Reforming Child Protection Services:

The Effect of Differential Response on Subsequent Maltreatment for Reported Children

([Click here for the most recent version](#))

Caitlyn Keo*

October 31, 2023

Abstract

Traditionally, *all* 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 for 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.

Key Words: Child Welfare, Child Protection Investigations, Reforms, Public Policy

JEL Codes: H00, J13, J18

1. Introduction

Traditionally, child protection services (CPS) agencies respond to all 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. 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 and 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. 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 AR. Ceteris paribus, race and ethnicity also plays 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. 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.

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 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

---

[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).

[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.

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 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, 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 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.

---

[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 been found in Appendix D.

[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.

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.

3. Data and Descriptive Analysis

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

---

[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.

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 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 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 which years each state reports into the national system.

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,

---

[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.

[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.

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).

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 state-year log of the child population, 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.

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

State-level treatment 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]

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.,

---

[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 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 C1 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 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.

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 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.

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 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 C3. 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).

4. Empirical Strategy

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 aggregate 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 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 tends 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).

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 C4).[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.

5. Results

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 4 shows the corresponding event study with 5

---

[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.

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 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.

Since DR policies are intended to allow CPS to work with lower risk families by providing 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 most significant 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].

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. DR is a particularly interesting, existing policy

because it 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 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.

5.3 Heterogeneous Effects Across Baseline Characteristics: Reason for the Report

A CPS maltreatment report could list up to four reasons for the allegation. The treatment effects are disaggregated by the reason for the maltreatment report in Table 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.

15

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. Child 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].

5.4 Robustness Checks

Robustness to the Choice of Estimator

Appendix Table B1 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 B1 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. Primarily, the awareness is that 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 requires 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 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 B3 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 B4. 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 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).

---

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

6. Summary and Concluding Remarks

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.

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).

6.3 Job Market Paper Note: Other Forthcoming Work on Differential Response

Whereas all reports are traditionally investigated, a key premise of this study is that DR reforms introduce a new step in decision making for whether to place the report on the AR track or the IR track. Evaluating the effect of the IR track or the AR track on safety as measured by re-reporting outcomes is complicated due to selection. Notably, both the outcome (a re-report) and the child-level treatment are influenced by a child's risk and other possible confounders. The unobserved differences in characteristics of higher risk reports and lower risk reports should be critically considered. These unobservable confounders may not be directly known by the agency at the time of decision making, or the researcher in retrospective analysis. The seemingly simple decision-making scenario turns out to be surprisingly complex, with important child welfare and policy implications.

An empirical approach that models endogenous selection mechanisms and accounts for a reported child's characteristics is hereby needed. With these considerations, in a separate paper, I assess whether there is an association between CPS decision-making and use of AR and IR tracks and subsequent maltreatment outcomes. See Appendix E for a theoretical model for child welfare decision making.

References

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

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).

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*, *15*(4), 1345–1372.

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.

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

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

Callaway, B., & Sant'Anna, P. H. C. (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.

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*.

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*.

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

de Chaisemartin, C., & D'Haultfœuille, 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

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?*

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

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

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.

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.

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

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.

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*.

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

Paxson, C., & Waldfogel, J. (1999). *Parental Resources and Child Abuse and Neglect*. *89*(2), 239–244.

Paxson, C., & Waldfogel, J. (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

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.

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). *bootest: Stata Module To Provide Fast Execution Of The Wild Bootstrap With Null Imposed*. Stata Command.

Roodman, D., MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2019). Fast and Wild: Bootstrap Inference in Stata using Boottest. *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). IPUMS USA: Version 9.0 [Dataset]. In *Minneapolis, MN: IPUMS*.

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

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*.

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

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.

