Effects of Enhanced Legal Aid in Child Welfare: Evidence from a Randomized Trial of Mi Abogado*

Ryan Cooper† Joseph J. Doyle‡ Andrés Hojman§

December 16, 2024

Abstract

Children spend years in foster care, and there are concerns that bureaucratic hurdles contribute to unnecessarily long stays. Mi Abogado (My Lawyer) was introduced in Chile to enhance legal aid for foster children to accelerate family reunification. In a novel approach to policy making, the Chilean government randomized the introduction of the program for children living in institutions to evaluate effects on child wellbeing. Using registry data, we find that exposure to the program significantly reduced the duration of foster care without increasing subsequent maltreatment and placements, leading to substantial cost savings to the child protection system that are over four times higher than the cost of the program itself. The exposure also decreased criminal justice involvement and improved school attendance. The results show that investing in improvements to foster care services can improve child wellbeing in a cost-effective way.

*Acknowledgments: This study is due to the efforts of the Experimental Policy Initiative of the Chilean Budget Office. We thank Rema Hanna and John Friedman for valuable comments and advice. We also thank Josefa Aguirre, Anna Aizer, Bocar Ba, Jason Baron, Patrick Bayer, Pablo Celhay, Janet Currie, Emilio Depetris-Chauvin, Maria Fitzpatrick, Francisco Gallego, Joseph Hotz, Caroline Hoxby, Jeanne Lafortune, Doug Miller, Derek Neal, Roberto Rigobon, and Michael Whinston, and as well as participants in seminars at Duke University, the Frisch Centre for Economic Research, Georgetown University, MIT Sloan, the NBER Program on Children, the NBER Program on Law and Economics, Pontificia Universidad Católica de Chile, Stanford University, the University of Bergen, the University of Notre Dame, and the University of Oslo, for helpful suggestions. Catalina Bravo, Nicolás Cabezas, Carolina De Iruarrizaga, Benjamín Echecopar, Raimundo Eyzaguirre, Julián García, and Antonia Sanhueza provided excellent research assistance. Francisca de Iruarrizaga provided valuable insights into the child protection system in Chile. Verónica Pincheira and the rest of the team from the Ministry of Justice provided valuable support and contributions to the evaluation project. We are grateful to the Supreme Court of Chile, and to Fabiola González, Omar Manriquez, and Ricardo Tucas for their excellent and generous work in constructing the justice-system databases. We thank Karina Vega from Servicio Nacional de Menores for help in facilitating and preparing their data. We thank the Studies Department of the Ministry of Education for making available data from the education system. We thank the MIT Sloan Latin America Office for financial support of this collaboration. Thanks to ANID for its support through FONDECYT Iniciación project 11221167. RCT ID: AEARCTR-0004160.

†University of Chicago and Experimental, ryancooperb@uchicago.edu

‡MIT Sloan and NBER. Corresponding author: jjdoyle@mit.edu; 77 Massachusetts Avenue, E62-518, Cambridge, MA 02139; (617) 452-3761.

§Pontificia Universidad Católica de Chile, School of Government, andreshojman@uc.cl,

1 Introduction

Child protective services (CPS) are remarkably common worldwide. In the US, 37% of children will be investigated for child abuse or neglect during their childhood, and in both high- and low-income countries, roughly 5% of youth spend some time in foster care during their childhood (García and Hamilton-Giachritsis, 2014; Fallesen et al., 2014; Kim et al., 2017; Rouland and Vaithianathan, 2018; Yi et al., 2020). Further, children involved with CPS are particularly vulnerable. They experience 2-3 times higher childhood mortality rates and 7 times higher rates of depression and anxiety, compared to children with similar observable characteristics (Johnson-Reid et al., 2007; Turney and Wildeman, 2016). Once children are in foster care, the primary aim of family courts is to rehabilitate and reunify families or secure adoptive homes (Becker et al., 2007; Ryan and Gomez, 2016; Konijn et al., 2019). The act of child removal, combined with rehabilitative services such as drug treatment and mental health services, form a powerful intervention that can have large effects, both positive and negative, on child and family wellbeing (Grimon, 2021; Bald et al., 2022; Baron and Gross, 2022).

Family rehabilitation activities typically take two years to complete, and there are serious concerns that bureaucratic delays harm child wellbeing (Miller, 2004; Farber et al., 2009; Miller et al., 2020). A large literature finds mixed results when investigating the correlation between length of stay in foster care and child outcomes, possibly due to selection bias (Dworsky et al., 2013; Okpych and Courtney, 2014; Bender et al., 2015; Font et al., 2021). Despite the ubiquity and policy interest in child welfare systems, there is little evidence of the causal effects of efforts to improve them (Blome and Steib, 2008; Hunter et al., 2014).

In this paper, we examine the impact of Mi Abogado (My Lawyer), a program that was initiated in 2017 in Chile that aims to protect the rights of foster children by improving their legal representation. The program provides foster children with a comprehensive support system to improve case management, including access to a lawyer with a reduced caseload, a psychologist, and a social worker. Program goals include reducing bureaucratic frictions and increasing access to services in order to speed children's exit from foster care. The Experimental Policy Initiative (EPI) at the Chilean Budget Office proposed that the introduction of the program for children living in residences (group homes) be structured as a large-scale randomized controlled trial (RCT) as a means to evaluate it. This evaluation was further enhanced by Chile's high-quality registry data, offering a means to track outcomes for children involved in the study.

By design, the treatment group was recommended to the family court for entry into the program, resulting in a 60% increase in program exposure compared to the control group over the following two years. Intent-to-treat estimates show that this greater exposure accelerated exit from foster care residences, with an increase in the number of days children live with their biological or adoptive family of 6 days per quarter, or 30% more than the control group's mean.

An innovation in this paper is that we can test whether a program that aims to reduce bureaucratic frictions and length of stay improves additional barometers of child wellbeing: measures of child safety, criminal justice involvement, and school attendance. We find that child safety was not significantly impacted, as indicated by four measures: subsequent child protection investigations, foster care re-entry, criminal victimization, and hospitalization. Meanwhile, the treatment group experienced a 30% reduction in criminal justice involvement over the two years following randomization compared to the control group, including a reduction in reports of violent crimes. We also find evidence that the program improved school attendance, and event studies show this increase is largely concentrated during months when school attendance was low across Chile. This provides suggestive evidence that the program improved attendance when it was more discretionary. Estimates of effects on school performance are imprecise and not statistically significant.

In terms of heterogeneity in treatment effects, we lack statistical precision to test differences across groups defined by regions or demographics. Notably, across the outcomes the treatment effect point estimates are larger for boys than for girls, although the differences are not statistically significant. More generally, crime improvements are concentrated among children with characteristics associated with higher crime rates.

The program is a bundle consisting of a lawyer, social worker, and psychologist, raising the question about which features of the bundle deliver these outcomes. A qualitative review commissioned by the Program highlights that while the lawyer is the main driver of the program, the psychosocial support is integral to ensuring the family is rehabilitated. Although the experimental design does not provide exogenous variation for the different elements of the bundle, we can examine how results vary with pre-determined characteristics that predict the use of these services. The results show that children predicted to have a higher share of psychosocial processes see the largest reduction in time in foster care, but they do not experience a larger decline in crime reports. This is consistent with the conclusion of the Program's qualitative assessment that the full bundle is important for family functioning and reunification. Meanwhile, the effects of the program appear to grow over time, consistent with increased exposure to the program improving the trajectory of child outcomes.

Our findings suggest that the program is highly cost-effective for the child welfare system. The reduction in length of stay in state custody results in substantial net savings: the estimate suggests that for every additional dollar spent on enhanced legal aid, the system would save nearly $4.50. If we included the cost of criminal justice involvement, the cost-benefit comparison would be even stronger. While we do not observe every welfare-relevant outcome of interest, the results demonstrate that improving the quality of case management for children in residential care can improve child wellbeing while lowering costs for the child protection system.

The paper has four main contributions. First, legal aid in child welfare is highly varied across jurisdictions, yet there is little causal evidence investigating how the structure of the aid affects outcomes for children and families (Haarberg, 2024). This paper contributes to closing this gap. Second, the results add to the literature that seeks to understand and improve the functioning of foster care systems more broadly. Effects of foster care placement on child wellbeing among marginal placements have been found to vary across time and space (Doyle, 2008, 2013; Bald et al., 2022; Gross and Baron, 2022; Helensdotter, 2024). A potential explanation for this heterogeneity may be the quality of foster care provided across these contexts, including differences in length of stay and case management. The current paper investigates how an effort to improve case management affects child wellbeing directly. Related, the paper contributes to the body of research that investigates the relationship between length of stay in foster care and child outcomes. This paper investigates a program that aims to rehabilitate families and reduce legal frictions so that children can be reunited faster. Further, the results suggest that speeding exit from care in our context improves child wellbeing in a highly cost effective way. Third, the paper adds to the growing interest in the role that legal aid can provide to improve wellbeing. Examples include legal aid to reduce evictions, improve health, and protect the rights of criminal defendants (Martinez et al., 2017; Tsai et al., 2017; Anderson et al., 2019; Harris, 2020; Cassidy and Currie, 2022; Hoynes et al., 2022; Shem-Tov, 2022; Anwar et al., 2023; Lacoe et al., 2023). Last, the paper demonstrates that large-scale randomized trials implemented as part of public policy reforms can equitably distribute new services while providing a rigorous way to evaluate their effectiveness (Muralidharan and Niehaus, 2017).

The remainder of the paper is organized as follows. Section 2 provides background information on legal aid in child welfare, foster care placement in Chile, and the intervention; Section 3 details the randomization and the empirical strategy; Section 4 describes the data; Section 5 reports the results; Section 6 interprets the results by exploring mechanisms, considering treatment effect dynamics, and examining the program's cost effectiveness; and Section 7 concludes.

2 Background

2.1 Child Protection in Chile

Child protective services (CPS) investigate allegations of child maltreatment and supervise foster care placements where children live temporarily with substitute families or institutional residences. CPS systems are found throughout the world (Duerr Berrick et al., 2023). Our context is Chile in 2019 when the child protection system was administered by the Servicio Nacional de Menores (SENAME).[1] Entry into the system begins with a report of maltreatment from reporters such as teachers, physicians, law enforcement, and families. The allegations that lead to foster care involve some form of neglect in 85% of cases, while 29% involve physical abuse, and 18% are related to sexual abuse.[2] Family courts then determine whether to place children in foster care and placement into foster care and supervise the cases over their stays. In 2021, there

[1]The child welfare system is currently administered by Mejor Niñez, Servicio Nacional de Protección Especializado a la Niñez y Adolescencia.

[2]Authors' calculations based on SENAME and judiciary data; categories are not mutually exclusive.

were approximately 11,000 children residing in foster care, a rate of 2.8 per 1000. This places Chile in the middle of the distribution in South America and lower than the rate in the U.S., which is closer to 6 per 1000 (Petrowski et al., 2017; SENAME, 2021). The most common placement type is with a foster family, often the child's extended family (kinship foster care). At the same time, residential care in institutions is also common, as in other countries in South America (Muñoz-Guzmán et al., 2015; Petrowski et al., 2017).

The intervention we study focuses on children in residential care. Residences are supervised by public and non-profit agencies, and their size varies from fewer than 10 to over 200 children, with the average (median) child living in an institution with 48(30) other children (Appendix Figure A.1). The average length of stay in Chile is relatively long by international standards at three years compared to fifteen months in the U.S., seventeen months in Germany, and twelve months in Portugal (De Iruarrizaga, 2016; Duerr Berrick et al., 2023). As in other countries, the goal for most children is family reunification, with adoption being relatively rare in Chile at 5% compared to 25%, in the U.S. (García and Hamilton-Giachritsis, 2014; Muñoz-Guzman et al., 2023) (Duerr Berrick et al., 2023).

In 2016, SENAME was the subject of a high-profile condemnation of the care and supervision provided within residences due to a large number of unexplained deaths over the prior decade. This included an investigation by a Congressional commission ("Comisión SENAME") and by the United Nations Committee on the Rights of the Child (USDoS, 2019). The scrutiny led to many policy changes. First, there was a push to reduce the reliance on residential care. In 2010, there were 15,497 children in foster care, including 12,350 (80%) in residential care with the remaining children placed in family settings. By 2021, there were 10,865 children in foster care, including 4,451 (41%) in residential care (SENAME, 2021). The share of the population in *residential* care in Chile is now similar to the global average and that of other countries in Latin America and the Caribbean in particular (Petrowski et al., 2017). Second, funding levels for residential care increased. Subsidies to residences had been US$300 per child per month, a level that was criticized as far lower than the estimated US$1,000 deemed necessary for high-quality supervision. In 2019, at the start of the intervention we are studying, the per-child subsidies had increased to approximately US$700 per child per month. Third, there were calls for improved legal representation. All foster children were nominally assigned a lawyer, but there were concerns that extremely high caseloads prevented them from providing high-quality case management. In order to explore ways of protecting the rights of foster children, the Ministry of Justice launched Mi Abogado as a pilot program in 2017 to enhance the legal aid provided to foster children.

2.2 Legal Aid in Child Protection

Child protection has parallels to criminal justice, involving allegations reported to authorities (often by reports that are mandatory for physicians and educators), investigations by child protective services, and a family court that holds hearings to oversee the process. In particular, family courts decide whether the child should be removed from home and placed in foster care. Once a child is in care, the goal of the case is typically family rehabilitation and reunification after the family has convinced the judge that it is safe to do so. If the court determines this is unlikely, then there is a process to terminate parental rights and seek an adoptive home. This reunification process takes place over several years (Bald et al., 2022).

Legal support for children in this process varies widely across jurisdictions. Children are often represented by a court-appointed special advocate (CASA) or a guardian-ad-litem who is usually an attorney (Sexton, 2018; Miller et al., 2020). The statutory goal for these advocates is usually to represent the "best interests of the child". There is little empirical work investigating the effects of different forms of legal aid on child outcomes (Pilkay and Lee, 2015; Haarberg, 2024). Orlebeke et al. (2016) implemented a randomized evaluation of additional training for 264 lawyers in Washington and Georgia. The training improved adherence to best practices, although no difference in the time to family reunification or adoption was detected. For a subset (older children in Washington), time in care was reduced. Rashid and Waddell (2019) studied the staggered rollout of mandates for legal representation across five states in the U.S., and they found that such a mandate did not change the disposition of cases, but it did decrease the time to adoption by 14%. Osborne et al. (2020) used propensity score methods and found that appointment of a CASA was associated with delays in family reunification, although CASAs in their context are often assigned to cases that are the most complex, which can confound comparisons (Cooley et al., 2019). Zinn and Peters (2015) evaluated an attorney-ad-litem program in Florida, where the advocate's role is to represent the child's expressed interests. Comparing children who were eligible to those deemed ineligible, they found that having a lawyer speeds exit to adoption after the adoption decision has been made. Meanwhile, parents are rarely represented, but evidence from matched comparisons in New York City and Washington State suggests that such representation can reduce the time in foster care (Courtney and Hook, 2012; Gerber et al., 2019). Despite interest among policymakers in learning how to improve the quality of foster care services, relatively few studies provide rigorous evidence on the effectiveness of foster care reforms (Bergström et al., 2020; Rushovich et al., 2021). The current study is an example of how randomization can be incorporated into a program scale-up to test its effectiveness.

2.3 Mi Abogado Program

The Mi Abogado program provides comprehensive legal aid by assigning each child a coordinated team consisting of a lawyer, a psychologist, and a social worker. The team's goals are to protect the children's rights, facilitate their return to family life (whether with their family of origin, extended family, or through adoption), and provide access to services intended to improve child wellbeing. Without Mi Abogado, children would not have access to a psychologist or a social worker as part of their case management team.

While the program's strategy is based on the team's work, the program is called "Mi Abogado" because it is largely focused on legal aid carried out by the lawyer. This includes improved preparation for hearings and attention to the timely achievement of case goals, in part through improved case management that connects families with rehabilitative services. One feature of the program is a lower caseload for the lawyer relative to usual practice. The nominal caseload of the lawyers is limited to 80, with a goal of fewer than 60. While data on caseloads is incomplete, as cases often remain open even when they are dormant, we can measure case assignment in data we describe in more detail below. We find that over the last 12 months of our observation period, the average Mi Abogado lawyer was assigned 130 cases compared to 309 case assignments among non-Mi Abogado lawyers (Appendix Figure B.1). The nominal caseload of the psychologist and the social worker is larger: limited to 240, with a goal of fewer than 180. As a summary measure, the lawyer's wages amount to 70% of the operational wages associated with the program. See Appendix B for a detailed description of the program and the roles of the team members.

The intervention begins when a family court judge assigns the child to the program. The program team then reviews the child's legal file and visits the residence to speak with the child and staff. Within the first 30 days of program initiation, the team is tasked with devising an interdisciplinary plan that involves a mental health evaluation, a diagnosis of social needs, and a legal strategy to overcome procedural hurdles. During the next three to six months, the team continues to meet with the child every month, as well as the residence staff and the family, in an effort to speed reunification. Once a child leaves residence and is reunited with family, the Mi Abogado program continues to monitor the child's welfare for 90 days to verify the quality of the family reestablishment.

The family court monitors compliance with the program's objectives. This tracking facilitates our ability to describe the program in detail, as our data include program interactions with the child, the residence, the family, and the court system. Figure 1 reports the average number of these processes carried out over the first year among children who participated in the program for at least one year. Documentary work is the most common, averaging 19 processes, followed by interacting with the residence staff (13). On average, the team or the lawyer meets with the child nine times over the first year and with the family four times. Processes labeled as court interactions are rare, although much of the court activity is categorized under documentary work.

With the lower caseload and access to the care team, administrators of the Mi Abogado program note that they were able to recruit relatively high-quality lawyers. As the program expands, the average quality of the lawyers may decline. That said, a lawyer in Chile is trained with a Bachelor's Degree instead of training as a Juris Doctor, and the program administrators note that there is a large pool of potential lawyers in this context.

Without the Mi Abogado program, children in residential care would rely solely on standard legal representation, which is often constrained by high caseloads and limited capacity for individualized attention; they would not have access to the interdisciplinary support led by a lawyer accompanied by a dedicated psychologist and social worker who construct a plan aimed at speeding family rehabilitation.

3 Empirical Strategy

3.1 A Pragmatic Randomized Controlled Trial

In 2019, the Mi Abogado program was expanded in the four most populous regions in Chile: Maule (population 1,118,947), Biobío (1,654,744), Valparaíso (1,935,455), and Metropolitan (7,915,199), which includes Santiago (INE, 2009). The program's assignment of children was managed by the Ministry of Justice, with the Experimental Policy Initiative at the Chilean Budget Office handling the randomization. This Initiative implemented a pragmatic randomized controlled trial to allocate the limited openings in an equitable manner and with the explicit interest in evaluating its effectiveness (Cooper et al., 2022). The trial was registered with the American Economic Association (AEA) Registry.[3]

The study focused on children aged 6 to 18 living in SENAME residences (i.e. group homes) across these four regions from January to February 2019, totaling 1,871 children. On March 30, 2019, the randomization assigned 581 children to treatment, with stratification by age (above or below 12 years), sex, and region. The number of available treatment slots and eligible children varied widely by region, affecting the proportion of children randomized into the treatment group: 32% in Metropolitan, 92% in Maule, 10% in Valparaíso, and 7% in Biobío. Appendix Table C.1 provides sample sizes for each region. Below, we discuss the empirical implications of this varying treatment propensity across strata.

Following randomization, the Experimental Policy Initiative forwarded the list of selected children to the Mi Abogado program, which then petitioned the family court judges to change the legal representation. These judges were ultimately responsible for determining the children who would participate in the program. In May 2019,

[3]https://www.socialscienceregistry.org/trials/4160.

the program randomly selected 51 children from the control group to be included in the treatment group. We include these children in the treatment group for the main results, although the manner in which they are included in the analysis does not affect the results, as shown below. Additionally, control group members gradually entered the program over time. The primary analysis uses intent-to-treat models and longitudinal data to compare differences in program engagement and child wellbeing measures over time.

3.2 Empirical Model

Given that we have longitudinal data on outcomes, our preferred way to measure the impacts of the program compares the treatment and control groups over time in event studies. These event studies trace the evolution in the difference in outcomes across treatment and control in the quarters before and after the randomization, providing a transparent way to estimate intent-to-treat effects over time. In particular, for child i in calendar quarter t and event time q, we estimate:

$$Y_{iq} = \alpha + \gamma T_i + \sum_{q \neq -1} \lambda_q \mathbb{1}\{Q_t = q\} + \sum_{q \neq -1} \theta_q \mathbb{1}\{Q_t = q\} \times T_i + \beta X_i + \varepsilon_{iq} \qquad (1)$$

where q is normalized as the number of quarters from the second quarter of 2019 (recall that the randomization occurred on the last day of the first quarter). T_i is an indicator the child was randomized to the treatment group, X_i represents randomization-strata indicators, and Q_t is the calendar quarter of the observation. The summation terms include indicators for each quarter in event time, and we are interested in the estimates of θ, the difference between the treatment and control groups in each quarter relative to the first quarter of 2019.[4] Confidence intervals are calculated using heteroskedasticty-robust standard errors, clustered at the child level.

