Engaging young mothers as partners in case management

THE IMPACTS OF CALIFORNIA'S ADOLESCENT FAMILY LIFE PROGRAM WITH POSITIVE YOUTH DEVELOPMENT AFTER ONE YEAR

April 2023

HHS Office of Population Affairs
Web: opa.hhs.gov | Email: opa@hhs.gov
Twitter: @HHSPopAffairs | YouTube: HHSOfficeofPopulationAffairs

Purpose statement

This report reports interim findings from the Adolescent Family Life Program with Positive Youth Development (AFLP-AFLP-PYD) program, an enhanced case management program for young parents operated by the state of California's Department of Public Health, Maternal, Child, and Adolescent Health division (MCAH). AFLP-AFLP-PYD is a one-year, highly structured program for adolescent mothers that incorporates intentional life planning, prescribed content on key topics, and motivational interviewing techniques during twice monthly visits. MCAH received a Pregnancy Assistance Fund (PAF) grant from the Office of Population Affairs (OPA) in 2011 to develop the program, and then a second grant in 2013 to expand the program across the state. Under contract with Mathematica, OPA used the planned expansion of the program to conduct a rigorous study of AFLP-AFLP-PYD program impacts. This report provides interim impacts of AFLP-AFLP-PYD about a year after the adolescent mothers had enrolled in the program. A future brief will examine the program's longer-term impacts.

Authors

Susan Zief, Mathematica

Andrew Langan, Mathematica

Jennifer Walzer, Mathematica

Emily Weaver, Mathematica

Lisa Dragoset, Mathematica

John Deke, Mathematica

Submitted to

Dr. Amy Farb, Evaluation Specialist, Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services

This report is in the public domain. Permission to reproduce is not necessary.

Suggested citation: Zief, S., A. Langan, J. Walzer, E. Weaver, L. Dragoset, and Deke, J. "Engaging Young Mothers as Partners in Case Management: The Impacts of California's Adolescent Family Life Program with Positive Youth Development After One Year." Washington, DC: Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2023.

Prepared for OPA under contract number:
H HSP233201450026A

Disclaimer

This publication was supported by Award No. HHSP233201450026A from the Office of Population Affairs (OPA). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of OPA or HHS.

Acknowledgements

Many organizations and people contributed in significant ways to this evaluation of the Adolescent Family Life Program with Positive Youth Development program. First, we acknowledge the invaluable support and guidance from the Office of Population Affairs (OPA) in the U.S. Department of Health and Human Services, and especially Amy Farb, the OPA project officer for the Positive Adolescent Futures study. We also thank Evelyn Kappeler, Alexandra Warner, and Callie Koesters for their support and assistance throughout this evaluation effort.

We are grateful to the invaluable assistance and partnership provided by Lissa Pressfield, Mary Campa and Sangi Kabadi at MCAH and Kathleen Tebb and University of California-San Francisco. Without their support and unwavering commitment to independent and objective evaluation, this study would not have been possible. Our thanks also to the study site administrators and their teams of dedicated case managers and supervisors who volunteered to be a part of the study and collaborated with us on all study enrollment and data collection procedures. The remarkable young mothers who work tirelessly as parents, and who also volunteered their time and energy to take the study surveys, also deserve our tremendous gratitude.

At Mathematica, we thank Ava Madoff, Laura Kalb, Kai Filipczak, Charles Tilley, and Juha Sohlberg who contributed their survey, programming, and data analysis expertise. Staff at Decision Information Resources, Inc. were invaluable to our follow-up data collection success. We are also grateful to Jean Knab, who provided helpful comments on our report. Brigitte Tran created the report graphics, Carol Soble edited the final report, and Sharon Clark produced it.

Susan Zief
Andrew Langan
Jennifer Walzer
Emily Weaver
Lisa Dragoset
John Deke

Contents

Purpose statement ii

Acknowledgements iii

I. Introduction 1

II. The Positive Youth Development Program 4

A. AFLP is an important resource for young parents and their families statewide 4

B. MCAH created a new version of AFLP called AFLP-PYD to promote better outcomes for young parents 4

C. AFLP-PYD is a shorter, more intensive program than AFLP and uses motivational interviewing 5

D. By building social competence and autonomy, AFLP-PYD aims ultimately to improve health and educational outcomes for young parents 6

III. Assessing Program Impacts: Design, Data Sources, and Analysis Methods 8

A. The AFLP-PYD impact evaluation includes two substudies, each with a distinct design 8

B. The two substudies differed with respect to the timing of random assignment relative to when individuals entered the program 9

C. To measure the impact of AFLP-PYD relative to AFLP, the evaluation team administered three surveys to the mothers in both groups 10

D. This report examines program impacts after one year of program enrollment, on outcomes that align with the AFLP-PYD logic model 11

1. Measures of program delivery: Exposure to information 11

2. Primary outcomes 12

3. Exploratory outcomes 13

E. Baseline characteristics of AFLP and AFLP-PYD study participants 13

F. There were some minor differences between the AFLP-PYD and AFLP groups at baseline 14

G. To estimate the impact of the new AFLP-PYD program, we calculated the difference between outcomes for AFLP-PYD and AFLP participants 15

IV. Impact Findings at 12-months 17

A. Measures of program delivery: Exposure to information 17

1. Mothers in the AFLP-PYD group received more home visits than mothers in the AFLP group, yet left the program in the first year at higher rates 17

2. AFLP-PYD decreased the likelihood that participants reported receiving information and services through their case manager 17

B. Primary outcomes 20

1. Mothers in the AFLP-PYD group exhibited the same amount of resiliency as mothers in the AFLP group 20
2. Mothers in AFLP-PYD received fewer linkages and referrals than mothers in AFLP, but reported greater access to a supportive adult 20
3. Mothers in AFLP-PYD were more likely than AFLP mothers to have engaged in unprotected sex in the past three months, but equally likely to be currently using birth control 20
4. Mothers in AFLP-PYD group had less knowledge about IUDs and slightly less knowledge of other methods of birth control methods than mothers in the AFLP group 21
5. Mothers in AFLP-PYD reported similar levels of depressive symptoms and health insurance as AFLP mothers, but more medical checkups 21
6. Mothers in AFLP-PYD were less likely than mothers in AFLP to be enrolled in high school or have a diploma or GED, and to have positive attitudes toward healthy relationships 21

C. Exploratory outcomes 23

1. Mothers in AFLP-PYD were less likely than mothers in AFLP be enrolled in a postsecondary institution, but equally likely to be employed and to have delayed a subsequent pregnancy 23
2. Children of AFLP-PYD mothers had fewer well-visits with a doctor and were less likely to have health insurance. 23

D. Results are robust to alternative analysis methods 23

V. Conclusion 25

Appendix A. Additional Details on Data Sources and Analysis Methods 27

Appendix B. Supplemental Analyses 46

References 65

Tables

II.1. A comparison of key features of AFLP and AFLP-PYD 5

III.1. Sample sizes and study consent rates, by substudy 10

IV.1. AFLP-PYD impacts on measures of program delivery at 12 months 19

IV.2. AFLP-PYD impacts on short-term outcomes and early indicators of long-term outcomes at 12 months 22

IV.3. AFLP-PYD impacts on exploratory outcomes at 12 months 24

A.1. Site-level random assignment sites 28

A.2. Survey response rates, by treatment group and survey mode at 12 months 31

A.3. Participant characteristics at study enrollment for 12 month follow-up survey respondents and non-respondents, by program group 31

A.4. Measures of program delivery 34

A.5. Short-term outcomes: resiliency and access to supportive adult items 36

A.6. Contraceptive knowledge measures, with correct answers shown 39

A.7. Exploratory outcomes 42

B.1. Baseline Equivalence for Overall Study 46

B.2. Baseline Equivalence for Individual-Level Study Participants 48

B.3. Baseline Equivalence for Site-Level Study Participants 49

B.4. Sensitivity Analyses Conducted 52

B.5. Results using alternate analysis methods 53

B.6. Results using site-level data for the site-level study 55

B.7. Results from additional sensitivity analyses 58

B.8. Impacts on program delivery measures for early versus later enrollers in the individual substudy 61

B.9. Sensitivity to alternative Bayesian interpretations 63

Figures

I.1. The federal evaluation of AFLP and AFLP-PYD in California 2

I.2. Participating sites 3

II.1. AFLP-PYD logic model 7

I. Introduction

Young women who give birth as adolescents confront obstacles on their path to self-sufficiency that are not faced by their non-parenting peers. They can struggle to finish their education and find a stable, well-paying job, potentially putting themselves at a greater risk of being socio-economically disadvantaged and relying on public assistance as adults (Maynard and Hoffman 2008; Perper, Peterson, and Manlove et al. 2010; Diaz and Field 2016; Harding et al. 2020). It is not surprising, therefore, that research shows that adolescent mothers face a higher risk of poor maternal health outcomes, such as postpartum depression, and that their children are more likely to have poor health, behavioral, and educational outcomes compared to children born to older mothers (Barnet et al. 1996; Maynard and Hoffman 2008).

A repeat pregnancy during adolescence compounds the risk of poor outcomes for both mother and child. Adolescent mothers who have rapid repeat pregnancies—that is, a pregnancy within 24 months of the previous pregnancy—are at significantly greater risk of a stillbirth or preterm birth than are mothers who delay a subsequent pregnancy (Conde-Agudelo et al. 2006). They are also less likely to stay in or complete high school, to work, or to maintain economic self-sufficiency, and their children are less likely to be prepared for school (Klerman 2004). Repeat pregnancy is common among adolescent mothers; about one in six adolescent mothers go on to have another child before the age of 20 (Dee et al. 2017).

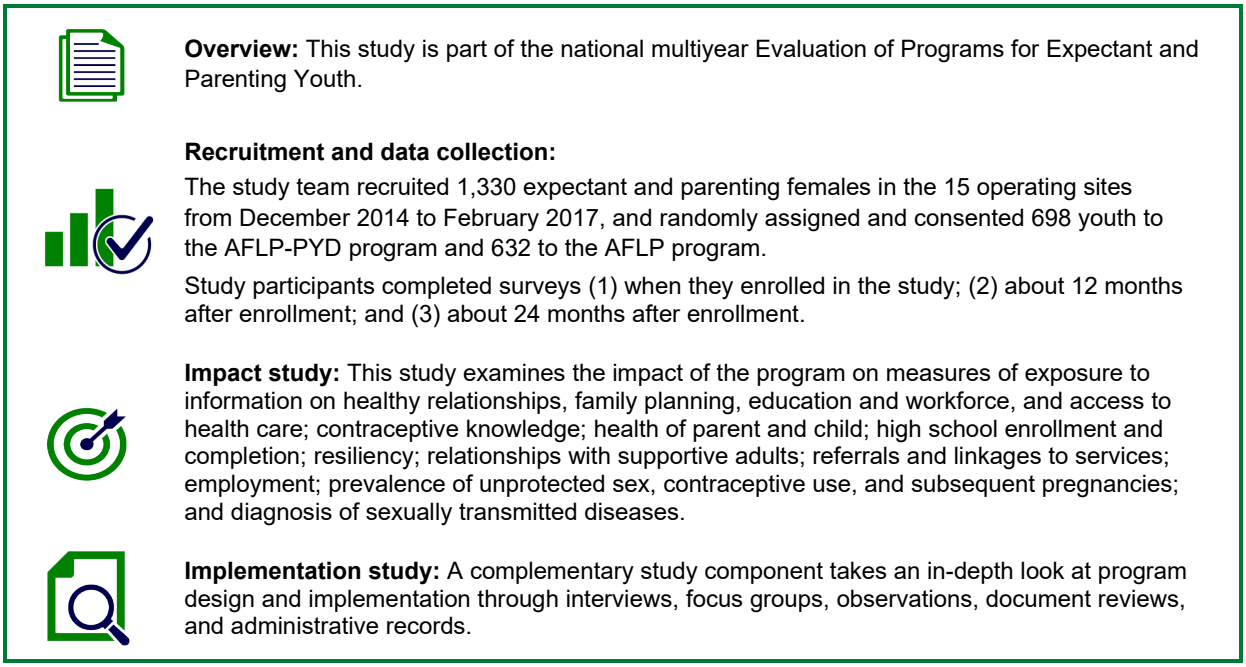
A small but growing body of evidence suggests that interventions for adolescent mothers can promote healthy birth spacing by providing a combination of individualized support services and improved access to effective contraception. For example, a randomized control trial of the Pathways Teen Mother Support Project in South Carolina—a program that included case management with a focus on family planning—found a lower rate of repeat pregnancy in treatment youth after two years in the program (McDonell, Limber, and Connor-Godbey 2007). More recently, a randomized controlled trial of the Teen Options to Prevent Pregnancy program for low-income adolescent mothers found that the program reduced rates of repeat pregnancy through a combination of one-on-one motivational interviewing sessions and facilitated access to contraceptive services (Rotz et al. 2016). In addition, in a recent randomized control trial of Steps to Success, a multicomponent home-visiting program with a focus on healthy birth spacing, in San Angelo, Texas, after one year in the program adolescent mothers were more likely than mothers in the control group to report using a long-acting reversible contraceptive method. Younger adolescent mothers in the program also reported a decreased incidence of unprotected sex compared to their counterparts in the control group (Rotz and Wood 2018).

In 2010, the Office of Population Affairs (OPA) in the Department of Health and Human Services (HHS), launched the Pregnancy Assistance Fund (PAF) to explicitly address the needs of adolescent parents. This unique program helps expectant and parenting adolescents improve their immediate outcomes, such as access to health care and education. Improvements in these outcomes are hypothesized to delay a subsequent pregnancy and improve the long-term well-being of adolescent mothers and their children.

With funding from the PAF, the state of California redesigned its existing Adolescent Family Life Program (AFLP) for expectant and parenting adolescents into a more intensive and structured intervention called AFLP with Positive Youth Development (AFLP-AFLP-PYD). The AFLP is a case management program for expectant and parenting adolescents developed by the California Department of Public Health, Center for Family Health, Maternal, Child and Adolescent Health (MCAH) Division. AFLP began in 1985 and primarily serves families with low incomes and high needs. The program is designed to help young parents delay repeat pregnancies, complete high school, improve their own health and the health of their children, and improve their support networks.

The recent enhanced version of the program—called AFLP-AFLP-PYD—began as a pilot in 2012. AFLP-PYD is intended to differ from the original AFLP by (1) using a prescribed set of structured activities and content to help young parents identify their strengths and use them to meet their goals and (2) requiring case managers to conduct two visits a month instead of one.

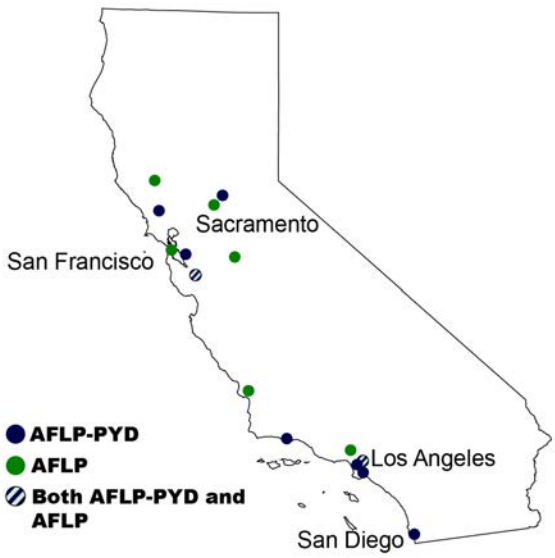
Figure I.1. The federal evaluation of AFLP and AFLP-PYD in California



MCAH agreed to participate in the Federal Evaluation of Selected Programs for Expectant and Parenting Youth (PEPY), funded by OPA and conducted by Mathematica. The evaluation was designed to examine the effectiveness of AFLP-PYD in influencing key outcomes, compared with the original program model, or business as usual (AFLP) (Figure I.1).

Working closely with MCAH, Mathematica recruited 13 agencies (operating in 15 locations or sites) to participate in the study.[1] Implementing agencies were geographically dispersed across the state and varied in terms of size, reach, and populations served (Figure I.2). In total, the evaluation team randomly assigned and enrolled 1,330 expectant or parenting adolescent mothers in the study, of whom 698 received AFLP-PYD and 632 received AFLP, or business as usual.

Figure I.2. Participating sites



This report is the second in a series on the implementation and impacts of AFLP-PYD. An earlier report described the implementation of both AFLP-PYD and its predecessor, AFLP (Asheer et al. 2020). The present report adds to these findings by examining the impact of AFLP-PYD compared to the business-as-usual condition, AFLP, beginning when the program was expanded throughout the state in 2014 through April 2018. It considers a variety of outcomes, measured about one year after participants began receiving program services. This timing coincides when most AFLP-PYD youth would have been exiting the program, whereas AFLP youth could have continued in the program for a total of 24 months. Outcomes examined in this report include the impact of the program on measures of exposure to information on healthy relationships, family planning, education and workforce, and access to health care; contraceptive knowledge; health of parent and child; high school enrollment and completion; resiliency; relationship with supportive adults; referrals and linkages to services; employment; prevalence of unprotected sex, contraceptive use, and subsequent pregnancies; and diagnosis of sexually transmitted diseases.

We have organized this report as follows. In Chapter II, we describe the AFLP-PYD program. In Chapter III, we discuss the impact study design, data sources, and analysis methods. In Chapter IV we present the findings on outcomes measured roughly 12 months after program enrollment. The report ends with a brief conclusion in Chapter V.

[1] Through a statewide RFP process, MCAH ranked agencies delivering AFLP based on the level of risk in their target communities and their need for adolescent sexual and reproductive health services. MCAH developed and used the California Adolescent Sexual Health Needs Index (CASHNI) to determine each applicant community's level of risk and need for services at the Medical Service Study Area (MSSA) level. This index allowed MCAH and others to focus available resources for primary and secondary adolescent pregnancy prevention programs on areas across the state with the greatest need for sexual and reproductive health services and supports. The identified sites for possible evaluation inclusion that were the highest-need sites that had not already participated in the AFLP-PYD pilot, and selected those that could provide the greatest number of sample members.

II. The Positive Youth Development Program

In this chapter, we briefly describe the Adolescent Family Life Program, which serves as the business-as-usual approach for the impact evaluation. We then describe the Positive Youth Development program that was introduced statewide as part of the rigorous impact evaluation and discuss how it differs from the AFLP model.

