Evaluation of Adolescent Pregnancy Prevention Approaches

OAH Evaluation Report

Impact Report from the Evaluation of Adolescent Pregnancy Prevention Approaches

Interim Impacts of Teen PEP in New Jersey and North Carolina High Schools

March 2016

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

OFFICE OF ADOLESCENT HEALTH

Purpose statement: This study reports interim findings from a large-scale demonstration project and evaluation of the Teen Prevention Education Program (Teen PEP), as implemented in New Jersey and North Carolina high schools. Teen PEP is an in-school, peer-to-peer sexual health promotion program that combines peer-led interactive workshops and peer-driven school-wide initiatives in an effort to reduce sexual risk behaviors and associated outcomes among high school students. The study reports interim impacts of the program on sexual risk behaviors and other short-term outcomes measured about 6 months after the program's conclusion. A future report will examine the program's longer-term impacts on sexual risk behaviors 18 months after the program's conclusion.

March 2016

Recommended citation:
Rotz, Dana, Brian Goesling, Molly Crofton, Christopher Trenholm, Jennifer Manlove and Kate Welti. "Interim Impacts of Teen PEP in New Jersey and North Carolina High Schools." Washington, DC: U.S. Department of Health and Human Services, Office of Adolescent Health, March 2016.

This document is available at http://www.hhs.gov/ash/oah/oah-initiatives/evaluation/federal-led-evaluation/ppa-study.html

Prepared for OAH under contract number: HHSP233201450030A

Authored by:

Dana Rotz
Brian Goesling
Molly Crofton
Christopher Trenholm
Mathematica Policy Research

Jennifer Manlove
Kate Welti
Child Trends

ACKNOWLEDGEMENTS

This evaluation was made possible only through the hard work and support of many people and organizations. First, we wish to acknowledge the support of staff from the Office of Adolescent Health (OAH) in the U.S. Department of Health and Human Services (HHS). We especially thank Amy Farb, the OAH project officer for the Evaluation of Adolescent Pregnancy Prevention Approaches (PPA) study, who has provided outstanding support and guidance throughout the project, and Alexandra Warner of OAH, who provided valuable support in preparing the report. We also thank staff from the Office of Planning, Research and Evaluation within the Administration for Children and Families, especially Seth Chamberlain, for their support and guidance during the early site recruitment and design phase of the evaluation.

Special thanks goes to staff from the Center for Supportive Schools (CSS) and HiTops, Inc. for their tireless efforts recruiting and retaining study schools and supporting program implementation and study data collection. We especially thank Sherry Barr of CSS, for her unparalleled commitment to and enthusiasm for the program and study. We also thank staff from Abt Associates, in particular Carolyn Layzer, for leading the study design and data collection efforts in North Carolina high schools.

At Mathematica Policy Research, we thank Melissa Thomas for her long-standing commitment to the study and for successfully leading a highly complex, extended study data collection effort. We also thank Jennifer Littel for editorial assistance and Jill Miller for producing the report. At Child Trends, we thank Heather Fish for helpful research assistance and Kristin Moore for valuable comments on the report.

Finally, we extend our greatest thanks and gratitude to the youth, teachers, and school and district staff who agreed to participate in the study. We hope that our report does justice to the time and effort they devoted to making the study possible.

Dana Rotz
Brian Goesling
Molly Crofton
Christopher Trenholm
Mathematica Policy Research

Jennifer Manlove
Kate Welti
Child Trends

This page has been left blank for double-sided copying.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

FIGURES

This page has been left blank for double-sided copying.

I. INTRODUCTION

Although teen pregnancy rates in the United States have decreased over the past 25 years (Martin et al. 2015), teenage pregnancy remains a serious concern. Teen parenting is linked to negative outcomes for both teen parents and their children. Teen mothers are less likely to graduate high school, have lower earnings, receive public assistance for longer periods, and are more likely to be single parents (Hoffman 2008; Perper et al. 2010). Children of teen mothers have worse educational, criminal, and health outcomes as well (Hoffman 2008). Furthermore, teen pregnancy is associated with high costs to the general public; in 2010, teen pregnancy and childbirth in the United States cost taxpayers more than $9.4 billion (National Campaign to Prevent Teen and Unplanned Pregnancy 2013).

Research evidence has been mixed, but many teen pregnancy prevention researchers and practitioners have identified peer-to-peer education as a potentially valuable strategy to reduce rates of teen pregnancy and associated sexual risk behaviors. The Teen Prevention Education Program (Teen PEP), developed and implemented by the Center for Supportive Schools (CSS) and HiTops, Inc., takes this approach. The in-school, peer-to-peer sexual health promotion program combines peer-led interactive workshops and peer-driven school-wide initiatives in an effort to reduce sexual risk behaviors and associated outcomes among high school students. In particular, the program selects and trains a group of 11th- and 12th-grade students to serve as peer educators in their schools. The students receive training on sexual health topics, communication, and leadership, preparing them to deliver a series of five sexual health education workshops to 9th-grade students and one family-night workshop for parents/guardians, family members, and the broader school community. By emphasizing student involvement in all aspects of the program, Teen PEP seeks to develop and promote a broad school-wide culture of positive, healthy youth development and reduce rates of adolescent health risk behaviors.

This report presents interim findings from a large-scale demonstration project and evaluation of Teen PEP as implemented in a sample of New Jersey and North Carolina high schools. The program's implementers had previously evaluated Teen PEP using pre-post methods and found encouraging results (Jennings et al. 2014; Princeton Center for Leadership Training n.d.a, n.d.b). Building on these results, we designed the present study using more rigorous research methods and a larger sample of schools. As described later in this chapter, we designed the evaluation to assess the impacts of the program on the 9th-grade students who received the Teen PEP workshops. This report describes the impacts of the program on student outcomes about six months after the workshops concluded. A future report will examine the program's longer-term impacts measured about 18 months after the workshops concluded. A series of earlier reports (Asheer et al. 2014; see also Layzer et al. 2014; Layzer and Rosapep 2012, 2013) presented findings from a corresponding implementation study of the program.

The evaluation has involved a unique partnership and collaboration among several organizations. In 2009, Teen PEP was selected as one of the first sites in the Evaluation of Adolescent Pregnancy Prevention Approaches (PPA), a major federal effort to expand available evidence on effective ways to prevent and reduce pregnancy and related sexual risk behaviors among teens in the United States. Mathematica Policy Research and its partners, Child Trends and Twin Peaks Partners, LLC, under contract with the Office of Adolescent Health (OAH) within the U.S. Department of Health and Human Services, are conducting the evaluation. The

original PPA study design called for evaluating the effects of Teen PEP among high schools in New Jersey. In 2010, CSS received a separate competitive grant from OAH to expand and evaluate Teen PEP in another state, North Carolina. The evaluation of the North Carolina program was originally designed by researchers from Abt Associates. Under the guidance of OAH, these planned evaluation activities were ultimately combined into a single, unified evaluation of Teen PEP across two states: New Jersey and North Carolina. Researchers from the PPA study team led the combined impact study and conducted an implementation study of Teen PEP in New Jersey high schools. Researchers from Abt Associates led an implementation study of Teen PEP in North Carolina high schools and supported data collection in that state for the combined impact study.

This report comprises five chapters. In the remainder of this chapter, we describe the Teen PEP curriculum, briefly review prior research on the effectiveness of programs incorporating peer-led and school-wide components, and summarize the research questions of interest for this study. Chapters II and III describe the study design, data, measures, and analytic methods. Chapter IV presents findings from the impact analysis, and Chapter V summarizes and discusses the implications of the results.

A. Teen PEP: Educating younger high school students through peer learning

Teen PEP has a long history within New Jersey high schools. In the mid-1990s, New Jersey's governor was looking for a way to combat HIV/AIDS that could potentially be implemented throughout the state. In preparation for a statewide conference on this topic, the New Jersey Department of Health (DOH) approached CSS and HiTops for assistance in developing a new approach to educating youth on HIV/AIDS prevention. Using DOH funding, the organizations developed a peer-led model that they introduced at the conference and began implementing in New Jersey communities. As a result of these implementation efforts, CSS and HiTops became aware that their approach needed to extend beyond HIV/AIDS prevention, to include a sustained and comprehensive school-based sexual education program. The organizations thus developed the Teen Prevention Education Program (Teen PEP). By 2009, the program was well established in more than 50 public high schools throughout New Jersey, with HiTops maintaining oversight of program content and CSS managing the operational aspects of program delivery. HiTops and CSS revised the curriculum in 2010 to increase its emphasis on risky sexual behaviors and pregnancy-prevention outcomes, and to better define the core required components.

More recently, CSS has worked to expand the reach of Teen PEP beyond the state of New Jersey. Since 2005, CSS has supported the use of Teen PEP in a small number of North Carolina schools, in partnership with the North Carolina Department of Health and Human Services, Teen Pregnancy Prevention Initiatives and the North Carolina Department of Public Instruction, Healthy Schools. CSS staff manage the North Carolina program, assisted by a full-time program-fidelity manager from HiTops. In 2010, CSS received a five-year federal grant from OAH to expand and evaluate program operations in North Carolina schools.

As currently implemented, Teen PEP's comprehensive sexual education curriculum is interactive and dynamic and aims to build strong connections among participating students, staff, and the school community. School stakeholders work with CSS program staff to implement the

program. At each school, a group of students with identified leadership potential, representative of the diverse nature of the student body, is selected through an application process to serve as peer educators. Faculty advisors then train these students to become leaders and role models. These youth leaders deliver small-group workshops to educate younger teens on making healthy choices and avoiding risky behaviors, with the ultimate goal of reducing teen pregnancy and transmission of HIV/AIDS and other sexually transmitted infections (STIs). The peer educators also lead a school-wide campaign to promote positive cultural change.[1] The 11th- and 12th-grade students selected by school personnel to be peer educators serve as the cornerstone of Teen PEP.

Peer educators begin their training with a mandatory retreat, designed to foster trust and cohesion between the students and faculty advisors. They then participate in either a 45-minute class each school day throughout the academic year or a 90-minute class each school day for one semester, usually as an elective course or as a replacement course for health or physical education. Peer educators earn a grade and credit toward their graduation requirements for their participation in the course. The course is designed to prepare students to deliver outreach workshops on sexual health topics to 9th-grade students, their parents, and other family members. Faculty advisors deliver 10 core units to the peer educators, incorporating experiential and activity-based learning. These units provide medically accurate comprehensive sexual education on topics such as overcoming gender roles and stereotypes, postponing sexual involvement, reproductive health, preventing pregnancy, HIV/AIDS and other STIs, and how using alcohol and drugs affects sexual decision making. The curriculum aims to dispel common myths, encourages youth to reach their own conclusions, and helps youth develop important life skills related to negotiation and refusal, communication, and problem solving. For example, for the unit on pregnancy prevention, students might conduct online research on the time and cost of raising a child, including buying diapers and arranging and paying for doctors' appointments.

The Teen PEP curriculum describes the five core workshops that peer educators deliver to 9th-grade participants in small groups. The 90-minute workshops correspond in content and format to the classroom-based course for the peer educators, focusing on topics most relevant to reducing risky behaviors:

1. **Let's Wait Awhile: Postponing Sexual Involvement.** Peer educators and workshop participants describe reasons why teens do and do not become sexually involved, and possible consequences of early sexual involvement; identify relationship qualities that are important to have before beginning a sexual relationship; and demonstrate negotiation and refusal skills.
2. **Later, Baby: Pregnancy Prevention.** Peer educators and workshop participants identify behaviors that put teens at risk for unintended pregnancy; identify solutions to barriers that get in the way of teens using condoms, practicing birth control, or seeking guidance at a family-planning clinic; describe at least three methods for preventing pregnancy; and identify the location of a nearby family-planning clinic.
3. **Don't Pass It On: Preventing Sexually Transmitted Infections.** Peer educators and workshop participants describe the identification, symptoms, treatment, and long-term

[1] For further details, see Asheer et al. (2014) and Layzer et al. (2014).

consequences of the most common STIs among teens; demonstrate a greater understanding of how STIs are spread; and identify personal strategies for preventing the spread of STIs.

4. **Break the Silence: HIV/AIDS Prevention.** Peer educators and workshop participants describe the two most common ways teens become infected with HIV/AIDS, identify behaviors that will decrease the risk of HIV infection, name strategies for reducing the risk of contracting HIV/AIDS, describe the steps to using a condom correctly, and increase motivation for using risk-reduction strategies.
5. **Sex on the Rocks: Alcohol, Other Drugs, and Sexual Decision Making.** Peer educators and workshop participants identify steps to decision making and the consequences of making sexual decisions under the influence of alcohol and other drugs. They also demonstrate refusal skills to resist the pressure to use alcohol and other drugs.

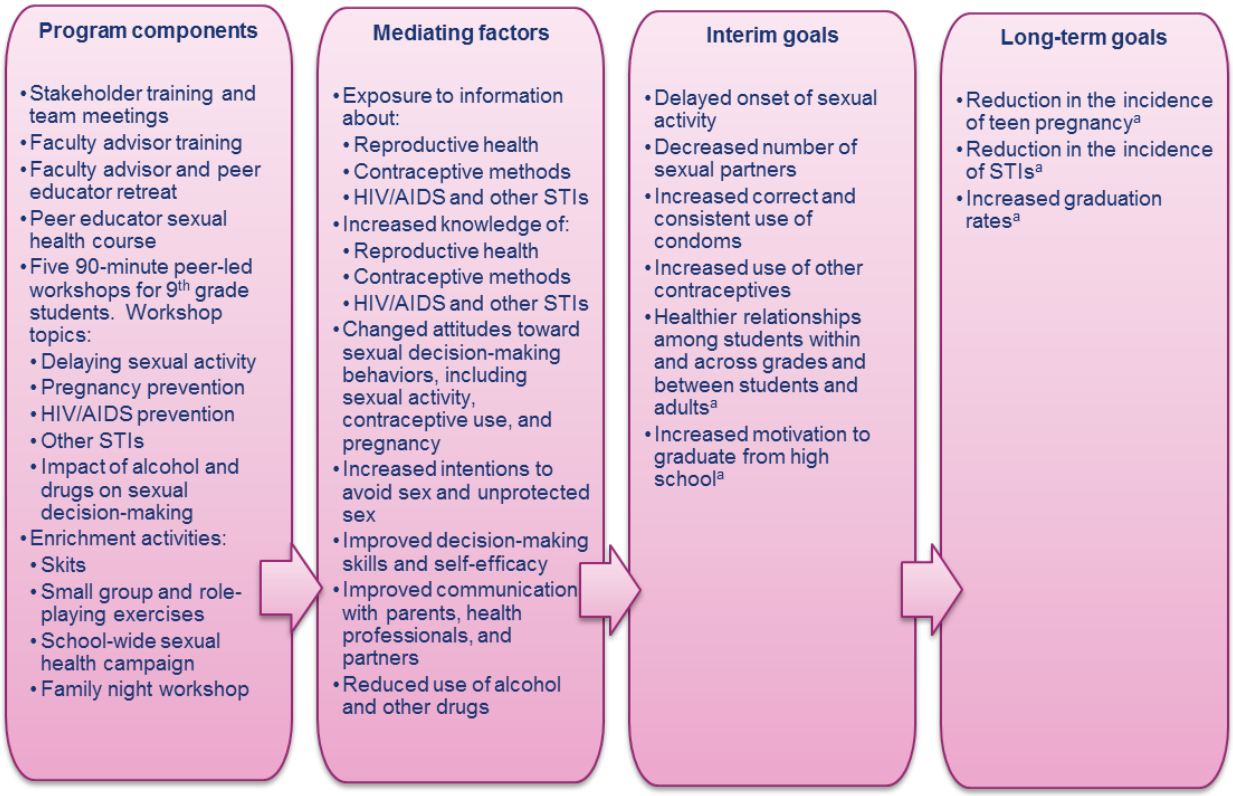
The peer educators also deliver a sixth workshop (**Talk to Me: A Family Night**) for parents/guardians, family members, and the broader school community. This workshop is designed to help parents or caregivers identify their personal attitudes and values regarding sexuality, become more comfortable talking about sex and sexuality with teens, and develop their understanding of how to initiate conversations about sex and sexuality with teens.

As with the classes that the peer educators take, the workshops they conduct are designed to engage and appeal to teens. Peer educators use accessible and plain language and use humor to convey main points and messages as part of the activities. Each workshop consists of skits, skill-building activities, and small-group discussions that the peer educators facilitate. Activities incorporate and emphasize communication with peers and parents, problem solving, decision making, and negotiation and refusal skills. For example, during the small-group sessions, peer educators answer specific questions from participants, present or reinforce key messages, and quiz students on what they have learned so far.

In addition to the workshops and family night, peer educators also lead a school-wide campaign to promote positive cultural change. This campaign reinforces workshop messages for participating 9th-grade students and can also help spread the Teen PEP messages to the broader school community. The Teen PEP curriculum does not prescribe the form that the campaign should take; students are asked to shape the campaign based on their own experiences of what activities are most influential and memorable.

These program components have the potential for both short- and long-term impacts on student outcomes (Figure I.1). The most immediate goal of Teen PEP is to increase student exposure to information on reproductive health, contraceptive methods, and STIs. In the short run, this exposure might affect possible mediating factors of sexual risk behavior, including student knowledge, attitudes, intentions, decision making and self-efficacy, communication, and other risky behavior (in particular, substance use). Changes in these mediating factors might then lead to changes in student sexual risk behaviors that the program targets. Among these are delayed sexual activity, increased correct and consistent use of condoms, increased use of other contraceptives, and reductions in the number of sexual partners. In the long run, such behavioral changes should lead to lower rates of both teen pregnancy and STIs, though changes in these long-term outcomes were not expected to emerge within the time frame of this study.

Figure I.1. Teen PEP intervention logic model



[a]Not assessed by this evaluation.

B. Prior research on peer-led programs to change behavior

Teen PEP seeks to achieve both its interim and long-term goals by leveraging the power of peer influence. Past research suggests that peers can play an important role in the choices teens make, including decisions about sexual activity and contraceptive use. For example, work by Cavanagh (2004) and Bearman and Brückner (1999) suggests that adolescents are more likely to have sex when they have peers who are sexually active, whereas adolescents with high-achieving peers are less likely to engage in sexual risk behaviors. Moreover, studies suggest that teens' perceptions of peer behaviors can be more predictive of their own actions than peer behaviors themselves (see Bearman and Brückner 1999; Killoren 2011; Majmudar 2005; and Miller et al. 1999). Thus, regardless of peers' actual behaviors, teens who perceive that their peers are sexually active might be more likely to become sexually active themselves.

Potentially because of these results, many teen pregnancy prevention programs, such as Teen PEP, have incorporated peer-led components, including establishing peer councils, engaging in peer-led activities, and including peers in program planning. Several have been rigorously evaluated but estimates of effectiveness have been mixed.[2]

[2] Many other programs not mentioned here include peer-led components as part of a multipronged approach to influence behavior. Evaluations of such programs typically do not allow researchers to estimate the impact of the specific peer-led component of the intervention.

Evaluations of a handful of programs have found that the peer-led approach can change student sexual behaviors. The Sisters, Informing, Healing, Living, Empowering (SiHLE) program, a peer-led community-based program, is geared toward African American female adolescents at high risk of contracting HIV and includes several sessions that one adult health educator and two peer mentors deliver. A random assignment evaluation of SiHLE found favorable impacts on measures of teen pregnancy and contraceptive use (DiClemente et al. 2004). Similarly, the Safer Choices program uses peer-led activities to supplement a two-year, school-based, teacher-led program. A cluster randomized controlled trial found that Safer Choices increased contraceptive use among sexually active high school students (Coyle et al. 1999, 2001; Basen-Engquist et al. 2001). Peer-led programs might also have larger impacts than adult-led programs with similar content. For example, Stephenson et al. (2008) demonstrated that RIPPLE, a peer-led sexual education program, was associated with reduced rates of pregnancy before age 18 when compared with a more standard teacher-led sexual education program. Peer leaders themselves can also change their behaviors after participating in program activities (for example, see Caron et al. 2004; Sieving et al. 2014).

But other studies of sexual education programs with peer-led components have found little impact of such interventions on behavior. In particular, many evaluations have demonstrated that youth prefer peer-led programs and that these interventions can influence adolescent knowledge, attitudes, and motivation to prevent pregnancy, but relatively few rigorous evaluations have found impacts on adolescent sexual activity and contraceptive use (Mellanby et al. 2000; Kim and Free 2008; Scott et al. 2014). This suggests that peer-led programs, such as Teen PEP, might face challenges in transforming impacts on intermediate outcomes into long-run behavioral change.

C. Prior research on school-wide cultural change

In addition to leading the educational sessions that form the core of Teen PEP, peer educators must develop and lead a school-wide campaign to emphasize and reinforce workshop messages. School-wide campaigns, including media campaigns, posters, staff training, and assemblies, have previously been shown to induce a wide range of positive behaviors among teens. Rigorous research has previously demonstrated that these strategies can improve nutrition (Nicklas et al. 1998), reduce alcohol-based risk taking (Duryea et al. 1984), and decrease bullying (Kärnä et al. 2011).

School- and community-wide activities have also been incorporated into effective, evidence-based teen pregnancy prevention programs. In addition to incorporating peer-led sessions (see Chapter I.B), the Safer Choices sexual health education program includes a peer leadership group to help plan school-wide activities and events designed to alter a school's normative culture. Evaluations of the program have found positive impacts both for teens who directly participated in Safer Choices (Coyle et al. 1999, 2001) and for students in intervention schools as a whole, including students who did not directly participate in the program (Basen-Engquist et al. 2001). This suggests that the school-wide component of Safer Choices might have impacted behavior above and beyond the other components of the program. As with Teen PEP's peer-initiated school-wide campaign, Teen Health Project incorpoates a peer-leadership council that plans community activities and events to reinforce the sexual education messages students receive during informational sesssions. Sikkema et al. (2005) found that the program as a whole

increased condom use among participants. Thus, the literature suggests that Teen PEP's school-wide component might reinforce its peer-led approach.

D. Research questions

This interim report examines the impacts of Teen PEP on 9th-grade students about six months after the program workshops. We focus specifically on Teen PEP's impacts on teen sexual behavior and on changes in intermediate factors that might respond more quickly to the program, including teens' exposure to information, knowledge, attitudes, and intentions. A future report will examine the program's longer-term effects about 18 months after the workshops.

The specific research questions we address in this interim report are:

- Does Teen PEP make students less likely to engage in sexual activity?
- Does Teen PEP reduce the incidence of unprotected sex?
- Is Teen PEP effective in increasing student exposure to information on reproductive health, contraceptive methods, and STI transmission and prevention?
- Does Teen PEP increase student knowledge of contraceptive methods and STI transmission and prevention?
- Do students receiving Teen PEP report different attitudes toward sexual activity, contraceptive use, and pregnancy?
- Do students receiving Teen PEP report increased intentions to avoid sexual activity or to use contraception when having sex?
- Does Teen PEP succeed in building students' decision-making skills and self-efficacy?
- Does Teen PEP increase communication about relationships or sexual health topics?
- Does Teen PEP lead to reductions in the use of alcohol and other drugs?

This page has been left blank for double-sided copying.

II. STUDY DESIGN

This study was originally designed as a cluster randomized controlled trial involving 30 schools spread across two states. Among the schools recruited for the study, just over half (n = 17) were randomly assigned to an "early" implementation group that could begin implementing Teen PEP the first year after enrolling in the study (the intervention group). The other schools (n = 13 schools) were randomly assigned to a "later" implementation group that had to delay implementing Teen PEP for at least 1.5 years (the comparison group). We planned to assess the impacts of Teen PEP by comparing student outcomes between the two groups of schools over an 18-month follow-up period.

However, as described in greater detail in this chapter, two factors worked to affect the validity of the original random assignment design. First, a relatively large number of schools ultimately dropped out of the study, especially from the New Jersey sample. Second, among those schools retained in the sample, we found relatively large differences in the baseline demographic and personal characteristics of students in the intervention and comparison groups. For these reasons, we ultimately applied quasi-experimental analysis methods to assess the effects of Teen PEP on student outcomes.