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.*

FIGURES

Figure 1. Types of Child Protection Services (CPS) structure
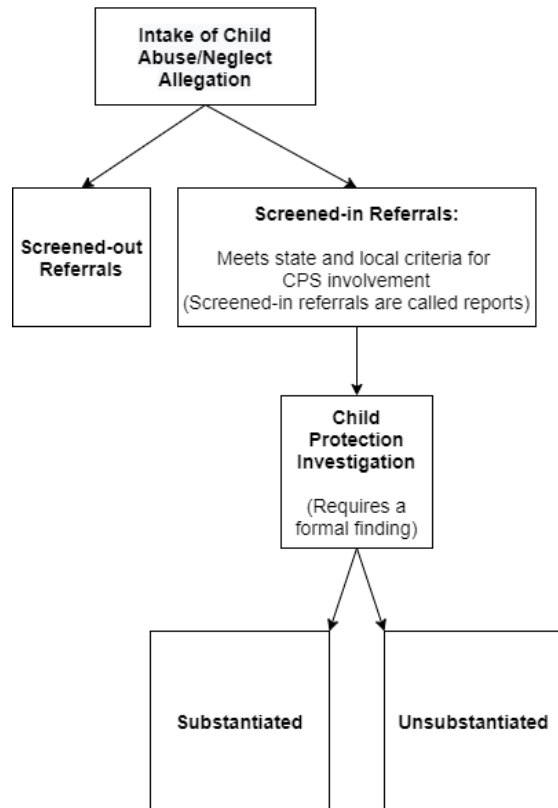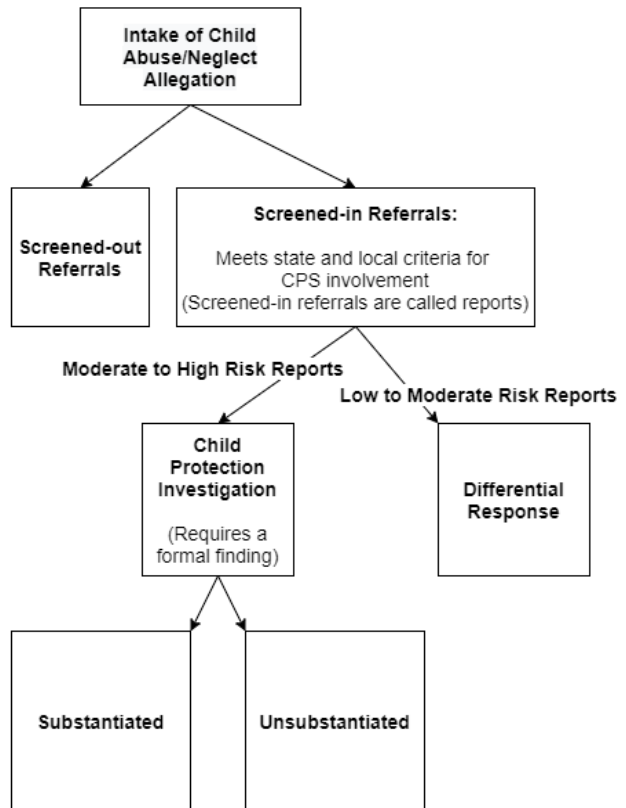
Notes. Author's creation. Panel A is a traditional one-track investigative CPS system. Panel B is a DR system that has two tracks, and *only* makes the track determination after a report is screened in.

