

**Supporting Justification for OMB Clearance of Teen Pregnancy Prevention
Replication Evaluation (OMB Control #0990-NEW)**

Part A: Justification for the Collection of Baseline Data

November 2011

(updated June 2012)

The Office of the Assistant Secretary for Planning and Evaluation (ASPE) in collaboration with the Office of Adolescent Health (OAH), Office of the Assistant Secretary for Health (OASH) in the U.S. Department of Health and Human Services (HHS) is overseeing the TPP Replication Study evaluation. The TPP Replication study is specifically designed to address the question “Do evidence-based program models, replicated and funded as part of the OAH Teen Pregnancy Prevention Program, demonstrate impacts on sexual risk behaviors that are comparable to the originally-reported impacts and are they effective in preventing teen pregnancy and reducing sexually transmitted infections?” This evaluation focuses on the replication of a small number of program models across multiple sites with the goals of determining the extent to which program impacts are replicated as well as addressing questions about the extent to which aspects of program implementation are associated with program impacts. In Fall 2011, ASPE awarded a contract to Abt Associates Inc. to conduct the evaluation.

For the purpose of this clearance, OAH is seeking OMB approval for baseline survey data collection for the *Teen Pregnancy Prevention (TPP) Replication Study*. The 60-day notice was published March 16, 2011 and requested approval for two teen pregnancy prevention studies: the TPP Replication Study and the Evaluation of Adolescent Pregnancy Prevention Approaches (PPA). However, the current request for clearance is for the TPP Replication Study only; the PPA baseline survey instrument has been submitted to OMB through ACF and was approved on August 17, 2011 (0970-0360). Therefore, the requested burden for this package has been substantially reduced since the 60-day notice was published.

OAH is overseeing and coordinating adolescent pregnancy prevention evaluation efforts as part of the Teen Pregnancy Prevention Initiative. In order to ensure that these Federal evaluation efforts across the Department are aligned, OAH is coordinating the submission of OMB Packages related to them. In support of these coordinated evaluation efforts, OAH has collaborated with other agencies that implement and evaluate teen pregnancy prevention and related issues in order to address a range of research and policy questions that complement rather than duplicate one another. These agencies include the Administration for Children and Families (ACF), the Office of the Assistant Secretary for Planning and Evaluation (ASPE), and the Centers for Disease Control and Prevention (CDC). HHS has created a Federal Teen Pregnancy Prevention Coordination Workgroup to develop and manage a coordinated strategy of HHS teen pregnancy prevention activities and evaluation efforts. The workgroup involves research and program staff from ACF, ASPE, CDC, and OAH. The workgroup has enabled the Department to collaborate on the new evaluation efforts and maximize the questions we can answer across the initiative, including the development of common core measures to be used across evaluation studies.

The baseline survey instrument for which clearance is requested here is almost identical to the baseline data collection survey instrument approved for the PPA evaluation study (0970-0360). The current Information Collection Request (ICR) is new because the baseline survey proposed will be used for the TPP Replication Study and has been tailored slightly for the different program models being evaluated.

HHS created a “core baseline instrument” to use across federal teen pregnancy prevention evaluation studies. The core baseline instrument identifies those items that the Federal Teen Pregnancy Prevention Coordination Workgroup agreed should be included on each baseline survey instrument administered as part of any federal teen pregnancy prevention evaluation study. The TPP Replication study baseline instrument for which clearance is requested here in large part duplicates the PPA baseline measure (including all core and many non-core items) which, as noted above, OMB has previously reviewed and approved. The instrument includes OAH grantee performance measures (also OMB-approved), and adds two items that have not previously been reviewed by OMB. Further description of the TPP Replication study baseline instrument for which approval is requested may be found at the end of A1. The TPP Replication Study baseline instruments can be found in Attachments D through G (items not previously reviewed and approved by OMB are highlighted in yellow).

A1. Circumstances Making the Collection of Information Necessary

For decades, policymakers and the general public have remained concerned about the prevalence of sexual activity among adolescents. Although adolescents today are waiting somewhat longer before having sex than they did in the 1990s, 60 percent of teenage girls and more than 50 percent of teenage boys report having had sexual intercourse by their 18th birthday.¹ Approximately one in five adolescents has had sexual intercourse before turning 15.² Rates of teenage pregnancy declined by 34 percent between 1991 and 2005 for teens aged 15-19, before rising 5 percent between 2005 and 2007.³ The rate of teen births again dropped between 2007 and 2009, falling 8 percent for teens aged 15-19.^{4 5} Preliminary data in

¹ Abma, J. C., G. M. Martinez, W. D. Mosher, and B. S. Dawson. “Teenagers in the United States: sexual activity, contraceptive use, and childbearing”, *Vital and Health Statistics*, vol. 23, no. 24, 2004, pp. 1-48.

² Albert, B., S. Brown, and C. Flannigan, eds. 14 and Younger: The Sexual Behavior of Young Adolescents. Washington, DC: National Campaign to Prevent Teen Pregnancy, 2003.