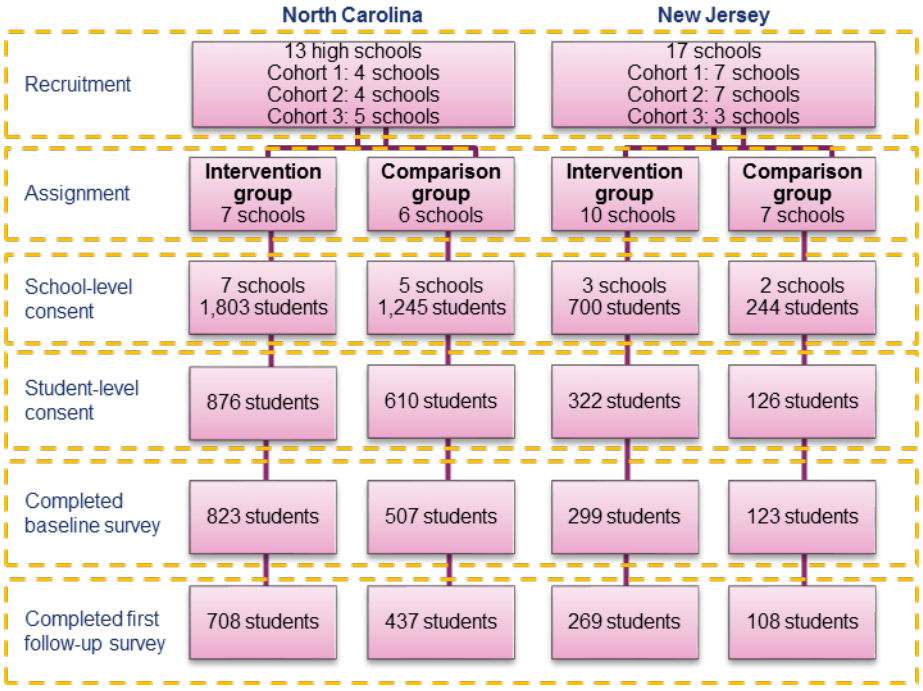
In the remainder of this chapter, we describe the recruitment and random assignment of the 30 study schools, the enrollment and retention of our student samples in these schools, and the intervention and comparison conditions. We describe the data, measures, and quasi-experimental analytic methods used to estimate program impacts later in Chapter III.

A. School recruitment

We recruited an initial sample of 30 schools in New Jersey and North Carolina, planning to implement the program over the course of three school years: 2011–2012, 2012–2013, and 2013–2014 (Figure II.1). We began the recruitment effort in New Jersey, seeking schools interested in implementing Teen PEP but not already doing so.[3] We later extended our recruitment efforts to North Carolina after CSS received a federal grant from OAH to support a demonstration project and evaluation of the program in schools outside of New Jersey. The North Carolina schools participating in the evaluation were some of the first in the state to receive the program, whereas more than 50 schools in New Jersey implemented Teen PEP before the evaluation. Three cohorts of schools were recruited for the evaluation, including 13 North Carolina schools and 17 New Jersey schools.

[3] The study design for this evaluation required recruiting schools that would be implementing Teen PEP for the first time.

Figure II.1. Flow of study schools and students



Source: Teen PEP study information system.

Within each cohort and state, schools were randomly assigned to the intervention and comparison conditions. In New Jersey, schools were further grouped into pairs or triplets based on school characteristics before random assignment in an effort to increase balance between study groups.[4] Schools in North Carolina were not grouped in this way.

A large number of schools ultimately dropped out of the study after random assignment but before administration of the baseline survey. Attrition was especially pronounced in the New Jersey sample. Of the 17 high schools initially recruited in New Jersey, 6 schools (35 percent) dropped out for various reasons, including concerns about the study survey, school closure, and a district-wide policy against evaluation. Teen PEP requires a planning period of several months, which typically takes place in the school year before implementation, necessitating a lengthy time between school recruitment and other study activities (in particular, student recruitment and student surveying), which might have exacerbated school concerns about participating in the evaluation. The evaluation also lost another 6 schools (35 percent) from the New Jersey sample because these had been randomized as part of the same matched pair or triplet as the attrited schools. Of the 13 high schools initially recruited in North Carolina, 1 school (8 percent) left the sample before any data collection. Appendix A provides details on the disposition of each school in our sample.

[4] In practice, this grouping may not have improved balance to a large degree. Each New Jersey cohort included only a small number of schools, with some schools being notably different from others in the cohort. As a result, matching schools into pairs before randomization only marginally improved balance.

For the current analysis, we thus used data from the 17 high schools that remained in the study—12 from North Carolina and 5 from New Jersey—as the basis for assessing the impacts of Teen PEP on student outcomes. This analysis sample includes 10 schools that were assigned to the intervention group and 7 schools that were assigned to the comparison group.

B. Student enrollment and retention

In both intervention and comparison schools, we recruited 9th-grade students for the evaluation. Students had to obtain the written permission of a parent or guardian before the baseline survey to participate in the study data collection. The study team offered individuals a $5 gift card for returning a completed consent form, regardless of whether their parent or guardian provided consent. In some schools (where allowed by the school stakeholder team), verbal consent was permitted as well. In these cases, the study team contacted parents and guardians via telephone and read the consent form aloud. The study team members then completed the form as indicated by the parent or guardian and sent a copy to the student's home. In intervention schools, participation in the study or study surveys did not impact whether a student could receive the program.

Table II.1 summarizes the consent and retention rates by state and assignment group. The study lost the largest number of students to the consent process. Just over two-thirds of students returned a consent form and slightly less than half of all students had their parents' consent to participate in the evaluation. Consent rates were very similar for individuals in intervention and comparison groups taken all together but varied across schools, from 37 to 67 percent.

Although consent rates limited sample sizes, participation in the baseline and first follow-up surveys was relatively high conditional on parental consent. In the sample of students who returned an affirmative consent form, 94 percent of the intervention group and 86 percent of the comparison group completed the baseline survey. Similarly, 86 percent of the intervention group and 84 percent of the comparison group completed the first follow-up survey. As with consent, these rates also varied widely by school, from 61 to 99 percent for the baseline survey and from 62 to 98 percent for the first follow-up survey. For further details on the consent and retention of students across all study schools, see Appendix A.

C. Intervention and comparison conditions

With the support of CSS and HiTops, all schools in the intervention group implemented Teen PEP during the prescribed time period, one school year for New Jersey schools and one semester for North Carolina schools (who used block scheduling). CSS and HiTops provided extensive support to intervention schools to facilitate the successful implementation of Teen PEP. Both organizations trained faculty advisors in what to look for when recruiting and selecting the peer educators who lead Teen PEP, how to prepare peer educators to conduct workshops, and how to observe the quality of the workshops. After implementation had begun, CSS and HiTops also offered additional training sessions to all participating schools that were delivering Teen PEP. CSS and HiTops further supplemented this training with technical assistance. Teen PEP staff monitored program delivery by attending and observing peer educator classes, peer-led workshops, and stakeholder team meetings. Staff also frequently offered written and oral feedback to ensure schools' adherence to the program model.

Table II.1. Student enrollment and retention by state and study group

	Number of students	Returned consent form		Consented		Completed baseline survey			Completed interim follow-up survey		
		Number	Percentage	Number	Percentage	Number	Percentage of all students	Percentage of consented students	Number	Percentage of all students	Percentage of consented students
North Carolina											
Intervention	1,803	1,239	69	876	49	823	46	94	740	41	84
Comparison	1,245	813	65	610	49	507	41	83	507	41	83
New Jersey											
Intervention	700	461	66	322	46	299	43	93	286	41	89
Comparison	244	174	71	126	52	123	50	98	110	45	87
All schools											
Intervention	2,503	1,700	68	1,198	48	1,122	45	94	1,026	41	86
Comparison	1,489	987	66	736	49	630	42	86	617	41	84

Source: Teen PEP study information system.

Findings from our accompanying implementation studies indicate that most key components of Teen PEP were implemented as intended, although some schools struggled with certain aspects of the program (Asheer et al. 2014; Layzer and Rosapep 2012, 2013). For example, some faculty advisors reported that delivering the activity-based learning model to peer educators was difficult, an issue that was compounded by many faculty members' lack of experience with classroom teaching and facilitating intensive, structured programs such as Teen PEP. Also, some schools struggled with the logistical demands of providing the small-group workshops to as many as 300 9th-grade students, which often required that peer educators deliver the same content multiple times. As a result, although most of the study schools implemented all of the scheduled Teen PEP workshops, in some cases attendance was low. For example, in one North Carolina school, only 62 percent of students attended four or more of the six Teen PEP workshops (including the Family Night, see Asheer et al. 2014). Finally, finding scarcely available classroom space for workshops impeded program operations in some cases. These issues were further intensified in some schools by limited buy-in and involvement of the stakeholder team.

During the intervention period, schools in the comparison group could implement any existing sexual health programs or curricula other than Teen PEP. Such programs varied across schools but especially across states. Students in most New Jersey schools take a health class that includes comprehensive sexual education. In contrast, few comparison schools in the North Carolina sample had exposure to comprehensive sexual education during their health class, despite a 2009 state requirement for schools to provide sexual education. Instead, students in North Carolina schools are more commonly exposed to abstinence-until-marriage curricula. For example, survey data collected from high school principals in North Carolina in 2010 indicate that only 21 percent of high schools had curricula that exposed students to condom-use topics within a required course (Mitchell and Greene 2011). The same data also show that most North Carolina health teachers had not received professional development covering HIV- or pregnancy-prevention information in the past two years, suggesting that educators might not have access to the most up-to-date information on these topics. Additionally, access to reproductive health services varied greatly by school. One school had a colocated clinic providing free health services, but another was more than 50 miles from such resources.

This page has been left blank for double-sided copying.

III. DATA, MEASURES, AND ANALYSIS

Our analysis of the impacts of Teen PEP draws on self-reported survey data collected from 9th-grade students in both the intervention and comparison schools at two time points: (1) a baseline survey administered in late fall or early winter of 9th grade, before the first Teen PEP workshop was delivered in the intervention schools and (2) an interim follow-up survey administered about 6 to 12 months later, in the fall of the students' 10th-grade year. Students received gift-card incentives for completing each survey (a $10 incentive if the student completed the survey at the time it was administered at the school and a $25 incentive if the student completed the survey over the phone). With these survey data, we can estimate the interim impacts of Teen PEP on the 9th-grade students receiving the workshops. A long-term, final follow-up survey of the same students is currently in the field.

A. Measures

To answer the study's research questions (see Section I.D), we constructed three groups of outcome measures: (1) measures of sexual behavior, (2) measures of student exposure to information on reproductive health and related topics, and (3) measures of other intermediate outcomes (knowledge, attitudes, intentions, decision making and self-efficacy, communication, and substance use).

1. Sexual behavioral outcomes

We used five outcome measures in two domains to determine whether Teen PEP had an impact on student sexual behavior:

Prevalence of sexual activity. We constructed three measures within this domain for each student: (1) a binary indicator equal to one if a student reported having sexual intercourse in the past three months (and zero otherwise), (2) a binary indicator equal to one if a student reported ever having sexual intercourse (and zero otherwise), and (3) a count measure of the number of lifetime sexual partners the student reported. Before analysis, we specified that the first variable would serve as a confirmatory outcome in this domain and be the chief measure of Teen PEP's success in reducing sexual activity. We chose this variable because it measures both primary and secondary abstinence.

Prevalence of unprotected sex. We constructed two measures within this domain, both focused on recent sexual activity: (1) a binary indicator equal to one if a student reported having intercourse without a condom in the past three months (and zero otherwise), and (2) a binary indicator equal to one if a student reported having intercourse without any effective birth control method in the past three months (and zero otherwise). Before analysis, we specified that the first variable would serve as a confirmatory outcome in this domain and be the chief measure of Teen PEP's success in reducing unprotected sex. We chose this variable because Teen PEP focuses on both pregnancy and STI-transmission prevention and gives particular attention to condom use as a means of prevention. Moreover, this measure is particularly important because of the relatively high prevalence of condom use among youth using any contraception.

2. Exposure to information

The interim follow-up survey contained one multipart question asking whether a student had received any information in the past 12 months on eight different topics relating to sexual activity and reproductive health. We used student responses to these individual items to assess whether Teen PEP increased student exposure to information on relationships, abstinence, methods of birth control, where to get birth control, STIs, talking about sex with a partner, saying no to sex, and how babies are made. To the extent we find evidence of impacts on behavioral outcomes, we would expect to find impacts on these informational measures as well.

3. Intermediate outcomes

We considered intermediate outcomes corresponding to six factors potentially mediating changes in sexual behavior: (1) student knowledge of the prevention of pregnancy and STI-transmission; (2) student attitudes toward birth control, sex, and pregnancy; (3) student intentions to engage in sexual activity and use contraception; (4) student decision-making skills and self-efficacy; (5) student communications with parents, health professionals, and partners about sexual health and related topics; and (6) student use of alcohol and marijuana. These measures are summarized in Table III.1 and described in greater detail in the remainder of this section. To the extent we find evidence of effects on behavioral outcomes, we would expect to find impacts on outcome(s) measuring one or more of these six mediating factors. Additionally, looking at these mediating factors could potentially enable us to project future changes in behavioral outcomes not yet observed. This is particularly valuable given the relatively young age and low rates of sexual initiation of the students in this study (see Section B of this chapter).

Knowledge. We constructed a single measure of knowledge within this outcome domain based on six survey questions related to contraceptive efficacy, pregnancy, and STI transmission. The knowledge index is the number of correct responses a student gave to the six items.

Attitudes. We created four measures of attitudes toward contraceptive use, sexual activity, and pregnancy to assess Teen PEP's impact within this domain:

Attitudes toward birth control. We constructed two measures of student attitudes toward birth control: an index of factors supporting birth control use and an index of barriers to using effective birth control. We used four survey items to construct the former measure. The survey asked students whether they agreed or disagreed with four statements related to support for using contraception, such as "Birth control is important to make sex safer" and "Condoms should always be used if a person your age has sexual intercourse." The five possible response categories ranged from "strongly agree" to "strongly disagree." We mapped the categorical responses to a five-point scale and averaged these numeric values across the four items to create a composite scale of general support for birth control. Higher values on the scale indicate stronger levels of support. We used five survey items to construct an index of barriers to using birth control. The survey asked students whether they agreed or disagreed with five statements concerning possible barriers to using contraception, such as "Condoms are a hassle to use" and "Birth control has too many negative side effects." The five possible response categories ranged from "strongly agree" to "strongly disagree." We mapped the categorical responses to a five-point scale and averaged these numeric values across the five items to create a composite scale of

perceived barriers to using contraception. Higher values on the scale indicate more perceived barriers.

Table III.1. Outcome variables by domain

Measure	Definition
Knowledge	
Knowledge of preventing STI transmission and pregnancy	Continuous index: sum of correct responses to six survey questions; variable ranges from 0 to 6, with higher values indicating greater knowledge. Students not responding to all six items have the index set to missing; otherwise, missing responses were counted as incorrect.
Attitudes	
Perceptions of factors supporting use of birth control	Continuous scale: average of responses to four survey questions; variable ranges from 1 to 5, with higher values indicating more supportive attitudes toward contraception (alpha coefficient = 0.73).
Perceptions of barriers to use of birth control	Continuous scale: average of responses to five survey questions; variable ranges from 1 to 5, with higher values indicating more perceived barriers toward contraception use (alpha coefficient = 0.71).
Negative views toward early sexual activity	Continuous scale: average of responses to four survey questions; variable ranges from 1 to 5, with higher values indicating less permissive attitudes toward early sexual activity (alpha coefficient = 0.75).
Would be very upset if got pregnant or got someone pregnant	Binary variable: equals 1 if a student reported he or she would be “very upset” if she became pregnant or he impregnated someone now; equals 0 if a student chose another response category (indicating he or she would be less upset or happy about a pregnancy).
Intentions	
Intend to have sex in the next year	Binary variable: equals 1 if a student reported he or she will “definitely” have sex in the next year if he or she has the chance; equals 0 if a student reported he or she will “probably”, “probably not” or “definitely not” do so.
Intend to use a condom if have sex in the next year	Binary variable: equals 1 if a student reported he or she will “definitely” use a condom if he or she has sex in the next year; equals 0 if a student reported he or she will “probably”, “probably not” or “definitely not” do so.
Intend to use any effective method of birth control if have sex in the next year	Binary variable: equals 1 if a student reported he or she will “definitely” use an effective method of contraception if he or she has sex in the next year; equals 0 if a student reported he or she will “probably”, “probably not” or “definitely not” do so.
Decision making and self-efficacy	
Perceptions of refusal skills	Continuous scale: average of four questions for boys and five questions for girls; variable ranges from 1 to 4, with higher values indicating greater perceived refusal skills (alpha coefficient = 0.78).

Measure	Definition
Thoughtfulness in decision making (follow-up)	Continuous scale: average of two question; variable ranges from 1 to 4, with higher values indicating more thoughtfulness in making decisions about sex (alpha coefficient = 0.74).
Believe could go to clinic to seek STI testing if needed	Scale-score variable: single survey item; variable ranges from 1 to 4, with higher values indicating greater belief that one could seek assistance.
Communication	
Frequency of discussions with parents about relationships or sexual health in past 12 months	Scale-score variable: based on one survey question; variable ranges from 0 to 10, with higher values indicating more communication.
Received very useful information from parents on relationships or sexual health in past 12 months	Binary variable: equals 1 if a student reported he or she received very helpful information from his or her parents on relationships, abstinence, birth control, or sexually transmitted diseases; equals 0 if the student otherwise responded to the survey question.
Spoke to health professional about sexual health in the past 12 months	Binary variable: equals 1 if a student reported he or she spoke with a doctor or nurse about sex, birth control, or sexually transmitted diseases in the past 12 months; equals 0 if a student reported otherwise.
Insufficient communication about sex with partner	Binary variable: equals 1 if a student had a partner in the past month and did not discuss at least one of six topics related to sexual activity "Often;" equals 0 if a student had frequent communication with his or her partner or did not have a partner.
Substance use	
Drank alcohol in past 30 days	Binary variable: equals 1 if a student reported drinking alcohol one or more times in the past 30 days; equals 0 if a student reported he or she did not do so.
Smoked marijuana in past 30 days	Binary variable: equals 1 if a student reported smoking marijuana one or more times in the past 30 days; equals 0 if a student reported he or she did not do so.

Note: See Appendix B for further details.

- **Views about having sex.** The survey asked students whether they agreed or disagreed with four statements about having sexual intercourse at their current age. For example, one of the statements read, "At my age, having sexual intercourse would create problems." Four response categories ranged from "strongly agree" to "strongly disagree." We mapped the categorical responses to a five-point scale and averaged these numeric values across the four items to create a composite scale of student views on early sexual activity. Higher values indicate less permissive attitudes toward early sexual activity.
- **Very upset about a pregnancy.** The survey asked female students, "If you got pregnant now, how would you feel?" and male students, "If you got someone pregnant now, how would you feel?" The five possible response categories ranged from "very happy" to "very upset." Students who indicated they would be "very upset" received a one for this binary measure; students who selected another response received a zero.

Intentions. We included three measures in this domain to assess whether Teen PEP influenced intentions to engage in risky sexual behavior in the near future. We constructed these binary measures to reflect whether, in the next 12 months, the student "definitely" intended to: (1) have sexual intercourse, (2) use a condom if he or she has sex, or (3) use any effective form of birth control if he or she has sex. The survey asked students whether they intended to engage in each activity. The four possible response options ranged from "yes, definitely" to "no, definitely not." We created binary indicators equal to one for each measure if the student said he or she "definitely" intended to engage in that activity and zero otherwise.

Decision making and self-efficacy. To assess Teen PEP's success in improving student decision making and related skills, we constructed outcomes capturing student refusal skills, thoughtfulness in decision making, and ability to seek sexual health services:

- **Refusal skills.** We constructed a scale measuring student perceptions of their refusal skills using two sets of survey questions. The first set of questions asked students to imagine they were alone with someone they liked very much and to assess the likelihood they could resist unwanted sexual contact. Boys and girls were asked specifically about avoiding sexual intercourse and someone touching their genitals. Girls were also asked about their perceived ability to avoid someone touching their chest. The second set of questions asked students whether they agreed or disagreed with two statements: "If my partner refused to use condoms, I could refuse to have sex" and "I would have sex now if someone I cared about pressured me to have sex." The four response categories for both sets of questions ranged from "strongly agree" to "strongly disagree." We mapped the categorical responses to a four-point scale, with higher numbers representing stronger refusal skills for all survey questions, and averaged these numeric values across the four or five survey items to create a composite scale.
- **Thoughtfulness in sexual decision making.** To assess students' thoughtfulness around decisions related to sexual activity, we constructed a scale based on two items in the follow-up survey. The survey asked, "When you have to make a decision about your sexual behavior, how often do you think of the consequences of each possible choice?" and "When you have to make a decision about your sexual behavior, how often do you first get as much information as you can?" Four response categories ranged from "never" to "very often." We mapped the categorical responses to a four-point scale and averaged these numeric values to create a composite measure, with higher values indicating more thoughtfulness in making decisions about sex.
- **Ability to seek health care.** The survey asked students whether they agreed or disagreed with a single survey item related to their ability to seek sexual health services: "I believe I could go to a clinic if I needed to get tested for HIV/AIDS or another sexually transmitted disease (STD)." The four response categories for this question ranged from "strongly agree" to "strongly disagree." We mapped the categorical responses to a four-point scale, with higher values indicating more agreement.

Communication. We created four measures of communication with parents, health care professionals, and partners about sexual health and related topics to measure the impacts of Teen PEP within this domain:

- **Communication with parents.** We created two measures of student communication with parents. The first outcome, which measured how often students communicated with their parents about sex, used a single question that asked students how many times they received information from their parents or other relatives about "relationships, abstinence, birth control, or sexually transmitted diseases" in the past 12 months. Categorical responses were coded to numerical values to create this measure. The second outcome is a binary variable capturing whether students received helpful information from their parents. The survey asked students where they had received "information on relationships, abstinence, birth control, or sexually transmitted diseases that was very helpful." If a student selected "parents and other relatives or family members" as a response, we set this binary variable to one; other responses were set to zero.
- **Communication with health professionals.** In the same battery of questions mentioned above, students were asked how many times they received information from a doctor, nurse, or clinic about "relationships, abstinence, birth control, or sexually transmitted diseases" in the past 12 months. We created a binary measure equal to one if their response indicated any communication with a health professional on these topics (and zero otherwise).
- **Communication with partner.** The follow-up survey asked students "In the last month, how often have you talked with your partner about each of the topics listed below?" Six topics related to sexual activity and boundaries were then listed, such as "birth control" and "what you feel comfortable doing sexually." Response categories included "often," "sometimes," and "never." We set this measure to one for students who did not discuss at least one of the six topics "often" with their partner and to zero for students who discussed at least one topic "often." The measure equals zero for students who indicated that they did not have a partner in the past month.

Substance use. The substance use domain includes indicators for whether a student had (1) consumed alcohol or (2) smoked marijuana at least one time in the past 30 days. The indicators were derived from survey questions that asked how many days in the past month the respondent used each substance. We set this binary variable to zero for students who did not engage in a given activity and one for students who engaged in the activity at any frequency.

B. Analytic approach

Our analytic approach was shaped by two key issues. First, as described in Chapter II, nearly half the schools randomized for the study left our sample before data collection.[5] This level of attrition compromised the validity of the original random assignment evaluation design; because schools left the study in a non-random manner, we cannot be confident that any differences between the study groups in our analytic sample at baseline are due only to chance. Second, as summarized in Appendix C, among the schools that were retained for the study, the baseline characteristics of student samples differed between the intervention and comparison groups. For example, students in intervention schools were 11 percentage points less likely to report having

[5] Note that even though only one school attrited from the subsample of North Carolina schools, many evidence reviews would still consider this a high level of attrition given that this resulted in a nontrivial difference in attrition rates for the intervention and comparison groups. See, for example, U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2013).

had sex in the three months before the baseline survey, compared to students in the comparison group.

To account for these analytic issues, we used a quasi-experimental (comparison group) design that restricted the study sample to a well-matched subset of the original sample. We conducted this matching using propensity-score methods. The resulting samples of intervention- and comparison-group students demonstrate improved balance. We further used regression analysis to strengthen the validity of our results.

