Center on Children and Families
at BROOKINGS

ESTIMATING THE EFFECTS AND COSTS
OF THREE PREGNANCY-PREVENTION PROGRAMS

Adam Thomas[*]
The Brookings Institution

March 2011

This technical report documents some of the key assumptions underlying the analysis described in Thomas (2011), which presents results from benefit-cost simulations of a mass media campaign encouraging contraceptive use, an evidence-based teen pregnancy prevention program, and an expansion in access to family planning services provided via Medicaid. These simulations are performed using FamilyScape, which is an agent-based simulation tool that allows the user to model the impacts of policy changes on family-formation outcomes. Each policy's effects are estimated by comparing the results of simulations that were conducted under FamilyScape's baseline assumptions to the results of simulations that were conducted using an alternative set of assumptions regarding the presumed effects of the policy in question on contraceptive use and/or sexual behavior. See Thomas and Monea (2009) for a thorough treatment of the simulation model's baseline assumptions. In this report, I detail the key assumptions that underpin each of the policy simulations. More specifically, I discuss my assumptions regarding the costs and effects of each simulated policy. I do not discuss here the way in which these policies' estimated benefits are monetized. Given the complexity of this topic, it is addressed in a separate report that is co-authored by Emily Monea.[1] I begin the discussion below by describing the simulations of a teen pregnancy prevention program, after which I discuss the simulations of expanded access to Medicaid-funded family planning services and a mass media campaign. I conclude by addressing a variety of technical issues that are relevant for all three sets of policy simulations.

---

[1] See Monea and Thomas (2010).

Teen Pregnancy Prevention Programs

In order to conduct a benefit-cost simulation of an effective teen pregnancy prevention intervention, one must first articulate a set of assumptions about what such an intervention's effects might be if it were implemented on a national scale. I develop the parameters for this simulation by synthesizing data on the effects of several small-scale interventions that were rigorously evaluated and whose evaluations showed that they successfully affected certain key behavioral outcomes. Evaluations were included in this synthesis if they met all of the following criteria:

- The intervention being evaluated must primarily have served high-school-aged adolescents.

- The intervention must have been evaluated using a well-constructed randomized research design.

- The intervention's evaluation must have found that it had a statistically-significant effect on contraceptive use.

- The intervention's evaluation must also have found that it had a statistically-significant effect on the frequency of intercourse and/or on the share of teens who were sexually active.

- Interventions were excluded from consideration if their effects were found to have faded over a relatively-short period of time.

- Interventions were also excluded if attempts to replicate them were generally unsuccessful (although they were not excluded if there has not yet been *any* attempt to replicate them).

- Interventions were also excluded if their evaluations *only* reported on their effects on outcomes such as pregnancy or childbearing; in order to simulate the impacts of a given program using FamilyScape, one must have information on its effects on such antecedent behaviors as sexual frequency and contraceptive use.

- Interventions were also excluded if their evaluations measured their effects on the incidence of pregnancy and/or childbearing (among other outcomes) and found that they had no such effects, since the prevention of pregnancy and childbearing are the yardsticks by which the benefit-cost simulations measure the cost-effectiveness of a given policy.

After canvassing the relevant literature and speaking with a number of experts in this area, I identified five programs that meet these criteria: *Becoming a Responsible Teen* (BART), *HIV Prevention for Adolescents in Low-Income Housing Developments* (HIVP), *Safer Choices* (SC), and two programs that were developed using a core curriculum designed by John and Loretta Jemmott: *Be Proud! Be Responsible!* (BPBR) and *¡Cuídate!* (CDT).[2] In the next several subsections, I discuss the manner in which each of these programs was implemented, the populations that they served, the magnitudes of their estimated impacts, and their estimated costs per participant. For reasons that are detailed in a later subsection, I focus primarily on each program's estimated effects as measured at its evaluation's most recent follow-up.