³ Hamilton, B.E., Martin, J.A., Ventura, S.J. (December, 2010). Births: Preliminary data for 2009. National vital statistics reports web release. Vol. 59 no 3. Hyattsville, MD: National Center for Health Statistics.

2009 indicate an overall teen birth rate for teens aged 15-19 of 39.1 per 1000⁶.

HHS is interested in identifying and evaluating approaches to reduce teen pregnancy, associated risk behaviors, and their consequences. One of the key policy questions is whether program that have demonstrated evidence of effectiveness can be replicated in new settings with positive impacts. Of the 31 programs on the HHS list of evidence-based programs, only one program model has been replicated and shown to have positive effects through a rigorous evaluation. The baseline data collection described in this ICR, combined with subsequent follow-up data collections, will provide important information to guide policy decisions aimed at replicating evidence-based programs.

Legal or Administrative Requirements that Necessitate the Collection

On December 19, 2009, the President signed the Consolidated Appropriations Act of 2010 (Public Law 111-117). Division D, Title II of the Act created the Teen Pregnancy Prevention Program, which is consistent with the Administration's interest in establishing an evidence-based program to prevent teen pregnancy. The Act provides \$110 million to fund this program within OAH, which is responsible for both program implementation and administration. The Teen Pregnancy Prevention Program is a two-tiered program that includes: (1) \$75 million for replicating evidence-based programs that have been proven effective through rigorous evaluation (Tier 1); and \$25 million for research and demonstration grants to develop and test additional models and innovative strategies (Tier 2).

In addition, Public Law 111-117, which set fiscal year (FY) 2010 appropriations levels, included the following language: "\$4,455,000 shall be available from amounts available under section 241 of the Public Health Service Act to carry out evaluations (including longitudinal evaluations) of adolescent pregnancy prevention approaches." The same language appropriated \$4,455,000 in FY 2011. These funds have been used to fund both the PPA evaluation and the TPP Replication Study. In addition to these funds, the FY 2012 Appropriations Act provided \$8.455 million in PHS evaluation funds, an increase of \$4 million over the FY 2011 level, which is also supporting longitudinal evaluations of teen pregnancy prevention approaches.

⁴ Hamilton, B.E., Martin, J.A., Ventura, S.J. (December, 2010). Births: Preliminary data for 2009. National vital statistics reports web release. Vol. 59 no 3. Hyattsville, MD: National Center for Health Statistics.

⁵

⁶ Hamilton BE, Martin JA, Ventura SJ. Births: Preliminary data for 2009. National vital statistics reports, Web release; vol 59 no 3. Hyattsville, MD: National Center for Health Statistics. 2010.

As previously mentioned, the TPP Replication Study is focused on evaluating replications of evidence-based program models funded through the OAH TPP Program Tier 1 replication grants. The PPA study is focused on evaluating untested and innovative program models funded through the OAH TPP Program Tier 2 research and demonstration grants as well as other funding streams.

To accomplish the objective of the appropriation, OAH seeks OMB approval of the baseline survey instrument.

Objectives of the TPP Replication Study

The goal of the TPP Replication Study is to determine the extent to which evidence-based program models that have been shown to be effective in an earlier trial, usually conducted by the program developer, demonstrate effects on adolescent sexual risk behavior and teenage pregnancy when they are replicated in similar and in different settings, for different populations, and over a longer time-period. The evaluation will help OAH provide guidance to program managers and state and local policymakers about evidence-based program models and about the factors necessary to support successful replication.

For this evaluation, HHS has identified three evidence-based program models that represent different approaches to the prevention of teenage pregnancy, and that are being widely replicated as part of the TPP Program and through other federal and state funding initiatives. The three program models are: Safer Sex, a clinic-based individualized intervention for sexually-active female youth; ¡Cuidate!, a culturally-sensitive small-group intervention aimed at Latino youth; and Reducing the Risk, a classroom-based sexual health curriculum that can also be implemented as an after-school program and in non-school settings. For each model, the agencies have identified three grantee replications, for a total of 9 replications. The nine vary in the scope of the replication (number of sites within a replication, number of youth served) and in the populations served. A good deal of variation can be expected in the settings in which the program is implemented. While one program model is implemented only in clinics, the others can be implemented in a variety of settings, including schools, churches, and other community-based settings that provide services to youth. The study will use a sample of approximately 8,550 across 9 grantee replication sites, a sufficient size to detect policy-relevant impacts of each of the program replications. In each of the replications selected, youth will be assigned to receive the intervention or to be part of a control group that does not receive it. In clinics and other community-based settings and in some school settings, individual youth will be randomly assigned. In the three sites where Reducing the Risk the program is being implemented as a classroom-based curriculum, the unit of random assignment will be classes within a school (for example, health or PE classes). In all cases, the

intervention will be delivered by grantee staff who are health educators, not by the regular class teacher, so that the issue of contamination when teachers deliver both the intervention to the treatment group as well as the regular class to the control group does not arise.

