Int. Journal on Child Malt. (2024) 7:175–204
https://doi.org/10.1007/s42448-023-00187-4

RESEARCH ARTICLE

Effects of a Voluntary Hosting Program for Child Welfare Involved Families

William Schneider[1] · **Mark F. Testa**[2] · **Stephen Budde**[3]

Accepted: 17 November 2023 / Published online: 6 December 2023

Abstract

The number of children who are removed into formal foster care in the USA remains stubbornly high. To address this concern, child welfare agencies are seeking safe ways of diverting low-risk cases from formal foster care to informal alternative care. However, little is known about the outcomes and wellbeing of children who spend some time in informal placements, particularly in the homes of unrelated caregivers. The current study reports results from a pre-registered, experimental evaluation of Safe Families for Children, a voluntary hosting program for children whose parents are being investigated for child maltreatment. Drawing on a Bayesian paradigm, it analyzes the effects of the hosting program in both formative and summative randomized controlled trials conducted in downstate Illinois. Findings indicate that the intervention deflects alleged and indicated victims of maltreatment from the formal foster care system to the voluntary alternative care of host families. The program demonstrates positive or null effects across a variety of child welfare outcomes, including subsequent episodes of indicated maltreatment and return to or maintenance in the parental home. Findings from this work contribute to the ongoing debate about the benefits and risks of informal non-kin care as a preventive alternative to the removal of children into formal foster care.

Keywords Child maltreatment · Foster care · Informal care · Child wellbeing

✉ William Schneider
ws16@illinois.edu

Mark F. Testa
mtesta@email.unc.edu

Stephen Budde
sbudde@jpa.org

[1] School of Social Work, University of Illinois, Urbana, IL, USA

[2] School of Social Work, University of North Carolina, Chapel Hill, NC, USA

[3] Juvenile Protective Association, Chicago, IL, USA

Introduction

Although the literature is mixed (Font & Gershoff, 2020), a range of research indicates that extended periods of time spent in formal foster care is associated, on average, with adverse consequences for children (Doyle, 2007, 2008). However, other work suggests that foster care may have positive effects on child wellbeing (e.g., Gross & Baron, 2022). The range of investigated outcomes include social-emotional development, educational attainment, involvement with the juvenile justice system, teen motherhood, and adult employment (Lawrence, Carlson, & Egeland, 2006; Pears et al., 2013). Despite the ongoing debate about the effects of formal foster care on child wellbeing, there is widespread agreement that avoidance of the need for foster care remains a primary objective of the child welfare system. Placing a child in formal foster care for an extended period can be stressful and painful for children and parents. Many children who remain long-term in formal foster care experience placement disruptions (Vreeland et al, 2020), which exacerbate behavior problems and compound the trauma they experienced from the original separation from their home (Rubin et al, 2007).

Background

Formal foster care differs from informal alternative care chiefly in terms of the legal authority that courts grant child protective services (CPS) agents over the care, control, and socialization of a child. Whereas with informal alternative care legal authority remains vested with parents who voluntarily delegate parental responsibilities to relatives, friends, or other acquaintances, formal foster care involves the judicial transfer of parental responsibilities to public agents who are expected to adhere to bureaucratic rules of due process, equal treatment, and licensing of caregivers when arranging the foster care of children. Given the proliferation of asymmetrical power relationships between individual persons and collective agents in modern societies, such as police, hospitals, and CPS authorities (Coleman, 1982), a major debate has arisen over whether it is preferable to regulate potentially intimidating relationships by introducing greater formality through due process, legal representation, and performance contracts or by delegating greater discretion to persons linked together through informal solidarities of kin relations, neighborhood organizations, faith-based communities, and voluntary associations. A crucial test of the superiority of formal organization over informal solidarities is whether the former demonstrates greater effectiveness than the later in accomplishing the broader child welfare purposes of child safety, family permanence, and socio-emotional wellbeing. It is this test of superiority, which the current study attempts to undertake.

Formal Foster Care

The thrust of recent developments has pulled public child welfare policy and practice away from formal foster care in the direction of informal alternative care. The

number of children in formal foster care decreased from a peak of 567,000 in 1999 (USDHHS, 2006) to 424,000 in 2019 (U.S. Department of Health and Human Services, 2021a, b). Approximately one-third of the children in formal foster care were placed with relatives (U.S. Department of Health and Human Services, 2021a, b).

During the same period from 1999 to 2109, the number of children under the informal alternative care of relatives and nonrelatives with no parent in the home remained fixed at approximately 2.5 million (U.S. Census Bureau, 2022). Nationally, approximately 647,000 children lived in the homes of non-related caregivers in 2019. A much larger number, 2,319,000 children, lived apart from their parents in the homes of grandparents and other relatives. Subtracting out the number of children in formal foster or kinship care leaves 2.5 million children in informal alternative care referenced above. These alternative care arrangements are informal in the sense that the legal custody of the children remains with their parents as compared to formal foster care where a court transfers legal custody to a governmental agency (Radel et al., 2016). The context of the variety of children's living arrangements is notable given recent trends in child welfare of placements among kith and kin providers, and new federal regulations around guardianship. In effect, many children—even those outside of the formal child welfare system—spend time living with nonparental figures while the legal custody remains with their parents.

One of the reasons for the reduced reliance on formal foster care in recent years is the distrust engendered, particularly within racial minority communities, by the 1994 to 1996 federal prohibitions on any postponement of the foster placement of a child until a suitable home could be found, which matched the child's race, color, or national origin. Because the racial and ethnic characteristics of available foster caregivers in most states do not match proportionately the characteristics of children being taken into foster care, more Black children than White children end up in homes that do not reflect their race, color, or national origin.

Recent research indicates that the cumulative risk of entering formal foster care during childhood is 5.0% for White children, which is nearly half as large as the percentage (9.1%) for Black children (Yi et al., 2020). Research on the effects of formal foster care is mixed (see Barth et al., 2020 for a review). However, other work indicates that these concerns may be overstated, particularly in regard to educational outcomes (Berger et al., 2015). Nevertheless, policymakers have sought to identify safe alternatives to formal foster care given the racial disproportionality in removal rates and the desireability of speedy family re-unification.

Informal Alternative Care

To reduce racial disparity and lower the cumulative risk of entering formal foster care, child welfare agencies have increasingly emphasized on the formal side, race-neutral procedures such as blind removals (Baron et al., 2022), and on the informal side the diversion of low-risk cases to alternative care, where children are placed privately by parents or voluntarily with the aid of child protective services (Testa, Hill & Ingram, 2020). The tilt toward alternative care arrangements can be framed by a broader body of research and practice exploring the importance of family and community engagement in child welfare. Interventions focused on community

engagement in child welfare often provide targeted prevention programs that take a strength-based approach to building indigenous networks of local support (Ali et al., 2020). These programs seek to deepen ties between child welfare agencies and local communities to increase informal sources of support for families in need or at risk of child removal (Herrenkohl et al., 2020). In this sense, community engagement models are often designed to be pro-active and tailored to the specific needs of a particular community in order to provide needed support and prevent child removals. Notably, community engagement programs often seek to strengthen communities that have been historically marginalized (Feely & Bosk, 2021). However, little is known about the outcomes and wellbeing of children who spend some time in informal placements, particularly with non-kin caregivers (Gupta-Kagan, 2020).

Although there has been growing interest in the use of informal placements in child welfare, beliefs about the value of informal placement alternatives divide between supporters who believe informal placements are less restrictive, voluntary actions that help limit the scope of coercive state involvement in private family life (Anderson, 2014; Dettlaff et al., 2020). Skeptics argue that so-called voluntary informal placements actually feed into an unjust, "hidden" foster care system that evades legal due process and judicial oversight (Gupta-Kagan, 2020; Redleaf, 2018). Most of the published debate centers on the diversion of children to informal kinship care. A range of work has indicated positive outcomes associated with kinship care (Wu, 2017), including fewer behavioral problems, mental health challenges, less placement disruptions, and higher reunification rates (Winkour et al., 2015; Wu, White, & Coleman, 2015). However, other work has found less positive results (Ehrle & Geen, 2002). Usual forms of kinship and fictive-kin care have linked families to caregivers who are known to them including, family, friends, teachers, or other acquaintances in a child's community. It may be that the positive effects of kinship and fictive-kin care also extend to informal volunteer caregivers who are not related but are intrinsically motivated by spiritual calling, biblical norms of hospitality, and sentiments of generalized beneficence. If the positive effects indeed generalize beyond kinship, the pool of available informal caregivers would greatly expand and increase the opportunities for diversion from formal foster care.