A. AFLP is an important resource for young parents and their families statewide

The California Maternal, Child, and Adolescent Health Division has administered AFLP, a case management program for expectant and parenting adolescents, since 1985. The program serves areas of California with the highest adolescent birth rates. A variety of community-based organizations, local health care agencies, and school districts across the state implement the program. AFLP was designed to support expectant and parenting adolescents through monthly case management visits over a two-year period. During their visits, case managers cover five key topics: (1) health, (2) nutrition, (3) education, (4) parenting, and (5) psychosocial skills. Although AFLP has no explicit or formal structured curriculum, it addresses several core objectives: delaying repeat pregnancies, completing high school, improving the health of the parent and child, and improving linkages and networks of support for expectant and parenting adolescents. The program enrolls adolescents who are younger than age 19 and are expecting or parenting; participants are primarily female, but male adolescents are also eligible. AFLP does not set specific income requirements, but it serves primarily low-income, high-need families. In addition, it does not place any restrictions on the number or age of the child(ren). Participants are no longer eligible for the program at age 19 (age 20 with an exception) or after 24 months of service.

B. MCAH created a new version of AFLP called AFLP-PYD to promote better outcomes for young parents

Using Pregnancy Assistance Fund program funding in 2010, MCAH launched an effort to redefine and standardize AFLP across the state to promote better outcomes for young parents. Drawing on evidence-informed positive youth development principles (Catalano et al. 2004; Lerner and Lerner 2009; Gloppen et al. 2010) and a framework that emphasizes youth resiliency and independence, the new AFLP-PYD program (1) prescribed a set of structured activities and content to help youth identify their strengths and use them to meet their goals and (2) required case managers to conduct two visits a month, instead of one a month, over a 12-month period. To accommodate the increased intensity of visits and to allow case managers more time with each client, case managers' maximum expected caseloads are about half of those expected for AFLP case managers: AFLP-PYD case managers are expected to serve about 20 to 25 mothers instead of the up to 40 allowable under AFLP. Critical elements of the AFLP-PYD approach include building competence for independent decision making, developing adolescents' confidence

through skill building, identifying and using adolescents' strengths and values to set and meet specific goals, and encouraging self-care and self-advocacy.

C. AFLP-PYD is a shorter, more intensive program than AFLP and uses motivational interviewing

As designed, AFLP-PYD presents a clear contrast with the AFLP program in terms of approach, methods, and structure (Table II.1). AFLP-PYD is a shorter but more intensive program than AFLP. AFLP-PYD case managers meet face-to-face with clients twice a month, for one year, with caseloads of 20 to 25 youth, compared to monthly visits for two years and caseloads of 40 youth for AFLP. AFLP-PYD requires case managers to use motivational interviewing to guide participating youth through a prescribed set of activities designed to set achievable goals for life planning and building self-sufficiency. Motivational interviewing relies on a collaborative conversation between the case manager and youth that empowers youth to identify their own reasons for behavior change. The content and activities follow a strengths-based approach to encourage the young parents to define and build on their strengths and successes to achieve the goals they set for themselves. MCAH requires AFLP-PYD case managers to complete MCAH-led trainings on the program and monthly technical assistance calls, while AFLP case managers complete supervisor-led, site-based trainings that can vary in intensity and methods based on a site's requirements.

Table II.1. A comparison of key features of AFLP and AFLP-PYD

	AFLP	AFLP-PYD
Length of program	24 months[a]	12 months[b]
Frequency of visits	Once a month	Twice a month
Maximum caseload	40 young parents	20–25 young parents
Structure	Loosely structured, with few required activities	Highly structured, with specific sequence of activities in four program phases
Methods	None required; depends on agency and case managers	Use of motivational interviewing and strengths-based approach
Required content topics	Health, nutrition, education, parenting, and psychosocial skills	Healthy relationships, family planning, education and workforce, and access to health care
Case manager training	Loosely structured, site-specific training from supervisor	Intensive, highly structured training led by MCAH

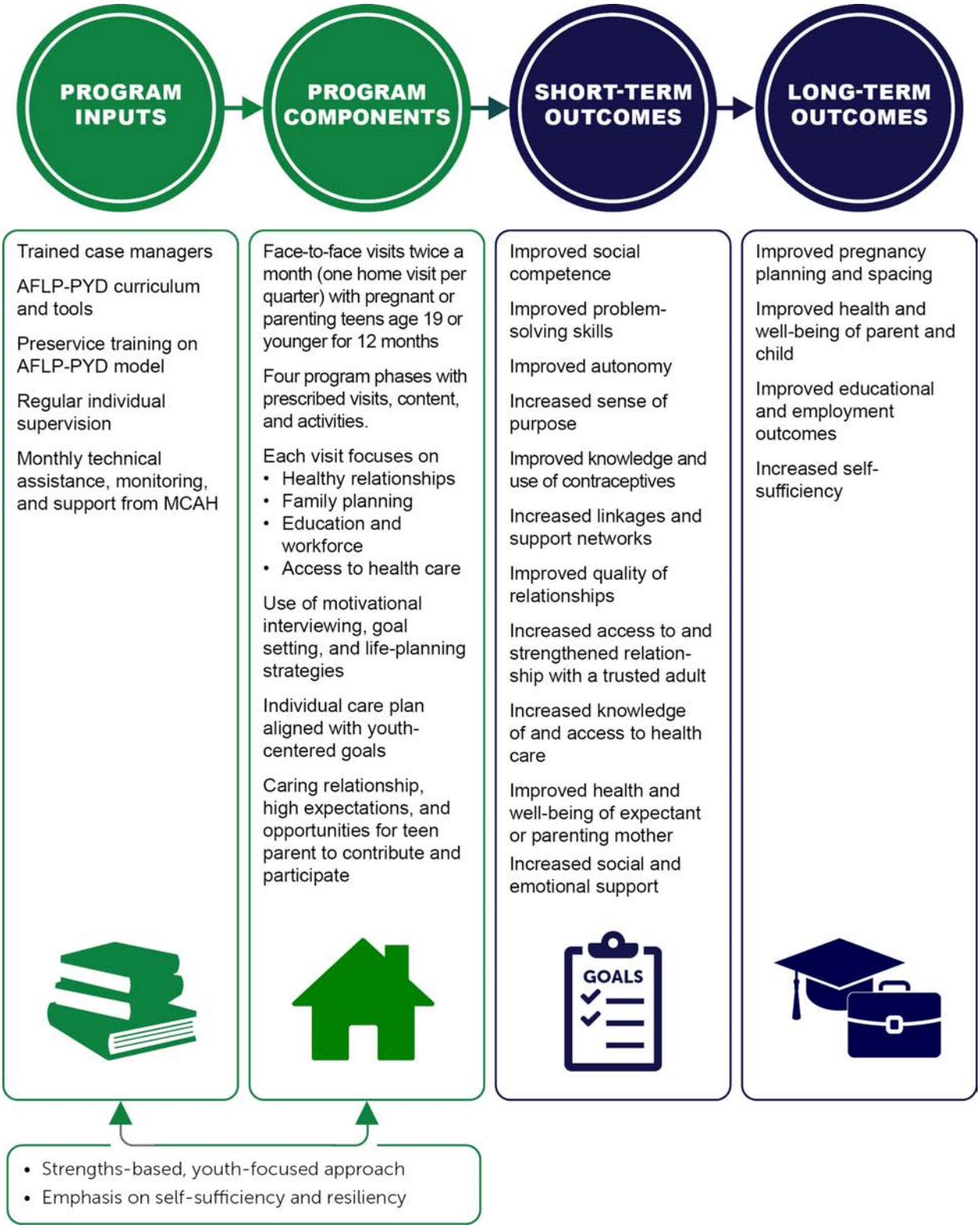
[a] AFLP clients may participate in the program until they reach age 19 or have been in the program for 24 months, regardless of the age of their child. Case managers could submit a waiver for participants to stay in the program longer.

[b] AFLP-PYD was designed to last 12 months, but participants could remain in the program longer if needed.

D. By building social competence and autonomy, AFLP-PYD aims ultimately to improve health and educational outcomes for young parents

The new AFLP-PYD model seeks to achieve several short- and long-term outcomes (Figure II.1). In the short term, the AFLP-PYD program aims to improve the youth's resiliency, knowledge and use of contraceptives, linkages and support networks, access to and relationship with trusted adults, knowledge of and access to health care, health and well-being, and social and emotional support. In the long term AFLP-PYD aims to prevent or delay repeat pregnancy, improve the health and well-being of the participant and child, improve education and employment outcomes, increase the participant's self-sufficiency, and improve linkages to services, community, and support networks that the participant can lean on during the transition out of the program to greater self-reliance.

Figure II.1. AFLP-PYD logic model



III. Assessing Program Impacts: Design, Data Sources, and Analysis Methods

In this chapter, we describe the study's design, data sources and outcome measures, study participants, and impact analysis methods.

A. The AFLP-PYD impact evaluation includes two substudies, each with a distinct design

The primary difference between the two substudies is the unit of assignment. In one substudy, comprising 2 sites, mothers[2] were individually randomized to receive either AFLP or AFLP-PYD. In the other substudy, comprising 13 sites, 7 were randomly assigned to deliver only AFLP-PYD and 6 to deliver only AFLP. We will examine both sets of results by substudy and the pooled results of the combined substudies.

In the first substudy, we randomly assigned *individuals* as part of the enrollment process to receive either AFLP-PYD or AFLP. Thus, we call the first substudy the individual-level study. The individual-level study involved two provider sites that could support individual random assignment because they had a sufficient number of case managers and clients to support both study conditions (AFLP-PYD and AFLP). Each site selected about half of its case managers to receive the state's AFLP-PYD training in fall 2014 and began study enrollment in late fall 2014 through January 2017. We encouraged sites to ensure that case managers assigned to AFLP-PYD were as experienced and qualified as those assigned to AFLP. It is still possible that different case managers were attracted to AFLP-PYD such that the estimated effect may reflect both differences in program services and differences in the type of case manager who prefers AFLP-PYD or AFLP.

In the second substudy, we randomly assigned 13 provider *sites*, each with a smaller number of case managers and clients than the first substudy, to AFLP-PYD or AFLP in summer 2015. We call the second substudy the site-level study because, rather than randomizing individuals, we randomized entire sites to one program or the other as a group. In particular, we grouped sites by geographic proximity and size and then conducted random assignment separately within each group. We assigned 7 sites to the AFLP-PYD group and 6 to the AFLP group. At each site, every eligible individual received the same programming (either AFLP-PYD or AFLP, depending on the site). The AFLP-PYD sites received training in early fall 2015, and participants began enrolling in fall 2015. Study enrollment for all sites ended by January 2017.

[2] Even though AFLP and AFLP-PYD served young parents—both mothers and fathers—the evaluation focused only on mothers.

B. The two substudies differed with respect to the timing of random assignment relative to when individuals entered the program

In both substudies, young mothers were referred to sites and then went through an initial eligibility screening process with the site enrollment manager, typically by telephone. The process was in place before the evaluation activities began. A variety of organizations, including schools or Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) centers, would refer mothers to sites, and sometimes mothers referred themselves. Youth were eligible to participate in the study if they (1) were interested in the program, (2) were pregnant or parenting, (3) had not been served at the site in the previous six months, (4) spoke either English or Spanish, (5) were between age 14 and 18 years, and (6) were not enrolled in Nurse Family Partnerships. When assessing eligibility and possible enrollment, staff at the study sites were instructed to describe both program experiences identically, without making distinctions between AFLP-PYD and AFLP.

The two substudies differed with respect to the timing of random assignment relative to when individuals entered the programs. In the individual-level study, we randomly assigned young mothers to either AFLP-PYD or AFLP *after* the site enrollment manager confirmed participant interest and eligibility, but *before* the mothers agreed to participate in the study. We conducted random assignment within blocks defined within site by pregnancy status and parenting status (ensuring a balance between the AFLP and AFLP-PYD groups with respect to these factors). Given that random assignment occurred after young mothers were referred and screened for eligibility, these steps could not have caused systematic differences between the AFLP-PYD and AFLP groups.

In the site-level study, we randomly assigned provider sites to either AFLP-PYD or AFLP *before* young mothers were referred to sites and screened for interest and eligibility; therefore, site assignment could have affected those processes, leading to systematic differences between the AFLP-PYD and AFLP groups. For instance, it is possible that enrollment managers considered the random assignment of sites when using their discretion to determine whether a young mother was offered enrollment in the study. However, we found no evidence supporting that hypothesis; the same proportion of eligible young mothers in both conditions received a first visit from a case manager (84 percent), and we found no large or statistically significant differences in baseline characteristics between AFLP and AFLP-PYD mothers in the site-level study. As in the individual-level study, mothers gave consent to participate after random assignment.

The study enrollment managers handled enrollment similarly for both substudies. After interest and eligibility screening, a study enrollment manager assigned the young mother to a case manager delivering the program to which she was randomly assigned (either individually or at the site level). The assigned case manager then tried to locate and conduct a first visit with each randomized mother, but case managers could not locate all mothers. For most who were found and visited, case managers were instructed to follow a script that described the program in general terms without differentiating between AFLP and AFLP-PYD. The script did not reveal

the group to which the young mother was assigned. As part of the scripted interaction, case managers sought consent from young mothers to participate in the study's data collection. Rates of consent were similar across program groups and substudies (Table III.1). Those who gave consent responded to the baseline survey in a telephone interview. If the young mother could not complete the baseline survey interview by telephone at that time, she either completed a hard copy survey or made an appointment to complete the survey in the future by telephone. After the young mothers either completed the baseline survey or made an appointment for survey completion in the near future, they received more information on the program to which they were assigned and were asked to provide consent for the programming. Declining programming did not preclude youth from participating in the study's data collection; only 45 of 1,330 youth in the study declined programming after the initial visit. Case managers then scheduled the next visit, and the study team prioritized completion of the baseline survey before that visit.

The overall sample size for the study was 1,330 mothers. Between the two substudies, 698 young mothers were assigned to AFLP-PYD and consented to participate, and 632 were assigned to AFLP and consented to participate (Table III.1). Consent attrition rates met the "low" threshold, according to the What Works Clearinghouse "cautious" boundary standards. The cautious boundary is used in studies with reason to believe that attrition may be more strongly related to outcomes.

Table III.1. Sample sizes and study consent rates, by substudy

	Number assigned to AFLP-PYD	Number assigned to AFLP
Individual-level study: Randomized, located, and consented individuals		
Randomized individuals	244	241
Individuals located for first visits	218 (89%)	221 (92%)
Consenting individuals	199 (82%)	202 (84%)
Site-level study: Sites randomized and individuals located and consented		
Sites	7	6
Randomized individuals	631	576
Individuals located for first visit	531 (84%)	484 (84%)
Consenting individuals in the randomized sites	499 (79%)	430 (75%)
Total consenting individuals	**698 (80%)**	**632 (77%)**

C. To measure the impact of AFLP-PYD relative to AFLP, the evaluation team administered three surveys to the mothers in both groups

The three surveys were (1) a baseline survey administered when mothers first enrolled in the study; (2) a one-year follow-up survey, administered approximately 12 months after enrollment; and (3) a two-year follow-up survey, administered approximately 24 months after enrollment. The evaluation team also acquired administrative data from MCAH's data system, LodeStar. The

administrative data provide information on the total number of visits received by youth, the total number of months enrolled, and the reason for each participant's exit from the program. The survey data provide more detailed information on self-reported services received and other outcomes.

The evaluation team designed the baseline survey as a computer-assisted telephone interview (CATI) that could be completed by using a study-provided cellular phone that allowed the youth to contact trained telephone interviewers in the Mathematica survey operations center during the first visit with their case manager. Overall, 86 percent of study youth completed the survey during the first visit, 8 percent of study youth completed the survey by telephone later, and 4 percent completed a hard copy and mailed it to our survey operations center. Two percent of study youth never completed a baseline survey. The survey included a broad range of measures of educational attainment, family background, and demographic and personal characteristics as well as measures of contraceptive knowledge and sexual risk behavior before a participant's current or recent pregnancy.

The evaluation team designed the follow-up surveys to be administered as both a web-based and CATI survey and conducted the first follow-up survey by using both approaches to ensure a high response rate. In total, 1,160 study participants completed the one-year follow-up survey for a response rate of 87 percent. Response rates were similar for the AFLP-PYD and AFLP groups (88 and 87 percent, respectively). A comparison of baseline characteristics between those who completed the one-year follow-up survey and those who did not showed several key differences, with noncompleters fitting a higher-risk profile. They were younger, less likely to be enrolled in school, reported fewer connections to services, were less knowledgeable about contraceptives, and engaged in riskier sexual behaviors. However, these differences were balanced between the AFLP-PYD and AFLP groups (see Table A.3 in Appendix A).

In this report, we use data from the baseline and one-year follow-up surveys. The analysis of the data from the two-year follow-up survey will appear in a future brief scheduled for fall 2020.

D. This report examines program impacts after one year of program enrollment, on outcomes that align with the AFLP-PYD logic model

Outcomes include (1) measures of program delivery, that is, exposure to information; (2) primary outcomes, including outcomes expected to change within 12-months and early indicators of long-term outcomes; and (3) exploratory outcomes, outcomes less directly related to the AFLP-PYD logic model and outcomes that may or may not have been affected at the time of the 12-month survey (Figure II.1). We summarize each of these outcome types in the discussion below. Appendix A presents detailed definitions of each outcome.

1. Measures of program delivery: Exposure to information

We examined measures in two domains related to the expected program delivery: (1) overall program exposure and (2) exposure to information related to program content. Within overall

program exposure, we examined the total number of visits between a sample member and her case manager, and program dropout, defined as either a sample member conveying the intent to stop programming prematurely to her case manager or becoming unresponsive to a case manager's attempt to contact. With respect to exposure to information related to program content, we examined the following measures of program content derived from the youth survey data: (1) attended a class about relationships/dating/marriage; (2) attended a class about parenting; (3) attended classes about health care for themselves; (4) attended classes about health care for baby; (5) received any information about methods of birth control or where to obtain contraceptives; (6) number of education-related services received, such as GED or college preparation activities; and (7) number of employment-related services received, such as career counseling. The first four of the measures were binary indicators that had a value of 1 if the mother attended a class on that topic administered either by her case manager or somewhere recommended by her case manager, and 0 otherwise.[3] The measure on information related to family planning was also a binary indicator no matter whether the content was administered by the mother's case manager or somewhere recommended by her case manager. It had a value of 1 if the mother received information about methods of birth control or where to obtain birth control, regardless of source.

