The Center for
State Child Welfare Data

Do Intensive In-Home Services Prevent Placement?:
A Case Study of Youth Villages' Intercept® Program

Scott Huhr and Fred Wulczyn

January 2020

CHAPIN HALL
AT THE UNIVERSITY OF CHICAGO

1313 E. 60th Street
Chicago, IL 60637
(773) 256.5100
fcda.chapinhall.org

Table of Contents

Introduction

In this report, we describe our assessment of whether Youth Village's Intercept program (previously known as YVIntercept) had a demonstrable impact on the likelihood of out-of-home placement for children at risk of placement who were referred to the program by the Tennessee Department of Children's Services (DCS). In the words of Youth Villages, Intercept is an integrated approach to in-home parenting skill development that offers a variety of evidence-based and best practices to meet the individualized needs of a family and young person (S. Hurley, personal communication, September 12, 2019). Intercept staff work with families with children who are at risk of either entry or re-entry into state custody (i.e. foster care) so as to prevent placement. DCS staff also refer children already in custody, with the goal of reducing time to reunification. Intercept prevention services generally last four to six months; reunification services generally last six to nine months. Among those at-risk children, only children who were referred to the program before their first placement were included in this analysis.

The study was commissioned by Bonnie Hommrich and Doug Swisher to learn whether those goals were/are being met. At the time, Bonnie Hommrich was the Commissioner of the Tennessee Department of Children's Services; Doug Swisher was the Associate Commissioner for Finance and Budget. Over the years, DCS had made a substantial investment in services designed to keep children from coming into care. The evaluation was seen as an opportunity to test whether Intercept was having the intended effect. The study was funded by Youth Villages.[1] Staff at the Center for State Child Welfare Data, which is located within Chapin Hall at the University of Chicago, carried out the study independently. Details of the approach taken are described in the sections that follow.

Program Background and Implementation

Under contract with DCS, Youth Villages has been providing Intercept since the late 1990s. The program goal then and now is to reduce the utilization of foster care (broadly defined to include congregate care in its various forms), either by preventing entry into care, reducing the time spent in care, and/or reducing the risk of re-entry. DCS allocates Intercept slots to counties based on recent trends in the number of children in custody in the county. DCS caseworkers, supervisors, or regional leadership make referrals to Intercept by phone, email, or online for youth at risk of coming into custody. Referrals are received by Intercept regional leadership or placement staff; the family is contacted within 24 hours and an assessment is scheduled at the family's earliest convenience. The completed assessment is sent to DCS staff for approval; once approval is received, Intercept services begin, usually within 24 hours of approval. If the family is Medicaid-eligible, services may be funded through TennCare (the state Medicaid program) rather than DCS; services may be funded through DCS until authorization for services is obtained from TennCare.

Regarding the program itself, Intercept employs Bachelor's- and Master's-level family intervention specialists, who are trained to engage families. Key features of the program include program intensity (meeting with families an average of three times weekly), low staff caseloads of 4 – 5 families, active 24/7 on-call structure, and structured weekly supervision and consultation from a licensed clinician who is an expert in the model. Program fidelity is monitored annually, key operational indicators are reviewed monthly, and post-discharge outcomes are gathered

[1] A few additional words about how the study was conducted are warranted. As mentioned, the work was commissioned by the Tennessee Department of Children's Services. The evaluation was funded by Youth Villages with resources provided to it by Blue Meridian Partners. Although we received data from Youth Villages and the state of Tennessee, the conduct of the study was carried out by Data Center staff without interference from either Tennessee DCS or Youth Villages. We consulted with both Youth Villages and DCS during the course of our analysis, to make sure our data processing rules were in accordance with accepted practices on the part of both parties. We met on a bi-weekly basis to discuss questions that emerged and the study findings were reviewed by Youth Villages' Evaluation Research Advisory Committee. That said, staff at the Data Center are solely responsible for the conduct of the analysis, interpretation of the findings, and preparation of the final report.

from youth and families for two years.

At its core, Intercept draws on a diverse set of evidence-based and research-informed interventions. In keeping with the use of evidence- and research-informed programs, Intercept is manualized. There are specific underlying principles that undergird the Intercept intervention and, in accordance with best practices, model adherence is measured using information derived from multiple sources including youth, parents, staff, program leadership, case record review, and an assessment of long-term outcomes.[2]

Data

The study uses administrative records within a quasi-experimental design. As opposed to a well-designed randomized clinical trial (RCT), quasi-experimental designs have to confront the problem of selection bias into treatment programs. To overcome the selection bias, we used multiple empirical strategies. Details of the approach adopted follows.

Overview

To conduct the study, we relied exclusively on administrative records provided to us by DCS and Youth Villages. DCS provided us with data from the Tennessee Family and Child Tracking System (TFACTS), the state's administrative data system (i.e., Statewide Automated Child Welfare Information System [SACWIS]). Data extracted from TFACTS included:

- Maltreatment investigations and related information including child-level characteristics and such characteristics of the maltreatment report as the report date, allegation type, perpetrator, disposition, and the caseworker responsible for managing the case. Because DCS operates a dual track Child Protective Services (CPS) system, we also took into account whether the CPS case was referred to the assessment track or the investigation track. Finally, we pulled from TFACTS the county where the child was living at the time of the investigation.
- Placement data that track if and when a young person enters care, how long they were in care, their reason for leaving (i.e., reunification, etc.), and other aspects of their placement history (e.g., number of moves, type of placement setting).
- Data from the Family Advocacy and Support Tool (FAST) which is an assessment that is completed by DCS caseworkers as part of the CPS process. DCS caseworkers use the FAST to document their concerns pertaining to family safety, financial resources, and conflict, among other domains.
- Caseworker role codes that identify the worker responsible for referring the child to services and the assignment dates

From Youth Villages, we received Intercept enrollment data. The enrollment data capture the onset of Intercept services and a stop date that indicates when services ended. The enrollment data also includes the TFACTS ID for each child so that the link between both data sets is unambiguous.

These data were linked together using the TFACTS ID. After the link was completed, we organized the data around the timing of the investigation. Each child in the file was the subject of an initial investigation. Based on the information in the file, this would have been the first investigation ever within the time span of the underlying data. In this particular case, we have the complete CPS history from 2003 forward, so there is very little left censoring.

To these data, we added two other event types. One was based on referral to Intercept; the other was based on

[2] Additional information about the program is available at: www.youthvillages.org/intercept

out-of-home care placement. Those events were collated into the CPS file based on the event date, a process that left us with a chronologically ordered set of service events. With these data, we were able to identify precisely when, in relation to the child's CPS history, referral took place and under what circumstances (i.e., number and type of prior reports, disposition, perpetrators, allegations, etc.). The event histories also gave us a refined way to define placement prevention, reunification, and re-entry populations given what had already happened when the referral to Intercept took place.

Study Sample

Sample Period. Our sample universe encompasses children with an initial CPS report between 1/1/2013 and 6/30/2018 in Tennessee. From the date of the initial CPS report (the start date), children were observed over the following time periods: (1) 1,800 days from start date, (2) start date to censor date (6/30/2018), (3) start date to the date the child turned 18 years old, or (4) start date to placement into out-of-home care, whichever came first. As such, the observation period for any individual youth is 1,800 days from the CPS report date or less because of censoring, reaching age 18, or placement.

