did not have a causal impact on re-entry rates.

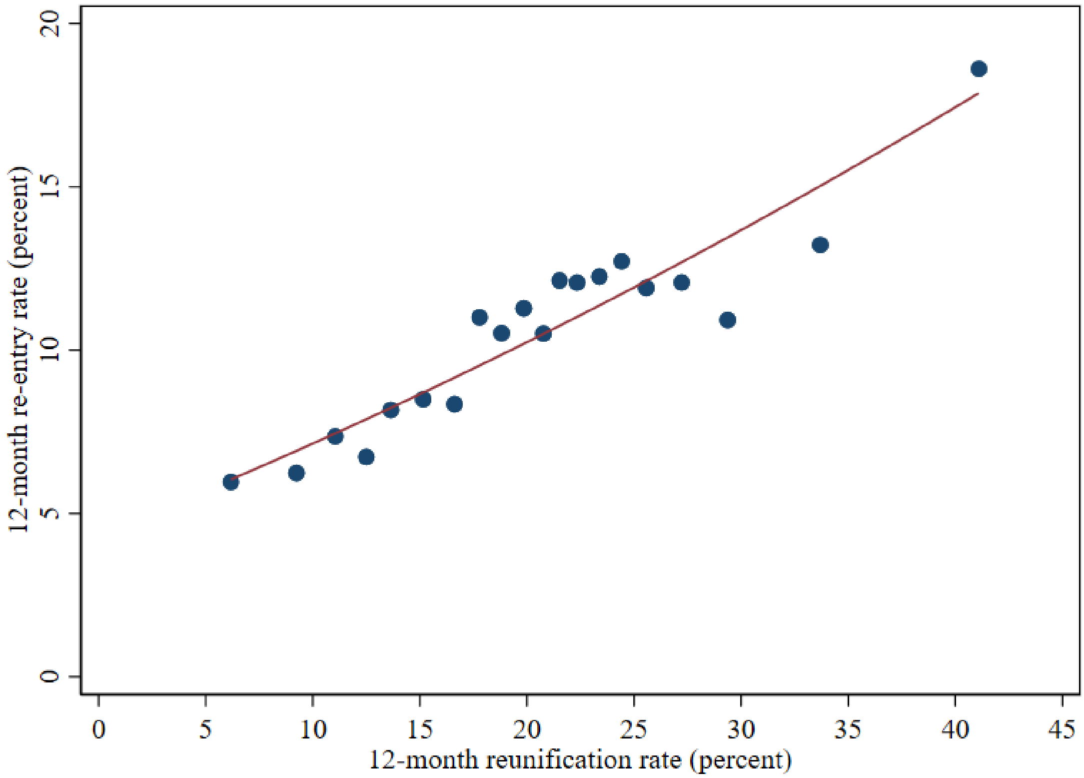
Moreover, there are some data limitations of the study. Given treatment (CFSR) timing variation, my model assesses the effect on earlier treated cohorts relative to later cohorts. For the earliest treated states, those treated by round 1 in 2001, I have only one pre-treatment observation. A larger limitation is that the states treated in the fourth and last year of each round have been a control group for all earlier treated cohorts, yet do not have a control group themselves. Additionally, while the state and year fixed-effects control for differences in definitions and practices between states and over time within states, it would be ideal to also control for county since child protection services is organized at the county-level. However, this is a limitation of the data since counties with fewer than 1,000 foster care children in a year are not identifiable in the public-use files.

Figure 3.1. Logic model



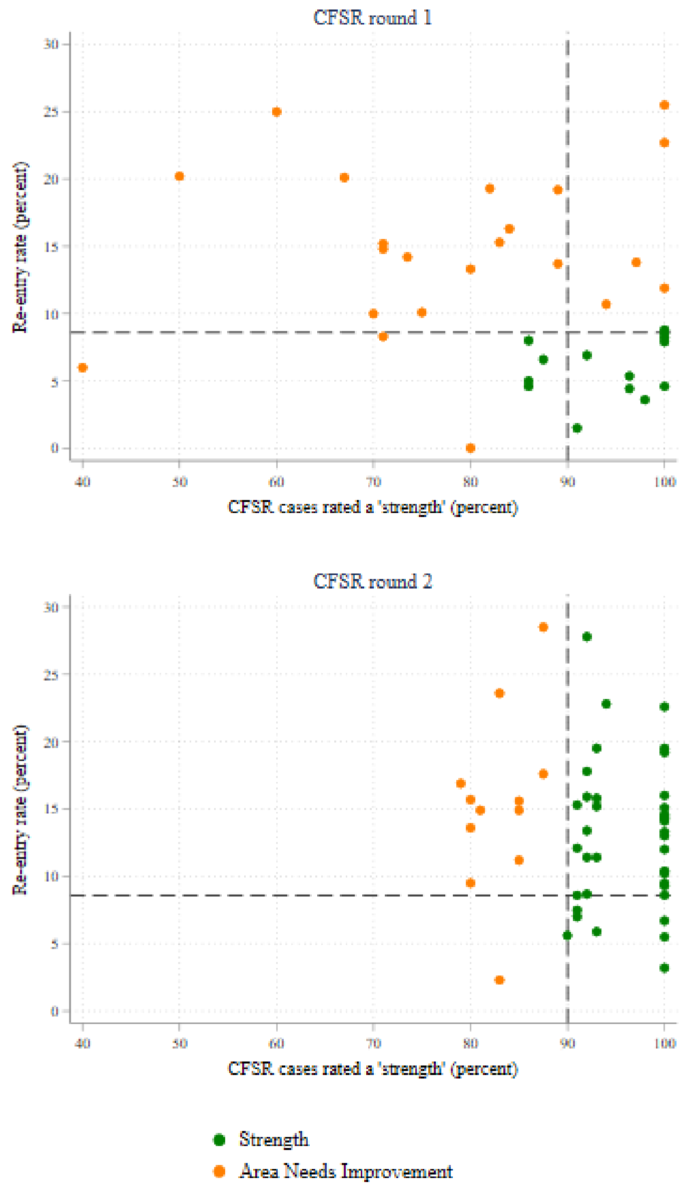
Source. Author's creation.

Figure 3.2. Correlation between the 12-months reunification rate and the 12-month re-entry rate



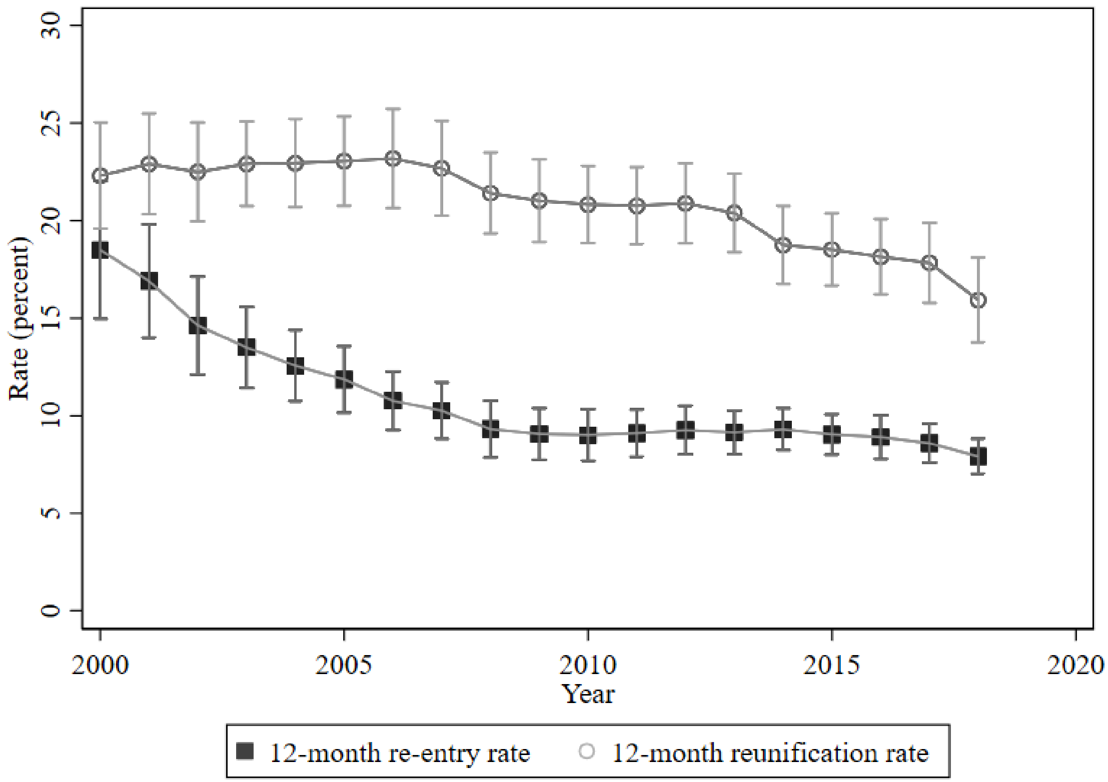
Notes. Nonparametric binscatter plot for state-year averages with no controls and quadratic line fit (Stepner, 2013). Correlation coefficient 0.261; standard error 0.093; t-statistic 2.82; p-value 0.005.

Figure 3.3. Two-way scatterplot with re-entry rate (percent) and CFSR cases rated a 'strength' (percent), by CFSR round



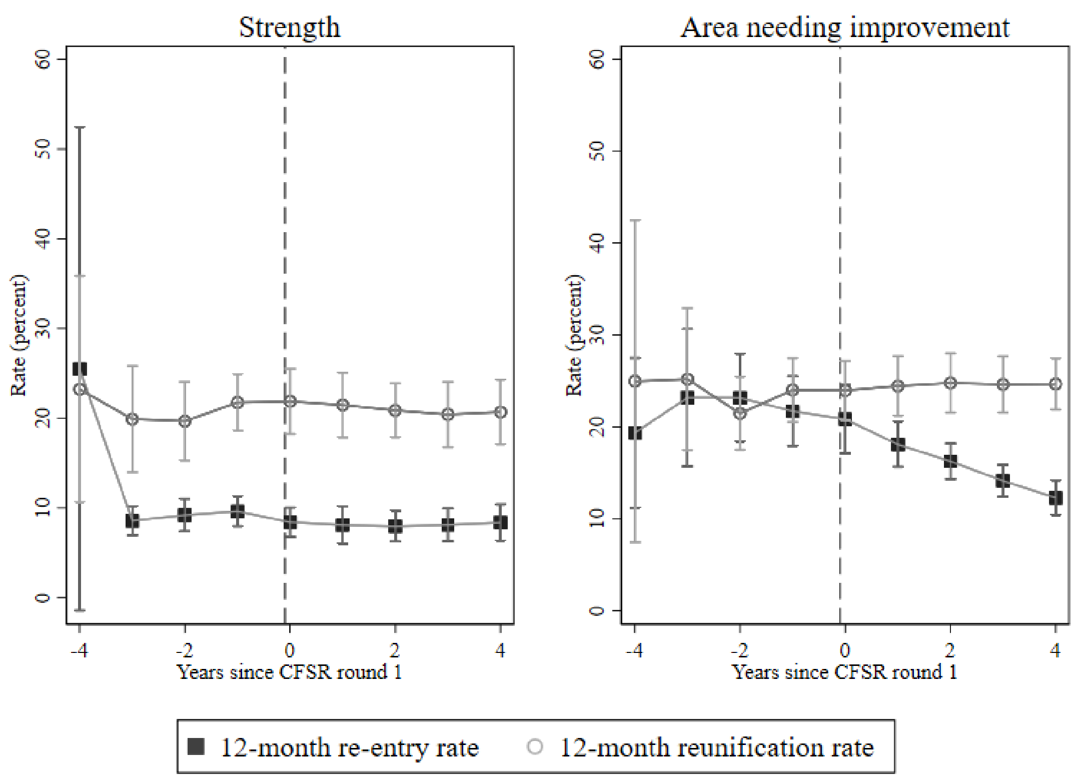
Notes. Each state is plotted for each round. The horizontal dashed line illustrates that the national standard for the re-entry rate was 8.6 percent. The vertical dashed line for the bottom histogram illustrates the 90 percent threshold for applicable reviewed cases that were rated a strength.

Figure 3.4. Trends in the average 12-month reunification rate and in the 12-month re-entry rate



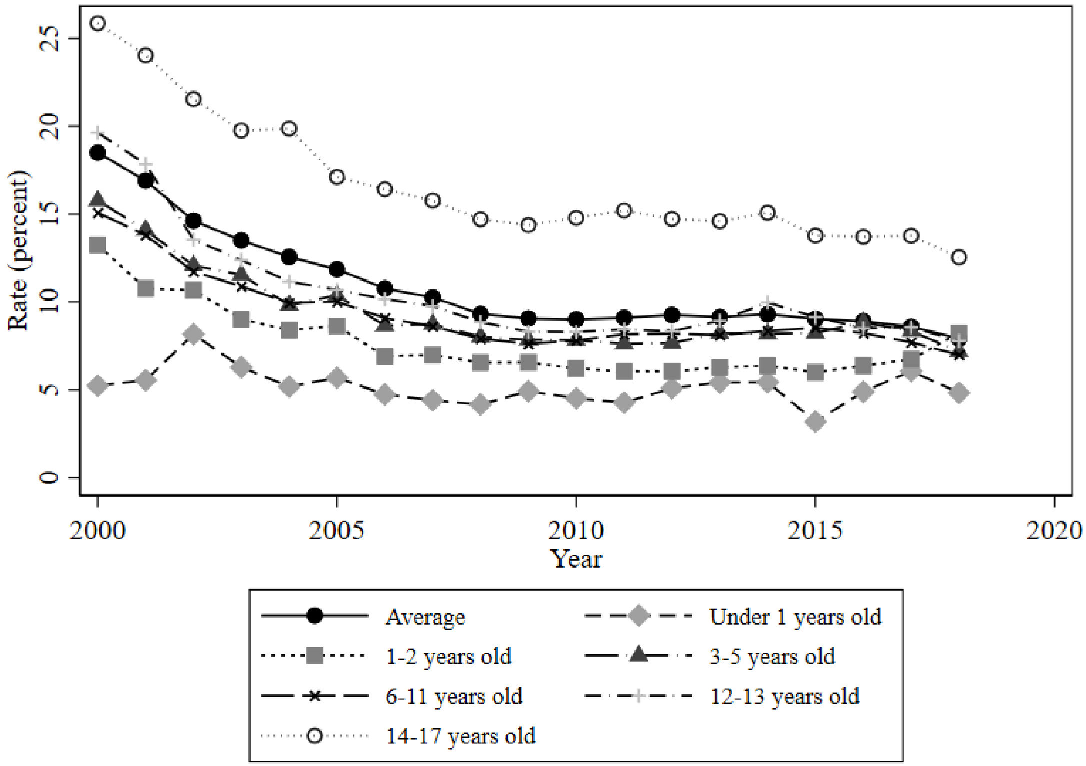
Notes. Average rates across states are shown annually. Bars indicate lower and upper bounds of the 95 percent confidence intervals.

Figure 3.5. Trends in the average 12-month reunification and foster care re-entry rates by strength or area needing improvement over year since CFSR round 1



Notes. During CFSR round 1, each state identified that re-entries were either a 'strength' or an 'area needing improvement.' Average rates are by the number of years since CFSR round 1 are shown separately: the left panel is averaged for states in which it was a 'strength' and in the right panel states in which it was an 'area needing improvement' have been averaged. Bars indicate lower and upper bounds of the 95 percent confidence intervals.

Figure 3.6. Trends in the 12-month foster care re-entry rate by age group at discharge



Note. Average state-year rates by age-group are shown.

Table 3.1. Summary of key variables

Variable	Mean	SD	Min	Max
Percent re-entering within 12 months	10.53	6.91	0.95	65.17
Percent reunified in 12 months	20.51	8.33	2.27	51.23
Percent re-entries 'area needing improvement' (ANI)				
CFSR Round 1 - re-entries ANI	54.00	49.86	0.00	100.00
CFSR Round 2 - re-entries ANI	24.00	42.73	0.00	100.00
Caseload characteristics				
Ln foster care caseload	9.10	0.96	6.84	11.92
Average months in care before discharge	19.92	6.79	7.63	70.90
Percent non-white	35.24	17.57	3.22	100.00
Percent female	47.96	2.10	35.82	52.57
Percent clinical disability	22.98	14.50	0.00	95.71
Percent physical abuse	14.46	8.65	0.00	68.32
Percent sexual abuse	5.51	4.78	0.00	48.75
Percent neglect	56.06	21.02	0.00	92.97
Percent drug abuse parent	24.57	14.41	0.00	66.48
Percent alcohol abuse parent	8.48	8.69	0.00	66.15
Percent parents in jail	7.52	5.48	0.00	66.05
Economic characteristics				
Avg log household income	11.03	0.16	10.66	11.40
Percent below poverty threshold	15.00	3.51	6.30	26.75
Unemployment rate (age 25-65)	0.05	0.02	0.02	0.12
N state-year obs.	1,000			

Note. Means are shown for 2000-2019 using a state-year panel.

Table 3.2. Estimates of the 12-month foster care re-entry and reunification rates

			One-year lagged effect	
	(1)	**(2)**	**(3)**	**(4)**
A.) 12-months re-entry rate (percent)				
CFSR Round 1	0.25	3.40**	-0.42	3.20***
	[-2.31, 2.90]	[0.77, 6.12]	[-2.51, 1.75]	[1.18, 5.24]
CFSR Round 2	-0.23	0.10	1.11	1.22
	[-1.37, 0.92]	[-1.28, 1.47]	[-1.62, 4.50]	[-1.50, 4.29]
CFSR Round 1 * re-entries need improvement		-5.86***		-6.67***
		[-8.88, -2.82]		[-9.50, -3.82]
CFSR Round 2 * re-entries need improvement		-1.04		-1.11
		[-4.09, 1.71]		[-4.10, 1.52]
Observations	1000	1000	1000	1000
Adjusted R-squared	0.711	0.729	0.712	0.744
B.) 12-months reunification rate (percent)				
CFSR Round 1	0.70	-0.09	-0.23	-0.98
	[-2.38, 3.99]	[-3.29, 3.47]	[-3.04, 2.55]	[-3.56, 1.61]
CFSR Round 2	-0.14	0.04	-1.54	-1.26
	[-1.83, 1.67]	[-1.49, 1.65]	[-3.53, 0.58]	[-3.39, 0.91]
CFSR Round 1 * re-entries need improvement		1.58		1.54
		[-2.17, 5.26]		[-2.04, 5.05]
CFSR Round 2 * re-entries need improvement		-0.86		-1.15
		[-3.47, 1.85]		[-3.70, 1.47]
Observations	1000	1000	1000	1000
Adjusted R-squared	0.722	0.723	0.723	0.724

Notes. Dependent variable in panel A is the 12-month foster care re-entry rate and in panel B is the 12-month reunification rate. Columns 3 and 4 lag the treatment effects by one year. All specifications include state and year fixed-effects and controls for caseload characteristics and economic conditions. Wild-cluster bootstrapped t-statistics are used to create confidence intervals (Cameron et al., 2008; MacKinnon at al., 2022; Roodman, 2022). 95% confidence intervals in brackets. The full set of coefficients is in appendix Table A3.3 and A3.4. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table 3.3. Estimates of the 12-month foster care re-entry rate by age-group at reunification

	(1) 0-17 years old	(2) Under 1 year old	(3) 1-2 years old	(4) 3-5 years old	(5) 6-11 years old	(6) 12-13 years old	(7) 14-17 years old
CFSR Round 1	3.40**	0.04	2.80***	2.76**	2.85**	3.07*	4.09**
	[0.77, 6.12]	[-2.91, 3.15]	[0.68, 5.04]	[0.22, 5.48]	[0.12, 5.78]	[-0.55, 6.60]	[0.42, 7.83]
CFSR Round 2	0.10	0.26	0.12	0.63	-0.08	-0.72	-0.11
	[-1.28, 1.47]	[-1.72, 2.36]	[-1.65, 1.89]	[-0.60, 1.89]	[-1.28, 1.08]	[-2.18, 0.70]	[-2.73, 2.44]
Interaction: CFSR Round 1 * re-entries need improvement	-5.86***	-1.98	-2.44	-4.34***	-4.69***	-7.79***	-7.66***
	[-8.88, -2.82]	[-5.19, 1.24]	[-5.61, 0.97]	[-7.34, -1.25]	[-7.46, -1.78]	[-12.20, -3.28]	[-11.80, -3.24]
Interaction: CFSR Round 2 * re-entries need improvement	-1.04	-1.09	-0.39	-0.10	-0.35	-0.38	-1.04
	[-4.09, 1.71]	[-4.03, 1.35]	[-2.51, 1.53]	[-2.88, 2.49]	[-2.87, 1.99]	[-4.00, 2.84]	[-5.23, 2.74]
Observations	1000	998	1000	1000	1000	1000	1000
Adjusted R-squared	0.729	0.153	0.460	0.541	0.556	0.607	0.754

Notes. Dependent variable is the 12-month foster care re-entry rate. All specifications include state and year fixed effects. CFSR rounds 1 and 2 are binary indicators for whether the state has received that round of review. Wild-cluster bootstrapped t-statistics to create confidence intervals (Cameron et al., 2008; MacKinnon et al., 2022; Roodman, 2022). 95% confidence intervals in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Chapter 4. The Decline in Family Child Care

With Elizabeth Davis

4.1 Introduction

Family child care (FCC) providers are a critically important part of the early care and education sector in the United States. Family child care providers are small, independent businesses where a provider cares for children in her home (nearly all FCC providers are women). While licensing and regulatory requirements for FCC differ across states, these home-based providers are an important source of child care services and are often more flexible and less expensive for parents than child care centers. In more than 20 states, the majority of licensed child care providers are in family homes rather than in child care centers (Haynie, 2019). Families rely especiallyon FCC providers in rural areas or in areas where access to other licensed childcare options, such as centers, is sparse (Haynie, 2019). Family child care providers are more likely than center-based care to be used by families with lower incomes and families needing care for infants and toddlers (Porter et al., 2010). Some FCC providers may also offer care at during evenings or weekends (Xue, 2017) or care for specific linguistic or cultural communities. Thus, the decline in the number of FCC providers has raised concerns about the availability and access to child care for families, particularly for parents who live in rural communities, need nonstandard care hours, or have infants and toddlers (Haynie, 2019).

In addition to supporting access and availability, FCC providers often provide a more affordable alternative to child care centers. Across all regions (Northeast, Midwest, South, and

West), average child care costs for infants and four-year-olds are lower at FCC providers than at center-based providers. For example, the average cost of infant care for a two-parent family in the Midwest is 12.5 percent of their family income at a center and 8.9 percent of their income at a FCC provider (Haynie, 2019). Even if new child care centers can make up for some of the gap in local licensed child care capacity as FCC providers close their businesses, the decrease in the number of FCC providers may reduce parents' options and the affordability of licensed child care.

Most states have experienced a substantial drop in the number of FCC providers, and this drop began long before the COVID-19 pandemic (Bromer et al. 2021b). According to Child Care Licensing Studies, between 2005 and 2017 the number of licensed small FCC providers fell nationally by 48 percent from 190,206 in 2005 to 99,614 in 2017 (Office of Child Care, 2019). Another source, the National Survey of Early Care and Education (NSECE), indicated that the number of licensed or regulated home-based providers declined 25 percent from roughly 121,000 in 2012 to 91,200 in 2019 (Datta et al., 2021). In Minnesota, the focus of this study, the number of FCC providers declined 33 percent from nearly 12,000 in January 2010 to fewer than 8,000 in December 2019.

Understanding changes in the supply of FCC providers is important because these providers directly influence local capacity (i.e., how many children can be cared for locally by licensed providers) and local average child care prices. This paper employs quantitative analysis of a panel dataset to examine factors associated with the decline in the number of FCC providers in Minnesota, complementing the growing literature of qualitative and survey analyses in other states (e.g., Dellano and Kaiser, 2017; Dwyer, 2019). Throughout this paper, we primarily observe provider 'exits' from the child care market which we interchangeably and synonymously also describe as provider business ‘closures’ and ‘closings.’ We focus on understanding the role of regulations and the economy in explaining FCC business closings prior to the COVID-19 pandemic, as the major disruptions to work and education associated with the pandemic altered the early care and education landscape in myriad ways. This study provides a better understanding of the long term trends in child care market that can help inform post-pandemic efforts to rebuild the supply of child care in local areas.