2. Primary outcomes

Within one year of program provision, AFLP-PYD intends to improve participants' resilience, social competence, problem-solving skills, autonomy, sense of purpose, and knowledge and use of contraceptives; provide increased linkages and support networks; improve quality of relationships; increase access to and strengthen relationships with a trusted adult; increase knowledge of and access to health care; and improve the health and well-being of expectant or parenting mothers. Each of these topics aims to improve participants' resiliency.

We examined two measures of resiliency: feelings about oneself and feelings about goals. We also examined the presence of a trusted adult or supportive relationship and the number of linkages and referrals received.

The three outcome variables—the two measures of resiliency and the presence of a trusted adult or supportive relationship—were scales constructed from a multi-item survey question developed by MCAH. Given that characteristics such as social competence, problem solving, autonomy, sense of purpose, and strength of relationships are difficult to summarize with a single question, the items instead capture various aspects of each. To summarize respondents' answers, we developed these three outcome variables based on a factor analysis of the survey responses; details about that analysis appear in Appendix A. Higher values indicate more "resilient" responses, or stronger adult relationships.

[3] In addition, although we use the shorthand "attended a class" for these four measures, respondents were actually asked whether they attended "any classes or sessions (individual or group) about" the topics listed.

The number of linkages and referrals received ranges from 0 to 4 and is based on whether the mother indicated that she received information—from either her case manager or at a place recommended by her case manager—about each of the following four topics: (1) how to get child care, (2) where to get food assistance, (3) where to find affordable housing, and (4) where to get counseling or treatment for depression or anxiety.

We also estimated impacts on several short-term outcomes from the logic model that may serve as early indicators of long-term outcomes. These impacts included the following binary outcomes: whether the participant had unprotected sex in the past 3 months, was currently using birth control, had depressive symptoms in the past 12 months, had a check-up in the 3 months after birth, had health insurance, and was enrolled in high school or had a diploma/GED. We also examined the degree to which mothers had knowledge of condoms, birth control pills, intrauterine devices (IUD), and other birth control methods. These knowledge variables range from 0 to 1 and indicate the proportion of items in each topic area that participants answered correctly. Finally, we examined whether mothers had positive attitudes toward healthy relationships. We calculated this outcome as the average value across six survey items about healthy relationships. Each item was rated on a scale of 1 to 4, with larger numbers indicating more positive attitudes.

3. Exploratory outcomes

For exploratory purposes, we also estimated impacts on additional outcomes. The outcomes are either less directly related to the AFLP-PYD logic model or are long-term outcomes that may or may not have been affected at the time of the 12-month survey. Exploratory outcomes included the following binary variables: enrolled in a postsecondary institution, employed post-high school, subsequent pregnancy, wants to wait at least three years before having another child, currently has an STD, had as many sick child visits as the mother wanted, and whether the child has health insurance. We also examined one continuous variable—the number of well child visits in the past year.

E. Baseline characteristics of AFLP and AFLP-PYD study participants

The study enrolled mothers who were either pregnant or parenting, with no exclusions on number of children or age of children. At the time of the baseline survey, 46 percent of participants were pregnant with their first child, 3 percent were both currently pregnant and already parenting a child, and 51 percent were parenting but not currently pregnant. Among pregnant participants, most were in their second or third trimester when they entered the program (11 percent were in their first trimester, 48 percent in their second trimester, and 40 percent in their third trimester). Among participants who were already parenting, 94 percent had one child. Among those parenting, their youngest child ranged in age from a newborn to age 4 years, with an average age of eight months. More than 80 percent of participants identified as Hispanic. The mothers ranged in age from 14 to 18 years at study enrollment, with an average age of 17 years across both groups. At the time of study enrollment, the vast majority (85 percent) were enrolled

in school or another education program. Nearly half of mothers reported that they had been suspended or expelled at one point in their lives, but fewer than 10 percent had repeated a grade. Participants were most knowledgeable about condoms, followed by birth control pills and least knowledgeable about long-acting reversible methods, IUDs, and implants.

At the time of study enrollment, more than half of study participants reported attendance in classes related to health care for themselves or their baby, and more than half reported participation in at least one education and employment service. This was true for both the AFLP-PYD and ALFP groups. Approximately 9 and 44 percent, respectively, reported attendance in classes related to relationships and parenting.

Study participants reported that they engaged in risky sexual behavior before pregnancy. Approximately one-third of participants did not use birth control at sexual initiation, which is higher than the national rate of less than 20 percent for youth age 17 and 18 years old as reported by the Centers for Disease Control.[4] The participants' rate is, however, lower than the rate of 55 to 58 percent reported in another similar PAF evaluation with pregnant and parenting adolescents.[5] About one in 10 participants had been diagnosed with a sexually transmitted infection (STI), which is lower than the national rate of one in four for sexually active young adults (age 15 to 24 years) reported by the Centers for Disease Control.[6]

F. There were some minor differences between the AFLP-PYD and AFLP groups at baseline

Besides her program assignment, many other factors could affect outcomes for a participating mother, such as her age, school achievement, or home environment. The random assignment process helps ensure that these factors are evenly distributed across participants in the AFLP-PYD and AFLP groups. We checked the balance by assessing the similarity between the two groups at the start of the program across 32 observable characteristics. These characteristics included demographic factors such as age, race/ethnicity, and language preference; measures of school experience, including current enrollment, highest grade completed, past grade retention, and past suspension or expulsion; pregnancy status; and baseline measures or correlates for 17 program outcomes.

We assessed both the magnitude and statistical significance of differences between the AFLP-PYD and AFLP groups. To assess the magnitude of differences on a standardized scale, we converted all differences into *effect sizes* (that is, we divide each difference by the standard deviation of the outcome). A difference was deemed statistically significant if its *p*-value was less than 0.10. In the site-level study, we detected no differences that met the threshold, but three

[4] Finer, L.B., and J.M. Philbin. "Sexual Initiation, Contraceptive Use, and Pregnancy among Young Adolescents." *Pediatrics*, vol. 131, no. 5, 2013, pp. 886–891.

[5] Baseline survey data on similar data points used in the PAF evaluation of Healthy Family, Healthy Futures in Houston, Texas, during the same time frame.

[6] See https://www.cdc.gov/std/stats18/adolescents.htm.

differences that did so in the individual-level study: Spanish language preference, knowledge of IUDs, and an indicator for recent attendance at classes on relationships/dating/marriage. In one of these cases—knowledge of IUDs—the statistical significance of the difference carried through to the combined sample. However, none of these differences had a large effect size (usually considered greater than 0.25 standard deviations). When testing equivalence for a large number of characteristics, some differences are likely to meet the 10 percent threshold simply by chance. Because these differences are relatively small in size and spread across several domains gives us confidence that they do not indicate systematic unobserved differences in the sample that could influence the results. Moreover, by controlling for all baseline characteristics in our main analysis, we further prevent these differences from skewing our estimates of AFLP-PYD's impacts. More details on the baseline equivalence analysis appear in Appendix B.

G. To estimate the impact of the new AFLP-PYD program, we calculated the difference between outcomes for AFLP-PYD and AFLP participants

In a random assignment study, we can see how much AFLP-PYD increased or decreased an outcome by subtracting the average outcome for AFLP participants from the average outcome for AFLP-PYD participants. When calculating this difference in outcomes, we used regression adjustment to account for the baseline characteristics described in the previous section. This means that even if, by chance, the AFLP-PYD and AFLP groups differed on some preprogram characteristics, the impact estimates reflect the impact of AFLP-PYD as compared to that of AFLP, rather than any preexisting differences between the two groups. As a check of our regression-adjusted impact estimates, we also calculated the impacts accounting only for treatment status and the block-group variables used in randomization (study site and pregnancy/parenting status for individual study participants). Appendix B presents the results of this analysis and other sensitivity analyses described below.

In the individual-level study, we used a straightforward ordinary least squares (OLS) model to estimate both baseline equivalence and impacts on 12-month follow-up outcomes. In the site-level study, we used an approach called random effects to account for the possibility that some sites may have had better or worse average outcomes, even before the study assigned whole sites to either the AFLP-PYD or AFLP group. We calculated the overall impact of AFLP-PYD as a weighted average of the impact estimates from the two substudies. Details about the weights appear in Appendix A.

We accounted for missing data by using a method called multiple imputation, which relies on non-missing response data to fill in missing values for one or more questions (Rubin 1987). In our main analysis, we used this method both for outcomes and for baseline characteristics. As a check of the results from the main analysis, we also estimated impacts by using three other methods. All three methods used actual responses for outcomes and excluded individuals with missing outcome data from the analysis of that particular variable. The first method was a complete case analysis that dropped individuals from the analysis if they had a missing value for

any baseline characteristic or for the outcome in question. Second, we used multiple imputation only for baseline characteristics (not for outcomes). Third, we imputed missing baseline characteristics to zero and added binary variables to the regression—one associated with each characteristic—indicating that a value had been "back-filled" in this way.[7] The results are not sensitive to the different methods we used to account for missing data (See Appendix B).

We used two approaches to interpret the evaluation results. First, we report the statistical significance of impact estimates. We report statistical significance because it is familiar to many readers. However, statistical significance is often misinterpreted (Wasserstein and Lazar 2016; Greenland et al. 2016). In addition to reporting the statistical significance of our impact estimates, we also report the probability that the program truly had a favorable (or unfavorable) impact given our estimates and prior evidence on the distribution of effects of home-visiting programs. This probability is called a Bayesian posterior probability.[8] In our discussion of the study's findings, we draw attention to results with probabilities greater than 70 percent. The prior evidence we used was a selection of moderate- or high-quality studies from the Home Visiting Evidence of Effectiveness (HomVEE) review—a large, rigorous, systematic review of interventions serving a disadvantaged population of new parents. The prior evidence shows that the effects of home-visiting programs tend to be small; we estimate that about 90 percent of effects are smaller in magnitude than 0.17 standard deviations.[9] Given that we report Bayesian posteriors, which are less sensitive than *p*-values to testing multiple outcomes within a single domain, we did not make a multiple comparison adjustment to our *p*-values.

[7] The practical effect of this is to create a new "no response" category for each variable—which may signify systematic differences from those who answered the question—without assuming a value for any missing data.

[8] When calculating the Bayesian posterior probability for a given outcome, we combine our prior evidence distribution with the impact and standard error estimates obtained for that outcome. We also take account for other impacts estimated using the same sample and analysis specification, and for correlations—estimated using a bootstrapping procedure—between impact estimates.

[9] Drawing on our meta-analysis of the HomVEE database we use a prior distribution that is normal with a mean effect size 0.01 and a standard deviation of 0.10. See Deke and Finucane (2019) for more information on The BASIE (BAyeSian Interpretation of Estimates) Framework.

IV. Impact Findings at 12-months

In this chapter, we describe the impacts of AFLP-PYD relative to AFLP on three sets of outcomes: (1) measures of program delivery, that is, exposure to information; (2) primary outcomes, including both short-term outcomes that are linked to the AFLP-PYD logic model and could reasonably be expected to shift in the 12 months between enrollment and follow-up, and early indicators of long-term outcomes; and (3) exploratory outcomes, which are either less related to the logic model or for which impacts may take more than 12 months to manifest. We discuss all findings with a Bayesian posterior probability of being favorable or unfavorable to AFLP-PYD of 95 percent or greater, and identify which are also statistically significant.

A. Measures of program delivery: Exposure to information

1. Mothers in the AFLP-PYD group received more home visits than mothers in the AFLP group, yet left the program in the first year at higher rates

The program model for AFLP-PYD called for an increased frequency of home visits, with 24 sessions between case managers and participants in the year after enrollment, instead of 12 sessions. As shown in Table IV.1, we found that AFLP-PYD increased the number of meetings in the first year by around 3 meetings, to 12.46 from an average of 9.14 for AFLP participants. Though the increase was statistically significant (p = .004) and had a high (99 percent) probability of being favorable to AFLP-PYD, it was smaller than needed to reach the intended 24 visits per year. Case managers also reported that it was difficult for youth to manage two visits a month (Asheer et al. 2020). Mothers in the AFLP-PYD group were also more likely than mothers in the AFLP group to drop out of the program in the first year. This impact was both statistically significant (p = .03) and had a 98 percent probability of being unfavorable to AFLP-PYD.

2. AFLP-PYD decreased the likelihood that participants reported receiving information and services through their case manager

We found that few mothers in both AFLP-PYD and AFLP received classes or services from their case manager. The great majority of both groups reported receiving no education services (63 percent of AFLP-PYD mothers and 62 percent of AFLP mothers) and no employment services (69 percent of AFLP-PYD mothers and 62 percent of AFLP mothers). For all seven measures of program delivery, we found that mothers in AFLP-PYD were less likely than mothers in AFLP to receive classes or services through their case manager. Two of the seven measures on information and service receipt were statistically significant—the decrease in information related to mother's health care (p = .08) and a lower rate of employment services received by AFLP-PYD participants as compared to AFLP (p = .04). However, for two of the seven measures, the probability that the impact was unfavorable to AFLP-PYD was greater than 95 percent, with impacts ranging from 1 to 6 percentage points fewer AFLP-PYD mothers receiving information than AFLP mothers, along with receiving fewer services on average.

These results could be expected given the findings from the implementation study (Asheer et al. 2020). One major focus of the AFLP visits was to get mothers the services they needed. During the implementation study, AFLP case managers discussed that they prepared for their visits by compiling the resources and referral information they would need for each mother, with materials typically focused on education, employment, and family planning services. In contrast, the AFLP-PYD visits were designed to emphasize goal-setting while building mothers' strengths, motivation, and self-sufficiency.

Still, we looked for other possible explanations for these findings. First, we observed that these negative impacts are primarily driven by statistically significant impacts in the individual-level substudy. While we also found large negative impacts of AFLP-PYD in the site-level substudy, they were not statistically significant.

Within the individual-level substudy, we then looked to see if the negative impacts were driven by the earlier enrollees, at the time when case managers were reporting more challenges implementing the new AFLP-PYD program, but found this was not the case. We asked MCAH to identify the AFLP-PYD and AFLP sites that they considered delivered the strongest programs and tested whether the negative impacts were being driven by the lower performing sites. However, the negative impacts were more pronounced in the higher performing sites. The negative impacts were also more pronounced among the youth who had a greater number of visits with case managers. Finally, recognizing that AFLP-PYD was designed to encourage self-sufficiency, we reconstructed the measure by removing the case manager as the source of the information, and the results were similar. Appendix B presents the results of these additional analyses.

Table IV.1. AFLP-PYD impacts on measures of program delivery at 12 months

Measure	AFLP-PYD mean	AFLP mean	Estimated impact	*p*-value	Probability the impact is favorable to AFLP-PYD	Probability the impact is unfavorable to AFLP-PYD
Number of meetings with a case manager	12.46	9.14	3.31***	0.00	**99%**	1%
Withdrew from MCAH programming (percent)	30	21	9**	0.03	2%	**98%**
Number of education-related services received (out of six)	0.45	0.56	-0.11	0.16	6%	94%
Number of employment-related services received (out of four)	0.31	0.47	-0.16**	0.04	1%	**99%**
Attended a class about relationships, dating, or marriage[a] (percent)	14	15	-1	0.75	11%	89%
Attended a class about parenting[a] (percent)	24	27	-3	0.35	7%	93%
Attended a class about health care for themselves[b] (percent)	21	27	-6*	0.08	2%	**98%**
Attended a class about health care for baby[b] (percent)	25	28	-4	0.22	8%	92%
Received information about methods of birth control or where to get birth control[c] (percent)	89	91	-2	0.42	8%	92%

Note: Probabilities greater than or equal to 70 percent are shown in bold. Values less than 70, but displayed as 70 percent due to rounding, are not bold.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

B. Primary outcomes

1. Mothers in the AFLP-PYD group exhibited the same amount of resiliency as mothers in the AFLP group

For both measures of resiliency—feelings about oneself and feelings about goals—the differences between the AFLP-PYD and AFLP groups were small and statistically insignificant (Table IV.2). While building resiliency was a primary goal of AFLP-PYD, this finding may reflect the fact that AFLP-PYD case managers found it challenging to translate their AFLP-PYD training into practice (Asheer et al. 2020). Although the training gave AFLP-PYD case managers a good foundation for the new model's approach and content, they felt that it did not provide sufficient guidance on how to implement the program on a day-to-day basis and, because the program was new, AFLP-PYD case managers did not have an opportunity to shadow more seasoned case managers to learn and observe in the field.

2. Mothers in AFLP-PYD received fewer linkages and referrals than mothers in AFLP, but reported greater access to a supportive adult

We found a negative impact of AFLP-PYD on the number of linkages and referrals received. This outcome measured the degree to which information was received through the case manager in four areas—child care, food assistance, affordable housing, and mental health services. AFLP-PYD participants reported that they received an average of 0.83 linkages or referrals (out of 4), compared to 1.08 among AFLP participants. Similar to the negative impacts on measures of program delivery described above, this finding could reflect the different focus of AFLP and AFLP-PYD visits, the fact that AFLP-PYD case managers found it difficult to follow the prescribed AFLP-PYD structure and content, and may also reflect that AFLP-PYD mothers left the program at higher rates in the first year. This impact was statistically significant ($p = .01$) and had a high probability (over 99 percent) of being unfavorable to AFLP-PYD. We also found a positive, nonsignificant impact of AFLP-PYD on the presence of a trusted adult or supportive relationships. In focus groups conducted during the implementation study, mothers in both study groups reported that they formed close bonds with their case manager and that the case manager often filled the need in their lives for a trusted adult (Asheer et al. 2020). It is possible that the greater number of AFLP-PYD visits during the first year helped AFLP-PYD participants feel a stronger connection.