To summarize the event-study findings, we compare the outcomes across two periods, before and after randomization. We estimate two models:

[4]The panel is balanced and the model does not include any continuous regressors. As a result, estimates of θ are identical if time-invariant controls or child-fixed effects are included. Similarly, because event time is zero in the second quarter of 2019 for all of the subjects, these point estimates are also identical if we add calendar-quarter or season fixed effects. These results are discussed in the robustness section below.

$$Y_i^{POST} = \alpha + \psi T_i + \beta_1 Y_i^{PRE} + \beta_2 X_i + \varepsilon_i \tag{2}$$

$$Y_{iq} = \alpha + \gamma T_i + \delta Post_q + \psi T_i Post_q + \beta X_i + \varepsilon_{iq} \tag{3}$$

Equation 2 is a cross-sectional regression of the subject's average outcome across the nine quarters in the post period on a treatment-group indicator, the average of the outcome across the quarters in the pre-period, and strata controls, X; ψ represents the intent-to-treat estimate. The difference-in-differences model described by Equation 3 also identifies the same ITT parameter (McKenzie, 2012). $Post_q$ is a variable that takes the value of 1 in all periods after randomization and 0 otherwise. The intent-to-treat parameter, ψ, is the coefficient on the interaction between the treatment indicator and $Post_q$, representing the average difference across the groups in the post-randomization period relative to the average difference in the pre-period. The estimate of γ provides a check on whether the treatment and control groups are similar in their average level of the outcome variable during the pre-randomization period, conditional on the strata controls. We report heteroskedasticity-robust standard errors for both models; for the longitudinal DD model they are clustered at the child level.[5]

In a context with no controls, the two models provide the same intent to treat estimate when β_1 is set to one, and McKenzie (2012) shows statistical power can vary across the models depending on the autocorrelation of the outcome.[6] In our setting with strata controls, the cross-sectional estimator includes them in a multivariate regression, while the DD estimator absorbs any difference in time-invariant covariates as in the event studies.[7]

We prefer the DD specification, as it is a standard summary of the main event-study estimates. Indeed, it simply collapses the event study into two periods: before and after

[5]The randomization was carried out at the child level, although siblings may receive attention from the program. This contributes to non-compliance with the treatment status. We explored using family-level models, but the family identifiers contain measurement error that we do not want to incorporate. Instead, we use intent-to-treat models of engagement and child outcomes to yield unbiased estimates of both the costs and benefits of offering the program.

[6]McKenzie (2012) compares the DD model to an ANCOVA model that is similar to our cross-sectional model as shown below.

[7]Similar to the event study models, the ITT point estimates are identical if time-invariant controls, child fixed effects, or calendar quarter indicators are included in the DD models. The invariance of these results are presented in robustness checks below.

the randomization. Further, the DD model provides more precise estimates across all the outcomes in our context. Third, the DD models absorb fixed characteristics of children, as demonstrated below. Nevertheless, we report estimates for both the cross-sectional and DD models for comparison, and the interpretation of the program's effects is similar regardless of the estimator.

There are concerns when estimating event studies and difference-in-differences models when the treatment evolves over time (de Chaisemartin and D'Haultfœuille, 2022). Here, there is a single "episode", where event time equals zero for everyone in the second quarter of 2019, which avoids issues related to staggered treatments. The event study estimates allow us to compare the costs and benefits of offering the program over the following two years (de Chaisemartin and D'Haultfœuille, 2024). To understand the evolution of the effects further, following the presentation of the main results we explore both complier characteristics over time and how effects of the treatment evolve with varying levels of exposure to the program.

4 Data Description

4.1 Data Sources

The analysis incorporates a wide range of outcomes that are visible longitudinally in registry data. The data are linked across administrative agencies in Chile using the child's social security number. For more detail on the data and the available time periods, see Appendix D.

First, we have child protection data from SENAME beginning in January 2017 through June 2021. This includes the dates of reports and their allegations. SENAME also oversees foster care placements, so we can observe the dates children enter and exit different care settings, including an identifier for each institution. We observe the disposition for children who exit foster care, including returning home or placement in an adoptive home. These data allow us to track whether children who exit the system subsequently re-enter care as a measure of child safety. They also include demographics, such as sex and age, and a measure of school delay, defined as the difference between age and the age expected for the child's grade. SENAME also linked hospitalization data from the Ministry of Health for children in their care each calendar year. The data include the child's last hospitalization each calendar year. We have data from 2019 for all children in the study, which allow us to see whether a child was hospitalized after randomization.[8]

Second, the Judiciary Registry data allow us to observe criminal justice involvement from 2009 to August 2021. This includes reports to the courts when a child is reported as the person who committed a crime. We restrict the sample to begin in April 2014, the first full quarter where we observe non-negligible crime-report rates given the age of the children.[9]

We also use the Judiciary Registry to observe victimization. This includes two main categories: children reported missing, which may be more likely for children in institutions, as residence staff are required to report children as missing if they are not in the residence at night; and children being reported as victims of a crime, which we use as a complementary measure of child safety along with the child protection reports. When SENAME data have missing allegations, and there is an open case involving child victimization at the time of the foster care placement, we use the victimization data to clarify the nature of the allegation.

The Judiciary Registry allows us to observe the associated lawyers for all children. We use this data to compute the number of cases assigned to lawyers as a proxy of their caseload. However, we do not estimate the caseload directly because cases usually stay "open" even after they become inactive. These data also include information on family court hearings and writs, which we use to measure court activity in the case.

Another set of outcomes comes from the Ministry of Education Registry between March 2017 and December 2019. We have monthly school attendance and annual school performance data, coded as the average performance across all subjects in a given year measured in nationwide percentiles. The COVID pandemic severely impacted most school activities beginning in March 2020; it is not possible to obtain grades for 2020.

Last, we observe information tracked by the program for those participating in Mi

[8] We obtained information on diagnoses associated with these hospitalizations, but for a subset of children in care in 2000. Appendix F provides more detail.

[9] Other common criminal justice outcomes such as conviction and incarceration are more difficult to observe in the Justice Registry, as some fields appear to be incomplete. For example, most reports are not accompanied by a guilty sentence in our data, partly because many cases are not closed.

Abogado from October 2018 to December 2020. This includes dates of participation and processes carried out.

4.2 Program Engagement Measures

We measure engagement using the Mi Abogado program data. First, we observe when a child enters the program, but we do not observe a program end date. When calculating days of participation, we rely on the program rules that the Mi Abogado team oversees a case for 90 days after exiting foster care. Days participating in the program is therefore measured as the exit date + 90 minus the program entry date.[10] Given that program exit is endogenous, our preferred engagement measure is one of exposure: days since entering the program. An advantage of the detailed program data is that we observe the processes performed for each case, including visits with the child, the residence staff, and the family.

4.3 Child Outcome Measures

The first child outcome we consider is whether children are living with their biological or adoptive family, which is a primary objective of child welfare agencies and courts. For a given calendar quarter, we calculate the days when children are living with their families by summing the days they have returned home and have not re-entered foster care. Children who exit foster care as adults (known as "aging out") are not recorded as living with family after they exit.[11] We also examine the number of days children are in different types of foster care as complementary outcomes.

One question is whether speeding the exit from foster care improves child wellbeing, particularly regarding the safety of the family environment. We measure child safety by observing whether a child returns to foster care, whether there is a new investigation for child maltreatment, and whether the child is identified as a victim of crime, which

[10] For those who age out of foster care, we do not include this 90-day supervision after exit from foster care in our calculation, consistent with program rules.

[11] In the jargon of child protection, these children have not achieved "permanency" by the time they leave foster care, which means they have not been adopted nor reunified with their biological family. In a small number of cases, children exit foster care alone or transition to another supervisory residence. These cases are also not coded as living with family. That is, children who exit foster care to live with family are considered to be living with their family unless they re-enter foster care.

is typically a form of child abuse. To the extent that the Mi Abogado program speeds the return home, program participants will have more time at risk for these outcomes. In addition, recall that the program provides services to children for 90 days after exit from foster care. This greater surveillance may lead to greater child maltreatment *reporting*, and the estimated effects of the program on reports may be biased upward relative to the effects of the program on actual child maltreatment. In addition to these direct child-protection measures, we measure whether a child was ever hospitalized in 2019 after randomization.

Another set of outcome measures covers criminal justice involvement and educational outcomes as barometers of whether the program is more likely to put children on a trajectory of improved wellbeing. In particular, we calculate the number of crimes reported each quarter, as this measures the intensity of criminal justice involvement. We also discuss the types of crimes reported. For educational outcomes, we focus on attendance in a given quarter. This is measured as the share of school days that the child attended school. We report effects on school performance as well.

5 Results

5.1 Balance

Table 1 reports baseline comparisons of observable characteristics across the treatment and control groups at the time of the randomization (March 30, 2019). The comparisons test whether the prescribed randomization was carried out faithfully, and they describe the system and the children involved. To calculate the mean for the treatment group, we regress each characteristic against a treatment group indicator along with strata controls: indicators for the region, sex, and age group. We then report the treatment group's conditional mean as the control group's mean plus the coefficient on the treatment-group indicator. We report p-values for the test that the coefficient on the treatment indicator is different from zero. Last, we report a p-value for an F-test of joint significance that all of the characteristics predict the treatment indicator, again controlling for randomization-strata controls.

The first two rows describe family-court activity that we observe in our data back to 2010. Both groups are similar, with approximately 2.9 writs filed per quarter and 0.2 hearings per quarter. The next row shows that children spent approximately 25 days with their families and 62 days in residential care per quarter since 2017.[12]

The criminal justice data show that the groups are comparable in the number of crime reports per quarter (0.03). In terms of other criminal justice measures, the two groups share similar rates of being reported missing (0.08) and reported as victims of abuse (0.02) before randomization. The education data show that the subjects are disadvantaged, and the treatment and control groups are comparable. The share of days attending school in 2017–2018 is 66%. The children are in the 27–29th percentile among those with grades available in 2018 and have similar rates of missing 2018 grade information, on the order of 35%.

The remaining rows show balance on other observable characteristics across the treatment and control groups. They average 1.4 siblings, and their school delay is about one year. The child maltreatment allegations that led to the placement in early 2019 are similar to those in the child welfare system as a whole: 85% involve some form of neglect, approximately 30% involve physical abuse, and 17% involve sexual abuse. Regarding demographics, the average age when the child first entered the child welfare system was 11, and the average age was 14 at the time of the randomization. This is somewhat older than the complete set of children in care, partly because participation in the evaluation was restricted to children at least age 6. When we calculate an F-statistic for the joint significance of all of these controls in predicting treatment, we fail to reject that they are not significant, with a p-value of 0.60. These comparisons confirm that the randomization resulted in treatment and control groups that are highly similar as designed.

5.2 Program Engagement

The treatment group was randomized to have access to the program, but participation depended on approval from a family court judge. There was also noncompliance as family court judges admitted some members of the control group over time. Regardless, the randomization of court petitions generated substantial variation in exposure to the

[12]Recall that eligibility for the program and the study required all of the subjects to live in a residence in early 2019. The remaining days are transitions between programs or family foster care.

program that we use to evaluate its effectiveness.

Figure 2 reports the difference in exposure and participation for the two groups over time.[13] Figure 2a examines the difference in exposure to the program using daily data. To do so, we residualize the data by regressing an indicator for having entered the program on the set of strata controls and a treatment indicator. We then compare the average residuals for the treatment and comparison groups on each date, which describes the shares of the treatment and control groups that had ever entered the Mi Abogado program, controlling for the randomization strata. The figure shows that a few weeks after the randomization, there was a sharp rise in program exposure among the treatment group compared to the control group. Control group members gradually entered the program over time until roughly 70% of the treatment group and 60% of the control group had participated in the program by December 2020.

The remaining figures report event-study estimates according to Equation 1, where the data are binned by calendar quarter. The difference in cumulative exposure (the number of days since beginning Mi Abogado) in shown in Figure 2b. This difference increased at a decreasing rate over time since randomization. After one year, the treatment group has 100 more days since first exposure to the program, increasing to 150 days by December 2020.

Figure 2c shows that the treatment group is 45 percentage points more likely to *participate* in the program during the first quarter after the randomization, and this difference in participation falls over time. The program processes provided to the subjects follow a similar pattern, with the treatment group experiencing four more program interactions in the first quarter after randomization, increasing to six in the second quarter, and the difference falls afterward (Figure 2d). Appendix Figure E.1 shows that this pattern of increased use of the program is observed for documentary work (rising to approximately 3 more filings per quarter for the treatment group), interacting with the child (up to 1 more visit per quarter), interacting with family (up to 0.25 more visits), and interacting with residence staff (up to 2 more visits). Appendix Figure E.2 shows that interdisciplinary processes involving more than one team member are consistently higher for the treatment group after randomization, while pro-

[13]Pre-period participation is not strictly zero because a small number of children were able to join the program during its pilot phase.

cesses carried out by a sole member of the team are elevated especially in the first or second quarter after randomization. Similarly, in the quarter after randomization, the treatment group had one more writ entered into the system compared to the control group (Appendix Figure E.3).

Finally, Table 2 reports differences in participation and exposure using the cross-sectional and difference-in-differences models, where the outcomes are measured in days per quarter for comparability.[14] Column (1) shows that the treatment group averages 17 more days of exposure per quarter in a cross-sectional model (s.e. = 2.3).

Columns (2)-(5) summarize the event studies using the two-period difference-in-differences specification. As with the other outcomes, the first row reports estimates of the difference in the pre-randomization mean of the outcome across treatment and control, and the estimates are small and not statistically significant as intended. The intent to treat estimate is given by the coefficient on the interaction term, Treatment x Post, and Column (2) shows that the treatment group has an estimated 20 more days of exposure per quarter (s.e. = 1.8). This is approximately 60% higher exposure compared to the control group's mean of 32 days over the post-randomization period. Column (3) shows that the effect of treatment is larger for girls by 5.4 days, but the difference is not statistically significant (p = 0.13). Last, column (4) reports results for days actually participating in the program each quarter, which reflects both program exposure while in foster care and time in foster care itself. On average, the treatment group has 13.5 more days of participation in the program each quarter in the post-period.[15]

5.3 Living with Family

A primary goal of the Mi Abogado program is to overcome unnecessary delays to accelerate family reunification. When evaluating its effectiveness, the first question is whether the program achieved this goal.

Figure 3 reports the event-study estimates. The figures show that in the pre-

[14] Aggregate differences are described in the cost-benefit discussion below.

[15] We have a proxy for the presence of a lawyer, and this is significantly higher for the treatment group (an analogous coefficient of 4.58 days per quarter, s.e. = 2.02). However, this may reflect higher data quality in measuring lawyer assignments for those entering Mi Abogado. Nevertheless, the increase in the proxy is another measure demonstrating greater exposure to the program among the treatment group.

randomization period, the difference between the groups is stable across these outcomes. After randomization, the difference between the groups grows before stabilizing after three calendar quarters. Figure 3a shows the difference in ever-living with family in a given quarter rises to approximately 7 percentage points compared to the quarter just before the randomization. A similar pattern is seen when measured as the number of days living with family (Figure 3b), and the reverse image is seen when measured as days living in SENAME residence (Figure 3c).[16] By and large, exiting foster care in our study period is unlikely to result from an adoption, as 3.5% of those who exit do so.

Table 3 reports the cross-sectional regression and difference-in-differences (DD) estimates. The intent-to-treat estimates represent an additional 3.61 days (s.e. = 2.14) in the cross sectional model, and 6.46 days (s.e. = 2.05) using the DD specification. The two point estimates are not statistically significantly different from each other (p = 0.15). This is 17-30% higher than the average of 22 days for the control group during the post-randomization period.

Column (3) indicates the intent-to-treat estimate is 8.9 additional days at home for boys, which is 4.3 days larger than the estimate for girls whose estimate is 4.6 more days living with family.[17] However, this difference between boys and girls is not statistically significant (p = 0.30). Column (4) shows that the treatment group is 7.65 percentage points more likely to ever live at home each quarter after randomization, or 30% higher than the control group's mean of 25.9%. Related, column (5) shows that the treatment group spends 4.6 fewer days in a SENAME residence each quarter, or 9% lower compared to the control group's mean of 54 days. With so much attention in the child welfare literature devoted to time in care, it is noteworthy that a legal aid intervention can substantially speed children through the system toward the goal of family reunification.

[16]The days in SENAME residence measure differs from living with family because there are alternatives, mainly living in family foster care and aging out of foster care.

[17]The difference relative to the control-group mean for girls is 19%, while for boys the point estimate represents a 50% increase compared to their control-group mean.

5.4 Child Safety

One question for programs aimed at accelerating family reunification is whether the effort results in premature exits and child-safety concerns. Figure 4 shows event-study results for three related measures of child safety measured in each quarter: the number of child-protection investigations opened, the number of reports of child victimization, and the number of foster care re-entries in a given quarter.[18] The event studies do not suggest an increase in these measures of subsequent child-protection involvement. These null effects are despite the fact that the treatment group is more likely to have returned home, where these outcomes can occur. Further, the lack of a detectable increase in safety concerns is despite the fact that the treatment group is more likely to participate in the Mi Abogado program, which provides greater surveillance once a child returns home.

Table 4 reports the cross-sectional and difference-in-differences estimates for these outcomes. In addition, we report cross-sectional results for the outcome of hospitalization after randomization during 2019. We find a null effect for each of these outcomes, although the standard errors do not rule out sizeable increases compared to the (relatively rare) means.[19] To gain precision and analyze the outcomes together, we calculated an index, which is a weighted average of the (standardized) safety outcome measures using inverse covariance weights to avoid adding weight to highly correlated measures (Anderson, 2008).[20] The results suggest that the treatment group has a small and statistically insignificant impact on the safety index. Across all of these measures, the point estimates are small, none are statistically significant, any increase is not sustained according to the event studies, and they are inconsistent in sign. Taken together, these findings suggest that child safety does not worsen for those in the treatment group.

[18] Recall that the hospitalization outcome is available only for the child's last hospitalization each calendar year. As a result, we are unable to trace hospitalization rates over time in event studies.

[19] The upper bound of the 95% confidence interval for protection cases is 0.013 (10% of the mean), for child victimization reports is 0.004 (17% of the mean), for foster care re-entry is 0.006 (67% of the mean), and for hospitalization is 0.03 (46% of the mean).

[20] The procedure standardizes the measures to have a mean zero and a standard deviation of 1, calculates the covariance matrix, and then calculates the weighted average.

5.5 Criminal Justice Involvement

In addition to measures of recidivism within child welfare, this project uses linked administrative data that allow us to investigate a wider set of child wellbeing measures. There is a close link between child welfare and juvenile delinquency (Hirsch et al., 2018; Cho et al., 2019), and criminal justice involvement can be used as a barometer for whether the intervention improves child outcomes as opposed to speeding children through the child welfare system "too quickly."

Figure 5 shows how the difference in the number of crime reports between the treatment and control groups changes with time. Before the randomization, the difference across the two groups is small and does not exhibit a pre-trend. After randomization, the difference becomes negative, as the treatment group has fewer crime reports than the control group by approximately 0.04 reports per quarter.

While the point estimates for any given quarter are not statistically significantly different from zero, the time pattern suggests that crime reports fell with greater access to the program. To summarize these findings and gain statistical precision, Table 5 reports the regression results. The cross-sectional and difference-in-differences models show that crime falls by about 0.038 crime reports per quarter. The cross-sectional model has a standard error of 0.020, while the difference-in-differences model has a standard error of 0.013. The rate of criminal justice involvement each quarter is relatively high, at 0.12 crime reports per quarter for the control group in the post-randomization period, and the treatment effect estimate implies a fall of 32% relative to this mean.

Columns (3)-(5) report the difference-in-differences results by type of crime. The estimates suggest a reduction in crime reports for property, violent, and "other" crimes, which include a range of offenses from vandalism to weapons possession.[21] These results show that crime reports fall for a wide range of crimes, including serious ones with more significant welfare implications.

We find heterogeneity by sex for crime reports. The number of crime reports for boys in the treatment group falls by 0.07, or 33% of the control-group mean for boys in the post-period (0.20). For girls, the estimated treatment effect is -0.012, representing a 19% decline compared to the mean number of crime reports per quarter for this group

[21] Drug crimes are very rare in our data due to how the Ministry of Justice treats these offenses (we observe a control group mean of 0.003).

in the post period (0.064).[22] This difference in the level of the treatment effect across boys versus girls is marginally significant (p = 0.06).

5.6 Schooling Outcomes

Another set of wellbeing measures is whether children are attending and performing well in school. Figure 6 reports the monthly event study for the attendance rate across the two groups and the average attendance rate across all children in Chile each month (shown with black diamonds). Summer months (for which there is no data) are not presented.

The event study results show that the difference in attendance rates before the randomization is close to zero, with average attendance somewhat lower for the treatment group than the control group. After randomization, we see that the difference in the attendance rate for the treatment group relative to the control group changes sign and is now positive on average.

We find larger differences across the treatment and control groups after randomization in months when the average attendance rate in the country is lower. For example, attendance was low across all students in June 2019, concurrent with a national teacher strike on June 3rd that lasted until July 9th, and this is a month where we see a relatively large difference in attendance between the treatment and control groups. We also observed a positive difference in the last two months, which were periods of lower-than-usual attendance due to social unrest. The difference in attendance across the two groups is largest when the decision to attend school appears more discretionary.

Table 6 reports the regression results using monthly data, excluding the months when school is not in session. Column (1) reports the cross-sectional estimate, which shows that attendance is higher after randomization for the treatment group by 2.6 percentage points. However, this difference is not statistically significant with a standard error of 0.17 percentage points. Column (2) reports a similar treatment effect using the difference-in-differences model of 2.9 percentage points, which is statistically significant with a standard error of 0.13. This is driven in large part by the difference during the teacher strike in June 2019. Relative to the pre-period, the difference is

[22]The ITT estimate for girls is the sum of the coefficients on treatment*post and treatment*post*female.