In the remainder of this section, we first describe the propensity score methods we used to define a matched study sample. We then show the baseline demographics and personal characteristics of the matched intervention and comparison groups. Next, we describe the regression models we used to estimate program impacts among the matched sample. We end this section by describing additional subgroup analyses and robustness checks we conducted with the matched sample.

1. Propensity score approach

We used propensity score methods to select a subset of intervention and comparison students for our analysis. Selecting a subgroup of individuals for analysis enables us to greatly decrease differences between the intervention and comparison groups at baseline and improve the causal validity of our impact estimates. However, this decrease in sample means that we have reduced statistical power and can make inferences only relevant to a smaller group of students. That is, our estimates are only valid for the students in our final analysis sample and not the entirety of the students enrolled in the study (who are themselves a subset of the students in study schools).

Although random assignment was conducted at the school level, we estimated the propensity score using data at the student level. Because our sample only contained 17 schools, it would have been infeasible to estimate propensity score regressions using school-level data. Other methods of school-level matching (for example, coarsened exact matching) would also have resulted in relatively poor balance between the intervention and comparison groups because of the small number of schools. Because of these constraints, we decided to use propensity score matching to select a group of students in comparison schools most similar to the students in Teen PEP schools and not focus on the level of random assignment.

The details of our propensity score approach follow the recommendations of Imbens (2015). We used a logit model and chose the precise propensity score regression specification using a stepwise procedure. First, we pre-specified that age, race, gender, state, cohort, and baseline measures of sexual risk behaviors (indicators for sex and sex without a condom in the past three months) be factors used in matching observations. We then iteratively added variables from most to least predictive of treatment status to the propensity score regression to determine which other variables should enter the model. We considered a rich set of demographic and background characteristics, baseline measures of outcomes, and related variables in this process (Table III.2). After determining the variables that we would use in the propensity score model, we also included in the model any first-order interactions of these variables that were highly predictive of treatment status. For all analyses, we imputed any missing covariates to their mean value to avoid limiting sample size.

Table III.2. Candidate and selected variables for propensity score regression

Variable	Included in propensity score regression?
Ever had sexual intercourse	
Number of lifetime sexual partners (as indicators for 0,1, 2, and 3 or more)	X
Had sexual intercourse without any effective method of birth control in the three months before survey	X
Ever had oral sex (cohort 1 only)	
Ever had nonpenetrative sex	X
Ever kissed member of opposite sex	
Received any information in past 12 months on	
Relationships	
Abstinence	
Methods of birth control	X
Where to get birth control	
STIs	X
Talking about sex with your partner	X
Saying no to sex	X
How babies are made	X
Knowledge of preventing STI transmission and pregnancy	X
Perceptions of factors supporting use of birth control	
Perceptions of barriers to use of birth control	X
Index of negative views toward having sex at current age	X
Would be very upset if got pregnant or got someone pregnant	X
Intend to have sex in the next year	
Intend to use a condom if have sex in the next year	
Intend to use any effective method of birth control if have sex in the next year	
Perceptions of refusal skills	X
Spoke to health professional about sexual health in past year	
Communication with parents about risky behavior	X
Drank alcohol in past 30 days	X
Smoked marijuana in past 30 days	X
Main language spoken at home is English	X
Biological mother living in home or main home	
Biological father living in home or main home	
Biological parents currently married	X
Biological parents divorced or separated	X
Report religion is very important in their life	X
Identify as lesbian, gay, bisexual, asexual, or questioning	
Ever drank alcohol	X
Ever smoked marijuana	X
Several interaction terms were also included in the model and are available upon request.	

We estimated the propensity score based on measures of age, race, gender, cohort, indicators for sex and sex without a condom in the past three months, and the variables indicated in

Table III.2.[6] We removed any students from our sample who had propensity scores less than 0.05 or greater than 0.95, to improve overlap and ensure our sample excluded individuals who are very different from the average individual in either study group (see Crump et al. 2009). After trimming based on the propensity score, the sample contained 714 intervention-group students and 471 comparison-group students.

We then matched each remaining intervention-group student to the comparison-group student in the same state with the closest propensity score. Comparison-group students could be matched to multiple intervention-group students. The matching comparison-group students became the comparison group we used in our analysis, with observations weighted based on the intervention group students they matched (weights for the intervention group were created so that each school implementing Teen PEP contributes equal weight to the analysis and the analysis thus yields the effects of Teen PEP in the average intervention school).[7]

After matching, our analytic sample contained 714 intervention-group students and 228 comparison-group students. The large reduction in the size of the comparison group highlights the differences in the samples at baseline; fewer than half of comparison-group students were sufficiently similar to the intervention group to be chosen for the propensity-score matched sample. This reduction in sample size resulted in a much better match between the study groups but came at the loss of statistical power. That is, the study groups are far more similar, but reductions in sample size decrease the likelihood that we will find statistically significant differences between the groups.

2. Baseline equivalence

The revised sample shows improved balance between the intervention and comparison groups. In the propensity-score matched sample, differences in demographics tend to be smaller than those for the full sample (Table III.3). For example, comparison-group students in the full sample were 18 percentage points more likely to be black than intervention-group students in the full sample. In the propensity-score matched sample, the magnitude of this difference shrinks to 7 percentage points. Likewise, the comparison group and intervention group in the propensity-score matched sample reported the same average age, whereas the full sample demonstrated a 0.3-year difference in average ages between the groups. The largest demographic difference in the revised sample pertains to gender composition, such that the intervention group contains relatively fewer females than the comparison group (a difference of 9 percentage points). However, the difference is not significant at the 5 percent level ($p = 0.07$).[8] We also found a large and statistically significant difference in the share of the sample that had ever smoked cigarettes across the study groups. In particular, 24 percent of the intervention group reported having ever smoked, compared to 43 percent of the comparison group (a difference of 19

[6] Note that we omitted state from the regression because our matching procedure required exact matching based on state.

[7] Comparison-group students matched to multiple intervention-group students received a weight equal to the sum of the weights of all matched students.

[8] In Section IV.4, we estimate impacts of Teen PEP by gender and found few differences, suggesting the difference in gender composition does not drive the results.

percentage points, $p < 0.01$). We controlled for this variable in all regression analyses (see Section B.3 of this chapter for details) to mitigate concerns that this might bias our results.

Table III.3. Demographic and background characteristics for study (propensity-score matched) sample

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Average age	15.1	15.1	0.0	0.67
Female	55.5	65.0	–9.5	0.07
Race				
Hispanic	24.9	26.7	–1.9	0.89
White, non-Hispanic	32.3	37.1	–4.8	0.64
Black, non-Hispanic	30.5	23.3	7.3	0.33
Other race/ethnicity	12.1	12.9	–0.8	0.86
Race missing	0.2	0.0	0.2	0.15
Main language spoken at home is English	87.8	88.6	–0.8	0.93
Biological mother living in home or main home	88.0	78.6	9.4	0.18
Biological father living in home or main home	49.2	56.5	–7.3	0.33
Biological parents currently married	45.5	45.5	0.0	1.00
Biological parents divorced or separated	27.1	29.1	–2.1	0.74
Report religion is very important in their life	40.3	39.0	1.4	0.86
Identify as lesbian, gay, bisexual, asexual, or questioning	13.4	9.8	3.7	0.18
Ever smoked cigarettes	24.4	43.3	–18.9**	<0.01
Ever drank alcohol	49.7	49.3	0.5	0.94
Ever smoked marijuana	26.5	23.3	3.2	0.44
Sample size	**714**	**228**		

Source: Teen PEP baseline survey.

Note: Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

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

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

Importantly, unlike the full sample, the study groups in the propensity-score matched sample exhibited similar rates of sexual risk behaviors at baseline (Table III.4). There is no significant difference in either of our confirmatory outcome variables: sexual activity in the past three months or sexual activity without a condom in the past three months. In both cases, the intervention group was 3 percentage points more likely to report these activities than the comparison group ($p > 0.22$). Individuals in the intervention group were slightly more likely to have ever had sex, but the difference was again small (1 percentage point) and statistically insignificant ($p = 0.80$). Differences in nonintercourse sexual activity (oral sex, nonpenetrative sex, and kissing) were also small and insignificant across the study groups ($p > 0.15$). Furthermore, the differences in the propensity-score matched sample were far smaller than those seen in the full sample. For example, the full-sample comparison group was 11 percentage points

more likely to report having had sexual intercourse in the past three months and 9 percentage points more likely to report having had sex without a condom in the past three months, compared to the full-sample intervention group (Appendix Table C.2).

Table III.4. Baseline measures of sexual risk behavior for study (propensity-score matched) sample

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Outcomes analyzed at follow-up				
Ever had sexual intercourse	25.1	23.7	1.4	0.80
Number of lifetime sexual partners	0.75	0.47	0.28	0.18
In the three months before survey				
Had sexual intercourse[a]	13.7	11.1	2.6	0.50
Had sexual intercourse without a condom[a]	9.0	6.3	2.8	0.22
Had sexual intercourse without any effective method of birth control	7.3	4.8	2.6	0.15
Other sexual behavior				
Ever had oral sex (cohort 1 only)	25.7	28.9	–3.2	0.15
Ever had nonpenetrative sex	52.2	55.0	–2.8	0.53
Ever kissed member of opposite sex	84.7	87.4	–2.7	0.45
Sample size	**714**	**228**		

Source: Teen PEP baseline survey.

Note: Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

[a]Designated as a confirmatory outcome.

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

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

Most, but not all, intermediate outcomes also demonstrate only small and insignificant differences in the propensity-score matched sample at baseline (Table III.5). Students in both study groups reported similar exposure to information on reproductive health, knowledge of pregnancy and STI transmission prevention, attitudes, decision making and self-efficacy, and communications with parents and health professionals. We found notable differences only in the intentions and substance use domains. In particular, intervention-group students were significantly more likely to intend to have sex in the next year (10 percent versus 6 percent, $p = 0.01$) and more commonly reported smoking marijuana in the past 30 days (16 percent versus 8 percent, $p = 0.01$). As discussed in the following section, our regression approach controlled for both marijuana use in the past 30 days and intentions to have sex (among other variables) to limit the differences' potential as a source of bias in the impact estimates.

Table III.5. Baseline measures of intermediate outcomes for study (propensity-score matched) sample

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Intermediate outcomes				
Received any information in past 12 months on				
Relationships	86.9	85.4	1.5	0.76
Abstinence	67.0	71.0	–4.0	0.47
Methods of birth control	49.9	49.9	0.0	1.00
Where to get birth control	42.6	43.3	–0.7	0.93
STIs	80.2	79.2	0.9	0.83
Talking about sex with your partner	51.2	51.2	0.0	1.00
Saying no to sex	74.7	75.6	–1.0	0.79
How babies are made	83.1	86.2	–3.1	0.53
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	2.8	3.0	–0.2	0.14
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.3	4.3	0.0	0.90
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.5	0.0	0.93
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	3.0	3.0	0.0	0.88
Would be very upset if got pregnant or got someone pregnant	53.1	61.3	–8.2	0.20
Intend to have sex in the next year	10.4	5.6	4.8**	0.01
Intend to use a condom if have sex in the next year	83.7	81.7	2.0	0.54
Intend to use any effective method of birth control if have sex in the next year	88.9	92.5	–3.6	0.11
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.2	3.2	–0.1	0.48
Spoke to health professional about sexual health in past year	30.0	31.7	–1.7	0.76
Communication with parents about risky behavior (average of six survey items; range 0–10)	2.9	3.2	–0.4	0.33
Drank alcohol in past 30 days	28.1	24.0	4.1	0.52
Smoked marijuana in past 30 days	15.8	8.2	7.6**	0.01
Sample size	**714**	**228**		

Source: Teen PEP baseline survey.

Note: Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

*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. Regression framework

We used regression analysis on our matched sample to estimate the impact of Teen PEP, controlling for a number of observed outcomes at baseline that might be a source of bias despite the use of propensity-score matching (Funk et al. 2011). For binary outcome measures (for example, sex without a condom in the past three months), we estimated impacts using logistic regression models.[9] When reporting results from these models, we calculated mean marginal effects to express the impact estimates as percentage-point differences between outcomes for the intervention and comparison groups. For all other outcomes, we estimated ordinary least squares regression models. In the regression models for all outcomes, we adjusted the standard errors to account for clustering at the school level.

Our regressions controlled for a baseline measure of the outcomes (or a close proxy from the same domain if one was available but the outcome was not measured at baseline), an indicator for whether this variable was missing at baseline, the linear terms controlled for in the propensity-score matching procedure, and any variables with differences at baseline in the propensity-score matched sample exceeding 0.05 standard deviations.[10] Missing covariates were again imputed to their mean value to avoid limiting sample size. In cases where more than 10 percent of observations had missing data for a variable, we also included in the regression a binary variable equal to one if an individual had an imputed response (and zero otherwise).

We adjusted the statistical significance tests (*p*-values) from our regression models to account for multiple hypothesis testing.[11] As discussed earlier in this chapter, our analysis used multiple outcomes to answer some of our key research questions. For example, we constructed four measures of attitudes and three measures of intentions. Unless we account for this multiplicity, it could increase the chances of making a false discovery and lead to spurious claims about the program's effectiveness. Researchers often declare a finding "statistically significant" if the probability of falsely rejecting the null hypothesis of no impact is less than 5 percent. However, when conducting separate tests arising from multiple outcomes, the probability of falsely rejecting the null hypothesis in at least one of them can be much higher than 5 percent. To correct for this increased probability, we applied a multiple hypothesis testing procedure outlined by Hothorn et al. (2008) and Schochet (2009). This procedure involves adjusting the reported *p*-value for each test to account for other tests conducted within the same "family" of related measures. This procedure yields a 5 percent false positive rate across outcomes within the same family. However, the procedure is less severe than other common adjustment methods, such as the well-known Bonferroni correction, because it also accounts for any correlation in test statistics among outcomes within the same family.

We made this adjustment separately for each of the nine groups of outcome measures described earlier in this chapter. That is, we adjusted the *p*-values accounting for multiple

[9] When examining the outcome for ever had sex, the logit model required that we restrict our sample to individuals who had not had sex at baseline to avoid perfect prediction of the outcome variable.

[10] See Table B.2 for a list of all possible covariates and Appendix B, Section E for a list of the covariates included in the regression model.

[11] See Appendix D, Section B for estimated *p*-values not adjusting for multiple hypotheses.

outcomes within each of the nine groups of measures, but we did not adjust for multiple outcomes measured across the different groups. We followed this approach because each group of outcomes aligns with a different and unique research question. We based our substantive conclusions for each question only on the corresponding group of outcome measures. The number of outcomes measured in other groups had no bearing on our substantive conclusions for each question and therefore does not warrant an additional adjustment for multiple hypothesis testing. Because we designated two outcomes as confirmatory outcomes before starting the analysis, we did not adjust these outcomes to correct for multiple comparisons. By predesignating these variables as of the highest importance, we have essentially placed each of them individually in a "group" and differently avoided concerns about multiple comparisons.

4. Subgroup and robustness analyses

Our main estimation approach used data for the full matched sample. However, we also conducted additional exploratory analyses for select subgroups of students. We examined impacts separately for boys and girls. We also separately examined impacts for students in North Carolina schools and for students who had not had sexual intercourse at the time of the baseline survey.[12] It is important to note that this study was not designed with these subgroup analyses in mind; therefore, this analysis should be considered exploratory. Additionally, the relatively small sample size available for each subgroup analysis means we have limited statistical power to detect impacts for these smaller sets of students.

We also conducted a series of robustness analyses. Our main estimation approach used the most reasonable decisions to produce the best possible estimates of the impacts of Teen PEP. Our robustness analyses explored the sensitivity of our results to these different analytic decisions. In particular, we looked at the following variations on our analytic approach:

1. Conducting regression analysis using the full analytic sample (that is, all individuals completing both a baseline and interim follow-up survey) without trimming or matching based on the propensity-score.
2. Conducting regression analysis using all individuals with a propensity score from 0.05 to 0.95 but without matching based on the propensity score.
3. Estimating separate propensity-score models for each state.
4. Using a more parsimonious set of controls in the regression analysis.
5. Using a linear probability model instead of a logit model to estimate impacts for binary outcome variables.
6. Ignoring corrections for clustering of standard errors and multiple comparisons.
7. Using alternative methods to construct key measures of sexual behavior.

Appendix D presents a more detailed description of these robustness analyses.

[12] There were an insufficient number of observations in the New Jersey subsample or the subsample of individuals who ever had sexual intercourse at baseline to produce meaningful subgroup results.

IV. RESULTS

Our study framework presented in Chapter I hypothesized favorable impacts of Teen PEP on several outcomes. First, the framework suggests that Teen PEP will increase student exposure to information on reproductive health, contraceptive methods, and the prevention of transmitting STIs. We expect that this increased information might then shape intermediate outcomes, including student knowledge, attitudes, and intentions. Changes in these, and other, mediating factors can then reduce risky sexual behavior among youth exposed to Teen PEP.

Our findings suggest that Teen PEP impacted student exposure to information and some intermediate outcomes but did not decrease student sexual risk behavior. Students in schools that implemented Teen PEP reported greater exposure to a wide range of reproductive health topics, including the major components of the Teen PEP curriculum. They also demonstrated improved knowledge of pregnancy and STI-transmission prevention, compared both to their own knowledge at baseline and to the knowledge of students in comparison schools. We further found some evidence that students in intervention schools reported stronger intentions to avoid unprotected sexual activity in the next year compared to those from comparison schools. Despite these observed impacts, we found no evidence that the program reduced sexual activity or unprotected sex at the time of the interim follow-up survey.

A. Teen sexual risk behavior

Our analyses found no statistically significant differences in the prevalence of sexual activity between the intervention and comparison groups (Table IV.1). At the time of our first follow-up survey, 32 percent of students attending Teen PEP schools reported having had intercourse in the three months before the survey, compared to 30 percent of students at comparison schools. The two-percentage point difference between groups is small and not statistically significant ($p = 0.83$). Similarly, we found no evidence of a statistically significant effect on the reported number of sexual partners (a difference of 0.07 partners, $p = 1.00$). Our findings for sexual initiation show higher rates of lifetime sexual activity for students in the intervention group than the comparison group (44 percent versus 37 percent). However, this difference is not statistically significant at the 5 percent level.

Similarly, there is no evidence that Teen PEP influenced the prevalence of unprotected sex at the time of our first follow-up survey (Table IV.1). Twenty-three percent of intervention-group students reported having engaged in sex without a condom during the three months before this survey, compared to 22 percent of comparison-group students. The one percentage-point difference is not statistically significant ($p = 0.93$). Finally, there is no significant difference in the share of students reporting having had sex without any effective method of contraception over the same time period (difference of 4 percentage points, $p = 0.16$).

Table IV.1. Impacts of Teen PEP on teen sexual risk behavior

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Confirmatory outcome				
Had sexual intercourse in the three months before survey	32.1	30.3	1.8	0.83
Other outcomes				
Ever had sexual intercourse[a]	44.0	37.1	6.9	0.15
Number of sexual partners	0.88	0.81	0.07	1.00
Unprotected sex				
Confirmatory outcome				
Had sexual intercourse without a condom in the three months before survey	22.9	22.2	0.6	0.93
Other outcome				
Had sexual intercourse without any effective method of birth control in the three months before survey	19.2	15.2	4.0	0.16
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Impacts are estimated using only the 533 intervention-group and 171 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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. Exposure to Teen PEP messages

Although we did not find evidence that Teen PEP influenced behavior by the time of the interim follow-up survey, we did find strong evidence that the program increased student exposure to a variety of sexual health topics (Table IV.2). Compared to students in comparison schools, students in intervention schools were significantly more likely to report receiving information in the past 12 months on abstinence (76 versus 54 percent), methods of birth control (71 versus 47 percent), where to get birth control (72 versus 45 percent), STIs (82 versus 73 percent), talking about sex with a partner (75 versus 49 percent), saying no to sex (85 versus 72 percent), and how babies are made (86 versus 79 percent). These significant differences range from 8 percentage points (receiving information on how babies are made) to 27 percentage points (receiving information on where to get birth control) and include topics related to all major components of Teen PEP. All *p*-values for tests of the significance of the differences in these outcomes were 0.01 or less.

Table IV.2. Impacts of Teen PEP on exposure to information

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Received any information in past 12 months on				
Relationships	79.2	76.5	2.7	1.00
Abstinence	75.7	54.3	21.4	<0.01**
Methods of birth control	70.9	47.0	23.9	<0.01**
Where to get birth control	72.1	45.4	26.7	<0.01**
STIs	82.3	73.4	8.9	0.01**
Talking about sex with your partner	74.6	48.8	25.8	<0.01**
Saying no to sex	85.2	72.4	12.8	<0.01**
How babies are made	86.4	78.6	7.8	0.01**
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

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

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

These large differences were made possible because of the limited exposure to information that youth appear to have had in the comparison schools where Teen PEP was not offered. Fewer than half of comparison-group students received information on methods of birth control or where to get birth control. Moreover, despite the focus that many North Carolina schools place on abstinence education, only 54 percent of comparison-group students reported receiving information on abstinence in the past 12 months. Instead, comparison group students most commonly received information on how babies are made (79 percent), relationships (77 percent), STIs (73 percent), and saying no to sex (72 percent).

C. Intermediate outcomes

Along with improving exposure to information on sexual risk behavior and related topics, we found that Teen PEP impacted several intermediate outcomes that might be tied to eventual behavior. In particular, there is evidence that Teen PEP increased student knowledge and improved student intentions to avoid unprotected sex. There is little evidence, however, that Teen PEP influenced outcomes within the attitudes, decision making and self-efficacy, communication, or substance use domains.

Knowledge. Teen PEP was associated with increases in knowledge of pregnancy and STI-transmission prevention (Table IV.3). At the interim follow-up, the survey asked students six questions about preventing pregnancy, transmission of HIV, and transmission of other STIs. On average, students from Teen PEP schools answered 3.5 questions correctly, whereas students in comparison schools answered 3.2 questions correctly. The difference of 0.4 questions was statistically significant ($p < 0.01$), indicating that Teen PEP improves student knowledge of these topics. However, both comparison and intervention students reported relatively low levels of knowledge of these key topics. This suggests substantial room for improvement across schools in student knowledge of sexual and reproductive health.

Table IV.3. Impacts of Teen PEP on knowledge

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)[a]	3.5	3.2	0.4	<0.01**
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]This index counts the number of correct responses to a series of six knowledge questions. Possible values range from 0 to 6, with higher values indicating a greater number of correct responses.

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

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

Attitudes. There is no evidence that Teen PEP was associated with changes in student attitudes toward having sex, using birth control, or becoming pregnant (Table IV.4). Students in both study groups reported relatively positive attitudes toward birth control. On a scale of perceptions of factors supporting the use of birth control ranging from 1 (most negative) to 5 (most positive), the mean at the interim follow-up was 4.4 for students from intervention schools and 4.3 for comparison schools. Students also reported relatively few barriers to birth control use. On a scale ranging from 1 (least barriers) to 5 (most barriers), students in the intervention and comparison groups reported scores of 2.5 and 2.6, respectively. Views toward sex at one's current age were also fairly permissive. On a scale of 1 (most permissive) to 5 (least permissive), the mean responses for our index of negative views toward sex at younger ages were 2.8 (intervention group) and 2.9 (comparison group). None of these means were significantly different across study groups ($p > 0.15$).

Although students had fairly positive attitudes toward birth control, attitudes toward pregnancy were mixed. Fifty-two percent of students in Teen PEP schools and 48 percent of students in comparison schools reported they would be "very upset" if they found out they were pregnant or had gotten a partner pregnant (difference of 4 percentage points, $p = 1.00$).

Intentions. Despite the lack of impact on attitudes about sex, pregnancy, and birth control, we did find some evidence that Teen PEP improved intentions to use birth control (Table IV.5). Eighty-three percent of students in Teen PEP schools reported they intended to use a condom if they had sexual intercourse in the next year, compared to 78 percent of students in comparison schools. The difference of 5 percentage points is marginally statistically significant ($p = 0.07$). Similarly, 89 percent of intervention-group students and 85 percent of comparison-group students reported they intended to use some effective method of birth control if they had sex over the same period. The difference of 4 percentage points is also marginally statistically significant ($p = 0.07$). Conversely, students attending intervention and comparison schools reported similar intentions to engage in sexual intercourse during the next year (difference of 2 percentage points, $p = 0.95$). Together, these outcomes suggest that Teen PEP could lead students to engage in unprotected sex less often in the future.