Teen Pregnancy Prevention Programs: Becoming a Responsible Teen (BART)
BART was developed and implemented jointly by Janet St. Lawrence and Education, Training, and Research Associates (ETR Associates) and was evaluated by St. Lawrence and her coauthors. Unless otherwise noted, the information contained in this subsection was taken from St. Lawrence et al. (1995). BART was implemented for a group of African-American Adolescents aged 14 to 18 who were recruited from among the patients at a local public-health center in a mid-sized southern city. Recruited subjects were randomly assigned either to the study's treatment or control group. Members of the treatment group participated in an eight-week "education plus skills training" intervention.[3] The intervention consisted of weekly sessions lasting between 90 to 120 minutes each. The first session lasted two hours and focused on providing participants with the sort of information on HIV-AIDS prevention that an adolescent "might encounter in a classroom, health care, or community

---

[2] There are several other programs that are prominently featured in much of the evaluation literature but that were excluded from consideration here because they failed to meet one or more of the criteria listed above. For example, the evaluation of the *Teen Outreach Program* did not measure the program's effects on coital frequency, sexual inactivity, or contraceptive use; *Seattle Social Development* was implemented only for grades one through six (although its effects were measured at later ages); *SiHLE's* effects on pregnancy faded by its twelve-month follow-up, and its evaluation did not measure the program's effects on coital frequency or sexual inactivity; *Reducing the Risk* was not evaluated using a randomized research design; one version of *Postponing Sexual Involvement* was not evaluated using a randomized research design, while another was implemented only for seventh graders; one version of *Focus on Kids* was implemented primarily for children who are younger than high-school-aged, and another version did not measure the program's effects on coital frequency or sexual inactivity; *Aban Aya* was implemented for middle-school students, and its evaluation did not measure the program's effects on coital frequency or sexual inactivity; *Making Proud Choices* was implemented only for females aged eleven to 13; and *CAS-Carrera* was excluded because of the very high costs of implementing it, and because of the difficulties that have been encountered in replicating the successful results of the original intervention. For more information on these programs – and, more specifically, on the information contained in this footnote – see Advocates for Youth (2008), Kirby (2007), and Suellentrop (2009).
[3] St. Lawrence et al. (1995), p. 224.

based setting."[4] The next seven sessions adopted a variety of behavior skills-training strategies, including the promotion of technical competency skills (e.g., teaching correct condom use and stressing that sexual abstinence is the only guaranteed way of avoiding contraction of a sexually-transmitted disease), social competency skills (e.g., teaching communication skills and assertiveness), and cognitive competency skills (e.g., promoting accurate recognition of sexual risk and helping to develop problem-solving strategies). Members of the control group participated only in the traditional HIV-AIDS education session (i.e., the first of the sessions described above).

As of the follow-up evaluation conducted one year after completion of the intervention, St. Lawrence and her coauthors found that members of the treatment group reported having used condoms during intercourse about 30 percent more often than did members of the control group. They found further that members of the treatment group were a little less than 65 percent as likely to report having engaged in sexual intercourse over the previous two months as were members of the control group. Both findings were significant at the .05 level. The intervention's evaluation does not present any results for its effects on the frequency of intercourse among those who remain sexually active.

BART has been replicated at least two times. One replication took place in a drug-rehabilitation center, and another took place in a juvenile reformatory. The program implemented for the first replication was similar to the original intervention, but, for the second replication, it was shortened by more than 50 percent. The results of the first of these two replications were qualitatively similar to those of the original intervention, while the second replication was found to have had little if any effect on the behavior of treatment-group members.[5]

The disappointing results of the second replication suggest that, in order for this and similar interventions to be taken to scale successfully, it is probably important that the more-broadly-implemented program maintain a high level of fidelity in replicating the small-scale program(s) upon which it is based. The cost of implementing such a program is thus a

---

[4] Ibid.
[5] Kirby (2007), St. Lawrence et al. (2002).