The Treatment Group. Referral to Intercept was established using linked administrative records received from Youth Villages and the Tennessee Department of Children's Services. The Intercept treatment group includes children who were referred to the Intercept program as described earlier. Eighty-seven percent of children who were referred to Intercept have a stop date and the average length of service is 95 days. Regardless of their level of participation, all children were included as part of the treatment group in the analysis. Thus, the analysis is an intent-to-treat (ITT) design. Among children referred to Intercept, some of them were referred after placement; however, only children who were assigned to Intercept before their first placement were included in the prevention sample, because the impact of Intercept on reunification and re-entry are not the subject of this study. Regarding fidelity, because we used an ITT design, we felt a detailed fidelity assessment was less important at this time. That said, Youth Villages' does maintain internal clinical and operational monitoring systems, which indicated that Intercept was delivered with fidelity during the years covered by this study. Going forward, a study of fidelity in light of outcomes achieved would provide valuable implementation insights.

Dependent Variable

Our analysis considers the likelihood of placement into out-of-home care given a referral to Intercept, a placement prevention program. We expect children referred to Intercept will have a lower likelihood of placement than similar children because of the services offered through Intercept. As described more fully below, the observation time is divided into discrete time intervals, with the risk of placement assessed during each interval. The discrete time model captures the fact that the likelihood of placement changes with the passage of time.

Independent Variables

The independent variables are clustered as into a set of child and family characteristics; a set that describes the maltreatment investigation; and, a set of case characteristics that describes the risk profile of the child and the family derived from the Family Advocacy and Support Tool (FAST) assessment. The independent variables are used for the exact matching and as covariates in the statistical model.

Child and Other Characteristics. The assembled data includes child characteristics together with data about the perpetrator type and county name. Child characteristics include race/ethnicity, which was divided into four categories: African American, White, other races and ethnicities, and unknown. Age was calculated using the first investigation or assessment date and date of birth, with the result divided into five groups: infants (age 0), children between the ages of 1 through 5 (Age 1), ages from 6 through 10 (Age 2), ages from 11 through 13 (Age 3), and ages from 14 through 17 (Age 4). Children age 18 and older were dropped from the sample, given restrictions on placement after a child turns age 18, if there has been no prior placement.[3] The perpetrator category captures the perpetrator's relationship to the child. Perpetrator categories include parents, family members, acquaintances, caretakers, and other, which we coded as 1 for parent and 0 for all others.

Investigation/Assessment. For each child, TFACTS records an event history. The event date establishes when the event happened. Tennessee operates a dual track CPS system. Reports of maltreatment are either assessed for services or investigated. We used the records for both types of reports. When an assessment is completed, the disposition falls into one of six categories: (1) Assessment completed, services required (ASR), (2) Assessment completed, services accepted (ASA), (3) Assessment completed, services refused (ASF), (4) Assessment completed, services not needed (ANN), (5) Assessment completed, and no disposition entered (ASU), and (6) Assessment still in process (ASZ).[4] Investigation types have four different categories: (1) Investigation completed, allegation(s) substantiated (SUB), (2) Investigation completed, allegation(s) unsubstantiated (UNSUB), (3) Investigation completed, no disposition entered (ICU), and (4) Investigation still in process (ICZ). The assessment and investigation events were transformed into binary variables: ASR and ASA were coded as 1 to indicate services were needed whereas the other assessment types (ASF, ANN, ASU, ASZ) were coded as 0. The investigations were coded as substantiated (SUB; 1) or other (UNSUB, ICU, ICZ; 0).

The Family Advocacy and Support Tool (FAST). To account for different family and child experiences and background, FAST data were used. Among FAST data fields, seven major variables that can affect placement decisions were included: family finance, family safety, sexual/physical abuse, emotional abuse, neglect, education, and developmental/mental health/substance use. The descriptions of those variables are shown in Table 1, which summarizes the actual descriptions in the FAST manual (Epstein & Lyons).

Table 1: FAST Data Field and Description

Data Field	Description
Family Financial Resources	Income and other sources of money available to family members (particularly caregivers) that can be used to address family needs
Family Safety	The degree to which family members are safe from being injured in the home. This describes whether individuals in the home present a danger to the child or youth
Sexual or Physical Abuse	Child or youth experience of sexual or physical abuse
Emotional Abuse	Whether the child or youth experienced emotional abuse or not
Neglect	Failure to provide adequate supervision and expectations and access to the basic necessities of life, including food, shelter, and clothing
Education	The child or youth's status with school; indicates whether the child or youth is experiencing school problems
Developmental/Mental Health/Substance Abuse	Learning disabilities/currently experiencing mental health issues/alcohol or drug use

These seven FAST variables have a 4-point scale with the following action levels: '0' (no evidence that action is needed), '1' (history, watchful waiting, or prevention), '2' (action needed), '3' (immediate or intensive action needed), or 'NA' (not applicable), as noted in the manual. The FAST is completed by DCS staff. Based on the data, binary variables were created and their definitions are noted in Table 2.

Table 2: FAST Variable Values and Descriptions

[3] A child above the age of 18 may be placed into foster care but only if there had been a prior placement (and exit). By definition, children returning to care cannot be part of the prevention study. Instead, they fall into the re-entry study.
[4] The number of ASU cases is very small.

Data Field	Value	Description
Family Finance	0	Family has the income and other sources of money available to family members
	1	Financial hardship (mild, moderate, or severe)
Family Safety	0	Family provides a safe home environment
	1	Danger or risk in neighborhood or home environment
Sexual or Physical Abuse	0	No evidence for sexual or physical abuse
	1	Evidence for suspicion or actually experienced sexual or physical abuse
Emotional Abuse	0	No evidence for emotional abuse
	1	Experienced emotional abuse
Neglect	0	No evidence for neglect
	1	Experienced neglect
Education	0	No evidence for school problems or not school aged
	1	Experienced school problems
Developmental/Mental Health/Substance Abuse	0	No evidence for developmental or mental health issues, or alcohol or drug usage
	1	Experienced developmental or mental health issues or used alcohol or drugs

Other Confounds

Although the approach we used to match children in the treatment group with children in the comparison group provides a recognized solution for the problem of selection bias, there are other confounds that interfere with sound study designs based on administrative records (Brookhart, Stürmer, Glynn, Rassen, & Schneeweiss, 2010). Among them, differences in administrative structure and differences in the way workers approach their job are especially important in the child welfare context. For example, there is considerable county variation in the underlying placement rates. Courts, which operate at the county level in Tennessee, have considerable discretion when it comes to whether a child will be placed into out-of-home care. Caseworkers also exhibit different decision-making thresholds relative to whether a child will be placed out of their home or referred to services while remaining in the home with their family. To address these two particular concerns, we added two variables to the analytical file.

County. Regarding county variation in placement rates, we use a random effects (i.e., a hierarchical) model. Described more fully below, the county random effects model allows the intercept in the model to vary at the county level. The random effect captures unmeasured differences in county characteristics. For this purpose, we used the county where the child was living at the time the investigation was started.