Table IV.4. Impacts of Teen PEP on attitudes toward sex and birth control

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)[a]	4.4	4.3	0.1	0.21
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)[b]	2.5	2.6	–0.1	0.15
Index of negative views toward having sex at current age (average of four survey items; range 1–5)[c]	2.8	2.9	0.0	1.00
Would be very upset if got pregnant or got someone pregnant	51.5	47.9	3.5	1.00
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]This scale averages responses to four questions on support for methods of protection. Possible values range from 1 to 5, with higher values indicating greater support. The inter-item reliability (alpha) equals 0.73.

[b]This scale averages responses to five questions on perceived barriers to methods of protection. Possible values range from 1 to 5, with higher values indicating greater perceived barriers. The inter-item reliability (alpha) equals 0.71.

[c]This scale averages responses to four questions on attitudes towards having sex at one's current age. Possible values range from 1 to 5, with higher values indicating more negative views. The inter-item reliability (alpha) equals 0.75.

*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 IV.5. Impacts of Teen PEP on intentions

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Intend to have sex in the next year	15.5	13.1	2.4	0.95
Intend to use a condom if have sex in the next year	82.7	77.8	5.0	0.07
Intend to use any effective method of birth control if have sex in the next year	88.9	85.0	3.9	0.07
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

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

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

Decision making and self-efficacy. Students in both the intervention and comparison schools reported relatively strong decision-making skills and high levels of self-efficacy (Table IV.6). But Teen PEP was not associated with any significant changes within this domain. For example, on a scale of ability to refuse unwanted sexual advances ranging from 1 (least skills) to 4 (greatest skills), the mean student response on the follow-up survey was 3.3 among intervention-group students and 3.2 among comparison-group students (difference of 0.0, p = 0.79). We found similarly positive but undifferentiated responses for other outcomes in this domain.

Table IV.6. Impacts of Teen PEP on decision making and self-efficacy

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)[a]	3.3	3.2	0.0	0.79
Thoughtfulness in decision making (average of two survey items; range 1–4)[b]	3.1	3.0	0.1	0.71
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)[c]	3.4	3.3	0.1	0.43
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]This scale averages responses to four (boys) or five (girls) questions on perceptions of refusal skills. Possible values range from 1 to 4, with higher values indicating greater perceived refusal skills. The inter-item reliability (alpha) equals 0.78.

[b]This scale averages responses to two questions on thoughtfulness in sexual decision making. Possible values range from 1 to 4, with higher values indicating more thoughtfulness. The inter-item reliability (alpha) equals 0.74.

[c]This scale-score variable is the response to a single question on ability to seek STI testing if needed. Possible values range from 1 to 4, with higher values indicating greater belief that one could seek testing.

*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 note that these findings are based only on student self-perceptions. We know from our analysis of impacts on exposure to program messages (see Table IV.2) that students in Teen PEP schools were significantly more likely to receive information on topics including saying no to sex. We cannot rule out the possibility that this increased exposure might have led to differences in more objective measures of decision making about sexual activity. However, despite any such differences, our findings show that Teen PEP did not add to the confidence students feel in their ability to make positive decisions.

Communication. The results for the communications domain closely mirror those for the decision making and self-efficacy domain. As seen in Table IV.7, Teen PEP was not associated with significant differences in measures of communications with parents, health professionals, or partners about sexual health. This occurs despite Teen PEP's focus on communication about sex, and the finding that Teen PEP increased student exposure to information on communication

about sex (that is, Teen PEP students were more likely than comparison group students to report receiving information on talking about sex and saying no to sex). As with the decision-making and self-efficacy domain, it is important to note that the students themselves reported all outcomes. It is feasible that Teen PEP could have led to changes in more objective or differently nuanced measures of communication and communication skills.

Substance use. One of five Teen PEP workshops explicitly focused on the impacts of alcohol and other drugs on sexual decision making, suggesting the program might affect teen substance use. However, Teen PEP was not associated with any significant changes within the substance use domain (Table IV.8). Thirty-one percent of students in Teen PEP schools reported drinking alcohol in the month before our follow-up survey, compared to 26 percent of comparison-group students. Similarly, 24 percent of intervention-group students and 23 percent of comparison group students reported smoking marijuana over the same period. Both differences (4 percentage points for alcohol, 1 percentage point for marijuana) were statistically insignificant ($p = 0.42$ and $p = 1.00$, respectively).

Table IV.7. Impacts of Teen PEP on communication

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)[a]	4.1	4.6	–0.4	0.13
Received very useful information from parents on relationships or sexual health in past 12 months	56.2	56.3	–0.1	1.00
Spoke to health professional about sexual health in the past 12 months	65.9	71.7	–5.7	0.14
Insufficient communication about sex with partner (0 if no partner)	19.8	20.6	–0.8	1.00
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]This scale-score variable is the response to a single question on communication with parents. Possible values range from 0 to 10, with higher values indicating more frequent communication.

*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 IV.8. Impacts of Teen PEP on substance use

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Drank alcohol in past 30 days	30.7	26.3	4.4	0.42
Used marijuana in past 30 days	23.9	22.5	1.4	1.00
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

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

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

D. Subgroup and robustness analyses

The results of the subgroup analyses for our confirmatory outcomes (whether a student reported having sexual intercourse in the three months before the interim follow-up survey and whether a student reported having sexual intercourse without a condom over this period) mirror those produced in our analysis of the full matched sample, with a small number of exceptions (Table IV.9). Among the subgroup of students who did not report having ever had sex at baseline, those in the intervention group had higher rates of both sexual activity and unprotected sex; however, neither difference is significant at the 5-percent level. Nineteen percent of students at Teen PEP schools who did not report having ever had sex at baseline reported having sex in the three months before the interim follow-up survey, compared to 12 percent of corresponding students in comparison schools (a difference of 6 percentage points, $p = 0.099$). Differences in rates of sex without a condom within this sample were smaller (13 percent versus 9 percent) and statistically insignificant (a difference of 4 percentage points, $p = 0.144$). For the remaining subgroups, we found no significant or notable differences in behavioral outcomes (Table IV.9). For North Carolina students, female students, and male students, impacts of Teen PEP on both confirmatory outcomes are small or moderate (less than 4 percentage points) and statistically insignificant ($p > 0.320$). Results are similar for other outcomes within the sexual behavior domains (see Appendix E for details).

Table IV.9. Impacts of Teen PEP on teen sexual risk behavior for select subgroups

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Students in North Carolina schools				
Had sexual intercourse in the three months before survey	31.7	27.9	3.8	0.320
Had sexual intercourse without a condom in the three months before survey	20.9	18.7	2.2	0.608
Sample size	**582**	**192**		

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Female students				
Had sexual intercourse in the three months before survey	31.7	30.3	1.4	0.898
Had sexual intercourse without a condom in the three months before survey	22.8	23.0	–0.2	0.963
Sample size	**401**	**130**		
Male students				
Had sexual intercourse in the three months before survey	31.8	31.3	0.5	0.953
Had sexual intercourse without a condom in the three months before survey	22.5	21.4	1.1	0.923
Sample size	**313**	**98**		
Students not reporting having ever had sexual intercourse at baseline				
Had sexual intercourse in the three months before survey	18.5	12.4	6.0	0.099
Had sexual intercourse without a condom in the three months before survey	12.8	8.9	3.8	0.144
Sample size	**533**	**171**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. All estimates are in percentages. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level.

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

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

Similarly, the results of our robustness analyses show that our findings from Section A of this chapter remained consistent under a number of alternative analytic decisions. Table IV.10 lists the impacts produced by each alternative set of assumptions we considered for our confirmatory outcomes (whether a student reported having sexual intercourse in the three months before the interim follow-up survey and whether a student reported having sexual intercourse without a condom over this period; for results for all outcomes, see Appendix D). In all but one case, impacts of Teen PEP on sexual behavior remain small to moderate and are statistically insignificant. In the one case of note, Teen PEP was associated with a marginally significant increase in one of the confirmatory outcomes (a 5 percentage point increase in the prevalence of sexual activity, $p = 0.086$). This robustness analysis considered an alternative method for constructing the behavioral outcomes, in which we dropped participants from our analysis if survey items related to sexual activity show a pattern of inconsistent responses.

Table IV.10. Robustness analysis of impacts of Teen PEP on teen sexual risk behavior

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Regression analysis using full analytic sample				
Had sexual intercourse in the three months before survey	34.7	34.5	0.2	0.959
Had sexual intercourse without a condom in the three months before survey	24.5	26.2	–1.7	0.619
Sample size	**977**	**545**		
Regression analysis using propensity-score trimmed sample				
Had sexual intercourse in the three months before survey	35.3	34.4	1.0	0.906
Had sexual intercourse without a condom in the three months before survey	24.5	25.5	–1.0	0.882
Sample size	**714**	**471**		
Estimate propensity score regression within state				
Had sexual intercourse in the three months before survey	31.7	29.4	2.2	0.715
Had sexual intercourse without a condom in the three months before survey	22.5	23.8	–1.3	0.869
Sample size	**714**	**228**		
Use regression with more parsimonious control variables				
Had sexual intercourse in the three months before survey	32.6	29.8	2.7	0.721
Had sexual intercourse without a condom in the three months before survey	23.0	22.1	0.9	0.926
Sample size	**714**	**228**		
Use linear probability model for binary variables				
Had sexual intercourse in the three months before survey	32.6	30.0	2.6	0.668
Had sexual intercourse without a condom in the three months before survey	22.8	22.4	0.4	0.951
Sample size	**714**	**228**		
Do not correct for school-level clustering				
Had sexual intercourse in the three months before survey	32.1	30.3	1.8	0.507
Had sexual intercourse without a condom in the three months before survey	22.9	22.2	0.6	0.802
Sample size	**714**	**228**		

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Alternative coding of contradictory information on sexual activity 1				
Had sexual intercourse in the three months before survey	32.3	30.4	1.9	0.832
Had sexual intercourse without a condom in the three months before survey	23.6	23.2	0.4	0.949
Sample size	**714**	**228**		
Alternative coding of contradictory information on sexual activity 2				
Had sexual intercourse in the three months before survey	32.3	28.2	4.1	0.184
Had sexual intercourse without a condom in the three months before survey	23.3	20.6	2.8	0.505
Sample size	**714**	**228**		
Alternative coding of contradictory information on sexual activity 3				
Had sexual intercourse in the three months before survey	33.3	28.3	5.0	0.086
Had sexual intercourse without a condom in the three months before survey	23.7	21.1	2.6	0.469
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Notes: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Appendix D provides details on the specification of each robustness analysis. All estimates are in percentages. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each intervention school equal weight. *P*-values are corrected for clustering at the school level (unless otherwise noted).

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

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

This page has been left blank for double-sided copying.

V. DISCUSSION AND CONCLUSIONS

This study presents interim impact findings from an ongoing evaluation of Teen PEP, a comprehensive sexual education program that leverages the power of peer influences to shape adolescent behavior. Teen pregnancy prevention experts have long viewed peer-led programs as a particularly promising approach for reducing teen pregnancy and associated sexual risk behaviors. But research on the effectiveness of these programs has produced mixed results.

The interim impact findings presented in this report suggest that Teen PEP succeeded in accomplishing some of its most proximal goals. Students in schools that implemented Teen PEP reported greater exposure to a wide range of reproductive health topics, including the major components of the Teen PEP curriculum. They also improved knowledge of pregnancy and STI/HIV-transmission prevention. We further found a marginally significant difference suggesting that students in intervention schools have stronger intentions to avoid unprotected sexual activity in the next year.

Despite these observed impacts on several key mediating outcomes, we found no evidence that Teen PEP led to decreases in the incidence of sexual activity or unprotected sex at the time of the interim follow-up survey, which was administered about six months after the program concluded. This finding holds under a wide variety of analytic assumptions and for different subgroups.

The results from this study should be viewed with two main limitations in mind. First, although this study was designed as a randomized controlled trial, we ultimately analyzed the data using quasi-experimental propensity-score matching methods. Before implementing propensity-score matching, the intervention and comparison groups demonstrated substantial and statistically significant differences on a number of characteristics, including large differences in past sexual experience. This lack of equivalence was likely due to the relatively high sample attrition. Additionally, given the relatively small number of diverse schools in our sample and that randomization was conducted at the school level, it is possible that the study groups would look different even in the absence of attrition. We attempted to mitigate such issues in New Jersey by randomly assigning matched pairs of schools; however, if schools within a matched pair were not similar (because a given school was very different from all possible matches), this might not have greatly improved balance. Secondly, although our propensity-score matching procedure improved the similarity of our intervention and comparison groups, it did so by sacrificing statistical power. This study was designed assuming a sample size of 2,778 students (Smith and Colman 2012), but our analytic sample included less than 1,000 students. This resulted in much less ability to detect impacts of Teen PEP on all outcomes. It should also be noted that the restriction in sample size decreases the generalizability of our estimates. That is, our results are only applicable to the sample of students included in our final analysis and not all students responding to the study surveys or all students in schools participating in the study.

Beyond these limitations, one possible explanation for our findings is that the interim follow-up survey was administered before meaningful changes in sexual risk behaviors had time to emerge. The statistically significant impacts on measures of exposure to program information and knowledge suggest that Teen PEP workshops were delivered as intended and that the 9th-grade students understood and retained the program messages. Marginally significant impacts of

Teen PEP on intentions to use contraceptives also suggest that the program might impact future contraceptive activity. However, in part because the program is based on a theory of peer influence and school-wide cultural change, it is possible that program impacts on behaviors only emerge over time, as students interact with and are influenced by their peers and the Teen PEP messages permeate school culture. Additionally, at the interim follow-up survey, rates of sexual activity were still relatively low, with only 37 percent of students in comparison schools having ever had sex. Statistically significant impacts of Teen PEP might emerge after a larger proportion of students become sexually active, as commonly occurs as students progress in high school (Kann et al. 2014). We will determine whether impacts change over time using data from our longer-term final follow-up survey (currently in the field).

Our findings might also reflect the specific environments in which the program was implemented. That is, the impacts of Teen PEP estimated here reflect the impacts of the program in the schools in our study and not that in the average school implementing Teen PEP. This distinction is important for both states involved in the analysis, though in different ways. In New Jersey, when the study team began recruitment, the program had already been widely available and implemented in many schools. The schools in our analysis had not previously chosen to use the program, perhaps signaling either that they already had an effective sexual education program in place, or that they had relatively less interest in implementing such a program. Either factor could cause the impact of Teen PEP in study schools to differ from the impact of Teen PEP in the average New Jersey school implementing the program. By contrast, our study schools in North Carolina were some of the first in the state to implement Teen PEP. As "early adopters," these schools naturally experienced some implementation challenges in starting the program, especially given cultural and political differences between North Carolina and New Jersey high schools. Moreover, estimated impacts reflect effects of the program in the first year it was implemented in all of the intervention schools. These impacts might change as school officials learn more about Teen PEP and how to best implement the program in their schools.

Finally, we should note that our findings apply only to the 9th-grade students who received the Teen PEP workshops, not the 11th- and 12th-grade students who served as peer educators. Our focus on the 9th-grade students was driven by the evaluation design and difficulty of identifying a suitable comparison group for the peer educators in comparison schools. However, there are several reasons to expect favorable program effects on these older students. For one, past research suggests peer-led programs might impact peer educators more than other students (Caron et al. 2004; Sieving et al. 2014). In addition, Teen PEP peer educators received the largest "dose" of the program, attending a daily class and sometimes delivering workshops repeatedly to several groups of 9th-grade students. Even if Teen PEP did not impact the sexual behavior of the 9th-grade students receiving the workshops, it might have had large and important impacts on the behavior of the older peer educators.

REFERENCES

Asheer, S., E. Kisker, and B. Keating. "Training Teens and Transforming School Culture Through Comprehensive Sex Education: An Implementation Study of Teen PEP." Princeton, NJ: Mathematica Policy Research, 2014.

Basen-Engquist, K., K. Coyle, G. Parcel, D.B. Kirby, S. Banspach, and S. Carvajal. "Schoolwide Effects of a Multicomponent HIV, STD, and Pregnancy Prevention Program for High School Students." *Health Education and Behavior,* vol. 28, no. 2, 2001, pp. 166–185.

Bearman, P., and H. Brückner. "Power in Numbers: Peer Effects on Adolescent Girls' Sexual Debut and Pregnancy." Washington, DC: The National Campaign to Prevent Teen Pregnancy, 1999.

Caron, F., G. Godin, J. Otis, and L.D. Lambert. "Evaluation of a Theoretically Based AIDS/STD Peer Education Program on Postponing Sexual Intercourse and on Condom Use Among Adolescents Attending High School." *Health Education Research,* vol. 19, no. 2, 2004, pp. 185–197.

Cavanagh, S.E. "The Sexual Debut of Girls in Early Adolescence: The Intersection of Race, Pubertal Timing, and Friendship Group Characteristics." *Journal of Research on Adolescence,* vol. 14, no. 3, 2004, pp. 385–312.

Coyle, K., K. Basen-Engquist, D. Kirby, G. Parcel, S. Banspach, J. Collins, E. Baumler, S. Carvajal, and R. Harrist. "Safer Choices: Reducing Teen Pregnancy, HIV, and STDs." *Public Health Reports,* vol. 116, suppl. 1, 2001, pp. 82–93.

Coyle, K., K. Basen-Engquist, D. Kirby, G. Parcel, S. Banspach, R. Harrist, E. Baumler, and M. Weil. "Short-Term Impact of Safer Choices: A Multicomponent, School-Based HIV, Other STD, and Pregnancy Prevention Program." *Journal of School Health,* vol. 69, no. 5, 1999, pp. 181–188.

Crump, R.K., V.J. Hotz, G.W. Imbens, and O.A. Mitnik. "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika,* vol. 96, no. 1, 2009, pp. 187–199.

DiClemente, R., G. Wingood, K. Harrington, D.L. Lang, S.L. Davies, E.W. Hook 3rd, M.K. Oh, R.A. Crosby, V.S Hertzberg, A.B. Gordon, J.W. Hardin, S. Parker, and A. Robillard. "Efficacy of an HIV Prevention Intervention for African American Adolescent Girls: A Randomized Controlled Trial." *Journal of the American Medical Association,* vol. 292, no. 2, 2004, pp. 171–179.

Duryea, E., P. Mohr, I.M. Newman, G.L. Martin, and E. Egwaoje. "Six-Month Follow-Up Results of a Preventive Alcohol Education Intervention." *Journal of Drug Education,* vol. 14, no. 2, 1984, pp. 97–104.

Funk, M.J., D. Westreich, C. Wiesen, T. Stürmer, M.A. Brookhart, and M. Davidian. "Doubly Robust Estimation of Causal Effects." *American Journal of Epidemiology,* vol. 173, no. 7, 2011, pp. 761–767.

Hoffman, S.D. "Kids Having Kids: Economic Costs and Social Consequences of Teen Pregnancy." Washington, DC: The Urban Institute Press, 2008.

Hothorn, T., F. Bretz, and P. Westfall. "Simultaneous Inference in General Parametric Models." *Biometrical Journal,* vol. 50, no. 3, 2008, pp. 346–363.

Imbens, G.W. "Matching Methods in Practice: Three Examples." *Journal of Human Resources,* vol. 50, no. 2, 2015, pp. 373–419.

Jennings, J.M., S. Howard, and C.L. Perotte. "Effects of a School-Based Sexuality Education Program on Peer Educators: The Teen PEP Model." *Health Education Research,* vol. 29, no. 2, 2014, pp. 319–329.

Kann, L., S. Kinchen, S.L. Shanklin, K.H. Flint, J. Kawkins, W.A. Harris, R. Lowry, E.O. Olsen, T. McManus, D. Chyen, L. Whittle, E. Taylor, Z. Demissie, N. Brener, J. Thornton, J. Moore, and S. Zaza. "Youth Risk Behavior Surveillance—United States, 2013. *Morbidity and Mortality Weekly Report Surveillance Summaries,* vol. 63, suppl. 4, 2014, pp. 1–168.

Kärnä, A., M. Voeten, T.D. Little, E. Poskiparta, E. Alanen, and C. Salmivalli. "Going To Scale: A Nonrandomized Nationwide Trial of the KiVa Antibullying Program for Grades 1–9." *Journal of Consulting and Clinical Psychology,* vol. 79, no. 6, 2011, pp. 796–805.

Killoren, S.E. "Family and Cultural Correlates of Mexican-Origin Youths' Sexual Intentions." *Journal of Youth and Adolescence,* vol. 40, no. 6, 2011, pp. 707–718.

Kim, C.R., and C. Free. "Recent Evaluations of the Peer-Led Approach in Adolescent Sexual Health Education: A Systematic Review." *Perspectives on Sexual and Reproductive Health,* vol. 40, no. 3, 2008, pp. 144–151.

Layzer, C., and L. Rosapep. "Year 2 Interim Evaluation Report: Teen PEP Cohort 1 Participant Experiences." Cambridge, MA: Abt Associates, 2012.

Layzer, C., and L. Rosapep. "Year 3 Interim Evaluation Report: Teen PEP Cohort 2 Participant Experiences." Cambridge, MA: Abt Associates, 2013.

Layzer, C., L. Rosapep, and S. Barr. "A Peer Education Program: Delivering Highly Reliable Sexual Health Promotion Messages in Schools." *Journal of Adolescent Health,* vol. 54, suppl. 3, 2014, pp. S70–S77.

Majmudar, D. "Explaining Adolescent Sexual Risks by Race and Ethnicity: Importance of Individual, Familial, and Extra-Familial Factors." *International Journal of Sociology of the Family,* vol. 31, no. 1, 2005, pp. 19–37.

Martin, J.A., B.E. Hamilton, and S.J. Ventura. "Births: Final Data for 2013." Hyattsville, MD: National Center for Health Statistics, 2015.

Mellanby, A.R., J.B. Rees, and J.H. Tripp. "Peer-Led and Adult-Led School Health Education: A Critical Review of Available Comparative Research." *Health Education Research,* vol. 15, no. 5, 2000, pp. 533–545.

Miller, K.S., R. Forehand, and B.A. Kotchick. "Adolescent Sexual Behavior in Two Ethnic Minority Samples: The Role of Family Variables." *Journal of Marriage and Family,* vol. 61, no. 1, 1999, pp. 85–98.

Mitchell, T.D., and M.W. Greene. "School Health Profiles in North Carolina Middle and High Schools." Atlanta, GA: Centers for Disease Control and Prevention, 2011.

Moulton, B.R. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics,* vol. 72, no. 2, 1990, pp. 334–338.

National Campaign to Prevent Teen and Unplanned Pregnancy. "Counting It Up: The Public Costs of Teen Childbearing." 2013. Available at http://thenationalcampaign.org/why-it-matters/public-cost. Accessed July 23, 2015.

Nicklas, T.A., C.C. Johnson, L. Myers, R.P. Farris, and A. Cunningham. "Outcomes of a High School Program to Increase Fruit and Vegetable Consumption: Gimme 5—A Fresh Nutrition Concept For Students." *Journal of School Health,* vol. 68, no. 6, 1998, pp. 248–253.

Perper, K., K. Peterson, and J. Manlove. "Diploma Attainment Among Teen Mothers." Washington, DC: Child Trends, 2010.

Princeton Center for Leadership Training. "Students Are More Likely to Postpone Sexual Intercourse After Participating in Teen PEP: Program Evaluation Findings, 2002-2003." Princeton, NJ: Princeton Center for Leadership Training, n.d.a.

Princeton Center for Leadership Training. "Teen Prevention Education Program (Teen PEP) Evaluation Findings: Impact on Student Outcomes 2001-2002 Academic Year." Princeton, NJ: Princeton Center for Leadership Training, n.d.b.

Schochet, P.Z. "An Approach For Addressing the Multiple Testing Problem in Social Policy Impact Evaluations." *Evaluation Review,* vol. 33, no. 6, 2009, pp. 539–567.

Schochet, P.Z., A. Duncan, and J.Q. Easton. "Technical Methods Report: The Estimation of Average Treatment Effects for Clustered RCTs of Education Interventions Technical Methods Report: The Estimation of Average Treatment Effects for Clustered RCTs of Education Interventions." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2009.