4.7 percentage points (s.e. = 0.026), or 11% of the mean. Further, the cross-sectional and DD estimates are not statistically significantly different from one another (p = 0.75). The estimated increase in attendance is substantial given that the mean attendance rate in the post period is 58% in our analysis sample among the control group. Column (3) shows that the effect is larger for boys, whose increase in attendance is 4.6 percentage points, but the difference across boys versus girls is not statistically significant (p = 0.28).[23] Taken together with the main results in the event study, it appears that access to the program increased school attendance, although the increase is concentrated during times when schooling was more discretionary.

Columns (4) and (5) report school performance results in 2019 in terms of grade percentile. These point estimates suggest a modest improvement for boys and a modest decline for girls, but the estimates are imprecise and not statistically significant.[24]

5.7 Results Summary

The results show that those in the treatment group had higher program engagement, returned home faster, had lower criminal justice involvement, and improved school attendance. We do not observe an increase in hospitalization or subsequent reports of maltreatment despite children in the treatment group being home for more days and under greater surveillance by the program. Taken together, it appears that providing this intensive case management had beneficial effects on wellbeing.

5.8 Heterogeneous Treatment Effects and Robustness Checks

5.8.1 Heterogeneous Treatment Effects

Understanding heterogeneous treatment effects could inform efforts to target the program. In addition, if the program improves outcomes for those at the highest or lowest risk of the outcomes, we learn about the types of cases with more malleable outcomes, which can help inform other programs aimed at improving child welfare. By comparing living-with-family results with the criminal justice and schooling outcomes, we can also

[23] The difference relative to the control-group mean for girls is 3%, while for boys the point estimate represents an 8% increase compared to their control-group mean.

[24] The estimation sample is an unbalanced panel with 2018 and 2019 percentile grades. The ITT estimate for the balanced sample is -0.65 compared to -0.71 reported in Table 6

learn whether these improvements coincide or are relatively independent.[25] We treat these comparisons as exploratory, as we often lack statistical precision when estimating differences in treatment effects across groups even without adjusting for multiple hypothesis tests.[26]

We first examined estimated effects for the randomization strata: sex, age group, and region (Table 7). We find marginally significant differences across the strata for crime reports. As noted above, boys have larger reduction in crime reports compared to girls (p = 0.06). Similarly, older children have a larger reduction in crime compared to those under the age of 12 (p = 0.03), likely due to the lower prevalence in crime for younger children.

The third set of randomization strata represents the four geographic regions.[27] We do not find statistically-significant differences in treatment effects across the four regions. Of the 12 point estimates for the 3 outcomes across 4 regions, they share the same sign with the exception of school attendance in Valparaíso. When we examine the results leaving out one region at a time, the results are always statistically significant for the duration in foster care; the crime estimates are always statistically significant except when we exclude Biobío; and the point estimates are qualitatively similar to the pooled estimates (and not statistically distinguishable) across all of the estimates. In particular, we find larger improvements in Biobío compared to the Metropolitan region in terms of living-with-family (p = 0.09), crime (p = 0.14), and school attendance (p = 0.49).

Next we systematically considered heterogeneity across different types of cases defined by their baseline characteristics. Appendix G.3 shows that there are not statistically-significant differences across the groups when we adjust for multiple hypothesis testing within families of characteristics. To gain precision and begin to un-

[25]We find that across the subgroups, the treatment group experiences substantially more exposure to the program as designed (see Appendix Table G.1).

[26]We explored machine learning methods to identify heterogeneous treatment effects, but existing methods are meant for settings where the share treated is closer to 0.5 (Chernozhukov et al. (2018); Athey and Wager (2019)). When we estimated the models, they failed a specification check: the estimates of heterogeneity were not predictive of actual heterogeneity. Indeed, the correlation between predicted and actual heterogeneity had the wrong sign and was imprecisely estimated, providing further evidence that these methods are not appropriate in our context.

[27]We find that the child characteristics are balanced across treatment and control in each of the four regions (Appendix Table G.2).

derstand the types of cases experiencing improvements, we construct two indexes of the propensity to return home or accrue crime reports. To implement this exploration, we first predicted which children were more likely to return home within one year of randomization and which were more likely to be reported for crimes over the same period.[28] Conditional on the other covariates, we found that returning home is associated with being female, living in the Maule region, entering care with an allegation of abuse as opposed to neglect, and being older at the time of first foster care entry; meanwhile, crime reports were positively related with being male, older, residing in Biobío, and having less time in residence at baseline.

We then divided the sample into categories based on the median of the predicted-outcome indexes. Table 7 shows that the only statistically-significant difference is for predicted crime: crime improvements are found for children more likely to have crime reports, such as older children and boys. Point estimates show that improvements are slightly larger for those predicted to remain in care longer across the three outcomes.

Another heterogeneity analysis worth noting is that the effects are more prominent in magnitude for children living in larger residences across all of our three outcomes. Larger residences have above-median number of children living in the facility in 2019, and they tend to be publicly run. These facilities house more complex cases. This finding is consistent with the largest impacts being concentrated among children with higher predicted crime reports.

5.8.2 Robustness Checks

The main results are robust to several empirical decisions. First, Appendix Table H.1 demonstrates that the difference-in-differences point estimates are identical if we include time-invariant controls, individual fixed effects, calendar-quarter fixed effects, or season fixed effects.[29] The clustered standard errors vary slightly across the empirical models.

Appendix Table H.2 provides additional robustness checks. The results are not

[28] Specifically, we regressed each outcome on the demographic and allegation characteristics (Appendix Table G.4). We estimated the predicted-outcome models using the control group, and within the control group the predicted outcome is calculated using a leave-out regression to avoid a child's outcome from informing his or her prediction following Abadie et al. (2018).

[29] As described in the appendix, the DD (and event-study) models already absorb these effects through the differencing in models where (1) the sample is balanced and (2) there are no continuous regressors).

sensitive to how we aggregate results across regions, whether re-weighting the data so that areas with more variance in treatment are not over-weighted, or taking a weighted average of the four regional estimates, where the weights are related to the relative precision of the estimates (Gibbons et al., 2019; Athey and Imbens, 2017).[30] Results are also not sensitive to the way the 51 children who were randomized to treatment when capacity became available in May 2019 are included. We also examine results in the first year after randomization so that they are unaffected by the COVID pandemic, with similar results.[31]

6 Interpretation

6.1 Exploring Mechanisms

6.1.1 Program Components

To explore mechanisms, we conducted interviews with the program administrators and benefited from a qualitative review of the program conducted in 2020 (FOCUS, 2020). The interviews and the report highlight that (1) the lawyer plays a dominant role in the program, and (2) the complementarity between the psychologist and social worker is seen as a crucial ingredient in the program's effectiveness. The lawyer is primarily responsible for legal tasks such as filing court briefs, interviewing and interacting with children, their families, and staff from residential facilities. Their involvement is thought to improve the efficiency of judicial processes and accelerate families' readiness for reunification. In addition, lawyers handle fewer cases compared to psychologists and social workers. As a result, "lawyers form the backbone of the program." Meanwhile, the psychologists and social workers conduct child and family assessments, which are meant to inform the lawyers' strategies and facilitate therapeutic referrals. They are seen as necessary to address the complex needs of institutionalized children and their

[30] When comparing the baseline estimates to the model that re-weights the estimates or aggregates them, the days living-with-family coefficient increases from 6.5 to 6.9 and 7.2, respectively; the crime report coefficient increases in magnitude from -0.037 to -0.043 and -0.052; and the attendance magnitude changes from 0.029 to 0.030 and 0.024.

[31] Last, McKenzie (2012) recommends an ANCOVA model similar to our cross-sectional estimator to incorporate longitudinal data. Appendix Table H.2 shows the results are the same as the cross-sectional estimates for completeness.

families. Further, they are thought to improve child outcomes by fostering supportive relationships that may shift children's perceptions of the legal system from adversarial to protective. Last, they raise awareness of non-legal issues, such as inadequate living conditions. The courts can then act to mitigate risk factors that can significantly affect the children's behavior.

To complement this qualitative analysis, we used the program's process-tracking data to investigate the roles of the different components. With these data, we can distinguish between interdisciplinary processes (more than one member involved), judiciary processes (only the lawyer is involved), and psychosocial processes (only the psychologist or the social worker is involved). As described above, 57% of processes are interdisciplinary while 13% are led solely by a psychologist or social worker (see Appendix Table B.1). We used the baseline characteristics to predict the share of each type of service used. Appendix Table I.1 reports that younger children and those with more time in care are more likely to receive interdisciplinary services rather than services provided solely by the lawyer. We then compared results for those with above/below median predicted shares of these types of services. The exercise shows that the impacts on living-with-family are larger for children predicted to use a high share of psychosocial processes, while the the effect on reducing crime reports is smaller for this group (Appendix Tables I.2–I.4).[32] While the results do not reveal a consistent pattern whereby particular services improve the full set of wellbeing measures, the results corroborate the administrators' view that the interdisciplinary team is important for the outcome they observe directly, namely family reunification.

We also constructed a measure that the child had any writs or hearings each quarter to gauge whether the program ensured a steady, minimal level of effort by the lawyer assigned to the case. Its mean level is close to 90% for both the treatment and control groups, with a slight increase for the treatment group after randomization that is not statistically significant. We take this as evidence that the program effects did not work through ensuring that each child had minimal legal representation.[33]

[32]We conducted a similar analysis investigating the different types of services across stakeholders: the child, the family, the residences, and the court system. This, too, did not reveal a straightforward pattern of results, perhaps due to these measures varying across different types of families in ways that can confound the comparisons.

[33]Another possible pathway to explain the program impacts is that Mi Abogado could have selected higher quality lawyers for the children. We attempted to estimate measures of Lawyer Value Added (risk-adjusted

6.1.2 Relationship between Time in Care and Criminal Justice Involvement

One motivation for the Mi Abogado program was to reduce time in foster care. It would be helpful to know if the crime-report reduction was due to the reduction in length of stay, the services received by children and families, or some combination of the two mechanisms. While measuring bureaucratic delays directly is challenging, we observe the time until living with family as a summary measure that reflects these delays.

The subgroup heterogeneity suggests that the two outcomes are not linked in this way. By and large, subgroups where treatment accelerated exit from foster care are not the same groups that experienced a reduction in crime reports or improved attendance. To investigate this further, we estimated a model of crime reports on treatment status while controlling for (endogenous) time in a residence in a given quarter as a mediation analysis. Our estimated effect of the program on crime reports is not affected by controlling for time in a residence (Appendix Table I.5). This again suggests that time in a residence is not driving the main crime results. A leading alternative explanation for reducing crime reports stems from stronger child and family rehabilitation services offered to children in foster care as part of the Mi Abogado program (Grimon, 2021; Gross and Baron, 2022; Baron and Gross, 2022).

We also tested whether time in care and crime are mechanically related through greater surveillance by residence staff. Appendix I.3 shows that crime does not rise or fall discontinuously with entry into or exit from a SENAME residence. This suggests that greater surveillance in residences is not driving the estimated effects of the program on crime reports.

---

mean outcomes measured at the lawyer level). The quality of the data on case end dates and exact dates of lawyer assignment result in imprecise measures of value added that are not predictive out of sample. Out of around 17,000 quarters/children since 2018, we are able to identify an active lawyer for 74% of them. Not surprisingly, our main impact results are not different if we control for this measure of lawyer effectiveness. Given the lack of precision and the potential endogeneity of lawyer assignment, we do not rely on this evidence in the interpretation.

6.2 Treatment Effect Dynamics

6.2.1 Effects across Time Since Program Exposure

These intent-to-treat estimates measure the effect of offering the program on child well being. A related question is how the effects evolve with program exposure as opposed to time since randomization. The answer would provide a better understanding of the sources of the intent-to-treat differences.

We address this question by estimating a recursive model, assuming that the dynamic impacts of the program are the same across program-entry cohorts. The intuition is as follows: we identify the effect of one quarter of exposure by relating our ITT estimate after one calendar quarter to the difference in exposure between treatment and control groups at that time. In the second calendar quarter after randomization, two effects occur simultaneously: (i) a second program-entry cohort begins its first quarter of exposure, and (ii) the first program-entry cohort enters its second quarter of exposure. We disentangle these effects by assuming that the impact of the first quarter of exposure is identical across cohorts (i.e. no cohort-specific dynamic effects), and we have an estimate of the impact of one quarter of exposure from the initial quarter after randomization. Applying this reasoning recursively, we identify the dynamic effects of varying program durations. Next, we explain this strategy in detail.

We can make progress if we assume that the environment and effectiveness of program exposure do not change over time. In the first year, this could be violated by seasonality in these outcomes. In the second year of the study period, this is most likely violated by the global pandemic. So, we view our dynamics analysis as speculative and more reliable over the first year after randomization.

For clarity, it helps to consider a more straightforward context. Suppose (i) all treated children entered the program at the same time, (ii) no control children participated, and (iii) there were no time shocks to the effects of the program so we can make comparisons across calendar time; then we could identify how treatment effects change with program exposure by observing the difference in outcomes between the treatment and control groups over time.

Those conditions are not met in practice, so we must impose some structure that relies on additional assumptions. In particular, we assume that (i) the treatment effects are homogeneous across cohorts (so, there is no difference between the treatment effects of compliers entering early and those entering late) and (ii) calendar time does not interact with treatment effects. Under those assumptions, our strategy calculates the difference in outcomes between the entire treatment and control groups period by period. This implies that the estimates are always made across comparable groups to minimize the risk of endogeneity. The idea behind the approach is that in the first observation quarter after randomization, we can identify the *effect of being exposed to the program for one quarter* by comparing outcomes and program exposure across the treatment and control groups. We can then use this estimate in the second observation quarter to predict the effect on children who are exposed for one quarter because they are entering the program at that time. The remainder of the difference across treatment and control identifies the *effect of being exposed to the program for two quarters.* Using this method recursively, we obtain identification of the dynamic effects of the program, which can depend nonlinearly on the number of periods since program exposure.[34]

More specifically, let the outcome in a given quarter for a given individual depend on the total time in the program. Let e^j_{iq} be an indicator that takes the value of 1 if an individual i has spent j quarters since their first exposure to the program up to quarter q (inclusive) and 0 otherwise. The quarters are defined such that the second quarter of 2019 is Quarter 1. For example, a given individual i entering the program in Quarter 1 will have $e^1_{i1} = 1$ when $q = 1$, $e^2_{i2} = 1$ when $q = 2$, and so on. People entering Quarter 2 (third quarter of 2019) will have $e^1_{i1} = 0$ when q=1, $e^1_{i2} = 1$ when $q = 2$, $e^2_{i3} = 1$ when $q = 3$, and so on. This definition of e^j_{iq} is simply equivalent to indicating program cohorts, with newer cohorts having participated in the program for less time.

We can let the effects of the program vary in a non-parametric way using the following specification, where β^j will capture the cumulative effect of having been exposed to the program for j quarters:

$$Y_{iq} = \alpha_q + \sum_{j\in\{1\dots q\}} \beta^j e^j_{iq} + v_{iq} \tag{4}$$

[34]This is similar in spirit to Cellini et al. (2010), who employ a recursive method to use estimated intent-to-treat estimates to trace dynamic treatment effects of school facility investments on housing prices.

To save notation, let $\Delta X_q \equiv E[X_q|T = 1] - E[X_q|T = 0]$, the simple difference of means between the treatment and the control group in quarter q, where T is an indicator of having been randomly assigned to the program.

From Equation 4, a simple difference in means of the observed outcomes in Quarter 1 implies:

$$\Delta Y_1 = \beta^1 \Delta e_1^1 \tag{5}$$

Then, the impact of one quarter in the program can be identified by the usual IV estimator, taking differences across the treatment and control groups: $\beta^1 = \frac{\Delta Y_1}{\Delta e_1^1}$.

Taking the difference in means in the second calendar quarter, we obtain:

$$\Delta Y_2 = \beta^2 \Delta e_2^2 + \beta^1 \Delta e_2^1 \tag{6}$$

This implies that the difference between the treatment and control groups in the second calendar quarter is given by (i) the effect of two periods in the program experienced by those who entered in the first quarter, and (ii) the effect of a single quarter in the program, experienced by those who entered in the second quarter. As the latter effect is already identified, we can plug in our estimate of β^1, solve for β^2, and identify it from the data:

$$\Delta Y_2 = \beta^2 \Delta e_2^2 + \beta^1 \Delta e_2^1 \tag{7}$$

$$\hat{\beta}^2 = \frac{\Delta Y_2 - \hat{\beta}^1 \Delta e_2^1}{\Delta e_2^2} \tag{8}$$

Using this method recursively, we can estimate the dynamics of the treatment effects for all periods. Standard errors are bootstrapped and incorporate the variance stemming from the calculation of the plug-in estimates.

The results are presented in Figure 7. The figure shows that the program's effects on both living-with-family and crime reports are almost linear as the effects of program exposure grow over the first two years.

As noted, however, caution is warranted in interpreting the results due to the assumptions required. In particular, the approach assumes that the effects of program exposure are unrelated to calendar time. In this case, the effects appear to level off around the fourth quarter after randomization and then begin to grow in magnitude again. That second increase is in the second quarter of 2020, when the COVID-19 pandemic began. To the extent that the quarters prior to the pandemic are more informative, then we can focus on the effects for the first year after randomization; they imply that the effects grow with time-since-exposure for at least one year.

6.2.2 Compliers over Time

A related complication when estimating the effects of the program over time is that the nature of the participants can change over time as well. We can also learn more about how judges adopted the program by tracking changes in the characteristics of compliers—those who were induced into the program due to their treatment status.

We estimate the average characteristics of compliers by observing the subgroups with larger first-stage estimates for program exposure relative to the first stage for the sample as a whole (Angrist and Pischke, 2008). Appendix Table J.1 shows that there is a strong first stage for all of the subgroups. The estimates imply that compliers were more likely to be younger, girls, and had a longer time in residence. As a summary measure, we find compliers are those with a lower predicted number of crime reports. We find similar patterns when we identify the types of children who are compliers over the first two years after randomization rather than just one year (Appendix Table J.2). This similarity suggests that changing complier characteristics does not complicate the interpretation of the evolution of outcomes in our context.

6.3 Cost-Effectiveness

Investing in quality-improvement programs may face budgetary hurdles, and evidence of a return on that investment can spur adoption. The intent-to-treat estimates measure benefits and costs of assignment to the treatment group, which offers a straightforward way to make these comparisons.[35]

6.3.1 Cost-Effectiveness

Table 8 summarizes the costs and benefits to SENAME. First, children in the treatment group participated in Mi Abogado for 91 more days, on average, compared to the

[35] We calculate the effects over the entire time period we observe, a total of nearly two years (641 days). All costs are in 2022 US dollars.

control group. Conversely, children in the treatment group were assigned to non-Mi Abogado lawyers for 99 fewer days. Mi Abogado has an average cost of \$4.99 per child per day, while non-MA lawyers cost \$2.73 per child per day, as calculated by the Interagency Roundtable according to the Ministry of Justice.[36] Overall, offering the program increased legal-aid costs by \$182 per child during the nearly two-year observation period.

Meanwhile, treated children spent 4.8 fewer days in government residences than the control group and 26.4 fewer days in private residences. The former cost the child protection agency \$67.27 per child per day, while the latter cost \$28.35 per child per day.[37] We also observe a small increase in days in family foster care as children leave residences for this placement setting.[38] In total, SENAME saved \$817 per child by offering them the program. The estimates imply that for every dollar of additional legal aid, SENAME saved \$4.50.[39]

While these savings are relevant to the budget of the child protection agency, from a societal perspective there are other costs and benefits to consider. A limitation of the analysis is that we do not observe the social services that may increase due to participation in Mi Abogado. However, our findings for criminal justice and schooling outcomes suggest that additional benefits are likely substantial. Appendix K provides estimates when including a measure of societal savings associated with the reduction in crime reports. After including crimes, the total benefits from the crime reduction plus savings to SENAME is \$4,152 per child from offering the program. While caution is warranted given the assumptions behind the calculations, the estimates reiterate that the program's benefits are likely greater than the savings that accrue to the child protection agency alone.

[36] \$4.99 per diem, or \$150 per child per month, is the average cost of the program, which we confirmed with the program staff and their regional reports. See Appendix B for more details.

[37] Costs for public (Centros de Reparación Especializada de Administración Directa, or CREAD) residence were obtained from program monitoring documents in 2020 and nonprofit (Organismos Colaboradores, or OCAS) residence and family foster care costs as established in Law 21140.

[38] We excluded the category of directly administered family foster care because it is rare. The difference in the number of days in SENAME care across these categories sums to 26 days over these two years, as expected based on the results in Table 3. The days with a lawyer differ from days in SENAME in part because the Mi Abogado program continues to aid children for 90 days after exit from SENAME.

[39] From Table 8, (Savings from fewer days in care) - Additional Legal-Aid Spending) / Additional Legal-Aid Spending = -817.58/(451.87-270.09) = 4.50.

7 Conclusion

Child protection involves far-reaching interventions into the lives of children and families, and more rigorous evidence is needed to inform efforts to increase the quality of foster care services. This study demonstrates that as new programs are introduced, the rollout can be structured in a way to evaluate their effects. Coupled with administrative data, we can examine the effects on a primary goal of the program—stable family reunification or adoption—and additional welfare-relevant outcomes for broader measures of wellbeing: safety, criminal justice, and schooling outcomes.

We find that the randomly assigned treatment group had 60% greater exposure to the program over the two years after the program's introduction. This additional treatment resulted in substantial increases in family reunification or adoption, no detectable decline in child safety, a decline in criminal justice involvement, and an improvement in school attendance. For all of these outcomes, the point estimates suggest greater improvements for boys. Both qualitative and quantitative evidence suggest that the combination of a lawyer with a psychologist and social worker connects children and families with services that improves their wellbeing. This enhanced legal aid appears highly cost effective to the child welfare system, as it reduces time in foster care.