Caseworker Assignment. TFACTS records caseworker assignments on a child-specific basis. Caseworkers have a unique ID, a role code, assignment start and stop dates, and a link to the ID attached to a child. Start date is the date that the caseworker is assigned to a child and stop date is the date that the caseworker stops working with a given child. Role code refers to the type of worker, including ongoing noncustodial worker, CPS assessment worker, and other designations. Because each TFACTS ID may be attached to multiple caseworker IDs (i.e., a child may have more than one assigned worker), we had to match children to the worker responsible for making decisions about service referrals. According to DCS officials, the on-going noncustodial worker has decision-making responsibility. To account for other assignments, we adopted decision rules for matching a caseworker and a child. Those rules are shown in Table 3.

Table 3: Decision Rules to Link TFACTS Child ID with Caseworker ID

Group	Description
Non-treatment Group	Use the most recent caseworker ID 1. Link using ongoing noncustodial worker first 2. If not linked, link using CPS assessment worker 3. If not linked, link using other most recent worker
Treatment Group	Use the most recent caseworker ID before the Intercept referral 1. Link using ongoing noncustodial worker first 2. If not linked, link using CPS assessment worker 3. If not linked, link using other most recent worker

In order to capture caseworker referral tendencies, we developed a separate statistical model in which we connected a worker to the children assigned to them. Technically, the referral to Intercept (yes/no) served as the dependent variable in a separate random effects logistic regression model with child and other covariates included.[5] From those models, we computed the Empirical Bayes (EB) residual which tells us the extent to which the worker's referral rate deviates from the adjusted average referral rate. When children share the same caseworker, they will have the same EB residual. The worker EB residuals were then added to the placement model as a way to adjust for the worker confound. Among the benefits, the adjustment for worker referral rate differences controls for unmeasured case characteristics that workers observe that contribute to why some children are referred to Intercept and others are not. When the worker EB residuals are added as covariates to the county random effects model, both sources of heterogeneity are accounted for in the impact analysis.

Methods

The study design addresses a set of interrelated challenges associated with the use of observational data in studies of intervention effects. Our first task addressed the matching process. As we show, the children referred to Intercept were, on average, from a higher-risk sub-sample of children. To identify a suitable comparison group, we followed closely the recommendations of the Title IV-E Prevention Services Clearinghouse Handbook for establishing a matched comparison group. The second issue relates to the problem of censoring. The opportunity to observe outcomes within the study sample varies by when the young person first came to the attention of DCS. In addition, some children are lost to follow-up because they are no longer age-eligible for placement services when the reach age 18. Last we had to contend with county context and the fact that county placement rates following a report to child protective services are a source of variation. Our specific solution to each of these issues is described below.

[5] The specifics of the model used to derive the EB residuals are available from the authors. That said, one additional detail is worth noting in this context. The EB residuals were derived from a separate random effects logistic regression model with referral to Intercept as the dependent variable. In building the model we had to contend with the fact that although children are nested within counties, workers are not, strictly speaking, nested within counties. That is, workers may serve children from more than one county. This creates a cross-classified data structure because children are, potentially, nested within two different entities - workers and counties. We tried to account for the cross-classified structure statistically, but the models did not converge. We adopted the two-step strategy to overcome this issue.

Matching Methodology

If the case mix and geographic differences are not taken into account then the unadjusted outcomes for Intercept referrals would likely show higher placement rates. Often times, the referral of more difficult cases (i.e., selection effects) in studies of treatment effectiveness is managed through random assignment within a randomized clinical trial (RCT). In this case, however, we only have observational data from an at-scale program rollout. To counteract the selection effect, we followed conventional methods to manage the problem of selective referral to a treatment program. As a general matter, the goal of these methods is to match the treatment and comparison groups as closely as possible on a set of observable characteristics so that we can say with reasonable certainty that the treatment and comparison groups are similar. Given those similarities, it is then possible to conclude that any positive treatment effect is because the treatment group received the service provided. Our specific approach to solving the problem of selection bias follows.

Baseline Equivalence

To measure program impact, we need to know what would have happened in the absence of program participation; however, because treated and non-treated cases cannot coexist (this is the "fundamental problem of causal inference" (Holland, 1986), the children who were treated have to be compared to a group of children who were not treated. This is where baseline equivalence becomes important. According to the Title IV-E Prevention Services Clearinghouse Handbook of Standards and Procedures (Handbook; Wilson et al., 2019), the methods used to establish baseline equivalence are those used by the What Works Clearinghouse (WWC). WWC was established by the Institute of Education Sciences of the U.S. Department of Education to evaluate existing research in education (https://ies.ed.gov/ncee/wwc/). Baseline equivalence deals with whether, based on observed characteristics, the intervention group and the comparison group are similar enough to each other. The Handbook explains the thresholds for baseline equivalence as follows:

> *"Specifically, baseline equivalence is assessed by examining baseline differences expressed in effect size (ES) units. Baseline effect sizes less than 0.05 are considered equivalent and no further covariate adjustments are required. Baseline effect sizes between 0.05 and 0.25 indicate that statistical adjustments in the impact models may be required. These baseline effect sizes are said to be in the adjustment range. Evidence of large differences (ES > 0.25) in demographic or socioeconomic characteristics can be evidence that the individuals in the intervention and comparison conditions were drawn from very different settings and are not sufficiently comparable for the review. Such cases may be considered to have substantially different characteristic confounds."*

According to the Handbook, if an effect size unit is less than 0.05, statistical adjustments are not required to examine program impacts; however, if it is in the adjustment range (between 0.05 and 0.25), statistical adjustments may be required to account for potential confounding effects. Effect size quantifies the difference between two groups. Effect size can be used to judge either covariate differences or impact size. However, effect size in this context, as explained in the Handbook, is not about treatment impact, but refers instead to a standardized mean difference between the treatment group and the comparison group. In other words, the extent to which the intervention group and the comparison group are similar at the baseline is determined by effect size units. Technically, this is calculated as the mean difference between the groups, divided by their pooled standard deviation.

Exact Matching

Propensity score matching (PSM) is used frequently to create equivalent treatment and comparison groups in quasi-experimental designs of treatment effects. Matching methods have been in use before PSM was introduced; however, the development of the propensity score changed the conversation regarding matching methods. The propensity score is defined as the probability of treatment assignment conditional on measured baseline covariates. The approach deals explicitly with violations of the ignorable (unconfounded) treatment assignment assumption, which means treatment assignment is independent of the potential outcomes given the covariates (Rosenblum & Rubin, 1983). As with other various matching strategies, PSM strives to reduce selection bias by constructing well-matched samples of the initial treatment and comparison groups. The assessment matched samples after PSM can be conducted by balance diagnostics that examine the distribution of baseline covariates.

One of the most common implementation strategies of PSM is 1:1 matching. This method selects the most similar untreated child (defined by propensity score) for each treated child. However, 1:1 matching is somewhat problematic because 1:1 matching throws out all unmatched sample members, which reduces statistical power. Further, when a large portion of unmatched data is thrown out, the generalization of the results is questionable. Also, multiple matches can reduce variance; however, using multiple controls for each unit can increase bias because when additional matched samples are used, their propensity scores are further away from the propensity score of the treated unit. This represents the typical trade-off between bias and efficiency. However, with the current data, the concern is side-stepped because we have a large number of children who were not referred and who have exactly the same covariates as children who were referred to Intercept.