Scott, M.E., A. Berger, S. Caal, S. Hickman, and K. Moore. "Preventing Teen Pregnancy Among Latinos: Recommendations from Research, Evaluation, and Practitioner Experience." Bethesda, MD: Child Trends, 2014

Sieving, R.E., A.L. McRee, M. Secor-Turner, A.W. Garwick, L.H. Bearinger, K.J. Beckman, B.J. McMorris, and M.D. Resnick. "Prime Time: Long-Term Sexual Health Outcomes of a Clinic-Linked Intervention." *Perspectives on Sexual and Reproductive Health,* vol. 46, no. 2, 2014, pp. 91–100.

Sikkema, K.J., E.S. Anderson, J.A. Kelly, R.A. Winett, C. Gore-Felton, R.A. Roffman, T.G. Heckman, K. Graves, R.G. Hoffmann, M.J. Brondino. "Outcomes of a Randomized, Controlled Community-Level HIV Prevention Intervention for Adolescents in Low-Income Housing Developments." *AIDS,* vol. 19, no. 14, 2005, pp. 1509–1516.

Smith, K., and S. Colman. "Evaluation of Adolescent Pregnancy Prevention Approaches: Design of the Impact Study." Princeton, NJ: Mathematica Policy Research, 2012.

Stephenson, J., V. Strange, E. Allen, A. Copas, A. Johnson, C. Bonell, A. Babiker, and A. Oakley. "The Long-Term Effects of a Peer-Led Sex Education Programme (RIPPLE): A Cluster Randomised Trial in Schools in England." *PLoS Medicine,* vol. 5, no. 11, 2008, e224.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. "What Works Clearinghouse: Procedures and Standards Handbook (Version 3.0)." March 2013. Available at http://whatworks.ed.gov. Accessed March 2, 2015.

APPENDIX A

CONSENT AND RETENTION RATES

This page has been left blank for double-sided copying.

This appendix provides more detailed information on consent, attrition, and response rates. We first discuss school-level attrition and then consider student-level consent and response rates.

Although 30 schools were randomly assigned as part of this evaluation, our analysis sample consisted of only 17 schools due to school-level attrition (Table A.1). Attrition was especially pronounced in the New Jersey sample. Of the 17 high schools initially recruited in New Jersey, 6 schools (35 percent) dropped out for various reasons, including concerns about the study survey, school closure, and a district-wide policy against evaluation. The evaluation also lost another 6 schools (35 percent) from the New Jersey sample because these had been randomized as part of the same matched pair or triplet as the attrited schools. Of the 13 high schools initially recruited in North Carolina, 1 school (8 percent) left the sample before any data collection.

Table A.1. School-level attrition by state and cohort

School	Random assignment group	Lost to follow-up?	Notes
		New Jersey	
Cohort 1			
NJ-1-A	Intervention	Yes	Randomized in a triplet with NJ-1-B and NJ-1-C. Dropped from the sample because NJ-1-C left study.
NJ-1-B	Intervention	Yes	Randomized in a triplet with NJ-1-A and NJ-1-C. Dropped from the sample because NJ-1-C left study.
NJ-1-C	Comparison	Yes	After being assigned to comparison group, the school left the study due to concerns that survey questions were too sensitive.
NJ-1-D	Comparison	Yes	Randomized in a pair with NJ-1-E. Dropped from the sample because NJ-1-E left study.
NJ-1-E	Intervention	Yes	Left study after being unable to implement program.
NJ-1-F	Comparison	Yes	Randomized in a pair with NJ-1-G. Dropped from the sample because NJ-1-G left study.
NJ-1-G	Intervention	Yes	School closed.
Cohort 2			
NJ-2-A	Intervention	Yes	Randomized in a triplet with NJ-2-B and NJ-2-C. Dropped from the sample because NJ-2-B left study.
NJ-2-B	Comparison	Yes	After being assigned to comparison group, the school left the study due to concerns that survey questions were too sensitive.
NJ-2-C	Intervention	Yes	Randomized in a triplet with NJ-2-A and NJ-2-B. Dropped from the sample because NJ-2-B left study.
NJ-2-D	Comparison	No	
NJ-2-E	Intervention	No	
NJ-2-F	Intervention	No	
NJ-2-G	Comparison	Yes	A district-wide policy prevents any school from participation in a study evaluation. The high school principal was unaware when he agreed to participate in Teen PEP that the evaluation would be compulsory, and rescinded the school's participation offer when the situation was clarified.

School	Random assignment group	Lost to follow-up?	Notes
Cohort 3			
NJ-3-A	Intervention	Yes	This school was dropped by the study team due to the large number of schools recruited in North Carolina and budget concerns.
NJ-3-B	Intervention	No	
NJ-3-C	Comparison	No	
NJ-3-D	Intervention	Yes	This is the same school as NJ-1-F that was part of cohort 1 and dropped. The school was re-randomized as part of cohort 3. This school was subsequently dropped from cohort 3 by the study team due to the large number of schools recruited in North Carolina and budget concerns.
North Carolina			
Cohort 1			
NC-1-A	Intervention	No	
NC-1-B	Comparison	No	
NC-1-C	Intervention	No	
NC-1-D	Comparison	Yes	The district contact informed the study team the day before the baseline survey was scheduled to be administered, requesting that the study team wait to administer the survey because of the school board's concerns. Repeated calls and emails to the district contact were not returned, and an agreement on steps to remedy the board's concerns could thus not be reached.
Cohort 2			
NC-2-A	Comparison	No	
NC-2-B	Comparison	No	
NC-2-C	Intervention	No	
NC-2-D	Intervention	No	
Cohort 3			
NC-3-A	Comparison	No	
NC-3-B	Comparison	No	
NC-3-C	Intervention	No	
NC-3-D	Intervention	No	
NC-3-E	Intervention	No	

Source: Teen PEP study information system.

Consent and survey response rates varied across the retained schools. Just over two-thirds of students in consenting schools returned a consent form, and slightly less than half of these students had their parents' consent to participate in the evaluation—49 percent for comparison schools and 48 percent for intervention schools (Table A.2). The consent rate varied across schools from 37 percent to 67 percent. Among consented students, a total of 91 percent completed the baseline survey—87 percent of comparison-group students and 94 percent of intervention-group students. Baseline survey completion rates varied across schools from 61 percent to 99 percent. Retention rates were high for the interim follow-up survey: 85 percent overall, 86 percent for intervention schools, and 84 percent for comparison schools. This retention rate varied across schools from 62 percent to 99 percent. In total, 79 percent of all students whose parents consented to the study completed both the baseline and interim follow-up surveys, including 82 percent of the intervention group and 74 percent of the comparison group.

Table A.2. Consent and retention rates by school and treatment status

School	Treatment status	School left sample	Sample size	Returned forms		Consented		Completed baseline survey		Completed follow-up survey		Completed baseline and follow-up surveys	
				N	%	N	%	N	% (of consent)	N	% (of consent)	N	% (of consent)
North Carolina	**All**		**3,048**	**2,052**	**67**	**1,486**	**49**	**1,338**	**90**	**1,248**	**84**	**1,145**	**77**
Cohort 1	*All*		*821*	*631*	*77*	*467*	*57*	*420*	*90*	*406*	*87*	*367*	*79*
NC-1-A	Intervention	No	255	147	58	108	42	102	94	96	89	91	84
NC-1-B	Comparison	No	300	244	81	200	67	166	83	171	86	142	71
NC-1-C	Intervention	No	266	240	90	159	60	152	96	139	87	134	84
NC-1-D	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Cohort 2	*All*		*1,027*	*611*	*59*	*419*	*41*	*370*	*88*	*358*	*85*	*322*	*77*
NC-2-A	Comparison	No	169	120	71	87	51	86	99	81	93	79	91
NC-2-B	Comparison	No	323	194	60	128	40	86	67	96	75	65	51
NC-2-C	Intervention	No	268	158	59	104	39	102	98	99	95	98	94
NC-2-D	Intervention	No	267	139	52	100	37	96	96	82	82	80	80
Cohort 3	*All*		*1,200*	*810*	*68*	*600*	*50*	*548*	*91*	*484*	*81*	*456*	*76*
NC-3-A	Comparison	No	163	90	55	72	44	69	96	62	86	60	83
NC-3-B	Comparison	No	290	165	57	123	42	108	88	98	80	91	74
NC-3-C	Intervention	No	228	187	82	134	59	124	93	115	86	109	81
NC-3-D	Intervention	No	271	191	70	116	43	101	87	72	62	65	56
NC-3-E	Intervention	No	248	177	71	155	63	146	94	137	88	131	85
New Jersey	**All**		**944**	**635**	**67**	**448**	**47**	**424**	**95**	**396**	**88**	**377**	**84**
Cohort 1			n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-A	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-B	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-C	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-D	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-E	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-F	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-1-G	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

School	Treatment status	School left sample	Sample size	Returned forms		Consented		Completed baseline survey		Completed follow-up survey		Completed baseline and follow-up surveys	
				N	%	N	%	N	% (of consent)	N	% (of consent)	N	% (of consent)
Cohort 2	*All*		*665*	*438*	*66*	*322*	*48*	*300*	*93*	*287*	*89*	*270*	*84*
NJ-2-A	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-2-B	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-2-C	Intervention	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NJ-2-D	Comparison	No	89	71	80	56	63	55	98	48	86	47	84
NJ-2-E	Intervention	No	284	219	77	156	55	144	92	131	84	123	79
NJ-2-F	Intervention	No	292	148	51	110	38	101	92	108	98	100	91
NJ-2-G	Comparison	Yes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Cohort 3			*279*	*197*	*71*	*126*	*45*	*124*	*98*	*109*	*87*	*107*	*85*
NJ-3-B	Intervention	No	124	94	76	56	45	55	98	47	84	46	82
NJ-3-C	Comparison	No	155	103	66	70	45	69	99	62	89	61	87
Total	**All**		**3,992**	**2,687**	**67**	**1,934**	**48**	**1,762**	**91**	**1,644**	**85**	**1,522**	**79**
	Intervention		**2,503**	**1,700**	**68**	**1,198**	**48**	**1,123**	**94**	**1,026**	**86**	**977**	**82**
	Comparison		**1,489**	**987**	**66**	**736**	**49**	**639**	**87**	**618**	**84**	**545**	**74**

Source: Teen PEP study information system.

n.a. = Not applicable.

This page has been left blank for double-sided copying.

APPENDIX B

DATA AND MEASURES

This page has been left blank for double-sided copying.

A. Survey design and administration

The survey instrument was designed to capture a broad range of measures of family background and demographic characteristics, views and attitudes, sexual activity and other youth risk behaviors, and intentions and aspirations. The PPA research team developed the survey, drawing on items found in well-established surveys such as the National Longitudinal Study of Adolescent Health, National Longitudinal Survey of Youth, Youth Risk Behavior Surveillance System, and National Survey of Family Growth. After compiling all relevant items from these surveys, we identified and prioritized those that best served the objectives of the PPA impact study. In some cases, we had to adapt the questions to fit our primary pencil-and-paper survey mode and the age range of our study sample. Most of these adaptations involved changing wording to make questions easier to understand or simplifying the response categories.

We designed the questionnaire so that only students who reported being sexually experienced were asked sensitive questions related to sexual activity. Specifically, the survey was split into three parts. All students completed Part A, which included only general questions about family background and demographic characteristics, views, attitudes, and knowledge. This part of the survey concluded with a single screening question about sexual experience. For the baseline and interim follow-up surveys administered to Cohort 1, this question was "Have you ever had sexual intercourse, oral sex, or anal sex?" In response to concerns about the sensitive nature of this question, we changed this screening question to "Have you ever had sexual intercourse?" for the baseline survey of Cohorts 2 and 3 and "Have you ever had sexual intercourse or oral sex?" for the interim follow-up survey of these cohorts. In all cases, students who answered "yes" to this screening question were instructed to complete Part B1 of the survey, which contained more detailed questions regarding sexual risk behaviors. Students who answered "no" to the screening question were instructed to complete Part B2 of the survey, which included an alternative set of questions. Both Parts B1 and B2 began with a question asking a student to confirm their answer to the screening question and instructed the students to either continue completing the form they selected (if the answer was confirmed) or switch to the alternative form (if not). Parts B1 and B2 of the survey were formatted to look indistinguishable, so that when administering the survey in a group setting, students could not tell which part of the survey other respondents were completing.

As is the case with any self-reported survey, the survey responses might be subject to reporting bias. For this study, we were primarily concerned with the questions relating to sexual behavior, intentions to engage in sexual activity, and attitudes about sex and contraceptive use. For these measures, the reporting bias might occur in either direction. On the one hand, students in the intervention group might be less likely to report risky sexual behaviors because they are embarrassed to admit to a behavior the program discourages. Such underreporting could lead to a spurious finding of *lower* sexual activity among students in the intervention group. On the other hand, especially because our study sample is relatively young, it is possible that the program made students in the intervention group better informed about sexual risk behaviors and therefore more likely to report their true involvement in these behaviors. Such an effect could lead to a spurious finding of *higher* sexual activity rates among students in the intervention group.

We made several different efforts to minimize these risks. To help encourage honest reporting, independent field staff trained and employed by the study team, not the school teachers or anyone else personally connected to the study participants, administered the survey. At the beginning of each survey administration, these staff reminded students that their answers would be kept confidential and encouraged them to respond truthfully. To help maintain their privacy, students were seated at a comfortable distance from their peers during survey administration and placed completed questionnaires in envelopes before handing them over to the survey staff. Questionnaires and return envelopes were labeled with a unique ID number with no personally identifiable information appearing on either. Questions were asked in an objective manner, and the survey instruments were pretested to ensure that questions were worded appropriately for the study sample.

B. Measures of sexual activity

As discussed in Chapter III, we used two confirmatory outcomes and three exploratory outcomes to assess whether Teen PEP reduced rates of risky sexual behavior. A binary (yes/no) indicator of whether a student reported having sexual intercourse in the past three months served as our primary measure of whether exposure to Teen PEP leads to a reduction in sexual activity. Other outcomes used to assess this research question were a binary indicator for whether a student reported ever having sexual intercourse and the number of lifetime sexual partners a student reported. A binary indicator of whether a student reported having sexual intercourse without using a condom in the past three months served as the primary measure of whether exposure to Teen PEP leads to a reduction in unprotected sex. We also used a binary indicator of whether a student reported having sexual intercourse without using any effective method of birth control in the past three months to gauge Teen PEP's impact within this domain.

We coded all measures in two steps and proceeded in parallel for outcomes measured in the baseline and the interim follow-up surveys. First, respondents who answered "no" to the screening question at the end of Part A of the questionnaire (discussed in Section A of this appendix) were coded as not engaging in any sexual activity. Second, respondents who answered "yes" to the screening question were asked a series of questions that we used to determine the values of each of the aforementioned outcomes:

1. The first questions are about sexual intercourse. By sexual intercourse, we mean a male putting his penis into a female's vagina. Have you ever had sexual intercourse? (Due to the different screening question asked in Section A, students in Cohorts 2 and 3 received slight variations on this question at baseline.)
2. Have you had sexual intercourse more than one time?
3. How many DIFFERENT PEOPLE have you ever had sexual intercourse with, even if only one time?
4. In the past 3 months, how many TIMES have you had sexual intercourse?
5. In the past 3 months, how many TIMES have you had sexual intercourse without using a condom?

6. In the past 3 months, how many TIMES have you had sexual intercourse without using any of these methods of birth control: condoms, birth control pills, the shot (Depo-Provera), the patch, the ring (NuvaRing), IUD (Mirena or Paragard), or implants (Implanon)?

For questions 4 to 6, students either checked a box that said "None" or filled in the number of times they engaged in an activity.

Students were instructed to skip any irrelevant questions (for example, students who had not had sex multiple times were not asked how many sexual partners they have had), although they often did not do so when filling out pencil-and-paper surveys. In all baseline surveys and the Cohort 1 interim follow-up survey, students who reported having had sex only one time were not asked to complete questions 3 to 6. In these cases, we assumed that the individual did not have sexual intercourse in the three months before the survey unless they filled out the survey to indicate otherwise.

We used question 1 to determine whether a student ever engaged in sexual intercourse, questions 2 and 3 to determine a student's number of lifetime sexual partners, question 4 to determine whether the student had sex in the three months before the survey, question 5 to determine whether the student had sex without a condom in the three months before the survey, and question 6 to determine whether the student had sex without any effective method of birth control in the three months before the survey.

In constructing these outcomes, we accounted for any observed inconsistent or discrepant responses across different items—for example, participants who reported both never having had sex and having had sex in the past three months. To resolve any inconsistent responses across the six questions, the screening question in Section A, and the confirmation of this question in Section B1 or B2, we developed the following set of rules and procedures:

- **Resolve inconsistencies in responses related to lifetime sexual activity.** We first examined inconsistencies among responses to the screening question, its confirmation, and questions about having ever had sexual intercourse, having had intercourse more than one time, the number of lifetime sexual partners, and the frequency of intercourse in the past three months. In the interim follow-up survey data, we found 98 cases in which two or more of these variables conflicted. We classified 64 of these responses as indicating that the student did not have sex, 27 as indicating that a student did have sex, and the remainder as missing because actual behavior could not be inferred based on a student's other responses. We sometimes also examined responses to questions on the recent frequency of sex without a condom or sex without any effective method of birth control to better understand a student's pattern of responses and choose the response most consistent with the preponderance of evidence the student provided. Additionally, we found 64 cases in which students responded that they had ever had sex at baseline (including cases for which there were some inconsistencies but the preponderance of evidence suggested an individual had been sexually active), but that they had never had sex at interim follow-up. We classified all of these individuals as having had sex at some point in their lifetime at both baseline and follow-up.
- **Resolve inconsistencies in responses related to recent sexual activity.** We then examined inconsistencies between responses to questions about having sexual intercourse, sex without

a condom, and sex without any effective method of birth control in the past three months. In the interim follow-up data, we found 26 cases in which two or more of these outcomes conflicted. We classified 25 of these responses as indicating that the student did have sex in the past three months, after confirming that the student reported having ever had sex. In the one remaining case, we could not determine the student's activity, so we set all three recent sexual activity responses to missing. We also found 24 cases with conflicting information on recent sex without a condom. We classified 11 cases as having engaged in this activity and the remaining 13 as not having done so. Finally, 8 individuals had conflicting responses related to recent sex without any effective method of birth control. We classified 7 of these respondents as having had sex without effective birth control and the remaining respondent as not having done so.

Appendix C explores the robustness of our results to these coding decisions for handling inconsistent responses.

C. Measures of program exposure

The baseline and interim follow-up surveys included a multipart question designed to assess whether students had received information on reproductive health and related topics. The question asked students whether they had received any information in the past 12 months on:

1. Relationships, dating, marriage, or family life.
2. Abstinence from sex.
3. Methods of birth control.
4. Where to get birth control.
5. Sexually transmitted diseases, also known as STDs.
6. How to talk to your partner about whether to have sex or whether to use birth control.
7. How to say no to sex.
8. How babies are made.

For each topic, students answered "yes" or "no". We used responses to this question to create a series of eight binary measures indicating whether a student had received the specified information. For each topic, we coded students who did not respond to the corresponding question as missing.

D. Intermediate outcomes

As discussed in Chapter III, we examined program impacts on six groups of intermediate outcomes: (1) student knowledge of the prevention of pregnancy and STI-transmission; (2) student attitudes toward birth control, sex, and pregnancy; (3) student intentions to engage in sexual activity and use contraception; (4) student decision-making skills and self-efficacy; (5) student communications with parents, health professionals, and partners about sexual health and related topics; and (6) student use of alcohol and marijuana. Outcomes in these domains included those derived from responses to a single survey item and from composites of multiple items.

We constructed all composite measures using principal-component factor analysis and reliability testing. To avoid any correlation between the impacts of Teen PEP and the process to construct the outcome measures, we conducted the factor analysis and reliability testing on the full analytic sample, using responses only from the baseline survey whenever possible. We conducted the factor analysis using the interim follow-up data for some measures in the communication and decision-making and self-efficacy domains because many survey items related to these outcomes were not available in the baseline survey.

To begin this process for each outcome, we conducted an initial, exploratory factor analysis of survey items related to a specific topic and aligning with the Teen PEP program model. We examined the results of this factor analysis to identify the specific questions to include together in different scales, typically excluding items with a factor loading of less than 0.5. If an item's factor loading was slightly under 0.5—but the item seemed conceptually relevant and did not substantially decrease reliability—we included it in a scale. Final decisions about which items to include in each index were driven by both conceptual reasons and by the data. In the remainder of this section, we provide more detail on the construction of these measures.

1. Knowledge

We constructed the measure of student knowledge of preventing pregnancy and STI-transmission from six survey questions, listed in Table B.1. For each item, students received a score of one for a correct answer and zero for an incorrect answer, response of "don't know," or missing response. We summed the individual-item scores for all students who answered at least one of the six questions to get the overall knowledge measure. If a student did not answer any question, the measure was set to missing.

Table B.1. Questions used to construct knowledge measure

Question	Response categories	Correct response
If condoms are used correctly and consistently, how much can they decrease the risk of pregnancy?	Not at all, a little, a lot, completely, don't know	A lot
If condoms are used correctly and consistently, how much can they decrease the risk of getting HIV, the virus that causes AIDS?	Not at all, a little, a lot, completely, don't know	A lot
If birth control pills are used correctly and consistently, how much can they decrease the risk of pregnancy?	Not at all, a little, a lot, completely, don't know	A lot
If birth control pills are used correctly and consistently, how much can they decrease the risk of getting HIV, the virus that causes AIDS?	Not at all, a little, a lot, completely, don't know	Not at all
If birth control pills are used correctly and consistently, how much can they decrease the risk of getting Chlamydia and gonorrhea?	Not at all, a little, a lot, completely, don't know	Not at all
Can you get a sexually transmitted disease or STD from having oral sex?	Yes, no, don't know	Yes

Source: Teen PEP baseline and interim follow-up surveys.

2. Attitudes

We created four scales of student attitudes toward sexual activity, pregnancy, and contraceptive use. These included measures of (1) perceptions of factors supporting the use of birth control, (2) perceptions of barriers to using birth control, (3) negative views toward early sexual activity, and (4) if a student would be very upset about a pregnancy.

We first attempted to measure student attitudes toward birth control. Factor analysis indicated that these attitudes should be measured as two distinct scales: one related to factors supporting birth control use and one related to barriers using birth control.

For the measure of general support for use of birth control, the survey asked respondents whether they agreed or disagreed with each of the following four statements:

1. Condoms should always be used if a person your age has sexual intercourse.
2. Condoms are important to make sex safer.
3. Birth control should always be used if a person your age has sexual intercourse.
4. Birth control is important to make sex safer.

For each statement, the possible response categories were "strongly agree," "agree," "neither agree nor disagree," "disagree," and "strongly disagree." We used responses to these items to create a composite scale of general support for birth control use. In constructing this scale, we omitted two items related to ease of access to condoms and birth control because Teen PEP did not seek to expand access to contraception. For all students responding to at least three of the four survey items, we mapped the categorical responses to a five-point scale and averaged these numeric values across the items with nonmissing responses to create a composite scale of general support for birth control. Higher values on the scale indicate stronger levels of support. We confirmed that the scale had suitable internal reliability (alpha coefficient = 0.69).

For the measure of perceived barriers to using birth control, the survey asked respondents whether they agreed or disagreed with each of the following five statements:

1. Condoms are a hassle to use.
2. Condoms decrease sexual pleasure.
3. Using condoms means you don't trust your partner.
4. Birth control is a hassle to use.
5. Birth control has too many negative side effects.

For each statement, the possible response categories were "strongly agree," "agree," "neither agree nor disagree," "disagree," and "strongly disagree." We used responses to these items to create a composite scale of barriers to birth control use. In constructing this scale, we omitted two items related to the morality of using condoms and birth control because Teen PEP did not seek to change students' moral values. For all students responding to at least four of the five survey items, we mapped the categorical responses to a five-point scale and averaged these numeric values across the items with nonmissing responses to create a composite scale of

perceived barriers to using birth control. Higher values on the scale indicate more substantial barriers. We confirmed that the scale had suitable internal reliability (alpha coefficient = 0.68).

For the measure of attitudes toward early sexual activity, the survey asked students whether they agreed or disagreed with each of the following statements:

1. Having sexual intercourse is a good thing for you to do at your age.
2. At your age right now, having sexual intercourse would create problems.
3. At your age right now, not having sexual intercourse is important for you to be safe and healthy.
4. At your age right now, it is okay for you to have sexual intercourse if you use birth control like a condom.