3. Mothers in AFLP-PYD were more likely than AFLP mothers to have engaged in unprotected sex in the past three months, but equally likely to be currently using birth control

Twenty-five percent of mothers in the AFLP-PYD group engaged in unprotected sex in the three months prior to the survey, compared to 21 percent of mothers in the AFLP group. This impact was not statistically significant but had a high probability (98 percent) of being unfavorable to AFLP-PYD. Similar percentages of mothers in both groups reported currently using birth control (73 percent for AFLP-PYD and 74 percent for AFLP).

4. Mothers in AFLP-PYD group had less knowledge about IUDs and slightly less knowledge of other methods of birth control methods than mothers in the AFLP group

For three of four types of birth control—condoms, IUDs, and other birth control methods—AFLP mothers answered a greater proportion of the survey items correctly than AFLP-PYD mothers. The probabilities that these impacts were unfavorable to AFLP-PYD were higher than 95 percent for knowledge about condoms and IUDs, but the differences between the AFLP-PYD and AFLP groups were small (fewer than one question, on average) and statistically significant only for IUDs. Even though AFLP-PYD was supposed to cover more structured discussions on family planning, the implementation study data suggest that AFLP case managers focused more on family planning during their visits (Asheer et al. 2020).

5. Mothers in AFLP-PYD reported similar levels of depressive symptoms and health insurance as AFLP mothers, but more medical checkups

Similar percentages of mothers in both groups reported depressive symptoms in the past 12 months (14 percent of AFLP-PYD mothers and 15 percent of AFLP mothers). Similar percentages in both groups reported health insurance coverage (89 percent of AFLP-PYD mothers and 91 percent of AFLP mothers). For both of these outcomes, the differences between the two groups were not statistically significant, nor did they have high probabilities of being favorable or unfavorable to AFLP-PYD. Mothers in the AFLP-PYD group were more likely than mothers in the AFLP group to have had a post-partum check-up in the 3 months after their child's birth (88 percent of AFLP-PYD mothers, versus 86 percent of AFLP mothers). However, this finding was not statistically significant.

6. Mothers in AFLP-PYD were less likely than mothers in AFLP to be enrolled in high school or have a diploma or GED, and to have positive attitudes toward healthy relationships

Eighty-three percent of mothers in the AFLP-PYD group were enrolled in high school or had earned a diploma or GED compared to 85 percent of mothers in the AFLP group. This finding was not statistically significant but had a high probability (96 percent) of being unfavorable to AFLP-PYD. A similar percentage of mothers in both groups reported positive attitudes toward healthy relationship, and while this finding was not statistically significant it had a high probability (95 percent) of being unfavorable to AFLP-PYD.

Table IV.2. AFLP-PYD impacts on short-term outcomes and early indicators of long-term outcomes at 12 months

Outcome	AFLP-PYD mean	AFLP mean	Estimated impact	*p*-value	Probability the impact is favorable to AFLP-PYD	Probability the impact is unfavorable to AFLP-PYD
Short-term outcomes						
Feelings about oneself[a]	-0.03	0.02	-0.05	0.57	8%	92%
Feelings about goals[a]	0.03	-0.01	0.03	0.69	25%	75%
Number of linkages and referrals received (out of four)	0.83	1.08	-0.25***	0.01	<1%	**>99%**
Presence of trusted adult/supportive relationships[a]	0.04	-0.03	0.08	0.23	29%	71%
Early indicators of long-term outcomes (percent unless otherwise noted)						
Had unprotected sex in past 3 months	25	21	5	0.15	2%	**98%**
Currently using birth control	73	74	-1	0.80	13%	87%
Knowledge of condoms[b]	68	69	-2	0.32	5%	**95%**
Knowledge of birth control pills[b]	53	53	0	0.98	25%	75%
Knowledge of IUDs[b]	39	42	-4*	0.09	2%	**98%**
Knowledge of other birth control methods[b]	42	44	-2	0.45	8%	92%
Had depressive symptoms in past 12 months	14	15	-1	0.66	22%	78%
Has health insurance	89	91	-2	0.44	12%	88%
Had check-up in the 3 months after child's birth	88	86	3	0.29	25%	75%
Enrolled in high school or earned a diploma/GED	83	85	-3	0.33	4%	**96%**
Positive attitudes toward healthy relationships[c] (scale 1–4)	3.30	3.32	-0.03	0.29	5%	**95%**

Note: Probabilities greater than or equal to 70 percent are shown in bold. Values less than 70, but displayed as 70 percent due to rounding, are not bold.

[a] This outcome is a factor variable that is roughly standard normal (mean = 0, standard deviation = 1). Positive values indicate more "resilient" responses.

[b] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

[c] This outcome was calculated as the average value across six survey items about healthy relationships. Each item was rated 1 through 4, with larger numbers indicating more correct responses.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

C. Exploratory outcomes

1. Mothers in AFLP-PYD were less likely than mothers in AFLP be enrolled in a postsecondary institution, but equally likely to be employed and to have delayed a subsequent pregnancy

Among mothers of high school graduation age or older, fewer than one-third in either group were enrolled in a postsecondary institution (26 percent of AFLP-PYD mothers and 30 percent of AFLP mothers). This impact was not statistically significant but had a high probability (94 percent) of being unfavorable to AFLP-PYD. In both the AFLP-PYD and AFLP groups about one-third of mothers were employed. Rates of subsequent pregnancy were also similar between the two groups (14 percent of AFLP-PYD mothers and 15 percent of AFLP mothers had a subsequent pregnancy). The percentages of mothers who wanted to wait at least three years before having another child were also similar (84 percent of AFLP-PYD mothers and 86 percent of AFLP mothers). For both the employment and subsequent pregnancy outcomes, the differences between the AFLP-PYD and AFLP groups were small and statistically non-significant.

2. Children of AFLP-PYD mothers had fewer well-visits with a doctor and were less likely to have health insurance.

The number of well visits for the child was nearly identical in the two groups (5.46 for the AFLP-PYD group and 5.79 for the AFLP group), though children in the AFLP-PYD group were slightly less likely than children in the AFLP group to have health insurance (94 versus 96 percent). Neither difference was statistically significant. On other measures of health and well-being, AFLP and AFLP-PYD mothers reported similar outcomes. Mothers in AFLP-PYD and AFLP were equally as likely to report an STD (5 percent of AFLP-PYD mothers and 6 percent of AFLP mothers). Similar percentages of mothers in both groups reported as many sick child visits as they wanted (90 percent for both groups).

D. Results are robust to alternative analysis methods

The results described above are robust to alternative analysis methods. As described in Chapter III, we probed the robustness of the results by using a variety of analysis methods. A full summary of the results for these analyses appears in Appendix B. In most cases, the alternative analyses yielded impact estimates of the same sign and similar magnitude as those of the main analysis. As describe above, for measures of program delivery, the exploratory analyses generally found more statistically significant negative impacts than the benchmark analysis.

Table IV.3. AFLP-PYD impacts on exploratory outcomes at 12 months

Outcome	AFLP-PYD mean	AFLP mean	Estimated impact	*p*-value	Probability the impact is favorable to AFLP-PYD	Probability the impact is unfavorable to AFLP-PYD
Enrolled in a postsecondary institution	26	30	-4	0.30	6%	94%
Employed post-high school	35	33	2	0.64	24%	76%
Subsequent pregnancy	14	15	-1	0.67	18%	82%
Wants to wait at least three years before having another child	84	86	-2	0.45	12%	88%
Has an STD	5	6	-1	0.62	18%	82%
Had as many sick child visits as mother wanted	90	90	0	0.89	15%	85%
Number of well visits for child	5.46	5.79	-0.32	0.35	8%	92%
Child has health insurance	94	96	-3	0.16	7%	93%

Note: Probabilities greater than or equal to 70 percent are shown in bold.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

V. Conclusion

Focusing on outcomes measured after 12 months, this report provides a first look at impacts of the AFLP-PYD enhancement to the AFLP program for pregnant and parenting youth in California. AFLP-PYD is intended to differ from the original AFLP by (1) using a prescribed set of structured activities and content to help young parents identify their strengths and use them to meet their goals and (2) requiring case managers to conduct two visits a month instead of one. By increasing social competence and independent decision making, AFLP-PYD aims to ultimately improve health and educational outcomes for young mothers.

For many of the measures we examined, outcomes for mothers served by the AFLP-PYD enhancement are similar to outcomes for mothers served by AFLP. This is perhaps not surprising, given that both groups participated in a case management program that aimed to ultimately delay repeat pregnancies and improve educational and health outcomes, and the implementation study did not detect notable differences in program delivery (Asheer et al. 2020). In addition, given that half of the mothers gave birth after beginning the program, case managers had fewer than 12 months to work with mother and child.[10]

AFLP-PYD was developed to increase youth resiliency, yet mothers in the AFLP-PYD and AFLP groups exhibited the same amount of resiliency. This finding may reflect the fact that AFLP-PYD case managers found it challenging to translate their AFLP-PYD training into practice and encountered several challenges as they began implementing the new program (Asheer et al. 2020). AFLP-PYD did increase the number of meetings with a case manager, as intended. However, the average number of visits was far less than the intended 24 per year. We found two additional positive impacts for AFLP-PYD mothers; they reported an increased presence of a supportive adult and a higher proportion of mothers reported having had a post-partum check-up in the three months after their child's birth.

Using tests of statistical significance and Bayesian posterior probabilities to interpret the magnitude of the differences, we identified several negative effects of AFLP-PYD on the mothers and their child. AFLP-PYD mothers attended fewer classes and sessions about parenting, health care for themselves and their baby, and received fewer education and employment related services. AFLP-PYD mothers received fewer linkages and referrals, and also withdrew from the program at higher rates. AFLP-PYD mothers were more likely to have unprotected sex and had less knowledge of some methods of birth control. AFLP-PYD mothers had less positive attitudes towards healthy relationships, were less likely to be enrolled in high school or have earned a diploma or GED, and less likely to be enrolled in a postsecondary institution. Their children had fewer well visits and were less likely to have health insurance.

[10] The proportion of pregnant and parenting youth at baseline were equivalent across the AFLP and AFLP-PYD groups. Therefore, this should not have influenced differences between AFLP and AFLP-PYD as much as it could have influenced the fact that the treatment being tested, AFLP-PYD, may not have had sufficient time to be implemented with mother and baby.

The AFLP-PYD program is intended to be implemented in 12 months. However, almost 40 percent of mothers ever randomized to AFLP-PYD were still receiving services after one year. Among mothers who were active in the program for the entire year, 90 percent of them were still participating. Since the program was ongoing for most AFLP-PYD youth still participating, it will be important to examine program impacts on all of these measures after they hypothetically should complete the program. We will report these impacts in an upcoming brief.

Appendix A. Additional Details on Data Sources and Analysis Methods

This appendix is a technical supplement to the impact report for the evaluation of the Positive Youth Development (AFLP-PYD) Program for Expectant and Parenting Teens in California. It provides additional detail on the evaluation's design, methods, and findings. The first section of the appendix describes the recruitment of adolescent mothers into the study. The second section describes the methods used to enroll and randomly assign mothers. The third section describes the survey administration procedures and response rates. The fourth section describes the program delivery measures and outcomes examined. The fifth section describes the methods used to analyze these measures and outcomes.

A. Study recruitment

Beginning in fall 2013, Mathematica worked with California Department of Public Health, Center for Family Health, Maternal, Child and Adolescent Health (MCAH) to identify sites that were (1) large enough to support individual random assignment, (2) located in counties with higher adolescent birth rates, and (3) not sites that had participated in a pilot study of AFLP-PYD. Four such sites were identified: (1) San Diego Adolescent Pregnancy Prevention Program (SANDAPP), (2) Planned Parenthood Mar Monte (PPMM), (3) Alta Med Youth Services – East LA, and (4) El Nido Family Centers. We recruited PPMM and AltaMed into the study for individual random assignment (substudy 1). Each site began study enrollment in late fall 2014. Although SANDAPP and El Nido had enough case managers and youth, the geographic dispersion of their service area and the assignment of case managers to specific regions in their service area made it difficult to implement individual random assignment. Therefore, we recruited El Nido for the site-level study (substudy 2), and randomly determined which of their two centers, which were geographically quite far from each other, would receive AFLP-PYD and which would remain AFLP. Finally, AltaMed had a small satellite location in Long Beach that could not support individual random assignment so we randomly determined which program it would provide (AFLP-PYD). These three locations (two El Nido programs and one AltaMed program) were the first sites of the second substudy.

In fall 2014, we began to recruit additional provider sites, including SANDAPP, for the second substudy. We looked for sites that had past evidence of enrolling at least 20 new clients per year that were located in counties with higher adolescent birth rates and were not AFLP-PYD pilot sites. We identified 10 additional providers for randomization. We recruited sites through summer 2015, and stratified sites by geographic proximity and size before random assignment. The AFLP-PYD sites received training in early fall 2015, and we began enrolling youth in these sites in fall 2015. Table A.1 below shows the sites randomly assigned within each stratum and their final enrollment counts.

Table A.1. Site-level random assignment sites

Random assignment strata	Site Name	Enrollment
1	El Nido Family Centers, South LA/Compton (AFLP-PYD)	157
	El Nido, Family Centers, San Fernando Valley (AFLP)	197
2	Tiburcio Vasquez Health Center (AFLP-PYD)	19
	Felton Institute Family Service Agency of San Francisco (AFLP)	27
3	Sonoma County Health Department (AFLP-PYD)	63
	Sutter Health Teen Programs (AFLP)	52
4	Community Action Partnership of Santa Barbara County (AFLP-PYD)	45
	Community Action Partnership of San Luis Obispo County (AFLP)	29
5	San Diego Adolescent Pregnancy Prevention Program (AFLP-PYD)	138
	Stanislaus County Health Services (AFLP)	106
6	Placer Country Health and Human Services (AFLP-PYD)	29
	Lake Family Resource Center (AFLP)	19
7	Alta Med Health Centers, Long Beach (AFLP-PYD)	48

B. Enrollment and Random Assignment

Young mothers entered both substudies in a similar manner, but the substudies differed in the timing of randomization relative to program enrollment. In the individual-level study, young mothers were first referred to sites. In the site-level study, provider sites were first randomized to either the AFLP-PYD or AFLP. Young mothers were then referred to sites in the same manner as in the individual-level study. A variety of organizations, including schools or WIC centers, could make these referrals. Referred young mothers then went through an initial eligibility screening process with the site enrollment manager, typically by telephone. Youth were eligible to participate in the study if they (1) were interested in the program, (2) were pregnant or parenting, (3) had not been served at the site in the prior six months, (4) spoke either English or Spanish, (5) were between ages 14 and 18, and (6) were not enrolled in Nurse Family Partnerships.

If youth were found eligible for the study, the site enrollment manager entered them into a web-based sample enrollment system. This system was used to randomly assign the young mothers to either AFLP or AFLP-PYD in the individual-level study. Random assignment was conducted within blocks defined by pregnancy status and parenting status, ensuring balance between the AFLP and AFLP-PYD groups with respect to these factors. For the site-level study, the sample enrollment system was used to enroll mothers in the study, but random assignment was not necessary since this had already occurred at the site level.

Across substudies, it was possible that eligible young mothers would not have been entered into the sample enrollment system and enrolled in the study. The site enrollment manager administered the AFLP Acuity Scale to youth during the first point of contact to assess youth needs and risk in many domains. The scale was used to prioritize youth who needed the programming the most if sites had full caseloads. If caseloads were full, the young mother would not have been entered into the sample enrollment system. At the discretion of the sites, some young mothers were offered the program but were not entered into the sample enrollment system due to their urgent need for services, such as young mothers who were homeless or suffering from interpersonal partner violence.

After entry into the sample enrollment system, and random assignment for the individual-level study, the study enrollment manager assigned the participant to a case manager delivering the program to which she was assigned. The assigned case manager then tried to locate and conduct a first visit with each youth mother, but not all could be found. For most who were found and visited, case managers were instructed to follow a script that described the program in general terms (without differentiating between AFLP and AFLP-PYD) and the study in detail. The script did not reveal the group to which the young mother was assigned. As part of the scripted interaction, case managers sought study consent from young mothers to participate in the study data collection. Those who consented were administered the baseline survey using a telephone interview during that first visit. After the young mothers completed the baseline survey, they received more information on the program to which they were assigned and were asked to provide consent for the programming; only 45 of 1,330 youth in the study declined programming. If the young mother could not complete the baseline survey interview at that time by telephone, she either completed the survey on hard copy or made an appointment to complete the survey in the future. Case managers then scheduled the next visit, and Mathematica prioritized completing the baseline survey before that visit.

C. Survey Administration and response rates

This study drew on information from three surveys: (1) a baseline survey, administered after consent; (2) a one-year follow-up survey, administered about 12 months after consent, and (3) a two-year follow-up survey, administered about 24 months after consent. This report uses data from the baseline and one-year follow-up surveys. A future brief will also use data from the two-year follow-up survey.

The evaluation team designed the surveys to capture a broad range of demographic and personal characteristics and outcomes across several topic areas: maternal education, family and relationships, resiliency, receipt of services, alcohol and drug use, maternal and child health, contraceptive knowledge, sexual behaviors, contraceptive use and intentions, and repeat pregnancy. The team drew most of the questions from surveys used in past evaluations with similar populations, such as the Personal Responsibility Education Program, Building Strong Families Evaluation, Youth Build Program Evaluation and Evaluation of Adolescent Pregnancy Prevention Approaches, as well as established surveys, such as the National Longitudinal Study

of Adolescent Health, the National Survey of Family Growth, the Youth Risk Behavior Survey, the Pregnancy Risk Assessment Monitoring System, and Power to Decide's[11] Fog Zone Survey. Additionally, the items measuring resiliency and adult relationships were developed by MCAH. Both the baseline and follow-up surveys were pretested with small groups of pregnant or parenting adolescents prior to administration, and the follow-up surveys were pretested with pregnant and parenting adolescents enrolled in the AFLP or AFLP-PYD program at sites not participating in the evaluation. The pretests resulted in minor changes to the wording of items to improve clarity and understanding with the target population.

Enrolled mothers were eligible to complete all surveys, regardless of past survey response. As described above, young mothers who consented to the study in their first visit with a case manager were asked to complete a baseline survey at the time of study enrollment, using a study-provided cellular phone that allowed the youth to contact telephone interviewers in our survey operations center; 86 percent of study youth completed the survey at this time. Eight percent of study youth completed the survey by telephone later, and four percent completed a hard copy and mailed it to our survey operations center. Two percent of study youth never completed a baseline survey. Young mothers were provided with a $25 gift card as a thank you for their participation.