critical consideration. BART's costs include the expenses incurred to train the staff who implement the program and the cost of the curriculum and training materials provided to, and used by, those staff members.[6] According to Child Trends (2009) and Manlove et al. (2004), the cost of training between 20 and 50 BART staff members was about $8,000. Child Trends (2009) also reports that the cost of materials per trainee was $60. The variation in the number of group leaders appears to be a function of the fact that groups were of differing sizes across sessions, which is to say that a greater number of facilitators were presumably required during some weeks than during others.[7] I assume here that a separate packet was necessary for every group leader, regardless of which sessions he/she led. I therefore assume the total cost for the program to be $8,000 + (50*$60) = $11,000. St. Lawrence (1995) reports that a total of 246 participants were assigned to the treatment group. Thus, I calculate the cost of the intervention per participant to be $11,000/246 ≈ $45, which is a little more than $70 in $2008.[8]

Teen Pregnancy Prevention Programs: HIV Prevention in Low-Income Communities (HIVP)
This program was developed, implemented, and evaluated by Kathleen Sikkema and her collaborators. The results of their evaluation were published in Sikkema et al. (2005); all information in this subsection was taken from that paper unless otherwise indicated. HIVP's evaluators randomly assigned subjects to one of three groups: 1) a control group whose members were invited to attend a standard community AIDS-education session that took place in the housing development in which they lived; 2) a "workshop-intervention" group whose members were invited to participate in two three-hour workshops that were designed to encourage participants to avoid risky sexual behavior; and 3) a "community-

---

[6] Note the cost of developing of BART's curriculum is not included among the items listed here. I was unable to find any estimates of this expense, and I assume that it is less relevant to the present discussion, given that the appropriate thought exercise relates to the question of how much it would cost to take an already-existing program(s) to scale.

[7] St. Lawrence et al. (1995) report that group sizes for BART were between five and 15.

[8] The evaluation of BART does not state the year in which it was implemented, but Manlove et al. (2004) write that it was implemented in "the early 1990's" (p. 11). I thus assume that the program was implemented in 1991 when inflating the cost of the program to $2008. Each participant in the study was also provided with a small stipend. Although participant stipends are common features of evaluations of this sort, I assume that, if such a program were taken to scale, stipends would not be provided to individuals participating in the intervention. I therefore do not consider these stipend expenses in estimating the program's cost. The assumption that participants would not be given stipends is one of several reasons why I argue in a later discussion that, if a program such as BART were implemented on a wider basis, its impacts would probably be smaller than those of the original, small-scale intervention.

6

intervention" group whose members were invited to attend sessions that were identical to those provided to the workshop-intervention group, and in whose housing developments the program's administrators distributed free condoms and conducted a variety of community-wide programs and parent workshops.[9] Participants in these three groups were all between the ages of twelve and 17 at baseline, and randomization was performed at the level of the housing development (i.e., all participants residing in housing development A were assigned to the control group, all participants in B were assigned to the community-intervention group, and so forth). In total, residents in 15 different housing developments participated in this experiment.

HIVP's evaluators found that there were no statistically-significant differences between members of the control and workshop groups in terms of the level of sexual activity or the likelihood of using contraception during intercourse. They did, however, find significant differences between the control and community-intervention groups. I therefore focus only on the program's estimated effects for the community-intervention group here, and I refer to this group simply as the "treatment group" for the remainder of this discussion.

AIDS-education sessions for adolescents in the control and treatment groups were completed within six months of baseline, and the community-level intervention was completed approximately 16 months after baseline (about ten months after completion of the AIDS-education sessions). The program's evaluators gathered follow-up data nine months after baseline (about three months after completion of the control group's AIDS-education sessions and about seven months before completion of the treatment group's community-level intervention) and 18 months after baseline (about twelve months after completion of the AIDS-education sessions and about two months after completion of the community-level intervention). They found that, as of the second of these follow-ups, treatment-group members who were sexually inexperienced at baseline were about 88 percent as likely to report having initiated sex as were sexually-inexperienced control-group members. They also found that a condom was reported to have been used at last sex about

_____

[9] Note that members of these groups were merely invited to attend the sessions described above. In fact, only about 15 percent of control-group members attended their group's sessions, about 87 percent of workshop-intervention group members attended their group's sessions, and about 86 percent of community-intervention group members attended their group's sessions.