Figure 2. Trends in re-reporting outcomes by year of the report
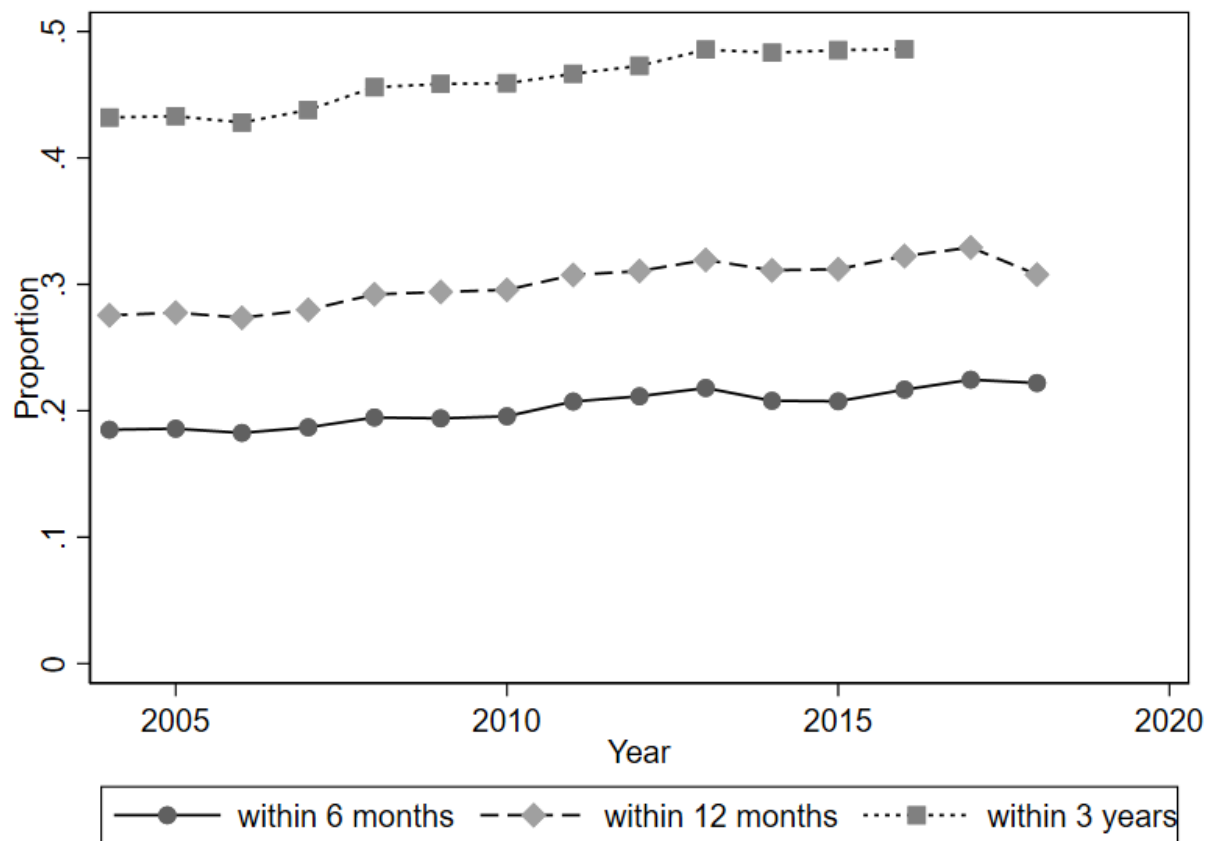
Notes. For my analytic sample of maltreatment reports, I show the proportion of report-child observations that are re-reported each year given different intervals of time.