Baseline surveys will be conducted with youth in both treatment and control groups before youth in the treatment group have been exposed to the intervention. In schools, the self-administered survey will be completed in a space that can accommodate small groups and assure privacy; in other settings, notably clinics where entry to the program is on a rolling basis, the survey will be completed in a room where the individual's privacy is protected.

Through the baseline and follow-up surveys HHS will address the following research questions:

- What are the impacts on adolescent sexual risk behavior and teen pregnancy rates when an evidence-based program is replicated?
- Do impacts vary for different youth populations (i.e., females vs. males, different age ranges, ethnicities)?
- Do impacts vary depending on the setting in which the program is replicated?
- Are impacts replicated across sites implementing a specific program model?

Major activities for the TPP Replication Study include the following:

- Selecting replication sites from the Teen Pregnancy Prevention Initiative grantees funded to replicate evidence based programs (Tier 1). All of the Tier 1 grantees are replicating "evidence-based" program models and are required to take steps to ensure fidelity to the model.
- Recruiting grantees to participate in a rigorous experimental evaluation and working with them to design and support a strong study.
- Collecting data on the research sample at baseline and at two subsequent time points (i.e. short-term and longer-term follow-up survey administration).
- Conducting a comprehensive implementation study in each replication site.
- Analyzing data and reporting the results.

The Baseline Survey

The proposed survey will be conducted with all study participants and contains a) a selected group of questions from the already-approved PPA baseline survey (including OAH performance measures) and (b) a small number of additional questions that address issues specifically appropriate to the program models being evaluated.

A2. Purpose and Use of the Information Collection

If this request is approved, the evaluation will collect baseline data on sample members' demographic characteristics, their sexual and other risk behavior, prior receipt of information related to reproductive health, and information on how they can be contacted later. These data will be obtained from a baseline survey administered to sample youth.

The data will serve several purposes. Identifying and contact information will help the study teams track sample youth throughout the evaluation, and locate them for follow-up if they have graduated, moved to another school, dropped out, or were absent for follow-up data collection. Baseline variables are also important in several ways for the analysis. They will be used to establish the baseline equivalence of the treatment and control groups on measurable variables and thus to confirm the integrity of the random assignment process. Baseline variables may be used to define subgroups for which impacts will be estimated, and to adjust impact estimates for the baseline characteristics of non-respondents to the follow-up survey. Many baseline variables will be measures of outcomes to be measured again at follow-up; their baseline values can be used to improve the precision of impact estimates through their inclusion as covariates in the impact models.

Baseline data will measure: teens' demographic and socioeconomic characteristics; dating experience; knowledge, attitudes, and expectations, including about sexual activity and contraception; stressors and supports; and school and community characteristics (as well as collect contact information). There are two versions of the baseline survey: one for the Safer Sex program model and one for the Cuidate! and Reducing the Risk program models. The two versions are nearly identical, although there are slight differences in the items in order to tailor the instrument to the program model and the target intervention audience. For example, the Safer Sex program model serves only sexually active females and therefore it is not necessary to include items targeted to males or to youth who are not sexually active. Furthermore, the Reducing the Risk and Cuidate! baseline has two versions: one for sexually active youth and one for youth who have never had sex. Sexually active youth will respond to questions about their sexual behavior whereas youth who have never had sex will answer questions unrelated to sex so that the survey length is equivalent for the different groups. This is important for settings in which the surveys are administered to a group. Attachment A is a table that provides:

- A question-by-question list with sources for each item;

- A crosswalk between the two versions of the TPP baseline survey and the PPA OMB-approved baseline survey indicating which items appear on which survey(s)
- A description of how the data will be used.

Attachment B lists the topics covered in the baseline instrument and our justification for their inclusion. A list of national surveys reviewed in developing the baseline survey instrument is provided in Attachment C together with detailed references for sensitive questions. The baseline survey instrument is broken into the following four pieces:

- Attachment D: Part A of the baseline survey to be used for Safer Sex sites;
- Attachment E: Part A of the baseline survey to be used for Cuidate! and Reducing the Risk sites;
- Attachment F: Part B1 of the baseline survey to be used with all Safer Sex sites and with youth who have ever had sex in the Cuidate! and Reducing the Risk sites; and
- Attachment G: Part B2 of the baseline survey to be used with youth who have never had sex in the Cuidate! and Reducing the Risk sites (all youth served by the Safer Sex program have had sex).

A3. Use of Improved Information Technology and Burden Reduction

The data collection plan reflects sensitivity to issues of efficiency, accuracy, and respondent burden. Where feasible, information will be gathered from existing data sources; the information being requested through surveys is limited to that for which the youth are the best or only information sources. For all surveys, both baseline and follow-up, state-of-the art technology will be used to reduce burden, improve comprehension and accuracy of responses, and ensure data security. All survey data will be collected via web-based Audio Computer-Assisted Self-Interview (ACASI), which has the capacity to capture and store data in real time, where each response to a question (as it is entered) is sent immediately to a central and secure database and no information is stored on local computers. This web-based ACASI technology has been successfully used in several large clinical trials, including studies that deal with drug use or exposure to HIV/AIDS. Research has demonstrated that surveys administered online are characterized by higher levels of self-disclosure, an increased willingness to answer sensitive questions and a reduction in socially desirable responses.