4.1.1 Factors Associated with the Decline in Number of Family Child Care Providers

Many factors may contribute to the decline in the number of FCC providers. The conceptual model developed by Bromer et al. (2021b) has three distinct sets of factors that influence providers at its core: 1) working conditions for the provider (e.g., long hours and isolation), 2) business sustainability (e.g., income and enrollment), and 3) provider experiences participating in local early childhood education systems (e.g., licensing, Quality Rating Improvement Systems, and child care subsidy programs). These core factors interact with the provider’s professional development supports and resources and the provider’s individual characteristics to contribute to a provider’s decision to stay in business or leave FCC. Consistent with the Bromer et al. (2021b) conceptual model, empirical research suggests that a number of factors contribute to the decision to close a FCC business. In surveys, providers report closing their doors for many different reasons, including regulatory burden, working conditions, and low financial returns (Bromer et al. 2021a).

Changes to licensing requirements may compound existing challenges and contribute to a provider’s decision to leave the field. While national health and safety standards for child care programs were initially developed in 1992, child care licensing regulations are determined and updated independently by each state and provide a range of standards for health and safety (Alkon et al., 2021). One oft-cited concern is that state regulations for licensed child care have been a contributing factor to the decline in the number of FCC providers. While they are often well-intended, regulations can have a major impact on the availability, price, and quality of care available in a community (Queralt and Witte, 1999).[25] At the same time, some evidence suggests that regulations are also associated with higher levels of quality in family child care (Raikes, Raikes and Wilcox, 2005).

Family child care businesses operate in a wider economic context, and changes in local economic conditions and the local child care market will influence their decisions to continue to operate. Changes in local economic conditions can affect family demand for child care as well as affect alternative employment options for the provider. When the local economy offers better employment alternatives in terms of pay and working conditions, FCC providers may be more likely to close their business to pursue these options. Evidence from Brown and Herbst (2022) shows that when unemployment rates are higher, child care workers in the public (e.g., Head Start) or the private (e.g., for-profit centers and non-profit churches) sector as well as some employed in home-based settings have lower turnover rates. Conceptual models of FCC provider turnover have cited perceived job alternatives, spouse income, and local unemployment rates as important

[25] A small number of studies have looked at the effects of regulations on licensed child care centers (e.g., Blau, 2007; Hotz and Xiao, 2011). Evidence from centers surveyed in four states in 1993 shows that the price and quality of child care did not change as a result of stricter licensing regulations (Blau, 2007). Using data from the Census of Services Industries for 1987, 1992, and 1997, Hotz and Xiao (2011) find that the imposition of regulations reduces the number of center-based child care providers, especially in lower income markets, yet increases the quality of services provided, especially in higher income areas.

moderating variables in a provider's decision to close their business (Deery-Schmitt and Todd, 1995). Competition from other providers, both public and private, may also influence closing decisions. For example, the opening of a child care center is likely to decrease demand for nearby FCC businesses. Expansion of school-based prekindergarten programs may also reduce the number of children attending home-based care. Recent studies have shown that while public prekindergarten programs may increase the overall capacity of the licensed child care market, they may distort the availability of private child care for younger children and have other consequences as well (Bassok, Fitzpatrick and Loeb, 2014; Brown 2018; Sipple, McCabe and Casto, 2020).

4.1.2 The Current Study

Understanding the determinants of provider closures and how new child care licensing regulations may contribute to the likelihood a provider leaves the profession helps inform policymakers and practitioners about promising reforms to help family providers stay in the licensed system. We specifically investigate the following research questions:

1. How has the rate of FCC providers closing in Minnesota varied over time?
2. How does exposure to new policies and licensing regulations change the rate at which FCC providers close their business?
3. What are the determinants of FCC closure?

To examine the risk of a FCC provider closing, we use data from Minnesota and compare findings from two methods: interrupted time series and survival analysis. We use a 10-year monthly panel of licensed FCC providers from the Minnesota Department of Human Services (DHS) linked with local area characteristics from the Minnesota Department of Employment and Economic

Development. Based on the prior literature, we hypothesize that three sets of factors will influence the risk of a provider closing her business: (1) the provider's own characteristics (e.g., education, race), (2) business characteristics (e.g., age of business, capacity, quality rating), and (3) external childcare factors (e.g., policy and licensing regulations, competition in the local child care market, local economic conditions). Our study focuses on two major changes in regulations that impacted licensed FCC providers in Minnesota. First, new health and safety regulations required additional trainings for FCC providers beginning in 2014. Second, expanded background check regulations for family care providers were announced in state legislation in 2017. These regulatory changes were seen as burdensome and intrusive to some in the FCC community.

This empirical study makes several important contributions to the literature. While a few studies provide evidence on how licensing regulations affected child care centers in the 1980s and 1990s (e.g., Blau, 2007; Hotz and Xiao, 2011), the impacts of regulations on FCC providers have been less well documented using quantitative methods. Due to the importance of FCC providers in the child care landscape, we seek to fill this void and discern the extent to which regulations and the local economy affect the decline in the number of FCC providers. Second, our data allow us to provide evidence for the rate and determinants of FCC provider closure using both non-parametric (Kaplan-Meier survival curves) and parametric models (discrete time survival models, and interrupted time series models).

Minnesota provides a useful case study of the decline in the supply of FCC providers. While most states have experienced declines in the number of FCC providers, data constraints in other states might pose an issue to directly assess whether these declines, and provider decisions to stay in business, may be attributed to changes in regulations in recent years or the role of the local economy. For Minnesota, we have a ten-year monthly panel of data on licensed FCC providers with detailed information on their characteristics and location in Minnesota, as well as distinct changes in regulations at two points in time. Given the high cost of child care in Minnesota (Child Care Aware of America, 2019), and high female labor force participation rate (MN DEED, 2022), the impact on families may be greater than in other states. Studying these trends in Minnesota, where the use of FCC and rate of closure differs between urban and rural parts of the state, also will provide insights into the extent to which new child care licensing regulations have differential effects across geography.

4.2 Data and Measures

4.2.1 Data Source and Sample

We constructed a 10-year monthly panel of licensed child care providers with data from the Minnesota Department of Human Services (DHS). The panel includes FCC providers with active licenses on the first day of each month from January 1, 2010, to December 1, 2019. Nearly 20,000 licensed FCC providers had active licenses between 2010 and 2019 in Minnesota, and more than half of those open since January 1, 2010, had opened prior to that date. In total, there were 1,154,795 provider-month observations and 19,605 unique licensed FCC providers.[26]

[26] While we have monthly data on active provider licenses for 2020 and later, we chose not to include those in this analysis, given that the COVID-19 pandemic may have changed the determinants and frequency of closings. In addition, during part of the pandemic, Minnesota DHS allowed providers to close temporarily without deactivating their license, so the closure rate as measured by active licenses was lower in 2021 than in pre-pandemic years.

4.2.2 Stock and Flow Samples for Duration Models

To study the probability of closure in each time period using duration analysis methods, our data are structured as provider-month observations beginning in January 2010. This format allows us to look at the length of time (in months) that each provider is in business, which we refer to as a "spell." Providers who opened their businesses prior to the start of our data (January 2010) and who were still operating their FCC business in January 2010 are considered an incoming "stock" sample. In the context of duration models, including a stock sample requires using methods that account for these providers who have spells that are left truncated (that is, that began prior to the start of the observation window). Including the stock sample is possible because the data includes the opening date for each licensed provider. The stock sample consists of 11,762 FCC providers whose average observed tenure (total spell length) in business is 15 years. Including the stock sample provides information about providers whose tenure in business is longer than the observation period of ten years. For example, one FCC provider has been operating for 33 years (opening in January of 1986 and still open at the end of 2019).

However, using only stock data would lead to over-representation of longer durations (Berger and Black, 1998). Therefore, we also include the new businesses that opened in January 2010 or later, called a "flow" or new entry-cohort sample. The flow data includes all licensed FCC providers who opened between 2010 and 2020, a total of 7,843 providers. These data capture many short tenures. Including only flow data in the study would exclude all providers in business for 10 years, an important and large segment of the market. While many duration studies use only flow (or entry-cohort samples), given the large fraction of FCC providers in business for more than 10 years, we instead use methods (described below) to account for the inclusion of both stock and flow samples.

Many of the providers in both the stock and flow samples are still in business at the end of our observation period (December 1, 2019), which means their spell length or duration is right-censored (i.e., we don't know the full length of time they will be operating). To understand the determinants of FCC closure, we, therefore, estimate survival analysis models (described below) to account for right-censoring, and using a "stock-and-flow" sample, adjusting the likelihood function to account for the providers who opened prior to 2010 based on a technique developed by Berger and Black (1998). Differences (and similarities) between the stock and flow samples are described below.

4.2.3 Measures of Exit or Closing

Using panel data on active provider licenses from the Minnesota DHS, a child care business closing or "exit" is defined as the last month when the license number appears in the data. In Minnesota, if a licensed provider stops serving children (i.e., closes), she is required to notify the licensing agency or county, and the license is no longer active. License ID numbers are unique to a provider, and do not change over time. Thus, for the main analysis, we define an exit month (equals one) as the last month that the license number appears in the active license data.

We also measure exits on an annual basis because about one-third of closings occur at a multiple of 12 months after the initial license was issued (12 months, 24 months, etc.). These exits may indicate that providers close their child care businesses when their license is due for renewal. While FCC license renewal dates may be up to 24 months after the initial license or last renewal date, training requirements are typically on an annual basis. Therefore, it seems likely that a provider who is planning to close may do so when their license needs to be renewed or when they must complete another round of training. An annual exit is defined (equals one) by the last year that the license number appears in the active license data.

4.2.4 Covariates

The probability that a FCC provider closes her business is expected to be related to internal (provider and business) factors and external factors (such as labor market conditions and regulations). Table 4.1 provides a detailed overview of the independent variables in this study including values and definitions. The provider characteristics are, for the most part, self-explanatory (education level, race/ethnicity). For business characteristics, a provider's licensed capacity is determined by the licensor based on factors such as whether they have staff (e.g., an assistant or co-provider), and the size of the physical space. Years of experience for providers is based on the tenure of the business (indicated by the business open date). Provider experience is operationalized in the survival analysis as 0-3 years of experience, 4-9 years of experience, 10-14 years of experience, and 15+ years of experience. In similar research on center-based child care, capacity and tenure were both important predictors of provider closures. Specifically, evidence showed that center-based providers run by a sole-proprietor or a non-profit, those that are older than four years, and those larger than 30 capacity have lower predicted odds of closure (Lam et al., 2013).

Other characteristics of the family child care business may influence how long it remains open. For example, we include an indicator for whether a provider has ever had their license revoked. We also include indicators for the provider's participation in government programs. We include the highest Quality Rating Improvement Systems (QRIS) rating that the provider has ever earned within the timeframe of the study: ratings range from one star to four stars, the highest rating. Since participation in Parent Aware (QRIS) ratings is optional in Minnesota, most providers remain unrated. Similarly, participation in the child care assistance program (CCAP) is also optional, and we control for an indicator for whether the provider accepts or is willing to accept CCAP (also known as child care subsidies).

Conceptual models of FCC provider turnover have cited perceived job alternatives, spouse income, and local unemployment rates as important confounding variables in a provider's decision to close their business (Deery-Schmitt and Todd, 1995). The external factors that we are able to control for in our analyses include measures of local child care market competition and local economic conditions. The capacity of nearby local child care providers (both centers and family child care) provides a measure of local competition, because these providers make up the local choice set of child care options that families may have available to them. For example, the opening of a center in the provider's zip code likely decreases demand for nearby FCC businesses. In this study, local competition is calculated for each provider by measuring the change in local child care capacity within the last 12 months by zip code. We also include a dummy variable for whether there is a school-based prekindergarten program locally available. The "Voluntary PreK Program" (VPK) in Minnesota expanded during this time period, which may have reduced the number of children attending family child care. Recent studies have shown that while public prekindergarten programs may increase the size of the licensed child care market, they impact the availability of private child care for younger children (Bassok, Fitzpatrick and Loeb, 2014; Brown 2018).

Changes in local economic conditions can affect demand for child care and on the supply side, may also affect alternative employment options for the provider. The county-month unemployment rate is included in the analyses to account for local economic conditions (MN DEED, 2021). Evidence shows that when unemployment rates are higher, child care workers across settings (public, private, and home-based) have higher levels of education and experience on average, and lower turnover rates (Brown and Herbst, 2022).

The key variables of interest indicate the timing of new licensing regulations. We consider two major changes in regulations that impacted licensed FCC providers in Minnesota. First, health and safety regulations which required additional trainings were implemented in July 2014. Second, expanded background check regulations were announced in January 2017. While the new background check requirements were not implemented immediately,[27] the announcement of these created much consternation among FCC providers.

4.3 Empirical Strategy

4.3.1 Duration Analysis: Hazard Functions and Discrete Time Survival Models

To answer the research questions concerning how long FCC businesses remain open and the factors related to business closings, we employ standard duration analysis methods (also known as "survival analysis"). Hazard-based duration models allow us to study the length of time providers stay in business, even when providers open at different points in time. In the duration

[27] Implementation Plan for 2017 Legislative Changes Impacting Licensed Family Child Care. Minnesota Department of Human Services. Updated March 1, 2018 (Minnesota Family Child Care Task Force Legislative Report 2021, 2021).

model, the closure of FCC providers (i.e., provider exit) is the outcome of interest. This outcome is right-censored for providers who have not closed by the end of our timeframe, December 2019. While our data allow us to overcome the left-truncation problem (given providers' open dates) for providers in business in January 2010, we must account for the inclusion of both a stock and flow sample. Duration analysis accounts for censoring and truncation and allows us to estimate the probability of closing by a certain month since opening.

Descriptively, Kaplan-Meier survival estimates show the probability of remaining in business by a certain month since opening. This statistic is calculated as the event (i.e., staying open), O_m, past a specific month m of operation, where T_m is the month at closure:

$$\mathrm{O_m} = \Pr(\mathrm{T_m} > m) \qquad (1)$$

We compare the survival function (the proportion remaining open) for the stock and flow samples.

Hazard functions, h_{im}, describe the conditional probability of individual FCC provider i closing at a particular month T_m for providers who have not yet closed ($T_m \geq m$):

$$\mathrm{h_{im}} = \Pr(\mathrm{T_m} | \mathrm{T_m} \geq \mathrm{m}) \qquad (2)$$

The hazard function (or hazard rate) represents the conditional probability of a provider ending a survival spell (exiting) after m periods, given that it survived for m - 1 periods (the elapsed duration of the survival spell). Therefore, the monthly hazard rate can be interpreted as the percentage of providers who closed that month given that they had survived up to that month.

Discrete time survival models allow us to analyze the outcome of interest conditional on time-invariant and time-varying covariates. We use a proportional odds logit model to investigate the determinants of FCC closure. One of the main advantages of discrete time logit model is that the baseline hazard does not need to be specified in order to estimate the relative odds that describe differences in the relative hazard between groups. This approach describes the relationship between the event of provider closure, as expressed by the hazard function, and our set of covariates. The model is specified as:

$$\frac{h(m,\ X)}{1-h(m,\ X)} = \left[\frac{\mathrm{h}_0(m)}{1-\mathrm{h}_0(m)}\right]\exp(\beta' X) \qquad (3)$$

where, $h(m,\ X)$ represents the discrete time survival function in month m for the baseline hazard rate is $\mathrm{h}_0(m,\ X)$. On the right-hand side of the equation, $h_0(m)$, is the baseline hazard and β is a vector of regression coefficients. The vector of covariates, *X*, includes controls for the provider, business, and external characteristics described above, as well as calendar year, calendar month (for seasonality), and county fixed effects.

4.3.2 Model of Annual Exits

As described below, the monthly survival function clearly shows a pattern of closings related to 12-month intervals. When a provider decides to close, we hypothesize that they may choose the timing of their closing to coincide with their next license renewal date. Therefore, we also analyze a model of annual provider exits, with each year counting 12 months from the initial license month (rather than calendar years). This approach tests whether the new regulations are associated with more exits allowing for more time for the provider to decide to close after they go into effect.

Annual provider closings are appropriately modeled as discrete time events, and therefore we estimate a complementary log-log (cloglog) model. The cloglog model is asymmetric and can be used when the event rate is very small or very large (Allison, 1995). We analyze the annual hazard of closing at a year (h_{ia}) for individual provider *i* where *a* is the specific age (or tenure) of the business. In the model, we can account for characteristics of providers, their businesses, and their environments through the vector X_{iat}. This vector includes both characteristics that are static (i.e., do not vary) and some that vary over t, calendar time, such as the local unemployment rate. These characteristics are incorporated into a complementary log-log model for the hazard of closing as:

$$h_{ia} = 1 - \exp\{-\exp[\theta(\mathrm{a}) - \beta X_{iat}\} \quad (4)$$

From this model we present hazard ratios which are exponentiated coefficients. Hazard ratios greater than one mean that a covariate increases the hazard of closing, while hazard ratios of less than one mean that a covariate decreases the hazard of closing.

4.3.3 Interrupted Time Series

While the main focus of this study is on provider exits, provider closures may not entirely drive the decline in the number of FCC providers. For instance, if the exit rate over time is relatively constant and the entry rate of new providers is declining, the observed phenomenon – a decline in the number of FCC providers over time – could occur. Appendix Figure A4.1 shows the annual family child care provider rate of exit and entry for 2010-2019. To study the entry and exit rate of FCC providers over time using interrupted time series (ITS), we calculate the statewide entry rate and the statewide exit rate over time.

Statewide, monthly exit rates are calculated as:

$$\text{Exit rate}_{\mathrm{t}} = \frac{\text{number of exiting providers}_{\mathrm{t}}}{\text{number of providers}_{\mathrm{t}}} \tag{5}$$

Statewide, monthly entry rates are calculated as:

$$\text{Entry rate}_{\mathrm{t}} = \frac{\text{number of new providers}_{\mathrm{t}}}{\text{number of providers}_{\mathrm{t-1}}} \tag{6}$$

where t is the month.

To help untangle whether regulations play a role in the decline of FCC providers as measured by changes in entry rates and exit rates, we use a standard interrupted time series (ITS) regression model (Huitema and McKean 2000; Linden and Adams 2011; Simonton 1977a; Simonton 1977b). This single-group ITS model is specified as:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \beta_4 W_t + \beta_5 W_t T_t + \varepsilon_t \tag{7}$$

where, Y_t is the outcome variable measured at each time point t. Our two outcome variables of interest are the monthly entry rates and the monthly exit rates. Then, T_t is the time since the start of the study (January 2010), X_t is a dummy (indicator) variable representing the first licensing regulation intervention, the health and safety regulations in July 2014 (preintervention period is 0, otherwise 1), and W_t represents the second licensing intervention, the background check regulations in January 2017 (preintervention period is 0, otherwise 1). Finally, X_tT_t and W_tT_t are the interaction terms.

In equation (7), β_0 represents the intercept or starting level of the outcome variable. β_1 is the slope or trajectory of the outcome variable until the introduction of the intervention. β_2 represents the change in the level of the outcome that occurs in the period immediately following the introduction of the health and safety regulations in 2014 (compared with the counterfactual, no regulatory changes). For the health and safety regulations in 2014, β_3 represents the difference between the preintervention and postintervention slopes of the outcome. β_4 represents the change in the level of the outcome that occurs in the period immediately following the introduction of the background check regulations in 2017 (compared with the counterfactual). For the background check regulations in 2017, β_5 represents the difference between preintervention and postintervention slopes of the outcome. We can rely on the p-values in β_2 to indicate an immediate treatment effect, or in β_3 to indicate a treatment effect over time (Linden and Adams 2011).

4.4 Results

4.4.1 Kaplan Meier Survival Curves: How Long Do They Stay in Business?

Figure 4.1 compares the Kaplan-Meier survival curves for the stock and flow samples of providers, showing the proportion of FCC businesses that "survive", or have not yet closed, by the length of time since opening. By one year (12 months) after opening, the proportion of providers who remain in business is 0.81 for providers in the flow sample compared to 0.82 for providers in the stock sample. The median flow provider is open for 49 months (just over four years). While the flow providers exit at a slightly faster rate than the more experienced stock providers, the gap between flow and stock providers is very small (0.01 - 0.04) and remains persistent over time. By three years (36 months), just slightly over half of family providers remain open for business: the proportion of providers who remain in business is 0.56 for flow providers compared to 0.57 of stock providers.

4.4.2 Characteristics of the stock and flow samples of providers

There are similarities as well as a few key differences between the stock providers who opened prior to 2010 and flow providers who opened between 2010 and 2020. Table 4.2 presents descriptive statistics of the independent variables by stock and flow providers. Family child care providers are distributed differently between the Twin Cities metropolitan area and greater Minnesota. As seen in the map of Minnesota in Appendix A, the Twin Cities metropolitan area is based on a seven-county definition of counties surrounding the centrally located "twin" cities themselves: Minneapolis (the largest city in Minnesota) and St. Paul (the capital city in Minnesota; Appendix Figure A4.2). While there are other metropolitan areas and cities in Minnesota, greater

Minnesota formally consists of 80 counties outside the Twin Cities metropolitan area. As seen in Table 4.2, 62 percent of stock providers are in greater Minnesota compared to 73 percent of flow providers.

Most licensed FCC providers in Minnesota have only a high school education and most providers are white: 58 percent of the stock providers and 33 percent of flow providers have a high school education; 77 percent of stock providers are white, and 46 percent of flow providers are white. The stock providers are generally licensed to provide care for more children than the flow providers: 81 percent of flow providers can provide care for 10 children. However, 36 percent of stock providers can provide care to 10 children, 41 percent can provide care to 11-13 children, and 21 percent can provide care to 14 children. A higher percentage of flow providers accept or are willing to accept subsidies from CCAP: 36 percent of stock providers accept CCAP compared to 60 percent of flow providers. Fewer than 1 percent of providers have participated in the Parent Aware Ratings. The low participation rate reflects the fact that the Parent Aware rating system is voluntary for providers and was rolled out over a multi-year period after 2012.

4.4.3 Cox Regression Results: What Factors are Related to Exits?

Table 4.3 shows the results for the Cox proportional hazard models. We include two specifications for comparison purposes. The model in Column 1 controls for the regulations and the main specification of covariates. Column 2 adds dummies to account for the Twin Cities Metro area and the interaction of the area with the two regulations. This allows for differential effects in the Twin Cities metropolitan area and in greater Minnesota (all counties outside the seven-county metropolitan area).

Looing first at Column 1, additional health and safety trainings positively are associated with a decrease in the hazard of exiting rather than an increase. In Column 2, overall the hazard of closure is lower in the Twin Cities (Minneapolis-St. Paul) metropolitan area compared to the rest of the state. The 2014 health and safety regulations are associated with a decrease in the hazard of exiting in both greater Minnesota, and in the Twin Cities metro area. Similarly, background check regulations are associated with a reduced hazard of exit for family providers. The announcement of changes to provider background checks shown in Column 1 are consistent with the theory that regulations can increase demand for FCC services by providing valuable quality assurance to consumers (families). However, based on Column 2, expanded background checks are associated with an higher hazard of exiting in the Twin Cities metro area relative to the greater Minnesota.

Beyond regulations, other changes in the market for child care and economic factors outside of the provider's control are associated with provider closings. In particular, in Column 1, the measure of increased local child care capacity (i.e., competition) is associated with a statistically significant increase in the hazard of closing across all specifications. The opening of a local public-school pre-kindergarten program is not significantly associated with a change in the hazard of closing. Moreover, local economic conditions matter, such as higher unemployment rates, are associated with lower likelihoods of FCC exit.

In addition to the important role of the economy and regulations, several provider and business characteristics were associated either provider persistence to stay in business or provider exit from the field. We find that across all specifications in Table 3, larger capacity FCC providers (11+ children) are less likely to exit than providers with a capacity for 10 children, and providers with a quality rating were also generally less likely to exit than those without a rating. Similarly, providers with more years of experience were less likely to close. Across specifications, we further find that providers who accept or are willing to accept child care subsidy (CCAP) payments are not more likely to close than those who do not.

Finally, some characteristics are associated with being more likely to exit the child care market. A provider is much more likely to exit if their license has been revoked. Moreover, relative to white providers, the hazard of closing is higher for black, Asian, American Indian, and providers whose race is unknown or missing from the Minnesota DHS licensing data.

4.4.4 Model of Annual Provider Exits

Looking back to Figure 4.1, there were noticeable annual drops in the number of providers still open at 12 months, 24 months, 36 months, and so on, indicating that there was a greater probability of a provider exiting and closing her business at those intervals. This pattern motivated the use of an annual model of provider closings.