The experimental design of a government policy at scale provides internally-valid estimates for a large number of children in Chile's foster care system. The estimation has limitations, however. First, the results may not apply to settings outside of Chile with better functioning legal structures or higher-quality foster care systems, although concerns over bureaucratic frictions and long stays in foster care are common worldwide. Second, we observe only a subset of welfare-relevant outcomes; we view the crime and education results as proxies for the broader construct of child wellbeing. In particular, we do not observe the social services received by foster children and their families, which affects the interpretation of societal costs and benefits. Third, the nature of the pragmatic randomized trial with unequal treatment and control group sizes limits our precision, especially when comparing effects across subgroups.

Despite these limitations, the results suggest that expanded legal aid is a reform that reduce time in foster care while improving child wellbeing more generally. Efforts to scale the program to be even larger would need to take into account effects on the quality of legal aid as more lawyers are recruited, along with the opportunity cost of the productive capacity of the legal team if employed elsewhere. Nevertheless, the results should add urgency to policy and practice that attempts to improve the quality of foster care.

References

Abadie, A., M. M. Chingos, and M. R. West (2018, 10). Endogenous Stratification in Randomized Experiments. *The Review of Economics and Statistics 100*(4), 567–580.

Anderson, J. M., M. Buenaventura, and P. Heaton (2019). The effects of holistic defense on criminal justice outcomes. *Harvard Law Review 132*, 819–893.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Anwar, S., S. Bushway, and J. Engberg (2023). The impact of defense counsel at bail hearings. *Science Advances 9*(18), eade3909.

Athey, S. and G. W. Imbens (2017). The econometrics of randomized experiments. In *Handbook of economic field experiments*, Volume 1, pp. 73–140. Elsevier.

Athey, S. and S. Wager (2019). Estimating treatment effects with causal forests: An application. *Observational Studies 5*(2), 37–51.

Bald, A., E. Chyn, J. S. Hastings, and M. Machelett (2022). The causal impact of removing children from abusive and neglectful homes. *Journal of Political Economy 130*, 1919–1962.

Bald, A., J. Doyle Jr., M. Gross, and B. Jacob (2022). Economics of foster care. *Journal of Economic Perspectives 36*(2), 223–246.

Baron, E. J. and M. Gross (2022, April). Is there a foster care-to-prison pipeline? evidence from quasi-randomly assigned investigators. Working Paper 29922, National Bureau of Economic Research.

Becker, M. A., N. Jordan, and R. Larsen (2007). Predictors of successful permanency planning and length of stay in foster care: The role of race, diagnosis and place of residence. *Children and Youth Services Review 29*(8), 1102–1113.

Bender, K., J. Yang, K. Ferguson, and S. Thompson (2015). Experiences and needs of homeless youth with a history of foster care. *Children and Youth Services Review 55*, 222–231.

Bergström, M., M. Cederblad, K. Håkansson, A. K. Jonsson, C. Munthe, B. Vinnerljung, I. Wirtberg, P. Östlund, and K. Sundell (2020). Interventions in foster family care: A systematic review. *Research on Social Work Practice 30*(1), 3–18.

Blome, W. W. and S. Steib (2008). An examination of oversight and review in the child welfare system. *Journal of Public Child Welfare 1:3*, 3–26.

Cassidy, M. T. and J. Currie (2022, March). The effects of legal representation on tenant outcomes in housing court: Evidence from new york city's universal access program. Working Paper 29836, National Bureau of Economic Research.

Cellini, S. R., F. Ferreira, and J. Rothstein (2010, 02). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design*. *The Quarterly Journal of Economics 125*(1), 215–261.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernández-Val (2018, June). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. Working Paper 24678, National Bureau of Economic Research.

Cho, M., W. Haight, W. S. Choi, S. Hong, and K. Piescher (2019). A prospective, longitudinal study of risk factors for early onset of delinquency among maltreated youth. *Children and Youth Services Review 102*, 222–230.

Cooley, M. E., H. M. Thompson, and M. L. Colvin (2019). A qualitative examination of recruitment and motivation to become a guardian ad litem in the child welfare system. *Children and Youth Services Review 99*, 115–124.

Cooper, R., J. Doyle, and A. Hojman (2022). Evaluacion de impacto del programa mi abogado. Technical report, DIPRES.

Courtney, M. E. and J. L. Hook (2012). Evaluation of the impact of enhanced parental legal representation on the timing of permanency outcomes for children in foster care. *Children and Youth Services Review 34*(7), 1337–1343.

de Chaisemartin, C. and X. D'Haultfœuille (2024, 02). Difference-in-Differences Estimators of Intertemporal Treatment Effects. *The Review of Economics and Statistics*, 1–45.

De Iruarrizaga, F. (2016). Rediseñando el sistema de protección a la infancia en chile. entender el problema para proponer modelos de cuidado alternativos y ayudar a la reunificación familiar. *Estudios Públicos 141*, 7–57.

de Chaisemartin, C. and X. D'Haultfœuille (2022, 06). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey. *The Econometrics Journal 26*(3), C1–C30.

Doyle, J. J. (2008). Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. *Journal of political Economy 116*(4), 746–770.

Doyle, Jr, J. J. (2013). Causal effects of foster care: An instrumental-variables approach. *Children and Youth Services Review 35*(7), 1143–1151.

Duerr Berrick, J., N. Gilbert, and M. Skivenes (2023, 02). *Oxford Handbook of Child Protection Systems.* Oxford University Press.

Dworsky, A., L. Napolitano, and M. Courtney (2013). Homelessness during the transition from foster care to adulthood. *American Journal of Public Health 103*(S2), 318–323.

Fallesen, P., N. Emanuel, and C. Wildeman (2014). Cumulative risks of foster care placement for danish children. *PLoS One 9*(10), e109207.

Farber, J., L. Bensky, and L. Alpert (2009). *The long road home: A study of children stranded in New York City foster care.* Children's Rights Organization.

FOCUS (2020). Levantamiento de línea base del programa mi abogado: Informe final. Technical report, FOCUS.

Font, S., L. M. Berger, J. Slepicka, and M. Cancian (2021). Foster care, permanency, and risk of prison entry. *Journal of Research in Crime and Delinquency 58*(6), 710–754.

García, M. and C. Hamilton-Giachritsis (2014). "in the name of the children": Public policies for children in out-of-home care in chile. historical review, present situation and future challenges. *Children and Youth Services Review 44*, 422–430.

Gerber, L. A., Y. C. Pang, T. Ross, M. Guggenheim, P. J. Pecora, and J. Miller (2019). Effects of an interdisciplinary approach to parental representation in child welfare. *Children and Youth Services Review 102*, 42–55.

Gibbons, C. E., J. C. Suárez Serrato, and M. B. Urbancic (2019). Broken or fixed effects? *Journal of Econometric Methods 8*(1), 1–12.

Grimon, M.-P. (2021). *Essays in Labor Economics and Child Welfare*. Ph. D. thesis, Harvard University.

Gross, M. and E. J. Baron (2022). Temporary stays and persistent gains: The causal effects of foster care. *American Economic Journal: Applied Economics 14*(2), 170–199.

Haarberg, F. L. (2024). What do we know about children's representation in child protection decisions? a scoping review. *Children and Youth Services Review 160*, 107588.

Harris, H. M. (2020). Building holistic defense: The design and evaluation of a social work centric model of public defense. *Criminal Justice Policy Review 31*(6), 800–832.

Helensdotter, R. (2024). Surviving childhood: Health and crime effects of removing a child from home. https://drive.google.com/file/d/1SVc4sGHy6UzyHmQOwTfkj7Oc_RZduf-N/view.

Hirsch, R. A., C. B. Dierkhising, and D. C. Herz (2018). Educational risk, recidivism, and service access among youth involved in both the child welfare and juvenile justice systems. *Children and Youth Services Review 85*, 72–80.

Hoynes, H. W., N. Maestas, and A. Strand (2022, March). Legal representation in disability claims. Working Paper 29871, National Bureau of Economic Research.

Hunter, D. R., P. A. Monroe, and J. C. Garand (2014). Understanding correlates of higher educational attainment among foster care youths. *Child Welfare 93*(5), 9–26.

INE (2009). *Anuario de estadísticas vitales.* Instituto Nacional de estadísticas. Ministerio de Salud. Servicio Nacional de Registro Civil e idenifcación.

Johnson-Reid, M., T. Chance, and B. Drake (2007). Risk of death among children reported for nonfatal maltreatment. *Child Maltreatment 12*(1), 86–95.

Kim, H., C. Wildeman, M. Jonson-Reid, and B. Drake (2017). Lifetime prevalence of investigating child maltreatment among us children. *American Journal of Public Health 107*(2), 274–280.

Konijn, C., S. Admiraalb, J. Baartb, F. van Rooijb, G.-J. Stamsb, C. Colonnesib, R. Lindauerc, and M. Assinkb (2019). Foster care placement instability: A meta-analytic review. *Children and Youth Services Review 96*, 483–499.

Lacoe, J., B. Fischer, and S. Raphael (2023). The effect of pre-arraignment legal representation on criminal case outcomes. Working Paper 31289, National Bureau of Economic Research.

Martinez, O., J. Boles, M. Muñoz-Laboy, E. C. Levine, C. Ayamele, R. Eisenberg, J. Manusov, and J. Draine (2017). Bridging health disparity gaps through the use of medical legal partnerships in patient care: A systematic review. *Journal of Law, Medicine and Ethics 45*(2), 260–273.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of Development Economics 99*(2), 210–221.

Miller, J. J., J. Donahue-Dioh, and L. Owens (2020). Examining the legal representation of youth in foster care: Perspectives of attorneys and attorney guardians ad litem. *Children and Youth Services Review 115*, 105059.

Miller, M. (2004). How do court continuances influence the time children spend in foster care? Technical report, Washington State Institute for Public Policy.

Muñoz-Guzmán, C., C. Fischer, E. Chia, and C. LaBrenz (2015). Child welfare in chile: Learning from international experiences to improve family interventions. *Social Sciences 4*(1), 219–238.

Muralidharan, K. and P. Niehaus (2017, November). Experimentation at scale. *Journal of Economic Perspectives 31*(4), 103–24.

Muñoz-Guzman, C., M. Cillero Bruñol, and M. Bernasconi (2023, 02). 399C20The Chilean Child Protection System. In *Oxford Handbook of Child Protection Systems*. Oxford University Press.

Okpych, N. J. and M. E. Courtney (2014). Does education pay for youth formerly in foster care? comparison of employment outcomes with a national sample. *Children and Youth Services Review 43*, 18–28.

Orlebeke, B., X. Zhou, A. Skyles, and A. Zinn (2016). *Evaluation of the QIC-ChildRep Best Practices Model Training for Attorneys Representing Children in the Child Welfare System*. Chapin Hall at the University of Chicago.

Osborne, C., H. Warner-Doe, M. LeClear, and H. Sexton (2020). The effect of casa on child welfare permanency outcomes. *Child Maltreatment 25*(3), 328–338.

Petrowski, N., C. Cappa, and P. Gross (2017). Estimating the number of children in formal alternative care: Challenges and results. *Child Abuse & Neglect 70*, 388–398.

Pilkay, S. and S. Lee (2015). Effects of court-appointed special advocate intervention on permanency outcomes of children in foster care. *Journal of Social Service Research 41*(4), 445–453.

Rashid, A. and G. R. Waddell (2019). Do lawyers increase the rate of adoption among foster children? Technical report, University of Oregon.

Rouland, B. and R. Vaithianathan (2018). Cumulative prevalence of maltreatment among new zealand children, 1998–2015. *American journal of public health 108*(4), 511–513.

Rushovich, B., A. Hebert, D. Crampton, and K. Malm (2021). Results from a randomized controlled trial of team decision-making. *Children and Youth Services Review 131*, 106263.

Ryan, T. N. and R. J. Gomez (2016). Trends in state budgets and child outcomes during and post child welfare class action litigation. *Children and Youth Services Review 62*, 49–57.

SENAME (2021). Acciones desde el ministerio de justicia y derechos humanos y el servicio nacional de menores para la implementación del servicio mejor niñez 2018-2021. Technical report, SENAME.

Sexton, V. (2018). Wait, who am i representing: The need for states to separate the role of child's attorney and guardian ad litem. *Georgetown Journal of Legal Ethics 31*(4), 831–846.

Shem-Tov, Y. (2022, 07). Make or Buy? The Provision of Indigent Defense Services in the United States. *The Review of Economics and Statistics 104*(4), 819–827.

Tsai, J., D. Jenkins, and E. Lawton (2017). Civil legal services and medical-legal partnerships needed by the homeless population: A national survey. *American Journal of Public Health 107*(3), 398–401. PMID: 28103065.

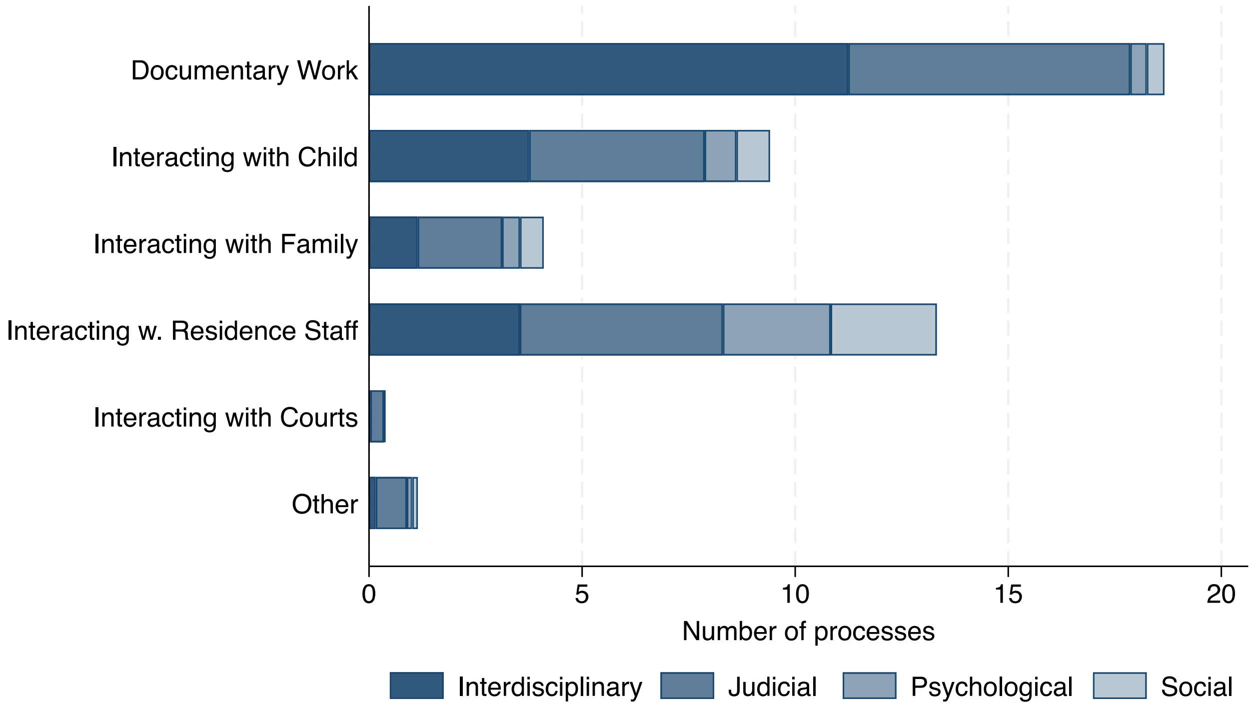
Turney, K. and C. Wildeman (2016). Mental and physical health of children in foster care. *Pediatrics 138*(5), 1–11.

USDoS (2019). Chile 2019 human rights report. Technical report, US Department of State, Bureau of Democracy, Human Rights, and Labor.

Yi, Y., F. R. Edwards, and C. Wildeman (2020). Cumulative prevalence of confirmed maltreatment and foster care placement for us children by race/ethnicity, 2011-2016. *American Journal of Public Health 110*(5), 704–709.

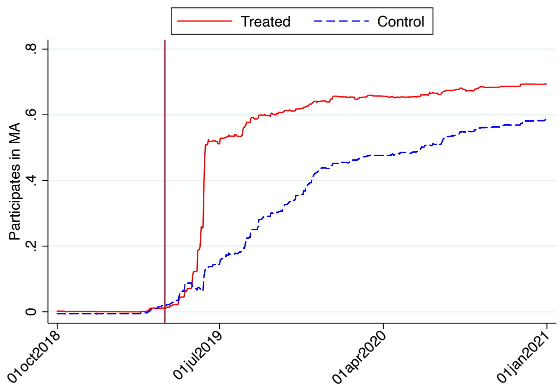
Zinn, A. and C. Peters (2015). Expressed-interest legal representation for children in substitute care: Evaluation of the impact of representation on children's permanency outcomes. *Family Court Review 53*(4), 589–601.

Figure 1: Mi Abogado Processes per Child per Year

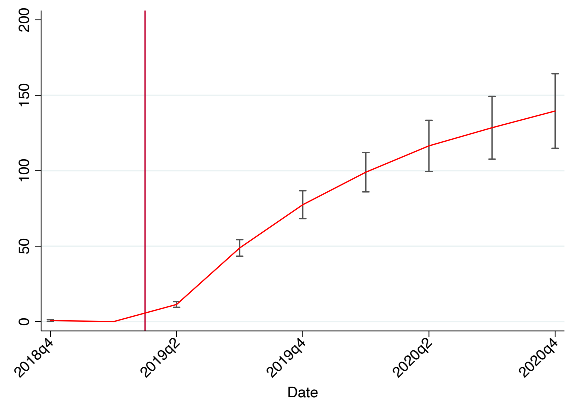


Note: This figure presents the average number of processes per child in their year after program initiation. The sample includes all Mi Abogado participants, not just those in our experimental sample, observed in the program for at least one year.

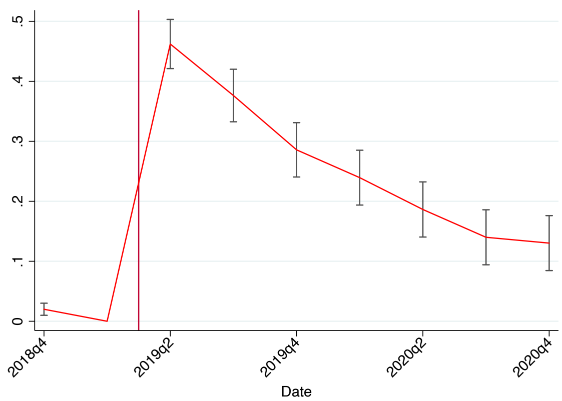
Figure 2: Mi Abogado Exposure and Engagement



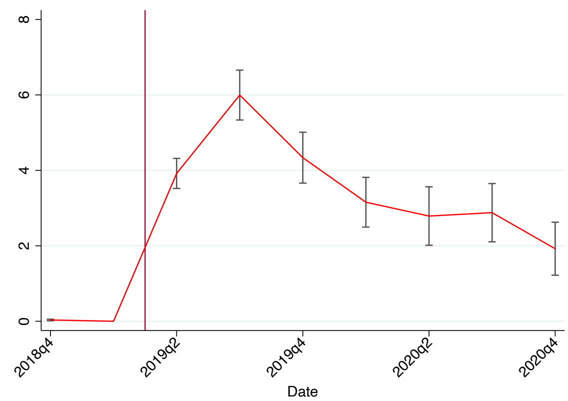
(a) Daily Program Exposure



(b) Cumulative Days of Exposure



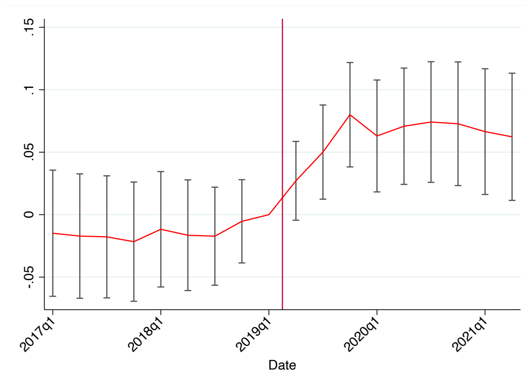
(c) Program Participation



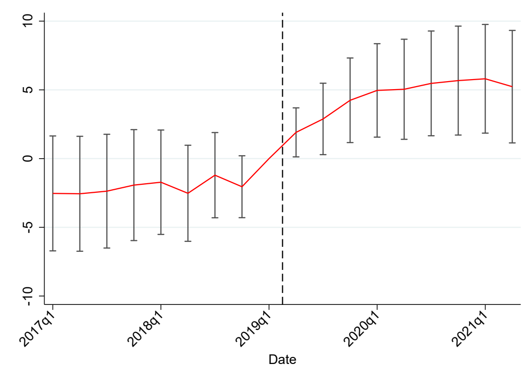
(d) Program Processes

Note: Figure 2a displays average program exposure (an indicator for ever having enrolled in the program by that date) for the treatment and control groups at the daily level, residualized of the randomization strata: sex, region and age group. Figures 2b-2d report event-study estimates binned at the calendar quarter level and described in Equation 1. Cumulative days of exposure measures the number of days since a child enrolled in the program. Program Participation measures whether the child ever participated that quarter. Program processes measure the number of Mi Abogado processes recorded each quarter. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization. Pre-period exposure and participation are not strictly zero because a small number of children were able to join the program during its pilot phase.