All sample members will be encouraged to complete the web-based survey, which will contain an audio option embedded in it. The strategy is

ideal for young survey respondents and reinforces the idea that no-one else will see or hear the survey questions. Once approved, the survey instrument will be translated into Spanish, so that respondents can choose the language in which they take it. In addition, English and Spanish versions of the survey will also be available in hard copy, for emergency use only if unanticipated technical glitches occur at the time of survey administration. Attachment H provides additional information on the use and administration of web-based ACASI surveys and research references.

A4. Efforts to Identify Duplication and Use of Similar Information

The information collection requirements for the evaluation have been carefully reviewed to determine what information is already available from existing studies and what will need to be collected for the first time. Although the information from existing studies adds to our understanding of teenage sexual risk behavior, HHS does not believe that the extant research literature provides robust evidence about the effectiveness of evidence-based programs (i.e., evidence from independent evaluation of the program or from more than one study) to meet the needs of policymakers and stakeholders aiming to reduce this behavior. The data collection for the evaluation is an essential step in providing this information.

HHS has created a Federal Teen Pregnancy Prevention Coordination Workgroup to develop and manage a coordinated strategy of HHS teen pregnancy prevention activities and evaluation efforts. The workgroup involves research and program staff from ACF, ASPE, CDC, and OAH. The workgroup has enabled the Department to collaborate on the new evaluation efforts and maximize the questions we can answer across the initiative, including the development of common core measures to be used across evaluation studies. We have collaborated to design research and evaluation efforts that will enable the Department to answer a range of research and policy questions that are complementary to, rather than duplicative of, one another. Specifically, we are interested in (1) adding to the evidence base by evaluating new and untested program models and innovative strategies; and (2) understanding how to effectively replicate and implement evidence-based program models and how to achieve impacts that were found in the original evaluations. The TPP Replication study addresses the latter research question. The federal evaluation strategy includes a combination of federal-led and grantee-led evaluation efforts described briefly below.

Federal-Led Evaluations: There are four federally managed evaluation studies that address unique questions about the implementation and effectiveness of a subset of HHS grantees.

- *Evaluation of Pregnancy Prevention Approaches (PPA):* An experimental evaluation study focused on assessing the implementation and impacts of innovative strategies and untested approaches for preventing teenage

pregnancy in seven sites. Three of the sites are from the TPP research and demonstration grantees, three sites are PREP Innovative Strategies grantees, and one is a non-federally funded site. Implementation reports are expected between March 2012 and October 2013 and internal short-term impact memos are expected between January 2014 and July 2015 across the sites. The contractor is Mathematica.

- *Teen Pregnancy Prevention (TPP) Replication Study Evaluation:* An experimental evaluation study that will examine the implementation and impacts of three TPP replications of three different evidence-based program models, for a total of 9 sites. The study will examine whether program models that were commonly chosen by replication grantees and widely used in the field can achieve impacts with different populations and settings. Implementation and short-term impact findings are anticipated in 2015. The contractor is Abt Associates.
- *CDC Community-Wide Evaluation:* A quasi-experimental evaluation study to examine the effects of integrating services, programs, and strategies. Initial impact findings are expected in 2016. The contractor is ICF Macro.
- *State PREP Multi-Component Evaluation:* This study will document program design and implementation within states and includes an experimental evaluation to assess the effectiveness of 4 or 5 selected programs. Preliminary descriptive findings are expected in 2013 and impact findings are expected in 2016. The contractor is Mathematica.

In addition, there are 41 grantee-led rigorous evaluations of both TPP and PREP Innovative Strategies replication and research and demonstration grants, supported by a federally sponsored evaluation technical assistance contractor (Mathematica). The contractor has reviewed each of the local evaluation designs to ensure they are rigorous and feasible and continues to provide ongoing technical assistance to grantees.

A5. Impact on Small Businesses or Other Small Entities

Programs in many sites may be operated by community-based organizations. The data collection plan is designed to minimize burden on such sites by providing staff from the evaluation contractor team to assist in group data collection. For respondents who do not complete the survey in a group setting, Abt Associates, (through its subcontractor for data collection, DIR) will provide passwords for web completion or will conduct telephone data collection, thus minimizing the need for extensive “sample pursuit” by program staff.

A6. Consequences of Collecting Information Less Frequently

Baseline data are essential to the conduct of a rigorous evaluation of pregnancy prevention programs, as the appropriations' language requires. The absence of baseline data would limit the ability to draw conclusions from the evaluation study. Furthermore, funding decisions about teen pregnancy prevention programs will continue to be based on insufficient and outdated information on program effectiveness.

A7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5

There are no special circumstances for the proposed data collection.

A8. Comments in Response to the Federal Register Notice and Efforts to Consult Outside the Agency

The 60-day notice was published in the Federal Register on March 16, 2011. The text is found in Attachment I. At this time there are no comments or responses to questions.