Figure 3. Predictors of an alternative response in periods with differential response
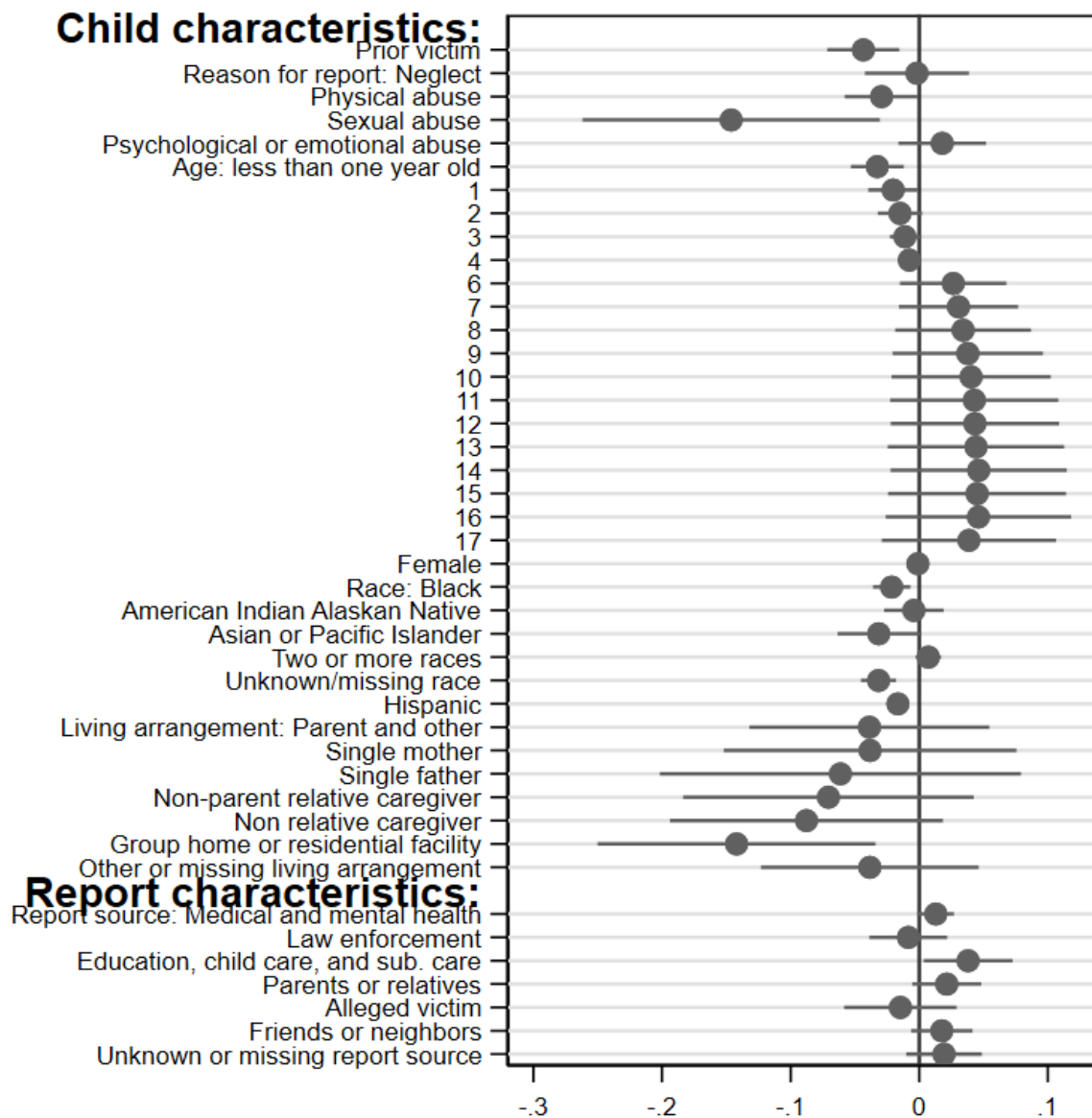
Notes. In this coefficient plot, the circles indicate point estimates and bars indicate the 95 percent confidence intervals. Omitted reference groups which are not shown include: non-prior victim, no neglect, no physical abuse, no sexual abuse, no psychological or emotional