Safe Families for Children

Safe Families for Children (SFFC) (https://safe-families.org/) recruits and oversees a network of unpaid host families with whom parents can voluntarily place their children in times of need. To that end, SFFC provides the option for temporary, voluntary, non-kin placement while allowing families of origin to retain custody of their children. This study reports results from an intent-to-treat evaluation of the SFFC program in Illinois. It tests the program's effectiveness in preventing judicial removal and the recurrence of maltreatment of children whose parents are investigated by Child Protective Services (CPS) compared to children who experience CPS interventions as usual.

SFFC: Theory of Change

SFFC's potential contributions to the overarching outcomes of child safety, family permanence, and socio-emotional wellbeing rest on the theory that voluntary hosting by families substitutes "bridging social capital"—or care from those outside of the child's immediate community, race/ethnicity, or social class—for the "bonding social capital"—or care provided from within a community—that may be lacking in the existing network of social support (Testa et al., 2010). In effect, programs such as SFFC seek to supplement existing ties within communities with additional voluntary resources from outside the community.

The motivational dynamics underlying the program align with the norms of spiritual generosity and hospitality illustrated in the story of Lydia found in the Bible. The program draws its inspiration from this parable, which is widely interpreted to encourage generosity and hospitality as a spiritual calling that "allows us to enter into dialogue with others from different cultural, racial, socioeconomic, and religious backgrounds" (Fleming, 2019: 52) By involving diverse families in regular interactions of reciprocal exchange and mutual trust, the ties that may develop between placing parents and hosting families can potentially continue well after family reunification and help lessen the hardships that many socially isolated families experience when trying to meet their family's needs.

The program is designed to build community, resources, and social networks. In essence, the theory of change underpinning SFFC is that by providing families with a voluntary alternative to formal removal, and which includes a range of supportive services, children will be less likely to enter into the formal foster care system and have subsequent contact with CPS. A range of prior research points to the importance of social networks as well as instrumental supports in child maltreatment prevention and intervention (Slack et al., 2003).

SFFC: Program Operation

SFFC is currently operating in over 40 sites across the USA. Despite its broad appeal, the effectiveness of the program had not been previously tested in the field under controlled experimental conditions. In 2013, the Illinois Department of Children and Families Services (DCFS) commissioned this evaluation to provide credible evidence of the program's effectiveness (or lack thereof) in preventing placement into formal foster care and the recurrence of child maltreatment.

After the Illinois study was commissioned, another randomized controlled field experiment was implemented in the North-East of England (Little, Warner & Baker, 2017). Even though the findings were suggestive of the program's value, the evaluation team at the Dartington Social Research Unit were able to enroll only 26 families out of the projected 360 families it estimated the experiment needed to detect with adequate statistical power a meaningful intervention effect. The evaluators acknowledged that 26 families was too low a number from which to draw reliable findings. They attributed the shortfall to lower than anticipated numbers of children whose primary caregiver requested a voluntary accommodation under the UK Children Act of 1989. They also attributed the lower than desired uptake to the evaluation

team's "overzealous methodological approach," which frustrated the desires of local authorities to avail all families of the hoped-for benefits of the program. Similar challenges beset the implementation and evaluation of SFFC in Illinois, which are detailed in the "Limitations" section below.

Almost all host families that participate in SFFC in Illinois are unrelated to the children.

Host families receive manualized training drawing on the Circle of Support curriculum, which includes information and resources such as supports to weather small crises, family coaching, and access to concrete resources. Host families also receive monitoring and management of support services while caring for children. SFFC is designed so that these resources and social networks can remain in place even after a child is returned to their family of origin.

Safety-Organized Practice and SFFC

The decision to refer a family to SFFC can be considered a form of "safety-organized practice" (Casey Family Programs, 2020; Meitner & Albers, 2012). The aim of the practice is to *preserve* children in the legal custody of their parents by developing a safety plan that insures children's safekeeping while the parents remediate the conditions that brought them to the attention of CPS or the agency clears them of the allegations of maltreatment. The plan is usually based on a formal safety assessment conducted by CPS, which assesses whether children are at serious risk of harm and in need of an immediate change in physical custody to insure their safety.

In Illinois, CPS uses the Child Endangerment Risk Assessment Protocol (CERAP; Fluke et al., 2001). Based on an itemized assessment of 16 safety threats, the investigator decides whether there is "clear evidence or other cause for concern" regarding the safety of any or all of the children. If the investigators in consultation with their supervisors deem one or more of the children to be "unsafe," protocol triggers the development of a safety plan. When the plan calls for a temporary change in the children's living situation, CPS may elect to delegate to a third party, such as SFFC, the primary responsibility of arranging the voluntary alternative care of the children. If the parents do not comply with the plan, the understanding is that one or more of the children may be removed from the home and placed into formal foster care.

Not all children referred to SFFC come from homes deemed "unsafe." Randomization to SFFC occurred after an investigator determined that a family could benefit from the program, including giving parents a reprieve from the daily demands of childrearing so they can arrange for remediating the conditions that brought them to the attention of CPS. Therefore, not all SFFC referrals involve immediate safety threats. Of the 49 children who were hosted by SFFC during the summative evaluation period (including crossovers), just 44% ($n = 18$) of the 41 with non-expunged records were deemed unsafe and in immediate need of a safety plan. However, given that randomization to SFFC and services as usual occurred after the investigator assessed that the family could benefit from the program, there is no reason to expect differential levels of risk between the treatment and control groups. The extent to

which referred children would actually have been removed and formally placed in the absence of SFFC is the subject of this evaluation.

Voluntary or Coerced

Even though Illinois Department of Children and Family Services (DCFS) rule and procedure declare that family cooperation with a safety plan is voluntary, some legal advocates allege that the threatened restrictions on a family's living arrangements are inherently coercive (Gupta-Kagan, 2020; Redleaf, 2018). Parents who agree to leave the home or place the children informally with relatives or SFFC host families are told they cannot modify the plan without risking removal of the children until they are cleared of the maltreatment allegations. Safety plans are generally implemented during a period of uncertainty while an investigation is underway. While this concern appears warranted for the approximately 40% of SFFC referrals that are made because the child is deemed unsafe, it is of less concern for referrals that are made for respite care or social support.

The Current Study

This study was approved by Institutional Review Boards (IRBs) at the Illinois Department of Children and Family Services (DCFS), the University of North Carolina at Chapel Hill (UNC), and the University of Illinois at Urbana-Champaign (UIUC) and was pre-registered on the Open Science Framework (https://osf.io/35ygk) in August of 2017.

Established in 2002 by LYDIA Home Association (https://www.lydiahome.org), a Chicago based, Christian nonprofit organization founded in 1916, SFFC partners with churches, ministries, and local community organizations to offer temporary 24-h care to children whose families are being investigated for alleged child abuse or neglect. During the course of investigation, families deemed appropriate for SFFC were allocated to either receive services as usual or be offered SFFC. Not all children referred to SFFC came from homes deemed "unsafe." Randomization to SFFC occurred after an investigator determined that a family could benefit from the program, including giving parents a reprieve from the daily demands of childrearing so they can arrange for remediating the conditions that brought them to the attention of DCFS. Most randomization assessments occurred within the first week of a reported allegation. Approximately 20% occurred a month or more after the allegation was reported. Therefore, not all SFFC referrals involved immediate safety threats.

In the current study, we extend preliminary work (Chen et al., 2020) to test the effectiveness of voluntary, unpaid non-kin care. We pre-registered three inter-related hypotheses for the evaluation:

1. Among child subjects investigated for maltreatment, the percentage taken into protective custody or later removed into foster care from day 1 to 24 months after allocation will be lower for the SFFC intervention groups compared to families who receive child protective services as usual (primary hypothesis).

2. Among child subjects investigated for maltreatment, the percentage who had a subsequent oral report of substantiated maltreatment within the first quarter and each subsequent quarter from the date of allocation to 24 months after the investigation will be statistically equivalent for both intervention and comparison groups.
3. Among child subjects investigated for child maltreatment, the percentage who are maintained in the custody of their parents or returned to their physical custody within 1 year after allocation to the treatment groups will be higher in the intervention group than the comparison group.