24 percent more often among treatment-group members than among control-group members. Both differences were statistically significant at the .05 level.

The authors did not report any results regarding the coital frequency of sexually-active treatment or control group members, there have been no attempts to replicate HIVP, and Sikkema informed me that her team did not keep records of the program's costs.

Teen Pregnancy Prevention Programs: Safer Choices (SC)

Safer Choices was developed, implemented, and evaluated by the staff of ETR Associates. The results of the primary evaluation of the intervention were published by Coyle et al. (2001). Unless otherwise indicated, all information contained in this subsection was taken from that paper. The evaluation of SC was conducted using a randomized controlled trial involving a total of twenty schools in Texas and California. Ten schools (five in each state) were randomly assigned to the treatment condition, and the other ten schools were assigned to the control condition. The intervention lasted for two academic years, and its evaluation relied on data that were collected on students who were ninth graders during its first year. SC was designed to change the cultures of the schools in which it was implemented. The intervention's administrators created school-wide Health Promotion Councils at each treatment school that were comprised of teachers, students, parents, administrators, and community representatives; implemented a 20-session classroom curriculum for participating students; formed peer clubs that hosted program-sponsored, school-wide activities; sponsored a number of activities for the parents of treatment-group members, including newsletters and parent-student homework sessions; and attempted to reinforce the message of the program within the larger community by increasing students' level of familiarity with, and access to, support services outside of the school environment. Students in control schools received a standard, five-session sex-education program and a limited number of other school-wide activities that varied from site to site.[10]

The most recent follow-up was conducted 31 months after the beginning of the first of the two school years over which the intervention was implemented (i.e., about ten months after the completion of the second of these two academic years). The evaluation showed that, as

---

[10] Kirby et al. (2004).

8

of that follow-up, members of the treatment group were significantly more likely than members of the control group to report having used contraception at last intercourse (odds ratio = 1.76; p < .05). SC's evaluators also found that, as of the last follow-up, there was no statistically-significant difference between members of the treatment and control groups in the odds that they reported having initiated sexual intercourse since the beginning of the intervention, although the difference between the two groups was in the desired direction (odds ratio = .83; p = .39). In a subsequent re-analysis of their data, however, the authors did find that there was a significant difference in the incidence of self-reported sexual initiation between Hispanic members of the treatment and control groups (odds ratio = .57; $p < .05$). They found no such differences for other race groups; nor did they find any such differences in analyses that were disaggregated by gender.[11] The program's evaluators also found no significant differences between sexually-active members of the treatment and control groups in the self-reported frequency of sexual intercourse in the previous three months (although, again, their parameter estimate was in the desired direction: odds ratio = .81; $p = .12$).

Thus, SC only barely meets the criterion stated earlier that each program included in this discussion must be found to have affected both contraceptive use and sexual frequency and/or the probability of being sexually active. However, given the limited number of studies that do meet these criteria, and since the re-analysis of SC data indicated that the intervention did have an effect on sexual activity among Hispanics, I include it in this synthesis. I would also note that SC's evaluators measure the program's effects using odds ratios. However, for some outcomes, they report enough information to allow me to transform these quantities into relative risk ratios, which are more comparable to the published evaluation results for most of the other programs included in this exercise. For the purposes of the synthesis below, I therefore transform the reported estimates of SC's effects into relative risk ratios whenever possible.

---

[11] Ibid.