abuse, age 5, race=White, 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 C3.

Figure 4. Event studies estimating the effects of differential response reforms on re-reporting outcomes
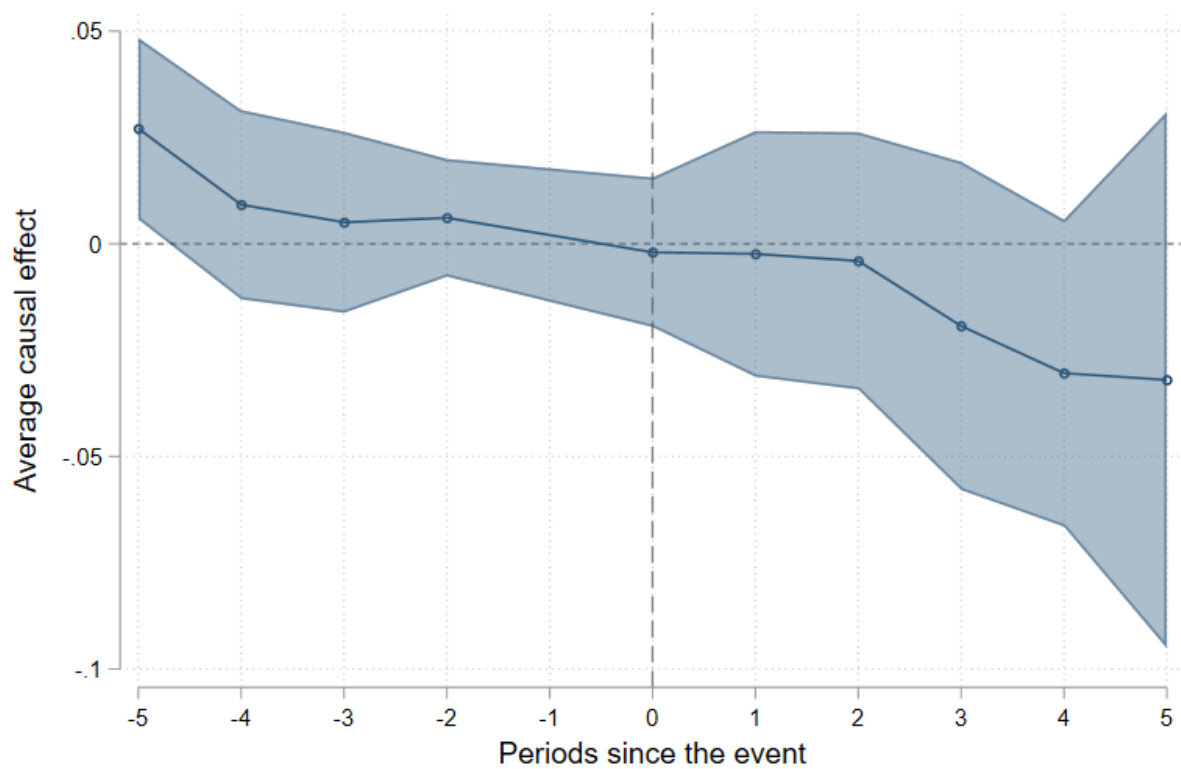
Notes. Unconditional Callaway and Sant'Anna event study with never treated comparison group. Year t-1 is omitted. Bars indicate the 95 percent confidence intervals.