For each statement, the possible response categories were "strongly agree," "agree," "disagree," and "strongly disagree." We used responses to these items to create a composite scale of negative views toward early sexual activity. In constructing this scale, we omitted one item related to the morality of having sex before marriage as the item did not relate to the Teen PEP curriculum. For all students responding to at least three of the four survey items, we mapped the categorical responses to a four-point scale and averaged these numeric values across the items with nonmissing responses to create a composite scale of negative perceptions toward early sexual activity. Responses to questions 2 and 3 were reverse-coded so that, for all individual items and the scale itself, higher values indicate less permissive attitudes. Results of our factor analysis indicated that the four items on attitudes toward early sexual activity loaded onto one scale. We further examined the internal reliability of the items in this scale and concluded it was high (alpha coefficient = 0.73).

The final measure in this domain is based on one survey question asked of boys and one asked of girls. To measure attitudes toward pregnancy, female students were asked, "If you got pregnant now, how would you feel?" and male students were asked, "If you got someone pregnant now, how would you feel?" The five possible response categories ranged from "very happy" to "very upset." Students who indicated they would be "very upset" received a one for this binary measure, and students who selected another response received a zero. Students who did not respond were coded as missing. We used factor analysis to determine whether this measure could be combined with those used for the scale of attitudes toward sexual activity; however, the analysis confirmed this question should not be included in the scale.

3. Intentions

We included three measures in this domain to assess whether Teen PEP influenced intentions to engage in risky sexual behavior in the near future. These measures are based on three survey questions:

1. Do you intend to have sexual intercourse in the next year?
2. If you have sexual intercourse in the next year, do you intend to use (or have your partner use) a condom?

3. If you have sexual intercourse in the next year, do you intend to use (or have your partner use) any of these methods of birth control: birth control pills, the shot (Depo-Provera), the patch, the ring (NuvaRing), IUD (Mirena or Paragard), or implants (Implanon)?

The response categories for all three questions were "yes, definitely"; "yes, probably"; "no, probably not"; and "no, definitely not." For each question, we constructed a binary measure comparing students who responded "yes, definitely" to students who responded otherwise. Students who did not respond to the question were coded as missing. If a student responded "yes, definitely" to question 2 (use of condoms) but not question 3 (use of any effective birth control method), we set the binary indicators for intention to use condoms and intention to use any effective birth control method to one.

Responses to questions about intentions to have sex and intentions to use birth control were treated as independent. That is, we did not use the response to the question which asked whether the student intended to have sex in determining the coding of the other intentions measures. The questions were phrased such that a student could logically reply to all items.

4. Decision making and self-efficacy

Three measures are included in this domain: (1) a refusal-skill scale, (2) a scale capturing thoughtfulness in sexual decision making, and (3) a scale-scored variable capturing student self-efficacy in seeking reproductive health care.

To measure refusal skills, we constructed a composite measure from up to five survey questions. For three of the survey questions, students were first asked to "imagine you are alone with someone you like very much." They were then asked:

1. How likely is it that you could stop them if they wanted to touch your private parts below the waist, meaning the parts of the body covered by underwear, and you did not want them to do that?
2. How likely is it that you could avoid having sexual intercourse if you didn't want to?

Girls were also asked:

3. How likely is it that you could stop them if they wanted to touch your chest and you did not want them to do that?

For each statement, the possible response categories were "not at all likely," "a little bit likely," "somewhat likely," and "very likely." The interim follow-up survey contained two additional, related items. Students were asked whether they agreed or disagreed with each of the following statements:

4. If my partner refused to use condoms, I could refuse to have sex.
5. I would have sex now if someone I cared about pressured me to have sex.

For each statement, the possible response categories were "strongly agree," "agree," "disagree," and "strongly disagree." We mapped the categorical responses to a four-point scale, with higher

numbers representing stronger refusal skills for all survey questions. Responses were averaged across all questions answered to create the index. At the baseline survey, we required that students answered all applicable questions to have a nonmissing value for this measure. For the follow-up survey, we coded a respondent as missing this index if he or she did not answer more than one of the four (for boys) or five (for girls) applicable survey items. Factor analysis confirmed that the items load strongly onto a single construct in both the baseline and interim follow-up data; the scale also has high internal reliability (alpha coefficient = 0.89 at baseline and 0.78 at interim follow-up).

We used two survey items to create a composite of students' thoughtfulness in making decisions about sexual activity:

1. When you have to make a decision about your sexual behavior, how often do you think of the consequences of each possible choice?
2. When you have to make a decision about your sexual behavior, how often do you first get as much information as you can?

These survey items, and thus this index, were available only for the interim follow-up survey. For each item, the possible response categories were "very often," "often," "not often," and "never." We mapped the categorical responses to a four-point scale and averaged these numeric values to create a composite measure ranging from 1 to 4, with higher values indicating more thoughtfulness in making decisions about sex. For students who did not answer one or both items, we set this scale to missing. The scale has high internal reliability (alpha coefficient = 0.74), and factor analysis confirmed that the two survey items load onto a single construct.

Finally, the interim follow-up survey asked students whether they agreed or disagreed with a single survey item related to their ability to seek sexual health services: "I believe I could go to a clinic if I needed to get tested for HIV/AIDS or another sexually transmitted disease (STD)." The four response categories for this question ranged from "strongly agree" to "strongly disagree." Factor analysis confirmed that this item should not be considered as part of either other scale within this domain. We mapped the categorical responses for this single item to a four-point scale, with 1 representing the least agreement and 4 representing the most agreement. We used the scaled-score variable to measure student self-efficacy in seeking reproductive health care. This variable is not available at baseline.

5. Communication

The survey items enabled us to create three measures of communication about sex and reproductive health topics in the follow-up data and two measures in the baseline data.

At both baseline and interim follow-up, we constructed a measure of communications with doctors, nurses, and other health care practitioners. For both surveys, we created a binary variable equal to one if an individual discussed sexual health topics with a doctor, nurse, or other clinic practitioner in the past 12 months. The questions on which these measures were based differed slightly across surveys. At baseline, the survey asked students, "In the past 12 months, have you spoken with a doctor or nurse about sex, birth control, or sexually transmitted diseases, also known as STDs?", to which they could respond "yes" or "no." At follow-up, the survey

asked students, "Thinking about the past 12 months, how many times did you get information on relationships, abstinence, birth control, or sexually transmitted diseases…[from a] doctor, nurse, or clinic?" and could select "never," "1–3 times," "4–9 times," and "10 or more times." In both cases, we coded individuals who did not respond to the question as missing this measure.

We used a single, multipart survey question at baseline to measure communication with parents about sex and risky behaviors. The survey asked students, "In the past 12 months, how many TIMES have you talked with at least one of your parents or guardians about…?" and then presented nine topics related to sex and risky behaviors:

1. How things are going with school work or with your grades.
2. A personal problem you were having.
3. How to have good romantic relationships.
4. Strategies for safe dating.
5. How to resist pressures to have sex.
6. Avoiding drugs and alcohol.
7. Pregnancy or birth.
8. Sexually transmitted diseases (also known as STDs), HIV, or AIDS.
9. Whether you should be having sex at this time in your life.

Response categories for each item included "never," "1–2 times," "3–9 times," and "10 or more times," which we coded at the midpoint of that category's range (0, 1.5, 6, and 10, respectively). We used the average responses to items 4 to 9 to create a composite scale ranging from 0 to 10, where higher values indicate more discussion with parents. The first three items were omitted to focus the index on discussions about risky behavior. For students who did not respond to two or more items, we set the scale to missing. A factor analysis confirmed that the items load strongly onto a single construct. The scale also has high internal reliability (alpha coefficient = 0.89).

At follow-up, this multipart survey item was not available. Instead, the survey asked students, "Thinking about the past 12 months, how many times did you get information on relationships, abstinence, birth control, or sexually transmitted diseases…[from] parents and other relatives or family members?" Students could select "never," "1–3 times," "4–9 times," and "10 or more times." We constructed the follow-up measure of communication with parents using the midpoints of the categories provided (0, 2, 6.5, or 10). We coded students who did not respond to the survey item as missing.

Additionally, for the interim follow-up data only, we created a binary measure of whether students received helpful information on sexual health or related topics from their parents. The survey asked students, "Thinking about the past 12 months, where did you get information on relationships, abstinence, birth control, or sexually transmitted diseases that was very helpful to you?" If a student selected "Parents and other relatives or family members" from a list of options, he or she received a one for this measure. All other students were coded to zero.

Finally, the interim follow-up survey contained a multipart item related to communicating with a partner about sex. Specifically, the follow-up survey asked students "In the last month, how often have you talked with your partner about each of the topics listed below?"

1. Expectations in the relationship.
2. Pregnancy.
3. Birth control.
4. Sexually transmitted diseases (STDs).
5. What you feel comfortable doing sexually.
6. What you do not feel comfortable doing sexually.

The three response categories included "often," "sometimes," and "never." Based on this question, we created a binary indicator measuring whether a student had insufficient communication with his or her partner. We set this measure to one for students who did not discuss at least one of the six topics "often" with a partner and to zero for students who discussed at least one topic "often." The measure equals zero for students who indicated that they did not have a partner in the past month. We coded students who did not indicate they had no partner and did not respond to any of the six survey items as missing for this measure. We additionally explored creating a scale based on categorical responses to the six different topics; however, candidate indices typically did not have high internal consistency.

6. Substance use

The two substance use measures focus on a student's use of alcohol or marijuana in the 30 days before completing the survey. Outcomes were based on two questions:

1. During the past 30 days, not including any times you just had a sip, on how many days did you have one or more alcoholic beverages?
2. During the past 30 days, on how many days did you use marijuana, also called weed or pot?

The possible response categories for each question were "More than 25 days," "5 to 25 days," "1 to 4 days," and "0 (zero) days". We used these responses to create two binary indicators (one for marijuana and one for alcohol) equal to one if the student reported using the substance one or more times in the past 30 days and zero if the student responded otherwise. We set measures for students who did not respond to these questions to missing.

E. Baseline covariates

As shown in Chapter III, we constructed a broad range of measures using data from the baseline survey to assess the equivalence of our study groups, select a propensity-score matched sample, and adjust for remaining differences between the intervention- and comparison-group students when producing impact estimates. We define these variables in Table B.2.

Table B.2. Baseline characteristics of study sample

Measure	Definition
Demographic and background characteristics	
Age	Continuous variable: approximate age when student completed baseline survey, calculated based on the date when the baseline survey was completed and the student's reported month and year of birth.
Female	Binary variable: equals 1 if a student is female; equals 0 if a student is male.
Hispanic	Binary variable: equals 1 if a student self-identifies as Hispanic or Latino/a; equals 0 otherwise.
White, non-Hispanic	Binary variable: equals 1 if a student self-identifies as white, non-Hispanic; equals 0 otherwise.
Black, non-Hispanic	Binary variable: equals 1 if a student self-identifies as black, non-Hispanic; equals 0 otherwise.
Other race/ethnicity	Binary variable: equals 1 if a student self-identifies as non-Hispanic and American Indian, Alaskan Native, Asian, Native Hawaiian, or Pacific Islander, or is of multiple race/ethnicities; equals 0 otherwise.
Race missing	Binary variable: equals 1 if a student's race/ethnicity is missing; equals 0 if a student's race/ethnicity is nonmissing.
Main language spoken at home is English	Binary variable: equals 1 if English is the main language spoken at the student's home; equals 0 if English is not the main language spoken at the student's home.
Biological mother living in home or main home	Binary variable: equals 1 if a student lives with his or her biological mother; equals 0 if a student does not live with his or her biological mother.
Biological father living in home or main home	Binary variable: equals 1 if a student lives with his or her biological father; equals 0 if a student does not live with his or her biological father.
Biological parents currently married	Binary variable: equals 1 if a student's biological parents are currently married to each other; equals 0 if a student's biological parents are not currently married to each other.
Biological parents currently divorced or separated	Binary variable: equals 1 if a student's biological parents were once married to each other but are now separated or divorced; equals 0 if a student's biological parents are not currently separated or divorced.
Report religion is very important in their life	Binary variable: equals 1 if a student reported that religion was "very" important in his or her life; equals 0 if a student reported that religion was "somewhat important" or "not at all important."
Identify as lesbian, gay, bisexual, asexual, or questioning	Binary variable: equals 1 if a student reported that he or she is a lesbian, gay, bisexual, asexual, or questioning; equals 0 if a student reported that he or she is heterosexual.
Ever smoked a cigarette	Binary variable: equals 1 if a student reported having ever smoked a cigarette; equals 0 if a student reported never having smoked a cigarette.
Ever drank alcohol	Binary variable: equals 1 if a student reported ever having more than a sip of an alcoholic drink; equals 0 if a student reported never having more than a sip of an alcoholic drink.
Ever smoked marijuana	Binary variable: equals 1 if a student reported having ever smoked marijuana; equals 0 if a student reported never having smoked marijuana.
Measures of sexual behavior	
Ever had sexual intercourse	Binary variable: equals 1 if a student reported that he or she has ever had sexual intercourse; equals 0 if a student reported that he or she has not had sexual intercourse.

Measure	Definition
Number of lifetime sexual partners	Count variable: equals the total number of sexual partners the participant has ever had.
Had intercourse in the past three months	Binary variable: equals 1 if a student reported that he or she had intercourse at least once in the three months before completing the survey; equals 0 if a student reported that he or she did not have intercourse in the three months before completing the survey.
Had intercourse without a condom in the past three months	Binary variable: equals 1 if a student reported that he or she had intercourse without a condom at least once in the three months before completing the survey; equals 0 if a student reported that he or she did not have intercourse without a condom in the three months before completing the survey.
Had intercourse without using any effective method of birth control in the past three months	Binary variable: equals 1 if a student reported that he or she had intercourse without any effective form of birth control at least once in the three months before completing the survey; equals 0 if a student reported that he or she did not have intercourse without any effective form of birth control in the three months before completing the survey.
Ever had oral sex	Binary variable: equals 1 if a student reported that he or she has ever had oral sex; equals 0 if a student reported that he or she has not had oral sex. This construct is only available at baseline for Cohort 1.
Ever had nonpenetrative sex	Binary variable: equals 1 if a student reported that he or she has ever had nonpenetrative sex; equals 0 if a student reported that he or she has not had nonpenetrative sex.
Ever kissed member of opposite sex	Binary variable: equals 1 if a student reported that he or she has ever kissed a member of the opposite sex; equals 0 if a student reported that he or she has not kissed a member of the opposite sex.
Exposure to information	
Received information about relationships	Binary variable: equals 1 if a student reported having received any information about “relationships, dating, marriage, or family life” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about abstinence	Binary variable: equals 1 if a student reported having received any information about “abstinence from sex” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about birth control methods	Binary variable: equals 1 if a student reported having received any information about “methods of birth control” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about where to get birth control	Binary variable: equals 1 if a student reported having received any information about “where to get birth control” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about STIs	Binary variable: equals 1 if a student reported having received any information about “sexually transmitted diseases, also known as STDs” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about talking with your partner about sex	Binary variable: equals 1 if a student reported having received any information about “how to talk to your partner about whether to have sex or whether to use birth control” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Received information about saying no to sex	Binary variable: equals 1 if a student reported having received any information about “how to say no to sex” in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.

Measure	Definition
Received information about how babies are made	Binary variable: equals 1 if a student reported having received any information about "how babies are made" in the past 12 months; equals 0 if a student reported having not received information on this topic in the past 12 months.
Knowledge	
Knowledge of preventing STI transmission and pregnancy	Continuous index: sum of correct responses to six survey questions; variable ranges from 0 to 6, with higher values indicating greater knowledge. For students who do not respond to all six items, index is set to missing; otherwise, missing responses were counted as incorrect.
Attitudes	
Perceptions of factors supporting use of birth control	Continuous scale: average of responses to four survey questions; variable ranges from 1 to 5, with higher values indicating more supportive attitudes towards contraception.
Perceptions of barriers to use of birth control	Continuous scale: average of responses to five survey questions; variable ranges from 1 to 5, with higher values indicating more perceived barriers toward contraception use.
Negative views toward early sexual activity	Continuous scale: average of responses to four survey questions; variable ranges from 1 to 5, with higher values indicating less permissive attitudes toward early sexual activity.
Would be very upset if got pregnant or got someone pregnant	Binary variable: equals 1 if a student reported he or she would be "very upset" if she became pregnant or he impregnated someone now; equals 0 if a student chose another response category (indicating he or she would be less upset or happy about a pregnancy).
Intentions	
Intend to have sex in the next year	Binary variable: equals 1 if a student reported he or she will "definitely" have sex in the next year if he or she has the chance; equals 0 if a student reported he or she will "probably", "probably not" or "definitely not" do so.
Intend to use a condom if have sex in the next year	Binary variable: equals 1 if a student reported he or she will "definitely" use a condom if he or she has sex in the next year; equals 0 if a student reported he or she will "probably", "probably not" or "definitely not" do so.
Intend to use any effective method of birth control if have sex in the next year	Binary variable: equals 1 if a student reported he or she will "definitely" use an effective method of contraception if he or she has sex in the next year; equals 0 if a student reported he or she will "probably", "probably not" or "definitely not" do so.
Decision making and self-efficacy	
Perceptions of refusal skills	Continuous scale: average of two questions for boys and three questions for girls; variable ranges from 1 to 4, with higher values indicating greater perceived refusal skills.
Communication	
Frequency of discussions with parents about relationships or sexual health in past 12 months	Continuous scale: based on seven survey questions; variable ranges from 0 to 10, with higher values indicating more communication.
Spoke to health professional about sexual health in the past 12 months	Binary variable: equals 1 if a student reported he or she spoke with a doctor or nurse about sex, birth control, or sexually transmitted diseases in the past 12 months; equals 0 if a student reported otherwise
Substance use	
Drank alcohol in past 30 days	Binary variable: equals 1 if a student reported drinking alcohol one or more times in the past 30 days; equals 0 if a student reported he or she did not do so.

Measure	Definition
Smoked marijuana in past 30 days	Binary variable: equals 1 if a student reported smoking marijuana one or more times in the past 30 days; equals 0 if a student reported he or she did not do so.

Our regression analysis controlled for the following subset of possible covariates, defined in Table B.2: a baseline measure of the outcome or a close proxy from the same domain (if available); an indicator for whether this variable is missing; cohort; state; whether the student had sex, had sex without a condom, or had sex without any effective method of birth control in the past three months; number of past sexual partners; whether the student ever had sex, ever had nonpenetrative sex, or ever kissed someone of the opposite gender; main language spoken at home; whether the student lived with their biological mother or biological father; gender; race; age; biological mother's and father's marital status; religiosity; identification as nonheterosexual; information received in the past 12 months on abstinence, how babies are made, talking about sex, saying no to sex, birth control methods, and STIs; knowledge of STI transmission and pregnancy prevention; perceived barriers to using birth control; attitudes toward sex; whether the student would be upset if she became pregnant or he got someone pregnant; whether the student intended to have sex in the next year, use a condom if he or she had sex in the next year, or use any effective method of birth control if he or she had sex in the next year; perceived refusal skills; communication with parents; whether the student ever drank alcohol, smoked marijuana, or smoked cigarettes, or did so in the past month; and indicators for students not reporting information on whether they ever engaged in nonpenetrative sex; whether the student ever drank alcohol, smoked marijuana, or smoked cigarettes, or did so in the past month; and whether the student identifies as heterosexual.

F. Rates of missing data

As described in Chapter III.B, for both the propensity-score regressions and the main regression analysis, we imputed any missing covariate values to their overall mean to maximize the size of the sample used in analysis. Table B.3 shows the rates at which each measure from the baseline survey used in this analysis was missing. We provide rates for the entire analytic sample (that is, all students who completed both the baseline and follow-up survey) and for the subset of students in the propensity-score matched sample. Table B.4 shows the rates at which each outcome measure from the interim follow-up survey was missing.

These tables demonstrate that missing data rates for most outcomes and covariates were relatively low and reasonably consistent between the intervention and comparison groups. For most measures, rates of missing data are less than 10 percent. Individuals tended to be least likely to respond to questions about substance use, their sexual orientation, and their race. But many of the sensitive, behavioral outcomes we focus on in this analysis had reasonably low item-specific nonresponse.

Table B.3. Rates of missing data in baseline survey

	Full analytic sample		Propensity-score matched sample	
	Intervention group	Comparison group	Intervention group	Comparison group
Demographic and background characteristics				
Age at baseline	2.56	2.57	1.75	2.52
Female	0.10	0.00	0.00	0.00
Race	13.20	11.56	8.77	9.10
Main language spoken at home	1.54	1.10	0.88	0.98
Biological mother living in home or main home	1.23	1.28	1.75	0.98
Biological father living in home or main home	1.23	1.28	1.75	0.98
Biological parents marital status	8.80	8.81	8.77	7.70
Religious importance	1.84	1.28	0.88	1.82
Identify as lesbian, gay, bisexual, asexual, or questioning	11.36	14.13	12.72	12.18
Ever smoked a cigarette	10.44	12.66	12.72	11.2
Ever drank alcohol	10.75	12.84	13.60	11.48
Ever smoked marijuana	11.05	13.03	12.28	11.76
Baseline measures of sexual behavior				
Ever had sexual intercourse	0.82	1.83	2.63	0.84
Number of lifetime sexual partners	3.07	6.24	5.26	3.50
In the three months before survey had sexual intercourse	1.74	2.39	3.07	1.96
In the three months before survey had sexual intercourse without a condom	1.74	2.39	3.07	1.96
In the three months before survey had sexual intercourse without any effective method of birth control	1.74	2.39	3.07	1.96
Ever had oral sex (cohort 1 only)	2.22	2.11	1.52	2.07
Ever had nonpenetrative sex	12.18	15.05	16.67	11.62
Ever kissed member of opposite sex	12.18	15.05	16.67	11.62
Baseline measures of exposure to information				
Received any information in past 12 months on				
Relationships	1.13	3.12	3.07	0.98
Abstinence	2.66	6.42	6.58	2.10
Methods of birth control	2.87	3.85	3.95	2.94
Where to get birth control	3.07	4.77	4.39	3.08
STIs	2.46	4.40	4.82	2.10
Talking about sex with your partner	2.76	6.06	5.70	2.80
Saying no to sex	2.46	4.04	4.82	1.96
How babies are made	2.76	3.85	3.95	2.80

	Full analytic sample		Propensity-score matched sample	
	Intervention group	Comparison group	Intervention group	Comparison group
Baseline measures of intermediate outcomes				
Knowledge of preventing STI transmission and pregnancy	1.13	2.20	1.32	0.84
Perceptions of factors supporting use of birth control	4.61	9.72	8.33	4.62
Perceptions of barriers to use of birth control	8.39	16.15	14.04	8.40
Index of negative views toward having sex at current age	3.07	5.87	4.39	2.80
Would be very upset if got pregnant or got someone pregnant	1.94	3.85	3.51	1.54
Intend to have sex in the next year	3.48	6.42	5.26	3.08
Intend to use a condom if have sex in the next year	4.20	7.52	8.77	3.64
Intend to use any effective method of birth control if have sex in the next year	3.89	7.89	7.89	3.50
Perceptions of refusal skills	6.35	10.64	9.21	5.46
Frequency of discussions with parents about relationships or sexual health in past 12 months	2.97	6.06	5.70	2.66
Spoke to health professional about sexual health in past year	10.03	13.21	13.6	10.22
Drank alcohol in past month	11.05	13.21	14.04	11.62
Smoked marijuana in past month	11.57	13.39	13.16	12.18

Source: Teen PEP baseline survey.