Compared to other PSM methods, exact matching has multiple benefits. It is the least controversial matching method and does not depend on model assumptions. As opposed to nearest neighbor matching, exact matching includes only exactly matched children and therefore there is no concern for bias and distance weight (i.e. inverse probability weight). Exact matching uses all matched children (one-to-many matchings) which increases efficiency without causing bias. Because of the additional efficiency (without sacrificing bias), imposing a limit on the number of matched children is not recommended. Even though exact matching is ideal in many ways, it does not always work because exact matches lead to many unmatched units when matching variables are high dimensional, which may increase bias (Imar, King & Stuart, 2008). However, we are in a good position to use exact matching because of the large number of children who were not referred to Intercept.

Due to the nature of exact matching, the children in the matched comparison group won’t be matched to multiple children in the treatment group who may have slightly different propensity scores. One complicating factor is that there are identical twins in the treatment group. That is, they have exactly the same set of covariates. The identical twins in the treatment group will be matched to identical twins in the comparison group. In this case, the matching becomes like non-parametric cell matching, which means many children who were treated are matched to many children who were not treated, each in a mutually exclusive cell. However, this is different from matching with slightly different children who received treatment. Therefore, the weighting process based on propensity score distance is not needed for exact matching. In the final dataset, the matched control children do not have duplicates because they were included only once as a mutually exclusive cell match.[6]

[6] The next question could be “do we need to use frequency weights for those cases?” Frequency weights are sometimes used in regression-based models when a comparison child is matched to different children in the treatment group more than once in nearest neighbor matching (and even in coarsened exact matching), because the matched children are not independent due to duplicates. However, the current exact match does not have duplicates, so it is not necessary to have frequency weights. Due to exact matching, the thresholds for baseline equivalence in the matched sample were satisfied automatically because we only included identical matches in the comparison group. Technically, effect sizes become zero in weighted estimates. Exact matching is a preferred method to demonstrate baseline equivalence as explained in the Handbook.

Covariates for Baseline Equivalence

According to the Handbook, a direct pre-test outcome variable must be used to assess baseline equivalence. Alternatively, if using direct pre-test data is not possible or feasible, or a suitable pre-test alternative is not available, baseline equivalence must be established on both race/ethnicity and socioeconomic status (SES). The Handbook also requires baseline equivalence to be demonstrated on child age for studies of programs for children and youth.

None of the children experienced placements previously and therefore direct pre-test data are not possible and also an alternative pre-test does not exist. As a result, baseline equivalence was established using exact matching with the covariates described in the data section: gender, race/ethnicity, age, perpetrator type, family finance, family safety, sexual/physical abuse, emotional abuse, neglect, education, developmental/mental health/substance use (see Table 7). Regarding SES, family finance from the FAST was used: family finance refers whether income and other sources of money available to family members (particularly caregivers) were a concern at the time of assessment. Family safety refers to whether the family home environment/neighborhood is safe or poses danger or risk (Epstein & Lyons, 2014).

Table 4 shows how many children were included in the study after exact matching. Ninety-four percent of the children referred to Intercept were included in the final analysis sample as were 46 percent of the children who were not referred to Intercept. As such, 6 percent of the children in the treatment group do not have matched comparisons and 54 percent of children in the comparison group were not matched to any treated child. See Table 7 in the results section for characteristics of the treatment sample both before and after matching.

Table 4: Sample Size Before and After Matching

	Before Matching	After Matching	Matching proportion
Intercept Group	1,899	1,778	94%
Comparison Group	187,080	86,254	46%

Censoring

Even if we have matched samples, the placement outcome depends on how much time has passed after the initial assessment or investigation. Usually, the first six months after assessment have the highest placement rate. Thereafter the likelihood of placement falls considerably. Moreover, for children referred during the latter time periods, there is less time to observe placement (i.e., the observations are censored). To manage this issue, the time from assessment or investigation date until the stop date was divided into six-month time intervals with one record per interval of time through the end of the observation window (1,800 days, censoring date (6/30/2018), the date the child reached age 18, or the placement date, whichever came first). The person-periods, as they are called, divide the total time of exposure into discrete intervals of time. For this research, six-month periods (180 days) were used as shown in Table 5. P-1 stands for the first six months, P-2 stands for the next six-month interval, etc. Person-periods were assessed until 1,800 days (P-10) at the maximum. Constructed this way, the approach allows us to use as much of the available data as possible.

All children have the first period (P-1) and the number of children decreases in subsequent person-periods as children are placed, reach age 18 or the observation period ends. If the length of observation within the sample period from the time of first CPS report until the observation ends (censoring date, maturity, and placement) is less than 181 days, that child will have only P-1. If more than 180 and less than 361 days, then only P-1 and P-2 are available. If more than 1,620 and less than 1,801 days, then P-1 through P-10 are available. For example, if a child has 500 days of length of exposure to placement, the child has three person-periods (P-1, P-2, and P-3).

If the child is placed at 200 days, the outcome for P-1 is coded as zero and the outcome for P-2 is coded as one, indicating that placement occurred during this particular person-period. As such, until a child is placed, the outcome for all prior person-periods is coded as zero. The placement outcome becomes a person-period specific outcome which means we are measuring the likelihood of being placed during specific intervals. In the analysis, each interval has its own placement probability and therefore the impact of the Intercept intervention is assessed after accounting for how much time has passed since the initial report. Even though the goal is to measure the overall impact of the Intercept regardless of person-periods, we do not assume that the impact will be uniform across all person-periods. In fact, we specifically examine whether the treatment effect is particularly strong soon after the CPS referral.

Table 5: Person-Periods by Treatment / Comparison Group

Interval	Person-Period	Original Comparison	Original Treatment	Final Comparison	Final Treatment
1-180 days	P-1	187,080	1,899	86,254	1,778
181-360 days	P-2	162,716	1,783	74,658	1,669
361-540 days	P-3	140,763	1,612	64,516	1,510
541-720 days	P-4	117,471	1,438	53,549	1,351
721-900 days	P-5	94,345	1,247	42,350	1,170
901-1080 days	P-6	70,921	1,076	30,908	1,006
1081-1260 days	P-7	52,549	909	22,153	850
1261-1440 days	P-8	42,422	748	17,797	699
1441-1620 days	P-9	32,845	573	13,694	531
1621-1800 days	P-10	22,904	433	9,499	401
More than 1800 days	P-11	13,160	243	5,350	226

Discrete-time Hazard Model

There are two major approaches for solving the problem of censored data: the Cox proportional hazard model and discrete-time hazard model (Singer and Willet, 2003). In contrast to Cox proportional hazard model, which uses one record per child, discrete time models divide time into intervals (six-month time intervals in this case), with one record per interval of time through the end of observation for a given child (i.e., placement or reach age 18). If censoring occurs, the outcomes of all person-periods are recorded as zero. Between the two methods, we opted for the discrete-time hazard model (DTHM) because it offers a number of advantages. First, the DTHM calculates the risk of placement for each person-period. Second, when testing for interaction effects that involve time (e.g., are children who are physically abused more likely than other children to be placed soon after the investigation starts), the DTHM offers more flexibility and transparency to test specific interactions than the Cox proportional hazard model. Third, in the event there are between-county differences in placement rates, the DTHM addresses the nested data structure in a straightforward manner.