Before contacting the young mothers for the one-year follow-up survey, we mailed two reminder postcards that asked for updated contact information at six and nine months after baseline. At the start of the follow-up survey data collection period, youth were mailed and emailed an invitation to complete the survey online. This contact also provided them with an option to call the survey operations center to complete the survey by telephone. During the next four weeks, nonresponders were emailed and texted reminders weekly. At the end of the initial four-week period, telephone interviewers began calling remaining nonresponders to try to complete the survey by telephone. For anyone who could not be reached by telephone, staff trained in locating techniques called alternate contacts provided at baseline and did database searches for updated contact information. In-person field locators worked any remaining noncomplete cases. In-person locators went to the address of the youth and, if they were available, asked them to complete the survey in person on a study provided tablet. Cases had approximately a 6-month window to complete the survey, reflecting the survey administration period in other recent large federal studies of older adolescents. Among the study youth, 39 percent completed the web survey, 35 percent completed the survey by telephone, and 13 percent completed a web survey with an in-person locator. Thirteen percent of youth never completed the 12-month follow-up survey. These response rates were similar across treatment groups (Table A.2). Study participants received a $25 thank you gift card for responding to the follow-up survey.

The baseline and one-year follow up survey procedures yielded high survey response rates. Of the 1,330 study participants, ninety-eight percent completed the baseline survey as part of the study enrollment process. In total, 1,160 individuals completed the one-year follow-up survey for

[11] Formerly the National Campaign to Prevent Teen and Unplanned Pregnancy

an overall response rate of 87 percent. Response rates were similar for the AFLP-PYD and AFLP groups (88 and 87 percent, respectively).

Table A.2. Survey response rates, by treatment group and survey mode at 12 months

	AFLP-PYD	AFLP	Total
Baseline Survey			
Completed (percent)	**98.6**	**96.7**	**97.7**
Telephone (at enrollment)	89.6	82.4	86.2
Telephone (post-enrollment visit)	5.3	10.0	7.5
Hardcopy	3.7	4.3	4.0
12 Month Follow-up Survey			
Completed (percent)	**87.8**	**86.5**	**87.2**
Web-survey	37.5	40.7	39.0
Telephone	38.1	32.0	35.2
In person (web)	12.2	13.9	13.0

A comparison of the characteristics of respondents to the one-year follow-up survey and nonrespondents indicates several differences based on response status (Table A.3). As might be expected, the survey non-responders appear to be more at-risk, they are younger, less likely to be in school or have a trusted adult in their life. However, within survey responders there were no large differences between mothers assigned to AFLP-PYD and AFLP.

Table A.3. Participant characteristics at study enrollment for 12 month follow-up survey respondents and non-respondents, by program group

Baseline measure	Survey respondents			Survey Non-respondents		
	AFLP-PYD mean	AFLP mean	Difference	AFLP-PYD mean	AFLP mean	Difference
Age at random assignment (years)	17.1	16.9	0.2	16.4	15.4	1.0
Race/Ethnicity (percent)						
Hispanic	85	84	1	75	62	13
Black Non-Hispanic	7	3	4	8	7	1
White Non-Hispanic	4	6	-2	7	11	-4
Other Race Non-Hispanic	3	4	-1	6	6	0
Spanish language preference (percent)	8	10	-2	14	14	0

Baseline measure	Survey respondents			Survey Non-respondents		
	AFLP-PYD mean	AFLP mean	Difference	AFLP-PYD mean	AFLP mean	Difference
Highest grade completed at random assignment (percent)						
Seventh grade or lower	1	2	-1	0	4	-4
Eighth grade	10	10	0	11	14	-4
Ninth grade	14	13	1	15	12	4
Tenth grade	24	27	-3	31	18	13
Eleventh grade	38	36	2	32	34	-2
Twelfth grade	12	9	2	8	6	2
Enrolled in school at time of random assignment (percent)	86	85	2	80	71	9
Ever repeated a grade in middle or high school (percent)	8	9	-1	8	19	-11
Ever been suspended or expelled from school (percent)	41	41	0	40	36	4
Pregnant at time of random assignment (percent)	47	48	0	52	45	7
Received information or classes in the past year (percent)						
Classes on relationships and dating	11	8	3	6	5	1
Classes on parenting	45	44	1	41	33	8
Classes on health care for mothers	59	56	2	45	54	-9
Classes on health care for children	65	63	2	52	56	-5
Number of education services received in the past year	0.94	0.93	0.01	0.89	0.86	0.04

Baseline measure	Survey respondents			Survey Non-respondents		
	AFLP-PYD mean	AFLP mean	Difference	AFLP-PYD mean	AFLP mean	Difference
Number of employment services received in the past year	0.39	0.35	0.04	0.49	0.45	0.05
Resiliency (scale 1 to 4)[a]	3.23	3.19	0.04	3.10	2.89	0.21
Presence of a trusted adult (scale 1 to 4)[a]	3.44	3.40	0.04	3.31	2.98	0.33
Used birth control in first sexual encounter (percent)	65	63	3	60	59	1
Percent of birth control knowledge questions answered correctly						
Knowledge of condoms	65	66	-1	60	59	2
Knowledge of birth control pills	51	52	-1	50	48	1
Knowledge of IUDs	39	40	-1	34	36	-2
Knowledge of other birth control	39	40	-1	35	35	0
Had depressive symptoms in past 12 months (percent)	16	16	0	22	12	11
Healthy relationships beliefs/attitudes (scale 1 to 4)	3.25	3.23	0.02	3.14	2.85	0.28
Would be upset to become pregnant again before age 20 (percent)	8	6	2	11	2	8
Currently has an STD (percent)	11	11	-1	14	14	0
Sample Size	613	547		85	85	

Source: Baseline survey conducted by Mathematica.

[a] At follow-up, resiliency and adult relationship outcomes were measured using factors generated from question batteries developed by MCAH. However, only a subset of these items were asked at baseline, so instead we averaged the answers related to each topic. Higher scores indicate responses corresponding to higher resiliency and stronger adult relationships.

IUD=intrauterine device, STD=Sexually transmitted disease

D. Outcome measures

To provide a comprehensive assessment of AFLP-PYD after one year of program enrollment, we assessed program impacts on outcomes that aligned with the AFLP-PYD logic model. These outcomes include (1) measures of program delivery, that is, exposure to information; (2) primary outcomes, including outcomes expected to change within 12-months and early indicators of long-term outcomes; and (3) exploratory outcomes, outcomes less directly related to the AFLP-PYD logic model and outcomes that may or may not have been affected at the time of the 12-month survey. Below we provide details on the construction of each of these measures.

1. Measures of program delivery: Exposure to information

We considered two primary outcomes related to program delivery: (1) overall program exposure and (2) exposure to information related to program content. Within overall program exposure we examined the total number of visits between a participant and her case manager, and program dropout, defined as either a participant conveying the intent to stop programming prematurely to her case manager or becoming unresponsive to a case manager's attempt to contact, based on administrative data from MCAH's data system, LodeStar. To assess exposure to program content, we constructed seven program content measures from the youth survey data: (1) attended a class about relationships/dating/marriage; (2) attended a class about parenting; (3) attended classes about health care for themselves; (4) attended classes about health care for baby; (5) received any information about methods of birth control or where to obtain contraceptives; (6) number of education-related services received, such as GED or college preparation activities; and (7) number of employment-related services received, such as career counseling. These items were adapted based on items from the evaluation surveys used in the Personal Responsibility Education Program and the Evaluation of the Impact of the Youth Build Program Follow-Up Survey. The item on exposure to family planning content was constructed differently from other information items because the survey item used to develop it did not ask specifically if content was delivered by the case manager or somewhere referred by the case manager. Table A.4 shows the individual survey items used to construct each measure.

Table A.4. Measures of program delivery

Measure	How the measure was defined
Overall program exposure	
Total number of visits	Total number of visits within the first 12 months from consent date. From MCAH administrative data system LodeStar.
Program dropout	Binary variable: One, within the first 12 months following the consent date, if the participant conveyed the intent to stop programming prematurely to her case manager or became unresponsive to a case manager's attempt to contact. From MCAH administrative data system LodeStar.

Measure	How the measure was defined
Exposure to information related to program content	
Attended a class about relationships, dating, or marriage	Binary variable: equal to one if respondent reported attending a class or session (individual or group) about relationships, dating, or marriage, and that class took place with the case manager or at the recommendation of the case manager.
Number of education-related services received	Number of education topics addressed by the case manager directly or through a recommendation by the case manager; variable ranges from 0 to 6, with higher values indicating exposure to more education topics. The education topics are: • GED preparation • Tutoring or outside help with school work • Preparation for a high school diploma • Standardized achievement test preparation for state or local tests • College preparation activities • Getting help finding financial aid
Number of employment-related services received	Number of workforce topics addressed by the case manager directly or through a recommendation by the case manager; variable ranges from 0 to 4, with higher values indicating exposure to more workforce topics. The workforce topics are: • Career counseling • Help finding or applying for a job training program • Job training • Help looking for or applying for a job
Attended a class about health care for baby	Multiple-item continuous scale variable: normalized count of topics related to the child's health care addressed by the case manager directly or through a recommendation by the case manager; variable ranges from 0 to 1, with higher values indicating exposure to more topics about the child's health care. These topics are: • How to get health insurance or apply for Medicaid for your baby • Where to get health care for your baby
Attended a class about health care for themselves	Multiple-item continuous scale variable: normalized count of topics related to the mother's health care addressed by the case manager directly or through a recommendation by the case manager; variable ranges from 0 to 1, with higher values indicating exposure to more topics about the mother's health care. These topics are: • How to get health insurance or apply for Medicaid for yourself • Where to get health care for yourself
Attended a class about parenting	Binary variable: One if respondent reported attending a class on parenting or how to care for your baby and the case manager led or recommended the class.
Received information about methods of birth control or where to get birth control	Binary variable: One if the respondent reported receiving information on birth control or where to get birth control from a doctor, nurse, case manager or clinic.

2. Primary outcomes

Within one year of program provision, AFLP-PYD intends to improve participants' social competence; improve participants' problem-solving skills; improve participants' autonomy, and sense of purpose; improve participants' knowledge and use of contraceptives; provide increased

linkages and support networks; improve quality of relationships; increase access to and strengthen relationships with a trusted adult; increase knowledge of and access to health care; and improve the health and well-being of expectant or parenting mothers. Below we provide details on how we developed each of these outcome measures.

3. Resiliency & Access to Supportive Adult Items

We examined two measures of resiliency: feelings about oneself and feelings about goals. We also examined the presence of a trusted adult or supportive relationship. The development of these items drew on a 27-item survey question, broken into the three separate questions in the survey instrument. These items were developed by MCAH through a review of twelve reliable scales.[12] To reduce the potential for respondent fatigue that could lead to lower quality data, more missing data or early breakoffs, respondents were divided randomly into three groups: one third were asked only about feelings about themselves, one third received only the questions about personal goals, and one third received neither group of questions. All three groups received the items about adult relationships. Factor analysis was conducted to generate three separate measures developed from these survey questions. Each group of questions contributed to a separate outcome measure.[13]

Table A.5. Short-term outcomes: resiliency and access to supportive adult items

Outcome	How the outcome was defined
Short-term outcomes	
Respondent feelings about themselves	Standard normal variable (constructed to have mean zero and standard deviation of one) for which positive values indicate more "resilient" responses. Computed based on factor analysis of the following survey items, for which respondents rated their agreement using a four-level scale, "Strongly Disagree", "Disagree", "Agree", or "Strongly Agree": • I have a positive attitude about myself • I am aware of my personal strengths • I use my strengths to solve my problems • I often feel that there is little I can do to change what happens to me • I usually make quick decisions based on what feels right in the moment • When I have a serious disagreement with someone I can talk calmly about it without losing control • I can resist doing something when I know that I shouldn't do it • My life has meaning
Respondent feelings about personal goals	Standard normal variable (constructed to have mean zero and standard deviation of one) for which positive values indicate more "resilient" responses. Computed based on factor analysis of the following survey items, for which

[12] Between the first and second follow-up surveys, MCAH determined that five items from the survey question about feelings about oneself should be removed from the item battery. Additionally, Mathematica determined that one item from the question about goals had low internal consistency when compared with other items from the same question. All six of these items were excluded from the factor analysis.

[13] Please contact the authors of this report for further information on scales used for development.

Outcome	How the outcome was defined
	respondents rated their agreement using a four-level scale, "Strongly Disagree", "Disagree", "Agree", or "Strongly Agree": • I am focused on preventing negative things from happening in my life • I set goals and think about what I need to do to reach those goals • When faced with a problem, I can usually find a solution • I think going to college is important for getting a good job • I am focused on achieving good and positive things in my life • I have a plan for achieving my future education or career goals • I have opportunities that are challenging and interesting
Increased access to and strengthened relationship with a trusted adult	Standard normal variable (constructed to have mean zero and standard deviation of one) for which positive values indicate more "resilient" responses. Computed based on factor analysis of the following survey items, for which respondents rated their agreement using a four-level scale, "Strongly Disagree", "Disagree", "Agree", or "Strongly Agree": • There is an adult who I can count on when things go wrong • There is an adult who helps me make good decisions • There is an adult who encourages me to do my best • There is an adult in my life who supports me with the plans and goals I have for the future • I know where to go to get support for the things I need • I express my ideas, concerns, and opinions with important people in my life (such as family, partner, or friend)

1. Linkages and referrals

The evaluation team developed a single item measuring linkages and referrals mothers received related to participation in AFLP-PYD or AFLP. The item is the count of topics related to linkages and referrals. It was developed based on the mother responding that she received information on the following topics within the past 12-months from her AFLP/AFLP-PYD case manager or through a recommendation by the case manager:

- How to get childcare for your baby
- Where to get food assistance and support for yourself and your baby
- Where to find affordable housing
- Where to get counseling or treatment for depression or anxiety

Values for this measure range from zero to four, with higher values indicating more linkages and referrals through the program.

2. Unprotected sex in past three months

To determine whether AFLP-PYD was successful in reducing rates of unprotected sex, the evaluation team constructed a binary (yes/no) indicator for whether the study participant reported

having had sex without using any effective contraceptive method in the three months before the survey. We constructed this variable using the two survey questions:

- In the past 3 months, have you had sexual intercourse? By sexual intercourse we mean a male putting his penis into a female's vagina.
- In the past 3 months, did you ever have sexual intercourse without using birth control, such as condoms, birth control pills, the patch, the shot, the ring, an IUD, or an implant?

These questions were adapted from the National Longitudinal Study of Youth, 1997 cohort (U.S. Bureau of Labor Statistics n.d.) and were used with pregnant and parenting adolescents on the evaluation of the Personal Responsibility Education Program.

The evaluation team constructed a binary (yes/no) indicator for having had unprotected sex. The indicator was equal to 1 if the mother reported having had sex in the three months before the survey without using any of the above listed methods of birth control. The indicator was equal to 0 if the mother reported that, in the three months before the survey, either she had not had sex at all or had not had sex without any of the above listed methods of birth control.

3. Currently using birth control

To determine whether AFLP-PYD was successful in increasing the use of contraceptives, the evaluation team constructed a binary (yes/no) indicator for whether the study participant currently used a birth control method. In particular, the survey asked mothers to report on whether they had used each of the following methods of birth control at any point in the past 12 months:

- Condoms
- Birth control pills
- The patch, such as Ortho Evra
- The shot, such as Depo-Provera or other injectable birth control
- The ring, such as NuvaRing
- An IUD, such as Mirena, Paragard, Kyleena, Liletta or Skyla
- An implant, such as Implanon or Nexplanon
- Any other type of birth control

Respondents were next asked whether they currently used each of the methods they had indicated using in the past 12 months. Survey items were adapted from the questions used on the Evaluation of Adolescent Pregnancy Prevention Approaches.

For the last category in this list—any other type of birth control—the survey instructed respondents to write in the method used. The evaluation team examined all such responses and ensured they should not be categorized as one of the provided methods.

4. Contraceptive Knowledge

The survey contained a series of true/false items, grouped into blocks around knowledge of condoms, knowledge of birth control pills, knowledge of intrauterine devices (IUDs) and knowledge of other birth control methods adapted from those used on Power to Decide's[14] Fog Zone Survey (2011) with a young adult survey population. Table A.6 provides details about the construction of each measure and on the specific item wording within each block. Answers of "don't know" or missing responses were counted as incorrect responses. If a mother skipped all items within the topic block, the evaluation team excluded her from the analysis of this outcome.

Table A.6. Contraceptive knowledge measures, with correct answers shown

Outcome	How the outcome was defined
Knowledge of condoms	Multiple-item continuous scale variable: variable ranges from 0 to 1 and indicates the proportion of the following six items about condoms answered correctly: • It is okay to use the same condom more than once. (FALSE) • Condoms have an expiration date. (TRUE) • When putting on a condom, it is important to leave space at the tip. (TRUE) • It is okay to use petroleum jelly or Vaseline as a lubricant when using latex condoms. (FALSE) • When using a condom, it is important for the man to pull out right after ejaculation. (TRUE) • Wearing two latex condoms will provide extra protection. (FALSE)
Knowledge of birth control pills	Multiple-item continuous scale variable: variable ranges from 0 to 1 and indicates the proportion of the following five items about birth control pills answered correctly: • Birth control pills are effective, even if a woman misses taking them for two or three days in a row. (FALSE) • Women should "take a break" from the pills every couple of years. (FALSE) • After a woman stops taking birth control pills, she is unable to get pregnant for at least two months. (FALSE) • In order to get the birth control pill, a woman must have a pelvic exam. (FALSE) • Birth control pills can reduce risk of getting a sexually transmitted disease or STD. (FALSE)

[14] Formerly the National Campaign to Prevent Teen and Unplanned Pregnancy

Outcome	How the outcome was defined
Knowledge of IUDs	Multiple-item continuous scale variable: variable ranges from 0 to 1 and indicates the proportion of the following six items about IUDs answered correctly: • Women who use IUDs cannot use tampons. (FALSE) • A woman can get an IUD without going to a doctor's office, clinic, or medical professional. (FALSE) • An IUD cannot be felt by a woman's partner during sex. (TRUE) • IUDs can move around in a woman's body. (FALSE) • An IUD is effective (prevents pregnancy) for at least 3 years. (TRUE) • Using an IUD will cause weight gain. (FALSE)
Knowledge of other birth control	Multiple-item continuous scale variable: variable ranges from 0 to 1 and indicates the proportion of the following four items about other forms of birth control answered correctly: • Women using the birth control shot, Depo-Provera, must get an injection about every 3 months. (TRUE) • Women using the vaginal ring, or NuvaRing, must have it inserted by a doctor or health care provider every month. (FALSE) • Long-acting methods like the implant (Implanon or Nexplanon) or an IUD (Mirena, ParaGard, or Skyla) cannot be removed early, even if a woman changes her mind about wanting to get pregnant. (FALSE) • Long-acting methods like the implant (Implanon or Nexplanon) or an IUD (Mirena, ParaGard, or Skyla) can make it more difficult to become pregnant in the future when a woman is no longer using them. (FALSE)

5. Improved health or well-being of the mother

We considered three outcomes related to the health and well-being of the mother: an indicator of the presence of depressive symptoms, receipt of postpartum checkup, and health insurance status of the mother.