Table B.4. Rates of missing data in interim follow-up survey

	Full analytic sample		Propensity-score matched sample	
	Intervention group	Comparison group	Intervention group	Comparison group
Sexual activity				
Ever had sexual intercourse	0.82	1.83	2.63	0.84
Number of lifetime sexual partners	3.07	6.24	5.26	3.50
In the three months before survey had sexual intercourse	1.74	2.39	3.07	1.96
In the three months before survey had sexual intercourse without a condom	1.74	2.39	3.07	1.96
In the three months before survey had sexual intercourse without any effective method of birth control	1.74	2.39	3.07	1.96
Exposure to information				
Received any information in past 12 months on				
Relationships	1.54	2.57	1.68	1.32
Abstinence	2.05	4.95	2.10	2.63
Methods of birth control	2.25	5.14	2.38	3.07
Where to get birth control	2.35	4.59	2.38	2.63
STIs	1.74	4.59	1.68	2.19
Talking about sex with your partner	2.66	5.69	2.66	3.51
Saying no to sex	2.05	4.40	1.96	3.95
How babies are made	2.56	3.49	2.52	2.63
Intermediate outcomes				
Knowledge of preventing STI transmission and pregnancy	0.92	0.92	0.84	0.00
Perceptions of factors supporting use of birth control	3.07	6.61	3.22	4.82
Perceptions of barriers to use of birth control	4.91	8.44	4.90	6.58
Index of negative views toward having sex at current age	2.35	4.77	2.66	4.82
Would be very upset if got pregnant or got someone pregnant	1.94	1.83	2.24	1.32
Intend to have sex in the next year	2.76	3.12	2.24	2.63
Intend to use a condom if have sex in the next year	1.94	4.40	1.68	2.63
Intend to use any effective method of birth control if have sex in the next year	1.94	3.30	1.96	1.75
Perceptions of refusal skills	4.30	8.07	4.90	7.02
Thoughtfulness in decision making	3.68	7.16	3.92	7.02
Believe could go to clinic to seek STI testing if needed	2.56	6.42	2.52	4.39

	Full analytic sample		Propensity-score matched sample	
	Intervention group	Comparison group	Intervention group	Comparison group
Frequency of discussions with parents about relationships or sexual health in past 12 months	16.17	25.50	17.23	24.56
Received very useful information from parents on relationships or sexual health in past 12 months	0.20	0.00	0.14	0.00
Spoke to health professional about sexual health in past year	16.07	26.24	17.23	25.44
Insufficient communication about sex with partner (0 if no partner)	1.84	2.94	1.68	3.07
Drank alcohol in past month	6.24	5.50	6.58	5.26
Smoked marijuana in past month	7.16	6.61	7.42	5.70

Source: Teen PEP interim follow-up survey.

This page has been left blank for double-sided copying.

APPENDIX C

CHARACTERISTICS OF ALL STUDENTS RESPONDING TO SURVEYS

This page has been left blank for double-sided copying.

To investigate the equivalence of our intervention and comparison groups before the implementation of propensity-score matching or trimming, we examined baseline demographic characteristics and measures of our outcomes among all survey respondents in the 17 high schools that were retained in the study. We estimated means for the intervention- and comparison-group students, weighting so that each school received equal weight. We tested whether the means were statistically different using t-statistics, correcting for clustering at the school level using Stata's cluster adjustment. The analysis included all students who consented and responded to both the baseline and interim follow-up surveys.

Attrition and the small number of schools randomly assigned likely explain most differences between the intervention and comparison groups. As detailed in Appendix A, school-level attrition was high in this evaluation. We lost 13 of 30 schools in the study during the period between random assignment and the baseline survey. The bias caused by this attrition is mitigated to some extent by our blocked random assignment design. In many cases, we grouped schools into pairs or triplets and conducted random assignment within these groups. When one school in a pair or triplet left the study, the other schools were dropped from our sample. If schools within a pair or triplet are very similar, this should reduce concerns about bias related to attrition. However, when grouping schools into pairs or triplets, we faced two major issues. First, the number of schools was relatively small, making it difficult to match the characteristics of schools within a pair or triplet. Second, we only had limited data available on school characteristics in order to form the matches. Thus, schools within a matched pair or triplet might have differed based on a number of characteristics.

Overall, we found that students in the intervention and comparison groups differed on many key characteristics (Table C.1). Students in intervention schools were 0.3 years older than those in comparison schools, a statistically significant difference ($p = 0.01$). Racial composition also varied between intervention and comparison schools. In intervention schools, 30 percent of students were Hispanic, 25 percent were black non-Hispanic, and 32 percent were white non-Hispanic. In comparison schools, the composition was 24 percent Hispanic, 43 percent black non-Hispanic, and 25 percent white non-Hispanic. The difference in share black is large (18 percentage points) and marginally statistically significant ($p = 0.09$). Students in intervention schools were also more likely to live with their biological mother and to live with their biological father, with both differences marginally statistically significant ($p \leq 0.10$). Students in the comparison group were also more likely to have reported having ever smoked cigarettes (difference of 9 percentage points, $p = 0.03$).

Table C.1. Demographic and background characteristics of full sample

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Average age	15.0	15.4	–0.3*	0.01
Female	55.4	57.2	–1.9	0.38
Race				
Hispanic	30.1	23.6	6.4	0.57
White, non-Hispanic	31.9	25.1	6.9	0.54
Black, non-Hispanic	24.9	42.9	–18.0	0.09
Other race/ethnicity	12.9	7.9	4.9	0.12
Race missing	0.2	0.5	–0.3	0.40
Main language spoken at home is English	82.3	84.2	–2.0	0.82
Biological mother living in home or main home	87.9	81.4	6.5*	0.07
Biological father living in home or main home	53.5	45.2	8.3*	0.10
Biological parents currently married	49.1	40.1	8.9*	0.08
Biological parents divorced or separated	27.1	24.1	2.9	0.41
Report religion is very important in their life	37.9	43.0	–5.0	0.22
Identify as lesbian, gay, bisexual, asexual, or questioning	12.1	16.7	–4.5	0.21
Ever smoked cigarettes	25.2	34.5	–9.3*	0.03
Ever drank alcohol	49.2	53.3	–4.1	0.20
Ever smoked marijuana	26.7	30.0	–3.3	0.41
Sample size	**977**	**545**		

Source: Teen PEP baseline survey.

Note: Estimates are percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

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

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

Students in different study groups also differed on measures of baseline sexual activity (Table C.2). In the three months before the baseline survey, students in comparison schools were 11 percentage points more likely to have had sex, 9 percentage points more likely to have had sex without a condom, and 8 percentage points more likely to have had sex without any effective method of birth control. All three differences are statistically significant ($p \leq 0.02$). Additionally, only 23 percent of intervention-group students had ever had sex, compared to 33 percent of comparison-group students, a statistically significant difference of 10 percentage points ($p = 0.05$). Individuals in the comparison group were also less likely to report having engaged in nonpenetrative sex or having kissed a member of the opposite sex (both significant differences, with $p = 0.03$).

Table C.2. Baseline measures of sexual risk behavior for full sample

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Outcomes analyzed at follow-up				
Ever had sexual intercourse	23.0	33.4	–10.3*	0.05
Number of lifetime sexual partners	0.72	1.21	–0.49	0.09
In the three months before survey				
Had sexual intercourse[a]	12.8	23.6	–10.8*	0.01
Had sexual intercourse without a condom[a]	9.0	17.8	–8.8*	0.02
Had sexual intercourse without any effective method of birth control	7.2	15.4	–8.2*	0.01
Other sexual behavior				
Ever had oral sex (cohort 1 only)	25.1	25.9	–0.8	0.66
Ever had nonpenetrative sex	49.1	62.5	–13.4*	0.03
Ever kissed member of opposite sex	82.6	89.3	–6.7*	0.03
Sample size	**977**	**545**		

Source: Teen PEP baseline survey.

Note: Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

[a]Designated as a confirmatory outcome.

*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 found statistically significant differences in the baseline measures of many intermediate outcomes (Table C.3). Students in intervention schools were significantly more likely to report being exposed to information on both STIs and how babies are made in the year before our survey (both differences of 8 percentage points with $p = 0.02$). Intervention students also reported significantly more negative views toward sexual initiation (a difference of 0.2 on a five-point scale, $p = 0.03$), and marginally significantly fewer barriers to contraceptive use (a difference of 0.1 on a five-point scale, $p = 0.09$). Students in the different study groups also reported different intentions to avoid sex or unprotected sex. For example, only 9 percent of individuals in the intervention group reported they intended to have sex in the next year, compared to 15 percent of comparison-group students (difference of 6 percentage points, $p = 0.02$). The baseline data additionally reveal that although the comparison group appears to be more likely to participate in risky sexual behavior, they are also more likely to communicate with health practitioners and parents about this behavior. For example, 36 percent of the comparison group spoke with a health professional about reproductive health in the past year, compared to 29 percent of intervention group students (a difference of 7 percentage points, $p = 0.04$). Finally, rates of nonsexual, risky behavior, namely alcohol and marijuana use, were also higher in the comparison group than in the intervention group. Students in the comparison group were 9 percentage points more likely to have drank alcohol ($p < 0.01$) and 7 percentage points more likely to have smoked marijuana ($p = 0.06$) in the 30 days before survey, compared to students in the intervention group.

Table C.3. Baseline measures of intermediate outcomes for full sample

Variable	Intervention group mean	Comparison group mean	Difference	*p*-value
Received any information in past 12 months on				
Relationships	86.0	85.7	0.3	0.87
Abstinence	68.1	62.8	5.3	0.15
Methods of birth control	49.8	49.4	0.4	0.92
Where to get birth control	42.3	42.8	–0.5	0.88
STIs	80.9	73.2	7.7*	0.02
Talking about sex with your partner	49.8	52.8	–3.0	0.46
Saying no to sex	73.5	72.1	1.4	0.71
How babies are made	84.8	76.8	8.0*	0.02
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	2.8	2.5	0.2	0.08
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.3	4.3	0.0	0.69
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.6	–0.1	0.09
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	3.0	2.8	0.2*	0.03
Would be very upset if got pregnant or got someone pregnant	54.1	49.9	4.2	0.39
Intend to have sex in the next year	8.6	14.5	–5.9*	0.02
Intend to use a condom if have sex in the next year	83.4	78.5	4.9*	0.03
Intend to use any effective method of birth control if have sex in the next year	89.1	85.1	4.1	0.07
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.1	3.0	0.1	0.36
Spoke to health professional about sexual health in past year	29.2	36.2	–6.9*	0.04
Communication with parents about risky behavior (average of six survey items; range 0–10)	2.7	3.3	–0.6*	0.01
Drank alcohol in past 30 days	25.2	34.6	–9.4**	<0.01
Smoked marijuana in past 30 days	14.7	21.5	–6.8	0.06
Sample size	**977**	**545**		

Source: Teen PEP baseline survey.

Note: Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. Estimates are weighted to give students in each school equal weight. *P*-values corrected for clustering at the school level.

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

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

APPENDIX D

ROBUSTNESS ANALYSIS

This page has been left blank for double-sided copying.

The main impact findings presented in Chapter IV of this report are derived from a particular set of analytic decisions, ranging from the data-cleaning procedures used to construct the outcome measures to the specification of the regression models. We made these decisions in accordance with established research standards and the particular features of our study design. However, we also investigated the sensitivity of our results to alternative analytic decisions. In this appendix, we present findings from three types of sensitivity tests. First, we examined the sensitivity of our results to the specification of the propensity score model and regression models used to estimate program impacts. We then considered how our results would change under alternative methods for calculating standard errors and statistical significance tests. We concluded by examining the robustness of our results to alternative data-cleaning procedures for the measures of sexual risk behavior.

Altogether, our robustness analysis confirms that Teen PEP did not lead to a decrease in risky teen behaviors but did increase knowledge and exposure to information. There is less evidence that the observed impacts of Teen PEP on intentions to avoid unprotected sex are robust to changes in specifications. Some evidence also suggests that Teen PEP could have led to earlier sexual initiation; however, this result does not hold across most analytic strategies, including our most-preferred strategy presented in Chapter III and IV.

A. Alternative model specifications

For our main findings, we used a multistep procedure to estimate impacts of Teen PEP. We first estimated a propensity-score regression using all individuals with baseline and interim follow-up data (the full analytic sample) and then removed any observations for which the propensity score was less than 0.05 or greater than 0.95 (the propensity-score trimmed sample). Next, we matched each individual in the intervention group to the student in the comparison group with the closest propensity score (generating the propensity-score matched sample). Finally, we obtained impact estimates using this sample of matched observations. This appendix section explores the robustness of our results to changes in this procedure.

First, we considered whether the regression results obtained for the propensity-score matched sample hold in the full analytic sample. That is, we examined the impacts of Teen PEP that are generated by conducting regression analysis on the full analytic sample. Table D.1 contains the results of this analysis. Overall, results are similar to our main findings: there is little evidence that Teen PEP led to a decrease in sexual behavior but strong evidence the program increased exposure to information and knowledge. We also found some evidence that Teen PEP might have changed some other intermediate outcomes; however, none of the effects estimated in the attitudes, intentions, decision making and self-efficacy, communications, and substance use domains is significant at the 5 percent level.

Table D.1. Impacts of Teen PEP on outcomes, full sample

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	34.7	34.5	0.2	0.959
Ever had sexual intercourse[b]	48.2	44.5	3.7	0.341
Number of sexual partners	1.2	1.1	0.1	1.000
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	24.5	26.2	–1.7	0.619
Had sexual intercourse without any method of birth control in the three months before survey	19.9	20.4	–0.5	0.925
Exposure to information				
Received any information in past 12 months on				
Relationships	79.7	71.6	8.2	0.080
Abstinence	75.6	55.1	20.5**	<0.001
Methods of birth control	71.1	51.0	20.1**	<0.001
Where to get birth control	72.6	48.3	24.3**	<0.001
STIs	81.3	69.4	11.9**	0.001
Talking about sex with your partner	74.9	51.9	23.0**	<0.001
Saying no to sex	83.8	65.0	18.8**	<0.001
How babies are made	84.8	78.3	6.5*	0.047
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.4	2.9	0.5**	<0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.526
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.6	0.0	1.000
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.8	0.0	1.000
Would be very upset if got pregnant or got someone pregnant	46.4	44.5	1.9	1.000
Intentions				
Intend to have sex in the next year	18.5	18.4	0.0	1.000
Intend to use a condom if have sex in the next year	81.9	77.5	4.4	0.213

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Intend to use any method of birth control if have sex in the next year	88.0	83.8	4.2	0.084
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.2	3.1	0.1	0.095
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.0	2.9	0.2	0.095
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.4	0.1	0.288
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.1	4.5	–0.4	0.081
Received very useful information from parents on relationships or sexual health in past 12 months	54.2	50.4	3.8	0.583
Spoke to health professional about sexual health in the past 12 months	65.7	68.3	–2.6	1.000
Insufficient communication about sex with partner (0 if no partner)	20.7	21.8	–1.1	1.000
Substance use				
Drank alcohol in past 30 days	27.8	24.2	3.6	0.389
Smoked marijuana in past 30 days	23.9	22.4	1.5	1.000
Sample size	**977**	**545**		

Source: Teen PEP first follow-up survey, full analytic sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 752 intervention-group and 373 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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 next considered the impacts of Teen PEP that would result from estimating the same regression but instead using the propensity-score trimmed dataset (Table D.2). Again, Teen PEP is associated with increases in knowledge and exposure to information but no change in sexual behavior. Using this specification, we did not find evidence that Teen PEP led to increases in intentions to use contraception, as we did in our main analysis. Somewhat counterintuitively, when we used the propensity-score trimmed dataset, we also found that Teen PEP was associated with a decrease in communication with parents. The decrease in communications contradicts Teen PEP's focus on improving student-parent conversations but is likely a replica of this analysis incompletely controlling for differences in the characteristics of the intervention and comparison groups. In all analyses that use the propensity-score matched sample and properly adjust for multiple comparisons, we found no impacts of Teen PEP on outcomes in the communication domain.

Table D.2. Impacts of Teen PEP on outcomes, propensity-score trimmed sample, no matching

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	35.3	34.4	1.0	0.906
Ever had sexual intercourse[b]	48.6	44.3	4.4	0.432
Number of sexual partners	1.2	1.1	0.1	1.000
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	24.5	25.5	–1.0	0.882
Had sexual intercourse without any method of birth control in the three months before survey	20.0	19.3	0.7	0.906
Exposure to Information				
Received any information in past 12 months on				
Relationships	80.0	72.3	7.7	0.092
Abstinence	75.1	55.0	20.2**	<0.001
Methods of birth control	71.8	50.8	20.9**	<0.001
Where to get birth control	72.6	47.6	25.0**	<0.001
STIs	81.2	68.9	12.3**	0.001
Talking about sex with your partner	74.8	51.9	22.9**	<0.001
Saying no to sex	84.0	66.5	17.5**	<0.001
How babies are made	85.2	78.7	6.5*	0.016
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.4	2.9	0.5**	<0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.0	1.000
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.5	0.0	1.000
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.8	0.0	1.000
Would be very upset if got pregnant or got someone pregnant	46.8	44.6	2.2	1.000
Intentions				
Intend to have sex in the next year	18.9	18.5	0.4	1.000
Intend to use a condom if have sex in the next year	81.1	77.0	4.2	0.358

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Intend to use any method of birth control if have sex in the next year	87.6	83.7	3.9	0.185
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.2	3.2	0.1	0.385
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.0	2.9	0.1	0.239
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.4	0.0	1.000
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.0	4.7	–0.7**	0.003
Received very useful information from parents on relationships or sexual health in past 12 months	54.2	52.3	1.9	1.000
Spoke to health professional about sexual health in the past 12 months	65.6	70.0	–4.4	0.425
Insufficient communication about sex with partner (0 if no partner)	20.5	21.7	–1.2	1.000
Substance use				
Drank alcohol in past 30 days	29.6	23.9	5.7	0.119
Smoked marijuana in past 30 days	24.1	22.9	1.2	0.992
Sample size	**714**	**471**		

Source: Teen PEP first follow-up survey, propensity-score trimmed sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 533 intervention-group and 333 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

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

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

Our main analytic approach used a single propensity-score regression to estimate the propensity score. We also explored estimating the propensity-score regression separately for students in North Carolina and New Jersey, as different factors might be important in predicting treatment status across states (Table D.3). We obtained very similar results using this method: Teen PEP is associated with increases in student exposure to information and knowledge but does not influence outcomes in any other domain. The only substantive deviation from our main findings is that the impacts of Teen PEP on student intentions to use condoms and other contraceptive methods are not marginally significant in this analysis.

Table D.3. Impacts of Teen PEP on outcomes, estimate propensity score models by state

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	31.7	29.4	2.2	0.715
Ever had sexual intercourse[b]	44.1	37.2	6.8	0.142
Number of sexual partners	0.9	0.7	0.2	0.246
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	22.5	23.8	–1.3	0.869
Had sexual intercourse without any method of birth control in the three months before survey	18.0	16.5	1.5	0.830
Exposure to information				
Received any information in past 12 months on				
Relationships	77.8	67.7	10.1	0.161
Abstinence	75.5	48.2	27.3**	<0.001
Methods of birth control	70.2	51.2	18.9*	0.020
Where to get birth control	70.3	47.1	23.2**	<0.001
STIs	81.1	67.2	13.8**	0.002
Talking about sex with your partner	74.5	50.4	24.0**	<0.001
Saying no to sex	84.4	67.5	16.9**	<0.001
How babies are made	82.7	78.5	4.2	0.714
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.6	3.0	0.6**	<0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.366
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.5	–0.1	1.000
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.8	0.0	1.000
Would be very upset if got pregnant or got someone pregnant	47.7	45.9	1.7	1.000
Intentions				
Intend to have sex in the next year	15.8	11.8	4.0	0.457
Intend to use a condom if have sex in the next year	81.3	79.1	2.2	1.000

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Intend to use any method of birth control if have sex in the next year	87.3	83.9	3.4	0.265
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.2	3.2	0.0	1.000
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.0	3.0	0.0	1.000
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.5	3.4	0.0	1.000
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.5	4.4	0.1	NA
Received very useful information from parents on relationships or sexual health in past 12 months	56.1	55.9	0.2	NA
Spoke to health professional about sexual health in the past 12 months	68.3	65.5	2.9	NA
Insufficient communication about sex with partner (0 if no partner)	20.7	22.6	–1.9	NA
Substance use				
Drank alcohol in past 30 days	29.4	28.3	1.1	1.000
Smoked marijuana in past 30 days	18.6	24.3	–5.7	0.368
Sample size	**647**	**437**		

Source: Teen PEP first follow-up survey, sample matched using propensity score estimated by state.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 485 intervention-group and 312 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

NA. Not available. *P*-value could not be calculated due to instability in the variance-covariance matrix of estimates.

*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 explored the sensitivity of our estimates to using a more parsimonious set of regression covariates. Our regression attempted to control for a large number of variables. Although this can lead to increased precision and decreased bias, there might be a concern that we are overfitting the regression model. Table D.4 thus contains the impact estimates obtained from using fewer covariates in the regression, in particular controlling only for a baseline measure of the outcome, an indicator for whether this variable is missing, and the linear variables used in the propensity-score matching procedure. Again, there is little change in the substantive conclusions implied by the analysis: Teen PEP does not impact either of our confirmatory outcomes, although it does increase exposure to information and knowledge.

This robustness analysis does present some evidence that Teen PEP might have adverse impacts on sexual initiation and communication about sex. When one limits the set of control variables, as in Table D.4, Teen PEP is associated with an increase in the share of students reporting having ever had sex and a decrease in the share of students who spoke recently with a health care professional about sexual health topics. The latter impact is likely a product of this specific regression specification because in no other case where standard errors are properly estimated do we find that Teen PEP was associated with a decrease in communications with health professionals.

The positive impact of Teen PEP on the share of students who have ever had sex is more suggestive. Even though this impact is only marginally statistically significant, several robustness analyses demonstrate similar results. In this case, 45 percent of students in the intervention group report ever having sex in the interim follow-up data, compared to 36 percent of the comparison group, a difference of 9 percentage points ($p = 0.065$).

Table D.4. Impacts of Teen PEP on outcomes, regression only controls for linear terms from propensity-score regression

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	32.6	29.8	2.7	0.721
Ever had sexual intercourse[b]	45.0	36.3	8.8	0.065
Number of sexual partners	0.9	0.8	0.1	0.719
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	23.0	22.1	0.9	0.926
Had sexual intercourse without any method of birth control in the three months before survey	19.2	15.2	4.0	0.267
Exposure to information				
Received any information in past 12 months on	.	.	.	.
Relationships	80.4	75.3	5.1	0.695
Abstinence	76.4	54.2	22.2**	<0.001
Methods of birth control	72.2	46.3	25.9**	<0.001
Where to get birth control	72.7	44.9	27.8**	<0.001
STIs	81.9	74.0	7.9*	0.022
Talking about sex with your partner	75.2	48.3	26.9**	<0.001
Saying no to sex	85.2	72.7	12.5**	<0.001
How babies are made	86.3	78.7	7.6**	0.006
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.5	3.2	0.4**	0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.114
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.6	–0.1	0.069
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.9	0.0	1.000
Would be very upset if got pregnant or got someone pregnant	50.9	48.6	2.4	1.000
Intentions				
Intend to have sex in the next year	15.6	13.0	2.6	0.820
Intend to use a condom if have sex in the next year	82.0	78.7	3.3	0.393
Intend to use any method of birth control if have sex in the next year	88.7	85.3	3.4	0.117
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.3	3.2	0.1	0.302

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.0	3.0	0.1	0.732
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.3	0.1	0.115
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.2	4.6	–0.4	0.270
Received very useful information from parents on relationships or sexual health in past 12 months	57.3	55.1	2.1	1.000
Spoke to health professional about sexual health in the past 12 months	65.0	72.6	–7.6*	0.039
Insufficient communication about sex with partner (0 if no partner)	19.6	20.8	–1.2	1.000
Substance use				
Drank alcohol in past 30 days	28.9	28.2	0.7	1.000
Smoked marijuana in past 30 days	22.6	24.0	–1.3	1.000
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics, including a baseline measure of the outcome or a close proxy from the same domain (if available); an indicator for whether this variable is missing; cohort; state; whether the student had sex, had sex without a condom, or had sex without any method of birth control in the past three months; number of past sexual partners; whether the student ever had nonpenetrative sex; main language spoken at home; gender; race; age; biological mother's and father's marital status; religiosity; information received in the past 12 months on how babies are made, talking about sex, saying no to sex, birth control methods, and STIs; knowledge of STI/HIV transmission and pregnancy prevention; perceived barriers to using birth control; attitudes toward sex; whether the student would be upset if he or she became pregnant or got someone pregnant; perceived refusal skills; communication with parents; whether the student ever drank alcohol or smoked marijuana or did so in the past month; and indicators for students not reporting information on whether they ever engaged in nonpenetrative sex and whether the student ever drank alcohol or smoked marijuana or did so in the past month. Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 533 intervention-group and 171 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

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

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

Finally, we explored using different regression models to estimate the impacts of Teen PEP. In our main analysis, we used logistic regression when analyzing binary outcomes. In this robustness analysis, we instead used a linear probability model. We also explored using a Poisson regression to look at impacts of Teen PEP on the one count measure that we analyze (number of sexual partners). The resulting regression specification did not converge. Thus, for this analysis, we chose to replace the number of sexual partners measure with an indicator for a student having more than one lifetime sexual partner. As for other binary outcomes, we used a linear probability model to estimate the impacts of Teen PEP in this robustness analysis.