As explained in earlier sections, the TPP Replication Study baseline survey is nearly identical to the PPA baseline survey that was approved by OMB in August, 2011. The development of the items in the survey was primarily done by Mathematica under the PPA contract, with the intention on the part of HHS that the surveys developed under that contract would form the basis for all subsequent federal evaluations of the TPP Initiative. In Section B5 we provide the names and contact information of persons consulted in the drafting and refinement of the baseline survey instrument, and a list of members of the Technical Work Group for the PPA evaluation who provided comments on a near-final draft instrument.

A9. Explanation of Any Payment or Gift to Respondents

The population targeted for the evaluation presents a challenge for the study, that is increased by the desire to measure long-term impacts of the program beyond the measures taken immediately at the end of the program that are typical in this research field. By design, the programs in this evaluation target youth who are at the highest risk for sexual risk behavior: inner-city youth in cities like St. Louis; low-income Latino youth, many of whose families are recent immigrants to the US; and young females ages 14-19 who are already sexually active and engaging in unprotected sex. These populations are more likely to drop out of school than their more advantaged counterparts and they are often extremely mobile, and hard to reach. To ensure that we achieve the required 80% response rate at the end of two years, it is important to take steps to attach them firmly to the study **at the outset** and to maintain that attachment over time. These steps are also essential to prevent differential attrition, leading to response bias, since

members of the control group are receiving no program services and are not in contact with program staff.

To this end, we propose to provide modest incentive payments to each participant at each survey point. These incentives will be uniform across program models, replication sites, as well as across the three survey time-points (baseline, short-term follow-up and long-term follow-up). All study participants who complete the baseline survey will receive a \$25 gift card; in a subsequent clearance request for the follow-up surveys, we will propose that the same incentive be offered for completion of those surveys. The payment is intended to encourage completion of the survey and, even more importantly, to reinforce the importance of subsequent surveys. In addition to regular efforts to track youth between survey points, the gift cards are intended to increase attachment to the study so as to keep attrition to a minimum and ensure that any attrition is not differential in favor of the control group.

We should point out that, although three of the sites are implementing Reducing the Risk as a classroom-based intervention, it may not be possible in these sites to administer the baseline survey in the classroom, given the size of the classes, our desire to ensure privacy for the respondents, and the need for laptops or tablets on which they will complete the survey. Our plan is to schedule small groups of study participants, in some cases during a class period (i.e., student is pulled out of the regular class), but more often at times when the student has free time or study hall. This strategy is highly dependent on cooperation from students in keeping scheduled appointments. We hope that the incentive payment will increase the level of cooperation at the initial time-point as well as increase the likelihood of retention at subsequent time-points.

To develop this strategy we reviewed the research literature on the problem of attrition in both panel and longitudinal surveys and the effectiveness of incentives to address the issue (Exhibit A9.1 below). We know of no experimental studies that compare the effects of different forms of incentives. Therefore, in selecting gift cards, we were guided by our IRB and the OAH grantees, all youth-serving organizations, who were unanimous in believing that gift cards would be the most effective form of incentive for their population. We are working with each grantee to identify the most appropriate gift card for youth in their area (Visa or Target for example).

What the research studies in Exhibit A9.1 demonstrate is that larger incentives (\$40) have greater effects than smaller (\$30, \$20), but that incentives generally have an impact on completion and retention rates. Some but not all of these studies focused on adolescents as opposed to adult respondents. Aside from this group of studies, most studies have chosen a single incentive level, so that we cannot with certainty attribute the completion and retention rates achieved to the incentive. However, ACF's

evaluation of Building Strong Families, conducted by Mathematica, found \$25 incentives for low-income youth (and a \$25 incentive for parents) effective in attaching the dyads to the study over time and achieving the required completion rates. In Abt’s multiple studies for the Corporation for National Service, a \$25 incentive has been the standard incentive used over the last two decades to achieve the desired completion rates with youth populations at all socio-economic levels, although the incentive has sometimes been raised to reach the hardest-to-reach youth. All these incentive payments were approved by OMB.

In settling on \$25 incentives, we attempted to balance the demonstrated effectiveness of greater incentives with the reasonableness of the total cost to the study. All of the youth populations targeted by the program interventions are high-risk and often highly mobile. We have, therefore, as noted in our revised submission, chosen to make the incentives uniform across replication sites and across time, since the challenges of retention are likely to be similar in all sites.