Regarding the cost of the intervention, Olaiya (2006) estimates that the annual, per-student cost of the first year of the program was $54, or about $80 in $2008.[12] I arrive at a roughly-equivalent estimate using comparable data that are reported on an itemized basis in Wang et al. (2000).[13] Wang and her coauthors estimate that the total annual cost of the program was $102,852, and Coyle et al. (2001) report that the treatment group had 1,983 members. Thus, my own back-of-the-envelope calculation suggests that the program's first-year, per-participant cost was $102,852/1,983 ≈ $52. Since SC is essentially a two-year program, I assume that it incurs a new facilitator-training cost every two years (see the later subsection on the assumed cost of the simulated program for additional discussion of my treatment of facilitator-training expenses). Given that Wang et al. (2000) and Olaiya (2006) express the annual cost of SC in terms of the expenses incurred to implement only the first year of the program, I calculate the full, two-year cost of the program using the itemized costs presented by Wang and her coauthors. Specifically, I assume that all costs other than facilitator training are incurred once per year, and that the cost of facilitator training is incurred once every two years. Wang et al.'s estimates indicate that, of the roughly $102,852 annual cost of the program, $54,619 was spent on facilitator training and the remaining $48,233 was spent on other expenses such as activity kits, teacher salaries, etc. I therefore assume that the full, two-year cost of the program was $54,619 + (2*$48,233) = $151,085, and I estimate that the full, two-year cost of the program per participant was $75 (i.e., ≈ $151,085/1983), or about $110 in $2008.

There have been no published evaluations of efforts to replicate SC.

Teen Pregnancy Prevention Programs: Be Proud! Be Responsible! (BPBR)
BPBR was developed, implemented, and evaluated by John and Loretta Jemmott and their colleagues. The results of the primary evaluation of the intervention were published by Jemmott et al. (1992); unless otherwise noted, all information presented in this subsection was taken from that paper. BPBR was initially implemented for a group of black males who were enrolled in the 10th, 11th, or 12th grades in the Philadelphia, PA area. The original BPBR

---

[12] Kirby et al. (2004) write that SC was implemented over the 1993-1994 and 1994-1995 academic years. I take 1994 to be the base year in inflating these costs to $2008.
[13] Olaiya's (2006) calculations are also based in large part on data presented in Wang et al. (2000).

curriculum was in fact tailored specifically for a black, inner-city, adolescent, male audience. Participants in the program's evaluation were recruited from a medical clinic, a high school, and a YMCA, all of which were located in West Philadelphia. These individuals were randomly assigned either to an AIDS risk-reduction condition (the intervention group) or to a career-opportunities condition (the control group). The AIDS intervention lasted for five hours on a single day and was designed to enhance participants' knowledge about issues related to AIDS and other sexually-transmitted diseases (STDs). According to BPBR's evaluation, it incorporated the use of "videotapes, games, exercises, and other culturally and developmentally appropriate materials."[14] Members of the control group participated in a session of equal length that focused on career planning.

The follow-up evaluation for BPBR was conducted three months after the intervention. Evaluators found at follow-up that, relative to members of the control group, treatment-group members reported having had sex on about 40 percent as many days over the previous three months and reported having used condoms more frequently during intercourse (on a five-point scale where 1 = "never" and 5 = "always," self-reported treatment-group and control-group scores were 4.4 and 3.5, respectively). Both differences were statistically significant at the .05 level. However, the authors also found no significant difference between the two groups in the likelihood of having engaged in intercourse at all, although the between-group difference was in the desired direction.