Study Design

The study was undertaken in two phases. We began with usability testing and formative evaluation designed to assess the construct validity of the SFFC logic model and theory of change (Chen, et al., 2020). Usability testing and formative evaluation included 276 families and their 597 children who were enrolled in the study prior to January 1, 2017. Review of the formative results, however, indicated substantial crossovers from the comparison to the intervention group in Cook County among other protocol violations. After several attempts at corrective action, it was decided to stop enrolling Cook County cases in the summative evaluation that was pre-registered on the Open Science Framework. Excluding them left 126 families and their 235 children in "downstate" Illinois counties. The summative evaluation was limited to downstate counties and included 99 families and their 216 children who were enrolled in the study for the 2-year period ending December 31, 2018.

The study tracked primary and secondary child welfare outcomes using public administrative data available to the investigators through June 30, 2020. The formative and summative evaluations did not meaningfully differ in programmatic substance. They were nearly identical in content, practice, and design, but they were undertaken sequentially at different times and increased precautions were taken to prevent treatment/control crossover in the summative evaluation. Notably, they drew from different samples across downstate Illinois, recruited during different years, and preceding participation in the formative study disqualified a family from participating in the summative study.

Procedure

Timing and Setting

The formative phase of the evaluation ran from the 4th quarter of 2015 to the 4th quarter of 2017, during which time annual intake into the Illinois foster care system was stable. The summative evaluation began in the 1st quarter of 2017 and ran through the end of 2019. During this period, annual intake into downstate foster care increased 43% from 3476 children in fiscal year 2017 to 4968 in fiscal year 2019.

Possible reasons for the 2017–2019 spike in annual downstate Illinois intake are explored in the "Discussion" section.

The SFFC evaluation built on DCFS's agreement to automate an unbiased allocation mechanism as part of its information system. In this way, a statistically equivalent group of families allocated to the intervention and comparison groups could be tracked in real time to measure differences in rates of foster care removal and repeat maltreatment. In effect, this is a "randomized encouragement design" (Holland, 1986), a fully valid type of RCT, which offers persons who are pre-allocated to the intervention, the option of participating in the treatment.

Families pre-allocated to the comparison group received treatment as usual, which could result in any of the following actions: dismissal of the allegation or, if indicated for maltreatment, closure of the case without further services, protective removal of the child for 48 h, the option of in-home family services, or formal placement of the child into foster care for an unspecified duration. The SFFC option was not offered to the comparison group. In this way, they were spared awareness of being denied services that the developers believed improved their chances of holding on to their children. Families pre-allocated to the intervention group were exposed to all of the usual treatment options plus the choice of participating in an SFFC host-family arrangement. Implementation of the evaluation focused on working with child protection staff who oversee and conduct investigations. Children and families were also referred to SFFC through services for intact families, which are provided after the investigation. These cases were excluded from the evaluation given the increased difficulty of monitoring compliance with group assignment and tracking child outcomes with the administrative data.

In the interests of concealing the allocation process and discouraging staff from tampering with assignments, the DCFS computer programmers installed a "behind the scenes," computerized routine for assigning families. Even though the routine was billed within the agency as a "randomizer," it was actually a binary, yes–no alternation routine that assigned every other referral to the comparison or intervention arm of the experiment. The process of forming study groups through alternation is a subset of an array of unbiased assignment methods, which include allocating subjects according to the day of the week, date of birth, medical record number, or the order in which subjects show up for services (Chalmers, 2011). Only supervisors had the security clearance to activate the "randomizer button," and they were geographically dispersed across the state. Referrals to SFFC were relatively rare occurrences (less than two in any office per week).

In addition to the automated alternation process, there was a second pathway for families to take up SFFC. Families could contact SFFC's intake team directly and have intake staff randomize them to the intervention or comparison group. Investigators generally knew the availability of SFFC for many years prior to the beginning of the study. It is likely that some families were aware of or informed of the program and sought services independently. This back-up assignment method was instituted after usability testing and the early phases of formative evaluation revealed that some investigators and families continued to contact SFFC directly as had been the practice in the past. Since treatment assignment was immediately known to SFFC staff because of direct contact by families, this second pathway was better

able to prevent treatment crossovers than the DCFS automated alternation process. Whereas 14% of the families that DCFS assigned to the comparison group managed to circumvent protocols and obtain hosting support, only one comparison case (1%) assigned by SFFC at intake crossed over to treatment.

Families who were randomized to the intervention group after contacting SFFC's offices were also far more willing to participate in the program (56%) than the families whom DCFS allocated to the intervention group (23%). Inspection of the reasons for incomplete compliance among parents randomized to the intervention group showed that among the 41 "no-shows," 15 (36%) had relatives step forward, 10 (25%) had their children removed by CPS within days of making the referral, 4 families (10%) withdrew their consent, and the remaining 12 (29%) were pulled back because other accommodations became available, a suitable host family could not be located, or the parents withdrew for other reasons.

Because no-shows are unlikely to be an unbiased subset of families assigned to the intervention group, the study retained all cases in the analysis of outcomes as though the families and children had completed the treatment to which they were originally assigned. This so-called intention-to-treat study design preserves the benefits of unbiased assignment for causal inference but yields an unbiased estimate *only of the effect of being assigned to treatment* and not of actually receiving the treatment. It was precisely this interest in learning about the practical impact of SFFC expansion on removal rates, regardless of family participation rates, which motivated the DCFS director at the time to commission a rigorous evaluation of the program.

Data and Measures

Data

The target population for the SFFC intervention is children who are the subjects of a report of child maltreatment that has been accepted for formal investigation by CPS. The primary problem that SFFC seeks to address is the removal of children either by taking them into state protective custody, which occurs temporarily while an investigation is ongoing, or judicially ordered placement into formal foster care. The key difference between foster care removal and informal hosting is that the former requires court action that transfers children to the legal custody of the CPS agency. In contrast, informal hosting leaves legal custody with the parents on the assumption that they will place the children voluntarily with an SFFC host family until the program or CPS agency deems it appropriate to restore them to the physical custody of their parents. We linked both SFFC participant families and families experiencing services-as-usual (SAU) to administrative data from the Illinois Department of Children and Family Services (DCFS) to examine outcomes.

Enrollment in the summative evaluation of SFFC lasted from January 1, 2017, to December 30, 2018. The enrollment period was originally scheduled to end in June of the following year, but was shortened at the request of the newly appointed DCFS

director. Outcomes were evaluated 12 months post-assignment to the comparison and intervention groups.

CONSORT Description of Sample

Figure 1 displays the Consolidated Standards of Report Trials (CONSORT) diagrams of the flow of SFFC referrals through the formative and summative stages of the enrollment process (Altman et al., 2001). A description of the excluded cases due to protocol deviations is provided in the "Limitations" section. After accounting for exclusions, there were 96 families and 235 children retained in the formative evaluation. There were 99 families and 216 children retained in the summative evaluation. The overall take up rate—or the number of families who were randomized to SFFC who chose to participate across the two studies was 32% (27% in the formative study and 37% in the summative study). Because families were not compelled to participate in SFFC, we estimate an intent-to-treat (ITT) effect. ITT yields an unbiased estimate of the effects of being assigned, but not the receipt, of treatment and is widely used in field experiments (Autor et al., 2022; Leventhal and Brookss-Gunn, 2003; List, 2022; Shadish et al., 2002).