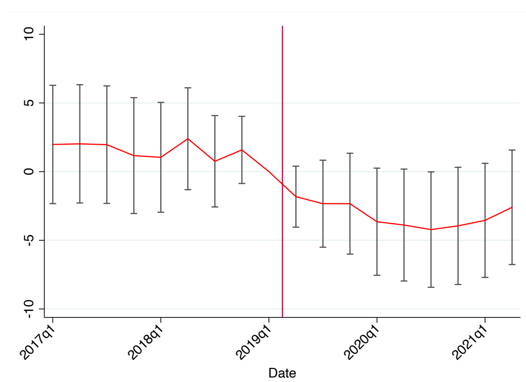
Figure 3: Living Arrangements



(a) Ever Living with Family



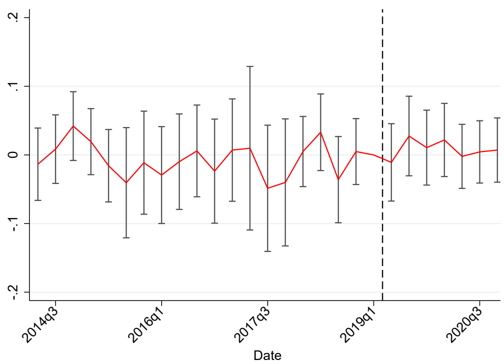
(b) Days Living with Family



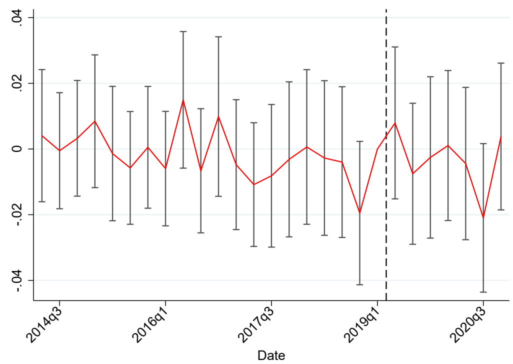
(c) Days in SENAME Residence

Note: These figures report event-study estimates described in Equation 1 of differences between the treatment and control groups for measures of living with a permanent (biological or adoptive) family and living in a SENAME residence. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

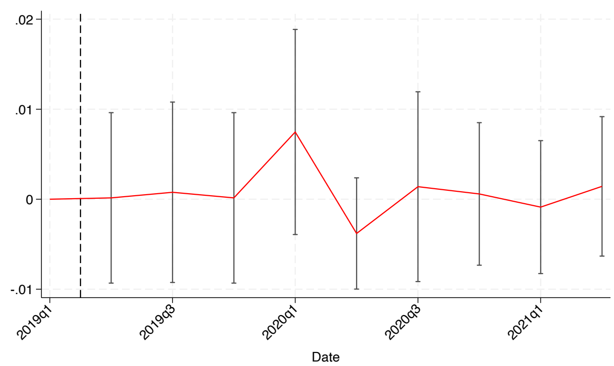
Figure 4: Child Safety Measures



(a) Child Protection Case Opening



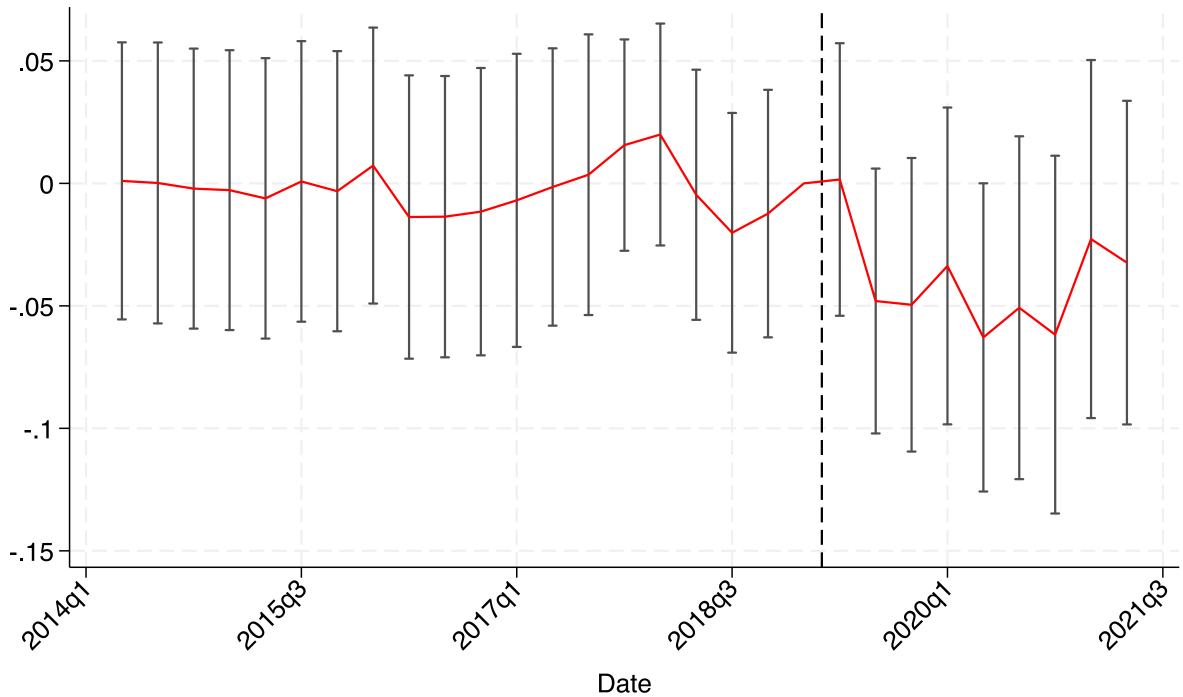
(b) Report of Child Victimization



(c) Foster Care Re-entry

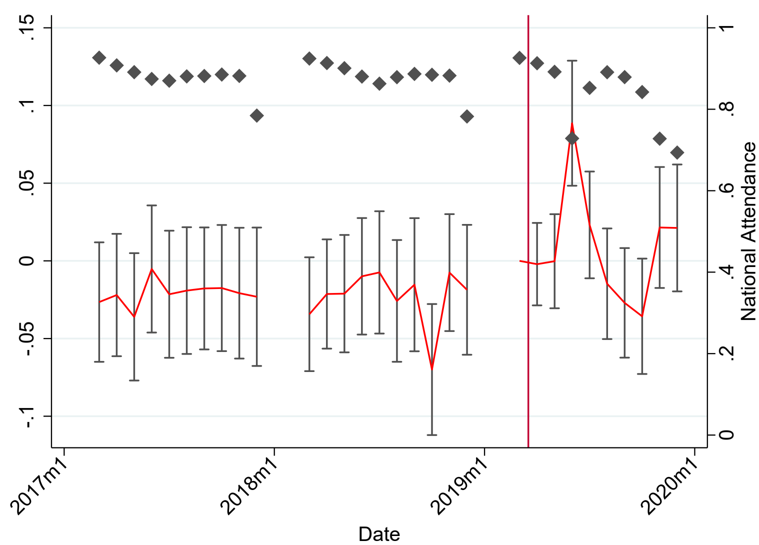
Note: These figures report event-study estimates described in Equation 1 of differences between the treatment and control groups for the number of child protection cases opened, number of criminal reports where the child is a victim, and number of foster care re-entries. For foster care re-entry, this is restricted to the post period, as all children are in foster care in Jan/Feb 2019. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

Figure 5: Crime Reports



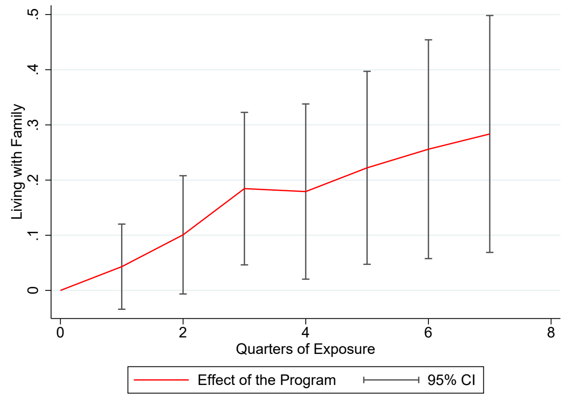
Note: This figure reports event-study estimates described in Equation 1 of differences between the treatment and control groups for the number of crime reports in a given quarter. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

Figure 6: School Attendance

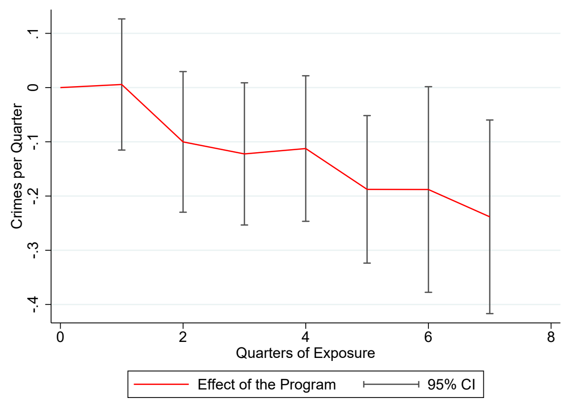


Note: This figure reports event-study estimates described in Equation 1 of differences between the treatment and control groups for the school attendance rate in each month rather than each quarter. The breaks in the event study represent the summer months when school is not in session. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization. Average attendance rates for all students in Chile each month are reported as black diamonds.

Figure 7: Treatment Effect Dynamics



(a) Living with Family



(b) Crime Reports

Note: These figures show estimated effects for each quarter of exposure to the program described in Equation 4. The first quarter is estimated using a Wald estimator. Subsequent quarters use the full sample and plug-in estimates from prior quarters as described in the text. Standard errors are calculated using bootstrap.

Table 1: Balance Across Baseline Measures

	Mean C	Mean T	SD	Dif	p-value
Writs/Qtr	2.80	3.02	2.48	0.23	0.16
Hearings/Qtr	0.19	0.21	0.17	0.01	0.24
Days Living with a Family/Qtr	26.20	24.27	31.06	-1.93	0.34
Days Living In a Residence/Qtr	61.57	63.99	31.66	2.42	0.24
Crime Reports/Qtr	0.03	0.04	0.14	0.01	0.28
Times Missing/Qtr	0.08	0.09	0.24	0.01	0.54
Times Victim of Abuse/Qtr	0.02	0.02	0.04	-0.00	0.61
School Percentage of Attendace	0.66	0.66	0.27	-0.01	0.76
Grades Percentile in 2018	26.87	28.73	24.28	1.86	0.34
Grades Percentile Missing	0.36	0.35	0.48	-0.01	0.65
Number of Siblings	1.40	1.31	1.98	-0.09	0.50
Delay in Schooling	1.01	1.15	1.36	0.14	0.12
Allegation: Sex Abuse	0.17	0.18	0.39	0.01	0.80
Allegation: Physical Abuse	0.26	0.30	0.45	0.04	0.23
Allegation: Neglect	0.85	0.83	0.36	-0.02	0.42
Age When First in Residence	10.83	10.73	3.69	-0.09	0.62
Age at Randomization	13.68	13.81	3.26	0.13	0.24
N of Children : 1871					
Joint F-Test p-value : 0.60					

Note: These baseline measures represent means in the pre-randomization period, and each measure's time period differs by data availability: writs and hearings from 2010, days in residence from 2017, days with family from 2017, criminal justice measures from 2014, and schooling for 2017-2018. The grades percentile measure is from 2018 and has a sample size of 1,222. Mean C is the mean for the control group. Mean T is calculated from a regression of the characteristic on a treatment indicator and strata indicators, and the coefficient on the treatment indicator is added to the control-group mean; SD is the control group standard deviation; Dif is the coefficient on the treatment indicator; and p-value is from the t-test for each coefficient. The F-test is for the joint significance of the measures in this table and is calculated from a model of treatment regressed on these covariates along with randomization-strata indicators.

Table 2: Mi Abogado Participation and Exposure

Dependent Variable:	(1) Days exposed to Mi Abogado/Qtr. Cross Section	(2) Days exposed to Mi Abogado/Qtr. Diff-Diff	(3) Days exposed to Mi Abogado/Qtr. Diff-Diff Including Sex Interaction	(4) Days participating in Mi Abogado/Qtr. Diff-Diff
Treatment Group	17 (2.29)***	-.407 (.268)	-.311 (.424)	-.419 (.269)
Participation Before Random	.447 (.0684)***			
Treatment x Post		20.3 (1.78)***	17.3 (2.73)***	13.5 (1.71)***
Post Randomization		31.5 (.967)***	35.8 (1.48)***	28.9 (.927)***
Female x Treat x Post			5.4 (3.6)	
Female x Post			-7.53 (1.94)***	
Female x Treat			-.167 (.548)	
Female			.0947 (.375)	
N	1,871	16,839	16,839	16,839
N of children	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188
Control Group Mean	32.266	32.266	32.266	29.639

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. The cross-sectional model is described in Equation 2, and heteroskedastic-robust standard errors are reported in parentheses. Difference-in-differences (Diff-Diff) models are described in Equation 3, and standard errors clustered at the child level are reported in parentheses. All models include strata indicators. Control Group Mean indicates the mean in the post-randomization period. Pre-period exposure and participation are not strictly zero because a small number of children were able to join the program during its pilot phase.

Table 3: Living Arrangements

Dependent Variable:	(1) Days Living w/ Family/Qtr. Cross Section	(2) Days Living w/ Family/Qtr. Diff-Diff	(3) Days Living w/ Family/Qtr. Diff-Diff Including Sex Interaction	(4) Ever Living w/ Family/Qtr. Diff-Diff	(5) Days Living in Residence/Qtr. Diff-Diff
Treatment Group	3.61	-2.52	-2.63	-.0305	1.77
	(2.14)*	(1.82)	(2.43)	(.0213)	(1.88)
Days w/Family Before Random	.0146				
	(.00273)***				
Treatment x Post		6.46	8.92	.0765	-4.58
		(2.05)***	(3.05)***	(.0236)***	(2.2)**
Post Randomization		-4.54	-5.82	-.0661	-7.22
		(1.22)***	(1.76)***	(.0142)***	(1.32)***
Female x Treat x Post			-4.28		
			(4.12)		
Female x Post			2.28		
			(2.44)		
Female x Treat			.15		
			(2.96)		
N	1,871	33,678	33,678	33,678	33,678
N of children	1,871	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188	1,188
Control Group Mean	21.667	21.667	21.667	0.259	54.351

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. The cross-sectional model is described in Equation 2, and heteroskedastic-robust standard errors are reported in parentheses. Difference-in-differences (Diff-Diff) models are described in Equation 3, and standard errors clustered at the child level are reported in parentheses. All models include strata indicators. Control Group Mean indicates the mean in the post-randomization period.

Table 4: Child Safety Measures

Dependent Variable:	(1) Protection Cases per Quarter Cross Section	(2) Protection Cases per Quarter Diff-Diff	(3) Victimization Cases per Quarter Cross Section	(4) Victimization Cases per Quarter Diff-Diff	(5) Reentry Cases per Quarter Cross Section	(6) Hospitalization Cases per Quarter Cross Section	(7) Child Safety Index Cross Section	(8) Child Safety Index Diff-Diff	(9) Child Safety Index Diff-Diff Including Sex Interaction
Treatment Group	-.0099	-.0152	-.00157	-.000593	.0027	.000032	.0044	-.0146	-.0335
	(.0115)	(.0102)	(.00408)	(.002)	(.00231)	(.0049)	(.034)	(.0109)	(.021)
Cases Before Random.	.00457		.0081						
	(.00198)**		(.00257)***						
Treatment x Post		.0196		-.00252				.00353	.0483
		(.0158)		(.0034)				(.0178)	(.0277)*
Post Randomization		-.0836		.00958				-.0186	-.0553
		(.0125)***		(.00238)***				(.014)	(.0216)**
Female x Treat x Post									-.0788
									(.0361)**
Female x Post									.065
									(.0282)**
Female x Treat									.0333
									(.0263)
N	1,871	50,517	1,871	54,259	1,871	1,871	1,871	54,259	54,259
N of children	1,871	1,871	1,871	1,871	1,871	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188	1,188	1,188	1,188	1,188	1,188
Control Group Mean	0.102	0.102	0.026	0.026	0.009	0.022	0.016	0.013	0.013

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. The dependent variables are the number of protection cases, the number of child victim reports, the number of foster care re-entries, and an indicator for having a hospitalization after randomization in 2019. Safety Index is a weighted average of the other measures in the row after being standardized as described in the text. ITT estimates are reported from the cross-sectional model described in Equation 2, and heteroskedastic-robust standard errors are reported in parentheses. Similarly, ITT estimates are reported from the difference-in-differences (Diff-Diff) models described in Equation 3, and standard errors clustered at the child level are reported in parentheses. All models include strata indicators. Control Group Mean indicates the mean in the post-randomization period.

Table 5: Crime Reports

Dependent Variable:	(1) Crime Reports/Qtr. Cross Section	(2) Crime Reports/Qtr. Diff-Diff	(3) Property Crimes Reports/Qtr.	(4) Violent Crimes Reports/Qtr.	(5) Other Crimes Reports/Qtr.	(6) Crime Reports/Qtr. Diff-Diff Including Sex Interaction
Treatment Group	-.0381 (.0197)*	.0104 (.0125)	.00756 (.00902)	.00275 (.00344)	.000258 (.00201)	-.000503 (.0148)
Crimes Before Random	.0635 (.00254)***					
Treatment x Post		-.0375 (.0134)***	-.0124 (.00607)**	-.0143 (.00582)**	-.0103 (.00523)*	-.0675 (.0283)**
Post Randomization		.0932 (.0102)***	.0226 (.00499)***	.0352 (.00409)***	.0326 (.00381)***	.155 (.0211)***
Female x Treat x Post						.0554 (.0299)*
Female x Post						-.109 (.0223)***
Female x Treat						.0186 (.0149)
Female						-.0418 (.00839)***
N	1,871	54,259	54,259	54,259	54,259	54,259
N of children	1,871	1,871	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188	1,188	1,188
Control Group Mean	0.125	0.125	0.037	0.048	0.037	0.125

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. The cross-sectional model is described in Equation 2, and heteroskedastic-robust standard errors are reported in parentheses. Difference-in-differences (Diff-Diff) models are described in Equation 3, and standard errors clustered at the child level are reported in parentheses. All models include strata indicators. Control Group Mean indicates the mean in the post-randomization period.

Table 6: School Attendance and Grades

Dependent Variable:	(1) School Attendance Cross Section	(2) School Attendance Diff-Diff	(3) School Attendance Diff-Diff Including Sex Interaction	(4) Grade Percentile Diff-Diff Unbalanced Panel	(5) Grade Percentile Diff-Diff Including Sex Interaction
Treatment Group	.0255	-.00588	-.0168	2.67	2.72
	(.0173)	(.0173)	(.023)	(1.54)*	(2.04)
Attendance Before Random	.0337				
	(.0011)***				
Treatment x Post		.0293	.0459	-.71	.654
		(.0131)**	(.0194)**	(1.41)	(2.43)
Post Randomization		-.0849	-.083	-.122	.0829
		(.00831)***	(.0122)***	(.849)	(1.32)
Female x Treat x Post			-.0282		-2.17
			(.0262)		(2.98)
Female x Post			-.00339		-.336
			(.0166)		(1.72)
Female x Treat			.019		-.129
			(.0257)		(2.38)
N	1,871	56,130	56,130	3,649	3,649
N of children	1,871	1,871	1,871	1,489	1,489
N Control Group	1,188	1,188	1,188	925	925
Control Group Mean	0.580	0.580	0.580	26.235	26.235

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. The cross-sectional model is described in Equation 2, and heteroskedastic-robust standard errors are reported in parentheses. Difference-in-differences (Diff-Diff) models are described in Equation 3, and standard errors clustered at the child level are reported in parentheses. Attendance is measured each month school is in session as the share of school days attended, and the Diff-Diff models use monthly data rather than quarterly data. Grade percentile Diff-Diff models use three academic years of data, 2017, 2018, and 2019, in an unbalanced panel. Control Group Mean indicates the mean in the post-randomization period.

Table 7: Treatment effects Across Subgroups

Heterogeneity Variable	Days Living with Family/Qtr			Crimes /Qtr			Attendance		
	ITT	SE	p	ITT	SE	p	ITT	SE	p
Male	8.92	3.05	0.00	-0.07	0.03	0.02	0.05	0.02	0.02
Female	4.64	2.77	0.09	-0.01	0.01	0.21	0.02	0.02	0.31
Difference	-4.28	4.12	0.30	0.06	0.03	0.06	-0.03	0.03	0.28
Age Older than 12	6.34	2.40	0.01	-0.04	0.02	0.02	0.03	0.02	0.06
Age Younger than 12	8.83	3.81	0.02	-0.00	0.00	0.12	0.01	0.02	0.55
Difference	2.49	4.50	0.58	0.04	0.02	0.03	-0.02	0.03	0.50
Metropolitan	2.89	3.71	0.44	-0.01	0.03	0.84	0.03	0.03	0.22
Region 5 Valparaíso	9.48	5.78	0.10	-0.05	0.02	0.03	-0.02	0.04	0.67
Difference vs. Metropolitan	6.59	6.87	0.34	-0.04	0.03	0.22	-0.05	0.05	0.30
Region 7 Maule	1.97	8.26	0.81	-0.10	0.08	0.23	0.01	0.04	0.84
Difference vs. Metropolitan	-0.92	9.06	0.92	-0.09	0.08	0.28	-0.02	0.05	0.65
Region 8 BioBío	18.15	8.26	0.03	-0.09	0.05	0.08	0.08	0.06	0.20
Difference vs. Metropolitan	15.26	9.05	0.09	-0.08	0.05	0.14	0.04	0.06	0.49
Low Predicted Crime	3.98	2.81	0.16	0.00	0.01	1.00	0.04	0.02	0.02
High Predicted Crime	7.89	2.98	0.01	-0.06	0.03	0.01	0.02	0.02	0.40
Difference	3.91	4.10	0.34	-0.06	0.03	0.02	-0.02	0.03	0.40
Low Predicted Permanency	9.00	3.19	0.00	-0.04	0.02	0.06	0.04	0.02	0.10
High Predicted Permanency	7.82	2.96	0.01	-0.03	0.02	0.05	0.02	0.02	0.41
Difference	-1.19	4.35	0.79	0.01	0.03	0.80	-0.02	0.03	0.48
Smaller Residences	3.30	2.81	0.24	-0.02	0.01	0.10	0.02	0.02	0.18
Larger Residences	10.83	2.99	0.00	-0.05	0.03	0.08	0.03	0.02	0.11
Difference	7.53	4.11	0.07	-0.03	0.03	0.28	0.01	0.03	0.62

Note: This table presents linear regression results using the longitudinal data. All models include strata indicators. Standard errors are clustered at the child level and used to calculate the p-values. The differences are reported using a fully saturated triple interaction model as the coefficient on treatment*post*subgroup-indicator. Predicted crime and predicted permanency use baseline characteristics as described in the text. Permanency is an indicator that the child is living with family (biological or adoptive) within one year of the randomization. Smaller and larger residences are defined based on the median of the number of children living in the facility in 2019.