There have been several attempts to replicate BPBR, many of them successful.[15] One replication of the program was implemented for both boys and girls and was found to have had positive impacts on the sexual behavior of both. A second replication was implemented in a suburban high school during a regular class day (the original intervention was implemented on a Saturday outside of a school setting) and was not found to have had any effect on sexual behavior. Kirby (2007) hypothesizes that the in-school replication may have been less successful than the original BPBR intervention for one or more of the following reasons: 1) instructors were prohibited from discussing some of the topics that were part of the original program's curriculum; 2) because the intervention was implemented in the

---

[14] Jemmott et al. (1992), p. 373.
[15] The information in this paragraph is drawn from Borawski et al. (2009) and from Kirby (2007).

middle of a school day, participants might have been tired or less attentive; 3) participation was mandatory, which may have had an effect on the average subject's willingness to participate fully in the program; and 4) its participants were, on average, older than were the participants in the original program It is also possible that the replication's effectiveness was hampered by the fact that it was implemented in a suburban setting, since the original program was developed specifically for inner-city minority youth.

BPBR was also modified into two other distinct interventions that were tailored for somewhat-different demographic groups than the one for which BPBR was originally implemented. One modified version, *Making Proud Choices!* (MPC), was originally developed for black males aged 11 to 13, also in the Philadelphia area. A second modified version, *¡Cuídate! (CDT)*, was developed for Latino youth in Philadelphia. Both replications were successful in reproducing many of the most promising results of BPBR, and CDT is described in more detail in the next section. MPC is not included as a distinct intervention in this synthesis because it primarily served youth who were not high-school aged.

Manlove et al. (2004) report estimates of the cost of BPBR. The authors write that the program's costs include $5,500 for facilitator training, $2,500 to cover the costs of travel for trained facilitators and training materials, and $295 for a curriculum package. I assume that one curriculum package is required per facilitator. The authors also state that the sizes of the groups for the intervention were between six and twelve. I therefore assume that the average group size was eight and that the total number of groups for the 85-member intervention sample was thus $85/8 \approx 11$. These estimates thus suggest that the cost of curricula for the intervention was 11*$295 = $3,245 and that the total cost of the program was $5,500 + $2,500 + $3,245 = $11,345. I use these data, then, to calculate that the average cost of the intervention per participant was $11,345/85 \approx $135, or about $150 in $2008.[16]

---

[16] These costs appear to be expressed in terms of current dollars for the year in which the report was published (2004); $135 in $2004 \approx $150 in $2008. As was the case for BART, part of the overall cost of BPBR involved the provision of a modest stipend to participants in the program, and I once again exclude stipend expenses when estimating the program's cost. See the equivalent footnote in the subsection summarizing BART for the reasoning behind this exclusion.

I calculate a separate cost estimate for BPBR using the data reported by United Way of Rochester (2009). According to this report, the training of facilitators cost $1,000 per day per facilitator plus travel expenses, and a complete set of BPBR curriculum materials cost $358. The report also states that recommended training for the program ranges from six to 26 hours. I assume that the midpoint of these two extremes – 16 hours – represents the amount of training that is necessary to implement the program successfully. Thus, I assume that two days' worth of training is necessary to prepare a group of facilitators, and I continue to assume that eleven facilitators were needed to implement the program. I assume further that a single person is needed to train eleven facilitators, and that the cost of travel for this training session is $2,500 (as per the estimate cited above from Manlove et al.'s study). Thus, I estimate the total cost of the program using United Way-Rochester estimates to be 2*$1,000 + $2,500 + $358*11 = $8,438. This estimate implies that the average cost of the program per participant was $8,438/85 ≈ $95 in $2008.[17] Given that my cost calculations using data from Manlove et al. (2004) and UW-Rochester produce differing estimates, I assume for the sake of this exercise that the cost of the program per participant is $120, which is roughly equal to the average of these two figures.