Table 4.4 shows the results for the annual complementary log-log models. Specifications for these annual models are similar to the specifications for the monthly Cox proportional hazards models. In Column 1, both the 2014 and the 2017 we see that the regulations are associated with a decrease in the hazard of exits. Like the results from Table 3, this result is consistent with the theory that regulations regarding background checks can increase demand for FCC services by providing valuable quality assurance to consumers. This evidence based on the first two regulation coefficients is salient and consistent between Table 3 for the monthly analysis and Table 4.4 for the annual analysis. Furthermore, the 2014 health and safety regulations were associated with a pronounced decrease in the hazard of exiting in the Twin Cities metro area, and unlike the monthly models, the hazard of exit following the 2017 expanded background checks was no longer positive and significant in the Twin Cities metro area.

The annual results in Table 4.4 for local economic and market conditions as well as provider and business characteristics were also similar to the monthly Cox model presented in Table 3, with a few exceptions. Whether a provider accepts or is willing to accept CCAP is now significant in the annual model at the one to five percent level. Again, the results were mostly consistent across the two approaches using either monthly or annual models, as seen in Tables 3 and 4 respectively.

4.4.5 Interrupted Time Series Results

The interrupted time series (ITS) analysis focuses on the timing of the licensing regulations as though they are treatments or interventions. This allows us to assess whether licensing regulations affect (or don't affect) FCC provider exit and entry rates. The ITS results are visually shown in Figure 4.2, where the two dashed vertical lines represent the introduction of the two sets of regulations: the health and safety regulations of July 2014 and the expanded background check regulations that were announced in January 2017. In Figure 4.2, the entry rates of the FCC providers are shown in the left panel and the exit rates of the FCC providers are in the right panel. The first set of regulations, health and safety regulations in 2014, is not associated with changes in either the entry or the exit rates for the FCC providers. The second set of regulations, background checks, is marginally associated with a reduction in the entry rate of the providers by -0.00006 (significant at the 10 percent level) but is not associated with a statistically significant effect on the exit rate. The estimated coefficients are shown in Table 4.5. Note that the numbers are small because the monthly entry and exit rates are quite small. Overall, the ITS results do not support the hypothesis that the regulations were associated with an increase in provider closings. As shown in the Appendix (Figure A4.3 and Table B4.1), the decrease in entry rates occurred primarily in greater Minnesota, with no significant effect in the Twin Cities metropolitan area. None of the estimated coefficients related to the policies were significantly associated with changes in exit rates. Note that the ITS model controls only for the monthly state-level unemployment rate, and no other characteristics of the provider or local market and economy.

4.5 Discussion

Regulations for FCC providers vary across states but typically focus on health and safety standards, child-staff ratios, and group size. The efficacy of these regulations in improving child outcomes in terms of health and safety is difficult to assess and some researchers and policymakers have raised concerns about the unintended consequences of stricter rules (Heeb and Kilburn, 2004). Providers may find it more costly to run a FCC business or may be unwilling or unable to comply with new regulations. Previous literature has found that providers report a negative impact from regulations, but most studies use self-reported qualitative information.

This study used administrative panel data to analyze whether the timing of these expanded regulations was related to increased provider closures in Minnesota. Our results are largely consistent using different statistical methods and modeling different timing structures (monthly and annual). The ITS approach found no significant change in exit rates after the regulatory changes, although there was evidence of a slight decrease in the rate of new entry of FCC providers, especially in greater Minnesota outside the Twin Cities metropolitan area. The ITS does not control for provider characteristics, however. Monthly and annual analysis of FCC provider licensing data allows us to assess the determinants of provider closures. The determinants that we consider throughout our analysis include licensing regulations foremost, as well as the characteristics of the local economy and other local child care market factors, and provider characteristics (including business characteristics). In the preferred specifications of the monthly and annual exit models (see Tables 3 and 4), evidence from our analysis of Minnesota providers suggests that regulations do impact provider decisions to participate and persist in the child care market: however, the 2014 change in health and safety regulations for FCC and the background check changes in 2017 were associated with a decrease in the hazard of exits, and not an increase. These results suggest the importance of examining regulation changes of different types and in different geographies, while controlling for other local conditions and provider characteristics.

Overall, the results did not strongly support our hypothesis that FCC closings would be related to the expansion of public-school pre-kindergarten; public VPK was not significantly associated with the likelihood of exiting in the annual specification or Column 2 of the monthly specification. In contrast to other states, Minnesota has expanded public school pre-kindergarten slowly, which may explain the contrast in results from places that saw impacts on FCC due to large rapid expansions of public prekindergarten (Bassok, Fitzpatrick and Loeb, 2014; Brown 2018; Sipple, McCabe and Casto, 2020).

Local economic conditions also influence the rate at which FCC providers close. In both the monthly and annual models, higher local unemployment was associated with a decrease in closings (Deery-Schmitt and Todd, 1995). Thus, when the local economy is stronger (i.e., there is lower unemployment), more FCC providers may close their business to seek other employment opportunities. This phenomenon could make access to child care more difficult for families, as the supply of local providers could decrease even though demand for child care may be increasing (due to local economic activity and more parents working).

The estimates for the other covariates mostly supported our hypotheses and were in concordance with prior studies. Family child care providers with more education and more experience were generally more likely to close, perhaps reflecting greater labor market alternatives. In both models, black, Asian, and American Indian, providers were more likely to close. Providers who accept child care subsidies were less likely to close in the annual model. Participation in the state's Quality Rating and Improvement System (QRIS), called Parent Aware, was associated with a lower likelihood of closing, regardless of rating level. Being rated may be more desirable for providers planning to stay in business longer, and the resources provided by the QRIS may help them to do so. In addition, being rated might support greater demand by parents or higher prices.

These findings, using a 10-year panel of administrative data and multiple estimation methods, provide support for qualitative studies which conclude that the factors leading to child care closings are multi-faceted. The administrative data and multivariate approach allow us to control for a number of factors simultaneously for the statewide population of FCC providers over ten years. Isolating the effects of regulatory changes is difficult, given that there may be concurrent changes in the local economy or other early care and education policies such as public pre-kindergarten expansion. Regulations and paperwork may be a salient burden to FCC providers, yet their business decisions will also reflect local economic conditions and personal circumstances.

4.5.1 Limitations

This study is the first we know of to use both survival analysis with a stock-and-flow sample of providers and interrupted time series to study FCC business closings. The fact that regulations were implemented at the same time across the state of Minnesota makes it difficult to develop a research design that would support causal interpretation. Therefore, our results should be interpreted as associations, that is, we provide evidence on whether regulations and other factors were associated with an increase in exits (or not). We have data on a number of business and personal characteristics but are lacking information on the age of the provider and of her children, if any. The age of the business is correlated with the provider's age, though, so may reflect the effect of retirements.

The timing of regulatory changes and of provider responses also present challenges to estimating the effect of a change on provider closings. A provider who has decided to close likely plans to do so a few months (or more) in the future. Thus, the annual exit model may provide a better approach than the monthly exit model. Based on other research, the factors affecting exits often interact with each other, and a provider may decide to close when these factors together reach a "tipping point" (Bromer et al., 2021a). Despite these issues, this study using administrative data complements the findings of qualitative research on providers' reasons for closing, which are largely based on self-reported reasons. Providers may blame regulations in some cases even though other factors, such as low enrollment or opportunities for better pay or benefits elsewhere also strongly influence their decisions.

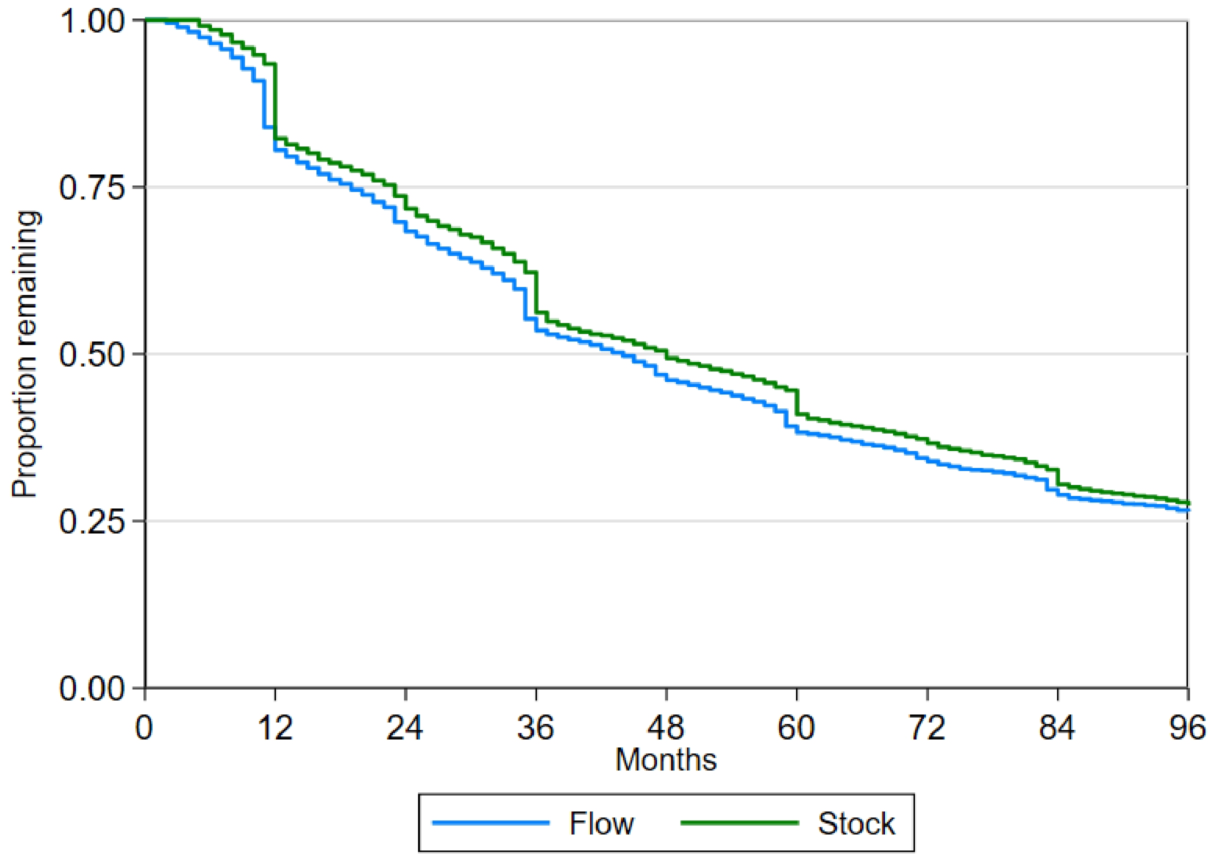
4.5.2 Conclusion

The decline in the number of FCC providers in Minnesota and nationally has raised concerns about the availability and affordability of child care, especially for families in rural areas, those who need more affordable care options, those who have infants or toddlers, or those who work non-standard hours. This study examined a number of potential factors that could influence the decision to close a home-based child care business, focusing on the role of regulations and local economic conditions. These findings may help policymakers and practitioners better understand the determinants of provider closures and the potential role of child care licensing regulations.

We find suggestive evidence that providers overall were less likely to exit after new regulations went into effect in Minnesota in 2014 corresponding to increased health and safety trainings, and similarly less likely to exit in 2017 after the expanded background check regulations. For example, looking at the monthly Cox proportional hazard models, we find that statewide health and safety regulations are associated with a statistically significant decrease in the hazard of provider exits (Table 3) An augmented specification with an interaction term on the regulation by region, allows us to see that there is some evidence of differential effects in greater Minnesota (outside of the seven counties in the Twin Cities metropolitan area), where the likelihood of monthly closures decreased slightly after the 2014 regulatory changes in greater Minnesota (Table 3). Additionally, the expanded background check regulations are associated with a reduced hazard of exit for providers. However, other factors, including local economic conditions and provider business characteristics are also associated with the likelihood of closing. The ITS analysis allows us to further consider that the regulations may also have impacted entries or business openings of licensed FCC providers, contributing to the overall decline in the number of FCC providers in Minnesota. Regulations can provide assurance of standards to families looking for care, potentially increasing demand.

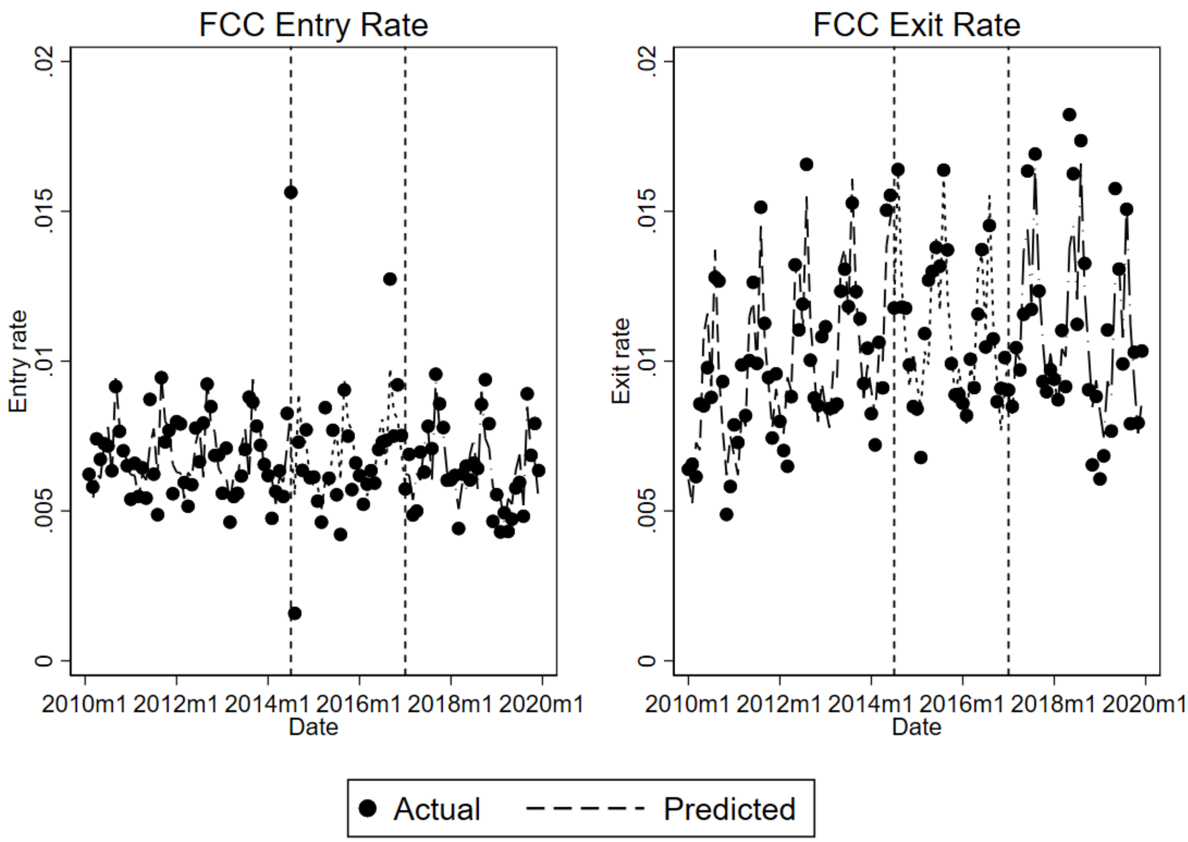
Qualitative research describing provider decisions to close reveals a complex and interrelated set of factors reported by providers (Bromer et al., 2021a). Identifying a single factor, or the most relevant factor, is challenging. While the role of regulatory burden should not be overstated given the importance of health and safety regulations, quantitative analyses such as this can help illuminate an understanding of the accumulation of factors that lead to FCC closings. Two avenues for future research are suggested: first, for quantitative analysis, causal identification strategies are needed, and second, for qualitative analysis, a focus on understanding the reasons why providers stay in business (despite the challenges of low returns, paperwork, and long hours). Future research should also investigate whether the rate and reasons for closures have changed since the COVID-19 pandemic. Family child care providers are an important segment of the child care market in the U.S., particularly for infants and toddlers and families in rural areas. Given the public investment in training and licensing costs for family child care providers, sustaining these businesses helps to support rural economies as well as family options.

Figure 4.1. The proportion of FCC providers who remained each month by stock and flow



Note. Stock providers opened prior to 2010 and flow providers opened between 2010 and 2020.

Figure 4.2. Interrupted time series of the FCC entry rate and exit rate



Notes. Minnesota statewide FCC entry and exit rates are shown. There are two sets of regulations represented by dashed vertical lines: the health and safety regulations of July 2014 and the background check regulations of January 2017. Controls for monthly state-level seasonally adjusted unemployment. Prais-Winsten regression coefficients are seen in Table 4.5.

Table 4.1. Definitions of covariates

Domain	Measure	Values	Definition
Provider characteristics	Education	High school degree 2-year college 4-year college Graduate degree	Highest education attainment is treated as time-invariant by using the record at the time of opening their FCC business.
	Race	White Black Asian American Indian Hispanic Multiracial	Race categories are mutually exclusive. Race is treated as time-invariant by using the record at the time of opening their business. This variable is time invariant.
Business characteristics	Capacity	9 or fewer children 10 children 11 - 13 children 14 children	Licensed capacity reflects the number of children the provider is allowed to provide care for at any given time. This variable is time invariant.
	Experience	0-3 years 4-9 years 10-14 years 15+ years	Experience or tenure (in years) is based on the provider's open year and the current year. Experience is time-varying and operationalized in 4 categories.
	Quality ratings	Unrated One star Two stars Three stars Four stars (highest)	Quality ratings reflect the highest ever achieved rating. Thus, this variable is time invariant. These ratings come from Parent Aware, Minnesota's Quality Rating and Improvement System.
	Revocation	1/0 integer that equals 1 post-receipt	Revocation of license is time varying. Revocation equals 0 until the provider receives a revocation, then when it 'turns on' it stays on (=1).
	Accepts CCAP	Accepts or is willing to accept CCAP Does not accept CCAP or declined to answer/missing response	A binary indicator for whether the provider ever accepts child care assistance program or that the provider never accepts CCAP, declined to answer, is missing a response. This variable is time invariant.

	Region	Twin Cities metropolitan area Greater Minnesota	The Twin Cities metro area is defined by the seven counties around Minneapolis and St. Paul. Greater Minnesota refers to the 80 Minnesota counties outside of the Twin Cities metropolitan area. See map in Appendix A.
External factors	Change in local child care capacity	Integer	A measure of accumulated change in new licensed capacity of providers within each zip code within the last 12 months. This variable is time varying.
	Voluntary pre-kindergarten	1/0 indicator that equals 1 post-implementation	This measure tells us whether voluntary pre-kindergarten (VPK) is offered within the zip code that year. VPK equals 0 until the provider's zip code offers VPK. Then when it 'turns on' it stays on (=1). This variable is time varying.
	Unemployment	Rate	County-level unemployment rate. This variable is time varying.
	Health and safety regulations of 2014	1/0 indicator that equals 1 post-implementation	Health and safety regulations require new training. When it 'turns on' it stays on (=1). This variable is time varying.
	Background check regulations of 2017	1/0 indicator that equals 1 post-implementation	Background check regulations require more stringent procedures for the provider and her household members. When it 'turns on' it stays on (=1). This variable is time varying.

Notes. This table of covariates refers to the definitions of variable construction for the monthly analysis. The covariate definitions for the annual analysis are the same for all time-invariant variables. For time-varying variables, there is one slight difference from the above descriptions such that the unemployment rate is the annual county-level unemployment rate collected from MN DEED rather than the monthly county-level unemployment rate. Otherwise, time-varying definitions for experience, revocation, changes in local child care capacity, and regulations are defined the same for the annual models as the monthly models.

Table 4.2. Descriptive provider statistics

	Stock		Flow		Difference		
	Mean	**SD**	**Mean**	**SD**	**Difference**		**T-statistic**
Region							
Twin Cities metro area	0.618	0.486	0.727	0.445	0.109	***	(16.255)
Greater Minnesota	0.382	0.486	0.273	0.445	-0.109	***	(16.255)
Education							
High school	0.582	0.493	0.327	0.469	-0.255	***	(36.553)
College 2-year	0.049	0.216	0.046	0.209	-0.003		(1.078)
College 4-year	0.190	0.393	0.155	0.362	-0.035	***	(6.502)
Graduate degree	0.009	0.096	0.014	0.117	0.004	**	(2.783)
Unknown or missing	0.170	0.375	0.459	0.498	0.290	***	(43.829)
Race							
White	0.771	0.420	0.461	0.499	-0.309	***	(45.249)
Asian	0.012	0.110	0.009	0.093	-0.004	*	(2.505)
Black	0.020	0.142	0.018	0.132	-0.003		(1.396)
Hispanic	0.008	0.089	0.008	0.090	0.000		(0.194)
American Indian	0.003	0.051	0.002	0.039	-0.001		(1.709)
Multiracial	0.006	0.076	0.009	0.092	0.003	*	(2.132)
Unknown or missing	0.180	0.384	0.494	0.500	0.314	***	(47.089)
Licensed capacity							
9 or fewer children	0.010	0.100	0.018	0.134	0.008	***	(4.645)
10 children	0.362	0.481	0.810	0.392	0.448	***	(71.561)
11-13 children	0.412	0.492	0.125	0.330	-0.287	***	(48.876)
14 children	0.216	0.412	0.047	0.211	-0.169	***	(37.778)
Experience							
Years of experience (max. for provider)	15.192	7.834	2.991	2.384	-12.2	***	(158.271)
CCAP							
Accepts or is willing to accept CCAP	0.356	0.479	0.597	0.491	0.241	***	(33.965)
Does not accept CCAP or declined to answer	0.644	0.479	0.403	0.491	-0.241	***	(33.965)
Parent Aware Rating (highest)							
Unrated	1.000	0.000	0.998	0.044	-0.002	***	(3.876)
One Star	0.000	0.000	0.001	0.023	0.001	*	(2.000)
Two Stars	0.000	0.000	0.001	0.025	0.001	*	(2.237)
Three Stars	0.000	0.000	0.000	0.011	0.000		(1.000)

Four Stars	0.000	0.000	0.001	0.025	0.001	*	(2.237)
Revocation	0.001	0.033	0.000	0.000	-0.001	***	(3.607)
Accumulated change in local capacity in last 12 months	0.010	0.649	9.786	20.748	9.776	***	(41.715)
VPK	0.000	0.000	0.054	0.226	0.054	***	(21.17)
Unemployment rate	8.655	1.675	4.854	1.890	-3.802	***	(144.311)
N	11,762		7,843		19,605		

Notes. Descriptive provider statistics are separately shown for stock and flow providers. Stock providers opened prior to 2010 and flow providers opened between 2010 and 2020. The difference in each characteristic between stock and flow providers is t-tested such that the statistical significance of the difference is reported as *p < 0.10. **p < 0.05. ***p < 0.01.

Table 4.3. Cox proportional hazard model predicting monthly family child care closures (hazard ratios)

	(1)	(2)
Licensing regulations		
Health and safety regulations (July 2014+)	0.809***	0.831***
	(0.025)	(0.029)
Background check regulations (2017+)	0.769***	0.738***
	(0.023)	(0.025)
Twin Cities Metro Area		0.597***
		(0.083)
Health and safety regulations (July 2014+) # Twin Cities Metro Area		0.922*
		(0.044)
Background check regulations (2017+) # Twin Cities Metro Area		1.145**
		(0.062)
Accumulated change in local capacity since the last renewal	1.00**	1.00
	(0.00)	(0.00)
Voluntary pre-kindergarten	1.076*	1.065
	(0.045)	(0.045)
Unemployment Rate	0.810***	0.810***
	(0.008)	(0.008)
Education (College 2-year omit.)		
High school	1.004	1.004
	(0.048)	(0.048)
College 4-year	1.144***	1.144***
	(0.058)	(0.058)
Graduate degree	1.598***	1.598***
	(0.154)	(0.154)
Unknown or missing	2.104***	2.104***
	(0.122)	(0.122)
Race (White omit.)		
Asian	1.493***	1.494***
	(0.129)	(0.129)
Black	1.648***	1.650***
	(0.110)	(0.110)
Hispanic	1.073	1.075
	(0.122)	(0.122)
American Indian	1.375*	1.375*

	(0.242)	(0.242)
Multiracial	1.021	1.023
	(0.131)	(0.131)
Unknown or missing	1.742***	1.745***
	(0.065)	(0.065)
Capacity (10 children omit.)		
<10 children	1.136	1.134
	(0.092)	(0.092)
11-13 children	0.680***	0.680***
	(0.015)	(0.015)
14 children	0.594***	0.593***
	(0.017)	(0.017)
Childcare assistance program participation (Does not accept CCAP/decline to answer/missing omit.)		
Accepts or is willing to accept CCAP	1.007	1.006
	(0.024)	(0.024)
Parent aware rating (unrated omit.)		
One Star	0.462***	0.462***
	(0.037)	(0.037)
Two Stars	0.343***	0.343***
	(0.027)	(0.027)
Three Stars	0.245***	0.245***
	(0.027)	(0.027)
Four Stars	0.289***	0.289***
	(0.028)	(0.028)
Ever received a revocation	8.965***	8.971***
	(0.342)	(0.342)
County fixed effects	Yes	Yes
Month fixed effects	Yes	Yes
N obs.	1,152,376	1,152,376

Notes. Coefficients are hazard ratios. Models also include license type, county and month fixed effects. Standard errors in parentheses. *p < 0.10. **p < 0.05. ***p < 0.01.