Random Effects Model

The between-county differences alluded to earlier pose important substantive and statistical concerns. For example, because counties operate in slightly different ways (e.g., the courts in Tennessee are organized at the county level), children with similar characteristics may have different placement rates because of administrative operating differences. Statistically, when the data are clustered in this way, standard errors are affected, which in turn affects the statistical significance of the parameter estimates. The random effects model applied to the DTHM provides a convenient way to manage this issue. The random effects model allows parameter estimates to vary (i.e., county differences in placement rates) and adjusts the standard errors.

The DTHM model is illustrated below using a hierarchical form with separate equations for the person- and county-levels. This follows the standard exposition on multi-level models (Raudenbush and Bryk, 2002).

Child level (level one): $\eta_{ijt} = \ln(h_{ijt} / (1 - h_{ijt})) = \beta_{0j} + \beta_{1j}X_{ij} + \beta_{2j}D_{ij} + \Sigma T_t P_{ijt}$

η_{ijt} is the log of the odds of the outcome (placed = 1) for child *i* in county *j* at discrete time *t*, h_{ijt} is the hazard of the outcome for child *i* in county *j* at time *t*, D_{ij} represents the Intercept indicator for child *i* in county *j* and X_{ij} represents child-level covariates for child *i* in county *j*. After linking a worker to the children assigned to them, a worker's EB residual was used as one of child level covariates. P_{ijt} is an indicator variable of discrete person-periods. T_t [t from 1 to 10] represents intercepts for different discrete time intervals, which forms the baseline hazard rate. As mentioned before, discrete-time intervals are constructed due to censoring.

County level (level two): $\beta_{0j} = \beta_{00} + \beta_{01}C_j + \gamma_{0j}$

For a county-level model, β_{0j} has a subscript *j*, which means each county has a unique intercept. For exposition purposes, β_{0j} includes county-level fixed variables, C_j, so that β_{0j} becomes the adjusted intercept for children in county *j*. β_{01} is the adjusted difference in the placement rate associated with county variable C_j. β_{00} refers to the overall intercept. However, when person-periods (P_{ijt}) are included along with the intercept, the intercept refers to the intercept (placement rate) for the missing person-period. Then, the *t* -1 person-period estimates are relative to β_{00}. Alternatively, for a no-intercept DTHM, which includes all *t* estimates, $T_1, T_2, \ldots T_{10}$ forms the baseline hazard rate. We used the no-intercept version of the DTHM with random effects.

In this model, γ_{0j} is a level-2 random variable and represents the adjusted average placement rate in county *j*. The presence of γ_{0j} changes the model to a random effects model. In terms of distributions, the county intercepts are assumed to be normally distributed with an expected value of zero. Therefore, the individual county placement intercepts are deviations from zero.

Combined Model (levels 1 and 2 together): $\eta_{ijt} = \ln(h_{ijt} / (1 - h_{ijt})) = \beta_{00} + \Sigma T_t P_{ijt} + \beta_{1j}X_{ij} + \beta_{2j}D_{ij} + \beta_{01}C_j + \gamma_{0j}$

The mixed or combined model is formed by algebraic substitution. As shown, the model contains fixed components (overall intercept, person-period intercepts, level 1 covariates, the program variable, and level 2 covariates) and one random component (γ_{0j}). The model used for the final analysis (Model 3 in Table 8) is below. Note that $\beta_{0j} = \gamma_{0j}$ in this case, due to the use of a DTHM without an intercept and county variables. SAS proc glimmix was used to conduct the analysis.

$$\eta_{ijt} = \ln(h_{ijt} / (1 - h_{ijt})) = \Sigma T_t P_{ijt} + \beta_{1j}X_{ij} + \beta_{2j}D_{ij} + \gamma_{0j}$$

Results

Descriptive Statistics

Within the observation period, 188,979 children were reported to DCS for the first time. Of those 1,899 children (1%) were assigned to Intercept; the remaining 187,080 children make up the potential comparison group. Among those children, the number of children placed from the treatment group and the potential comparison group were

229 (12.1%) and 14,833 (7.9%), respectively (see Table 6).

Table 6: Sample Size and Placement before Matching

	Total Children	Number of Children Placed	Percent Placed (Rate)
Comparison	187,080	14,833	7.9%
Intercept	1,899	229	12.1%

The higher placement rate found among young people referred to Intercept is indicative of the likelihood that referrals to Intercept may come from a higher risk sub-population within the larger group of children reported to DCS. If so, this would indicate a case mix difference. The higher placement rate may also reflect how the program was deployed geographically. Placement rates vary from one county in Tennessee to another. Because DCS asked Youth Villages to roll out the Intercept program in counties with higher placement rates, the baseline difference may reflect that implementation choice. Both issues – the case mix differences and the implementation strategy – require statistical adjustments, which we described earlier in the report.

Table 7 shows that older children were more likely to be referred to Intercept than younger children, as are male children. Regarding race/ethnicity, White children were more likely to be referred to Intercept. The likelihood of referral was higher among families where someone other than the parent was the perpetrator. Fewer Intercept children have a parent as a perpetrator (0.5% vs. 2.0%). Although children with substantiated investigations were less likely to be referred, referral percentages among children with ASR/ASA assessments were comparable.

Regarding FAST assessment, the children who were referred to Intercept had a higher percentage of family finance and family safety issues than members of the comparison group (0.9% vs. 1.4% and 0.8% vs. 1.8%, respectively). Also, the group referred to Intercept had more children impacted by sexual and physical abuse, emotional abuse, neglect, school problems (education), and developmental/mental health/substance use concerns. In particular, Intercept referrals were three times more likely to be affected by sexual/physical abuse and emotional abuse (0.7% vs. 2.2% and 0.7% vs. 2.9%, respectively). Also, children referred to Intercept were 12 times more likely to have developmental/mental health/substance use issues (0.3% vs. 3.6%). In sum, results from the FAST assessment show that the children referred to Intercept had more challenging and riskier individual and family characteristics, a fact that no doubt contributed to the higher baseline placement rate.

Table 7: Child and Family Characteristics Before and After Matching

		Before Matching			After Matching	
Characteristic	Value	Treatment Group	Comparison Group	Percent Served	Treatment Group	Comparison Group*
Age	Infants	34	28,296	0.1%	31	7,927
	1 to 5	179	56,309	0.3%	164	27,954
	6 to 10	489	52,153	0.9%	463	25,079
	11 to 13	615	25,890	2.3%	569	12,726
	14 to 17	582	24,432	2.3%	550	12,568
Gender	Male	1,053	92,564	1.1%	989	42,891
	Female	846	94,516	0.9%	789	43,363
Race/Ethnicity	African American	203	12,240	1.6%	167	3,295
	White	872	39,378	2.2%	841	14,683
	Other	128	7,972	1.6%	88	1,140
	Unknown[7]	696	127,490	0.5%	682	67,136
Perpetrator	Parent	601	122,133	0.5%	558	51,873
	Other	1,298	64,947	2.0%	1,220	34,381
Assessment: ASR/ASA	No	1,644	160,929	1.0%	1,551	81,430
	Yes	255	26,151	1.0%	227	4,824
Investigation: Substantiated	No	1,816	171,397	1.0%	1,717	85,403
	Yes	83	15,683	0.5%	61	851
Family Finance	No	1,350	148,355	0.9%	1,283	79,439
	Yes	549	38,725	1.4%	495	6,815
Family Safety	No	1,296	153,281	0.8%	1,246	81,951
	Yes	603	33,799	1.8%	532	4,303
Sexual/Physical Abuse	No	1,027	147,616	0.7%	981	75,269
	Yes	872	39,464	2.2%	797	10,985
Emotional Abuse	No	1,208	163,870	0.7%	1,164	82,533
	Yes	691	23,210	2.9%	614	3,721
Neglect	No	1,275	150,838	0.8%	1,219	81,124
	Yes	624	36,242	1.7%	559	5,130
Education	No	907	161,600	0.6%	857	78,404
	Yes	992	25,480	3.7%	921	7,850
Developmental Concerns/ Mental/Substance Use	No	510	149,494	0.3%	488	72,107
	Yes	1,389	37,586	3.6%	1,290	14,147
Total		1,899	187,080	1.0%	1,778	86,254

* As shown, the comparison group frequencies are unweighted. The weighted cell percents (not shown) for the comparison group are identical to the cell percents for the treatment group because of exact matching.