Teen Pregnancy Prevention Programs: ¡Cuídate! (CDT)

CDT was developed, implemented, and evaluated by Antonia Villaruel, John Jemmott, and Loretta Jemmott. The results of the primary evaluation of the intervention are reported in Villaruel et al. (2006), and, unless otherwise indicated, all information discussed in this subsection is taken from that paper. As was stated in the previous subsection, CDT was a modified version of BPBR and was designed specifically for a Latino adolescent audience. It was implemented for a group of Latino teenagers aged 13 to 18 in the Philadelphia metropolitan area. Subjects were recruited from three local high schools and various community organizations and were randomly assigned to treatment and control groups. Members of the treatment group participated in six one-hour HIV-prevention sessions that were conducted over consecutive days and that used small-group discussions, videos, interactive activities, and skill-building exercises to encourage both abstinence and condom

---

[17] Since the United Way report was published in 2009, I assume that its cost estimates are expressed in $2009. Since these costs would only be negligibly affected by adjusting them from $2009 to $2008, I do not do so here.

use as methods for avoiding contraction of STDs.[18] The program's evaluators write that the intervention "incorporated salient aspects of Latino culture, specifically familialism, or the importance of family, and gender-role expectations."[19] Members of the control group participated in a culturally-specific health-promotion intervention of similar length that addressed such issues as diet, exercise, smoking, and drug and alcohol use.

CDT's evaluators presented estimates of the program's effect on the proportion of treatment-group members who reported having had sex over the previous three months, who reported having used condoms consistently over the previous three months, and who reported having used a condom at last sex (they did not report on between-group differences in the frequency of intercourse). They found that all of these differences were in the desired direction (OR = .66, OR = 1.91, and OR = 1.45, respectively), but that only the differences for sexual activity and consistent condom use were statistically significant. At baseline, there were already substantial differences between treatment- and control-group members in the self-reported consistency of contraceptive use. Specifically, treatment-group members were about a third more likely than control-group members to report consistent condom use at baseline (there were no notable differences between the share of treatment and control group members who reported having had recent intercourse at baseline). Thus, the evaluation's estimate of CDT's effect on contraceptive use is somewhat suspect. I nonetheless include this estimate in the synthesis below for the sake of completeness, because the proportional difference between groups in the self-reported consistency of condom use did in fact increase somewhat after the start of the intervention, and because the program's estimated effects on this particular margin of behavior are qualitatively consistent with the findings for the other interventions that are considered in this exercise.

The odds ratios for the outcomes described above were calculated using data from the evaluation's three-month, six-month, and twelve-month follow-ups. Thus, these results roughly reflect average differences between the treatment and control groups over the first year after completion of the intervention. However, the authors also report marginal tabulations for these outcomes at each of the three follow-up periods. At the three-month

---

[18] On the number of sessions over which CDT was implemented, see Advocates for Youth (2008).
[19] Villaruel et al. (2006), p. 773.

follow-up, members of the treatment group were about 84 percent as likely as members of the control group to report that they had engaged in sexual intercourse over the previous three months, and the equivalent differences at the six-month and twelve-month follow-ups were about 86 percent and about 88 percent, respectively. Members of the treatment group were also about 66 percent, about 56 percent, and about 53 percent more likely than members of the control group to report having used condoms consistently over the previous three months at the three-month, six-month, and twelve-month evaluations, respectively. Thus, the program's estimated effects did not decay all that much during the first year after completion of the intervention.

There have been no published evaluations of any attempts to replicate *¡Cuídate!*. As was discussed above, however, there have been several published evaluations of various iterations of BPBR and MPC, both of which were based on the same core curriculum as CDT, and these interventions have often (but not always) been found to have been successful at changing sexual behavior. Given the similarities between CDT and BPBR, and since the materials for both programs are produced by the same publisher, I assume that the costs of implementing CDT are comparable to those for BPBR.