Table 8: Cost Benefit Analysis

	Mean T	Mean C	Dif	P-Value	Costs	Dif*Costs
A. Legal-aid Costs						
Days of Legal Aid in MA Program	296.51	205.95	90.57	0.00	4.99	451.87
Days of Legal Aid outside MA	76.76	175.74	-98.98	0.00	2.73	-270.09
B. Residence Costs						
Days in residence (public)	111.07	115.85	-4.78	0.48	67.27	-321.47
Days in residence (nonprofit)	281.17	307.54	-26.37	0.07	28.35	-747.58
C. Family Foster Care Costs						
Days in care (nonprofit)	11.88	6.88	5.00	0.26	13.94	69.69
Net SENAME Costs						-817.58

Note: Estimates are on a per-child basis. The means report days in the program, residence, or family foster care over our entire observation period after randomization, 641 days. Costs are calculated in 2022 US dollars. MA costs include a 90-days period of supervision after a child exits from SENAME care.

ONLINE APPENDIX

A Size Distribution of Residences **5**

B Mi Abogado Detailed Description **6**
B.1 Budget Components 6
B.2 Role of the Lawyer 7
B.3 Role of the Social worker 9
B.4 Role of the Psychologist 9
B.5 Distribution of Mi Abogado Services across Team Members 10

C Randomization by Region **11**

D Data Sources **12**

E Mi Abogado Processes: Event Studies **13**

F Hospitalization Diagnosis Results **16**

G Heterogeneity **17**
G.1 Subgroup Heterogeneity for Exposure and Participation 17
G.2 Balance by Region 18
G.3 Across Subgroups, Adjusting for Multiple Hypothesis Testing 19
G.4 Construction of Predicted-Outcome Indexes 20

H Robustness Checks **22**
H.1 Invariance to child or time-period fixed effects 22
H.2 Additional robustness checks 23

I Exploring Mechanisms **25**
I.1 Heterogeneity by Predicted Usage of Services from Different Mi Abogado Team Members 25
I.2 Mediation Analysis 30
I.3 Crime Reports Around Entry and Exit from Foster Care 31

J Dynamics: Complier Characteristics Over Time **33**

K Cost-Benefit Analysis with Crime Outcomes **35**

Appendix Figures

A.1 Number of Children in Each Residence 5
B.1 Number of Cases Assigned in the Last 12 Months at Endline, Mi Abogado vs Non-Mi Abogado 8
E.1 Mi Abogado Processes by Stakeholders 13
E.2 Mi Abogado Processes by Team Members 14
E.3 Impacts on Quarterly Writs Submitted 15
I.1 Event Study on Entering Residences 31
I.2 Event Study on Exiting Residences 32

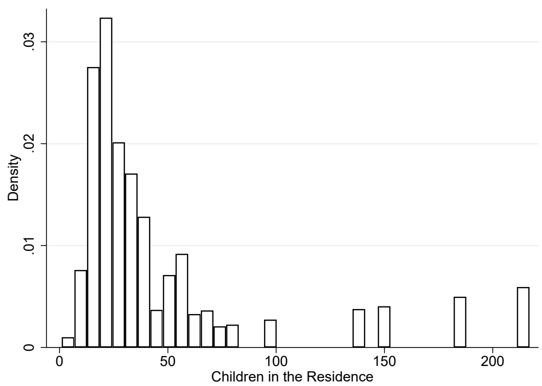
Appendix Tables

B.1 Processes share, by team member 10
C.1 Randomization by Region 11
D.1 Data Description 12
F.1 Hospitalization cases in 2019 Cross-Section Estimates 16
G.1 Participation Impacts by subgroups 17
G.2 Balance by Region 18
G.3 Treatment effects Across Subgroups 19
G.4 Models to Calculate Predicted Outcomes 21
H.1 Robustness to individual and time fixed effects 22
H.2 Robustness Checks (Diff-Diff Estimates) 23
H.3 Robustness Checks (Cross-Section Estimates) 24
I.1 Models to Predict Share of Processes by Mi Abogado Team Member . . 26
I.2 Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, Living with Family 27
I.3 Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, Crime Reports 28
I.4 Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, School Attendance 29
I.5 Days in Residence as a Mediator 30
J.1 Complier Characteristics Q1 2020 33
J.2 Complier Characteristics Q1 2021 34
K.1 Cost Benefit Analysis with Crime Outcomes 35

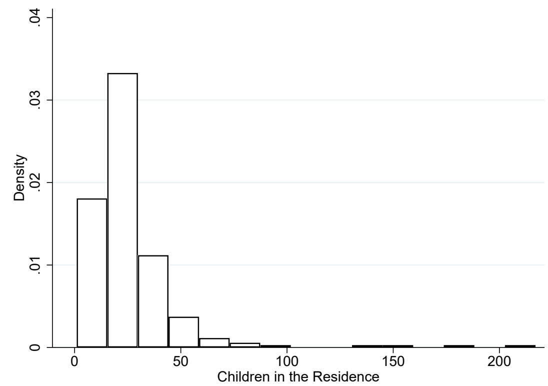
A Size Distribution of Residences

These figures display the distribution of the number of children in each residence at the time of randomization. Panel (a) is weighted by the number of children in each residence; Panel (b) is unweighted.

Figure A.1: Number of Children in Each Residence



(a) By Child.



(b) By Residence.

Note: This figure shows the distribution of the number of children in residence. Panel (a) is built using children as the unit of observation. Panel (b) is built using the residence as the unit of observation.

B Mi Abogado Detailed Description

This section describes the Mi Abogado program based on qualitative interviews with program administrators and staff along with a qualitative assessment of the program (FOCUS, 2020). The description highlights the interdisciplinary team members and their actions, the program's public and open recruitment process, its flexible criteria for lawyer selection, and its emphasis on psychological evaluation.

Unlike the status quo, where lawyers are constrained by a high caseload without the psychosocial support of the additional team members, the Mi Abogado program is designed to perform tasks that accelerate family reunification. These tasks include:

- **Diagnosis of Children's Situations**: Each child is diagnosed by the psychosocial-judiciary team, focusing on the urgency and prioritization of legal decisions. This includes interviewing or observing the child within the first month.
- **Legal Strategy Development**: Incorporating psychosocial aspects, the team develops a strategy for legal representation, ensuring:
  - Establishment of case-specific legal objectives.
  - Incorporation of feedback from relevant actors.
  - Comprehensive documentation of the strategy in the child's file.
- **Family Engagement**: Actions with the child's family or significant adults are guided by the legal strategy, including communication about legal plans and collaboration in monitoring and decision-making processes.
- **Intersectoral Coordination**: Teams ensure that responsible parties for the child's care access other relevant public services. This duty is supervised by the program's Regional Coordination leadership.
- **Procedural Processes**: Execution of the legal strategy in family and other courts, with all actions recorded in the child's file.
- **Post-Care Follow-Up**: The Technical Unit ensures regional teams supervise the implementation of court decisions, with follow-up extending for a minimum of three months after discharge from substitute care.
- **Program Exit Criteria**: Assessment of whether legal strategy objectives have been met and cases processed is conducted, with reasons for discharge such as family reunification, adoption or reaching adulthood.

B.1 Budget Components

According to regional reports of the program, the largest component of costs is allocated to the lawyers at 47%, growing to 53% if we include the regional coordinator who is also a lawyer; next is office space (14%); social workers and psychologists (12% for each category); other staff (6%); and travel (3%).

B.2 Role of the Lawyer

Lawyers, selected through a competitive process, are responsible for processing cases, developing legal strategies in collaboration with the psychosocial team, and ensuring judicial efficiency. Their caseloads are managed to allow for focused attention on each child, reflecting the program's adaptation to the realities of expansion and regional diversity.

B.2.1 Functions

- Legal strategy development and management of legal actions.
- Attendance at hearings and conducting interviews.
- Striving for decisions favorable to the child's interests.
- Informing the child and relevant adults about case progress.
- Participation in case analysis meetings and supporting complementary projects.
- Maintaining detailed records of all procedures.

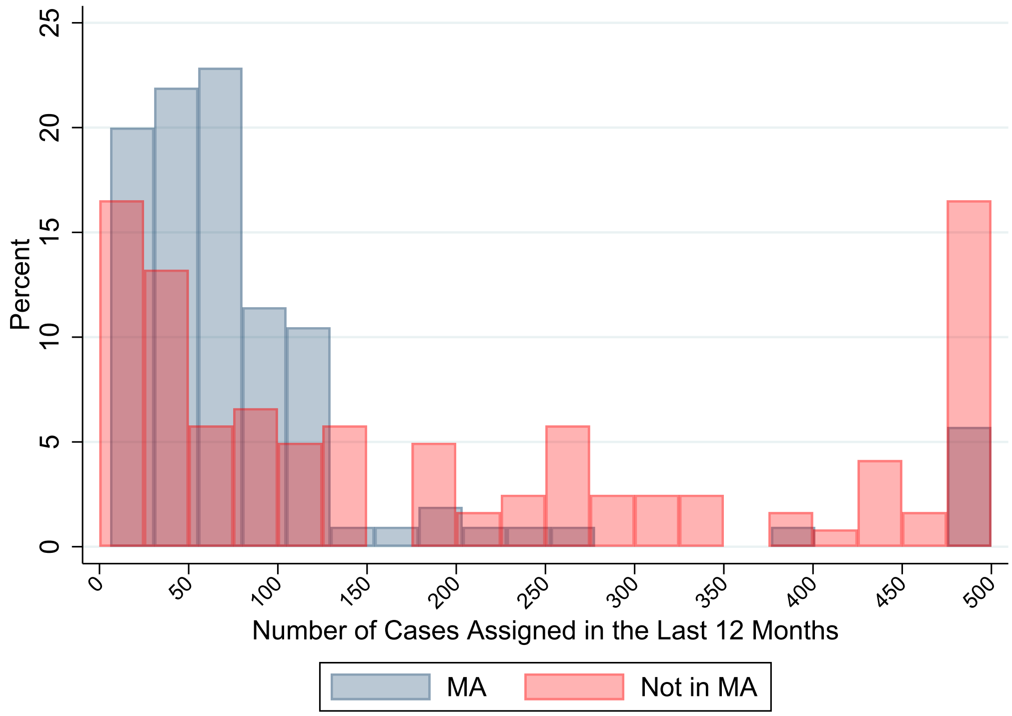
B.2.2 Training and Experience

The program looks for qualified lawyers to lead the child's team. Ideal candidates will have a specialization in human rights, child and adolescent rights, criminal law, criminal procedural law, family law, or similar. With experience in litigation before family courts, in ordinary and extraordinary procedures; before criminal courts; and before the superior courts of justice, with knowledge in prevention, promotion, protection, and restitution of rights, threat, and violation of rights and crimes committed against children. With experience in work, coordination, and articulation in the inter-institutional and intersectoral network. With skills for conflict resolution and interventions in crises. Experience in interviews with children in situations of high complexity is desirable.

B.2.3 Case Assignment Comparison

One goal of Mi Abogado is to limit the caseload for lawyers in its program. Our data include high-quality measures of case assignment, but the end of legal representation is not well recorded. This complicates measurement of caseload for any lawyer at a point in time. This appendix shows the distribution of the number of cases assigned in the last 12 months of our data for the Mi Abogado lawyers and non-Mi Abogado lawyers. The mean number of case assignments over this period among Mi Abogado lawyers is 130; for non-Mi Abogado lawyers the mean is 309.

Figure B.1: Number of Cases Assigned in the Last 12 Months at Endline, Mi Abogado vs Non-Mi Abogado



Note: This figure shows the annual case-assignment distribution for lawyers in the Mi Abogado (MA) program and lawyers not in the MA program. This proxy for caseload is built with the number of new cases that lawyers were assigned in last twelve months of our sample period. This variable is truncated at 500 cases. The mean number of cases assigned to MA lawyers over the course of the year is 130 compared to 309 for non-MA lawyers.

B.3 Role of the Social worker

The program employed professional social workers hired full time with the goal is to cap their caseload at 200 children.

B.3.1 Functions

- Responsible for delivering social support to the program team in problems associated with serious violations of rights.
- Conduct home visits, interviews, and works with the child's networks, as strictly required by the legal strategy, and in permanent coordination with professionals of complementary projects to the program, when appropriate.
- Conduct interviews or observations with the children, family, or others involved.
- Contribute to the elaboration of the diagnosis of the judicial situation and the development and execution of the legal strategy of each child. Record all the actions performed.
- Note that the social worker aids the lawyer in constructing a plan to help the child and the family reunite, as opposed to direct delivery of services.

B.3.2 Training and Experience

A qualified social worker with specialized training in family and childhood matters, desirable training in criminal law or child abuse, experience working with children in violation of rights, and health and education networks. Desirable experience in interviews with children in situations of high complexity.

B.4 Role of the Psychologist

The program employed professional psychologists hired full time with the goal is to cap their caseload at 200 children.

B.4.1 Functions

- Assess the child's mental health is entering the program by pre-existing reports.
- Assistance in emergencies or crises of the child in the context of the hearing, when appropriate.
- Contribute to elaborating the diagnosis of the judicial situation and legal strategy of each child.
- Permanent coordination with the network involved. Conduct interviews or observations with the children, family, or others involved that correspond and must move if necessary.
- Record all the actions performed and incorporated required verifiers.

- Other functions specific to the work methodology and legal strategy adopted by the program.
- Note that the social worker aids the lawyer in constructing a plan to help the child and the family reunite, as opposed to direct delivery of therapeutic services.

B.4.2 Training and Experience

Qualified psychologist with specialized training in family and childhood matters, desirable training in the field of criminal law to child abuse, and experience in working with children in situations of violation of rights.

B.5 Distribution of Mi Abogado Services across Team Members

Table B.1: Processes share, by team member

Type of Processes	Share
Interdisciplinary	0.574
Judicial (only lawyer participates)	0.293
Only psychologist participates	0.066
Only social worker participates	0.067
Psychologist or Social Worker	0.134

Note: This table reports the share of processes among Mi Abogado participants. The Interdisciplinary category is used whenever two or more of the members of the team participate in the process.

C Randomization by Region

The Mi Abogado program was rolled out as a stratified randomized trial. The strata were sex, age group (above and below 12 years old), and each of the four most populous regions in Chile. This was a pragmatic trial that aimed to spread the limited number of Mi Abogado lawyers across the cases via randomization. Each region had a different number of eligible children and a different number of available lawyers. As a result, the share of the eligible children assigned to the treatment group varied widely across regions ranging from 7% to 92% as shown in the Table C.1. The paper describes the concerns that arise when the share treated varies across regions and offers a set of robustness checks in Appendix Section H.2.

Table C.1: Randomization by Region

	N Total	N Treatment	Share Treatment
Valparaíso	419	42	0.10
Maule	451	413	0.92
Biobío	378	28	0.07
Metropolitan	623	200	0.32

Note: This table shows the number and share of children randomized to the treatment group across the four regions.

D Data Sources

A feature of the pragmatic randomized trial in Chile is the ability to link subjects to a wide range of outcomes throughout the country's registry data. We were able to secure linkages from the Judiciary Registry, including crime report outcomes and victimization as a measure of child safety; SENAME (child protection) data to measure days in care and identify when children leave care to live in a permanent family; Mi Abogado program data, which allows us to measure program participation and describe the program's processes in detail; and Ministry of Education data, which provides measures of attendance and school performance in the form of overall grade percentiles.

In addition, SENAME collects hospitalization data from the Ministry of Health of Chile (MINSAL). Each year, they collect data on all hospitalizations of children in their care and construct a dataset that includes the last hospitalization for each child. We obtained data for hospitalizations after March 31, 2019, the date of the randomization. We also obtained 2019 data that includes the type of diagnosis, but only for the subset of children who were in care in 2020.

Below is a table that describes the years of data we use in the analysis, along with the number of children in the study represented in each data source.

Table D.1: Data Description

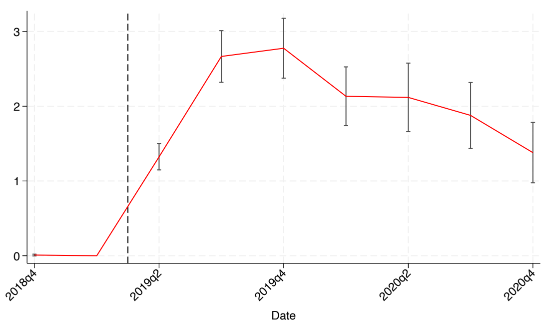
Source	Variable	Period use	Obs
Judiciary Registry	Crime reports	April 2014 - June 2021	1871
	Protection Cases	April 2014 - December 2020	1871
	Missing	April 2014 - June 2021	1871
	Victimization	April 2014 - June 2021	1871
	Allegations	April 2014 - June 2021	1871
	Writs	January 2010 - December 2020	1871
	Hearings	January 2010 - December 2020	1871
	Lawyer Assignment	January 2018 - February 2021	1871
SENAME (SENAINFO)	Days living with family	January 2017 - June 2021	1871
	Days living in residence	January 2017 - June 2021	1871
	Days living in family foster care	January 2017 - June 2021	1871
	Age at entry in residence	January 2017 - June 2021	1871
	Allegations	January 2017 - June 2021	1871
	Dispositions (exit reasons)	January 2017 - June 2021	1871
	Length of stay in residence	January 2017 - June 2021	1871
	Delay in School	January 2017 - June 2021	1871
	Hospitalization Dates (via Ministry of Health)	April 2019 - December 2019	1871
	Hospitalization Diagnoses (via Ministry of Health)	April 2019 - December 2019	1345
Mi Abogado	Participation in Mi Abogado program	October 2018 - December 2020	1871
	Days in Mi Abogado program	October 2018 - December 2020	1871
	Mi Abogado processes	October 2018 - December 2020	1871
Ministry of Education	Grades	March 2017 - December 2019 (annual measure)	1616
	School Attendance	March 2017 - December 2019	1871

Note: This table shows the sources of information used to construct each variable, the period available for each set of information, and the number of observations (children) that each source includes. Hospitalization dates are for the last hospitalization in the calendar year; diagnoses are available for children who remained in SENAME care into 2020.

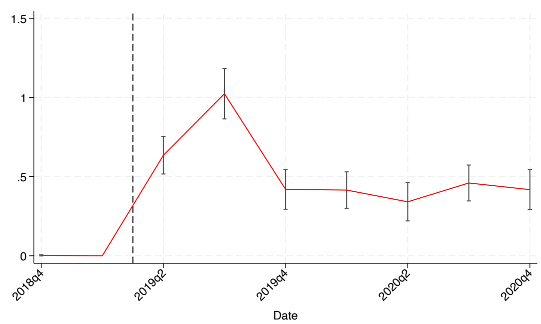
E Mi Abogado Processes: Event Studies

The Mi Abogado program data provide detailed measures of each process delivered. We do not observe such detailed information for those who are not part of the program. The figures below show event studies that represent the difference between treatment and control in the average number of each process type per child over time. In the two quarters after randomization, the largest relative increase in services is documentary work, followed by interactions with residence staff, interactions with family, and then interactions with the child. Similarly, we show an event study regarding the average number of processes by different team members, and for the number of writs issued.