Exhibit A9.1: REFERENCES ON THE EFFECT OF INCENTIVES IN LONGITUDINAL/MULTI-MODE SURVEYS
<p><i>Impact of incentives on initial and subsequent response rates of adult survey takers</i> Goldenberg, Karen L., David McGrath, and Lucilla Tan. 2009. “The Effects of Incentives on the Consumer Expenditure Interview Survey.” Proceedings of the Survey Research Methods Section, American Statistical Association (ASA). Accessed via http://www.amstat.org/sections/srms/proceedings/allyears.html</p> <p>An incentives experiment was conducted in the Consumer Expenditure (CE) Quarterly Interview Survey to determine whether offering prepaid incentives of \$20 or \$40 prior to the first interview would improve response rates in the current wave and subsequent 4 waves. Offering \$40 significantly increased response rates 4.5% compared with offering no incentive and the effect, while smaller, persisted across all five interviews. The \$20 incentive increased response rates 2.2% in the first wave compared with no incentive, although this difference was not statistically significant.</p>
<p><i>Impact of incentives on attrition from a multi-modal panel study of teenagers</i> Jäckle, Annette and Peter Lynn. 2007. Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias. Institute for Social & Economic Research (ISER) Working Paper. Accessed at https://www.iser.essex.ac.uk/publications/working-papers/iser/2007-01.pdf</p> <p>This working paper considered the cumulative effects of conditional and unconditional incentives in a multi-mode (mail and telephone) panel study of teenagers in the UK. Unconditional incentives significantly reduced attrition in a multi-mode panel study, with no impact on attrition bias, regardless of mode or type of incentive. The results suggest that incentives are also effective in maintaining sample sizes in a panel study.</p>
<p><i>Impact of incentives on response rates, sample composition and attrition bias.</i> Laurie, Heather, and Peter Lynn. 2009 . “The Use of Respondent Incentives on Longitudinal Surveys.” Chapter 12 in Peter Lynn (ed.) <i>Methodology of Longitudinal Surveys</i>. Hoboken, NJ: John Wiley and Sons.</p> <p>Chapter 12 provides a comprehensive review of the literature on incentives in longitudinal surveys, including the effect of incentives on response rates, sample composition and bias, and data quality.</p>
<p><i>Impact of incentives on response rates and attrition rates for adult survey-takers</i> Mack, Stephen, Vicki Huggins, Donald Keathley, and Mahdi Sundukchi. 1998. “Do Monetary</p>

Exhibit A9.1: REFERENCES ON THE EFFECT OF INCENTIVES IN LONGITUDINAL/MULTI-MODE SURVEYS

Incentives Improve Response Rates in the Survey of Income and Program Participation?" *JSM Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association, 529-34.

This paper describes incentive experiments undertaken by the U.S. Census Bureau in the Survey of Income and Program Participation (SIPP), a high-burden, face-to-face panel-design interview survey, to deal with rising nonresponse to government surveys in the 1990s. The SIPP research demonstrated that incentive effects for large, interview-administered government surveys were similar to those for non-government surveys, and that these effects continued to hold through the 6th interview wave two years after an incentive was provided.

Impact of incentives on attrition rates in an adult panel study

Martin, Elizabeth, Denise Abreu, and Franklin Winters. 2001. "Money and Motive: Effects of Incentives on Panel Attrition in the Survey of Income and Program Participation." *Journal of Official Statistics* 17 (2): 267-284.

This paper describes an experiment that compared the effects of offering a prepaid incentive of \$20, \$40, or no incentive on panel attrition in a household survey. Both \$20 and \$40 significantly improved conversion rates of prior non-interviews compared to offering no incentive, particularly for households with higher poverty rates.

Impact of initial incentives on initial and subsequent response rates in a longitudinal study

Rodgers, Willard. 2011. "Effects of Increasing the Incentive Size in a Longitudinal Study." *Journal of Official Statistics* 27 (2): 279-299.

In this study, participants in one wave of a longitudinal study were offered \$20, \$30, or \$50. Offering the highest incentive of \$50 showed the greatest improvement in response rates and also had a positive impact on response rates for the next four waves.

Impact of promised incentives on refusal conversion in a panel study

Zagorksy, Jay L. and Patricia Rhoton. 2008. "The Effect of Promised Monetary Incentives on Attrition in a Long-Term Panel Survey." *Public Opinion Quarterly* 72 (3): 502-513.

In a face-to-face longitudinal study of women, promised incentives of up to \$40 had a positive effect on response rates in panel members who had previously participated in the survey but had previously refused to participate in the current wave.

A10. Assurance of Confidentiality Provided to Respondents

HHS has embedded protections for privacy in the study design. Data collection will occur only if informed consent is provided by a parent or legal guardian if the respondent is a minor or by respondents themselves if they are 18 or older. For the Safer Sex replication sites, the contractor will seek a waiver of parental permission from its IRB. Federal regulations permit the IRB to approve research without parent permission "if the IRB determines that a research protocol is designed for conditions or for a subject population for which permission is not a reasonable requirement to protect the subjects, provided an appropriate mechanism for protecting the children who will participate as subjects in the research is substituted and provided further that the waiver is not inconsistent with federal, state or local law". In sites such as the clinics that will implement Safer Sex, where adolescents can consent to treatment and procedures, such as contraceptive services,

pregnancy and disease testing, without parental knowledge, the contractor will seek such a waiver to protect the privacy of the adolescent. Youth themselves will be required to provide written assent. Both forms will explain the data being collected, and its use. The forms will also state that answers will be kept private, that youths' participation is voluntary, and that they may refuse to participate at any time. Participants and their parents/guardians will be told that, to the extent allowable by law, individual identifying information will not be released or published; rather, data collection will be published only in summary form with no identifying information at the individual level. Identifying and contact information will be stored in secure files, separate from survey and other individual-level data.