Table 7 also shows the post-match treatment and comparison group composition. In order to establish baseline equivalence, we used exact matching. Exact matching matches each child referred to Intercept to one or more comparison group children who have the same combination of covariate values. No comparison children are matched twice due to the nature of exact matching. After exact matching, the final sample consists of 1,778

[7] Two observations are worth noting with respect to unknown race. First, as children moved further into the system – i.e., children were substantiated and referred to services or placed into foster care, the number of children with an unknown race category dropped substantially. Very few children who were placed had their race/ethnicity listed as unknown. Second, not knowing a child's race or ethnicity does not mean, for our purposes here, that no information is provided. To the extent that unknown race is correlated with lower risk, then unknown is useful from the perspective of matching and statistical adjustment. For additional information, see Footnote 8.

children referred to Intercept and 86,254 comparison children. In order to show who was included in the final treatment sample, the characteristics of the treatment and comparison samples are included in the final columns of Table 7.

Average Treatment Effect

The results of the random effects model are found in Table 8, which displays model coefficients and their standard errors, p-values, and odds ratios (O.R.). Odds ratios greater than one are associated with an increased likelihood of placement in out-of-home care. Odds ratios smaller than one are associated with a lower likelihood of placement.

Estimates from four different models are shown. Across all the models, the estimate of interest is the Intercept parameter, which represents the treatment effect. Model 1 (all 1,899 treated and 187,080 comparison children) shows the parameter estimates before exact matching was employed. Models 2, 3 and 4 are based on exact matching. The Intercept Tx. effect in Model 1 and 2 refers to the program's overall impact (i.e., the average treatment effect), which assumes that the baseline hazard rates for person-periods are all equal. Models 3 and 4 introduce the person-period specific intercepts, so the assumption of equivalent baseline hazard rates is dropped. Model 4 includes the interaction terms between person-periods and the Intercept program.

Regardless of model type, the demographic characteristics of age, gender, and race/ethnicity are associated with statistically significant effects. Babies (infants younger than 1 year old) serve as the reference category, which means the rate of placement for other ages is evaluated relative to the rate at which babies are placed. Babies are more likely to be placed than other children. Male children are more likely to be placed. White children (reference category) are less likely to be placed than African American children and children of other races/ethnicities. The models in Table 8, the race unknown category was used as a dummy variable together with African American, White, and Other, which includes children of Hispanic origin. For unknown race, we imputed values for African American, White, and Other. However, to preserve the unique information linked to children whose race was listed as unknown, we retained the unknown dummy variable in the models even after imputation[8]. When race/ethnicity is unknown, they are far less likely to be placed. We suspect that race/ethnicity was not recorded if the case was managed out of the system quickly before certain information could be collected.

Regarding perpetrator type and assessment/investigation type, children whose parent(s) were the perpetrator were slightly more likely to be placed. Children whose assessment outcome is ASR or ASA, in contrast to other assessment outcomes, show a higher likelihood of placement. Also, children whose investigation was substantiated show a higher likelihood of placement.

Regarding the Intercept treatment effect, the treatment parameter indicates that Intercept lowered placement rates among the children referred to the program as compared to a similar group of children who were not referred to Intercept. The Intercept treatment effects are similar in Models 1, 2, and 3 and all of them are statistically significant. Compared to Model 1, Model 2 incorporates exact matching as the basis for constructing the comparison group; Models 3 and 4 both include the person-periods in addition to exact matching; Model 4 includes interaction terms that test whether the Intercept treatment effect is sensitive to discrete time increments following initial contact with DCS. Based on the Model 3, after exact matching, the likelihood of placements of children who were referred to Intercept is 53 percent lower than those who were not referred Intercept.

[8] Multiple imputations (20 imputations here) for unknown race were conducted using SAS PROC MI with a fully conditional specification (FCS) method which assumes the existence of a joint distribution for all variables (The MI procedure, SAS Institute). Within FCS, the regression is for continuous variables and thus we used the discriminant function for the nominal categorical race variable. The parameter estimates from each data set were combined using SAS PROC MIANALYZE.

Table 8: Treatment Effect Estimates

	Model 1 Before exact mathing			Model 2 After exact matching w/out Person Periods			Model 3 After exact matching w/ Person Periods				Model 4 After exact matching w/ Person Periods and Interactions			
Effect	Estimate	s.e.	Pr > |t|	Estimate	s.e.	Pr > |t|	Estimate	s.e.	Pr > |t|	O.R.	Estimate	s.e.	Pr > |t|	O.R.
Model Intercept	-1.740	0.067	.0001	-1.877	0.160	<.0001								
Intercept Tx. Effect	-0.717	0.084	<.0001	-0.792	0.095	<.0001	-0.762	0.083	<.0001	*0.47*	-0.493	0.093	<.0001	0.61
*Intercept Tx. Effect*P-1*											-1.499	0.278	<.0001	0.22
*Intercept Tx. Effect *P-2*											-0.306	0.228	0.182	0.74
Person-Periods														
P-1							-3.125	0.151	<.0001		-3.076	0.151	<.0001	
P-2							-3.698	0.155	<.0001		-3.695	0.156	<.0001	
P-3							-3.642	0.156	<.0001		-3.663	0.156	<.0001	
P-4							-3.881	0.161	<.0001		-3.904	0.161	<.0001	
P-5							-3.729	0.161	<.0001		-3.753	0.161	<.0001	
P-6							-3.516	0.162	<.0001		-3.541	0.162	<.0001	
P-7							-3.431	0.166	<.0001		-3.459	0.166	<.0001	
P-8							-3.548	0.172	<.0001		-3.577	0.173	<.0001	
P-9							-3.599	0.181	<.0001		-3.627	0.181	<.0001	
P-10							-3.321	0.184	<.0001		-3.351	0.184	<.0001	
Age														
Infants < 1	Reference													
1 to 5	-0.792	0.036	<.0001	-0.752	0.146	<.0001	-1.006	0.136	<.0001	0.37	-1.004	0.136	<.0001	0.37
6 to 10	-1.756	0.038	<.0001	-1.678	0.151	<.0001	-1.809	0.139	<.0001	0.16	-1.807	0.139	<.0001	0.16
11 to 13	-1.576	0.041	<.0001	-1.261	0.152	<.0001	-1.343	0.140	<.0001	0.26	-1.342	0.140	<.0001	0.26
14 to 17	-1.708	0.043	<.0001	-1.433	0.153	<.0001	-1.101	0.141	<.0001	0.33	-1.100	0.141	<.0001	0.33
Gender														
Females	Reference													
Male	0.147	0.022	<.0001	0.243	0.048	<.0001	0.222	0.043	<.0001	1.25	0.222	0.043	<.0001	1.25
Race/ Ethnicity														
Whites	Reference													
African Am.	0.188	0.036	<.0001	0.356	0.075	<.0001	0.330	0.068	<.0001	1.39	0.330	0.068	<.0001	1.39
Unknown	-4.075	0.039	<.0001	-3.845	0.090	<.0001	-3.533	0.090	<.0001	0.03	-3.535	0.090	<.0001	0.03
Other	-0.110	0.036	0.003	0.223	0.109	0.040	0.321	0.100	0.001	1.38	0.319	0.100	0.001	1.38