Depressive symptoms. The survey instrument asked if, in the past twelve months, the respondent was so sad or hopeless for two or more weeks in a row that she stopped doing some usual activities. Using this question the evaluation team constructed a binary (yes/no) variable indicating whether the participant had depressive symptoms. This question is from the National Youth Risk Behavior Standard High School Survey.[15]

Health insurance for mother. The survey instrument asked mothers whether they currently have any type of health insurance. Using this question, the evaluation team constructed a binary (yes/no) variable indicating mother's health insurance status.

Mother received postpartum checkup. The survey instrument asked mothers whether they received a checkup for themselves from a doctor, nurse or other health care worker in the 3

[15] Available at https://www.cdc.gov/healthyyouth/data/yrbs/questionnaires.htm.

months after the birth of their child. Using this question, the evaluation team constructed a binary (yes/no) variable indicating whether a woman received a postpartum checkup.

6. Continued high school attainment

The evaluation team constructed a single primary outcome measuring the mother's education attainment. The measure used data from three survey items:

- Are you currently enrolled in any type of school or education program? If you are currently on summer vacation, a semester break or taking a short break to have your baby but plan to return to school, please answer "yes."
- Are you enrolled in a GED program or a post high school vocational training program?
  - Yes, a GED program
  - Yes, a vocational training program (post high school)
  - No
- Do you have any of these degrees or certificates?
  - A high school diploma
  - A GED
  - A certificate or license, for example, from a vocational training program
  - An associate degree from a two-year college or community college
  - A bachelor's degree from a four-year college
  - None of these

These questions were adapted from the survey used in the Evaluation of the School Dropout Demonstration Assistance Program (SDDAP).

The evaluation team constructed a binary (yes/no) indicator that was equal to 1 if the mother was enrolled in school, enrolled in a GED program, or had a high school diploma or GED. If a mother skipped all of the questions, the evaluation team excluded her from the analysis of this outcome.

7. Positive attitudes towards healthy relationships

We developed a single item measuring mothers' attitudes toward healthy relationships. The outcome measure was created using the following survey item, in which respondents were asked how much they strongly disagree, disagree, agree or strongly agree with each statement, with agree being the highest value.

The statements were:

- In a good relationship, you don't always get your own way
- There are times when hitting or pushing is okay in a relationship
- A good relationship is based on mutual respect, not just sex
- People who make their own partner jealous deserve to be hit or pushed
- It would be easy to trust your partner, even when you're apart
- Avoiding a disagreement with your partner is always better than talking about your problem

This survey item was adapted from a very similar measure used with pregnant and parenting adolescents on the evaluation of the Personal Responsibility Education Program. Each statement was given a rating of 1 through 4, with larger numbers indicating more correct responses. The study team reverse-coded the second, fourth, and sixth statements in the series. We then calculated the average rating across the six statements. Thus, the final outcome variable ranged from 1 to 4, with higher values indicating more positive attitudes toward healthy relationships.

4. Exploratory outcomes

We also estimated impacts on exploratory outcomes, which are outcomes that are either less directly related to the AFLP-PYD logic model, or are long-term outcomes may not yet have been affected at the time of the 12-month survey. Table A.7 provides details about these exploratory outcomes.

Table A.7. Exploratory outcomes

Outcome	How the outcome was defined
Improved education and employment outcomes	
Enrolled in a postsecondary institution	Binary variable: One if the respondent was enrolled in a post-secondary institution at the time of the 12-month survey. Respondents were excluded from the analysis of this outcome if they responded to the follow-up survey prior to July of the year in which they turned 18 (for respondents with birthdays in August or earlier) or 19 (for respondents with birthdays in September or later).
Employed post–high school	Binary variable: One if the respondent reported having a high school diploma or a GED and being employed at the time of the 12-month survey or in the 12 months prior. Respondents were excluded from the analysis of this outcome if they responded to the follow-up survey prior to July of the year in which they turned 18 (for respondents with birthdays in August or earlier) or 19 (for respondents with birthdays in September or later)
Delay of subsequent pregnancy and reduction in health risk behaviors	
Subsequent pregnancy	Binary variable: One if respondent reported having another pregnancy since the birth of the focal child
Wants to wait at least three years before having another child	Binary variable: One if the respondent reported wanting to wait at least three years until having another child or does not wish to have another child

Outcome	How the outcome was defined
Improved health and well-being of parent and child	
Has an STD	Binary variable: One if the respondent indicated having been told they have a sexually transmitted disease in the past year. Respondents were excluded from the analysis of this outcome if they indicated in their baseline survey response that they had been diagnosed with a chronic STD (defined as herpes, human papilloma virus, or HIV/AIDS)
Had as many sick child visits as mother wanted	Binary variable: One if the respondent reported that the focal child when sick was seen by a doctor or nurse as many times as the mother wanted
Number of well visits for child	Count variable: The number of regular checkups the child received
Child has health insurance	Binary variable: One if the respondent reported that the child has health insurance

E. Analysis Methods

1. Impact analysis

The individual-level study and site-level study required different analytic methods due to their different designs.

In the individual-level study, we used a straightforward ordinary least squares (OLS) model, based on the following equation, to estimate impacts of the intervention:

$$(1)\quad y_i = \alpha^{Indiv} + \tau^{Indiv} * T_i + \beta^{Indiv} * X_i + \varepsilon_i$$

Here, y_i is an outcome for individual i. The estimated impact of AFLP-PYD is represented by the coefficient τ on the indicator variable T_i, which equals one for AFLP-PYD group participants and zero for AFLP group participants. To improve the precision of our impact estimates we included X_i, a vector of individual-level covariates that includes program site, pregnancy/parenting status at time of random assignment, and all baseline measures (see Table B.1). Finally, ε_i is an individual-level error term with a mean of zero. We estimated heteroskedasticity-robust standard errors (White, 1980).

In the site-level study, we replaced OLS with a random effects maximum likelihood model, using Stata's "mixed" command, based around the following specification:

$$(2)\quad y_{ij} = \alpha^{Cluster} + \tau^{Cluster} * T_j + \beta^{Cluster} * X_{ij} + \mu_j + \varepsilon_{ij}$$

The variables in equation (2) have the same meanings as in equation (1), except now every individual i is associated with a site j, and program group status (AFLP-PYD versus AFLP) is determined at the site level. Because this model is estimated separately from the individual-level study model, the coefficient values are indexed with the superscript "Cluster", since they may

differ from the individual-level study estimates. Additionally, in a random effects model, there is assumed to be a random site-specific effect, μ_j, which is drawn from a probability distribution independent of the other variables. Since program group status was randomly assigned at the site level, the site-specific effects should be uncorrelated with that status, allowing us to estimate a treatment effect from results in these locations.

Although we analyzed the two substudies separately, we combined the impact estimates and standard errors using an inverse variance weighted average to obtain results for the full sample. Weights were obtained using a priori variance estimates and the actual sample sizes for the two substudies. This led to slightly greater weight being placed on the individual-level study results, with approximately 60 percent of the weight placed on these results and 40 percent on the site-level study.

For both outcome and control variables we dealt with missing responses (including for individuals who responded to the baseline survey, but not the follow-up survey) using multiple imputation by chained equations. This method uses non-missing response data to predict the values of missing variables (Rubin, 1987). We used an imputation algorithm called predictive mean matching, or PMM. For an observation that is missing the value of some variable *Z*, PMM uses non-missing data for other variables to find several similar-looking observations with non-missing values of *Z*, and then randomly chooses one of them to "donate" its value to the first observation.

Multiple imputation works best when there are no systematic reasons for data to be missing. This could be the case if, for instance, a random subset of respondents is chosen to answer additional questions, as with some of our follow-up questions about attitudes and resiliency. It could be violated if certain groups are systematically less likely to respond to a question—such as if especially low-performing students were less likely to report their grades. As a check, we also analyze our results using several other methods, described in Appendix B below.

2. Baseline equivalence analysis

In assessing baseline equivalence between the AFLP-PYD and AFLP groups, we used regression models similar to those described in equations (1) and (2), except that now the variable y represents a baseline variable for which equivalence is being assessed and the vector X includes only variables used to stratify the sample—program site and pregnancy/parenting status at time of random assignment. We did not use multiple imputation when assessing baseline equivalence. Instead, we dropped observations that were missing the baseline variable for which equivalence was being assessed.

3. BASIE (BAyeSian Interpretation of Estimates)

For nearly 100 years, the concept of statistical significance has been used to draw attention to some findings over others. However, misinterpretations of statistical significance—and *p*-values—are so widespread that in 2016 the American Statistical Association issued a statement

on the subject (Wasserstein & Lazar, 2016). Although readers often think statistical significance means that there is a high probability an intervention had an effect, that conclusion may be incorrect and can lead to serious misinterpretation of study findings. Similarly, a lack of statistical significance does not necessarily mean that there is a low probability an intervention had an effect.

By contrast, BASIE (Deke and Finucane 2019) avoids *p*-value misinterpretation by calculating the probability that an intervention truly had a favorable effect, given an impact estimate and prior evidence about how common it is for interventions to have meaningful effects. This probability is called a *Bayesian posterior probability*. A common concern with Bayesian methods is that they can be subjective—instead of drawing on prior *evidence*, they sometimes rely on prior *beliefs* about the effects of an intervention (de Finetti, 1974; Kaplan, 2019). The BASIE framework avoids this concern by drawing on prior evidence rather than the researcher's beliefs about the intervention's effects, and by defining probability in terms of relative frequencies, not personal belief.

To calculate the probability that an intervention truly had a meaningful effect under the BASIE framework, a researcher needs to know (1) the impact estimate and standard error for the intervention that was evaluated and (2) how common it is for *generally similar* interventions to have meaningful effects. The prior evidence tells us how common it is to achieve effects of different magnitudes, such as how common it is to achieve positive effects or effects greater than 0.20 standard deviations. Effect estimates that are similar to the prior evidence are judged to be more credible; effect estimates that are very different are deemed less credible.

To form the prior distribution for this study, we used evidence gathered from moderate- or high-quality studies reviewed for the Home Visiting Evidence of Effectiveness (HomVEE) review. This large, rigorous, systematic review provides effectiveness evidence for interventions serving disadvantaged populations of new parents. The prior evidence shows that the effects of home visiting programs tend to be small—we estimate that about 90 percent of effects are smaller in magnitude than 0.17 standard deviations. Specifically, based on our meta-analysis of the HomVEE database, for all outcomes, we generated out posterior probabilities using a prior distribution that is normal with a mean effect size 0.01 and a standard deviation of 0.10 effect size units for any given outcome.

Bayesian posterior probabilities are calculated using Bayes' Rule. Bayes' Rule is shown in equation 3. Equation 3 shows the posterior density function of a true parameter θ conditional on the data y, where P is a probability density function. In this equation, the prior is $P(\theta)$, the likelihood is $P(y|\theta)$, and $P(y)$ is a normalizing constant (which does not depend on θ). In the context of an experimental impact study, the parameter of interest is the true effect and the data are summarized by the impact estimate and its standard error.

$$(3) \quad P(\theta|y) = \frac{P(\theta)P(y|\theta)}{P(y)}$$

Appendix B. Supplemental Analyses

This appendix presents results from supplemental analyses for the impact evaluation of the Positive Youth Development (AFLP-PYD) Program for Expectant and Parenting Teens in California. The first section presents findings from assessing whether the AFLP-PYD and AFLP groups were similar at baseline. The second section present findings from sensitivity analyses.

A. Baseline Equivalence

In addition to analyzing the impacts of AFLP-PYD relative to AFLP, we compared the two participant groups prior to the intervention, to check for systematic differences between them. While we expect the groups to be identical due to random assignment, we checked this by assessing the similarity between the two groups at the start of the program. In the site-level study (Table B.3), no statistically significant differences were detected on any of the baseline measures, but in the individual study (Table B.2), three measures had statistically significant differences: Spanish language preference, knowledge of IUDs, and an indicator for having recently attended classes on relationships/dating/marriage. The statistical significance of the difference in knowledge of IUDs carried through to the combined sample (Table B.1).

The effect size column in Tables B.1 through and B.3 shows the magnitude of the difference between the treatment and control groups, in terms of the standard deviation of the variable shown in each row. All effect size differences between the two groups were less than 0.25 standard deviations.

Table B.1. Baseline Equivalence for Overall Study

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Age at random assignment (years)	17.4	17.3	0.0	0.78	0.03	1,298
Race/ethnicity (percent)				0.66		
Hispanic	89	84	5		0.14	1,321
Black non-Hispanic	4	3	1		0.03	1,321
White non-Hispanic	4	5	-1		-0.04	1,321
Other race non-Hispanic	3	5	-2		-0.12	1,321
Spanish language preference	7	9	-3	0.51	-0.09	1,330
Highest grade completed at random assignment (percent)				0.58		
Grade 7 or lower	1	2	-1		-0.09	1,294
Grade 8	8	10	-2		-0.07	1,294

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Grade 9	14	12	2		0.05	1,294
Grade 10	24	27	-3		-0.07	1,294
Grade 11	38	38	0		0.01	1,294
Grade 12	14	10	4		0.14	1,294
Enrolled in school or education program at time of random assignment (percent)	86	85	1	0.78	0.02	1,297
Ever repeated a grade in middle or high school (percent)	7	9	-2	0.43	-0.06	1,300
Ever been suspended or expelled from school (percent)	40	45	-5	0.22	-0.11	1,297
Pregnant at time of random assignment (percent)	46	46	0	0.95	0.00	1,297
Received information or classes in the past year (percent)						
Classes on relationships and dating	8	11	-3	0.37	-0.09	1,294
Classes on parenting	45	45	0	0.94	0.01	1,294
Classes on health care for mothers	59	59	0	0.98	0.00	1,295
Classes on health care for children	65	66	-1	0.79	-0.02	1,295
Number of education services received in the past year (out of six)	0.96	0.96	0.00	0.97	0.00	1,296
Number of employment services received in the past year (out of four)	0.41	0.40	0.01	0.82	0.02	1,296
Resiliency (scale 1–4)	3.26	3.26	0.00	0.93	-0.01	1,299
Presence of a trusted adult (scale 1–4)	3.46	3.47	-0.01	0.82	-0.02	1,300
Used birth control in first sexual encounter (percent)	67	65	2	0.59	0.05	1,294
Share of questions on birth control knowledge answered correctly						
Knowledge of condoms[a]	68	69	-1	0.61	-0.04	1,297
Knowledge of birth control pills[a]	54	56	-2	0.52	-0.06	1,297
Knowledge of IUDs[a]	40	44	-4	0.09	-0.14	1,297

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Knowledge of other birth control methods[a]	41	42	-1	0.52	-0.04	1,329
Had depressive symptoms in past 12 months (percent)	18	16	2	0.52	0.04	1,298
Healthy relationships beliefs/attitudes (scale 1–4)	3.30	3.31	-0.01	0.66	-0.04	1,299
Currently has an STD (percent)	11	12	-1	0.58	-0.04	1,293

[a] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

Table B.2. Baseline Equivalence for Individual-Level Study Participants

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Age at random assignment (years)	17.5	17.4	0.1	0.53	0.09	392
Race/ethnicity (percent)				0.42		
Hispanic	93	90	3		0.04	392
Black non-Hispanic	1	1	1		0.10	392
White non-Hispanic	2	2	0		0.04	392
Other race non-Hispanic	2	5	-3		-0.13	392
Spanish language preference	4	8	-4	0.07	-0.16	401
Highest grade completed at random assignment (percent)				0.41		
Grade 7 or lower	1	2	-1		-0.09	389
Grade 8	6	9	-3		-0.10	389
Grade 9	12	11	2		0.04	389
Grade 10	23	28	-5		-0.13	389
Grade 11	41	39	2		0.05	389
Grade 12	17	12	5		0.15	389
Enrolled in school or education program at time of random assignment (percent)	84	83	1	0.73	0.03	392
Ever repeated a grade in middle or high school (percent)	6	9	-2	0.40	-0.06	392
Ever been suspended or expelled from school (percent)	41	47	-6	0.22	-0.12	391
Pregnant at time of random assignment (percent)	43	40	3	0.18	-0.03	392
Received information or classes in the past year (percent)						

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Classes on relationships and dating	7	13	-6	0.07	-0.19	392
Classes on parenting	46	46	0	0.96	0.01	392
Classes on health care for mothers	61	59	3	0.58	0.07	392
Classes on health care for children	65	66	-1	0.81	0.00	392
Number of education services received in the past year (out of six)	0.96	0.97	-0.01	0.97	-0.02	392
Number of employment services received in the past year (out of four)	0.43	0.40	0.02	0.81	0.02	392
Resiliency (scale 1–4)	3.28	3.25	0.03	0.38	0.10	392
Presence of a trusted adult (scale 1–4)	3.47	3.45	0.01	0.78	0.03	392
Used birth control in first sexual encounter (percent)	67	66	1	0.80	0.02	391
Share of questions on birth control knowledge answered correctly						
Knowledge of condoms[a]	70	71	-1	0.72	-0.04	392
Knowledge of birth control pills[a]	57	60	-3	0.29	-0.09	392
Knowledge of IUDs[a]	42	48	-7	0.02	-0.22	392
Knowledge of other birth control methods[a]	43	45	-2	0.51	-0.06	401
Had depressive symptoms in past 12 months (percent)	18	16	3	0.50	0.05	392
Healthy relationships beliefs/attitudes (scale 1–4)	3.33	3.32	0.01	0.86	0.02	392
Currently has an STD (percent)	10	12	-1	0.64	-0.05	392