The results are very similar to those seen for the main analysis with one exception: Teen PEP is associated with a marginally significant increase in the share of students who report having ever had sex at interim follow-up (Table D.5). In the comparison group, 33 percent of students reported having ever had sex, compared to 40 percent of intervention-group students (difference of 6 percentage points, $p = 0.079$). This provides further evidence that Teen PEP might increase early sexual initiation.

Table D.5. Impacts of Teen PEP on outcomes, linear probability model used for binary outcomes

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	32.6	30.0	2.6	0.668
Ever had sexual intercourse	39.7	33.4	6.3	0.079
Have had multiple sexual partners	14.1	17.4	–3.3	0.284
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	22.8	22.4	0.4	0.951
Had sexual intercourse without any method of birth control in the three months before survey	19.3	15.1	4.2	0.081
Exposure to information				
Received any information in past 12 months on				
Relationships	79.2	76.7	2.6	1.000
Abstinence	77.1	54.4	22.6**	<0.001
Methods of birth control	71.2	47.5	23.6**	<0.001
Where to get birth control	72.4	45.8	26.6**	<0.001
STIs	82.0	74.2	7.8**	0.013
Talking about sex with your partner	74.9	49.1	25.8**	<0.001
Saying no to sex	86.5	72.4	14.1**	<0.001
How babies are made	87.2	78.4	8.8*	0.018
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.5	3.2	0.4**	0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.215
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.6	–0.1	0.154
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.9	0.0	1.000
Would be very upset if got pregnant or got someone pregnant	51.2	48.5	2.7	1.000
Intentions				
Intend to have sex in the next year	15.7	13.2	2.5	0.780
Intend to use a condom if have sex in the next year	83.3	77.4	5.8	0.100
Intend to use any method of birth control if have sex in the next year	89.4	84.8	4.6	0.086
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.3	3.2	0.0	0.786

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.1	3.0	0.1	0.706
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.3	0.1	0.426
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.1	4.6	–0.4	0.135
Received very useful information from parents on relationships or sexual health in past 12 months	55.8	56.6	–0.8	1.000
Spoke to health professional about sexual health in the past 12 months	65.8	71.4	–5.5	0.105
Insufficient communication about sex with partner (0 if no partner)	19.9	20.5	–0.5	1.000
Substance use				
Drank alcohol in past 30 days	30.6	26.6	4.0	0.550
Smoked marijuana in past 30 days	23.5	23.2	0.3	1.000
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

*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. Alternative estimates of standard errors and *p*-values

For the main findings presented in Chapter IV of this report, we adjusted the statistical significance tests (*p*-values) to account for two statistical issues. First, we adjusted the standard errors to account for the cluster random assignment design. That is, because schools (and not students) were randomly assigned to the intervention and comparison groups, regression error terms might be correlated across students in the same school (Schochet et al. 2009). More generally, there is always concern about correlations in regression error terms when one variable (in this case, treatment status) does not vary across individuals in a specific group (in this case, those in the same school; see Moulton 1990). Second, we also adjusted our *p*-values to correct for multiple hypothesis testing within domains, using a procedure outlined by Hothorn et al. (2008) and Schochet (2009).

To examine the sensitivity of our results to these adjustments, we estimated comparable regression models under three alternative conditions: (1) no adjustment for multiple hypothesis testing, (2) no adjustment for clustering, and (3) no adjustment for clustering *or* multiple-hypothesis testing. The results of these analyses (Table D.6) show that both corrections are important. When one does not adjust standard errors as needed, estimated impacts of Teen PEP on many outcomes become statistically significant. When one ignores one or both of these important adjustments, Teen PEP is associated with a statistically significant or marginally significant increase in the share of students reporting ever having sexual intercourse but, conversely, significant or marginally significant improvements in many outcomes in the mediating factor domains, including those related to knowledge, attitudes, and intentions. The program is also associated with significant decreases in communication with parents and health professionals when one ignore the multiple comparison correction. Given the rich literature on how to correctly estimate standard errors, these estimates largely suggest the importance of properly accounting for clustering and multiple comparisons in estimating program impacts, rather than decreasing the credibility of our main findings.

Table D.6. Impacts of Teen PEP on outcomes, different methods to estimate standard errors

Variable	Intervention group mean	Comparison group mean	Difference	p-value: Ignoring multiple comparisons	p-value: Ignoring clustering	p-value: Ignoring both
Sexual activity						
Had sexual intercourse in the three months before survey[a]	32.1	30.3	1.8	0.830	0.507	0.507
Ever had sexual intercourse[b]	44.0	37.1	6.9	0.074	0.056	0.016*
Number of sexual partners	0.88	0.81	0.07	0.600	1.000	0.520
Unprotected sex						
Had sexual intercourse without a condom in the three months before survey	22.9	22.2	0.6	0.933	0.802	0.802
Had sexual intercourse without any method of birth control in the three months before survey	19.2	15.2	4.0	0.161	0.155	0.155
Exposure to information						
Received any information in past 12 months on						
Relationships	79.2	76.5	2.7	0.372	1.000	0.399
Abstinence	75.7	54.3	21.4	<0.001**	<0.001**	<0.001**
Methods of birth control	70.9	47.0	23.9	<0.001**	<0.001**	<0.001**
Where to get birth control	72.1	45.4	26.7	<0.001**	<0.001**	<0.001**
STIs	82.3	73.4	8.9	0.001**	0.179	0.012*
Talking about sex with your partner	74.6	48.8	25.8	<0.001**	<0.001**	<0.001**
Saying no to sex	85.2	72.4	12.8	<0.001**	0.007**	<0.001**
How babies are made	86.4	78.6	7.8	0.001**	0.104	0.005**
Knowledge						
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.5	3.2	0.4	0.001**	<0.001**	<0.001**
Attitudes						
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.053	0.312	0.060
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.5	2.6	–0.1	0.037*	0.223	0.039*

Variable	Intervention group mean	Comparison group mean	Difference	p-value		
				Ignoring multiple comparisons	Ignoring clustering	Ignoring both
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.9	0.0	0.382	1.000	0.342
Would be very upset if got pregnant or got someone pregnant	51.5	47.9	3.5	0.384	1.000	0.326
Intentions						
Intend to have sex in the next year	15.5	13.1	2.4	0.315	0.900	0.285
Intend to use a condom if have sex in the next year	82.7	77.8	5.0	0.023*	0.236	0.060
Intend to use any method of birth control if have sex in the next year	88.9	85.0	3.9	0.026*	0.351	0.098
Decision making and self-efficacy						
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.3	3.2	0.0	0.261	1.000	0.365
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.1	3.0	0.1	0.234	0.658	0.202
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.3	0.1	0.141	0.895	0.283
Communication						
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.1	4.6	–0.4	0.033*	0.772	0.174
Received very useful information from parents on relationships or sexual health in past 12 months	56.2	56.3	–0.1	0.980	1.000	0.980
Spoke to health professional about sexual health in the past 12 months	65.9	71.7	–5.7	0.035*	0.599	0.131
Insufficient communication about sex with partner (0 if no partner)	19.8	20.6	–0.8	0.855	1.000	0.801

Variable	Intervention group mean	Comparison group mean	Difference	p-value		
				Ignoring multiple comparisons	Ignoring clustering	Ignoring both
Substance use						
Drank alcohol in past 30 days	30.7	26.3	4.4	0.211	0.382	0.172
Smoked marijuana in past 30 days	23.9	22.5	1.4	0.625	1.000	0.610
Sample size	**714**	**228**				

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain, unless otherwise specified.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 533 intervention-group and 171 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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. Alternative data-cleaning procedures

As described in Appendix B.B, our analysis of the self-reported survey data uncovered some inconsistent or discrepant responses to the questions on sexual risk behaviors. For example, it was possible for a participant to report having not had sex in the past three months but having had sex without a condom over the same period. For the main impact findings presented in this report, we accounted for these discrepancies when creating our outcome measures by considering the preponderance of evidence across all relevant questions in the survey (see Appendix B.B for a more detailed description). However, we also examined the sensitivity of our results to three alternative methods for cleaning the data:

- **Alternative coding 1:** Coding a participant as having engaged in a specific behavior if *any* survey item indicates he or she did so.
- **Alternative coding 2:** Coding a participant as *not* having engaged in a specific behavior if any survey item indicates he or she did *not* do so.
- **Alternative coding 3:** Dropping a participant from the analysis if the survey items show a pattern of inconsistent responses.

The results of these analyses showed that our findings might be sensitive to how we resolve discrepancies in the survey data (Table D.7-D.9). In particular, each alternative suggested a pattern of impacts on sexual behavior that deviates from that seen in our main results. Using the first coding, Teen PEP is associated with a 12 percentage point increase in the share of students who report ever having had sex ($p = 0.009$). Using the second or final coding, Teen PEP is associated with a 6 percentage point increase in the share of students who report having had sex without using any effective method of birth control in the past three months ($p = 0.022$ and $p = 0.020$, respectively). When we use the final coding method, the program is also associated with marginally significant increases in the share of students who report having had sex within the past three months (5 percentage points, $p = 0.086$) or having ever had sex (8 percentage points, $p = 0.060$).

Table D.7. Impacts of Teen PEP on outcomes, alternative coding of sexual behaviors 1

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	32.3	30.4	1.9	0.832
Ever had sexual intercourse[b]	50.7	39.0	11.6**	0.009
Number of sexual partners	0.9	0.8	0.1	1.000
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	23.6	23.2	0.4	0.949
Had sexual intercourse without any method of birth control in the three months before survey	19.3	15.4	3.9	0.264
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 519 intervention-group and 166 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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 D.8. Impacts of Teen PEP on outcomes, alternative coding of sexual behaviors 2

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	32.3	28.2	4.1	0.184
Ever had sexual intercourse[b]	38.3	32.9	5.4	0.306
Number of sexual partners	0.9	0.7	0.2	0.181
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	23.3	20.6	2.8	0.505
Had sexual intercourse without any method of birth control in the three months before survey	19.8	14.2	5.6*	0.022
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 570 intervention-group and 177 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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 D.9. Impacts of Teen PEP on outcomes, alternative coding of sexual behaviors 3

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	33.3	28.3	5.0	0.086
Ever had sexual intercourse[b]	42.7	34.4	8.3	0.060
Number of sexual partners	0.9	0.7	0.2	0.168
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	23.7	21.1	2.6	0.469
Had sexual intercourse without any method of birth control in the three months before survey	19.9	14.4	5.5*	0.020
Sample size	**714**	**228**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 570 intervention-group and 177 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

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

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

This page has been left blank for double-sided copying.

APPENDIX E

SUBGROUP-SPECIFIC ESTIMATES

This page has been left blank for double-sided copying.

In addition to examining all outcomes for all students, we also considered the impacts of Teen PEP separately within the North Carolina sample, for female students, for male students, and for students who did not report having ever had sexual intercourse at baseline.

Estimates within the North Carolina sample (Table E.1) deviate slightly from the findings for the sample of students from both states. Within the smaller sample, Teen PEP is still associated with increases in student exposure to information and knowledge. But for North Carolina students, there is also some evidence of adverse impacts of Teen PEP in other domains. In particular, at interim follow-up, compared to students in comparison schools, students in intervention schools exhibited less negative views toward having sex at their current age (difference of 0.1 points on a 5-point scale, $p = 0.015$), and were less likely to have discussed sexual health topics with a health professional in the past 12 months (difference of 9 percentage points, $p = 0.024$). Students in the intervention schools were also more likely to report having ever had sex (difference of 8 percent points), although this difference was not statistically significant at the 5-percent level.

Table E.1. Impacts of Teen PEP on outcomes, North Carolina students only

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	31.7	27.9	3.8	0.320
Ever had sexual intercourse[b]	42.7	35.0	7.7	0.091
Number of sexual partners	1.0	0.8	0.2	0.304
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	20.9	18.7	2.2	0.608
Had sexual intercourse without any method of birth control in the three months before survey	17.0	14.4	2.5	0.425
Exposure to information				
.Received any information in past 12 months on				
Relationships	79.1	75.6	3.5	1.000
Abstinence	76.5	54.0	22.6**	<0.001
Methods of birth control	71.3	50.5	20.8**	0.001
Where to get birth control	72.3	48.7	23.7**	<0.001
STIs	80.6	71.6	9.0	0.080
Talking about sex with your partner	73.4	48.7	24.8**	<0.001
Saying no to sex	83.5	67.6	15.9**	<0.001
How babies are made	83.9	73.5	10.4**	0.005
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.7	3.3	0.4**	0.004
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.0	1.000
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.4	2.5	–0.1	0.792
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.8	2.9	–0.1*	0.015
Would be very upset if got pregnant or got someone pregnant	53.4	46.3	7.1	0.216
Intentions				
Intend to have sex in the next year	16.7	13.6	3.2	0.971
Intend to use a condom if have sex in the next year	79.4	77.5	1.9	1.000
Intend to use any method of birth control if have sex in the next year	86.7	84.1	2.6	0.552
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.2	3.3	0.0	1.000
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.0	3.0	0.0	1.000

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.4	0.0	1.000
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.2	4.7	–0.5	0.120
Received very useful information from parents on relationships or sexual health in past 12 months	56.3	58.2	–2.0	1.000
Spoke to health professional about sexual health in the past 12 months	61.4	70.2	–8.8*	0.024
Insufficient communication about sex with partner (0 if no partner)	18.8	27.5	–8.8	0.360
Substance use				
Drank alcohol in past 30 days	30.5	25.3	5.2	0.267
Smoked marijuana in past 30 days	22.7	19.2	3.5	0.425
Sample size	**582**	**192**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 429 intervention-group and 142 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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 estimated the impacts of Teen PEP separately within subsamples of female (Table E.2) and male (Table E.3) students. There is little evidence that Teen PEP led to changes in sexual behavior within either subsample. But estimates in the intermediate outcome domains suggest that students of different genders might have taken away different messages from their exposure to Teen PEP. In particular, Teen PEP is associated with only a small and insignificant impact on the knowledge of female students (difference of 0.2 questions answered correctly, $p = 0.429$) but a larger and statistically significant impact on the knowledge of male students (difference of 0.7 questions answered correctly, $p < 0.001$). Conversely, Teen PEP was associated with marginally significant increases in intentions to use a condom among female students (difference of 6 percentage points, $p = 0.077$) but no corresponding change for male students (difference of 3 percentage points, $p = 1.000$). Interestingly, although Teen PEP was not associated with significant changes in attitudes within the sample as a whole, both gender subsamples provide some evidence that the program was associated with increases in positive perceptions toward birth control. The impacts are 0.1 points for girls and 0.2 points for boys, both on a 5-point scale ($p = 0.064$ and $p = 0.027$, respectively). Finally, the male subsample demonstrates one counterintuitive result not seen in the female subsample or for the sample as a whole. At odds with Teen PEP's focus on improving parent-student communication, male students in Teen PEP schools spoke with their parents about sexual health and related topics less often than male students at comparison schools (difference of 1.1 on an 11-point scale, $p = 0.012$).

Table E.2. Impacts of Teen PEP on outcomes, female students only

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	31.7	30.3	1.4	0.898
Ever had sexual intercourse[b]	42.6	38.8	3.8	0.661
Number of sexual partners	0.6	0.5	0.1	0.634
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	22.8	23.0	–0.2	0.963
Had sexual intercourse without any method of birth control in the three months before survey	17.8	15.4	2.4	0.675
Exposure to information				
Received any information in past 12 months on				
Relationships	81.7	80.7	1.0	1.000
Abstinence	81.2	60.0	21.2**	<0.001
Methods of birth control	75.1	53.1	22.1**	<0.001
Where to get birth control	77.7	58.9	18.7**	<0.001
STIs	88.5	76.0	12.5**	0.013
Talking about sex with your partner	76.8	51.5	25.3**	<0.001
Saying no to sex	90.6	87.8	2.8	1.000
How babies are made	91.5	87.8	3.7	0.237
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.4	3.2	0.2	0.429
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.064
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.4	2.5	–0.1	0.109
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.9	3.0	–0.1	0.268
Would be very upset if got pregnant or got someone pregnant	56.0	56.1	–0.2	1.000
Intentions				
Intend to have sex in the next year	10.0	8.7	1.2	1.000
Intend to use a condom if have sex in the next year	84.5	78.5	6.0	0.077
Intend to use any method of birth control if have sex in the next year	90.2	86.3	3.9	0.118
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.5	3.5	0.0	1.000
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.2	3.1	0.0	1.000

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.4	0.0	1.000
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.6	4.7	–0.2	1.000
Received very useful information from parents on relationships or sexual health in past 12 months	61.8	63.3	–1.6	1.000
Spoke to health professional about sexual health in the past 12 months	72.9	81.2	–8.3	0.270
Insufficient communication about sex with partner (0 if no partner)	17.0	16.9	0.1	1.000
Substance use				
Drank alcohol in past 30 days	28.7	23.6	5.1	0.100
Smoked marijuana in past 30 days	26.9	26.6	0.3	1.000
Sample size	**401**	**130**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 308 intervention-group and 100 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

*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 E.3. Impacts of Teen PEP on outcomes, male students only

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	31.8	31.3	0.5	0.953
Ever had sexual intercourse[b]	43.0	35.7	7.2	0.214
Number of sexual partners	1.3	1.1	0.2	0.977
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	22.5	21.4	1.1	0.923
Had sexual intercourse without any method of birth control in the three months before survey	19.8	15.3	4.5	0.286
Exposure to information				
Received any information in past 12 months on				
Relationships	75.4	69.3	6.1	1.000
Abstinence	67.8	45.8	22.0**	0.006
Methods of birth control	62.8	37.5	25.3**	0.001
Where to get birth control	63.6	20.5	43.1**	<0.001
STIs	75.2	65.2	10.0	0.592
Talking about sex with your partner	69.8	47.1	22.7*	0.011
Saying no to sex	76.6	47.8	28.8**	<0.001
How babies are made	79.4	62.0	17.4**	0.009
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.7	3.0	0.7**	<0.001
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.3	4.2	0.2*	0.027
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.6	2.6	0.0	1.000
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.6	2.7	–0.1	0.766
Would be very upset if got pregnant or got someone pregnant	43.7	35.9	7.8	0.586
Intentions				
Intend to have sex in the next year	24.5	18.6	5.9	0.269
Intend to use a condom if have sex in the next year	79.9	77.2	2.7	1.000
Intend to use any method of birth control if have sex in the next year	86.4	83.1	3.4	1.000
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	2.8	2.8	0.1	0.785
Thoughtfulness in decision making (average of two survey items; range 1–4)	2.9	2.7	0.2	0.306

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.3	3.3	0.0	1.000
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	3.4	4.5	–1.1*	0.012
Received very useful information from parents on relationships or sexual health in past 12 months	47.7	44.9	2.8	1.000
Spoke to health professional about sexual health in the past 12 months	54.2	56.5	–2.3	1.000
Insufficient communication about sex with partner (0 if no partner)	22.8	28.2	–5.4	1.000
Substance use				
Drank alcohol in past 30 days	34.9	29.2	5.8	0.685
Smoked marijuana in past 30 days	18.9	16.1	2.7	0.925
Sample size	**313**	**98**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific non-response limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

[b]Impacts are estimated using only the 225 intervention-group and 71 comparison-group students who had not had sex at baseline (the comparison and intervention group means reflect the share of students who ever had sex by follow-up, unconditional on sexual initiation at baseline).

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

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

Finally, we examined impacts of Teen PEP within the subsample of students who did not report having ever had sex at baseline (Table E.4). Impacts on knowledge and exposure to information are similar to those seen for the sample as a whole. However, within this subsample of previously abstinent students, Teen PEP was associated with a marginally significant increase in sexual activity within the past three months. Eighteen percent of students at Teen PEP schools who did not report having ever had sex at baseline reported having sex in the three months before the interim follow-up survey, compared to 13 percent of corresponding students in comparison schools. The 6 percentage point difference is marginally statistically significant, though just barely so ($p = 0.099$).

Table E.4. Impacts of Teen PEP on outcomes, students who did not report having ever had sexual intercourse at baseline only

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Sexual activity				
Had sexual intercourse in the three months before survey[a]	18.5	12.4	6.0	0.099
Ever had sexual intercourse	19.5	12.6	6.9	0.152
Number of sexual partners	0.4	0.2	0.1	0.381
Unprotected sex				
Had sexual intercourse without a condom in the three months before survey	12.8	8.9	3.8	0.144
Had sexual intercourse without any method of birth control in the three months before survey	11.6	8.1	3.5	0.251
Exposure to information				
Received any information in past 12 months on				
Relationships	78.4	75.7	2.7	1.000
Abstinence	76.1	57.1	19.0**	<0.001
Methods of birth control	69.7	43.1	26.5**	<0.001
Where to get birth control	70.6	40.9	29.7**	<0.001
STIs	81.9	69.8	12.0**	0.002
Talking about sex with your partner	71.3	43.3	28.0**	0.001
Saying no to sex	85.6	71.3	14.4**	0.001
How babies are made	85.6	77.9	7.6*	0.024
Knowledge				
Knowledge of preventing STI transmission and pregnancy (number of correct answers to six survey items; range 0–6)	3.5	3.2	0.3**	0.004
Attitudes				
Perceptions of factors supporting use of birth control (average of four survey items; range 1–5)	4.4	4.3	0.1	0.286
Perceptions of barriers to use of birth control (average of five survey items; range 1–5)	2.4	2.6	–0.1	0.130
Index of negative views toward having sex at current age (average of four survey items; range 1–5)	2.9	3.0	0.0	0.723
Would be very upset if got pregnant or got someone pregnant	56.2	53.9	2.3	1.000
Intentions				
Intend to have sex in the next year	9.1	6.0	3.1	NA
Intend to use a condom if have sex in the next year	84.0	84.1	–0.1	NA
Intend to use any method of birth control if have sex in the next year	89.3	88.6	0.7	NA
Decision making and self-efficacy				
Perceptions of refusal skills (average of four [boys] or five [girls] survey items; range 1–4)	3.4	3.3	0.1	0.466

Variable	Intervention group mean	Comparison group mean	Difference	p-value
Thoughtfulness in decision making (average of two survey items; range 1–4)	3.1	3.0	0.1	0.779
Believe could go to clinic to seek STI testing if needed (single survey item; range 1–4)	3.4	3.3	0.1	0.103
Communication				
Frequency of discussions with parents about relationships or sexual health in past 12 months (single survey item; range 0–10)	4.0	4.3	–0.3	0.494
Received very useful information from parents on relationships or sexual health in past 12 months	53.2	54.0	–0.8	1.000
Spoke to health professional about sexual health in the past 12 months	61.7	65.8	–4.1	0.650
Insufficient communication about sex with partner (0 if no partner)	18.6	19.9	–1.3	1.000
Substance use				
Drank alcohol in past 30 days	26.8	19.3	7.5	0.100
Smoked marijuana in past 30 days	18.3	16.1	2.2	0.891
Sample size	**533**	**171**		

Source: Teen PEP first follow-up survey, propensity-score matched sample.

Note: Impact estimates and group means are regression-adjusted to account for baseline characteristics (see Appendix B for a full list of covariates). Estimates are weighted to give students in each intervention school equal weight. Estimates are in percentages unless otherwise specified. Item-specific nonresponse limits sample size for some outcomes. *P*-values are corrected for clustering at the school level and for multiple comparisons within domain.

[a]Designated before analysis as a confirmatory outcome.

NA. Not available. *P*-value could not be calculated due to instability in the variance-covariance matrix of estimates.

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

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

PPA

Evaluation of Adolescent
Pregnancy Prevention Approaches

http://www.hhs.gov/ash/oah/

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

OFFICE OF ADOLESCENT HEALTH