Table 4.4. Complementary log-log model predicting annual family child care closures (hazard ratios)

	(1)	(2)
Licensing regulations		
Health and safety regulations (July 2014+)	0.84***	0.88***
	(0.03)	(0.03)
Background check regulations (2017+)	0.82***	0.75***
	(0.03)	(0.03)
Twin Cities Metro Area		0.00***
		(0.00)
Health and safety regulations (July 2014+) # Twin Cities Metro Area		0.90**
		(0.04)
Background check regulations (2017+) # Twin Cities Metro Area		1.09
		(0.06)
Accumulated change in local capacity since the last renewal	1.00	1.00
	(0.00)	(0.00)
Voluntary pre-kindergarten	1.07	1.05
	(0.05)	(0.05)
Unemployment Rate	0.83***	0.93***
	(0.01)	(0.01)
Education (College 2-year omit.)		
High school	0.91*	1.10*
	(0.05)	(0.06)
College 4-year	0.96	1.19***
	(0.05)	(0.07)
Graduate degree	1.31***	1.62***
	(0.13)	(0.16)
Unknown or missing	1.88***	2.36***
	(0.12)	(0.15)
Race (White omit.)		
Asian	1.33***	1.32***
	(0.12)	(0.12)
Black	1.36***	1.42***
	(0.11)	(0.12)
Hispanic	1.01	1.02
	(0.12)	(0.12)
American Indian	1.44**	1.39*

	(0.24)	(0.25)
Multiracial	0.94	0.94
	(0.13)	(0.13)
Unknown or missing	1.71***	1.78***
	(0.07)	(0.07)
Capacity (10 children omit.)		
<10 children	1.11	1.13
	(0.10)	(0.10)
11-13 children	0.69***	0.69***
	(0.02)	(0.02)
14 children	0.57***	0.58***
	(0.02)	(0.02)
Childcare assistance program participation (Does not accept CCAP/decline to answer/missing omit.)		
Accepts or is willing to accept CCAP	0.92***	0.95**
	(0.02)	(0.02)
Parent aware rating (unrated omit.)		
One Star	0.75***	0.74***
	(0.08)	(0.08)
Two Stars	0.59***	0.59***
	(0.06)	(0.06)
Three Stars	0.56***	0.56***
	(0.08)	(0.08)
Four Stars	0.60***	0.61***
	(0.08)	(0.08)
Ever received a revocation	10.07***	9.90***
	(0.91)	(0.91)
County fixed effects	Yes	Yes
Month fixed effects	Yes	Yes
N obs.	1,152,339	1,152,339

Notes. Coefficients are hazard ratios. Models also include license type, county, and month fixed effects. Standard errors in parentheses. *p < 0.10. **p < 0.05. ***p < 0.01.

Table 4.5. Interrupted time series models predicting entry and exit rates

	Entry rates	**Exit rates**
Intercept	-0.0000246	-0.0000650
	(0.0000598)	(0.0000681)
Time(β1)	-0.0006195	-0.0006972
	(0.000616)	(0.000698)
Intervention one (β2) (Change in level)	0.0000615	0.0000207
	(0.0000665)	(0.0000755)
Time After Int. (β3) (Change in slope)	-0.0004968	0.0007177
	(0.0007276)	(0.0008238)
Intervention two (β4) (Change in level)	-0.0000762	-0.0000292
	(0.0000353)	(0.0000405)
Time After Int. (β5) (Change in slope)	-0.0000246*	-0.0000650
	(0.0000598)	(0.0000681)
N	120	120
R-squared	0.4712	0.7973
Durbin–Watson statistic	2.19145	1.90941

Notes. Prais-Winsten regression controls for monthly state-level seasonally adjusted unemployment. Standard errors in parentheses. *$p < 0.10$. **$p < 0.05$. ***$p < 0.01$.

Bibliography

Akin, Becci A. (2011). Predictors of Foster Care Exits to Permanency: A Competing Risks Analysis of Reunification, Guardianship, and Adoption. Children and Youth Services Review, 33(6), 999-1011.

Alaska Citizen Review Panel. (2015). *Alaska Citizen Review Panel Annual Report 2015*. www.crpalaska.org

Alkon, A., Rose, R., Hazard, K., & Moser, D. (2021). National Health and Safety Standards: Family Child Care Homes Compared With Child Care Centers. Journal of Pediatric Health Care, 35(1), 5-15.

Allison, Paul D. (1995). Survival Analysis Using the SAS System: A Practical Guide. Cary, NC: SAS Institute.

Antle, B. F., Barbee, A. P., Christensen, D. N., & Sullivan, D. J. (2009). The Prevention of Child Maltreatment Recidivism through the Solution-Based Casework Model of Child Welfare Practice. *Children and Youth Services Review*, *31*(12), 1346–1351.

Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How Much Should We Trust Staggered Difference-in-Differences Estimates? *Journal of Financial Economics*, *144*(2), 370–395.

Baron, J., Doyle Jr, J., Emanuel, N., Hull, P., & Ryan, J. P. (2023). *Racial Discrimination in Child Protection* (NBER Working Paper).

Bassok, D., Fitzpatrick, M., & Loeb, S. (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. Journal of Urban Economics, 83, 18-33.

Berger, M. C., & Black, D. A. (1998). The duration of Medicaid spells: An analysis using flow and stock samples. Review of Economics and Statistics, 80(4), 667-675.

Berger, L. M., Font, S. A., Slack K. S., & Waldfogel, J. (2017). Income and Child Maltreatment in Unmarried Families: Evidence from the Earned Income Tax Credit. Review of Economics of the Household, 1345-72.

Berger, L. M., & Waldfogel, J. (2004). Out-of-Home Placement of Children and Economic Factors: An Empirical Analysis*. *Review of Economics of the Household*, *2*(4), 387–411.

Bertanha, M., McCallum, A. H., Payne, A., & Seegert, N. (2022). Bunching Estimation of Elasticities using Stata. The Stata Journal, 22(3), 597-624.

Blau, D. M. (2007). Unintended consequences of child care regulations. Labour Economics, 14(3), 513-538.

Borusyak, K., Jaravel, X., & Spiess, J. (2023). *Revisiting Event Study Designs: Robust and Efficient Estimation*.

Bromer, J., Porter, T., Melvin, S., & Ragonese-Barnes, M. (2021a). Family Child Care Educators' Perspectives on Leaving, Staying, and Entering the Field: Findings from the Multi-State Study of Family Child Care Decline and Supply. Herr Research Center, Erikson Institute.

Bromer, J., Melvin, S., Porter, T., & Ragonese-Barnes, M. (2021b). The shifting supply of regulated family child care in the U.S.: A literature review and conceptual model. Chicago, IL: Herr Research Center, Erikson Institute.

Brown, D., & De Cao, E. (2018). The Impact of Unemployment on Child Maltreatment in the United States. 2018-04. ISER Working Paper Series.

Brown, J. H. (2018). Does public pre-k have unintended consequences on the child care market for infants and toddlers?. Princeton University Industrial Relations Section Working Paper, 626.

Brown, J. H., & Herbst, C. M. (2022). Childcare over the Business Cycle. Journal of Labor Economics, 40(S1), S429-S468.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-In-Differences with Multiple Time Periods. Journal of Econometrics, 225(2), 200-230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, *90*(3), 414–427.

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372.

Cheng, T. C. (2010). Factors Associated with Reunification: A Longitudinal Analysis of Long-Term Foster Care. Children and Youth Services Review 32 (10): 1311-16.

Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. The Quarterly Journal of Economics, 126(2), 749-804.

Child Care Aware of America. (2019). The US and the High Price of Child Care: An Examination of a Broken System. Child Care Aware of America.

Child Welfare Information Gateway. (2003). *Decision-Making in Unsubstantiated Child Protective Services Cases*. www.childwelfare.gov

Child Welfare Information Gateway. (2014). *Differential Response to Reports of Child Abuse and Neglect*. www.childwelfare.gov

Children's Bureau. (2019). National Child Abuse and Neglect Data System (NCANDS) Child File, FFY 2004-2019 [Dataset]. In *National Data Archive on Child Abuse and Neglect*.

Choi, M. J., Kim, J., Roper, A., LaBrenz, C. A., & Boyd, R. (2021). Racial Disparities in Assignment to Alternative Response. *Children and Youth Services Review*, *125*.

Congress, U.S. (1997). The Adoption and Safe Families Act of 1997.

Cuomo, A. M., & Carrión, G. (2011). *Differential Response in Child Protective Services in New York State Implementation, Initial Outcomes and Impacts of Pilot Project Report to the Governor and Legislature*.

Datta, A. R., Milesi, C., Srivastava, S., & Zapata-Gietl, C. (2021). NSECE Chartbook—Homebased Early Care and Education Providers in 2012 and 2019: Counts and Characteristics (NSECE Chartbook OPRE Report #2021-85; pp. 1–40). Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

David Scharfenberg. (2015, November 17). DCF Ends its 2-Tier Child-Welfare Monitoring Process. *The Boston Globe*.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. American Economic Review, 110(9), 2964-2996.

De Chaisemartin, C., & D'haultfoeuille, X. (2023). Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey. *Econometrics Journal*, *26*(1), C1–C30. https://doi.org/10.1093/ectj/utac017ï

Deery-Schmitt, D. M., & Todd, C. M. (1995). A conceptual model for studying turnover among family child care providers. Early Childhood Research Quarterly, 10(1), 121-143.

Dellano, D., & Kaiser, J. (2017). Family Child Care In New Jersey: Challenges and Opportunities. *Advocates for Children of New Jersey*.

Doyle Jr., J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. The American Economic Review, 97(5), 1583-1610.

———. (2008). Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care. Journal of Political Economy, 116(4), 746-70.

Dwyer, T. (2019). The Decline of Family Child Care in Maine. M.S.Ed. in Early Childhood Research Projects. 43.

Fluke, J., Jones Harden, B., Jenkins, M., & Ruehrdanz, A. (2010). *Research Synthesis on Child Welfare Disproportionality and Disparities*.

Ginther, D. K., & Johnson-Motoyama, M. (2017). "Do State TANF Policies Affect Child Abuse and Neglect? Working Paper.

Goodman-Bacon, A. (2021). "Difference-in-Differences with Variation in Treatment Timing." Journal of Econometrics, 225(2), 254-277.

Goodman-Bacon, A., Goldring, T., & Nichols, A. (2022). bacondecomp: Stata Module to Perform a Bacon Decomposition of Difference-in-Differences Estimation. Stata Comand.

Grimon, Marie-Pascale. (2020). Effects of the Child Protection System on Parents. Dissertation.

Gross, M., & Baron, E. J. (2022). Temporary Stays and Persistent Gains: The Causal Effects of Foster Care. American Economic Journal: Applied Economics, 14(2), 170-199.

Haynie, K, (2019). Checking in on the educational landscape: 2019 state fact sheet. Child Care Aware of America.

Heckman, J. J., & Honoré, B. E. (1990). The Empirical Content of the Roy Model. *Econometrica: Journal of the Econometric Society*, *58*(5), 1121–1149.

Heeb, R., & Kilburn, R. (2004). The effects of state regulations on childcare prices and choices. Working Paper. Labor and Population Program. Santa Monica, CA: RAND.

Hotz, V. J., & Xiao, M. (2011). The impact of regulations on the supply and quality of care in child care markets. American Economic Review, 101(5), 1775-1805.

Horwitz, S. M., Hurlburt, M. S., Goldhaber-Fiebert, J. D., Palinkas, L. A., Rolls-Reutz, J., Zhang, J., Fisher, E., & Landsverk, J. (2014). Exploration and Adoption of Evidence-Based Practice by US Child Welfare Agencies. Child Youth Service Review, 39, 147-52.

Huitema, B. E., & Mckean, J. W. (2000). Design specification issues in time-series intervention models. Educational and Psychological Measurement, 60(1), 38-58.

Hussey, J. M., Chang, J. J., & Kotch, J. B. (2006). Child Maltreatment in the United States: Prevalence, Risk Factors, and Adolescent Health Consequences. *Pediatrics*, *118*(3), 933–942.

Jedwab, M., & Shaw, T. V. (2017). Predictors of Reentry into the Foster Care System: Comparison of Children with and without Previous Removal Experience. Children and Youth Services Review, 82, 177-84.

Johnson-Motoyama, M., Ginther, D. K., Phillips, R., Beer, O. W. J., Merkel-Holguin, L., & Fluke, J. (2023). Differential Response and the Reduction of Child Maltreatment and Foster Care Services Utilization in the U.S. From 2004 to 2017. *Child Maltreatment*, *28*(1), 152–162.

Jones, Annette Semanchin. (2015) "Implementation of Differential Response: A Racial Equity Analysis." Child Abuse & Neglect: 73-85.

Kaplan, C., & Rohm, A. (2010). *Ohio Alternative Response Pilot Project Final Report of the AIM Team*.

Kimberlin, S. E., Anthony, E. K. & Austin, M. J. (2009). Re-Entering Foster Care: Trends, Evidence, and Implications. Children and Youth Services Review, 31(4), 471-81.

Kleven, H. J., & Waseem, M. (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence From Pakistan. The Quarterly Journal of Economics, 128(2), 669-723.

Koh, E., & Testa, M. F. (2011). Children Discharged from Kin and Non-Kin Foster Homes: Do the Risks of Foster Care Re-Entry Differ? Children and Youth Services Review, 33(9), 1497-1505.

Lam, M., Klein, S., Freisthler, B., & Weiss, R. E. (2013). Child center closures: Does nonprofit status provide a comparative advantage? Children and Youth Services Review, 35(3), 525-534.

Lemon, K., D'Andrade, A., & Austin, M. (2008). Understanding and Addressing Racial/Ethnic Disproportionality in the Front End of the Child Welfare System. *Journal of Evidence-Based Social Work*, *5*(1–2), 9–30.

Linden, A., & Adams, J. L. (2011). Applying a propensity score-based weighting model to interrupted time series data: improving causal inference in programme evaluation. Journal of evaluation in clinical practice, 17(6), 1231-1238.

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Cluster-Robust Inference: A Guide to Empirical Practice. *Journal of Econometrics*, *232*(2), 272–299.

Mackinnon, J. G., & Webb, M. D. (2017). Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, *32*(2), 233–254.

Mackinnon, J. G., & Webb, M. D. (2018). The Wild Bootstrap for Few (Treated) Clusters. *The Econometric Journal*, *21*(2), 114–135.

Martin, M., & Connelly, D. D. (2015). *Achieving Racial Equity: Child Welfare Policy Strategies to Improve Outcomes for Children of Color*.

McDonald, T., Bryson, S., & Poertner, J. (2006). Balancing Reunification and Reentry Goals. Children and Youth Services Review, 28(1), 47-58.

Minnesota Department of Employment and Economic Development (MN DEED). (2021). Monthly unemployment rates by Minnesota county. [Data set].

Minnesota Department of Employment and Economic Development (MN DEED). (2022). Labor Force Participation Rates.

Minnesota Family Child Care Task Force Legislative Report 2021. (2021).https://www.lrl.mn.gov/docs/2021/mandated/210195.pdf

National Data Archive on Child Abuse and Neglect. (2019). AFCARS Foster Care Annual File Code Book. Ithaca, New York.

National Quality Improvement Center. (2011). *Differential Response in Child Protective Services: A Literature Review*.

Parolini, A., Shlonsky, A., Magruder, J., Eastman, A. L., Wulczyn, F., & Webster, D. (2018). Age and Other Risk Factors Related to Reentry to Care from Kin Guardian Homes. Child Abuse & Neglect, 79, 315-324.

Paxson, C., & Waldfogel, J. (1999). "Parental Resources and Child Abuse and Neglect." The American Economic Review, 89(2), 239-44.

———. (2003). Welfare Reforms, Family Resources, and Child Maltreatment. Journal of Policy Analysis and Management, 22(1), 85-113.

Piper, K. A., Vandervort, F., Schunk, S., Kelly, C., & Holzrichter, J. (2019). *Issues in Differential Response: Revisited*. http://www.centerforchildpolicy.org

Porter, T., Paulsell, D., Del Grosso, P., Avellar, S., Hass, R., & Vuong, L. (2010). A review of the literature on home-based child care: Implications for future directions. Princeton, NJ: Mathematica Policy Research.

Putnam-Hornstein, E., Prindle, J., & Hammond, I. (2021). Engaging Families in Voluntary Prevention Services to Reduce Future Child Abuse and Neglect: A Randomized Controlled Trial. *Prevention Science*, *22*(7), 856–865.

Office of Child Care. (2019). The Decreasing Number of Family Child Care Providers in the United States. Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Queralt, M., & Witte, A. D. (1999). Childcare regulations: A method to pursue social welfare goals? Children and Youth Services Review, 2(21), 111-146.

Raikes, H. A., Raikes, H. H., & Wilcox, B. (2005). Regulation, subsidy receipt and provider characteristics: What predicts quality in child care homes? Early Childhood Research Quarterly, 20(2), 164-184.

Raissian, K. M., & Bullinger, L. R. (2017). Money Matters: Does the Minimum Wage Affect Child Maltreatment Rates?. Children and Youth Services Review, 72, 60-70.

Rittenhouse, K. (2023). *Income and Child Maltreatment: Evidence from a Discontinuity in Tax Benefits* (Dissertation).

Roodman, D. (2022). boottest: Stata Module To Provide Fast Execution Of The Wild Bootstrap With Null Imposed. Stata Command.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and Wild: Bootstrap Inference In Stata Using boottest. The Stata Journal, 19(1), 4-60.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., & Sobek, M. (2019). Integrated Public Use Microdata Series USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS.

Sant'Anna, P. H. C., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122.

Shaw, T. V. (2006). Reentry into the Foster Care System after Reunification. Children and Youth Services Review, 28(11), 1375-90.

Shipe, S. L., Uretsky, M. C., LaBrenz, C. A., Shdaimah, C. S., & Connell, C. M. (2022). When Families, Organizational Culture, and Policy Collide: A Mixed Method Study of Alternative Response. *Children and Youth Services Review*, *139*.

Simonton, D. K. (1977a). Cross-sectional time-series experiments: some suggested statistical analyses. Psychological Bulletin, 84(3), 489.

Simonton, D. K. (1977b). Creative productivity, age, and stress: a biographical time-series analysis of 10 classical composers. Journal of personality and social psychology, 35(11), 791.

Sipple, J. W., McCabe, L. A., & Casto, H. G. (2020). Child care deserts in New York State: Prekindergarten implementation and community factors related to the capacity to care for infants and toddlers. Early Childhood Research Quarterly, 51, 167-177.

Stepner, M. (2013). binscatter: Stata Module to Generate Binned Scatterplots. Boston College Department of Economics. Stata Command.

Swahn, M. H., Whitaker, D. J., Pippen, C. B., Leeb, R. T., Teplin, L. A., Abram, K. M., & McClelland, G. M. (2006). Concordance between Self-Reported Maltreatment and Court Records of Abuse or Neglect Among High-Risk Youths. *American Journal of Public Health*, *96*(10), 1849–1853.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

University of Kentucky Center for Poverty Research. (2023). *UKCPR National Welfare Data, 1980-2021 [Dataset]*.

Weigensberg, E., Islam, N., Knab, J., Grider, M., Page, J., & Larson, A. (2022). State Child Abuse and Neglect (SCAN) Policies Database 2019-2021 [Dataset]. In *National Data Archive on Child Abuse and Neglect.*

Wulczyn, F., Parolini, A., Schmits, F., Magruder, J., & Webster, D. (2020). Returning to foster care: Age and other risk factors. Children and Youth Services Review, 116, 105166.

Xue, C. (2017). Caring 24/7: Family child care as a solution to the paradox of nonstandard schedules and child care accessibility in the post-welfare reform era. CARING, 24(7), 2.

Appendices

Appendix to Chapter 2

Appendix 2A. Analytic Sample Construction of Reported Children in NCANDS

In my analytic sample construction for maltreatment reports, I first make systematic state-year exclusions, then I make child exclusions across states. Note that NCANDS is a voluntary system. Therefore, the data on reported children begins an unbalanced panel of state-years due to some the fact that some states did not submit data to NCANDS for some fiscal years.

➢ *State-year exclusion criterion*

Systematic exclusions of state-year pairs are made to fit a feasible context of the studying DR and re-reporting. While the analysis throughout this study refers to ‘year’ as the ‘year of the report,’ in creating the sample, state-years exclusions are based on NCANDS Child File fiscal years.

- To be able to identify maltreatment re-reporting, I need to be able to link children across reports. The first set of state-year exclusions are based on state-years that were not linkable by child ID. Breaks in the child ID links were determined by NDACAN statisticians and shared with the author in the form of a Child File linking grid Excel file. I drop state-years in which no reported children can be linked to reports in preceding years (i.e., 0% success rate for linking). Then, the state-year pairs that occur prior to a year with child ID linkage break for that state are dropped as well to ensure continuity.

- Second, I exclude all report-child observations from Alaska, Nevada, and New Jersey due to an incorrect use of alternative response codes or implementation of a program that was inconsistent with DR core components. This comes from Johnson-Motoyama et al. (2023), who had validated their characterization of DR using longitudinal information collected by the QIC-DR regarding actual DR program implementation and direct outreach to states. Given their reasoning for placing states in the control group (non-DR reforming) in their paper, I choose to exclude these states from my analyses.

- The third set of exclusions is for states that have DR systems that make determinations of track at the time of screening to screen out to AR. This is because NCANDS data do not provide the appropriate universe for analysis of those states. It would be more appropriate to use data on referrals to CPS, rather than data on CPS reports.

- The fourth set of exclusions is for states that are always treated. This exclusion is done to avoid issues that arise for comparisons between already-treated observations as control and later-treated observations as treatment. The comparison would specifically introduce bias to the conventional TWFE estimate when the treatment effects are heterogeneous (Baker et al., 2022; Borusyak et al., 2023; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021)

- The fifth set of exclusions is for states that discontinue the use of DR over this time frame.

➢ *Child exclusion criterion*

Then for child-level exclusions, I limit my sample to only consist of reported children ages 0 – 17 at the time of report. I also drop observations where the report disposition is categorized as 'other' or 'unknown/missing.' Based on the maltreatment allegation type, I also drop observations where 'neglect,' 'physical abuse,' 'emotional or psychological abuse,' or 'sexual abuse' were not at least one of the (up to four) documented reasons for the report. Again, these four types of child maltreatment are the most recognized and documented. The reports that were dropped had 'medical neglect,' 'no alleged maltreatment,' 'other,' or 'unknown or missing' documented as the reasons(s) for the report. Note, these reasons are not universally used across all jurisdictions.