[a] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

Table B.3. Baseline Equivalence for Site-Level Study Participants

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Age at random assignment (years)	17.2	17.2	0.0	0.89	-0.12	906
Race/ethnicity (percent)				0.72		906
Hispanic	83	75	8		0.07	929
Black non-Hispanic	7	7	0		0.16	929

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
White non-Hispanic	6	9	-3		-0.17	929
Other race non-Hispanic	3	5	-2		-0.01	929
Spanish language preference	11	11	0	0.98	-0.02	929
Highest grade completed at random assignment (percent)				0.88		905
Grade 7 or lower	1	2	-1		-0.07	905
Grade 8	11	13	-2		-0.02	905
Grade 9	16	14	2		0.02	905
Grade 10	27	27	0		0.00	905
Grade 11	35	37	-2		0.00	905
Grade 12	10	7	2		0.04	905
Enrolled in school or education program at time of random assignment (percent)	87	88	0	0.96	0.03	905
Ever repeated a grade in middle or high school (percent)	9	10	-1	0.77	-0.12	908
Ever been suspended or expelled from school (percent)	39	43	-4	0.59	0.05	906
Pregnant at time of random assignment (percent)	50	55	-4	0.59	-0.01	905
Received information or classes in the past year (percent)						
Classes on relationships and dating	10	8	2	0.70	0.21	902
Classes on parenting	45	44	1	0.82	0.04	902
Classes on health care for mothers	56	60	-4	0.40	-0.04	903
Classes on health care for children	64	65	0	0.90	-0.01	903
Number of education services received in the past year (out of six)	0.95	0.94	0.00	0.99	0.00	904
Number of employment services received in the past year (out of four)	0.39	0.39	0.00	0.96	0.05	904
Resiliency (scale 1–4)	3.23	3.28	-0.05	0.28	-0.07	907
Presence of a trusted adult (scale 1–4)	3.45	3.49	-0.04	0.55	0.00	908
Used birth control in first sexual encounter (percent)	66	63	4	0.62	0.02	903

Measure	AFLP-PYD mean	AFLP mean	Difference	*p*-value	Effect size	N
Share of questions on birth control knowledge answered correctly						
Knowledge of condoms[a]	64	65	-1	0.72	-0.10	905
Knowledge of birth control pills[a]	50	50	0	0.99	-0.03	905
Knowledge of IUDs[a]	37	37	0	1.00	0.02	905
Knowledge of other birth control methods[a]	37	37	0	0.90	-0.02	928
Had depressive symptoms in past 12 months (percent)	17	16	0	0.92	0.01	906
Healthy relationships beliefs/attitudes (scale 1–4)	3.25	3.30	-0.04	0.49	-0.07	907
Currently has an STD (percent)	12	12	-1	0.77	-0.02	901

[a] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

B. Sensitivity Analyses

We conducted three types of sensitivity analyses. First, we assessed the robustness of our findings to changes in analytical methods and outcome definitions. Second, we explored potential mechanisms or drivers of our results. Third, we considered two alternate Bayesian interpretations of our evidence.

1. Alternate analysis methods and outcome definitions

As described in Chapter III, the main impact analysis accounted for missing data using multiple imputation. We re-estimated the impacts of AFLP-PYD using several alternative methods. These methods are described in Table B.4 below. The results from each of these sensitivity analyses are shown in Table B.5. Each row in Table B.4 corresponds to a column in Table B.5, in the order shown.

Table B.4. Sensitivity Analyses Conducted

Sensitivity analysis	Description
Complete case analysis	Regression equations included only individuals who provided answers to all relevant baseline and follow-up survey items.
Multiple imputation of covariates only	Missing data for baseline covariates were filled in using multiple imputation. For outcome variables, only actual survey responses were used. Any individuals lacking responses to the relevant follow-up survey items were excluded from the analysis.
Missing covariate data back-filled and indicated	For observations missing baseline covariate data, missing values were "back-filled" with zero, and a binary variable was included in the regression corresponding to each covariate, indicating when an observation had been back-filled. Observations were excluded from the analysis of any outcomes for which the respondent did not provide data, whether because of non-response or in the case of the resiliency items because a question was not presented to the respondent.
No covariates	Results were calculated without regression controls for baseline covariates. The only covariates included in the regressions were treatment status and randomization stratum. Observations were excluded from the analysis of any outcomes for which the respondent did not provide data, whether because of non-response or in the case of the resiliency items because a question was not presented to the respondent.
Site-level data	For the site-level study sample, this analysis averages all individuals' responses within the same site, generating a single "composite" observation from each randomly assigned site. This approach has been shown to reduce the likelihood of "false positive" results, and weights each site equally regardless of the number of participants it contributes. We then analyze the resulting 13 observations (one for each site) in a single OLS regression, using heteroskedasticity-robust standard errors as in the main analysis. Because of the reduced number of observations, we include only the treatment indicator and a single additional baseline control variable. We selected a different control variable for each outcome, choosing the baseline variable that was most closely conceptually related to the outcome. (Deke 2016). This method affects only the results in the site-level study. These new results are averaged with the main analysis results from the individual-level study to create the combined impact estimate for the study overall.

Our findings are robust to different analytical approaches. In some instances, the alternate approaches yielded more statistically significant results, largely varying with the method of handling missing data. However, the impact estimates and Bayesian posterior probabilities are similar across all sensitivity analyses. Because the sensitivity analysis of using site-level data affects only the site-level substudy, for additional reference, the last column of Table B.6 presents results from that analysis for the site-level substudy only, compared to the site-level results from the main analysis. Bayesian posteriors may differ across specifications and

substudies due to differences in impact point estimates and standard errors, or correlations between impact estimates.

Table B.5. Results using alternate analysis methods

Outcome measure	Main analysis	Complete case	Covariate imputation only	Back-filled covariates	No covariates	Site-level data
Program delivery measures (percent unless otherwise noted)						
Number of meetings with a case manager	3.31*** (99%)	3.29*** (98%)	3.31*** (99%)	3.22*** (98%)	2.71*** (>99%)	3.24*** (98%)
Withdrew from MCAH programming	9** (2%)	8** (3%)	9** (2%)	9** (2%)	11*** (<1%)	9** (<1%)
Number of education-related services received (out of six)	-0.11 (6%)	-0.11 (8%)	-0.12* (5%)	-0.13* (4%)	-0.11 (4%)	-0.12 (2%)
Number of employment-related services received (out of four)	-0.16** (1%)	-0.17** (2%)	-0.17** (2%)	-0.18** (1%)	-0.15** (1%)	-0.15* (<1%)
Attended a class about relationships, dating, or marriage[a]	-1 (11%)	-1 (16%)	-1 (12%)	-1 (12%)	-1 (10%)	-1 (8%)
Attended a class about parenting[a]	-3 (7%)	-3 (14%)	-3 (9%)	-4 (6%)	-2 (6%)	-3 (5%)
Attended a class about health care for themselves[b]	-6* (2%)	-7** (2%)	-7** (2%)	-7** (1%)	-5 (3%)	-8** (<1%)
Attended a class about health care for baby[b]	-4 (8%)	-5 (10%)	-5 (8%)	-5* (7%)	-4 (6%)	-2 (12%)
Received information about methods of birth control or where to get birth control[c]	-2 (8%)	-2 (12%)	-1 (11%)	-2 (9%)	-2 (6%)	-2 (5%)
Short-term outcomes						
Feelings about oneself[d]	-0.05 (8%)	0.03 (29%)	-0.01 (15%)	-0.03 (15%)	-0.06 (7%)	-0.14 (3%)
Feelings about goals[d]	0.03 (25%)	-0.04 (18%)	-0.03 (14%)	-0.05 (13%)	0.05 (22%)	-0.01 (18%)
Number of linkages and referrals received (out of four)	-0.25*** (<1%)	-0.29*** (<1%)	-0.28*** (<1%)	-0.29*** (<1%)	-0.23** (<1%)	-0.24** (<1%)
Presence of trusted adult/supportive relationships[d]	0.08 (29%)	0.11* (74%)	0.09 (55%)	0.10 (61%)	0.06 (16%)	0.08 (43%)
Early indicators of long-term outcomes (percent unless otherwise noted)						
Had unprotected sex in past 3 months	5 (2%)	7* (4%)	5 (4%)	6* (3%)	3 (3%)	5 (<1%)

Outcome measure	Main analysis	Complete case	Covariate imputation only	Back-filled covariates	No covariates	Site-level data
Currently using birth control	-1 (13%)	-2 (18%)	-1 (14%)	-1 (14%)	-1 (9%)	-2 (8%)
Knowledge of condoms[e]	-2 (5%)	-1 (17%)	-1 (9%)	-1 (13%)	-1 (3%)	-2 (2%)
Knowledge of birth control pills[e]	0 (25%)	0 (36%)	0 (25%)	0 (29%)	0 (15%)	0 (28%)
Knowledge of IUDs[e]	-4* (2%)	-4* (6%)	-3 (4%)	-3 (6%)	-3 (3%)	-4* (1%)
Knowledge of other birth control methods[e]	-2 (8%)	-1 (13%)	-2 (10%)	-1 (10%)	-1 (10%)	-2 (7%)
Had depressive symptoms in past 12 months	-1 (22%)	-2 (40%)	-1 (26%)	-1 (25%)	-1 (15%)	0 (20%)
Has health insurance	-2 (12%)	-2 (21%)	-2 (17%)	-2 (17%)	-2 (9%)	-3 (10%)
Had check-up in the 3 months after child's birth	3 (25%)	3 (42%)	3 (34%)	3 (37%)	4* (28%)	3 (34%)
Enrolled in high school or earned a diploma/GED	-3 (4%)	-3 (6%)	-3 (4%)	-3 (4%)	-2 (4%)	-1 (10%)
Positive attitudes toward healthy relationships[f] (scale 1–4)	-0.03 (5%)	-0.03 (6%)	-0.02 (8%)	-0.03 (5%)	-0.02 (7%)	-0.01 (13%)
Exploratory outcomes (percent unless otherwise noted)						
Enrolled in a postsecondary institution	-4 (6%)	-5 (8%)	-4 (8%)	-4 (7%)	0 (13%)	-5 (6%)
Employed post-high school	2 (24%)	3 (43%)	3 (29%)	3 (32%)	5 (31%)	0 (20%)
Subsequent pregnancy	-1 (18%)	0 (24%)	-1 (21%)	-1 (22%)	-1 (17%)	-1 (22%)
Wants to wait at least three years before having another child	-2 (12%)	-3 (15%)	-2 (11%)	-2 (12%)	-1 (16%)	-2 (13%)
Has an STD	-1 (18%)	-2 (26%)	-1 (20%)	-1 (20%)	0 (19%)	0 (20%)
Had as many sick child visits as mother wanted	0 (15%)	0 (25%)	1 (17%)	1 (19%)	0 (16%)	0 (9%)
Number of well visits for child	-0.32 (8%)	-0.31 (11%)	-0.33 (9%)	-0.44 (5%)	-0.29 (10%)	-0.40 (9%)
Child has health insurance	-3 (7%)	-3 (9%)	-3 (7%)	-3 (7%)	-3 (8%)	-3 (6%)

Note: Probability outcome is favorable to AFLP-PYD is displayed in parentheses under impact estimate.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[d] This outcome is a factor variable that is roughly standard normal (mean = 0, standard deviation = 1). Positive values indicate more "resilient" responses.

[e] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

[f] This outcome was calculated as the average value across six survey items about healthy relationships. Each item was rated 1 through 4, with larger numbers indicating more correct responses.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

Table B.6. Results using site-level data for the site-level study

	Site-level study	
Outcome measure	**Main analysis**	**Sensitivity analysis: Site-level data**
Program delivery measures (percent unless otherwise noted)		
Number of meetings with a case manager	3.52* (74%)	3.35 (73%)
Withdrew from MCAH programming	5 (<1%)	7 (<1%)
Number of education-related services received (out of six)	-0.08 (<1%)	-0.10 (<1%)
Number of employment-related services received (out of four)	-0.15 (<1%)	-0.12 (2%)
Attended a class about relationships, dating, or marriage[a]	3 (75%)	2 (68%)
Attended a class about parenting[a]	-6 (<1%)	-6 (<1%)
Attended a class about health care for themselves[b]	2 (50%)	-4 (<1%)
Attended a class about health care for baby[b]	1 (77%)	5 (>99%)
Received information about methods of birth control or where to get birth control[c]	-2 (<1%)	-2 (<1%)

Outcome measure	Site-level study	
	Main analysis	Sensitivity analysis: Site-level data
Short-term outcomes		
Feelings about oneself[d]	-0.17 (6%)	-0.37** (<1%)
Feelings about goals[d]	0.01 (51%)	-0.08 (<1%)
Number of linkages and referrals received (out of four)	0.00 (71%)	0.01 (96%)
Presence of trusted adult/supportive relationships[d]	0.10 (97%)	0.12 (98%)
Early indicators of long-term outcomes and exploratory outcomes (percent unless otherwise noted)		
Had unprotected sex in past 3 months	-2 (47%)	-3 (49%)
Currently using birth control	-3 (<1%)	-6 (<1%)
Knowledge of condoms[e]	-3 (<1%)	-4 (<1%)
Knowledge of birth control pills[e]	-1 (29%)	-1 (7%)
Knowledge of IUDs[e]	-2 (25%)	-4 (25%)
Knowledge of other birth control methods[e]	-3 (18%)	-4 (5%)
Had depressive symptoms in past 12 months	-2 (93%)	0 (14%)
Has health insurance	-2 (13%)	-3 (12%)
Had check-up in the 3 months after child's birth	-2 (1%)	0 (9%)
Enrolled in high school or earned a diploma/GED	-6* (<1%)	-2 (26%)
Positive attitudes toward healthy relationships[f] (scale 1–4)	-0.04 (<1%)	-0.01 (47%)
Enrolled in a postsecondary institution	-3 (<1%)	-4 (<1%)
Employed post-high school	0 (40%)	-3 (30%)
Subsequent pregnancy	-1 (23%)	-1 (24%)
Wants to wait at least three years before having another child	2 (83%)	2 (92%)

Outcome measure	Site-level study	
	Main analysis	Sensitivity analysis: Site-level data
Has an STD	-1 (38%)	1 (<1%)
Had as many sick child visits as mother wanted	1 (6%)	1 (<1%)
Number of well visits for child	-0.90** (<1%)	-1.08*** (<1%)
Child has health insurance	-3 (<1%)	-2 (<1%)

Note: Probability outcome is favorable to AFLP-PYD is displayed in parentheses under impact estimate.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[d] This outcome is a factor variable that is roughly standard normal (mean = 0, standard deviation = 1). Positive values indicate more "resilient" responses.

[e] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

[f] This outcome was calculated as the average value across six survey items about healthy relationships. Each item was rated 1 through 4, with larger numbers indicating more correct responses.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

We also calculated impact estimates using alternative definitions for certain outcomes (results not shown). Specifically, we calculated impacts on birth control use separately for three types of birth control (condoms, long-acting reversible contraception, and hormonal birth control methods). We also examined high school enrollment and GED outcomes separately instead of combined together into a single outcome variable, as in the main analysis. Finally, as mentioned in Chapter IV, we relaxed the restriction on program delivery measures that information or services be delivered or recommended by a participant's case manager. None of these sensitivity analyses yielded results that differed meaningfully or substantially from our main results.

2. Potential drivers of results

In addition to the alternate analysis methods mentioned above, we conducted other sensitivity tests to help probe potential mechanisms or drivers of our results. First, we examined whether impacts of AFLP-PYD were different for sites that implemented the program with high versus low fidelity. MCAH provided information on the technical assistance needs among the sites,

both in the AFLP-PYD and AFLP groups and at the sites with individual random assignment. We then calculated three separate AFLP-PYD impacts for high-, medium-, and low-technical assistance need sites (corresponding to low, medium, and high program fidelity, respectively). To calculate these impacts, we created three fidelity-level indicators, and interacted each with the treatment indicator in a single estimating equation. In other words, the high-fidelity group impact estimates compare participants at the highest-fidelity AFLP-PYD program sites against those in the highest-fidelity AFLP sites, and likewise among the medium- and low-fidelity sites. Table B.7 shows the results of this analysis—only the high-fidelity impact estimates are displayed, but the estimates for the other two groups are similar. Results for the high-fidelity sites were similar to those from the main analysis.

Second, we examined impacts separately for mothers that received high, medium, and low dosages of the AFLP-PYD or AFLP program. If some AFLP-PYD or AFLP respondents dropped out of the program early or missed a substantial number of visits with their case manager then the program impacts could have differed among those with a high "dose" of service versus those with a medium or low dose. We used individual-level data on how long each participant stayed enrolled in their program to split the sample into three dosage groups of roughly equal size, corresponding to the group enrolled for the shortest time, the longest time, and an intermediate group. We then calculated separate impacts in each group. In both the AFLP-PYD and AFLP groups, more than one third of participants continued participating in the program for 12 months after enrollment, so the high-dosage impact estimate shows the differences between those participants who participated in AFLP-PYD for a full year or more against those who participated for a year or more in AFLP.[16] As shown in Table B.7, even when looking only at these high-dosage participants our results are similar to those in the main analysis. This suggests that the results we find in the main sample are not due to systematic differences in program dosage across the two groups of participants.

Third, we calculated impacts excluding the Alta Med East LA site. According to MCAH, this site had particularly high fidelity in implementing both AFLP-PYD and AFLP. Results excluding this site were similar to results from the main analysis (Table B.7).

Table B.7. Results from additional sensitivity analyses

Outcome measure	Main analysis	High-fidelity sites	High-dosage participants	Exclude Alta Med East LA
Program delivery measures (percent unless otherwise noted)				
Number of meetings with a case manager	3.31*** (99%)	3.07*** (99%)	6.55*** (>99%)	3.35*** (98%)
Withdrew from MCAH programming	9** (2%)	8 (3%)	1 (32%)	9* (10%)

[16] Our dosage results stayed substantively the same when dividing the sample into dosage quartiles instead of terciles, or when replacing months enrolled with the cumulative hours spent meeting with a case manager as a measure of program dosage.