Table 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. 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 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

APPENDICES

Appendix A. 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 B. Robustness Checks

Table B1. 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. 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 B3. 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 B4. 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.

Figure C1. Differential response implementation across U.S. states as of 2019
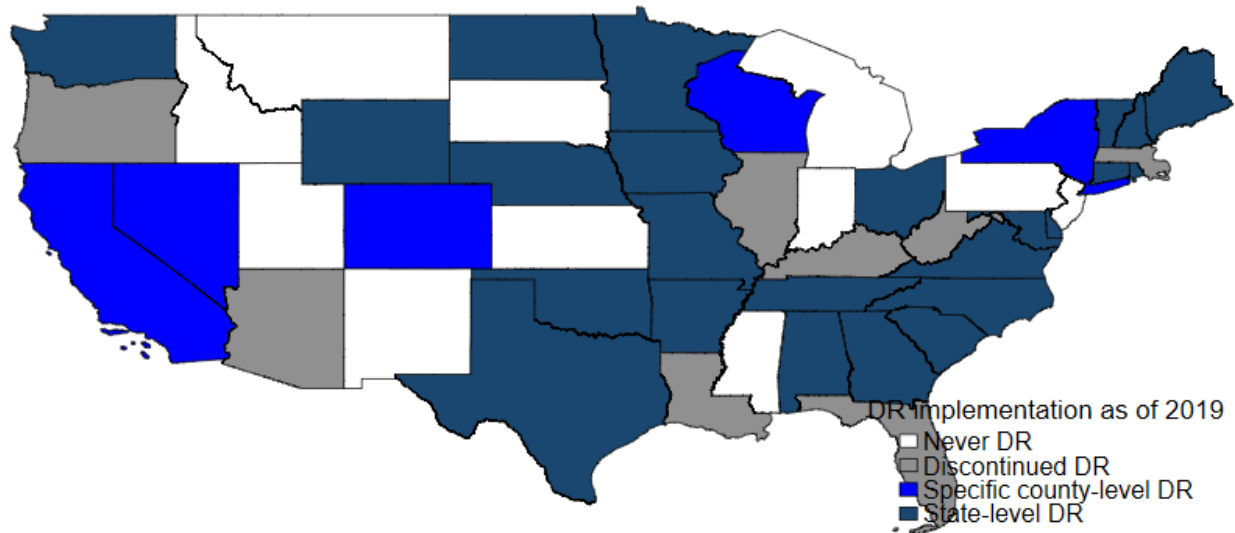
Notes. Differential response is a popular CPS reform. The states with statewide and specific county-level DR in 2019 were identified using the SCAN Policies Database 2019 (Weigensberg et al., 2022). The states that previously implemented DR but discontinued it were identified using other sources. By 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. Differential response treatment status by year for states in my analytic sample
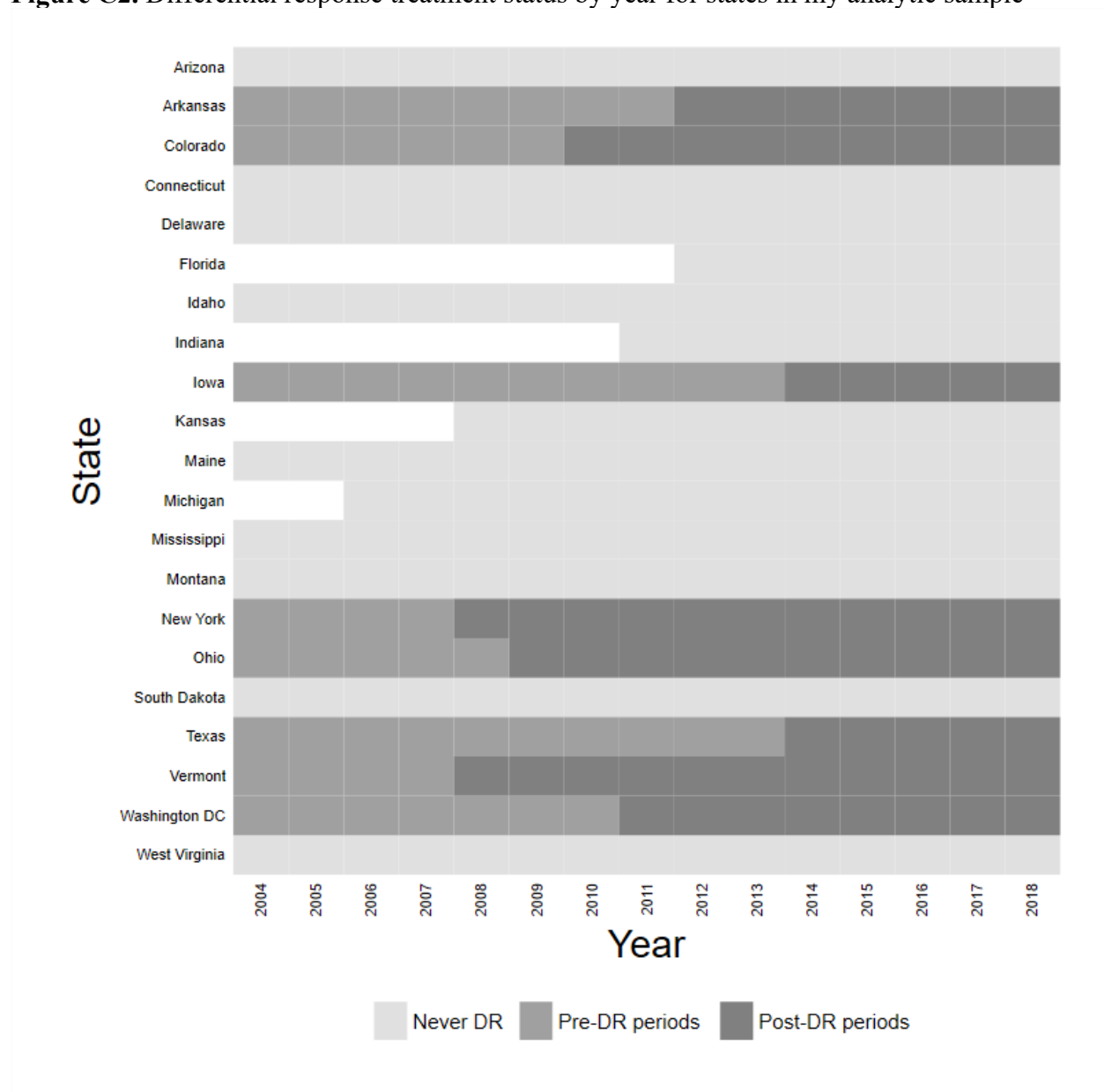
Notes. There are 21 states in the sample, and 8 of these states adopt DR between 2004 and 2019. These are specifically states that do not discontinue the use of DR before 2019, and states that make the track determination (AR or IR) after a report is screened in. The white cells indicate state-year pairs that are dropped from the sample (See Appendix A).