A copy of the parental consent form for program participants and a copy of the youth consent form are found in Attachment J.

A11. Justification for Sensitive Questions

Many of the measures in the baseline survey ask for information of a sensitive nature because the programs we will be evaluating are designed specifically to reduce sexual activity and associated risk behaviors among teens. Comprehensive measures of behavior are included because they will provide more accurate representations of teen sexual behavior, and the responses will significantly supplement the knowledge currently available on program effectiveness. Attachment B provides the justification for these and other questions and Attachment C provides detailed references.

Sensitive questions are drawn from previously-successful youth surveys and evaluations (see Attachment C). The items have been carefully selected, and we have been guided by past experience in determining whether or not the benefits of measures may outweigh concerns about the heightened sensitivity among sample members, parents, and program staff to specific issues. Although these questions are sensitive, they are commonly and successfully asked of youth similar to those who will be in the study, and all of these specific survey questions have been pretested among a diverse group of teens without any concerns raised about the questions' sensitivity. Most of the sensitive items related to sexual activity will be asked only of sample members who report being or having been sexually active.

A12. Estimates of Annualized Burden Hours and Costs

The baseline information collection does not impose a financial burden on youth respondents. Respondents will not incur any burden other than the time spent answering the questions contained in the survey.

Exhibit A12.1 summarizes the reporting burden on study participants. Enrollment for the TPP evaluation will take place over two years, so the annualized burden is based on half (4,275) of the expected sample (8,550).

Questionnaire response times were estimated from pretests with student respondents and from prior experience. The annual burden for questionnaire response is estimated from the total number of completed questionnaires proposed and the time required to complete the questionnaires. The total annual burden is expected to be 2,138 hours.

Exhibit A12.1. Reporting Burden on Study Participants

Respondent	Form Name	Annual Number of Respondents	Number of Responses per Respondent	Average Burden Hours per Response	Total Annual Burden Hours
Sexually Active Youth	Part A Safer Sex	1425	1	0.25	356.25
Sexually active and non-sexually active youth	Part A Reducing the Risk and Cuidate!	2850	1	0.25	712.50
Sexually active youth	B1 All Safer Sex youth and youth who have <u>ever</u> had sex in the Cuidate! and Reducing the Risk sites	2138	1	0.25	534.50
Non sexually active youth	Part B2 of the baseline survey to be used with youth who have <u>never</u> had sex in the Cuidate! and Reducing the Risk sites	2138	1	0.25	534.50
Total		(8551)	1	0.25	2137.25

Table A.12.2 Estimated Response Costs

Respondent	Form Name	Annual Number of Respondents	Number of Respondents 18 years and older	Average Hourly Wage of Respondents (18 and older)	Total Annual Response Cost
Sexually Active Youth	Part A Safer Sex	1425	998	\$7.25	\$1,809.37
Sexually active and non-sexually active youth	Part A Reducing the Risk and Cuidate!	2850	570	\$7.25	\$1,033.41
Sexually active youth	B1 All Safer Sex youth and youth who have <u>ever</u> had sex in the Cuidate! and Reducing the Risk sites	2138	1,378	\$7.25	\$2,498.31
Non sexually active youth	Part B2 of the baseline survey to be used with youth who have <u>never</u> had sex in the Cuidate! and Reducing the Risk sites	2138	190	\$7.25	\$344.47
Total		(8551)	3,136		\$5,685.56

The total annual response cost is \$5,685.5

A13. Estimates of Other Total Annual Cost Burden to Respondents and Record Keepers

These information collection activities do not place any additional cost on respondents.

A14. Annualized Cost to the Federal Government

This clearance request is specifically for collecting data at baseline. The total estimated cost to the government for the TPP Replication Study baseline data collection is \$1,600,000. Because baseline data collection will be carried out over two years, in order to achieve adequate sample sizes, the estimated annualized cost to the government for the baseline data collection is \$800,000.

A15. Explanation for Program Changes or Adjustments

No program adjustments are anticipated based on this data collection.

OMB gave approval on August 31, 2009 under a generic clearance (0970-0355) to conduct pre-tests of the baseline instrument. The PPA contractor, Mathematica Policy Research, Inc., conducted the pre-test and took the results into account - as well as advice from experts in the field - in redrafting the instrument. On August 17, 2011, the baseline survey instrument was approved (under 0970-0360).

HHS now seeks OMB approval for the baseline survey for the TPP Replication Study, a survey that is nearly identical to the approved PPA baseline survey (new items for the TPP Replication Study baseline survey are noted in the table in Attachment A). The data will be used for the impact analysis. Approval for follow-up surveys will be requested in a subsequent submission (the 60-day notice for the first follow-up survey was published at the same time as the baseline survey on March 16, 2011). The 60-day notice for a request for approval of the Implementation Study for the TPP Replication Study was published on September 23, 2011.