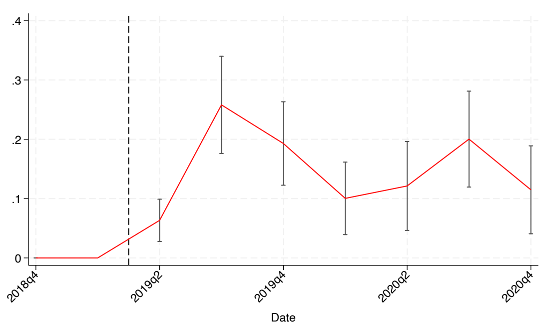
Figure E.1: Mi Abogado Processes by Stakeholders



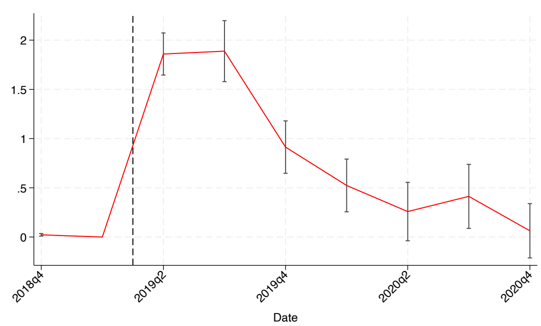
(a) Documentary Work



(b) Interacting with Child



(c) Interacting with Family

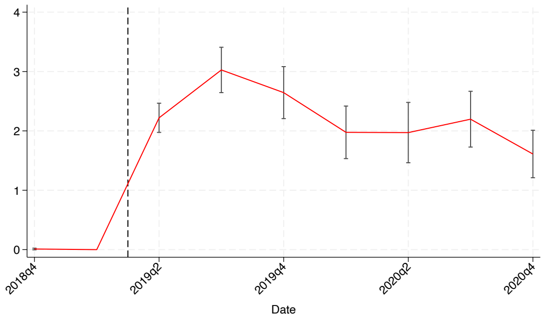


(d) Interacting with Residence Staff

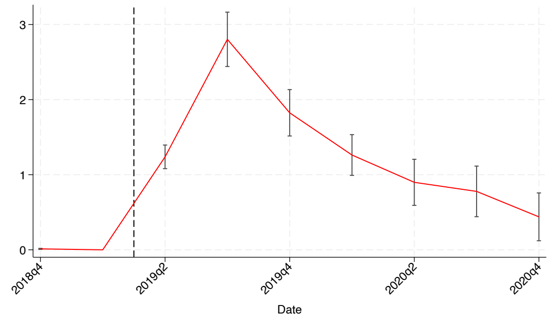
Note:

Note: These figures report event-study estimates described in Equation 1 of differences between the treatment and control groups for measures of the number of Mi Abogado process by type of process. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

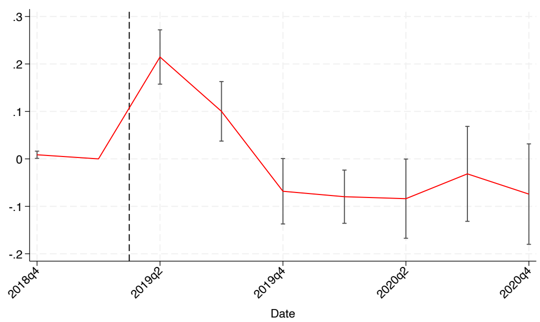
Figure E.2: Mi Abogado Processes by Team Members



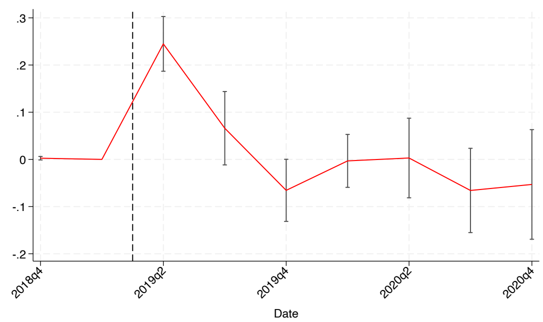
(a) Interdisciplinary Participation



(b) Only Lawyer Participation



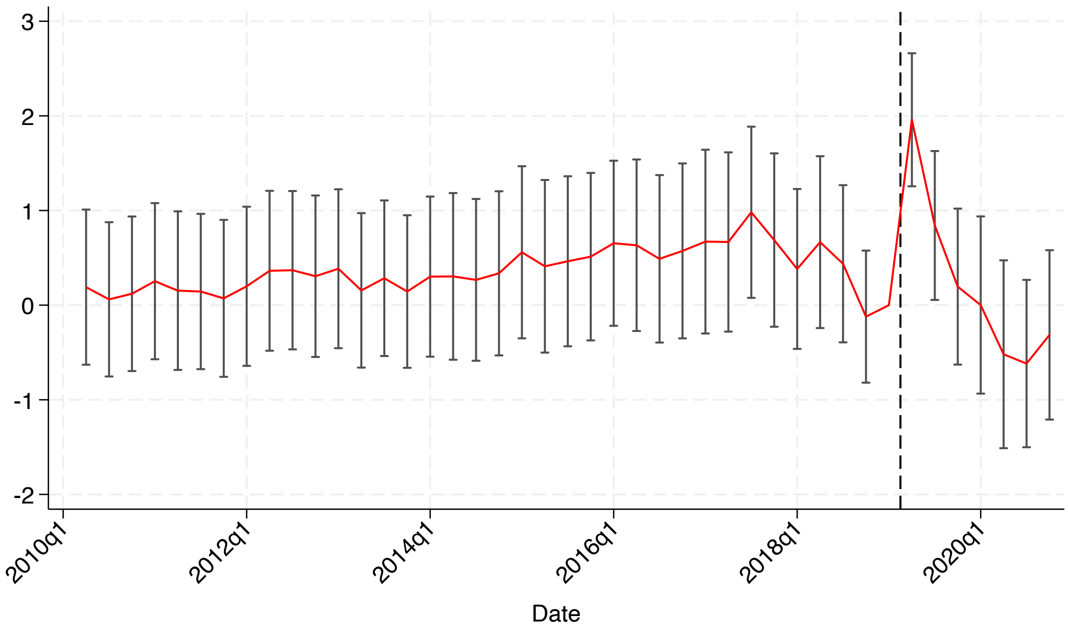
(c) Only Psychologist Participation



(d) Only Social Worker Participation

Note: Note: These figures report event-study estimates described in Equation 1 of differences between the treatment and control groups for measures of the number of Mi Abogado process by team member. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

Figure E.3: Impacts on Quarterly Writs Submitted



Note: Note: These figures report event-study estimates described in Equation 1 of differences between the treatment and control groups for measures of the number of writs submitted to the court. Confidence intervals are calculated using standard errors clustered at the child level. The vertical line shows the time of randomization.

F Hospitalization Diagnosis Results

Recall that we obtained data on the last hospitalization for each child during 2019, as well as data on diagnoses for those who remain in care in 2020. We observe that the treatment group is 4.5 percentage points (s.e.=3.0) less like to be in care in 2020. As a result, we focus on hospitalization in 2019 in the main results.

Table F.1 reports results by type of diagnosis. We do not observe statistically significant differences in hospitalization across these different types of admissions.

Table F.1: Hospitalization cases in 2019 Cross-Section Estimates

Dependent Variable:	(1) Child Hospit.	(2) Infection Diseases	(3) Chronic Diseases	(4) Respiratory Diseases	(5) Child Injuries	(6) Mental Diseases	(7) Addiction Diseases	(8) Other Diseases
Treatment Group	.000032	-.0015	-.00073	.00017	.0027	-.000077	-.00043	-.000039
	(.0049)	(.002)	(.00057)	(.0011)	(.0023)	(.0027)	(.0015)	(.0021)
N	1,871	1,871	1,871	1,871	1,871	1,871	1,871	1,871
N of children	1,871	1,871	1,871	1,871	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188	1,188	1,188	1,188	1,188
Control Group Mean	0.022	0.003	0.001	0.001	0.002	0.009	0.003	0.003

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table presents linear regression results. Heteroskedastic-robust standard errors are reported in parentheses. Each column show estimates for a model of a dummy indicating if there is a case within the period of interest on treatment variable. Data includes the last hospitalization case that occurred in 2019 after randomization, and whose diagnosis was available in dataset "Atenciones2020", i.e. it considers the period 04/01/2019 to 12/31/2019. All models include strata indicators.

G Heterogeneity

G.1 Subgroup Heterogeneity for Exposure and Participation

Table 7 reports heterogeneous treatment effects across subgroups of interest for the main outcomes; this table demonstrates that exposure to the program is higher for the subgroups as well.

Table G.1: Participation Impacts by subgroups

Heterogeneity Variable	Exposure				Participation			
	Mean	ITT	SE	p	Mean	ITT	SE	p
Male	36.65	17.26	2.73	0.00	31.07	10.24	2.63	0.00
Female	30.04	22.66	2.35	0.00	25.25	15.16	2.23	0.00
Difference	-6.60	5.40	3.60	0.13	-5.82	4.92	3.45	0.15
Age Older than 12	32.54	23.93	2.18	0.00	27.93	14.03	2.01	0.00
Age Younger than 12	48.22	10.93	3.08	0.00	43.23	8.43	3.07	0.01
Difference	15.68	-13.00	3.77	0.00	15.30	-5.59	3.67	0.13
Metropolitan	35.45	19.41	3.12	0.00	29.72	14.02	3.04	0.00
Region 5 Valparaíso	27.57	13.02	6.47	0.04	25.71	13.83	6.44	0.03
Difference	-7.88	-6.39	7.18	0.37	-4.01	-0.19	7.12	0.98
Region 7 Maule	36.57	14.13	6.49	0.03	28.61	9.99	6.00	0.10
Difference	1.12	-5.28	7.20	0.46	-1.11	-4.03	6.73	0.55
Region 8 BioBío	31.34	11.46	7.92	0.15	28.37	0.06	8.78	0.99
Difference	-4.11	-7.95	8.51	0.35	-1.35	-13.96	9.29	0.13
Low Predicted Crime	39.40	19.77	2.45	0.00	34.41	15.26	2.40	0.00
High Predicted Crime	33.69	20.13	2.60	0.00	28.65	9.63	2.37	0.00
Difference	-5.70	0.36	3.57	0.92	-5.76	-5.63	3.38	0.10
Low Predicted Permanency	38.92	23.01	2.86	0.00	33.38	18.34	2.99	0.00
High Predicted Permanency	24.30	26.48	2.38	0.00	18.50	18.34	2.16	0.00
Difference	-14.62	3.46	3.72	0.35	-14.87	0.00	3.69	1.00
Smaller Residences	38.79	17.48	2.41	0.00	32.86	11.18	2.30	0.00
Larger Residences	35.13	23.33	2.66	0.00	29.04	14.63	2.57	0.00
Difference	-3.66	5.85	3.59	0.10	-3.82	3.45	3.44	0.32
Pre-Covid Period	24.75	25.18	1.65	0.00	21.49	21.58	1.63	0.00
Covid Period	48.42	13.69	1.94	0.00	40.09	1.48	2.15	0.49
Difference	23.67	-11.48	0.98	0.00	18.59	-20.10	1.38	0.00

Note: Exposure is the number of days since first entering the Mi Abogado program each quarter, while participation is the number of days actually in the program each quarter.This table presents linear regression results using the longitudinal data. All models include strata indicators. Standard errors are clustered at the child level and used to calculate the p-values. The differences are reported using a fully saturated triple interaction model as the coefficient on treatment*post*subgroup-indicator. Our prediction models using baseline characteristics for Crime and Permanency are presented in Table G.4. Permanency is an indicator that the child is living with family (biological or adoptive) within one year of the randomization.

G.2 Balance by Region

Table 7 reports heterogeneous treatment effects across regions, and this table demonstrates that covariates appear to be well balanced in each of the regions, as designed.

Table G.2: Balance by Region

	F	p-value
All regions	0.87	0.60
Region=5 (Valparaíso)	0.86	0.62
Region=7 (Maule)	1.18	0.28
Region=8 (Bío-Bío)	1.31	0.19
Region=13 (Metropolitan)	0.99	0.47

Note: This table shows, for all regions together, and for each of them separately, the the F-statistic and associated p-value for the test of joint significance of the baseline characteristics predicting treatment, controlling for randomization-strata indicators. The table is analogous to Table 1.

G.3 Across Subgroups, Adjusting for Multiple Hypothesis Testing

We systematically explore for heterogeneous effects for each of the baseline characteristics in Table 1. We report Romano-Wolf adjusted p-values to account for multiple hypothesis testing and control the family-wise error rate (FWER). While simple p-values from standard difference-in-difference estimates indicate the significance of individual tests, they do not account for the increased risk of Type I errors when conducting multiple comparisons. By applying the Romano-Wolf adjustment, we reduce the probability of false positives across the set of hypotheses, providing more robust and reliable statistical inferences in the presence of multiple treatments or interactions.

Table G.3: Treatment effects Across Subgroups

Heterogeneity Variable	Days Living with Family/Qtr				Crimes /Qtr				Attendance			
	ITT	p	GRWp	ARWp	ITT	p	GRWp	ARWp	ITT	p	GRWp	ARWp
Female	-4.28	0.30	0.30	0.96	0.06	0.06	0.06	0.42	-0.03	0.28	0.28	0.99
Age Younger than 12	2.49	0.58	0.58	1.00	0.04	0.03	0.03	0.73	-0.02	0.50	0.50	1.00
Region 5 Valparaíso	4.37	0.48	0.79	1.00	-0.00	0.88	0.88	0.99	-0.05	0.23	0.62	0.99
Region 7 Maule	-4.29	0.62	0.79	1.00	-0.06	0.45	0.56	0.89	-0.02	0.63	0.62	1.00
Region 8 BioBío	11.80	0.17	0.43	0.88	-0.06	0.21	0.56	0.89	0.05	0.39	0.62	0.99
Number of Siblings	-2.21	0.60	0.76	1.00	0.01	0.64	0.88	0.97	0.01	0.80	0.98	1.00
Delay in Schooling	-6.07	0.16	0.52	0.88	0.00	0.90	0.90	0.99	-0.02	0.42	0.92	1.00
Time in Residence	-4.55	0.22	0.52	0.90	0.05	0.04	0.16	0.42	0.01	0.59	0.98	1.00
Age When First in Residence	2.59	0.53	0.76	1.00	-0.07	0.01	0.04	0.18	-0.01	0.79	0.98	1.00
Larger Residences	7.53	0.07	0.30	0.60	-0.03	0.28	0.57	0.89	0.01	0.62	0.98	1.00
Allegation Sex Abuse	-1.23	0.81	0.97	1.00	0.02	0.35	0.73	0.96	0.02	0.62	0.84	1.00
Allegation Physical Abuse	8.83	0.04	0.15	0.50	-0.02	0.54	0.73	0.96	0.01	0.65	0.84	1.00
Allegation Neglect	0.51	0.94	0.97	1.00	0.03	0.27	0.73	0.93	-0.03	0.41	0.78	1.00
High Predicted Crime	3.91	0.34	0.56	0.96	-0.06	0.02	0.03	0.26	-0.02	0.40	0.64	1.00
High Predicted Permanency	-1.19	0.79	0.79	1.00	0.01	0.80	0.80	0.99	-0.02	0.48	0.64	1.00

Note: This table presents linear regression results for days living with family, crimes, and attendance using difference-in-difference models. All models include strata indicators. Standard errors are clustered at the child level and are used to calculate the p-values. Each row presents results for a model using a fully saturated triple interaction, represented as treatment*post*subgroup-indicator, where the indicators are dummy variables related to the specific category. For "Number of Siblings," "Delay in School," "Time in Residence," and "Age When First in Residence," we construct a dummy variable that indicates whether the value for the child is above the median. The third and fourth columns of each outcome variable present the Romano-Wolf adjusted p-values described above. Two are reported: GRWp presents the adjustment for grouped variables (GRWp), where the groups are defined by closely related measures and indicated by the horizontal lines in the table; the second one presented is for all variables together (ARWp).

G.4 Construction of Predicted-Outcome Indexes

Table 7 reports heterogeneous treatment effects across children that vary in their predicted permanency and crime-report outcomes. This table reports the models used in calculating those predictions based on the child's baseline characteristics. Specifically, we regressed each outcome on the demographic and allegation characteristics. For the treatment group, we estimated the relationship between these characteristics and the outcomes using the control group. Within the control group, the predicted outcome is calculated using a leave-out regression to avoid a child's outcome from informing his or her prediction Abadie et al. (2018).

Table G.4: Models to Calculate Predicted Outcomes

Dependent Variable:	(1) Living with Family One Year After Randomization	(2) Number of Crimes First Year After Randomization
Female	0.075	-0.528
	(0.025)***	(0.073)***
Region=5	-0.009	-0.031
	(0.029)	(0.086)
Region=7	0.141	0.080
	(0.070)**	(0.205)
Region=8	-0.045	0.257
	(0.030)	(0.089)***
Low Age Stratum	0.058	-0.117
	(0.051)	(0.149)
Number of Siblings	-0.005	-0.024
	(0.006)	(0.018)
Delay in Schooling	0.011	0.042
	(0.009)	(0.027)
Allegation: Sex Abuse	-0.067	-0.184
	(0.033)**	(0.097)*
Allegation: Physical Abuse	-0.053	0.046
	(0.028)*	(0.083)
Allegation: Neglect	-0.195	0.156
	(0.034)***	(0.100)
Time in Residence	-0.009	-0.176
	(0.017)	(0.051)***
Age When First in Residence	0.015	0.041
	(0.006)**	(0.018)**
Age at Randomization	-0.000	0.049
	(0.009)	(0.027)*
Constant	0.246	0.072
	(0.138)*	(0.403)
N	1,188	1,188
R^2	0.069	0.137

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table shows models predicting living with family and number of crime reports. The models for the two outcomes are estimated using only the control group. For Living with Family we use the living with family status exactly one year after randomization. For Crimes we use the number of crimes within the first year after randomization.

H Robustness Checks

H.1 Invariance to child or time-period fixed effects

The event study estimates compare the difference between treatment and control compared to the omitted time period (event time = -1). Similarly, the difference-in-differences estimates in the results tables compare the difference in the treatment group compared to the control group in the post-randomization period to the pre-randomization period. Given that the sample is balanced and there are no continuous regressors, these differences absorb child fixed characteristics, such as time-invariant controls.

Intuition can be gained from considering a standard DD model and de-meaning the data at the child level to absorb any fixed effects. The main explanatory variable post*treat increases from 0 to 1 for the treatment group in the post period. When it is de-meaned, it goes from -0.5 to +0.5: the variation is identical to the non-de-meaned variable.

Table H.1 demonstrates that the results do not change when including child fixed effects or time-invariant controls. With clustered standard errors, they are also similar with and without child fixed effects in our context. Similarly, calendar quarter fixed effects also yield the same results because there is no variation across treatment and control given the single event date. The event-study point estimates for the difference across the groups in any given quarter are also identical with or without time-invariant controls or child fixed effects.

Table H.1: Robustness to individual and time fixed effects

Specification	Days Living w Family		Crimes /Qtr		Attendance	
	ITT	SE	ITT	SE	ITT	SE
Main Estimate	6.456754	2.054193	-0.037504	0.013418	0.029285	0.013088
Main Estimate with Strata	6.456754	2.054346	-0.037504	0.013418	0.029285	0.013089
Adding Controls	6.456754	2.054437	-0.037504	0.013418	0.029285	0.013089
Adding Individual FE	6.456754	2.113688	-0.037504	0.013655	0.029285	0.013312
Adding Quarter FE	6.456754	2.054681	-0.037504	0.013421	0.029285	0.013092
Adding Season FE	6.456754	2.054285	-0.037504	0.013418	0.029285	0.013088

Note: This table reports results for the living-with-family, crime, and attendance estimates with different specifications. The first row presents our main estimates. Strata indicators are included in the no controls and controls specifications, and they are not included in the fixed effects specifications because they are constant over time. The additional controls include: number of siblings, delay in schooling, age at randomization, indicators for type of allegation: sexual abuse, physical abuse and/or neglect, days in residence prior to randomization, and age when first in residence. The fifth row adds quarter fixed effects. The sixth row adds season fixed effects. Seasons are defined as Summer from January to March, Fall from April to June and Winter from July to September. Standard errors are clustered at the child level.

H.2 Additional robustness checks

Combining Regional Effects. The share randomized to treatment varies across the regional strata because the randomized evaluation is part of the roll-out of the program and capacity varied across these regions. Linear regression places more weight on the areas with more variance in treatment. Table H.2 shows that results are similar if we weight the regression to undo this weighting (Gibbons et al., 2019). Results are also similar if we estimate the effects separately by region and then compute the weighted average, where the weights reflect the precision of the region-specific estimates, as described in (Athey and Imbens, 2017).

Second Randomization (Replacements) Recall from the text that a group of 51 children who were initially not drawn into the treated group on the randomization date were randomized into treatment; they were referred to as "replacements" in May 2019 as the program expanded. For all other results shown in the paper, these newly-randomized subjects are treated as part of the treatment group and contribute to the non-compliance at the beginning of the post-randomization period. Table H.2 shows that the main results are similar regardless of whether these children are included in the control group instead, or not included in the analysis.

Table H.2: Robustness Checks (Diff-Diff Estimates)

Specification	Days Living w Family		Crimes /Qtr		Attendance	
	ITT	SE	ITT	SE	ITT	SE
Main Estimate	6.457	2.054	-0.038	0.013	0.029	0.013
Inverse Weighting	6.916	2.100	-0.044	0.013	0.030	0.013
Athey-Imbens	7.225	3.157	-0.052	0.024	0.024	0.020
Replacements as Controls	6.331	2.094	-0.038	0.013	0.028	0.013
Replacements as Missings	6.550	2.108	-0.039	0.014	0.030	0.013
Exclude COVID Period	7.322	2.632	-0.044	0.023		

Note: This table reports results for the living-with-family, crime, and attendance results with different specifications. The first row presents our main diff-in-diff estimates. The second row weights our estimates by the inverse of the variance of the treatment variable, as suggested by Gibbons et al. (2019). The third row aggregates the regional treatment effects as suggested in Athey and Imbens (2017). The rows referring to replacements treat the 51 children randomized to treatment in May 2019 either as controls or dropped from the analysis. The last row drops the observations after the onset of the COVID pandemic; this is not applicable for the Attendance result as that outcome did not extend to the COVID period in our main analysis. Standard errors are clustered at the child level.

The paper describes results presented by McKenzie (2012) who suggested an ANCOVA model that is very similar to the cross-sectional model we report. To verify that the models produce the same results, we produce them in Table H.3.