Appendix 2B. Robustness Checks

Table B2.1. Robustness to the choice of estimator

	ATT	ATT by risk tercile		
		Low risk tercile	Middle risk tercile	High risk tercile
	(1)	(2)	(3)	(4)
Panel A.) Main specification: Callaway and Sant'Anna (unconditional; never treated)				
Differential Response	-0.02	-0.02*	-0.03**	-0.03*
	[-0.059, 0.019]	[-0.038, 0.003]	[-0.059, -0.008]	[-0.056, 0.004]
Panel B.) Callaway and Sant'Anna (unconditional; not yet treated)				
Differential Response	-0.01	-0.02*	-0.03**	-0.02
	[-0.050, 0.022]	[-0.036, 0.003]	[-0.060, -0.003]	[-0.048, 0.012]
Panel C.) Callaway and Sant'Anna (conditional; doubly robust; never treated)				
Differential Response	-0.04	0.01	0.03	0.02
	[-0.142, 0.060]	[-0.057, 0.080]	[-0.086, 0.141]	[-0.127, 0.163]
Panel D.) Two-way fixed effects (unconditional)				
Differential Response	0.02	0.01	0.02	0.02**
	[-0.010, 0.050]	[-0.016, 0.029]	[-0.010, 0.045]	[0.002, 0.031]
Panel E.) Two-way fixed effects (conditional)				
Differential Response	0.02*	-0.00	0.02**	0.02***
	[-0.003, 0.051]	[-0.008, 0.006]	[0.006, 0.052]	[0.005, 0.045]
Observations	15,660,574	4,202,890	4,239,644	4,263,808
Mean: outcome	0.46	0.35	0.47	0.60
Mean: AR | DR	0.11	0.11	0.13	0.09

Notes. The dependent variable is a re-report within 3 years. Conditional models include covariates for child, report, and local characteristics. Models use regression outcome, unless doubly robust is specified. The Callaway and Sant'Anna models use never treated comparison groups or not-yet treated comparison groups. Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table B2.2. Robustness to the outcome definition: length of time for a re-report

	ATT	ATT by risk tercile		
		Low risk tercile	Middle risk tercile	High risk tercile
	(1)	(2)	(3)	(4)
Panel A.) DV = Re-reported within 3 years				
Differential Response	-0.02	-0.02*	-0.03**	-0.03*
	[-0.059, 0.019]	[-0.038, 0.003]	[-0.059, -0.008]	[-0.056, 0.004]
Mean: outcome	0.46	0.35	0.47	0.60
Panel B.) DV = Re-reported within 12 months				
Differential Response	-0.01	-0.01	-0.02	-0.02
	[-0.048, 0.022]	[-0.037, 0.015]	[-0.056, 0.009]	[-0.055, 0.025]
Mean: outcome	0.30	0.23	0.31	0.41
Panel C.) DV = Re-reported within 6 months				
Differential Response	-0.01	-0.01	-0.01	-0.01
	[-0.039, 0.022]	[-0.033, 0.015]	[-0.043, 0.016]	[-0.049, 0.032]
Mean: outcome	0.21	0.15	0.20	0.28
Observations	15,660,574	4,202,890	4,239,644	4,263,808
Mean: AR | DR	0.11	0.11	0.13	0.09

Notes. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the Callaway and Sant'Anna model (unconditional model; never treated comparison group). Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table B2.3. Robustness to the outcome definition: re-report with substantiation

	ATT	ATT by risk tercile		
		Low risk tercile	Middle risk tercile	High risk tercile
	(1)	(2)	(3)	(4)
Panel A.) DV = Re-reported with substantiation within 3 years				
Differential Response	-0.02	-0.01**	-0.01	-0.01
	[-0.059, 0.019]	[-0.024, -0.000]	[-0.026, 0.004]	[-0.037, 0.016]
Mean: outcome	0.11	0.07	0.11	0.17
Panel B.) DV = Re-reported with substantiation within 12 months				
Differential Response	-0.01	-0.01	-0.01	-0.01
	[-0.048, 0.022]	[-0.013, 0.001]	[-0.014, 0.004]	[-0.024, 0.012]
Mean: outcome	0.08	0.05	0.07	0.12
Panel C.) DV = Re-reported with substantiation within 6 months				
Differential Response	0.00	-0.00	-0.00	-0.00
	[-0.007, 0.007]	[-0.008, 0.002]	[-0.010, 0.005]	[-0.017, 0.010]
Mean: outcome	0.05	0.03	0.05	0.08
Observations	15,660,574	4,202,890	4,239,644	4,263,808
Mean: AR | DR	0.11	0.11	0.13	0.09

Notes. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the Callaway and Sant'Anna model (unconditional model; never treated comparison group). Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

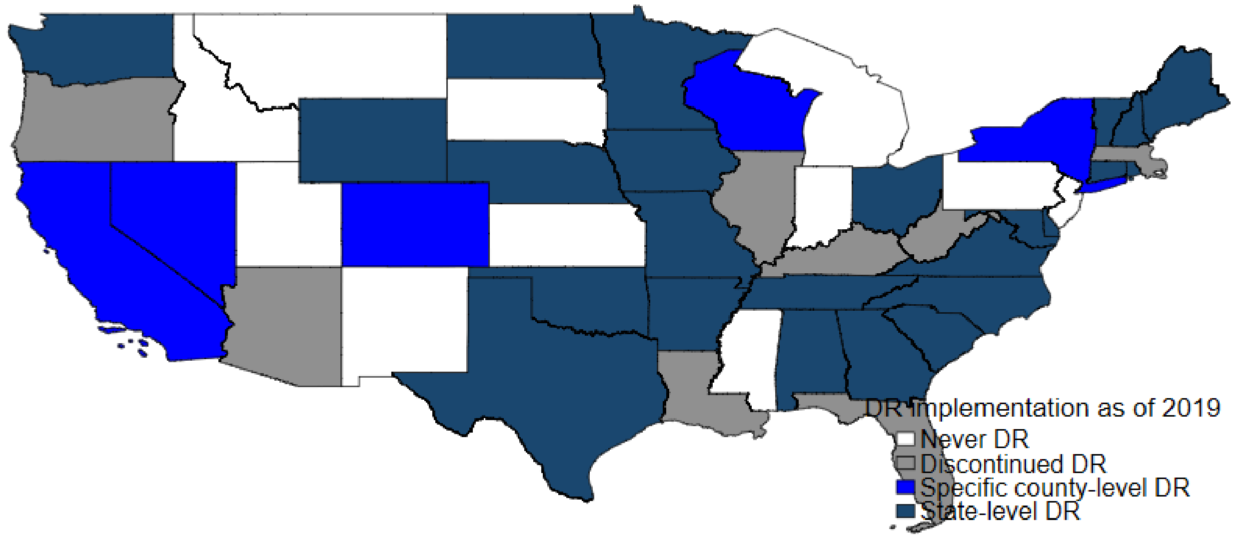
Table B2.4. Robustness to the sample

	ATT	ATT by risk tercile		
		Low risk tercile	Middle risk tercile	High risk tercile
	(1)	(2)	(3)	(4)
Panel A.) Main specification				
Differential Response	-0.02	-0.02*	-0.03**	-0.03*
	[-0.059, 0.019]	[-0.038, 0.003]	[-0.059, -0.008]	[-0.056, 0.004]
Observations	15,660,574	4,202,890	4,239,644	4,263,808
Mean: outcome	0.46	0.35	0.47	0.60
Mean: AR | DR	0.11	0.11	0.13	0.09
Panel B.) Excluding states with county-level DR systems (i.e., New York and Colorado)				
Differential Response	0.01	-0.01	-0.04*	-0.02
	[-0.028, 0.038]	[-0.048, 0.023]	[-0.087, 0.004]	[-0.055, 0.023]
Observations	11,958,518	4,069,860	4,069,831	4,069,829
Mean: outcome	0.45	0.33	0.47	0.58
Mean: AR | DR	0.14	0.15	0.17	0.13
Panel C.) Excluding small, masked counties				
Differential Response	-0.02	-0.02*	-0.03***	-0.03**
	[-0.063, 0.014]	[-0.036, 0.003]	[-0.055, -0.012]	[-0.060, -0.005]
Observations	13,012,436	3,585,609	3,541,880	3,401,670
Mean: outcome	0.45	0.34	0.48	0.61
Mean: AR | DR	0.09	0.11	0.12	0.09

Notes. The coefficients shown are the weighted average (by group size) of all available group-time average treatment effects estimates from the Callaway and Sant'Anna model (unconditional model; never treated comparison group). Wild bootstrapped t-statistics are used to create confidence intervals and determine significance. The 95% confidence intervals are in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

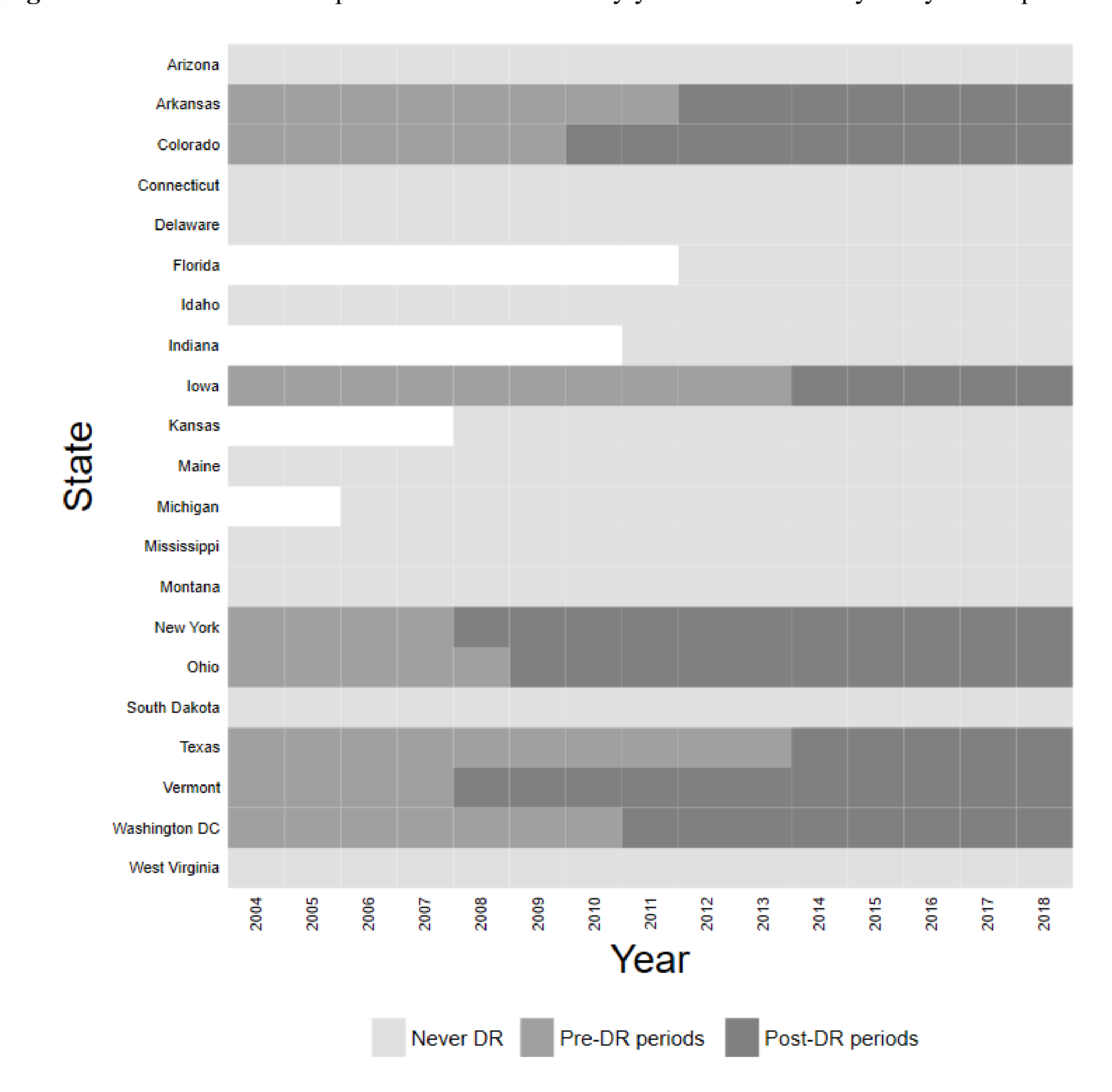
Appendix 2C. Supplementary Figures and Tables Noted in Text

Figure C2.1. Differential response implementation across U.S. states as of 2019



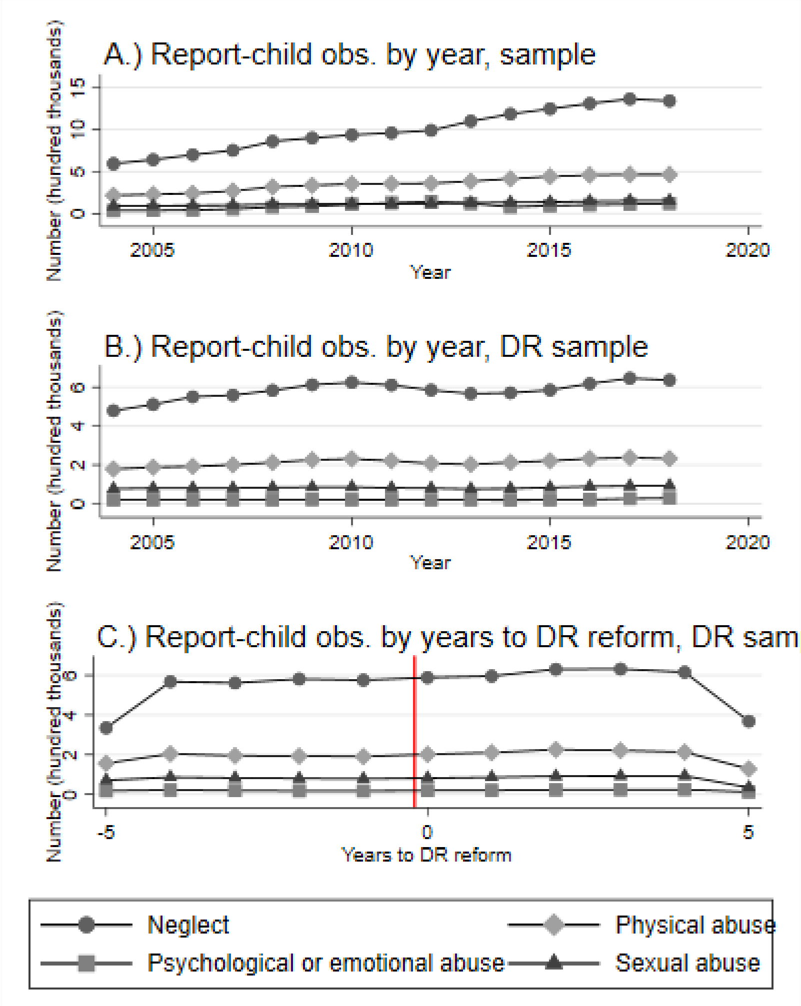
Notes. Differential response is a popular CPS reform. The states with statewide and specific county-level DR in 2019 were identified using the SCAN Policies Database 2019 (Weigensberg et al., 2022). The states that previously implemented DR but discontinued it were identified using other sources. By 2019, Alaska (not shown) had discontinued DR, Hawaii (not shown) had state-level DR, and the District of Columbia (not shown) had discontinued DR. Discontinuation occurs for numerous reasons including funding.

Figure C2.2. Differential response treatment status by year for states in my analytic sample



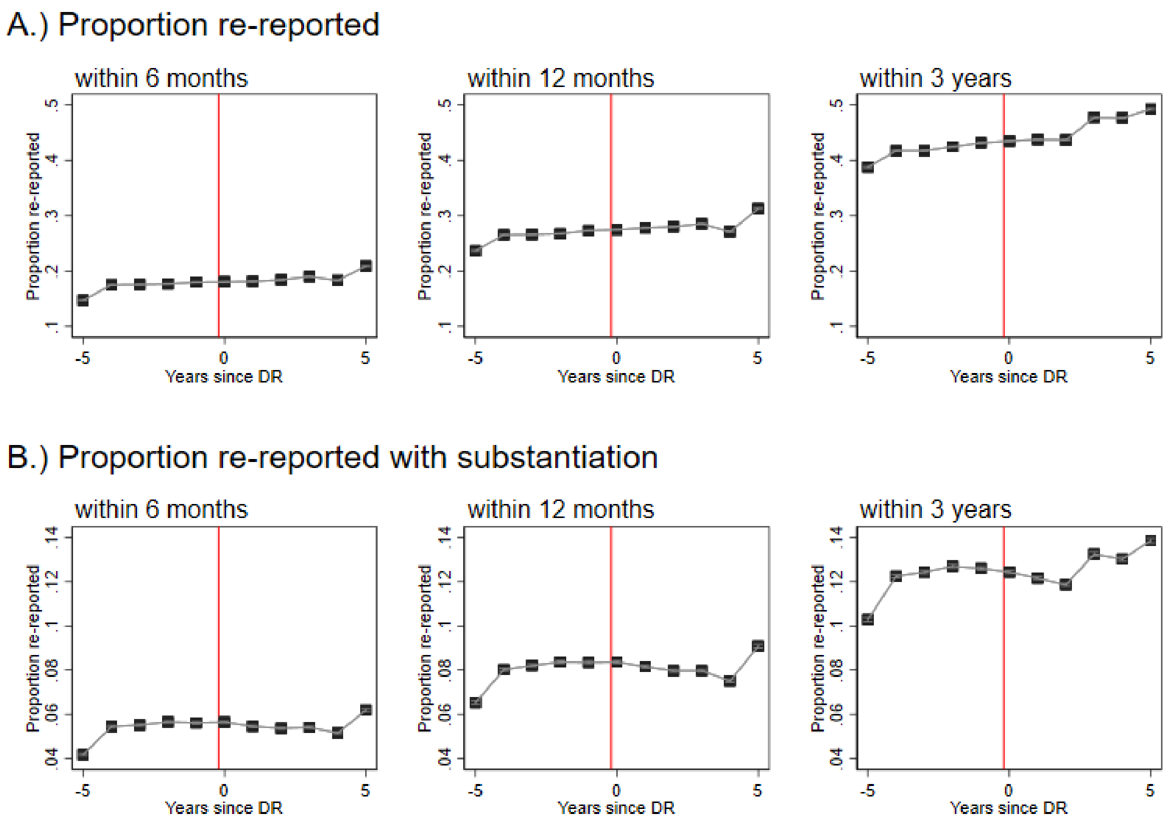
Notes. There are 21 states in the sample, and 8 of these states adopt DR between 2004 and 2019. These are specifically states that do not discontinue the use of DR before 2019, and states that make the track determination (AR or IR) after a report is screened in. The white cells indicate state-year pairs that are dropped from the sample (See Appendix A).

Figure C2.3. Number of annual reports-child observations by reasons for report



Notes. In panel A, the number of annual report-child observations are pooled across all sampled states (N=21 states) by the reason for the report. In panels B and C, the annual sums are pooled across the 8 states that adopt DR between 2004 and 2019 and do not discontinue in this time frame. In panel C, year 0 represents the relative year to DR reform.

Figure C2.4. Proportion of reports-child observations that are re-reported by year to DR reform



Notes. The average proportions of reports-child observations that are re-reported are pooled across the 8 states by the relative year to DR reform. These 8 states adopt DR between 2004 and 2019 and do not discontinue in this time frame.

Table C2.1. Characteristics of report-child observations in states that adopt DR policies between 2004 and 2019

	Pre-DR reform		Post-DR reform		Post-DR reform										
	IR pathway		IR *or* AR pathway		IR pathway		AR pathway		Difference: (2) - (1)			Difference: (4) - (3)		
	(1)		(2)		(3)		(4)		(5)			(6)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Diff.*		*SE*	*Diff.*		*SE*
Child characteristics														
Prior victim	**0.24**	0.43	**0.30**	0.46	**0.31**	0.46	**0.22**	0.41	**0.06**	**	(0.02)	**-0.09**	**	(0.04)
Prior victim missing	**0.00**	0.06	**0.00**	0.07	**0.00**	0.04	**0.03**	0.17	**0.00**		(0.01)	**0.03**		(0.04)
Reason for Report														
Neglect	**0.75**	0.44	**0.79**	0.40	**0.81**	0.39	**0.70**	0.46	**0.05**		(0.05)	**-0.10**	*	(0.05)
Physical abuse	**0.30**	0.46	**0.27**	0.44	**0.26**	0.44	**0.36**	0.48	**-0.03**		(0.06)	**0.10**	*	(0.05)
Sexual abuse	**0.13**	0.34	**0.10**	0.29	**0.11**	0.31	**0.00**	0.06	**-0.03**	**	(0.01)	**-0.10**	**	(0.04)
Psychological/emotional abuse	**0.03**	0.17	**0.03**	0.17	**0.03**	0.16	**0.05**	0.22	**-0.00**		(0.00)	**0.02**	**	(0.01)
Age at report	**7.20**	4.99	**7.62**	5.07	**7.56**	5.08	**8.09**	4.92	**0.43**	*	(0.21)	**0.52**		(0.83)
Female	**0.51**	0.50	**0.50**	0.50	**0.50**	0.50	**0.48**	0.50	**-0.01**	**	(0.00)	**-0.02**	***	(0.00)
Female missing	**0.00**	0.06	**0.01**	0.07	**0.01**	0.07	**0.00**	0.06	**0.00**	**	(0.00)	**0.00**	*	(0.00)
Race														
White	**0.66**	0.47	**0.57**	0.50	**0.56**	0.50	**0.63**	0.48	**-0.09**	***	(0.03)	**0.07**		(0.06)
Black	**0.20**	0.40	**0.24**	0.43	**0.24**	0.43	**0.21**	0.41	**0.04**	**	(0.01)	**-0.04**		(0.02)
American Indian, Alaskan Native	**0.00**	0.05	**0.00**	0.06	**0.00**	0.06	**0.00**	0.05	**0.00**		(0.00)	**-0.00**		(0.00)
Asian or Pacific Islander	**0.01**	0.08	**0.01**	0.11	**0.01**	0.11	**0.01**	0.09	**0.00**	*	(0.00)	**-0.00**		(0.00)
Two or more races	**0.03**	0.17	**0.04**	0.21	**0.04**	0.20	**0.06**	0.24	**0.02**	*	(0.01)	**0.02**	*	(0.01)
Unknown/missing race	**0.10**	0.30	**0.13**	0.34	**0.14**	0.35	**0.09**	0.28	**0.04**		(0.02)	**-0.05**		(0.04)
Hispanic ethnicity	**0.32**	0.47	**0.27**	0.44	**0.29**	0.45	**0.17**	0.37	**-0.05**		(0.04)	**-0.12**	***	(0.03)
Hispanic ethnicity missing	**0.14**	0.35	**0.17**	0.37	**0.17**	0.37	**0.17**	0.38	**0.03**		(0.05)	**0.01**		(0.03)
Living arrangements														
Both parents	**0.12**	0.33	**0.29**	0.45	**0.29**	0.45	**0.29**	0.45	**0.16**	***	(0.03)	**0.00**		(0.08)
Parent and other	**0.04**	0.20	**0.10**	0.31	**0.10**	0.30	**0.15**	0.36	**0.06**		(0.04)	**0.06**		(0.04)
Single mother	**0.07**	0.26	**0.14**	0.34	**0.13**	0.34	**0.18**	0.39	**0.06**		(0.04)	**0.05**		(0.03)

Single father	**0.01**	0.10	**0.02**	0.13	**0.02**	0.13	**0.02**	0.15	**0.01**		(0.01)	**0.01**		(0.01)
Non-parent relative caregiver	**0.01**	0.11	**0.02**	0.15	**0.02**	0.14	**0.02**	0.15	**0.01**		(0.01)	**0.00**		(0.00)
Non-relative caregiver	**0.02**	0.12	**0.02**	0.12	**0.02**	0.13	**0.00**	0.07	**0.00**		(0.00)	**-0.01**	***	(0.00)
Group home or residential facility	**0.00**	0.06	**0.00**	0.05	**0.00**	0.06	**0.00**	0.02	**-0.00**		(0.00)	**0.00**	*	(0.00)
Other or missing living arrangements	**0.72**	0.45	**0.41**	0.49	**0.43**	0.49	**0.32**	0.47	**-0.31**	***	(0.08)	**-0.11**		(0.10)
Report characteristics														
Report source														
Social services personnel	**0.10**	0.30	**0.14**	0.34	**0.14**	0.35	**0.11**	0.31	**0.04**	***	(0.01)	**-0.03**	**	(0.01)
Medical and mental health personnel	**0.11**	0.31	**0.12**	0.32	**0.12**	0.32	**0.11**	0.31	**0.01**		(0.01)	**-0.01**		(0.01)
Law enforcement	**0.15**	0.36	**0.16**	0.36	**0.16**	0.37	**0.15**	0.36	**0.01**		(0.01)	**-0.01**		(0.01)
Education, child care, substitute care personnel	**0.16**	0.37	**0.17**	0.38	**0.17**	0.38	**0.22**	0.41	**0.01**	**	(0.01)	**0.05**	*	(0.02)
Parents or relative	**0.18**	0.38	**0.15**	0.36	**0.15**	0.35	**0.16**	0.37	**-0.03**		(0.02)	**0.02**		(0.01)
Alleged victim	**0.00**	0.05	**0.00**	0.04	**0.00**	0.04	**0.00**	0.05	**-0.00**		(0.00)	**0.00**		(0.00)
Friends or neighbors	**0.06**	0.24	**0.05**	0.22	**0.05**	0.22	**0.05**	0.22	**-0.01**		(0.01)	**0.00**		(0.00)
Other, unknown, or missing	**0.23**	0.42	**0.21**	0.41	**0.22**	0.41	**0.20**	0.40	**-0.02**		(0.02)	**-0.02**		(0.01)
Administrative CPS outcomes														
Substantiated	**0.30**	0.46	**0.27**	0.44	**0.31**	0.46	**0.00**	0.00	**-0.03**		(0.02)	**-0.31**	***	(0.02)
Re-reported														
within 6 mo.	**0.16**	0.37	**0.20**	0.40	**0.20**	0.40	**0.18**	0.38	**0.04**	***	(0.01)	**-0.03**	**	(0.01)
within 12 mo.	**0.25**	0.43	**0.30**	0.46	**0.30**	0.46	**0.27**	0.44	**0.04**	***	(0.01)	**-0.03**	**	(0.01)
within 3 yrs.	**0.40**	0.49	**0.47**	0.50	**0.47**	0.50	**0.45**	0.50	**0.06**	***	(0.01)	**-0.02**		(0.01)
Re-reported w. sub.														
within 6 mo.	**0.05**	0.22	**0.06**	0.24	**0.06**	0.24	**0.03**	0.17	**0.01**		(0.01)	**-0.04**	***	(0.01)
within 12 mo.	**0.08**	0.27	**0.09**	0.28	**0.09**	0.29	**0.04**	0.20	**0.01**		(0.01)	**-0.05**	***	(0.01)
within 3 yrs.	**0.12**	0.32	**0.13**	0.34	**0.14**	0.35	**0.07**	0.25	**0.01**		(0.01)	**-0.07**	***	(0.01)
Observations	4,888,868		6,410,877		5,686,047		724,830		11,299,745			6,410,877		

Notes. Report-child observations are pooled from 2004 to 2019. In columns (5) and (6), the statistical significance of the difference between the two sets of mutually exclusive groups of observations comes from separate regressions of the group indicator (e.g., post-DR =1) on each characteristic. Cluster robust standard errors are clustered by state.