Figure C3. Number of annual reports-child observations by reasons for report
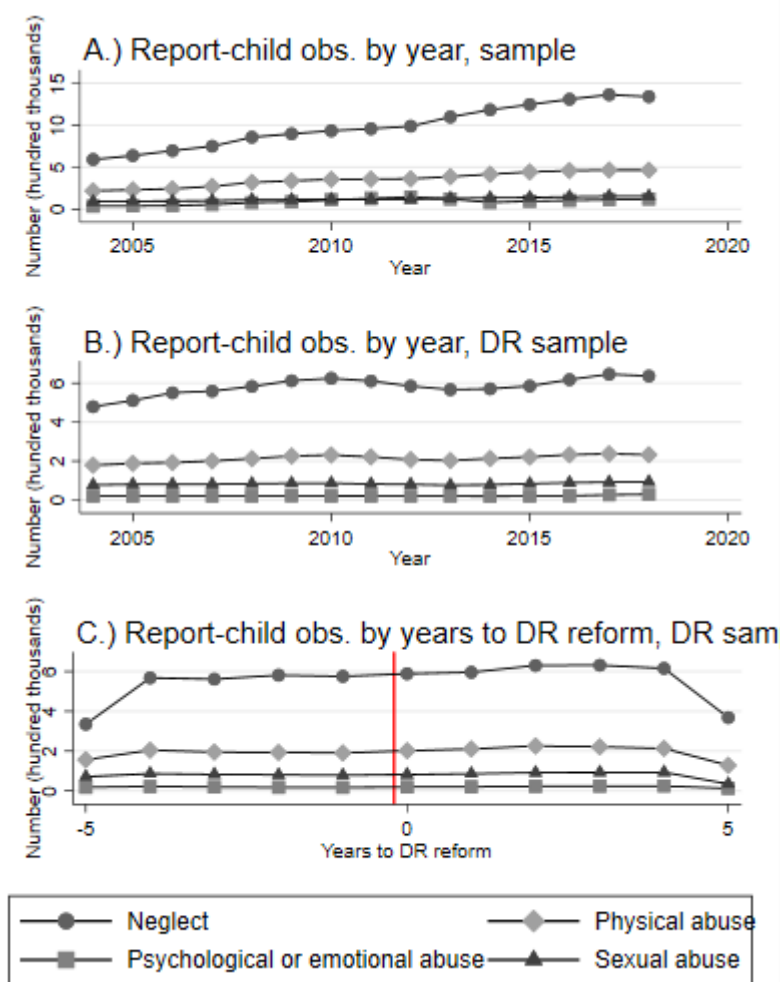
Notes. In panel A, the number of annual report-child observations are pooled across all sampled states (N=21 states) by the reason for the report. In panels B and C, the annual sums are pooled across the 8 states that adopt DR between 2004 and 2019 and do not discontinue in this time frame. In panel C, year 0 represents the relative year to DR reform.

Figure C4. Proportion of reports-child observations that are re-reported by year to DR reform

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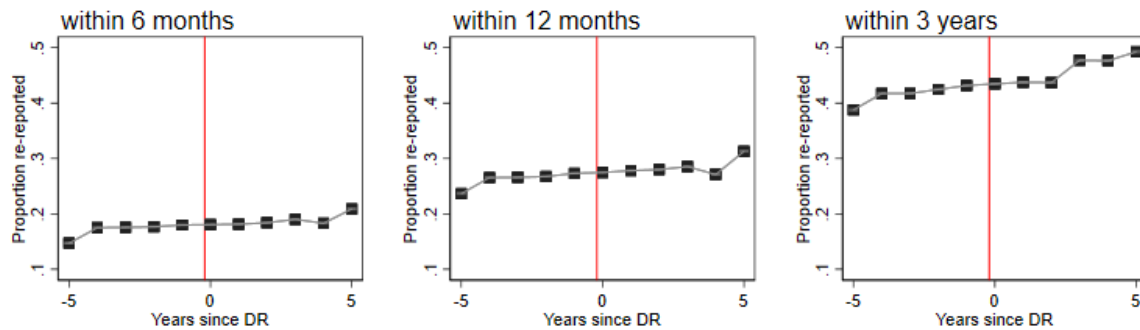
B.) Proportion re-reported with substantiation
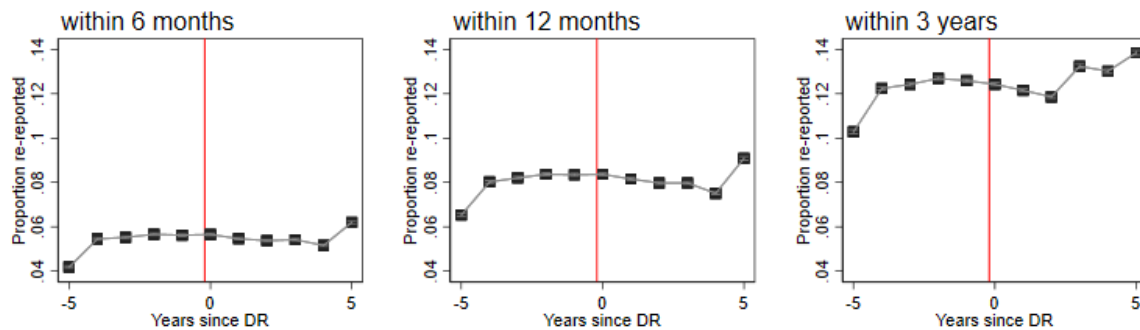
Notes. The average proportions of reports-child observations that are re-reported are pooled across the 8 states by the relative year to DR reform. These 8 states adopt DR between 2004 and 2019 and do not discontinue in this time frame.