	Model 1 Before exact mathing			Model 2 After exact matching w/out Person Periods			Model 3 After exact matching w/ Person Periods				Model 4 After exact matching w/ Person Periods and Interactions			
Effect	Estimate	s.e.	Pr > |t|	Estimate	s.e.	Pr > |t|	Estimate	s.e.	Pr > |t|	O.R.	Estimate	s.e.	Pr > |t|	O.R.
Perpetrator														
Other perpetrators	Reference													
Parent	0.237	0.027	<.0001	0.022	0.060	0.712	-3.533	0.090	<.0001	0.03	0.020	0.055	0.714	1.02
Assessment														
Other	Reference													
ASR or ASA	1.472	0.028	<.0001	1.514	0.070	<.0001	1.318	0.061	<.0001	3.74	1.318	0.061	<.0001	3.73
Investigation														
Unsubstantiated	Reference													
Substantiated	2.385	0.032	<.0001	1.449	0.137	<.0001	1.316	0.119	<.0001	3.73	1.314	0.119	<.0001	3.72
FAST														
Finance	0.258	0.025	<.0001	0.118	0.068	0.083	0.019	0.059	0.752	1.02	0.018	0.059	0.763	1.02
Safety	0.454	0.026	<.0001	0.495	0.068	<.0001	0.399	0.059	<.0001	1.49	0.400	0.059	<.0001	1.49
Sexual/Physical	0.137	0.027	<.0001	0.247	0.063	<.0001	0.195	0.056	0.001	1.21	0.193	0.056	0.001	1.21
Emotional	0.215	0.029	<.0001	0.198	0.075	0.008	0.127	0.065	0.051	1.14	0.130	0.065	0.046	1.14
Neglect	0.735	0.025	<.0001	0.642	0.068	<.0001	0.476	0.060	<.0001	1.61	0.474	0.060	<.0001	1.61
Education	0.478	0.029	<.0001	0.479	0.059	<.0001	0.345	0.054	<.0001	1.41	0.344	0.054	<.0001	1.41
Developmental	0.472	0.027	<.0001	0.518	0.060	<.0001	0.407	0.056	<.0001	1.50	0.408	0.056	<.0001	1.50
Caseworker														
EB Residual	0.118	0.021	<.0001	0.166	0.042	<.0001	0.124	0.038	0.001		0.125	0.038	0.001	
EB*EB	-0.103	0.020	<.0001	-0.104	0.038	0.006	-0.095	0.033	0.004		-0.096	0.033	0.004	

Caseworker's EB residual measures the worker's propensity to refer children to Intercept services as compared to the Intercept referral rate as observed for all workers in the county. The statistical significance of this parameter estimate suggests that a caseworker's tendency to refer to Intercept program is associated with placement outcomes. The EB parameter detects the linear trend, whereas the EB * EB parameter indicates that the relationship is not strictly linear.

Differential Effect

The average treatment effect captures the overall treatment effect of the Intercept program across all person-periods. In other words, the approach assumes the treatment effect is, by and large, the same within each person-period. To test this assumption, the interaction between the Intercept Program and P-1 and P-2 was used. Because only the first two person-periods (P-1 and P-2) were used for interaction terms, the main effect of the Intercept intervention now indicates an average treatment effect of P-3 through P-10. The impact of the first and second interaction terms (-1.499 for P-1 interaction and -0.306 for P-2 interaction) are relative to the baseline Intercept impact (-0.493). The baseline for YV*P-1 and YV*P-2 is the Intercept effect from P-3 through P-10, rather than the comparison group. Therefore, to contrast with the comparison group, logits have to be adjusted by adding the logit for Intercept, as shown in Table 9.

Table 9: Differential Treatment Effects

Effect	Logit	Recalculated logit	Odds ratio
Intercept Program	-0.493		0.61
Intercept Program*P-1	-1.499	-1.992 = [(-0.493) + (-1.499)]	0.14
Intercept Program*P-2	-0.306	-0.799 = [(-0.493) + (-0.306)]	0.45

The odds ratios in Table 9 indicate that Intercept has the most powerful placement prevention effect in P-1, which corresponds to the period of time that stretches from the initial investigation/assessment until six months later. Despite the lower odds ratio in P-2 than the odds ratio of the average from P-3 through P-10, the difference between P-2 and the remaining person-periods was not statistically significant. These findings indicate that the Intercept treatment effect is generalized over all person-periods but more pronounced in the first six months (odds ratio = 0.14) after the first investigation/assessment.

Sustained Effect

From the analysis of average treatment effects and differential (i.e., period-specific) effects, we found that (1) the Intercept program decreased the likelihood of placement overall, and (2) that the impact is more pronounced in the months soon after the initial investigation or assessment (i.e., during the first person-period). The next analysis considers whether the program impact is sustained after treatment is completed.[9] For this analysis we identified the group of children listed as having completed treatment in Youth Villages' administrative data. Of the

[9] As with other observational studies, a study of sustained effects has to contend with the problem of self-selection. Parents and children who finish an intervention may have different motivations or characteristics as compared to the parents and children who did not finish the intervention. These differences may be correlated with the outcome, which clouds the interpretation of the findings. Though we used the same approach to matching children from the treatment group with children from the comparison group, we note that the results are not as clear cut in part because we cannot arbitrarily assign an end-of-treatment date to children in the comparison group. Period-specific comparisons are, therefore, not possible. We tried multiple approaches to the analysis and the results are consistent: for children in the treatment group, the risk of placement following the end of treatment is lower than the risk for children in the comparison group.

1,899 children who were to referred to Intercept, 1,614 children completed the program. Out of the 1,614 children, 1,513 children were included in the study of sustained effects after conducting exact matching. We refer to these children as Completers in Table 10. These children were matched to the original comparison group using the same exact matching methods described earlier. Out of 86,254 comparison children, 81,325 children matched exactly and were used as members of the comparison group.

For the analysis of sustained effects, we used the same post-investigation/assessment 6-month person-period structure as before. Our first assessment considers the likelihood of placement for Completers as compared to the matched children from the comparison group across all person-periods. An example may help clarify the specific comparison being carried out. Child A is a member of the Treatment Group, referred to Intercept two-months after the start of the investigation. Treatment was completed within four months. The assessment of the sustained effect begins, then, in Month 7 (i.e., the second person-period). The risk of placement in the second post-investigation/assessment person-period, for children who completed the service during the first person-period, is compared with the likelihood of placement for children in the comparison group who are at-risk of placement in the second person period and so on. In other words, the risk of placement is assessed at a comparable period of time post investigation/assessment for both the treatment and comparison groups.