Measures

Foster Care Deflection

The primary outcome is the time to the event of protective custody or court-sanctioned removal of the child into public foster care. In an emergency, DCFS

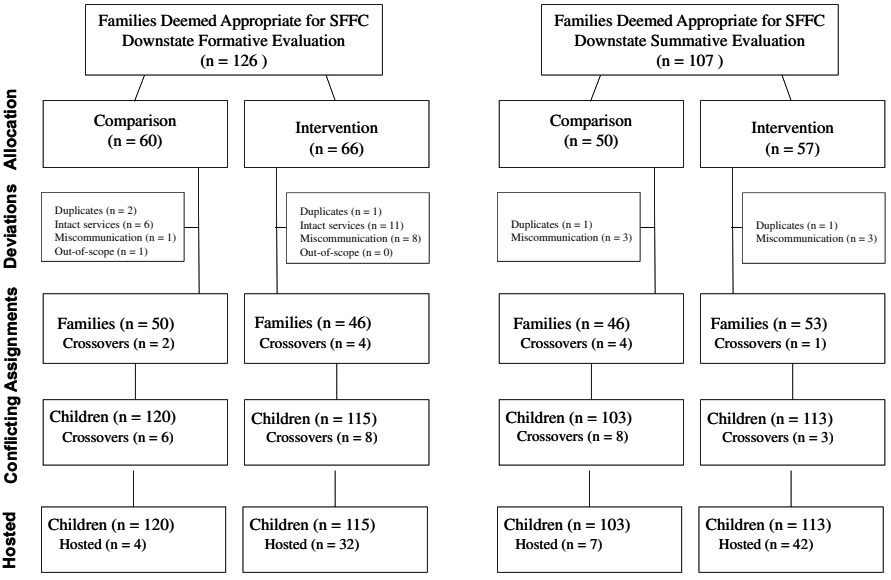

Fig. 1 CONSORT flow diagram of SFFC referrals, downstate Illinois

investigators have the authority to remove children from the physical custody of their parents without a court order or warrant. However, protective custody lapses within 48 h of removal unless a State's Attorney approves the filing of a petition with the juvenile court to retain the child in state custody.

Recurrence of Maltreatment

Each new episode of alleged child maltreatment receives a State Central Registry (SCR) number and sequence letter (A thru ZZ), which uniquely identify each new household configuration and the history of prior reports (starting with A) for that specific household configuration. In effect, the letters A through ZZ demarcate each new instance of allegation of child maltreatment, beginning with the first instance, A. In addition to the SCR number, each family member receives a person-specific identifier, which enables the tracking of subsequent reports of maltreatment and findings of indicated maltreatment for each child regardless of household changes. The federal indicator of recurrence of maltreatment measures whether a CPS agency is successful in preventing subsequent maltreatment of a child who was the subject of a prior indicated report. The follow-up interval for making this determination is 12 months after the initial report, excluding re-reports within 14 days of the previous report. This exclusion rule helps eliminate those false (positive) indicators of recurrence, which are based on allegations registered later but pertain to the original maltreatment report.

Family Support and Stabilization (Permanence)

One component of the theory of change, which underlies the SFFC intervention, is that host families will become the "functional equivalent" of a supportive social network that may be lacking in a parent's life. By offering the prospect of an ongoing supportive relationship beyond the date that SFFC restores the child to full parental custody, it is anticipated that after a year proportionately fewer children will have spent time in formal foster care compared to the time they might have spent in the absence of support from SFFC. For some parents with minimal social support, having children in SFFC may enable them to meet an urgent but temporary service need (e.g., housing, inpatient treatment) without having to place children in foster care. Table 1 shows that the families allocated to the different treatment groups are well balanced in terms of the date the families were allocated to the intervention and comparison groups. Therefore, the permanency outcome can be measured by ascertaining the whereabouts of the children (e.g., never removed, reunified back with their parents, or still in foster care) 12 months after their enrollment in the study.

Table 1 Families allocated to treatment groups by quarter, downstate Illinois

Quarter	Comparison *N*	Intervention *N*	Comparison %	Intervention %
Formative evaluation				
2016 1	6	7	5.5%	5.7%
2016 2	25	20	22.7%	16.3%
2016 3	17	20	15.5%	16.3%
2016 4	12	19	10.9%	15.4%
Summative evaluation				
2017 1	5	11	4.5%	8.9%
2017 2	14	9	12.7%	7.3%
2017 3	7	6	6.4%	4.9%
2017 4	1	5	0.9%	4.1%
2018 1	1	2	0.9%	1.6%
2018 2	10	10	9.1%	8.1%
2018 3	6	10	5.5%	8.1%
2018 4	6	4	5.5%	3.3%

Analytic Strategy

We draw on three analytic techniques: observed differences in proportions, hazard models pooling data from the formative and summative evaluations, and Bayesian analysis. The pre-registered analysis plan specified the use of transition (hazards) ratios to examine the effects of the offer to participate in the intervention (intent-to-treat). The advantage of this approach is that transition ratios make full use of the date and time information that are stored in most administrative data systems. For example, not only can it be determined whether a child was removed into foster care but also how quickly it occurred after allocation. The downside is that transition models are less familiar to a general audience than observed proportions or odds of children deflected from foster care. To that end, we start with the observed differences in the proportions and odds of deflections from foster care during the formative and summative evaluation periods.

Second, we estimate transition ratios. The transition ratio, also referred to as a hazard rate, takes into account the amount of time elapsed before the outcome or event of interest happens. Events that never happen during an evaluation period, i.e., the child is never removed, are said to be "censored" observations. Follow-up stops at the date of last observation. The transition ratio indicates the proportionate amount that the risk of experiencing a specific outcome or event is expected to vary at any particular time.

In this study, a transition ratio less than 1.0 implies that assignment to SFFC lowers the chances of an event's occurring at any particular time, whereas a ratio greater than 1.0 implies that SFFC raises the chances. Estimates close to 1.0 imply

no difference between SFFC and SAU in the chances of the event. A transition ratio below 0.5 is considered practically important for the primary outcome of deflection. A transition ratio below 0.5 signifies that assignment to SFFC reduces the estimated risk of removal by more than 50% on any particular day after assignment. The reason the transition ratio is favored over the odds ratio is that it facilitates updating prior information in light of newer information from replication studies without needing to take into account different durations of follow-up intervals (DePanfilis & Zuravin, 1998).

Last, we employ a Bayesian paradigm for updating beliefs in light of accumulated evidence—our preferred model. Prior beliefs ("priors" for short) under this paradigm refer to one's subjective sense of where the truth lies in advance of seeing the experimental results. For example, supporters of temporary informal placement begin with the prior belief that SFFC, on average, is superior to the status quo with regard to the primary outcome of family permanence and no worse with respect to the secondary outcome of child safety. Detractors start out with the opposite prior that diversion to informal placements, on average, is detrimental to family integrity and potentially unsafe for children.

Posterior beliefs refer to one's subjective sense of where the truth lies after seeing the results from an experiment or, preferably, a series of experimental trials. The Bayesian paradigm provides a set of rules for integrating experimental results from various times and settings. In the SFFC evaluation, we integrate findings from two independent trials conducted at different times—the formative and summative studies. The Bayesian paradigm offers a more comprehensive summary of the program's effectiveness by using the data from the formative evaluation as a prior that is updated with the findings from the summative evaluation to build cumulatively toward an understanding of program effectiveness as a posterior. That the formative and summative studies are nearly identical and undertaken at different time points is ideal for the Bayesian approach because it allows us to condition on a prior study that differs very little from the summative study, providing greater inference than is traditionally available (Gelman et al., 2021). It also allows us to assess how possible temporal changes in removal practices may affect the external validity of the findings.

The Bayesian models simulate 1 million separate intervention studies in which the formative and summative samples for each of the four outcomes were used to construct the data likelihood distributions.[1] The Bayesian posterior distribution incorporates the formative data into the data likelihood from the summative sample. A *p*-value was obtained from this posterior distribution to test the null hypothesis of non-superiority. The Monte Carlo estimate of Prob(I > C) is the proportion of the total number of simulations with I > C among the 1 million simulations (Chen & Fraser, 2017). In our primary models, we estimate one-tailed tests to assess whether the intervention is superior to services as usual.

[1] We are indebted to Dr. Ding-Geng Chen, professor in biostatistics at Arizona State University, for writing the R code and assisting us in estimating the frequentist and Bayesian *p*-values and probabilities of superiority.

The combined take-up rate of the two studies was 37% (Fig. 1). This take-up rate is in line with a wide range of other experimental studies from the economics, sociology, and psychology literatures that draw on intent-to-treat analyses (e.g., Autor et al., 2022; Karlan, et al., 2017; Leventhal and Brookss-Gunn, 2003). It is important to note that take-up rates in this range can be employed in intent-to-treat analyses because intent-to-treat estimates the effect of being offered treatment, rather than the effect of treatment itself, while still preserving the benefits of random assignment in terms of causal inference (List, 2022; Shadish et al., 2002). In this sense, intent-to-treat analysis provides insight into the expected benefits of the program should it be implemented more broadly, since no program can expect take-up rates at near universal levels in the real world.