Table C1. Characteristics of report-child observations in states that adopt DR policies between 2004 and 2019

	Pre-DR reform IR pathway (1)		Post-DR reform IR *or* AR pathway (2)		Post-DR reform IR pathway (3)		Post-DR reform AR pathway (4)		Difference: (2) - (1) (5)			Difference: (4) - (3) (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)

	(1)		(2)		(3)		(4)		(5)			(6)		
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. Characteristics of report-child observations that are <u>investigated</u> in state-years with no DR reform (conventional investigative CPS structure) <u>and</u> state-years that are post-DR reform (CPS with DR reform)

	No DR reform: *Only IR* (Pre-DR *or* never DR reform) (1)		**Post-DR reform**: *Selected to be IR* (2)		**Difference: (2) - (1)** (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)

	(1)		(2)		(3)		
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 C3. 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 C4. 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 C5. Alternative response utilization rates by state-year

State	Year of the report														
	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
Idaho	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Indiana										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
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 D. 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 E. Differential Response and Modelling the New Step in Child Welfare Decision Making

CPS agencies respond to reports of child maltreatment based on some (formal or informal) assessment of risk. To provide a simple characterization of a reported child's risk, I denote 'true' maltreatment risk with the term σ, where $\sigma \geq 0$. Then, I define the 'perceived' risk from the maltreatment report that the CPS agency receives as:

$$s = \sigma + \mu, \tag{1}$$

where μ represents idiosyncratic noise from the initial report. Intuitively, the perceived risk assessment is uncertain. When $\mu \neq 0$, the perceived risk (s) may be greater than or less than the true risk (σ) depending on the direction of μ.

Traditionally, the status quo is that *all* CPS reports are investigated regardless of perceived risk for future maltreatment. This is true in all periods without DR reforms.[17] However, in periods reformed for DR (henceforth referred to as a post-DR period), the jurisdiction has a new binary choice problem. Let D_{IR} denote a binary decision variable indicating whether a reported child is investigated. Specifically, I use a standard binary decision setup and I write D_{IR} as:

$$D_{IR} = \begin{cases} 1 & \text{the child is placed on the IR track, and} \\ 0 & \text{the child is placed on the AR track.} \end{cases}$$

Whereas it is always true that $D_{IR} = 1$ in periods without DR reforms, in post-DR reform periods, the perceived risk index (which is a function of true risk) drives the decision to investigate. I write the selection equation as:

$$D_{IR} = 1[s \geq 0]. \tag{2}$$

In equation (2), agencies the post-DR period will generally choose to place the child on the IR track ($D_{IR} = 1$) if the perceived risk, s, is considered of moderate to high risk. Otherwise, the agency will choose to place the child on the AR track ($D_{IR} = 0$). Equation (2) models (self-) selection into the IR and AR tracks in a way that is inspired by a Roy model (Heckman & Honoré, 1990; Roy, 1951). Endogeneity exists in this context because the severity of a child's report and their presumed risk is expected to be related to both CPS decisions (e.g., track placement) and the possibility of subsequent re-reporting.

---

[17] Periods without DR reforms include all periods for jurisdictions that never reform for DR, and pre-DR periods for jurisdictions that eventually reform for DR.