Table C2.2. Characteristics of report-child observations that are investigated in state-years with no DR reform (conventional investigative CPS structure) and state-years that are post-DR reform (CPS with DR reform)

	No DR reform: *Only IR* (Pre-DR *or* never DR reform)		**Post-DR reform**: *Selected to be IR*		**Difference: (2) - (1)**		
	(1)		(2)		(3)		
	Mean	*SD*	*Mean*	*SD*	*Diff.*		*SE*
Child characteristics							
Prior victim	**0.28**	0.45	**0.31**	0.46	**0.03**		(0.05)
Prior victim missing	**0.02**	0.13	**0.00**	0.04	**-0.02**		(0.02)
Reason for Report							
Neglect	**0.76**	0.43	**0.81**	0.39	**0.04**		(0.08)
Physical abuse	**0.28**	0.45	**0.26**	0.44	**-0.02**		(0.05)
Sexual abuse	**0.09**	0.29	**0.11**	0.31	**0.02**		(0.03)
Psychological or emotional abuse	**0.09**	0.29	**0.03**	0.16	**-0.06**	**	(0.03)
Age at report	**7.42**	5.05	**7.56**	5.08	**0.15**		(0.19)
Female	**0.50**	0.50	**0.50**	0.50	**-0.00**		(0.00)
Female missing	**0.00**	0.07	**0.01**	0.07	**0.00**		(0.00)
Race							
White	**0.66**	0.47	**0.56**	0.50	**-0.10**	*	(0.05)
Black	**0.20**	0.40	**0.24**	0.43	**0.04**		(0.03)
American Indian, Alaskan Native	**0.01**	0.10	**0.00**	0.06	**-0.01**		(0.00)
Asian or Pacific Islander	**0.01**	0.07	**0.01**	0.11	**0.01**	*	(0.00)
Two or more races	**0.04**	0.21	**0.04**	0.20	**-0.00**		(0.01)
Unknown/missing race	**0.08**	0.26	**0.14**	0.35	**0.06**	*	(0.03)
Hispanic ethnicity	**0.22**	0.41	**0.29**	0.45	**0.07**		(0.04)
Hispanic ethnicity missing	**0.17**	0.38	**0.17**	0.37	**-0.00**		(0.04)
Living arrangements							

Both parents	**0.20**	0.40	**0.29**	0.45	**0.09**		(0.10)
Parent and other	**0.13**	0.33	**0.10**	0.30	**-0.03**		(0.04)
Single mother	**0.13**	0.34	**0.13**	0.34	**-0.00**		(0.04)
Single father	**0.02**	0.14	**0.02**	0.13	**-0.00**		(0.01)
Non-parent relative caregiver	**0.02**	0.15	**0.02**	0.14	**-0.00**		(0.01)
Non-relative caregiver	**0.01**	0.11	**0.02**	0.13	**0.00**	*	(0.00)
Group home or residential facility	**0.00**	0.06	**0.00**	0.06	**-0.00**		(0.00)
Other or missing living arrangements	**0.49**	0.50	**0.43**	0.49	**-0.06**		(0.13)
Report characteristics							
Report source							
Social services personnel	**0.11**	0.31	**0.14**	0.35	**0.03**		(0.03)
Medical and mental health personnel	**0.12**	0.32	**0.12**	0.32	**0.00**		(0.01)
Law enforcement	**0.16**	0.37	**0.16**	0.37	**-0.00**		(0.02)
Education, child care, substitute care personnel	**0.17**	0.38	**0.17**	0.38	**-0.00**		(0.01)
Parents or relative	**0.17**	0.38	**0.15**	0.35	**-0.03**	***	(0.01)
Alleged victim	**0.00**	0.06	**0.00**	0.04	**-0.00**		(0.00)
Friends or neighbors	**0.06**	0.24	**0.05**	0.22	**-0.01**	*	(0.01)
Other, unknown, or missing	**0.20**	0.40	**0.22**	0.41	**0.02**		(0.02)
Administrative CPS outcomes							
Substantiated	**0.22**	0.42	**0.31**	0.46	**0.08**	**	(0.03)
Re-reported							
within 6 mo.	**0.21**	0.41	**0.20**	0.40	**-0.01**		(0.02)
within 12 mo.	**0.31**	0.46	**0.30**	0.46	**-0.01**		(0.03)
within 3 yrs.	**0.46**	0.50	**0.47**	0.50	**0.00**		(0.03)
Re-reported w. sub.							
within 6 mo.	**0.05**	0.21	**0.06**	0.24	**0.02**		(0.01)
within 12 mo.	**0.07**	0.25	**0.09**	0.29	**0.02**		(0.02)

within 3 yrs.	**0.11**	0.31	**0.14**	0.35	**0.03**	(0.02)
Observations	12,772,738		5,686,047		18,458,785	

Notes. Report-child observations are pooled from 2004 to 2019. Column (1) shows average characteristics of reported children in state-years with no DR (i.e., cases from pre-DR state-years and never DR states). Column (2) shows average characteristics of reported children in state-years with DR currently in place (i.e., post-DR). In column (3), the statistical significance of the differences between the mutually exclusive groups of observations comes from separate regressions of the group indicator (e.g., post-DR =1) on each characteristic. Cluster robust standard errors are clustered by state.

Table C2.3. Descriptive regressions predicting an alternative response in periods with DR reform

	DV = Alternative Response Track		
	(1)	(2)	(3)
Child characteristics			
Prior victim	-0.04***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)
Reasons for report			
Neglect	-0.00	0.00	-0.00
	(0.02)	(0.02)	(0.02)
Physical abuse	-0.03*	-0.03*	-0.03**
	(0.01)	(0.01)	(0.01)
Sexual abuse	-0.15**	-0.14**	-0.15**
	(0.05)	(0.05)	(0.05)
Psychological/emotional abuse	0.02	0.02	0.02
	(0.01)	(0.02)	(0.01)
Child age at report (omit. age 5)			
Child age: 0	-0.04***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)
Child age: 1	-0.02**	-0.02*	-0.02**
	(0.01)	(0.01)	(0.01)
Child age: 2	-0.02*	-0.01	-0.01*
	(0.01)	(0.01)	(0.01)
Child age: 3	-0.01*	-0.01*	-0.01*
	(0.01)	(0.00)	(0.00)
Child age: 4	-0.01*	-0.01*	-0.01*
	(0.00)	(0.00)	(0.00)
Child age: 6	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)
Child age: 7	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)
Child age: 8	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)
Child age: 9	0.04	0.04	0.04
	(0.02)	(0.02)	(0.02)
Child age: 10	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)
Child age: 11	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)
Child age: 12	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)
Child age: 13	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)

Child age: 14	0.05	0.05	0.05
	(0.03)	(0.03)	(0.03)
Child age: 15	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)
Child age: 16	0.05	0.05	0.05
	(0.03)	(0.03)	(0.03)
Child age: 17	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)
Sex (omit. male)			
Female	-0.00*	-0.00	-0.00*
	(0.00)	(0.00)	(0.00)
Race (omit. White)			
Black	-0.04**	-0.04**	-0.02**
	(0.02)	(0.01)	(0.01)
American Indian, Alaskan Native	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.01)
Asian or Pacific Islander	-0.04**	-0.04**	-0.03*
	(0.01)	(0.01)	(0.01)
Two or more races	0.01	0.01	0.01
	(0.00)	(0.00)	(0.00)
Unknown/missing race	-0.04**	-0.04**	-0.03***
	(0.01)	(0.01)	(0.01)
Ethnicity (omit. non-Hispanic)			
Hispanic	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)
Living arrangements (omit. Both parents)			
Parent and other	-0.03	-0.03	-0.04
	(0.04)	(0.04)	(0.04)
Single mother	-0.04	-0.04	-0.04
	(0.05)	(0.05)	(0.05)
Single father	-0.06	-0.06	-0.06
	(0.06)	(0.06)	(0.06)
Non-parent relative caregiver	-0.07	-0.07	-0.07
	(0.05)	(0.05)	(0.05)
Non relative caregiver	-0.09*	-0.09*	-0.09*
	(0.05)	(0.04)	(0.04)
Group home or residential facility	-0.15**	-0.15**	-0.14**
	(0.05)	(0.05)	(0.05)
Other or missing living arrangements	-0.04	-0.04	-0.04
	(0.04)	(0.03)	(0.04)
Report characteristics			
Report source (omit. social services personnel)			
Medical and mental health personnel		0.01*	0.01*

		(0.01)	(0.01)
Law enforcement		-0.00	-0.01
		(0.01)	(0.01)
Education, child care, substitute care personnel		0.04**	0.04**
		(0.02)	(0.01)
Parents or relative		0.02	0.02
		(0.01)	(0.01)
Alleged victim		-0.02	-0.01
		(0.02)	(0.02)
Friends or neighbors		0.02	0.02
		(0.01)	(0.01)
Other, unknown, or missing		0.02	0.02
		(0.01)	(0.01)
Month of report (omit. January)			
February		0.00	0.00
		(0.00)	(0.00)
March		0.00	0.00
		(0.00)	(0.00)
April		0.00	0.01
		(0.00)	(0.00)
May		0.00	0.00
		(0.00)	(0.00)
June		0.00	0.00
		(0.00)	(0.00)
July		0.01	0.01
		(0.00)	(0.00)
August		0.01*	0.01*
		(0.00)	(0.01)
September		0.01*	0.01*
		(0.00)	(0.01)
October		0.01*	0.01*
		(0.01)	(0.01)
November		0.01	0.01*
		(0.01)	(0.01)
December		0.01	0.01
		(0.01)	(0.01)
Local demographic and economic factors	No	No	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	5,306,355	5,306,355	5,306,355
Adjusted R-squared	0.15	0.15	0.18

Notes. Report-child observations from periods with DR (i.e., post-DR) are used. Year and state fixed effects are binary indicators. The ordinary least squares regressions use cluster robust standard errors that are clustered by state. Standard errors are shown in parentheses. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table C2.4. Descriptive regressions of a re-report in periods without DR reform

	DV = Re-reported						DV = Re-reported w. substantiation					
	within 6 mo.		within 12 mo.		within 3 yrs.		within 6 mo.		within 12 mo.		within 3 yrs.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Child characteristics												
Prior victim	0.11***	0.11***	0.15***	0.14***	0.18***	0.18***	0.05***	0.05***	0.06***	0.06***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reasons for report												
Neglect	0.03***	0.04***	0.05***	0.05***	0.06***	0.06***	0.01***	0.02***	0.02***	0.02***	0.02***	0.03***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Physical abuse	0.02***	0.01***	0.02***	0.01***	0.03***	0.02**	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sexual abuse	0.00	0.00	0.00	-0.01	-0.01	-0.02**	0.00	0.00	-0.00**	0.00	-0.01***	-0.01**
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	-0.01	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Psychological/emotional abuse	0.02***	0.01	0.02***	0.01	0.02**	0.01	0.01***	0.00	0.01***	0.00	0.01***	0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age at report (omit. age 5)												
Child age: 0	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.02***	0.01***	0.02***	0.02***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 1	0.01*	0.00	0.01**	0.01*	0.02***	0.02***	0.01***	0.01***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 2	0.00	0.00	0.00	0.00	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 3	0.00	0.00	0.00*	0.00	0.01***	0.01***	0.00***	0.00***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 4	0.00	0.00	0.00	0.00	0.01***	0.01***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 6	-0.00***	-0.00***	-0.00***	-0.00***	-0.01***	-0.00**	-0.00*	-0.00*	-0.00***	-0.00***	-0.00*	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 7	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***	-0.00***	-0.00***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Child age: 8	-0.02***	-0.01***	-0.02***	-0.02***	-0.03***	-0.03***	-0.00***	-0.00***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 9	-0.02***	-0.02***	-0.03***	-0.03***	-0.04***	-0.04***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 10	-0.03***	-0.03***	-0.04***	-0.04***	-0.05***	-0.04***	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 11	-0.03***	-0.03***	-0.04***	-0.04***	-0.05***	-0.04***	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 12	-0.03***	-0.02***	-0.04***	-0.03***	-0.05***	-0.04***	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 13	-0.03***	-0.02***	-0.03***	-0.03***	-0.05***	-0.04***	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 14	-0.03***	-0.02***	-0.04***	-0.03***	-0.06***	-0.06***	-0.01***	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 15	-0.04***	-0.03***	-0.06***	-0.05***	-0.13***	-0.12***	-0.01***	-0.01***	-0.02***	-0.02***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child age: 16	-0.06***	-0.06***	-0.10***	-0.09***	-0.24***	-0.23***	-0.02***	-0.02***	-0.03***	-0.03***	-0.06***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Child age: 17	-0.12***	-0.11***	-0.21***	-0.20***	-0.39***	-0.38***	-0.03***	-0.03***	-0.05***	-0.05***	-0.09***	-0.10***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Sex (omit. male)												
Female	0.00***	0.00***	0.01***	0.01***	0.01***	0.01***	0.00***	0.00***	0.00***	0.00***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Race (omit. White)												
Black	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03**	-0.01***	-0.01***	-0.01***	-0.01***	0.00	-0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
American Indian, Alaskan Native	-0.01	-0.02**	-0.01	-0.02**	-0.01	-0.01	0.01**	0.00	0.01*	0.00	0.02**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Asian or Pacific Islander	-0.07***	-0.07***	-0.10***	-0.10***	-0.16***	-0.16***	-0.02***	-0.02***	-0.02***	-0.02***	-0.04***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Two or more races	0.02***	0.02***	0.03***	0.03***	0.04***	0.04***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Unknown/missing race	-0.06***	-0.06***	-0.08***	-0.08***	-0.11***	-0.11***	-0.02***	-0.02***	-0.03***	-0.02***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Ethnicity (omit. non-Hispanic)												
Hispanic	-0.02***	-0.02***	-0.03***	-0.03***	-0.03***	-0.04***	-0.00**	-0.00***	0.00	-0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Living arrangements (omit. Both parents)												
Parent and other	0.03***	0.04***	0.04***	0.05***	0.03***	0.05***	0.00	0.00	0.00	0.00	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Single mother	0.02***	0.03***	0.03***	0.04***	0.03***	0.04***	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Single father	0.03***	0.04***	0.03***	0.04***	0.02**	0.04***	-0.00*	-0.01***	-0.01**	-0.01***	-0.01*	-0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Non-parent relative caregiver	0.00	0.01	0.00	0.02*	0.00	0.02	-0.01**	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Non relative caregiver	0.00	-0.01	0.00	-0.01	-0.01	-0.01	-0.02**	-0.02**	-0.02**	-0.02***	-0.03***	-0.03***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Group home or residential facility	0.05***	0.04	0.06**	0.05	0.08***	0.07***	-0.01	-0.01*	-0.01	-0.02*	-0.02**	-0.02**
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Other or missing living arrangements	0.02	-0.01	0.02	-0.02	0.00	-0.02	0.00	0.00	0.00	0.00	-0.01*	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Report characteristics												
Report source (omit. SS personnel)												
Medical and mental health personnel	-0.01*	-0.01**	-0.01*	-0.01**	0.00	-0.01	-0.00***	-0.00**	-0.01***	-0.01**	-0.00**	0.00
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Law enforcement	-0.02***	-0.03***	-0.03***	-0.03***	-0.02***	-0.03***	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Education, CC, sub. care	0.00	0.00	0.01*	0.01	0.01***	0.01*	-0.00***	-0.00**	-0.00*	-0.00*	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Parents or relative	0.01**	0.01	0.01***	0.01**	0.02***	0.02***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Alleged victim	0.03***	0.03***	0.04***	0.03***	0.03***	0.03***	-0.00*	-0.01***	-0.00**	-0.01***	-0.00*	-0.01***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Friends or neighbors	0.00	0.00	0.01	0.01	0.01	0.01	-0.00***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other, unknown, or missing	0.01	0.00	0.01*	0.01	0.01**	0.00	-0.00***	-0.00***	-0.01***	-0.00**	-0.01***	-0.00*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Month of report (omit. January)												
February	-0.00***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
March	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.00**	-0.00**	-0.00*	-0.00*	-0.00*	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
April	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.00*	-0.00*	-0.00**	-0.00**	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
May	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.00**	-0.00**	-0.00**	-0.00**	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
June	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.00***	-0.00***	-0.00**	-0.00**	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
July	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.00***	-0.00***	-0.00***	-0.00***	-0.01***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
August	-0.01**	-0.01**	-0.00**	-0.00**	-0.01*	-0.01*	-0.00***	-0.00***	-0.00**	-0.00**	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
September	0.00	0.00	0.00	0.00	0.00	0.00	-0.00***	-0.00***	0.00	-0.00*	0.00	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
October	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00**	-0.00*	-0.00**	-0.00*	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
November	0.01***	0.01***	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00*	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
December	0.00**	0.00**	0.00	-0.01*	0.00	0.00	0.00	0.00	-0.00**	-0.00***	-0.00*	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Local demographic and economic factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

State fixed effects	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Year fixed effects	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Observations	10,354,219	10,354,219	10,354,219	10,354,219	8,874,866	8,874,866	10,354,219	10,354,219	10,354,219	10,354,219	8,874,866	8,874,866
Adjusted R-squared	0.048	0.041	0.062	0.054	0.094	0.082	0.018	0.016	0.023	0.019	0.031	0.025

Notes. Report-child observations from periods without DR are used. Periods without DR are periods in states that never adopt DR and pre-DR periods for states that eventually adopt DR between 2004 and 2019. Year and state fixed effects are binary indicators. The ordinary least squares regressions use cluster robust standard errors that are clustered at the state level. Standard errors are shown in parentheses. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table C2.5. Alternative response utilization rates by state-year

	Year of the report														
State	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Arizona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Arkansas	0	0	0	0	0	0	0	0	0	0.08	0.11	0.12	0.13	0.16	0.18
Colorado	0	0	0	0	0	0	0.01	0.06	0.11	0.17	0.19	0.19	0.17	0.2	0.19
Connecticut	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Delaware	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Washington DC	0	0	0	0	0	0	0	0.01	0.04	0.14	0.38	0.40	0.44	0.44	0.44
Florida									0	0	0	0	0	0	0
Idaho	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Indiana								0	0	0	0	0	0	0	0
Iowa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kansas					0	0	0	0	0	0	0	0	0	0	0
Maine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Michigan			0	0	0	0	0	0	0	0	0	0	0	0	0
Mississippi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Montana	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
New York	0	0	0	0	0	0.02	0.04	0.06	0.08	0.08	0.07	0.07	0.07	0.07	0.05
Ohio	0	0	0	0	0	0	0.05	0.16	0.19	0.26	0.35	0.41	0.42	0.43	0.41
South Dakota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Texas	0	0	0	0	0	0	0	0	0	0	0	0.03	0.07	0.09	0.12
Vermont	0	0	0	0	0	0.19	0.29	0.28	0.29	0.32	0.36	0.37	0.35	0.32	0.28
West Virginia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 2D. Differential Response Implementation Differs Across Jurisdiction

While there are common features of DR systems across jurisdictions, DR is not a uniformly defined reform, and its implementation varies. Four main differences in implementation stand out. I describe these differences using information from the SCAN Policies Database 2019 (Weigensberg et al., 2022).

First, in some states, the determination for the appropriate track (IR or AR) for maltreatment referrals is made *at the time of screening to screen out to AR*, whereas in other states, a determination is made *after a report is screened in*. Among the 30 states with DR in 2019, 17 states make the track determination only after a report is screened in, six states make the determination only *at the time of screening to screen out to AR*, six states allow for both (determinations at screening *and* after a report is screened in), and one state uses another process.

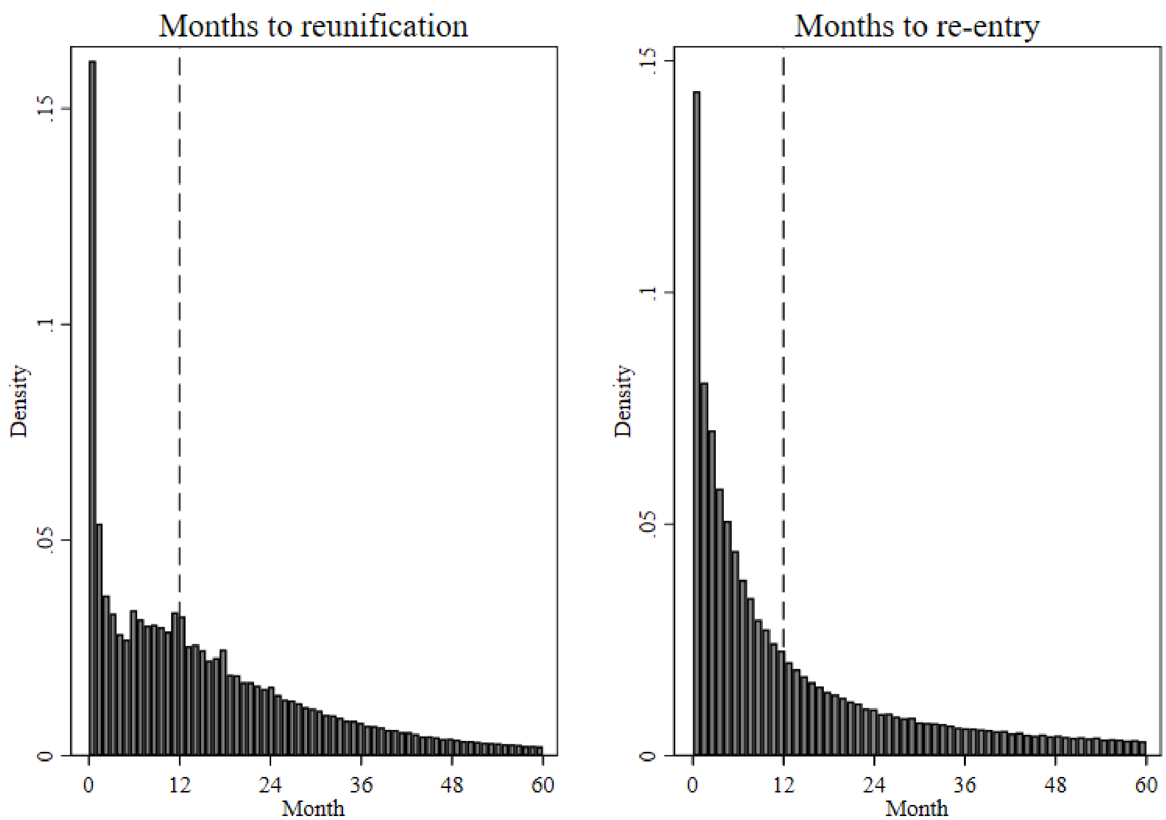
Second, some jurisdictions restrict the types of child maltreatment that are eligible for AR. In 2019, among the 30 states with DR, 27 states restricted eligibility for AR based on maltreatment type. In 24 DR states, sexual abuse cases are ineligible for AR. In 18 DR states, physical abuse cases are ineligible for AR. In 9 DR states, neglect cases are ineligible. Other report types that may be ineligible include child fatality (25 states), substance-exposed newborns (4 states), abandoned infant cases (11 states), or cases based on some other pre-defined ineligibility (13 states).