Results

Table 2 provides descriptions of the formative and summative evaluation samples by treatment group for downstate Illinois outside of Cook County. The table indicates that the geographical spread is well balanced between treatment groups for the final formative and summative samples, which omit duplicate assignments, intact family referrals, and other protocol deviations (see the "Limitations" section for discussion). The distribution of original allocations looked similar. The same is true for case status and the findings on the percentage of reports that were subsequently indicated for maltreatment. At the conclusion of the observation period, similar proportions of records that exceeded statutory retention limits were expunged. Because a report that is indicated for maltreatment is an important predictor of removal from parental custody, the finding at the conclusion of the investigation is treated as a potential confounder of the effect of SFFC on the primary outcome of deflection from foster care.

An imbalance that arose by chance during summative evaluation is report sequence. The intervention group has a much lower percentage of sequence A (initial reports) reports than the comparison group (41.6% vs. 62.1%). Because sequence A reports are associated with a lower risk of removal, the imbalance could obscure the true effect of SFFC on the primary outcome of deflection from foster care. Therefore, we treat it as a potential confounder in the analyses.

Household (HH) allegation refers to the percentage of homes that included at least one indicated finding of physical or sexual abuse. The remainder of children are coded as neglect allegations even if they were not directly involved in the maltreatment investigation (e.g., the sibling of a substance-exposed infant). The treatment group differences for HH allegation and the mean ages of the children at referral—approximately 5 years old—are ignorable at conventional levels of statistical significance. On the other hand, the race/ethnicity distributions exhibit significant imbalances that could threaten the conclusions of our Bayesian analyses. Therefore, we also control for children's racial and ethnic origins in our analyses.

Table 2 Baseline characteristics of children at formative and summative phases by treatment group, downstate Illinois

Characteristic	Formative Sample ($N=235$)				Summative Sample ($N=216$)			
	Number		Percentage		Number		Percentage	
	Comparison	Intervention	Comparison	Intervention	Comparison	Intervention	Comparison	Intervention
Geography								
Central	54	55	45.0%	47.8%	36	45	35.0%	39.8%
Northern	57	52	47.5%	45.2%	54	48	52.4%	42.5%
Southern	9	8	7.5%	7.0%	13	10	12.6%	8.8%
Report sequence								
A	79	78	65.8%	67.8%	64	47	62.1%	41.6%
B	18	9	15.0%	7.8%	9	27	8.7%	23.9%
C	8	9	6.7%	7.8%	4	16	3.9%	14.2%
D or higher	15	19	12.5%	16.5%	26	23	25.3%	25.2%
Finding								
Indicated	81	85	67.5%	73.9%	71	71	68.9%	62.8%
Unfounded	39	30	32.5%	26.1%	32	42	31.1%	37.2%
Status								
Closed	90	88	75.0%	76.5%	87	90	84.5%	79.6%
Expunged	30	27	25.0%	23.5%	16	23	15.5%	20.4%
HH Allegation								
Abuse	42	30	35.0%	26.1%	31	46	30.1%	40.7%
Neglect	78	85	65.0%	73.9%	72	67	69.9%	59.3%
Race/ethnicity								
Black	70	47	58.3%	40.9%	43	39	41.7%	34.5%
White	31	49	25.8%	42.6%	38	53	36.9%	46.9%
Latinx	13	16	10.8%	13.9%	11	16	10.7%	14.2%

Table 2 (continued)

Characteristic	Formative Sample ($N=235$)				Summative Sample ($N=216$)			
	Number		Percentage		Number		Percentage	
	Comparison	Intervention	Comparison	Intervention	Comparison	Intervention	Comparison	Intervention
Multiethnic	6	3	5.0%	2.6%	11	5	10.7%	4.4%
Age at referral	Mean		Standard deviation		Mean		Standard deviation	
	5.9	6.8	5.1	5.3	5.9	5.4	5.3	4.8

Table 3 Results from tests of differences in proportions and odds ratios: is the SFFC intervention superior to services as usual (SAU) within a year of assignment in downstate Illinois?

Phase		Deflection from foster care		No protective custody	
		Proportions		Proportions	
Formative	SFFC(I)	$y_{1I}/n_{1I}=99/115=0.861$		$y_{1I}/n_{1I}=108/115=0.939$	
	SAU(C)	$y_{1C}/n_{1C}=77/120=0.642$		$y_{1C}/n_{1C}=99/120=0.825$	
Summative	SFFC(I)	$y_{2I}/n_{2I}=81/113=0.717$		$y_{2I}/n_{2I}=104/113=0.920$	
	SAU(C)	$y_{2C}/n_{2C}=63/103=0.612$		$y_{2C}/n_{2C}=78/103=0.757$	
		Odds ratios		Odds ratios	
Pooled	OR, CI*	2.223	[1.278, ∞]	3.443	[1.639, ∞]
Formative	OR, CI*	3.455	[1.480, ∞]	3.273	[0.981, ∞]
Summative	OR, CI*	1.607	[0.776, ∞]	3.704	[1.456, ∞]
Pooled frequentist *p*-value		0.009		0.004	
Formative frequentist *p*-value		0.008		0.053	
Summative frequentist *p*-value		0.141		0.011	
Bayesian *p*-value		0.009		0.003	
Formative Freq. Prob(I > C)		0.992		0.937	
Summative Freq. Prob(I > C)		0.860		0.990	
Bayesian Prob(I > C)		0.992		0.997	

Table 3 (continued)

Phase		Permanence at 1 year		No recurrence of maltreatment	
		Proportions		Proportions	
Formative	SFFC(I)	$y_{1I}/n_{1I} = 102/115 = 0.887$		$y_{1I}/n_{1I} = 104/115 = 0.904$	
	SAU(C)	$y_{1C}/n_{1C} = 81/120 = 0.675$		$y_{1C}/n_{1C} = 109/120 = 0.908$	
Summative	SFFC(I)	$y_{2I}/n_{2I} = 84/113 = 0.743$		$y_{2I}/n_{2I} = 105/113 = 0.929$	
	SAU(C)	$y_{2C}/n_{2C} = 68/103 = 0.660$		$y_{2C}/n_{2C} = 94/103 = 0.913$	
		Odds ratios		Odds ratios	
Pooled	OR, CI*	2.199	[1.252, ∞]	1.084	[0.401, 2.930]
Formative	OR, CI*	3.778	[1.619, ∞]	0.954	[0.225, 4.048]
Summative	OR, CI*	1.491	[0.710, ∞]	1.257	[0.337, 4.688]
Pooled frequentist *p*-value		0.011		0.873	
Formative frequentist *p*-value		0.005		0.949	
Summative frequentist *p*-value		0.188		0.733	
Bayesian *p*-value		0.009		0.834	
Formative Freq. Prob(I > C)		0.995		0.525	
Summative Freq. Prob(I > C)		0.814		0.634	
Bayesian Prob(I > C)		0.991		0.583	

*1-sided 95% CIs except for no recurrence of maltreatment, which is a 2-sided 95% CI

Note: *SFFC* Safe Families for Children, *SAU* services as usual. The frequentist *p*-values for chi-square tests of the difference in proportions and the proportional hazard rate are adjusted for the clustering of children within family units

Deflection from Foster Care

Turning to the main findings from the analysis of DCFS administrative data, the first set of results examines the differences in proportions and odds ratios for all of the outcomes specified in the pre-registered analysis plan. Table 3 displays the proportions and odds of deflection from the date of allocation to treatment groups up to a year later. Drawing on an intent-to-treat framework, it shows that the proportions of children deflected from foster care were consistently higher among children assigned to the intervention group compared to the comparison group.

Focusing first on the odds ratio for deflection (the primary outcome), the parameter estimate for the pooled data indicates that children assigned to the intervention group were 2.223 times as large as the odds of deflection from foster care for the comparison group. Separating the study into its two phases, children assigned to the intervention group were 3.455 times as large as the odds of deflection for the comparison group during the formative evaluation period. The ratio fell to 1.607 times as large during the summative evaluation period. Unlike the frequentist *p*-values for the larger odds ratio estimated with the pooled and formative data, the *p*-value for the smaller odds ratio estimated during summative evaluation indicates that the odds of the SFFC's superiority is not distinguishable statistically from 1 (no difference) at the conventional 0.05 level.

Bayesian Analysis

As Gerber and Green (2012) note in their discussion of the Bayesian paradigm, one rather superficial way to interpret the change in significance levels from the formative to summative evaluation is the following: "The first study showed a statistically significant effect, but the second study…did not; because the initial results failed to replicate, the experimental effect cannot be considered robust" (Gerber & Green, 2012: 360). However, as they note, it is inefficient to ignore totally the strong effect observed during formative evaluation. Even though the summative experiment failed to produce a statistically significant estimate, this estimate is in the expected direction and is not a precisely estimated zero. Second, there is no good reason to ignore entirely the formative experiment when forming a judgement about the claim addressed by the summative experiment.