Outcome measure	Main analysis	High-fidelity sites	High-dosage participants	Exclude Alta Med East LA
Number of education-related services received (out of six)	-0.11 (6%)	-0.29* (3%)	-0.17 (15%)	-0.07 (17%)
Number of employment-related services received (out of four)	-0.16** (1%)	-0.30** (1%)	-0.22** (3%)	-0.14* (7%)
Attended a class about relationships, dating, or marriage[a]	-1 (11%)	1 (42%)	0 (37%)	0 (19%)
Attended a class about parenting[a]	-3 (7%)	-10 (7%)	-5 (14%)	0 (19%)
Attended a class about health care for themselves[b]	-6* (2%)	-12* (4%)	-7* (11%)	-3 (11%)
Attended a class about health care for baby[b]	-4 (8%)	-5 (27%)	-5 (26%)	-1 (17%)
Received information about methods of birth control or where to get birth control[c]	-2 (8%)	-2 (17%)	-1 (24%)	-1 (17%)
Short-term outcomes				
Feelings about oneself[d]	-0.05 (8%)	-0.12 (16%)	-0.03 (37%)	-0.04 (17%)
Feelings about goals[d]	0.03 (25%)	-0.03 (20%)	0.06 (66%)	0.03 (22%)
Number of linkages and referrals received (out of four)	-0.25*** (<1%)	-0.50*** (<1%)	-0.21* (13%)	-0.15 (9%)
Presence of trusted adult/supportive relationships[d]	0.08 (29%)	0.03 (27%)	0.12 (85%)	0.10 (34%)
Early indicators of long-term outcomes (percent unless otherwise noted)				
Had unprotected sex in past 3 months	5 (2%)	10* (3%)	6 (22%)	3 (11%)
Currently using birth control	-1 (13%)	-7 (4%)	-4 (38%)	1 (26%)
Knowledge of condoms[e]	-2 (5%)	2 (58%)	1 (70%)	-3* (4%)
Knowledge of birth control pills[e]	0 (25%)	0 (34%)	1 (69%)	0 (22%)
Knowledge of IUDs[e]	-4* (2%)	-6 (3%)	-3 (19%)	-2 (15%)
Knowledge of other birth control methods[e]	-2 (8%)	-4 (5%)	0 (57%)	-1 (22%)
Had depressive symptoms in past 12 months	-1 (22%)	-2 (59%)	-2 (70%)	-1 (23%)
Has health insurance	-2 (12%)	-10** (4%)	0 (53%)	2 (26%)

Outcome measure	Main analysis	High-fidelity sites	High-dosage participants	Exclude Alta Med East LA
Had check-up in the 3 months after child's birth	3 (25%)	-2 (16%)	1 (59%)	7** (35%)
Enrolled in high school or earned a diploma/GED	-3 (4%)	-3 (5%)	4 (65%)	-1 (12%)
Positive attitudes toward healthy relationships[f] (scale 1–4)	-0.03 (5%)	-0.02 (15%)	0.00 (53%)	-0.02 (13%)
Exploratory outcomes (percent unless otherwise noted)				
Enrolled in a postsecondary institution	-4 (6%)	-10 (14%)	-3 (46%)	-4 (17%)
Employed post-high school	2 (24%)	-4 (22%)	3 (65%)	6 (32%)
Subsequent pregnancy	-1 (18%)	2 (28%)	-2 (71%)	-5 (31%)
Wants to wait at least three years before having another child	-2 (12%)	-3 (16%)	4 (82%)	-2 (23%)
Has an STD	-1 (18%)	-4 (34%)	-1 (58%)	-3 (26%)
Had as many sick child visits as mother wanted	0 (15%)	2 (29%)	1 (64%)	0 (24%)
Number of well visits for child	-0.32 (8%)	0.16 (45%)	-0.12 (51%)	-0.55 (13%)
Child has health insurance	-3 (7%)	-7** (6%)	-3 (34%)	0 (22%)

Note: Probability outcome is favorable to AFLP-PYD is displayed in parentheses under impact estimate.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[d] This outcome is a factor variable that is roughly standard normal (mean = 0, standard deviation = 1). Positive values indicate more "resilient" responses.

[e] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

[f] This outcome was calculated as the average value across six survey items about healthy relationships. Each item was rated 1 through 4, with larger numbers indicating more correct responses.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

Fourth, we controlled for funding source in the regression model. Services for some mothers were supported by funds from another state program called CalLearn, which might have increased case managers' incentives to deliver information and classes, but otherwise would not have affected the services rendered. Therefore, as another sensitivity analysis we included a regression control for CalLearn funding. This additional control did not qualitatively affect our findings (results not shown).

Finally, Table B.8 presents results for the sensitivity analysis described in Chapter IV that looked only at information and service delivery outcomes in the individual-level study. Namely, we examined impacts separately for early versus later enrollers, to account for possible challenges when first implementing the program. Our results did not suggest that late enrollers had better program delivery impacts than early enrollers—if anything, the results suggest the opposite.

Table B.8. Impacts on program delivery measures for early versus later enrollers in the individual substudy

Outcome measure	Main analysis	Early enrollees only	Late enrollees only
Program delivery measures (percent unless otherwise noted)			
Number of education-related services received (out of six)	-0.13 (16%)	-0.11 (26%)	-0.15 (27%)
Number of employment-related services received (out of four)	-0.17* (3%)	-0.20 (23%)	-0.02 (39%)
Attended a class about relationships, dating, or marriage[a]	-4 (15%)	0 (35%)	-8 (29%)
Attended a class about parenting[a]	-1 (35%)	5 (42%)	-9 (34%)
Attended a class about health care for themselves[b]	-10** (2%)	0 (38%)	-25*** (18%)
Attended a class about health care for baby[b]	-7* (17%)	-11 (15%)	-9 (30%)
Received information about methods of birth control or where to get birth control[c]	-2 (17%)	-1 (31%)	-4 (31%)

Note: Probability outcome is favorable to AFLP-PYD is displayed in parentheses under impact estimate.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

* Significantly different from zero at the 0.10 level, two-tailed test.

** Significantly different from zero at the 0.05 level, two-tailed test.

*** Significantly different from zero at the 0.01 level, two-tailed test.

3. Alternate interpretation of findings

In this section, we do not change the analytical methodology, but rather how we interpret the findings. First, we examine sensitivity to changing the prior distribution. Second, we compare the traditional point estimate to a Bayesian *shrunken* estimate, which is an alternative program effect estimate that considers both study data and prior evidence.

As described in Appendix A, we used evidence from the HomVEE systematic review to form a prior distribution. Here, we examine how sensitive posterior probabilities are to the width of the prior distribution by using a prior distribution informed by evidence from the Institute for Education Studies' What Works Clearinghouse (WWC). The WWC includes evidence regarding a wide range of educational interventions. Our analysis of the WWC yields a wider range of intervention effects than what we see with HomVEE. The standard deviation of intervention effects from the WWC is 0.23 effect size units, as opposed to 0.10 for HomVEE. In this sensitivity analysis, we hold the mean of the prior distribution constant at 0.01 – we only change the standard deviation of the prior distribution.

Sensitivity to the prior distribution is shown in the second column of Table B.9. Compared to our main estimate's prior, using a wider prior distribution makes impact estimates further from zero seem more credible than they otherwise would. For example, the probability of a favorable impact on Presence of trusted adult/supportive relationships is 87 percent when using the HomVEE-informed prior distribution and is 90 percent when using the wider WWC-informed prior distribution.

The shrunken impact estimate is reported in the third column of Table B.9**.** This sensitivity analysis does not affect the Bayesian posterior probability, but yields a slightly different impact estimate. This shrunken impact estimate is essentially a weighted average between the impact estimate calculated using study data and the center of the prior distribution. Because the prior distribution is centered near zero, a shrunken estimate has the effect of moderating both very positive and very negative impact estimates.

Table B.9. Sensitivity to alternative Bayesian interpretations

Outcome measure	Main analysis	Alternate prior standard deviation	Bayesian-adjusted impacts
Program delivery measures (percent unless otherwise noted)			
Number of meetings with a case manager	3.31 (99%)	3.31 (>99%)	1.29 (99%)
Withdrew from MCAH programming	9 (2%)	9 (2%)	3 (2%)
Number of education-related services received (out of six)	-0.11 (6%)	-0.11 (5%)	-0.08 (6%)
Number of employment-related services received (out of four)	-0.16 (1%)	-0.16 (1%)	-0.10 (1%)
Attended a class about relationships, dating, or marriage[a]	-1 (11%)	-1 (9%)	-2 (11%)
Attended a class about parenting[a]	-3 (7%)	-3 (7%)	-2 (7%)
Attended a class about health care for themselves[b]	-6 (2%)	-6 (2%)	-3 (2%)
Attended a class about health care for baby[b]	-4 (8%)	-4 (7%)	-2 (8%)
Received information about methods of birth control or where to get birth control[c]	-2 (8%)	-2 (7%)	-1 (8%)
Short-term outcomes			
Feelings about oneself[d]	-0.05 (8%)	-0.05 (7%)	-0.05 (8%)
Feelings about goals[d]	0.03 (25%)	0.03 (23%)	-0.03 (25%)
Number of linkages and referrals received (out of four)	-0.25 (<1%)	-0.25 (<1%)	-0.14 (<1%)
Presence of trusted adult/supportive relationships[d]	0.08 (29%)	0.08 (29%)	-0.02 (29%)
Early indicators of long-term outcomes (percent unless otherwise noted)			
Had unprotected sex in past 3 months	5 (2%)	5 (2%)	3 (2%)
Currently using birth control	-1 (13%)	-1 (12%)	-2 (13%)
Knowledge of condoms[e]	-2 (5%)	-2 (4%)	-1 (5%)
Knowledge of birth control pills[e]	0 (25%)	0 (24%)	-1 (25%)
Knowledge of IUDs[e]	-4 (2%)	-4 (2%)	-2 (2%)
Knowledge of other birth control methods[e]	-2 (8%)	-2 (6%)	-2 (8%)

Outcome measure	Main analysis	Alternate prior standard deviation	Bayesian-adjusted impacts
Had depressive symptoms in past 12 months	-1 (22%)	-1 (21%)	1 (22%)
Has health insurance	-2 (12%)	-2 (10%)	-1 (12%)
Had check-up in the 3 months after child's birth	3 (25%)	3 (26%)	-1 (25%)
Enrolled in high school or earned a diploma/GED	-3 (4%)	-3 (3%)	-2 (4%)
Positive attitudes toward healthy relationships[f] (scale 1–4)	-0.03 (5%)	-0.03 (4%)	-0.02 (5%)
Exploratory outcomes (percent unless otherwise noted)			
Enrolled in a postsecondary institution	-4 (6%)	-4 (6%)	-3 (6%)
Employed post-high school	2 (24%)	2 (23%)	-1 (24%)
Subsequent pregnancy	-1 (18%)	-1 (17%)	1 (18%)
Wants to wait at least three years before having another child	-2 (12%)	-2 (12%)	-1 (12%)
Has an STD	-1 (18%)	-1 (16%)	1 (18%)
Had as many sick child visits as mother wanted	0 (15%)	0 (15%)	-1 (15%)
Number of well visits for child	-0.32 (8%)	-0.32 (8%)	-0.28 (8%)
Child has health insurance	-3 (7%)	-3 (7%)	-1 (7%)

Note: Probability outcome is favorable to AFLP-PYD is displayed in parentheses under impact estimate.

[a] This measure is a binary indicator with a value of 1 if the mother attended a class about that topic either administered by her case manager or at a place recommended by her case manager, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[b] This measure is a normalized count of topics related to health care addressed by the case manager either directly or through a recommendation by the case manager. The variable ranges from 0 to 1, with higher values indicating exposure to more topics about health care. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[c] This measure is a binary indicator with a value of 1 if the mother reported receiving information about birth control methods or where to get birth control from a doctor, nurse, case manager, or health clinic, and 0 otherwise. It has been rescaled from 0 to 100 to show the percentage of mothers who attained this outcome.

[d] This outcome is a factor variable that is roughly standard normal (mean = 0, standard deviation = 1). Positive values indicate more "resilient" responses.

[e] This outcome ranges from 0 to 1 and has been rescaled from 0 to 100 to show the percentage of items answered correctly. The number of items was six for condoms and IUDs, five for birth control pills, and four for other birth control methods.

[f] This outcome was calculated as the average value across six survey items about healthy relationships. Each item was rated 1 through 4, with larger numbers indicating more correct responses.

References

Asheer, S., B. Keating, J. Crowley, and S. Zief. "Engaging Young Mothers as Partners in Case Management: An Implementation Study of California's Adolescent Family Life Program with Positive Youth Development." Washington, DC: Office of Population Affairs, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services, 2020.

Barnet, B. A. Joffe, A.K. Duggan M.D. Wilson, and J.T. Repke. "Depressive Symptoms, Stress, and Social Support in Pregnant and Postpartum Adolescents." *Archives of Pediatrics & Adolescent Medicine*, vol. 150, no. 1, 1996, pp. 64–69.

Catalano, R.F., M.L. Berglund, J.A.M. Ryan, H.S. Lonczak, and J.D. Hawkins. "Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs." *Annals of the American Academy of Political and Social Science*, vol. 594, no. 1, 2004, pp. 98–124.

Conde-Agudelo, A., A. Rosas-Bermúdez, and A.C. Kafury-Goeta. "Birth Spacing and Risk of Adverse Perinatal Outcomes: A Meta-analysis." *JAMA*, vol. 295, no. 15, 2006, pp. 1809-1823.

Covington, R., B. Goesling, C. Trenholm, J. Manlove, and K. Welti. "Interim Impacts of the AIM 4 Teen Moms Program." Washington, DC: Office of Adolescent Health, U.S. Department of Health and Human Services, 2015.

Dee, D.L., K. Pazol , S. Cox, R.A. Smith, K. Bower, M. Kapaya, A. Fasula, et al. "Trends in Repeat Births and Use of Postpartum Contraception among Teens—United States, 2004–2015." *Mortality and Morbidity Weekly Report*, vol. 66, no. 16, 2017, pp. 422–426. doi: http://dx.doi.org/10.15585/mmwr.mm6616a3.

Deke, John. "Design and Analysis Considerations for Cluster Randomized Controlled Trials That Have a Small Number of Clusters." Evaluation Review, vol. 40, no. 5, Oct. 2016, pp. 444–486, doi:10.1177/0193841X16671680.

Deke, J., & Finucane, M. (2019). Moving beyond statistical significance: the BASIE (BAyeSian Interpretation of Estimates) Framework for interpreting findings from impact evaluations (OPRE Report 2019-35). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Diaz, C.J., and J.E. Field. "The Effect(s) of Teen Pregnancy: Reconciling Theory, Methods, and Findings." *Demography*, vol. 53, no. 1, 2016, pp. 85–116.

Finer, L.B., and J.M. Philbin. "Sexual Initiation, Contraceptive Use, and Pregnancy among Young Adolescents." *Pediatrics*, vol. 131. no. 5, 2013, pp. 886–891.

Gloppen, K.M., C. David-Ferdon, and J. Bates. "Confidence as a Predictor of Sexual and Reproductive Health Outcomes for Youth." *Journal of Adolescent Health,* vol. 46, suppl. 3, 2010, pp. S42–S58.

Harding, J.F., J. Knab, S. Zief, K. Kelly, and D. McCallum. "A Systematic Review of Programs to Promote Aspects of Teen Parents' Self-sufficiency: Supporting Educational Outcomes and Healthy Birth Spacing." *Maternal and Child Health Journal*, 2020.

Kaye, K., K. Suellentrop, and C. Sloup. *The Fog Zone: How Misperceptions, Magical Thinking, and Ambivalence Put Young Adults at Risk for Unplanned Pregnancy*. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy, 2009.

Klerman, L.V. "Another Chance: Preventing Additional Births to Teen Mothers." Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy, 2004.

Lerner, R.M., and J.V. Lerner. "The Positive Development of Youth: Comprehensive Findings from the 4-H Study of Positive Youth Development." Chevy Chase, MD: National 4-H Council, December 2009. Available at https://www.researchconnections.org/childcare/resources/28375.

Maynard, R.A., and S. Hoffman. "The Costs of Adolescent Childbearing." In *Kids Having Kids: Economic Costs and Social Consequences of Teenage Pregnancy*, edited by Saul Hoffman and Rebecca Maynard. Washington, DC: Urban Institute Press, 2008.

McDonell, J.R., S.P. Limber, and J. Connor-Godbey. "Pathways Teen Mother Support Project: Longitudinal Findings." *Children and Youth Services Review*, vol. 29, no. 7, 2007, pp. 840-855.

National Longitudinal Study of Youth, 1997 cohort (U.S. Bureau of Labor Statistics n.d.)

Perper, K., K. Peterson, and J. Manlove. "Diploma Attainment among Teen Mothers. Fact Sheet. Publication 2010-01." Child Trends, 2010.

Rotz, D., D. Luca, B. Goesling, E. Cook, K. Murphy, and J. Stevens. "Final Impacts of the Teen Options to Prevent Pregnancy Program", Washington, DC: Office of Adolescent Health, U.S. Department of Health and Human Services, 2016.

Rotz, Dana, and Robert G. Wood. "Enhancing a Home Visiting Program to Address Repeat Adolescent Pregnancy: The Early Impacts of Steps to Success." OPRE Report # 2018-60. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2018.

Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, Inc., 1987.

Wasserstein, R., and N. Lazar. "The ASA Statement on *p*-Values: Context, Process, and Purpose." *The American Statistician*, vol. 70, no. 2, 2016, pp. 129–133.

What Works Clearinghouse. "Standards Handbook (version 4.1)." Washington, DC: Institute of Education Sciences, 2020.

White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, vol. 48, no. 4, 1980, pp. 817–838.

Positive Adolescent Futures Study

OASH Office of Population Affairs

Office of Population Affairs | Website: http://www.hhs.gov/opa/ Email: opa@hhs.gov Phone: (240) 453-2846