A16. Plans for Tabulation and Publication and Project Time Schedule

1. Analysis Plan

Before estimating impacts, HHS will conduct two analyses of the data from the baseline survey. First, HHS will use the data to describe the study sample and help define subgroups of policy interest. This step will enable HHS to compare the characteristics of youth in the study with youth nationwide and provide guidance on how the study sample and findings might generalize to a broader policy setting. Second, HHS will assess whether random assignment resulted in similar baseline characteristics of youth, on average, for the treatment and control groups.

To estimate program impacts, HHS will compare the outcomes of treatment and control group members in each site. Random assignment ensures that there are no systematic differences on measurable variables between the treatment and control group at the point of randomization. This ensures that any differences in their outcomes can be attributed with some confidence to the impacts of the intervention (and not to other factors, such as selection bias).

While the simple treatment/control mean outcome comparisons provides an unbiased estimate of true impact, HHS will estimate regression models that control for variation across the sample in baseline measures. Control variables will both increase statistical precision of the impact estimates for a given sample size, reduce the sample size requirements of the study for a given Minimum Detectable Effect size, and reduce attrition bias from missing

data, i.e., for a given sample size, regression adjusted estimates will have smaller standard errors.

For replications in which individual sample members are randomized to treatment or control, HHS will estimate an equation like equation (1) below. In equation (1), β_1 is the overall treatment effect, known as the Intent-to-Treat effect of the program:

$$(1) Y_i = \beta_0 + \beta_1 T_i + \beta_2 D_i + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i ,$$

Where:

Y_i is the outcome of interest (e.g. consistent condom use) for student i .

T_i is a dummy variable equal to 1 if student i was assigned to the treatment group

D_i is a clinic dummy (which accounts for blocking by clinic)

X_{mi} is the m^{th} baseline characteristic or control variable for student i (e.g. =1 for males).

The coefficient on the treatment dummy, β_1 , is the primary coefficient of interest. For an unfavorable outcome (e.g. teen pregnancy), a negative and statistically significant coefficient would be interpreted to mean that the program was effective in reducing the rate of that outcome. Impact estimates will be reported as standardized effect sizes.

For replications in which classrooms are randomized to treatment or control, we propose to estimate a regression model that accounts for the clustering of students within classrooms. The clustering of students within classrooms increases the variance of the impact estimates. Two methods are often used to correct standard errors for clustering: cluster-robust standard errors and multilevel modeling. Because we believe that readers of the teen pregnancy prevention literature will be more familiar with multilevel modeling, HHS will take that approach. Hierarchical Linear Modeling (HLM) has the added advantage that it enables the researcher to estimate what portion of the variance is attributable to each level of the model, which would be useful information in the design of future teen-pregnancy prevention evaluations.

Equations (2a) and (2b) provided a stylized version of the model we will use to estimate program impacts when classrooms are randomized:

$$(2a) \quad \text{Level 1:} \quad Y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \dots + \beta_{kj} X_{kij} + \varepsilon_{ij}$$

$$(2b) \quad \text{Level 2:} \quad \beta_{0j} = \gamma_1 T_j + \mu_j,$$

Where at level 1 (the individual level):

Y_{ij} is the outcome of interest (e.g. sex in prior 90 days) for student i in classroom j .

X_{kij} is the k^{th} baseline characteristic or control variable for student i in classroom k (e.g. =1 for males).

β_{0j} is the mean value of the outcome measure in classroom j

ε_{ij} is the residual error for student i from classroom j , which is assumed to be independently and identically distributed.

At level 2 (the classroom level):

T_j is a dummy variable equal to 1 if class j was assigned to the treatment group

γ_1 is the coefficient of interest, which represents the estimated impact of treatment

μ_j is the residual error for classroom j , which is assumed to be independently and identically distributed.

The coefficient on the treatment dummy, γ_1 , is the primary coefficient of interest. As with the previous model, for an undesirable outcome (e.g. teen pregnancy), a negative and statistically significant coefficient would be interpreted to mean that the program was effective at reducing the prevalence of that outcome. As before, impact estimates will be reported as standardized effect sizes.

After estimating these regression models for each replication, HHS will then compute pooled impact estimates across the replications for each of the three program models. In creating pooled impact estimates, HHS will weight the replications based on the number of individuals in the treatment group in that replication. This will produce estimates of the impact of each program for the average person who received the intervention as part of the three replications conducted under this evaluation.

2. Time Schedule and Publications

The TPP Replication Study evaluation will be conducted over a six-year period that began in Fall 2010 with a feasibility and design study. The contractor for the feasibility and design study (Abt Associates) assisted HHS with the identification of program models and replications and recruited the sites selected by HHS beginning in spring 2011. The baseline data collection will take place over a two-year period beginning in June 2012 and ending in Fall 2013. Follow-up data collections are projected to occur between February 2013 and December 2015. The implementation study will be conducted between Fall 2012 and Fall 2014. No formal publications are planned from the baseline data collection.

A17. Reason(S) Display of OMB Expiration Date is Inappropriate

All instruments will display the OMB number and the expiration date.

A18. Exceptions to Certification for Paperwork Reduction Act Submissions

No exceptions are necessary for this information collection.