Because the program registered for summative evaluation is fundamentally the same as the program tested during formative evaluation, rather than simply pooling the data, the Bayesian paradigm offers a systematic way of accumulating evidence to provide a comprehensive summary of a program's effectiveness. When the formative results are updated in light of the new information from the summative evaluation, the *p*-value shows a highly significant value of 0.009. We also make use of a Monte-Carlo, simulation-based statistical model (Chen & Fraser, 2017) to assess the extent to which the program is superior to SAU. The results indicate that the likelihood that SFFC is truly superior, Bayesian Prob(I > C), in deflecting children from foster care is 99% certain.

Transition Rate Analysis

Table 4 displays the transition ratios observed during both the formative and summative evaluation periods for the three outcomes that could occur at any moment after assignment. Protective custody could be taken immediately upon assignment or up to a year later, after which the observation is censored at 366 days. Because the follow-up period for assessing family permanence is fixed at 1 year after assignment, it is omitted from Table 4.

The results in Table 4 show that the ratio of the daily probability of removal relative to the comparison group fell below the threshold of 0.5 that the pre-registered analysis plan set for assessing practical significance, but it averaged a little higher during summative evaluation. Controlling for the imbalances in the distribution of children's racial and ethnic background, indicated and sequence A reports does not alter the results to any appreciable degree. In particular, the effects of the Black and White binary indicators on all four outcomes of interest are not significant at conventional statistical levels. Controlling for the races of Black and White children, in fact, yields an estimated stronger intervention effect on the primary outcome of foster care removal. The same holds true for the secondary outcome of permanence at 1 year. The differences in the estimated effects on the other secondary outcomes are negligible whether or not controls for race/ethnicity are included in the proportional rate models.

In summary, the *p*-values calculated from both the pooled frequentist and Bayesian perspectives indicate that the transition ratio observed for foster care removal statistically is distinguishable from one (no difference) at a significance level of $p < 0.004$ for a one-tailed test. Separating the evaluation periods, the transition ratio is statistically significant at the 0.007 level in the formative evaluation but rises slightly higher than the conventional 0.05 level ($p < 0.077$). The second set of coefficients listed in the table are the logs of the transition ratio. They convey the same results as the transition ratios.

Protective Custody

Table 3 indicates that the intervention effect on the transition ratio for protective custodies was highly significant in both practical and statistical terms. The transition ratio was lower than the practical importance threshold of 0.5 and the pooled frequentist and Bayesian *p*-value were statistically significant at the 0.003 and 0.002 levels, respectively. Particularly noteworthy is that the intervention effect is stronger in the summative evaluation. The transition ratio is statistically significant at the 0.007 level in the summative evaluation but is slightly higher than the conventional level in the formative evaluation ($p < 0.055$). The same conclusion was supported by the proportions and odds ratios reported in Table 4. Again controlling for the other potential confounders did not appreciably alter the results.

Table 4 Results from tests of proportional transition ratios: is SFFC intervention superior to services as usual?

	Observed outcomes within a year of assignment									
	Foster care removal			Protective custody			Recurrence of maltreatment			
	Transition ratios			Transition ratios			Transition ratios			
Phase	$Exp(\lambda_{(t)})$	1-sided CI		$Exp(\lambda_{(t)})$	1-sided CI		$Exp(\lambda_{(t)})$	2-sided CI		
Pooled	0.478	[−∞, .755]		0.315	[−∞, .621]		0.924	[0.361, 2.368]		
Formative	0.335	[−∞, .691]		0.338	[−∞, 1.032]		1.057	[0.273, 4.093]		
Summative	0.602	[−∞,1.081]		0.297	[−∞, .667]		0.797	[0.229, 2.768]		
	Coefficients			Coefficients			Coefficients			
Phase	$\lambda_{(t)}$	SE	1-sided CI	$\lambda_{(t)}$	SE	1-sided CI	$\lambda_{(t)}$	SE	2-sided CI	
Pooled	−0.737	0.276	[−∞, .281]	−1.156	0.412	[−∞,−.476]	−0.079	0.477	[−1.020, 0.861]	
Formative	−1.093	0.437	[−∞,−.370]	−1.086	0.676	[−∞, .031]	0.056	0.686	[−1.298, 1.409]	
Summative	−0.508	0.354	[−∞, .078]	−1.212	0.489	[−∞,−.405]	−0.227	0.631	[−1.472, 1.018]	
Pooled frequentist *p*-value	0.004			0.003			0.869			
Formative frequentist *p*-value	0.007			0.055			0.935			
Summative frequentist *p*-value	0.077			0.007			0.720			
Bayesian *p*-value	0.004			0.002			0.834			
Formative Freq. Prob(I > C)	0.993			0.945			0.532			
Summative Freq. Prob(I > C)	0.924			0.993			0.641			
Bayesian Prob(I > C)	0.996			0.998			0.583			

Note: The frequentist *p*-values for *t* tests of the difference in proportional transition ratios are adjusted for the clustering of children within family units

Permanence of Care at 1 Year

Table 3 also shows the proportions and odds ratios for children who were consistently maintained in their own home or who at 1 year after allocation to treatment were back living with their families. The pooled data yield a statistical significance level of $p < 0.001$. However, the effect size is not equivalent across evaluation phases. During the formative evaluation period in downstate Illinois, the odds of permanence at 1 year were 3.778 times as large for children assigned to SFFC than children assigned to SAU. The permanency odds ratio observed during summative evaluation declined to 1.491 times as large. While the intervention effect was in the desired direction, the *p*-value was no longer statistically significant. Nonetheless, when the weaker effect is conditioned on the much stronger prior effect in the Bayesian model, the probability that SFFC is superior to SAU—Bayesian Prob(I > C)—exceeds 99%. In this model, the inclusion of controls for the other potential confounders yielded estimates of even stronger intervention effects.

Recurrence of Maltreatment

Table 3 displays the differences in the proportions of children assigned to each treatment group who did not experience a recurrence of maltreatment within a year of their prior indicated report. Both the proportions and associated odds ratios are approximately equivalent in spite of the fact that a significantly higher proportion of children assigned to the SFFC program were deflected from protective custody and foster care. The high *p*-values for the difference in proportions and the odds ratios of recurrence point to statistically similar recurrence rates in both treatment groups. The absence of significant differences in the transition ratios to recurrence reinforces this conclusion (see Table 4). The low probability that the intervention's effect is superior to SAU (0.583) means that the diversion of indicated victims of maltreatment to voluntary alternative care does not put them at any greater risk of repeat maltreatment than SAU.

The expectation prior to formative evaluation had been that SFFC referrals would reduce repeat maltreatment. By the time we pre-registered the summative evaluation; however, we thought a more realistic expectation was that SFFC exposes children to no greater danger than SAU. Consistent with concerns that hidden foster care could increase child safety issues for some children (Gupta-Kagan, 2020), our revised reasoning is that deflection of children from formal foster care could expose them to higher risks of repeat victimization than SAU by maintaining them in potentially abusive situations or by reunifying them too quickly before their families had been adequately helped. Unlike the other measures, we report the two-sided 95% CI because in light of concerns over the safety of children diverted to the hidden foster care system, there is no a priori reason to believe that one response is safer than the other is. Again, the inclusion of additional control variables did not alter the findings.

Discussion

The current study indicates that the SFFC program potentially provides benefits to children in the forms of deflection from the foster care system and increased family permanence. More importantly, it does not appear that the program prolongs children's absence from their family home and poses no greater risk to their safety compared to SAU. Together, the findings indicate that the intervention works well as designed, with no apparent significant risks to children compared to services as usual.

These outcomes are particularly welcome in Illinois, in which there is a deeply entrenched pattern of delaying the discharge of children from formal foster care to permanent homes. It is one of the systemic injustices that the federal *B.H.* consent decree (B.H. v. Smith, 1989) aims to remedy. The SFFC program was one of several initiatives that DCFS rolled out and evaluated under a supplemental implementation plan that the United States District Court for the Northern District of Illinois Eastern Division approved pursuant to the *B.H. v. Smith* Consent Decree. Despite repeated pledges by successive DCFS administrations to hasten the tempo of permanence, annual permanency rates have changed little over the past decade.