Third, among the 30 DR states in 2019, 24 states have eligibility for AR that is determined using a formal risk determination process, whereas in 6 states eligibility is not informed by risk determination. However, the tools differ even across states that use risk determination. Some examples of tools include structured decision-making or a standard assessment tool.

Fourth, for families involved with CPS, the efficacy of DR in addressing the underlying reasons for the maltreatment report likely depends on the criterion for who receives referrals to community services. In some jurisdictions, all cases on the AR track receive referrals to services (i.e., universal referrals to services). In other jurisdictions, referrals are selectively made when families express interest (i.e., opt-in referrals to services), when there is a determination of risk (i.e., meet the criterion for referrals to services), or via some other process.

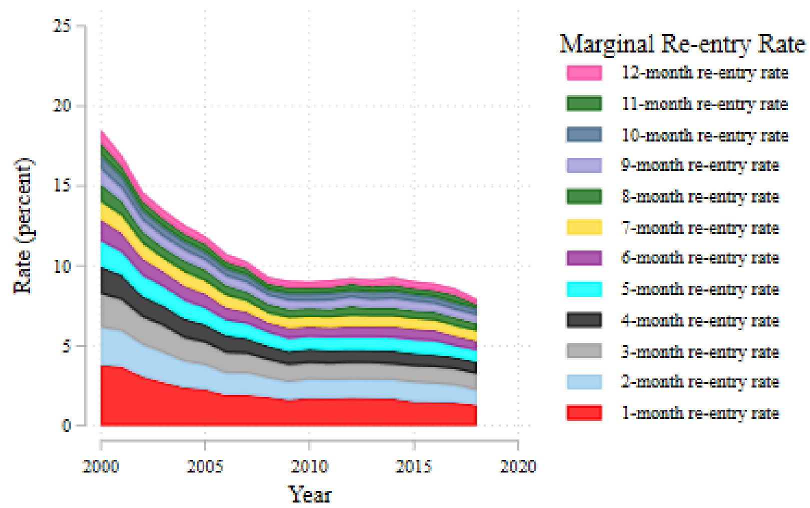
Appendix to Chapter 3

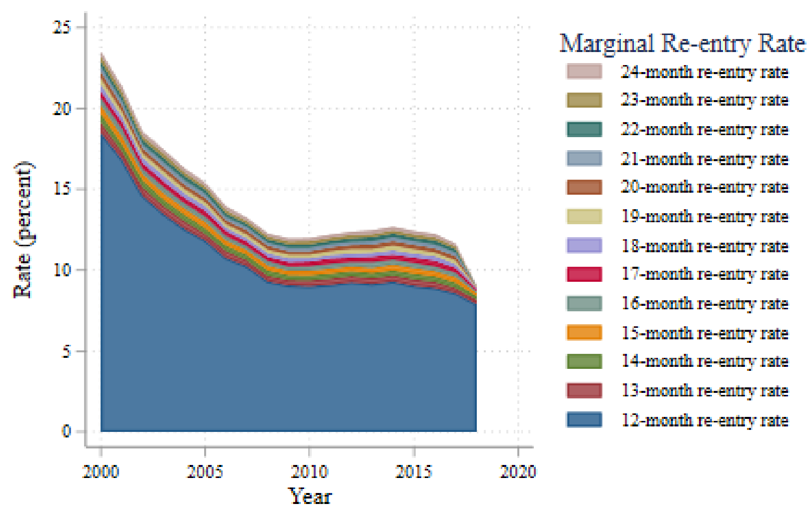
Figure A3.1. Histograms of months to reunification and months to foster care re-entry



Notes. Histograms of time (in months) between entering foster care to family reunification on the left and discharge from foster care and foster care re-entry on the right. Figures are based on child-level AFCARS data pooled across all states and years (2000-2019) with one observation per child-case. The x-axes show through 60 months (5 years).

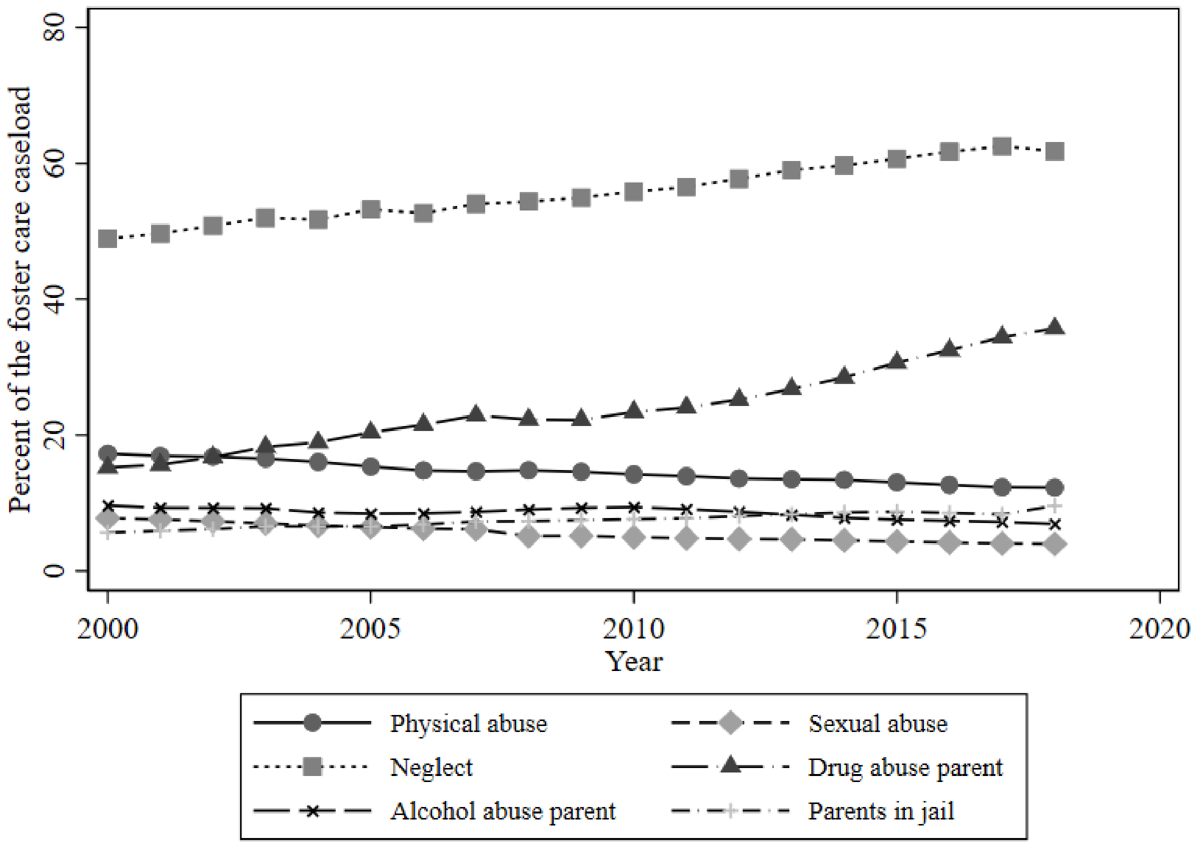
Figure A3.2. Marginal monthly re-entry rates





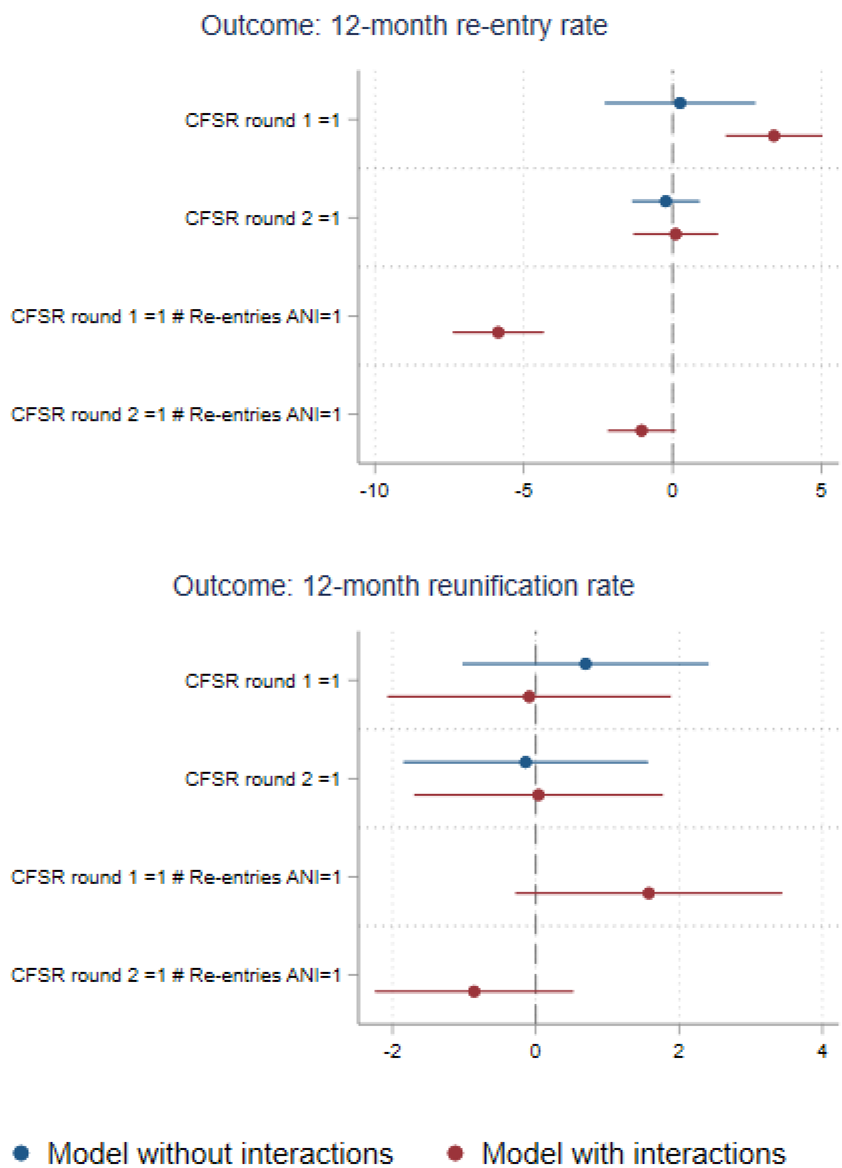
Notes. Average marginal re-entry rates are shown by year for 2000-2018. Each color shows the average marginal change in the re-entry rate when extending the observation period by an additional month. Note, these re-entry rates are calculated by changing the observation period in equation (1). To assess the cumulative re-entry rate over time, we sum up the relevant rate with the rates below. For example, for the 3-month re-entry rate, we add the marginal 1-month re-entry rate with the marginal 2-month and the marginal 3-month re-entry rates. The stacked area chart on top shows the marginal monthly re-entry rates for 1-12 months, whereas the stacked area chart below shows the marginal monthly re-entry rates for 12-24 months.

Figure A3.3. Trends in the average state foster care caseload by reasons for removal



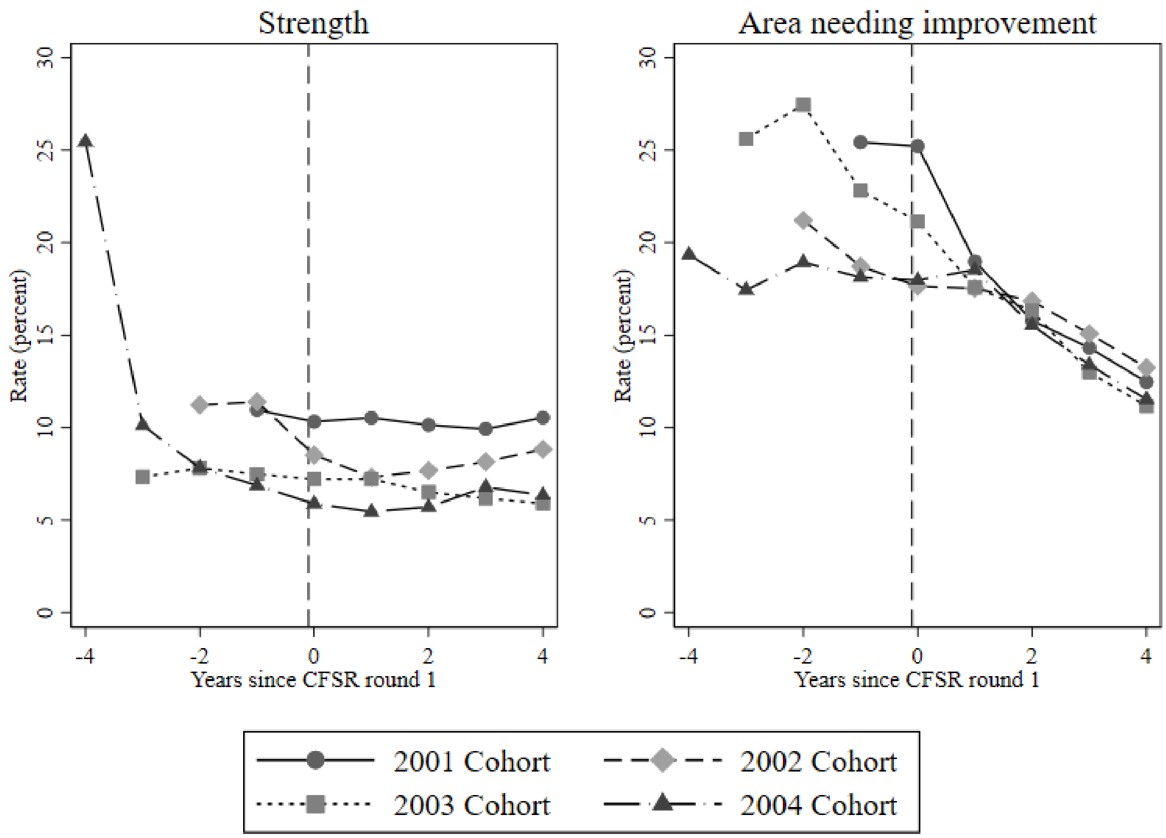
Notes. Foster care caseloads by reasons for removal are averaged across state-years for 2000-2018. These reasons for removal are not mutually exclusive and each case can have up to three reported reasons. For example, a child could have been removed from their household and placed in foster care due to neglect and drug abuse of their parent(s).

Figure A3.4. Coefficient plots showing estimates of the 12-month foster care re-entry rate and the 12-month reunification rate



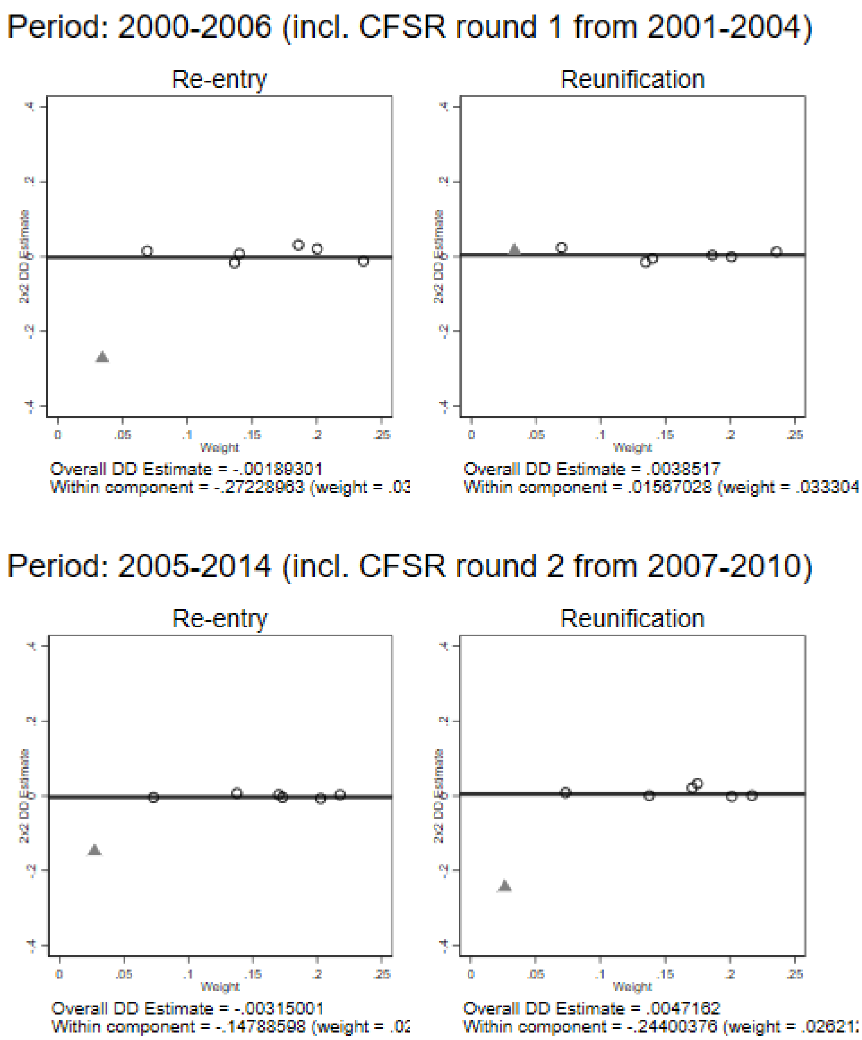
Notes. The main coefficients predicting the 12-month re-entry rate are shown (for columns 1 and 2 from appendix table A3.3) in the top coefficient plot. The main coefficients predicting the month reunification rate are shown (for columns 1 and 2 from appendix table A3.4) in the bottom coefficient plot. Bars indicate the 95 percent confidence intervals.

Figure A3.5. Trends in the average 12-month re-entry rates by strength or area needing improvement over year since CFSR round 1 by cohort treatment year



Notes. During CFSR round 1, each state identified that re-entries were either a 'strength' (i.e., re-entries were below federal standards) or an 'area needing improvement' (i.e., re-entries were above federal standards). Average rates are by the number of years since CFSR round 1 are shown separately: the left panel is averaged for states in which it was a 'strength' and in the right panel states in which it was an 'area needing improvement' have been averaged.

Figure A3.6. Difference-in-differences decomposition for CFSR rounds 1 and 2 (separately) and re-entry and reunification (separately)



Notes. The figure plots each 2x2 DD components against their weight for CFSR rounds 1 and 2 (separately) and re-entry and reunification (separately) using the Bacon (2018) decomposition theorem. The open circles are 'timing-group' terms, and the shaded triangles are 'within estimates' terms. Below each panel, I note the average DD estimate, as well as the 2x2 estimate and the weight on 'within component.' The two-way fixed effects estimate is shown in each panel as the horizontal line and equals the average of the y-axis values weighted by their x-axis value.

Table A3.1. Year of review for each CFSR round by state

State	Rounds 1	Round 2
Alabama	2002	2007
Alaska	2002	2008
Arizona	2001	2007
Arkansas	2001	2008
California	2002	2008
Colorado	2002	2009
Connecticut	2002	2008
Delaware	2001	2007
Florida	2001	2008
Georgia	2001	2007
Hawaii	2003	2009
Idaho	2003	2008
Illinois	2003	2009
Indiana	2001	2007
Iowa	2003	2010
Kansas	2001	2007
Kentucky	2003	2008
Louisiana	2003	2010
Maine	2003	2009
Maryland	2004	2009
Massachusetts	2001	2007
Michigan	2002	2009

Minnesota	2001	2007
Mississippi	2004	2010
Missouri	2004	2010
Montana	2002	2008
Nebraska	2002	2008
Nevada	2004	2009
New Hampshire	2003	2010
New Jersey	2004	2009
New Mexico	2001	2007
New York	2001	2008
North Carolina	2001	2007
North Dakota	2001	2008
Ohio	2002	2008
Oklahoma	2002	2007
Oregon	2001	2007
Pennsylvania	2002	2008
Rhode Island	2004	2010
South Carolina	2003	2009
South Dakota	2001	2008
Tennessee	2002	2008
Texas	2002	2008
Utah	2003	2010
Vermont	2001	2007
Virginia	2003	2009
Washington	2004	2010

West Virginia	2002	2008
Wisconsin	2003	2010
Wyoming	2002	2008

Table A3.2. Summary of key variables during review year by CFSR round

	Round 1		Round 2	
Variable	Mean	SD	Mean	SD
Percent re-entering within 12 months	15.11	9.91	9.65	5.35
Percent reunified in 12 months	22.99	8.68	24.00	43.14
Percent re-entries 'area needing improvement'	54.00	50.35	22.54	8.51
Caseload characteristics				
Ln foster care caseload	9.13	1.01	9.14	0.96
Average months in care before discharge	20.24	7.78	20.56	6.81
Percent non-white	39.72	19.01	35.81	17.27
Percent female	47.93	2.39	47.80	2.21
Percent clinical disability	19.60	12.50	23.68	12.42
Percent physical abuse	16.59	9.86	14.19	9.68
Percent sexual abuse	7.27	7.06	5.14	2.61
Percent neglect	51.41	20.82	54.22	20.91
Percent drug abuse parent	17.33	13.60	21.94	12.88
Percent alcohol abuse parent	9.15	8.78	8.88	10.51
Percent parents in jail	6.33	5.11	7.31	5.21
Economic characteristics				
Avg log household income	11.02	0.16	11.04	0.16
Percent below poverty threshold	12.60	3.15	15.61	2.95
Unemployment rate (age 25-65)	0.05	0.01	0.05	0.02
N state-year obs.	50		50	

Note. Averages across states are shown.