Table H.3: Robustness Checks (Cross-Section Estimates)

Specification	Days Living w Family		Crimes /Qtr		Attendance	
	ITT	SE	ITT	SE	ITT	SE
Cross-Section	3.609	2.136	-0.038	0.020	0.026	0.017
ANCOVA	3.609	2.209	-0.038	0.020	0.026	0.018

Note: This table reports the results of living with family, delinquency, and attendance with different specifications. The first row presents cross-sectional estimates. The second row shows estimates from an ANCOVA model suggested by McKenzie (2012). This model controls for period dummies and uses standard errors clustered at the child level.

I Exploring Mechanisms

I.1 Heterogeneity by Predicted Usage of Services from Different Mi Abogado Team Members

The Mi Abogado program is a bundle of services, beginning with legal advice and services from the lawyer, as well as services provided by the psychologist and social worker. As noted in the text, administrators believe the interdisciplinary nature of the program is an important ingredient for the program's success. Using the program processes data, we categorized the processes according to whether they are interdisciplinary (more than one member of the team participates on them), judicial (only the lawyer participates), or psychosocial (the psychologist or the social worker are the only ones participating in the process). We construct, for each individual, the share of each of those three types of processes. Then, we predict those shares using the baseline characteristics. Finally, we explore heterogeneity by whether the child is above or below the median of the predicted shares of each type.

The tables below present the estimates that predict the different mix of services, followed by the heterogeneity results.

Table I.1: Models to Predict Share of Processes by Mi Abogado Team Member

	Total Processes	Share: Interdisciplinary	Share: Judicial	Share: Psychosocial
Female	6.830	-0.014	0.021	-0.007
	(2.747)**	(0.008)*	(0.007)***	(0.005)
Region=5	-45.758	0.008	0.079	-0.087
	(3.891)***	(0.011)	(0.010)***	(0.008)***
Region=7	20.237	0.091	0.052	-0.142
	(3.487)***	(0.010)***	(0.009)***	(0.007)***
Region=8	-10.946	0.090	-0.006	-0.084
	(3.795)***	(0.011)***	(0.010)	(0.007)***
Low Age Stratum	-25.152	0.051	-0.047	-0.004
	(5.532)***	(0.016)***	(0.015)***	(0.011)
Number of Siblings	0.094	0.003	-0.003	0.000
	(0.652)	(0.002)	(0.002)*	(0.001)
Delay in Schooling	0.274	-0.002	-0.001	0.003
	(0.982)	(0.003)	(0.003)	(0.002)
Allegation: Sex Abuse	-1.773	-0.010	0.017	-0.007
	(3.515)	(0.010)	(0.009)*	(0.007)
Allegation: Physical Abuse	5.425	-0.009	-0.002	0.011
	(2.875)*	(0.008)	(0.008)	(0.006)*
Allegation: Neglect	11.630	-0.030	0.012	0.018
	(4.661)**	(0.013)**	(0.013)	(0.009)**
Time in Residence	1.782	0.021	-0.013	-0.008
	(2.066)	(0.006)***	(0.006)**	(0.004)**
Age When First in Residence	0.453	0.003	-0.001	-0.002
	(0.721)	(0.002)	(0.002)	(0.001)
Date of Birth	0.019	-0.000	0.000	0.000
	(0.003)***	(0.000)***	(0.000)***	(0.000)
Constant	-249.435	0.871	-0.059	0.188
	(58.117)***	(0.168)***	(0.157)	(0.113)*
N	1,271	1,271	1,271	1,271
Number/Share	84.420	0.594	0.268	0.138
R^2	0.224	0.141	0.095	0.289

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table shows our predictive models for the total number and shares of Mi Abogado processes by team-member participation.

Table I.2: Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, Living with Family

	(1)	(2)	(3)	(4)
Dependent Variable:	Days Living w/Family/Qtr.	Days Living w/Family/Qtr.	Days Living w/Family/Qtr.	Days Living w/Family/Qtr.
Heterogeneity Variable:	Predicted Processes Total Processes	Predicted Processes Share: Interdisciplinary	Predicted Processes Share: Judicial	Predicted Processes Share: Psychosocial
Treatment x Post	12.154	4.990	8.503	-0.479
	(3.624)***	(3.250)	(3.149)***	(2.743)
Treatment Group	-4.412	-0.467	-1.369	-0.242
	(2.710)	(2.480)	(2.184)	(2.884)
Post Randomization	-3.136	-9.520	0.624	3.908
	(1.490)**	(1.565)***	(1.546)	(1.870)**
Above Median in Prediction	-1.714	-11.952	19.352	8.865
	(2.231)	(2.232)***	(2.581)***	(2.117)***
Above Median x Post x Treatment	-4.986	-2.439	1.072	9.570
	(4.612)	(4.272)	(4.224)	(4.285)**
Above Median x Post	-4.124	12.209	-12.307	-14.128
	(2.605)	(2.471)***	(2.480)***	(2.448)***
Above Median x Treatment	0.284	-2.877	-5.599	-3.025
	(3.590)	(3.360)	(3.243)*	(3.792)
N	33,678	33,678	33,678	33,678
Mean if Share Above Median	22.370	25.673	24.033	19.884
Mean if Share Below Median	24.488	21.184	22.826	26.978

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table reports results for living-with-family estimates based on the predicted share of processes by the team member(s). The prediction is summarized by a binary variable indicating whether the observation is above or below the median of the prediction. We use those binary indicators as controls and as heterogeneity variables. All models include strata indicators.

Table I.3: Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, Crime Reports

	(1)	(2)	(3)	(4)
Dependent Variable:	Crimes /Qtr.	Crimes /Qtr.	Crimes /Qtr.	Crimes /Qtr.
Heterogeneity Variable:	Predicted Processes Total Processes	Predicted Processes Share: Interdisciplinary	Predicted Processes Share: Judicial	Predicted Processes Share: Psychosocial
Treatment x Post	-0.041	-0.021	-0.029	-0.064
	(0.022)*	(0.020)	(0.024)	(0.019)***
Treatment Group	0.009	0.010	0.023	0.019
	(0.023)	(0.013)	(0.018)	(0.023)
Post Randomization	0.108	0.078	0.102	0.119
	(0.012)***	(0.014)***	(0.015)***	(0.016)***
Above Median in Prediction	-0.049	-0.019	0.002	-0.018
	(0.012)***	(0.017)	(0.012)	(0.010)*
Above Median x Post x Treatment	0.030	-0.040	-0.005	0.046
	(0.031)	(0.027)	(0.029)	(0.028)*
Above Median x Post	-0.045	0.037	-0.021	-0.043
	(0.023)**	(0.020)*	(0.020)	(0.021)**
Above Median x Treatment	-0.011	-0.002	-0.039	-0.015
	(0.024)	(0.024)	(0.020)*	(0.027)
N	54,259	54,259	54,259	54,259
Mean if Share Above Median	0.071	0.125	0.081	0.097
Mean if Share Below Median	0.151	0.097	0.141	0.125

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table reports results for crime reports based on the predicted share of processes by the team member(s). The prediction is summarized by a binary variable indicating whether the observation is above or below the median of the prediction. We use those binary indicators as controls and as heterogeneity variables. All models include strata indicators.

Table I.4: Heterogeneity by Share of Predicted Processes by Mi Abogado Team Member, School Attendance

Dependent Variable:	(1) Attendance /Qtr.	(2) Attendance /Qtr.	(3) Attendance /Qtr.	(4) Attendance /Qtr.
Heterogeneity Variable:	Predicted Processes Total Processes	Predicted Processes Share: Interdisciplinary	Predicted Processes Share: Judicial	Predicted Processes Share: Psychosocial
Treatment x Post	0.009	0.031	0.037	0.036
	(0.025)	(0.022)	(0.022)*	(0.017)**
Treatment Group	0.023	-0.014	-0.025	-0.018
	(0.026)	(0.020)	(0.022)	(0.031)
Post Randomization	-0.093	-0.074	-0.093	-0.095
	(0.010)***	(0.011)***	(0.011)***	(0.013)***
Above Median in Prediction	0.077	0.012	-0.060	0.040
	(0.021)***	(0.024)	(0.023)***	(0.020)**
Above Median x Post x Treatment	0.014	0.008	-0.019	-0.008
	(0.030)	(0.028)	(0.028)	(0.028)
Above Median x Post	0.023	-0.027	0.019	0.016
	(0.017)	(0.017)	(0.017)	(0.017)
Above Median x Treatment	-0.040	0.020	0.048	0.018
	(0.034)	(0.033)	(0.030)	(0.038)
N	56,130	56,130	56,130	56,130
Mean if Share Above Median	0.655	0.606	0.637	0.601
Mean if Share Below Median	0.554	0.602	0.571	0.607

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table reports results for school attendance based on the predicted share of processes by the team member(s). The prediction is summarized by a binary variable indicating whether the observation is above or below the median of the prediction. We use those binary indicators as controls and as heterogeneity variables. All models include strata indicators.

I.2 Mediation Analysis

We next explore whether the improvement in the crime-report and schooling-attendance outcomes are related to the reduction in the time in care. We do this first via a mediation analysis by controlling for the number of days in residence. This is speculative as the days in residence is endogenous. When we control for the number of days in residence, we continue to find a similar reduction in crime reports for the treatment group.(Table I.5. This suggests that the improvement in crimes and attendance are related to the family rehabilitation and other services facilitated by the legal team rather than simply duration in foster care.

Table I.5: Days in Residence as a Mediator

Dependent Variable:	(1) Crime Reports/Qtr. Usual Estimate	(2) Crime Reports/Qtr. Control for Residences	(1) School Attendance/Qtr. Usual Estimate	(2) School Attendance/Qtr. Control for Residences
Treatment x Post	-.0375	-.0407	.0291	.0339
	(.0134)***	(.0148)***	(.0127)**	(.0128)***
Treatment Group	.0104	.0197	-.00295	-.00533
	(.0125)	(.0229)	(.0165)	(.0158)
Post Randomization	.0932	.0634	-.103	-.115
	(.0102)***	(.01)***	(.0081)***	(.00802)***
In Residence		-.0003		.00134
		(.0001)***		(.00011)***
N	54,259	33,678	22,452	22,452
N of children	1,871	1,871	1,871	1,871
N Control Group	1,188	1,188	1,188	1,188
Control Group Mean	0.125	0.125	0.580	0.580

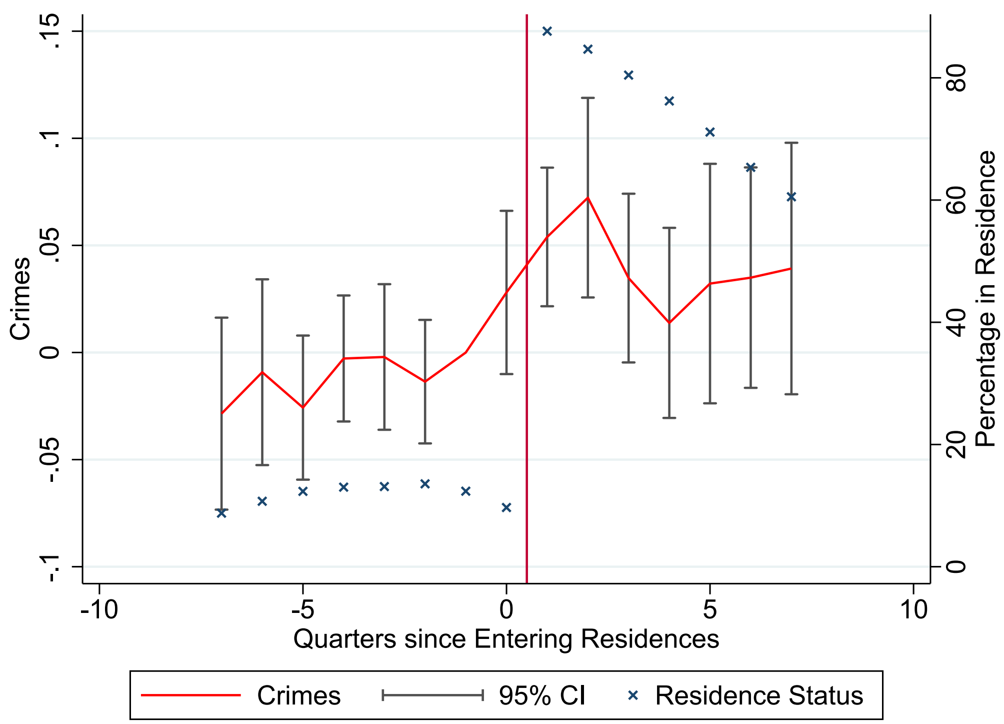
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: This table reports models with and without controlling for the number of days spent in residence each quarter. Note that attendance is measured at the same frequency as number of days in residence: quarterly, as opposed to the main attendance results, which are presented at the monthly level. Control Group Mean indicates the mean in the post-period. Standard errors are clustered at the child level. All models include strata indicators.

I.3 Crime Reports Around Entry and Exit from Foster Care

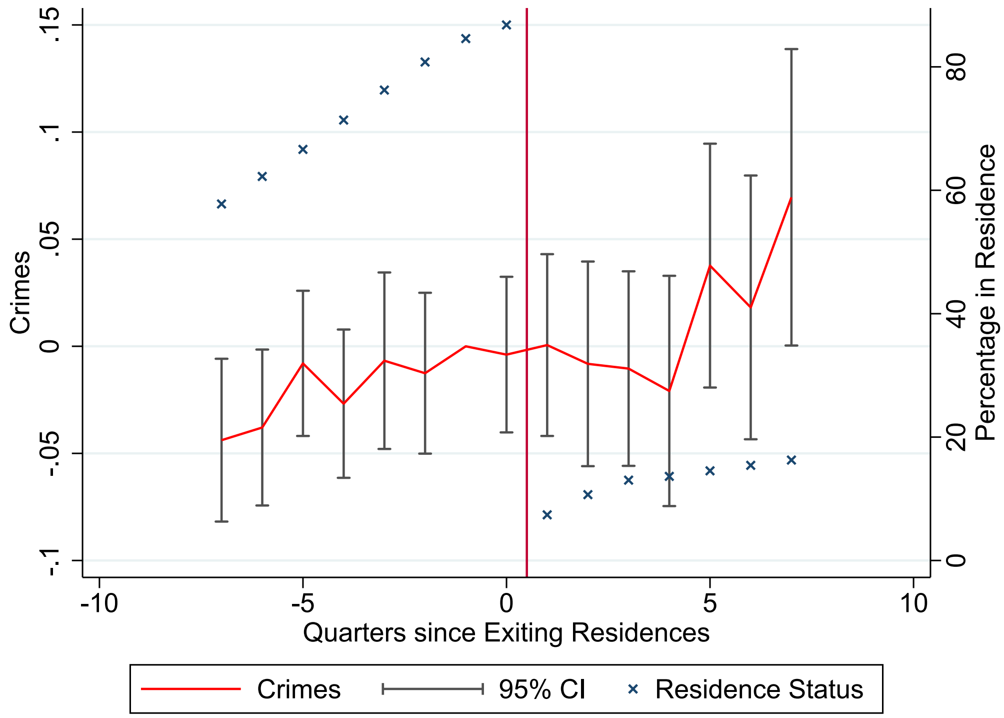
One mechanism that would explain the drop in crime reports found for the treatment group would be a reduction in delinquency while in residence due to altercations or greater surveillance. Below are event studies that trace the rate of crime reports before and after entry or exit from the facility. These time series are suggestive as they do not include a control group. Rather, we wish to examine whether there is a sharp change in crime reports upon entry or exit. Figure I.1 shows that crime reports begin to rise just prior to entering the facility and remain at an elevated level. This suggests that entry is correlated with problematic behavior on the part of the child. Figure I.2 shows that crime reports do not fall once children exit residences, which suggests that a change in surveillance does not account for the main results.

Figure I.1: Event Study on Entering Residences



This chart shows an event study for crime reports before and after entry into a residence. The omitted period is two quarters before entering residence, to test whether crime reports precipitate entry. The vertical line is the moment of entering residences. The estimates are obtained from a regression of crime reports on indicators of the number of quarters since the child entered a residence, age indicators (to control for the increase in crime that comes with age), and child fixed effects. Confidence intervals are calculated using standard errors clustered at the child level.

Figure I.2: Event Study on Exiting Residences



This chart shows an event study for crime reports before and after entry into a residence. The omitted period is two quarters before exiting residence, to test whether crime reports are changing just prior to exit. The vertical line is the moment of entering residences. The estimates are obtained from a regression of crime reports on indicators of the number of quarters since the child entered a residence, age indicators (to control for the increase in crime that comes with age), and child fixed effects. Confidence intervals are calculated using standard errors clustered at the child level.

J Dynamics: Complier Characteristics Over Time

To explore non-compliance over time, it is useful to describe the complier characteristics and how they may evolve over time. Consider the effect of being assigned to the treatment group on the number of days since first exposure to the Mi Abogado program over the first year after randomization. This is a "first stage" estimate that describes how the treatment group received greater access to the program. The relative likelihood of a complier characteristic is simply the ratio of the first stage for that group divided by the overall first stage (Angrist and Pischke, 2008).

Table J.1 reports the estimates when the first stage for days of exposure is estimated through Q1 2020; the next table shows the analogous estimates when the first stage is estimated through Q1 2021. We find that compliers are more likely to have larger sibling groups, more time in residence, are younger, and are female. As a summary, when predicted permanency is above median, and predicted crime is below median, the child is more likely to be a complier (i.e. cases more prone to having positive outcomes). The same pattern is found when we analyze the first stage across the first two years rather than the first.

Table J.1: Complier Characteristics Q1 2020

X	Above-Median First Stage (1)	(1)/Overall First Stage	Mean X Below Median	Mean X Above Median	N
Number of Siblings	80.656	1.07	0.18	3.49	1,871
	(15.017)				
Delay in Schooling	72.958	0.97	0.26	2.64	1,871
	(14.613)				
Time in Residence	86.204	1.15	2.96	4.45	1,871
	(11.732)				
Age When First in Residence	57.983	0.77	7.53	13.79	1,871
	(11.975)				
Age at Randomization	68.336	0.91	10.82	16.23	1,871
	(11.368)				
Gender(Girl)	87.905	1.17	0.00	1.00	1,871
	(11.106)				
Predicted Permanency	76.694	1.02	0.06	0.09	1,871
	(11.256)				
Predicted Crimes	67.642	0.90	0.19	1.09	1,871
	(11.173)				
	Overall First Stage	Compliers			
Full Sample	75.226	929			
	(8.571)				

Note: This table reports the first-stage coefficient for days of Mi Abogado exposure (days since first participating) by the end of Q1 2020 when the characteristic, X, is greater than the median for the sample at baseline. It then reports the ratio of this first stage to the overall first stage, along with the mean of the characteristic when it is below its median and when it is above its median.

Table J.2: Complier Characteristics Q1 2021

X	Above-Median First Stage (1)	(1)/Overall First Stage	Mean X Below Median	Mean X Above Median	N
Number of Siblings	119.116	1.02	0.18	3.49	1,871
	(31.089)				
Delay in Schooling	122.634	1.05	0.26	2.64	1,871
	(31.528)				
Time in Residence	136.349	1.17	2.96	4.45	1,871
	(24.627)				
Age When First in Residence	91.276	0.78	7.53	13.79	1,871
	(26.504)				
Age at Randomization	113.730	0.98	10.82	16.23	1,871
	(25.368)				
Gender(Girl)	147.395	1.27	0.00	1.00	1,871
	(23.775)				
Predicted Permanency	134.709	1.16	0.06	0.09	1,871
	(24.460)				
Predicted Crimes	106.272	0.91	0.19	1.09	1,871
	(24.522)				
	Overall First Stage	Compliers			
Full Sample	116.371	1100			
	(18.094)				

This table reports the first-stage coefficient for days of Mi Abogado exposure (days since first participating) by the end of Q1 2021 when the characteristic, X, is greater than the median for the sample at baseline. It then reports the ratio of this first stage to the overall first stage, along with the mean of the characteristic when it is below its median and when it is above its median.

K Cost-Benefit Analysis with Crime Outcomes

To obtain the cost of crime estimates, we assigned each type of crime a cost. Thus, the total social savings from the program are the estimated treatment effects on the different types of crimes summed up over the post-randomization period (641 days) multiplied by the total average costs of each type of crime. The average cost for each type of crime is calculated based on estimates in Miller et al. (2021). We apply a deflation factor equal to the ratio of Chile's per capita GDP to the United States' (0.20) to place the estimates in US dollar terms.

Table K.1: Cost Benefit Analysis with Crime Outcomes

	Mean T	Mean C	Dif	P-Value	Costs	Dif*Costs
A. Legal-aid Costs						
Days of Legal Aid in MA Program	296.51	205.95	90.57	0.00	4.99	451.87
Days of Legal Aid outside MA	76.76	175.74	-98.98	0.00	2.73	-270.09
B. Residence Costs						
Days in residence (public)	111.07	115.85	-4.78	0.48	67.27	-321.47
Days in residence (nonprofit)	281.17	307.54	-26.37	0.07	28.35	-747.58
C. Family Foster Care Costs						
Days in care (nonprofit)	11.88	6.88	5.00	0.26	13.94	69.69
Net SENAME Costs						-817.58
D. Crime						
Property	0.35	0.43	-0.08	0.08	1,698.43	-129.09
Violent	0.35	0.48	-0.13	0.00	24,507.68	-3,202.08
Substance	0.03	0.03	-0.00	0.82	2,196.56	-3.87
Net Criminal Justice Costs						-3,335.04
Total						-4,152.61

Note: Estimates are on a per-child basis, and the observation period is 641 days. Costs are calculated in 2022 US dollars based on estimates in Miller et al., 2021. MA costs include a 90-days period of supervision after a child exits from SENAME care.