Comparative data show that Illinois registers the lowest rate of family reunification and the longest median duration of formal foster care among all 50 states and the District of Columbia. Data released by the U.S. Children's Bureau for 2017 showed that Illinois ranked lowest in the percentage of children discharged during their first year in foster care to the physical custody of birth parents, relatives, guardians, and adoptive parents—10.6%. The next lowest percentage was 24.9% for DC and the weighted average for the rest of the USA was 39.0% (USDHHS, 2020).

The low rate of family reunification in Illinois is partly related to its low rate of removal per 1000 children, which brings in the children from families that are the most challenging to reunify. Nonetheless, risk-adjustment for confounding factors over which states have limited control does not elevate Illinois's ranking on most indicators of child and family permanence. Detailed state-by-state performance indicators released by the U.S. Children's Bureau in 2020 show that Illinois's rank as the state with the lowest percentage of children reunified within a year of removal remained unchanged after making risk adjustments for the state's foster care entry rate and age at entry (U.S. Children's Bureau., 2020). The same holds true for permanence in 12 months for children in foster care for 12 to 24 months.

Not surprisingly given this context, of the 132 children in the SFFC formative evaluation sample who were taken into foster care for 8 days or more, only nine (7%) were either back home ($n=7$) or discharged to private kinship care ($n=2$) within 12 months. There were no significant differences between the intervention and comparison group in duration of formal foster care. Most of the children in the intervention group, who were removed ($>80\%$), were still in foster care 18 months after enrollment. The insight from the equivalence in duration of formal foster care is that SFFC succeeded in reducing foster care not by shortening the time children stay in formal foster care but by preventing their entry into care in the first place.

Skeptics of the expanded use of informal placements with kin or non-kin raise the objection that SFFC still involves removal of children from their families of origin. How

is this any different from formal removal? It was precisely in anticipation of this objection that we included the secondary outcomes of family permanence at 1 year and recurrence of maltreatment. The findings from the pooled-data and Bayesian analyses indicate that the odds of living with parents at 1 year were significantly larger for children assigned to SFFC than children assigned to SAU. If the formative results had instead shown no difference or lower odds of permanence at 1 year for the SFFC group, it is doubtful the Department would have proceeded with a summative evaluation of null findings. Instead, it would have concluded that SFFC afforded children no greater safety and family permanence than business as usual. However, the formative results suggested otherwise. Even though the program still temporarily removes children from their families of origin, the positive effects on permanence at 1 year indicate that temporary removal through SFFC may be superior to formal removal as usual.

Differential Timing

There was a spike in child removals in downstate Illinois during the period that the summative evaluation was being fielded, which interrupted a stable trend line that extended back as far as 2000. At the statewide level, Illinois's rate of removing children from parental custody rose from one of the lowest in the nation at 1.52 per 1000 children in 2016 to 2.82 per 1000 children in 2020. The rise put the state's removal rate on a par with the rest of the nation during the height of the COVID-19 lockdown. The spike in removals resulted in Illinois being the only state to experience double-digit percentage growth (12%) during the pandemic (USDHHS, 2022). The growth in the size of the *B.H.* plaintiff class strained all parts of the DCFS system. Therefore, it is plausible that removal practices became more risk-adverse during this period.

The absence of a state budget during fiscal years 2016, 2017, and part of 2018 and the accompanying freeze on public spending deprived community-based agencies of the public resources that helped keep struggling families intact despite social, economic, and public health hardships (Durbin, 2021). In addition, a series of high-profile child murders in the state between 2017 and 2019 heightened public concerns, as recounted in the opinion pages of the *Chicago Tribune* that DCFS "too often leans toward keeping families together, even in crisis."[2] Case level factors, especially the opioid crisis, likely also have contributed to the rising intake rates. Between 2013 and 2018, both fatal and non-fatal opioid overdose rates more than doubled.[3] Together, these macro factors may have influenced the context in which investigators had to operate.

As shown in Table 3, the loss of statistical significance of the odds of permanence at 1 year was due entirely to changes in the odds of removal after the children had been deflected from protective custody. The effect sizes for the secondary outcome of no protective custody are statistically significant and equally as large ($OR > 3.0$) in both phases of the evaluation. What appears to have changed are removal

[2] (Chicago Tribune, Nov. 1, 2019).

[3] The Illinois Department of Public Health Opioid Data Dashboard (http://idph.illinois.gov/opioiddatadashboard/) shows that fatal overdoses per 10,000 population increased from .88 in 2013 to 2.28 in 2018. Non-fatal overdoses increased during that time frame from 4.55 to 10.73.

practices at the later stages of child protective surveillance but only for children and families allocated to the SFFC program. In the comparison group, the proportion of children deflected from foster care was approximately the same in both phases (≈ 61% and 64%, respectively). In contrast, the proportion of deflections in the SFFC group declined from 86 to 72%, which diminishes the effect size estimated in the summative evaluation and accounts for the loss of statistical significance. Further, the change in the SFFC effect on deflections carries over to the odds of permanence at 1 year. Whereas the proportion of children residing with their parents or other family members at 1 year after allocation remains nearly identical in the comparison group (≈ 68% and 66%, respectively), the proportion in the intervention group declines from 88.7% in the formative evaluation to 74.3% in the summative evaluation. While still in the desired direction, the difference in proportions between the comparison and intervention groups in the summative evaluation declines to 14.4 percentage points, which is no longer statistical distinguishable from zero.

Another possible explanation for the weakened SFFC effects on deflection and permanence at 1 year may be that there are unmeasured differences in implementation between the two phases. Nevertheless, given the range of analytic methods we employ and the qualitative consistency of findings across these methods, we believe that the results are not significantly biased.

In sum, our findings indicate potential positive impacts of SFFC. Although we cannot test potential pathways, it is possible that the program truly works as intended: tapping into a reservoir of bridging social capital, which supplies an alternative to formal foster care and sustains safe, temporary living arrangements for children, which are key elements for the decentering of formal foster care.

Limitations

This study has a number of limitations. Administrative data restrictions did not permit us to assess the specific mechanisms through which the intervention could work. Nonetheless, the "black box" results indicate that the configuration of voluntary care, active social and material support, and continued engagement after the child returns home is most effective, the program demonstrates success in reducing the least desirable child welfare outcomes. The statistical analyses suggest, however, that these potential benefits may be sensitive to variations in settings and timing. In spite of the apparent weakening of the effect sizes due perhaps to placement environments growing more risk adverse, the net benefits still compare favorably to formal foster care.

A number of protocol deviations from the logic model hindered our ability to test the external validity of the formative findings as originally planned. The least consequential deviation involved direct referrals from intact family workers, which were not part of the original study protocol. Intact family workers served families in which the assessment of risk to the children was not severe enough to warrant removal of the children at the time of case opening. Because these cases were not subject to the DCFS alternation routine, they should not have been randomized by SFFC. A more serious deviation concerned the sizable uptick in direct referrals from

CPS investigators, which by-passed the automated allocation routine. By the end of the formative evaluation period, the percentage of crossovers from the comparison to treatment in Cook County had already surpassed the 10% threshold, which necessitated excluding Cook County observations from the summative evaluation. Other minor deviations that required dropping additional cases from the allocated samples are itemized in Fig. 1 CONSORT diagrams.

The extent to which the results are generalizable to other geographic regions and populations outside downstate Illinois is unknown. Given that over 90% of SFFC hosts in Illinois are White, SFFC professionals in Illinois and nationally have noted they have faced concerns from minority professionals about implementing SFFC in urban predominately minority communities. Our corollary finding that children's racial or ethnic origin was not disproportionately associated with adverse outcomes in any of the four areas of safety and permanence tracked in the study should not be misinterpreted as showing that systemic racism and discrimination are unimportant concerns (Baron et al., 2022). Given the history of racism in our country and the considerable racial disparities in the chances of involvement in the child welfare system, it is essential to acknowledge and study issues related to race and religion in the implementation of SFFC and assessment of its potential disparate impact.

Conclusion

The accumulated evidence gathered from this investigation shows that SFFC had a positive impact on the primary outcome of deflecting children from foster care and on the secondary outcome of living with parents or relatives without state oversight at 12 months after enrollment. The chances of removal were significantly lower for children receiving SFFC compared to services as usual. There were also no concerning differences in the recurrence of maltreatment between the two treatment groups.