Evidence regarding the average sustained effects is presented in Model 5 of Table 10. The average sustained treatment effect covers all time periods available for follow-up. The odds ratio of the average sustained effect is 0.43 ($p < .0001$), which suggests that regardless of how long after treatment ends, the risk of placement is lower for the treatment group than it is for the comparison group.

Model 5 does not differentiate the length of post treatment completion (i.e., the sustained effect twelve-months post treatment completion). To address those specific concerns, we constructed an additional model with Completers divided into two groups: (1) children with twelve or fewer months of post-treatment follow-up and (2) children with thirteen or more months of post-treatment follow-up. As shown in Model 6 of Table 10, these groups are labeled Completers – 12 or fewer months and Completers – 12 or more months, respectively.

The results, found in Model 6 of Table 10, indicate that among children with no more than 12 months of post-treatment follow-up, the risk of placement, as measured by the odds ratio, is lower (.17) than the risk faced by children in the comparison group at comparable points in time. For children with 13 or more months of follow-up time, the risk of placement is also lower than the risk facing members of the comparison group at comparable points in time.

Table 10: Estimates for Sustained Treatment Effect

	Model 5				Model 6			
Effect	Estimate	s.e.	Pr > |t|	O.R.	Estimate	s.e.	Pr > |t|	O.R.
Intercept								
Comparison group	Reference				Reference			
Completers -	-0.836	0.090	<.0001	0.43				
12 or fewer months					-1.791	0.208	<.0001	0.17
13 or more months					-0.499	0.098	<.0001	0.61
Person-Periods	Reference							
P-1	-3.084	0.154	<.0001		-3.045	0.154	<.0001	
P-2	-3.641	0.159	<.0001		-3.599	0.159	<.0001	
P-3	-3.604	0.160	<.0001		-3.629	0.160	<.0001	
P-4	-3.830	0.164	<.0001		-3.856	0.165	<.0001	
P-5	-3.685	0.165	<.0001		-3.713	0.166	<.0001	
P-6	-3.500	0.167	<.0001		-3.529	0.167	<.0001	
P-7	-3.366	0.170	<.0001		-3.397	0.170	<.0001	
P-8	-3.571	0.179	<.0001		-3.605	0.179	<.0001	
P-9	-3.626	0.189	<.0001		-3.658	0.189	<.0001	
P-10	-3.299	0.191	<.0001		-3.333	0.191	<.0001	
Age								
Infants < 1	Reference				Reference			
1 to 5	-1.015	0.140	<.0001	0.36	-1.013	0.140	<.0001	0.36
6 to 10	-1.852	0.143	<.0001	0.16	-1.850	0.143	<.0001	0.16
11 to 13	-1.368	0.143	<.0001	0.25	-1.368	0.143	<.0001	0.25
14 to 17	-1.138	0.144	<.0001	0.32	-1.138	0.144	<.0001	0.32
Gender								
Female	Reference				Reference			
Male	0.226	0.045	<.0001	1.25	0.225	0.045	<.0001	1.25
Race/Ethnicity								
Whites	Reference				Reference			
African American	0.372	0.075	<.0001	1.45	0.370	0.075	<.0001	1.45
Unknown	0.372	0.075	<.0001	1.45	-3.608	0.095	<.0001	0.03
Other	0.291	0.107	0.007	1.34	0.290	0.107	0.007	1.34
Perpetrator								
Other perpetrators	Reference				Reference			
Parent	0.003	0.058	0.960	1.00	0.001	0.058	0.987	1.00
Assessment								
Other	Reference				Reference			
ASR or ASA	1.339	0.064	<.0001	3.81	1.339	0.064	<.0001	3.82
Investigation								
Other dispositions	Reference				Reference			
Substantiated	1.373	0.128	<.0001	3.95	1.368	0.128	<.0001	3.93
FAST								
Finance	0.033	0.062	0.601	1.03	0.031	0.062	0.616	1.03
Safety	0.378	0.063	<.0001	1.46	0.378	0.063	<.0001	1.46
Sexual/Physical	0.224	0.059	0.000	1.25	0.221	0.059	0.000	1.25
Emotional	0.074	0.069	0.281	1.08	0.077	0.069	0.259	1.08
Neglect	0.462	0.063	<.0001	1.59	0.460	0.063	<.0001	1.58
Education	0.369	0.057	<.0001	1.45	0.367	0.057	<.0001	1.44
Developmental	0.397	0.059	<.0001	1.49	0.398	0.059	<.0001	1.49
Caseworker								
EB Residual	0.112	0.039	0.004		0.113	0.039	0.004	
EB*EB	-0.094	0.035	0.008		-0.095	0.035	0.007	

In sum, the average sustained effect, across all time periods post-treatment is in the expected direction and statistically significant. Placement risk is lower among children in the treatment group regardless of how long after treatment ends. Moreover, there is a particularly strong effect within first 12-months post-treatment.

Conclusion

To evaluate the impact of the Intercept program on reducing the likelihood of out-of-home placement, multiple empirical strategies were employed to overcome the selection bias inherent in quasi-experimental designs. In particular, to establish baseline equivalence exact matching was used. Also, the discrete-time hazard model and the random effects model were employed with controls for county-level random effects and differences in worker referral rates.

- When it comes to the program impacts, the average treatment effect was assessed first. Among children referred to Intercept the risk of placement was 53 percent lower than the children in the comparison group.
- In addition, we analyzed whether program effects are similar across person-periods. The findings indicate that the Intercept treatment effect was generalized over all person-periods, but more pronounced in the first six months after the first investigation/assessment.
- Sustained effects were assessed as well. The findings indicate that the sustained effect is more pronounced with first 12-months and also persists beyond twelve months.
- We can reasonably conclude, therefore, that not only does the Intercept program reduce the likelihood of placement, but the program has sustained effects, beyond the end of treatment.

References

Brookhart, M. A., Stürmer, T., Glynn, R. J., Rassen, J., & Schneeweiss, S. (2010). Confounding control in healthcare database research. *Medical Care*, 48, S114–S120.

Epstein, R.A., & Lyons, J.S. (2014*). Family Advocacy and Support Tool*. Praed Foundation.

Holland, P.W. (1986), Statistics and causal inference. *Journal of the American Statistical Association*, Vol. 81, No. 396, pp. 945-960.

Imai, K., King, G. & Stuart, E. A. (2008), Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society Series A,* 171(2): 481-502.

SAS Institute. *The MI procedure*. Retrieved from:| https://support.sas.com/documentation/onlinedoc/stat/141/mi.pdf

Raudenbush, S., & Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Newbury Park, Ca: Sage.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, Volume 70, Issue 1, pp. 41–55.

Singer, J.D & Willet, J.B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* Oxford University Press, pp. 325-406.

Stuart, E.A. (2010), Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, Volume 25, Number 1, pp. 1-21.

Wilson, S. J., Price, C. S., Kerns, S. E. U., Dastrup, S. D., & Brown, S. R. (2019). *Title IV-E Prevention Services Clearinghouse Handbook of Standards and Procedures*, version 1.0, OPRE Report # 2019-56, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

What Works Clearinghouse Procedures and Standards Handbook, Version 3.0, Institute of Education Sciences (IES), the U.S. Department of Education. Retrieved from https://ies.ed.gov/ncee/wwc.