Table A3.3. Estimates of the foster care re-entry rates (percent), 2000-2018

			One-year lagged effect	
	(1)	(2)	(3)	(4)
CFSR Round 1	0.25	3.40**	-0.42	3.20***
	[-2.31, 2.90]	[0.77, 6.12]	[-2.51, 1.75]	[1.18, 5.24]
CFSR Round 2	-0.23	0.10	1.11	1.22
	[-1.37, 0.92]	[-1.28, 1.47]	[-1.62, 4.50]	[-1.50, 4.29]
CFSR Round 1 - re-entries ANI		12.58***		13.62***
		[6.83, 20.10]		[7.88, 19.90]
CFSR Round 2 - re-entries ANI		-9.40		-9.12
		[-21.9, 3.15]		[-21.7, 3.23]
Interaction: CFSR Round 1 * re-entries need improvement		-5.86***		-6.67***
		[-8.88, -2.82]		[-9.50, -3.82]
Interaction: CFSR Round 2 * re-entries need improvement		-1.04		-1.11
		[-4.09, 1.71]		[-4.10, 1.52]
Percent reunified in 12 months	0.07	0.08	0.07	0.09
	[-0.09, 0.22]	[-0.07, 0.23]	[-0.08, 0.23]	[-0.06, 0.23]

Ln foster care caseload	-4.91	-5.42*	-4.91	-5.36*
	[-12.40, 0.96]	[-13.4, 0.50]	[-12.40, 1.01]	[-13.40, 0.55]
Average months in care before discharge	-0.36***	-0.28**	-0.35***	-0.25**
	[-0.66, -0.09]	[-0.56, -0.05]	[-0.66, -0.08]	[-0.51, -0.04]
Percent non-white	-0.01	0.02	-0.01	0.03
	[-0.08, 0.07]	[-0.04, 0.09]	[-0.09, 0.07]	[-0.03, 0.09]
Percent female	-0.69***	-0.75***	-0.68***	-0.75***
	[-1.10, -0.20]	[-1.17, -0.28]	[-1.11, -0.20]	[-1.18, -0.27]
Percent clinical disability	0.03	0.03	0.03	0.03
	[-0.01, 0.06]	[-0.01, 0.06]	[-0.01, 0.06]	[-0.01, 0.06]
Percent physical abuse	-0.12	-0.12*	-0.12	-0.12*
	[-0.30, 0.06]	[-0.27, 0.01]	[-0.29, 0.07]	[-0.24, 0.01]
Percent sexual abuse	0.10	0.13	0.10	0.13
	[-0.18, 0.73]	[-0.11, 0.75]	[-0.18, 0.71]	[-0.11, 0.69]
Percent neglect	-0.00	-0.01	-0.00	-0.02
	[-0.04, 0.05]	[-0.05, 0.04]	[-0.04, 0.05]	[-0.06, 0.03]
Percent drug abuse parent	0.08	0.08	0.08	0.08
	[-0.06, 0.25]	[-0.05, 0.26]	[-0.06, 0.25]	[-0.05, 0.27]

Percent alcohol abuse parent	0.13	0.12	0.13	0.12
	[-0.06, 0.38]	[-0.04, 0.36]	[-0.06, 0.38]	[-0.04, 0.33]
Percent parents in jail	-0.01	-0.02	-0.01	-0.03
	[-0.34, 0.31]	[-0.36, 0.22]	[-0.35, 0.32]	[-0.37, 0.23]
Avg log household income	-3.63	-1.12	-4.41	-0.53
	[-30.70, 21.40]	[-27.10, 21.90]	[-31.30, 20.00]	[-26.00, 21.20]
Percent below poverty threshold	-0.55	-0.47	-0.57	-0.43
	[-1.35, 0.27]	[-1.29, 0.36]	[-1.36, 0.28]	[-1.27, 0.40]
Unemployment rate (age 25-65)	0.48**	0.36**	0.47**	0.23
	[0.07, 0.89]	[0.01, 0.73]	[0.09, 0.86]	[-0.09, 0.54]
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1000	1000	1000	1000
Adjusted R-squared	0.711	0.729	0.712	0.744

Notes. The dependent variable is the 12-month state-year re-entry rate. CFSR rounds 1 and 2 are binary indicators for whether the state has received that round of review. Columns 2 and 4 include interactions for whether the state has undergone that round of the CFSR and whether re-entries were identified as a strength or an area needing improvement that round. Wild-cluster bootstrapped t-statistics to create confidence intervals (Cameron et al., 2008; MacKinnon et al., 2022; Roodman, 2022)/ 95% confidence intervals in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table A3.4. Estimates of the 12-month reunification rate (percent), 2000-2018

			One-year lagged effect	
	(1)	(2)	(3)	(4)
CFSR Round 1	0.70	-0.09	-0.23	-0.98
	[-2.38, 3.99]	[-3.29, 3.47]	[-3.04, 2.55]	[-3.56, 1.61]
CFSR Round 2	-0.14	0.04	-1.54	-1.26
	[-1.83, 1.67]	[-1.49, 1.65]	[-3.53, 0.58]	[-3.39, 0.91]
CFSR Round 1 - re-entries ANI		4.58		3.75
		[-6.91, 15.40]		[-9.72, 15.80]
CFSR Round 2 - re-entries ANI		8.74*		8.34
		[-1.72, 19.50]		[-2.66, 19.50]
CFSR Round 1 * re-entries need improvement		1.58		1.54
		[-2.17, 5.26]		[-2.04, 5.05]
CFSR Round 2 * re-entries need improvement		-0.86		-1.15
		[-3.47, 1.85]		[-3.70, 1.47]
Ln foster care caseload	2.82*	2.79*	2.75	2.65*
	[-0.44, 6.34]	[-0.34, 6.16]	[-0.51, 6.25]	[-0.44, 6.03]

Average months in care before discharge	-0.48***	-0.49***	-0.48***	-0.50***
	[-0.77, -0.20]	[-0.78, -0.20]	[-0.78, -0.20]	[-0.79, -0.19]
Percent non-white	-0.07	-0.07	-0.07	-0.08
	[-0.19, 0.07]	[-0.20, 0.07]	[-0.20, 0.07]	[-0.20, 0.07]
Percent female	-0.92	-0.92	-0.93	-0.92
	[-1.91, 0.25]	[-1.92, 0.24]	[-1.92, 0.25]	[-1.93, 0.24]
Percent clinical disability	-0.02	-0.02	-0.01	-0.02
	[-0.06, 0.03]	[-0.06, 0.03]	[-0.06, 0.04]	[-0.06, 0.03]
Percent physical abuse	0.17**	0.18**	0.18***	0.18***
	[0.026, 0.43]	[0.032, 0.45]	[0.035, 0.43]	[0.042, 0.46]
Percent sexual abuse	-0.05	-0.05	-0.05	-0.05
	[-0.26, 0.24]	[-0.30, 0.27]	[-0.26, 0.24]	[-0.29, 0.25]
Percent neglect	-0.04	-0.04	-0.04	-0.04
	[-0.11, 0.01]	[-0.11, 0.01]	[-0.11, 0.01]	[-0.11, 0.01]
Percent drug abuse parent	-0.05	-0.05	-0.05	-0.05
	[-0.19, 0.07]	[-0.18, 0.07]	[-0.19, 0.07]	[-0.18, 0.07]
Percent alcohol abuse parent	-0.08	-0.08	-0.08	-0.08
	[-0.26, 0.06]	[-0.26, 0.07]	[-0.26, 0.06]	[-0.26, 0.07]

Percent parents in jail	-0.11	-0.10	-0.11	-0.11
	[-0.33, 0.20]	[-0.31, 0.19]	[-0.33, 0.20]	[-0.31, 0.19]
Avg log household income	19.10	19.74	19.49	20.16
	[-5.75, 41.40]	[-6.08, 41.30]	[-5.81, 41.60]	[-5.60, 41.80]
Percent below poverty threshold	0.01	0.03	-0.01	0.01
	[-0.74, 0.76]	[-0.70, 0.74]	[-0.75, 0.74]	[-0.72, 0.73]
Unemployment rate (age 25-65)	0.46	0.47	0.48	0.51
	[-0.35, 1.26]	[-0.35, 1.20]	[-0.32, 1.29]	[-0.31, 1.26]
State fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1000	1000	1000	1000
Adjusted R-squared	0.722	0.723	0.723	0.724

Notes. The dependent variable is the 12-month state-year reunification rate. CFSR rounds 1 and 2 are binary indicators for whether the state has received that round of review. Columns 2 and 4 include interactions for whether the state has undergone that round of the CFSR and whether re-entries were identified as a strength or an area needing improvement that round. Wild-cluster bootstrapped t-statistics to create confidence intervals (Cameron et al., 2008; MacKinnon et al., 2022; Roodman, 2022). 95% confidence intervals in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level.

Table A3.5. Estimates of the 12-month foster care re-entry rate (percent) by age-group at reunification, 2000-2018

	(1) 0-17 years old	(2) Under 1 year old	(3) 1-2 years old	(4) 3-5 years old	(5) 6-11 years old	(6) 12-13 years old	(7) 14-17 years old
CFSR Round 1	3.40**	0.04	2.80***	2.76**	2.85**	3.07*	4.09**
	[0.77, 6.12]	[-2.91, 3.15]	[0.68, 5.04]	[0.22, 5.48]	[0.12, 5.78]	[-0.55, 6.60]	[0.42, 7.83]
CFSR Round 2	0.10	0.26	0.12	0.63	-0.08	-0.72	-0.11
	[-1.28, 1.47]	[-1.72, 2.36]	[-1.65, 1.89]	[-0.60, 1.89]	[-1.28, 1.08]	[-2.18, 0.70]	[-2.73, 2.44]
CFSR Round 1 - re-entries ANI	12.58***	8.53**	8.22***	11.84***	10.91***	15.53***	19.61***
	[6.83, 20.10]	[2.34, 16.30]	[3.96, 12.70]	[5.59, 18.50]	[4.70, 18.30]	[7.71, 26.00]	[12.10, 29.80]
CFSR Round 2 - re-entries ANI	-9.40	-0.24	-13.94*	-10.65	-7.01	-13.56**	-9.07
	[-21.90, 3.15]	[-12.20, 12.00]	[-34.90, 1.42]	[-28.10, 3.18]	[-19.50, 3.84]	[-25.90, -0.58]	[-22.20, 4.36]
Interaction: CFSR Round 1 * re-entries need improvement	-5.86***	-1.98	-2.44	-4.34***	-4.69***	-7.79***	-7.66***
	[-8.88, -2.82]	[-5.19, 1.24]	[-5.61, 0.97]	[-7.34, -1.25]	[-7.46, -1.78]	[-12.20, -3.28]	[-11.80, -3.24]
Interaction: CFSR Round 2 * re-entries need improvement	-1.04	-1.09	-0.39	-0.10	-0.35	-0.38	-1.04
	[-4.09, 1.71]	[-4.03, 1.35]	[-2.51, 1.53]	[-2.88, 2.49]	[-2.87, 1.99]	[-4.00, 2.84]	[-5.23, 2.74]
Percent reunified in 12 months	0.08	-0.10*	0.11	0.03	-0.00	0.04	0.08
	[-0.07, 0.23]	[-0.19, 0.00]	[-0.04, 0.28]	[-0.12, 0.19]	[-0.12, 0.12]	[-0.12, 0.20]	[-0.11, 0.25]

Ln foster care caseload	-5.42*	-2.46	-7.20*	-6.20	-4.32	-4.85	-7.05**
	[-13.40, 0.50]	[-5.90, 0.57]	[-19.20, 0.45]	[-16.80, 0.83]	[-11.60, 1.03]	[-11.4, 0.86]	[-14.0, -0.70]
Average months in care before discharge	-0.28**	0.02	-0.13	-0.24*	-0.25**	-0.29**	-0.30**
	[-0.56, -0.05]	[-0.18, 0.20]	[-0.40, 0.12]	[-0.54, 0.04]	[-0.51, -0.03]	[-0.62, -0.03]	[-0.63, -0.01]
Percent non-white	0.02	-0.05	0.02	0.05	0.05	0.04	0.05
	[-0.04, 0.09]	[-0.30, 0.16]	[-0.06, 0.10]	[-0.02, 0.14]	[-0.01, 0.12]	[-0.04, 0.12]	[-0.06, 0.16]
Percent female	-0.75***	-0.39	-0.41***	-0.19	-0.25	-0.72***	-0.84
	[-1.17, -0.28]	[-0.98, 0.07]	[-0.92, -0.09]	[-0.65, 0.15]	[-0.67, 0.05]	[-1.22, -0.32]	[-1.60, 0.22]
Percent clinical disability	0.03	0.03**	0.00	0.03*	0.03*	0.05***	0.05
	[-0.00, 0.06]	[0.00, 0.06]	[-0.03, 0.03]	[-0.00, 0.07]	[-0.00, 0.07]	[0.01, 0.09]	[-0.03, 0.15]
Percent physical abuse	-0.12*	-0.06	-0.15**	-0.12**	-0.13*	-0.09	-0.12
	[-0.27, 0.01]	[-0.51, 0.38]	[-0.24, -0.03]	[-0.23, -0.00]	[-0.22, 0.01]	[-0.30, 0.11]	[-0.34, 0.06]
Percent sexual abuse	0.13	0.01	0.07	0.08	0.15	0.19	0.16
	[-0.11, 0.75]	[-0.54, 0.69]	[-0.18, 0.57]	[-0.16, 0.52]	[-0.08, 0.65]	[-0.14, 1.21]	[-0.14, 0.96]
Percent neglect	-0.01	-0.01	0.01	-0.00	-0.01	-0.02	-0.06
	[-0.05, 0.04]	[-0.06, 0.03]	[-0.02, 0.04]	[-0.05, 0.06]	[-0.06, 0.06]	[-0.08, 0.05]	[-0.12, 0.01]
	0.08	0.01	0.11	0.12	0.03	0.04	0.09

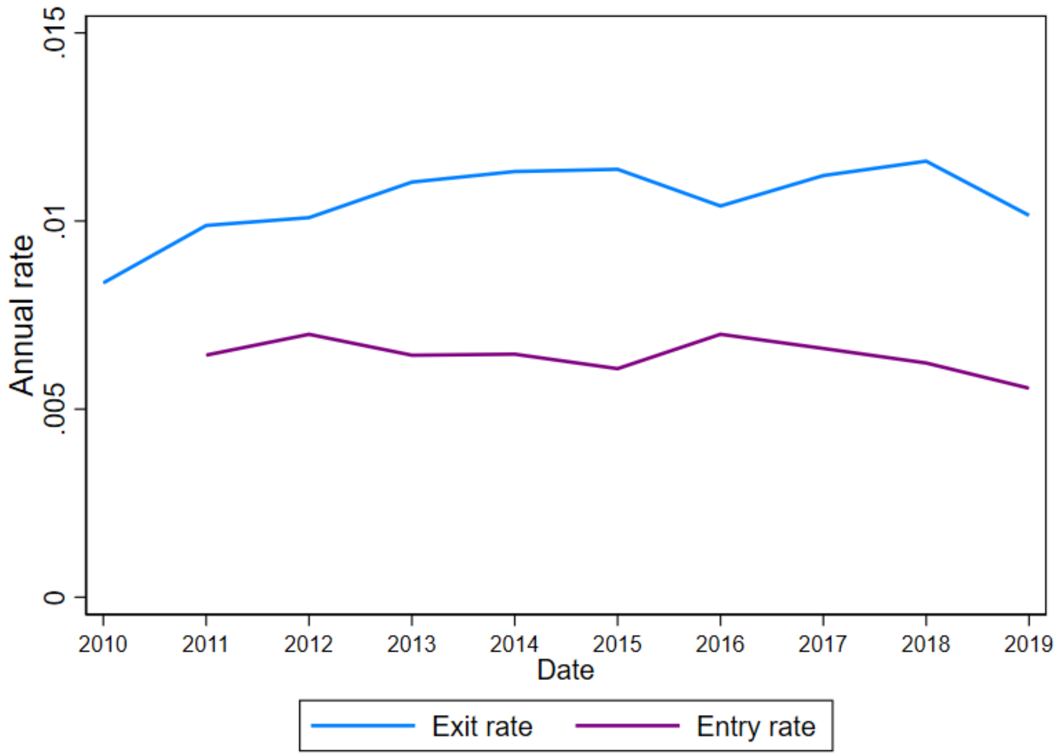
Percent drug abuse parent	[-0.05, 0.26]	[-0.09, 0.12]	[-0.06, 0.40]	[-0.05, 0.39]	[-0.12, 0.21]	[-0.11, 0.22]	[-0.05, 0.26]
Percent alcohol abuse parent	0.12	0.11	0.11	0.07	0.10	0.12	0.15
	[-0.05, 0.36]	[-0.15, 0.39]	[-0.03, 0.27]	[-0.10, 0.26]	[-0.11, 0.29]	[-0.16, 0.55]	[-0.05, 0.49]
Percent parents in jail	-0.02	-0.19**	-0.10	-0.15*	-0.07	-0.03	0.16
	[-0.36, 0.22]	[-0.51, -0.01]	[-0.48, 0.12]	[-0.58, 0.00]	[-0.49, 0.12]	[-0.50, 0.49]	[-0.32, 0.37]
Avg log household income	-1.12	-13.78	-1.74	-0.91	2.85	-5.92	-7.93
	[-27.10, 21.90]	[-36.80, 17.50]	[-32.30, 21.40]	[-33.80, 23.40]	[-28.70, 28.80]	[-46.90, 29.20]	[-37.10, 21.20]
Percent below poverty threshold	-0.47	-0.18	-0.36	-0.36	-0.08	-0.71	-0.59
	[-1.29, 0.36]	[-1.19, 0.94]	[-1.22, 0.50]	[-1.23, 0.53]	[-0.92, 0.79]	[-1.90, 0.39]	[-1.68, 0.55]
Unemployment rate (age 25-65)	0.36**	-0.28	0.45**	0.35*	0.33	0.47*	0.09
	[0.01, 0.73]	[-0.99, 0.37]	[0.03, 0.84]	[-0.06, 0.75]	[-0.11, 0.73]	[-0.05, 0.99]	[-0.46, 0.64]
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1000	998	1000	1000	1000	1000	1000
Adjusted R-squared	0.729	0.153	0.460	0.541	0.556	0.607	0.754

Notes. The dependent variable is the 12-month foster care re-entry rate. CFSR rounds 1 and 2 are binary indicators for whether the state has received that round of review. Wild-cluster bootstrapped t-statistics to create confidence intervals (Cameron et al., 2008;

MacKinnon et al., 2022; Roodman, 2022). 95% confidence intervals in brackets. Significance levels: * 10 percent level, ** 5 percent level, *** 1 percent level

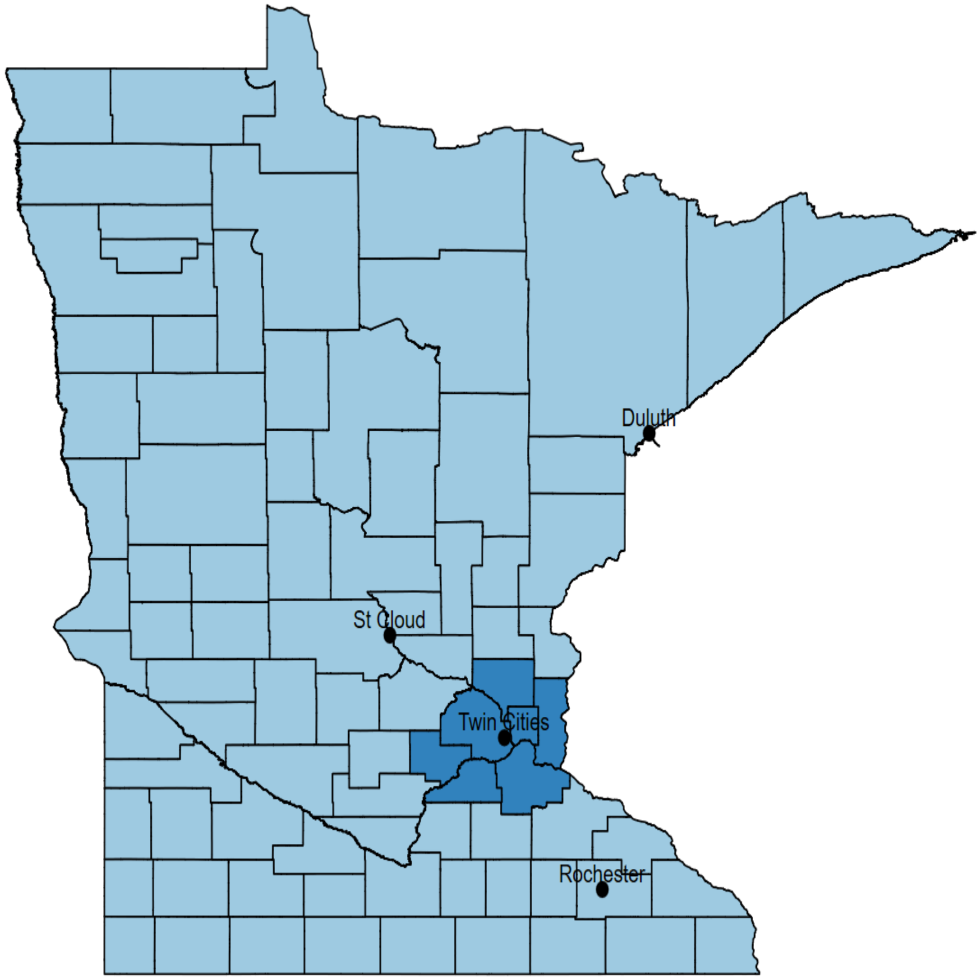
Appendix to Chapter 4

Figure A4.1. The annual family child care provider rate of exit and entry



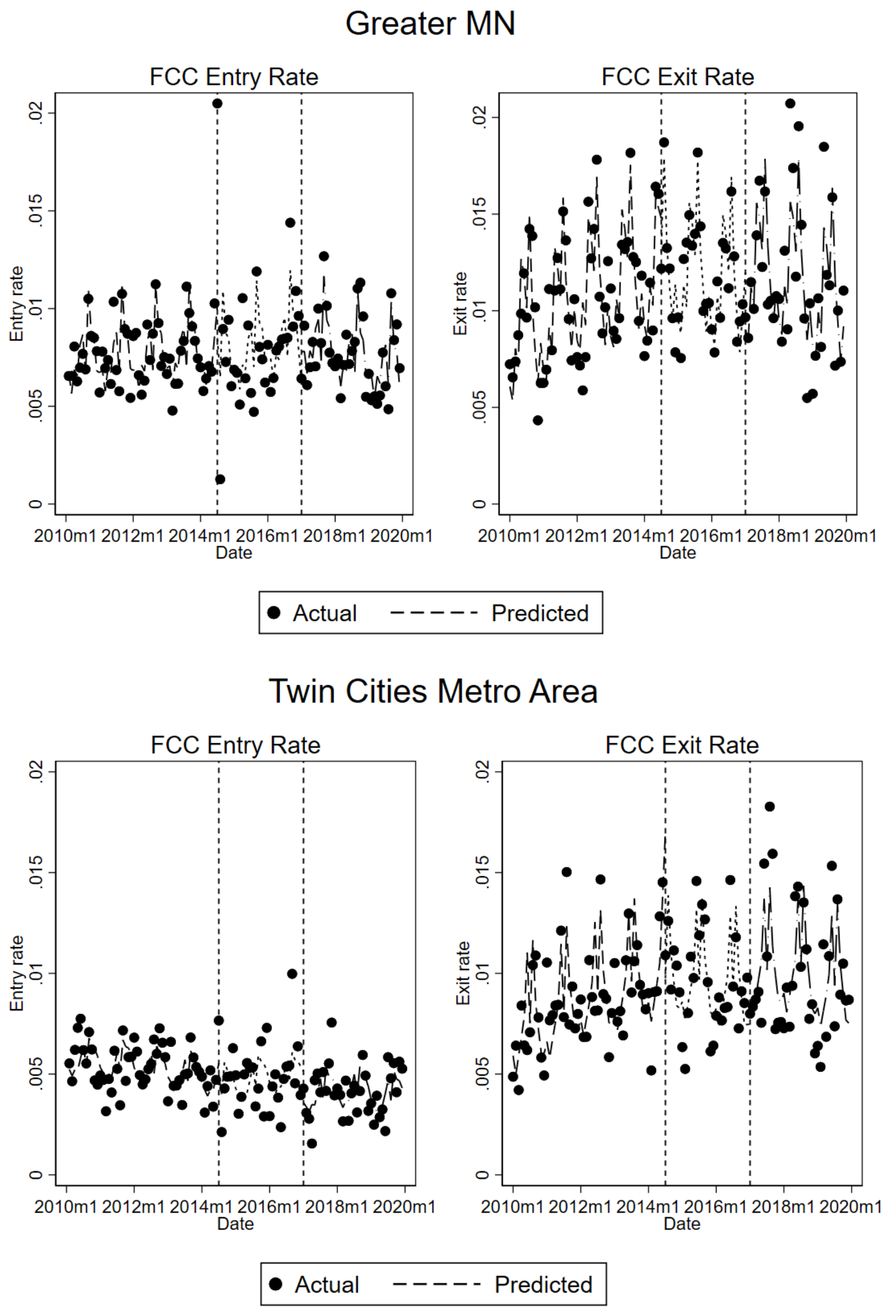
Note. Minnesota statewide FCC entry and exit rates are shown annually for 2010-2019.

Figure A4.2. Map of Minnesota



Note. There are 87 counties in Minnesota. The darker blue shows the Twin Cities (Minneapolis-St. Paul) metro area given the seven-county definition. The lighter blue shows greater Minnesota (all counties outside of the seven-county metro area).

Figure A4.3. Interrupted time series of the FCC entry rate and exit rate by region



Notes. FCC entry and exit rates are shown separately for greater Minnesota and the Twin Cities (Minneapolis-St. Paul) metro area. There are two sets of regulations represented by dashed vertical lines: health and safety regulations (July 2014) and background check regulations (January 2017). Controls for monthly state-level seasonally adjusted unemployment. Prais-Winsten regression coefficients are seen in Appendix Table B4.1.

Table B4.1. Interrupted time series of the FCC entry rate and exit rate by region

	Greater Minnesota		Minneapolis-St. Paul Metro Area	
	Entry rates	**Exit rates**	**Entry rates**	**Exit rates**
Intercept	-0.0002412	0.03451	0.0063703	0.0147874
	(0.0098012)	(0.0104577)	(0.0069165)	(.011078)
Time(β1)	0.0000685	-0.0001212	-0.0000298	-4.25E-06
	(0.0000865)	(0.0000916)	(0.0000611)	(0.0000971)
Intervention one (β2) (Change in level)	-0.0002314	-0.001074	-0.0001432	-0.0006412
	(0.0007741)	(0.0008204)	(0.0005478)	(0.000869)
Time After Int. (β3) (Change in slope)	-0.0000369	0.000086	0.0000258	-0.0000149
	(0.0000852)	(0.00009)	(0.0000601)	(0.0000953)
Intervention two (β4) (Change in level)	0.0005534	-0.0005173	-0.0003805	0.0010003
	(0.0010855)	(0.0011409)	(0.0007639)	(0.0012087)
Time After Int. (β5) (Change in slope)	-0.0000975**	-0.0000492	-6.35E-06	-0.000046
	(0.0000434)	(0.0000467)	(0.0000309)	(0.0000494)
N	120	120	120	120
R-squared	0.44	0.77	0.37	0.61
Durbin–Watson statistic	2.24	1.95	2.10	1.96

Notes. Prais-Winsten regression controls for monthly state-level seasonally adjusted unemployment. Standard errors in parentheses. *$p < 0.10$. **$p < 0.05$. ***$p < 0.01$.