Future Directions

The Bayesian paradigm utilized in this evaluation supplies a conceptual framework for interpreting research findings and setting future goals for the accumulation of sufficient empirical information to enable decision makers with different priors to reach evidence-informed posterior judgements about the net benefits of deflection programs like SFFC. To gain a better understanding for how prior beliefs are transformed into posterior beliefs using the Bayesian paradigm, it is helpful to compare the following limiting cases. Among staunch critics of the hidden foster care system, their prior is that the true effect of SFFC on positive child welfare outcomes is negative or zero. Updating their priors with evidence may require much larger estimates of effect sizes and statistical significance than what this study alone presently offers. Conversely, both the undecided and true believers in the value of deflecting children from formal foster care may more readily accept the positive results of this evaluation as supportive of the posterior belief that the SFFC model can be an effective and less costly alternative to services as usual.

Even though findings from the current study are supportive of continued use of SFFC, future research should endeavor to follow families for a longer follow-up period, assess implementation fidelity (e.g., how often do SFFC host families provide ongoing support to the family after children return home?), and undertake qualitative research to better understand the specific aspects of the SFFC interventions that families felt were the most important to their success. Similarly, widespread adoption of the SFFC model should be undertaken with deliberation given to interests in recruiting a diverse and well-trained network of host families. The nature, extent, and impact of racial equity concerns about SFFC from minority professionals and other child welfare professionals should be examined in future implementation evaluations. These concerns need to be heard and addressed openly on an ongoing basis. Further, more work is needed to understand how well the program's results generalize from downstate Illinois to other areas and other time periods.

Although we find positive effects on the targeted child welfare outcomes, the current study cannot shed light on the potential mechanisms. It may be that temporary hosting works by connecting families to networks of volunteer temporary hosts and social networks. However, it may also be that the program provides parents with temporary respite, allowing them the time and flexibility to obtain safe housing, attend substance use treatment, or find employment. Additional research is required to better understand the pathways through which the intervention works.

Data Availability Data for this study come from administrative records from the Illinois Department of Children and Family Services and are restricted access.

Declarations

Competing Interests The authors declare no competing interests.

References

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., & Lang, T. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine, 134*(8), 663–694.

Autor, D., Cho, D., Crane, L. D., Goldar, M., Lutz, B., Montes, J., Peterman, W. B., Ratner, D., Villar, D., & Yildirmaz, A. (2022). An evaluation of the Paycheck Protection Program using administrative payroll microdata. *Journal of Public Economics, 211*, 104664.

B.H. v. Smith 5 F. Supp. 1387 (N.D. Ill. 1989), 128 F.R.D. 659 (N.D. Ill. Dec. 19, 1989); 715 F. Supp. 1387 (N.D. Ill. May 30, l989), 984 F. 2d 196 (7th Cir. 1993), 49 F.3d 294 (7th Cir. 1995), reh'g and reh'g en banc denied, (Apr. 7, 1995).

Baron, E. J., Goldstein, G. Ee., & Ryan, J. (2022). The push for racial equity in child welfare: Can blind removals reduce disproportionality. *Journal of Policy Analysis and Management.*

Chalmers, I. (2011). Why the 1948 MRC trial of streptomycin used treatment allocation based on random numbers. *Journal of the Royal Society of Medicine, 104*(9), 383–386.

Chen, D. G., & Fraser, M. (2017). A Bayesian approach to sample size estimation and the decision to continue program development in intervention research. *Journal of the Society for Social Work and Research, 8*(3), 457–470.

Chen, D.-G., Testa, M. F., Ansong, D., & Brevard, K. C. (2020). Evidence building and information accumulation: Using the Bayesian paradigm to advance child welfare intervention research. *Journal of the Society for Social Work and Research, 11*(3), 483–498.
Coleman, J. S. (1982). *The asymmetric society*. Syracuse University Press.
DePanfilis, D., & Zuravin, S. J. (1998). Rates, patterns, and frequency of child maltreatment recurrences among families know to CPS. *Child Maltreatment, 3*(1), 27–42.
Dettlaff, A. J., Weber, K., Mendleton, M., Boyd, R., Bettencourt, B., & Burton, L. (2020). It is not a broken system, it is a system that needs to be broken; the upend movement to abolish the child welfare system. *Journal of Public Child Welfare, 14*(5), 500–517.
Fleming, J. B. (2019). Spiritual generosity: Biblical hospitality in the story of Lydia (Acts 16:14–16, 40). *Missiology, 47*(1), 51–63.
Fluke, J., Edwards, M., Bussey, M., Wells, S., & Johnson, W. (2001). Reducing recurrence in child protective services: Impact of a targeted safety protocol. *Child Maltreatment, 6*(3), 207–218.
Font, S. A., & Gershoff, E. (2020). *Foster care and "best interests of the child": Integrating research, policy, and practice*. Springer.
Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. W.W. Norton & Company.
Gupta-Kagan, J. (2020). America's hidden foster care system. *Stanford Law Review, 72*(841), 841–913.
Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945–960.
Karlan, D., Savonitto, B., Thuysbaert, B., & Udry, C. (2017). Impact of saving groups in the lives of the poor. *Proceedings of the National Academy of Sciences, 114*(12), 3079–3084.
Leventhal, T., & Brookss-Gunn, J. (2003). Children and youth in neighborhoods contexts. *Current Directions in Psychological Sciences, 12*(1), 27–31.
List, J. A. (2022). *The voltage effect: How to make good ideas great and great ideas scale*. Random House.
Meitner, H., & Albers, M. (2012). *Introducing safety-organized practice*. National Council on Crime and Delinquency, Children's Research Center.
Radel, L., Bramlett, M., Chow, K., & Waters, A. (2016). Children living apart from their parents: Highlights from the National Survey of Children in Nonparental Care. *U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.*
Redleaf, D. L. (2018). *They took the kids last night: How the child protection system puts families at risk* (Davis Library). Santa Barbara, California : Praeger, an imprint of ABC-CLIO, LLC, [2018]. https://catalog.lib.unc.edu/catalog/UNCb9406645
Rubin, D. M., O'Reilly, A. L., Luan, X., & Localio, A. R. (2007). The impact of placement stability on behavioral well-being for children in foster care. *Pediatrics, 119*(2), 336–344.
Shadish, W. R., Cook, TDd., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton.
Slack, K. S., Holl, J. L., Lee, B. J., McDaniel, M., Altenbernd, L., & Stevens, A. B. (2003). Child protective intervention in the context of welfare reform: The effects of work and welfare on maltreatment reports. *Journal of Policy Analysis and Management, 22*(4), 517–536.
Testa, M., Bruhn, C. & Helton, J. (2010). Comparative safety, stability, and continuity of children's placements in informal and formal substitute care arrangements. In M.B. Webb, K. Dowd, B.J. Harden, J. Landsverk & M.F.Testa, (Eds.). *Child welfare and child well-being: New perspectives from the national survey of child and adolescent well-being* (pp. 159–191). Oxford University Press.
U.S. Census Bureau. (2022). Historical living arrangements of children. https://www.census.gov/data/tables/time-series/demo/families/children.html
U.S Department of Health and Human Services (2006). The AFCARS Report FY 1998 through FY 2002. https://www.acf.hhs.gov/sites/default/files/documents/cb/afcarsreport12.pdf
U.S. Children's Bureau. (2020). Child Welfare Outcomes, 2018.
U.S. Department of Health and Human Services. (2021a). The AFCARS Report FY 2019. https://www.acf.hhs.gov/sites/default/files/documents/cb/afcarsreport27.pdf
U.S. Department of Health and Human Services. (2021b). CFSR Round 3 Statewide Data Indicators Workbook. https://www.acf.hhs.gov/cb/report/cfsr-round-3-statewide-data-indicators-workbook
Vreeland, A., Ebert, J. S., Kuhn, T. M., Gracey, K. A., Shaffer, A. M., Watson, K. H., Gruhn, M. A., Henry, L., Dickey, L., Siciliano, R. E., & Anderson, A. (2020). Predictors of placement disruptions in foster care. *Child Abuse & Neglect, 99*, 104283.

Wu, Q. (2017). The relationship between kinship diversion and child behavior problems. *Child Welfare, 95*(3), 61–85.

Yi, Y., Edwards, F. R., & Wildeman, C. (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.