Teen Pregnancy Prevention Programs: Programmatic Duration and Persistence of Effects
In order to develop a set of assumptions about what the effects and costs might be of a representative example of this diverse group of programs if it were implemented on a broader scale, one must decide how best to standardize their evaluation results. Before doing so, however, one must first set forth a clear thought experiment in terms of the duration of the program to be simulated and the persistence of its effects. I assume that estimates of an intervention's impacts as measured at the most recent follow up provides information on the durability of its effects. Table 1 thus summarizes information described in previous subsections regarding the approximate length of each intervention and the timing of its most recent evaluation follow-up. Roughly speaking, these programs can be grouped into two categories: those whose implementation was completed in a relatively short period of time and for which the final follow-up was conducted about a year after the program's completion (BART and CDT) and those for which this was not the case (HIVP, SC, and BPBR). For the two programs in the former category, I make the simplifying

assumption that the effects observed twelve months after the intervention's completion (and not much more than twelve months after the start of the intervention, since they were both relatively short) represent the average of its effects over a two-year period. This assumption is consistent with a variety of scenarios, including ones in which the program's effects remain constant over time or in which its effects fade after completion of the intervention at a roughly consistent rate as time passes.

Table 1. Durations of Interventions and Follow-Up Periods
for Studies Used to Parameterize Teen Pregnancy Prevention Benefit-Cost Simulations

Intervention	Approximate Duration of Intervention	Approximate Timing of Most Recent Follow-Up
Becoming a Responsible Teen	Eight Weeks	Twelve Months after Completion of the Intervention
HIV Prevention for Adolescents in Low-Income Housing Developments	*Community Intervention:* Ten Months *Preceding Workshops (also available to control developments):* Two Weeks; Administered over Six-Month Period	Two Months After Completion of the Intervention (18 Months After Baseline)
Safer Choices	Two Academic Years (≈ 21 months)	Ten Months After Completion of the Intervention (≈ 31 Months After Baseline)
Be Proud! Be Responsible!	One Day	Three Months After Completion of the Intervention
¡Cuídate!	One Week	Twelve Months After Completion of the Intervention

I also assume that the effects of the much-longer HIVP and SC peaked sometime around the programs' completion (16 and 21 months after the start of the intervention for HIVP and SC, respectively), that their effects were likely to have been comparatively strong while the intervention was ongoing, that those effects began to fade after reaching their peak levels, and that the program's effects as measured two months after completion of the intervention for HIVP (18 months after the start of the intervention) and ten months after completion of the intervention for SC (two years and seven months after the start of the

intervention) can therefore be considered to be rough proxies for their average effects over the two-year period that elapsed after the point in time when they began. Data from a period of similar length are not available for BPBR, which lasted only one day and for which the final follow-up took place only three months after the intervention. However, because the evaluation of BPBR does not report enough data to allow for any kind of extrapolation of its effects to a later point in time – and since its evaluation results are comparable to those of these other programs only if they are assumed to correspond to equivalent periods of time – I assume here that the program's effects as measured at the three-month follow-up roughly approximate its average effects over a two-year period.[20] Thus, for the purposes of the benefit-cost simulation, I make the simplifying assumption that the simulated program's effects will remain constant over a two-year period, and I estimate the magnitude of these effects by synthesizing these five interventions' evaluation results as measured at the most recent follow-up. I now turn to the task of synthesizing these interventions' evaluation results.

Teen Pregnancy Prevention Programs: Synthesis of Evaluation Findings
In the discussion that follows, I focus on these five programs' effects as measured at the most recent follow-up. Table 2 summarizes each one's key characteristics. The table makes clear that these programs' evaluations often measured their effects differently. For example, most of the estimated impacts reported in the SC evaluation are expressed as odds ratios (OR) rather than as relative risk ratios (RRR). This distinction has important implications for the practical implications of these estimates. Relative risk ratios can be interpreted as reflecting the proportional difference in key behaviors between two groups. For example, a RRR of 1.5 implies that members of the treatment group engage in the behavior in question (say, contraceptive use) 50 percent more often than do members of the control group. However, an OR has no such simple interpretation. In order to transform an OR into a more-readily-interpretable quantity, one must have data on sample members' baseline behavioral attributes. As is discussed below, the SC evaluation does in fact provide enough information to allow me to transform some of its reported odds ratios into relative-risk

---

[20] The plausibility of this assumption is bolstered by the evaluation results for CDT, which might be considered to be a "sister program" to BPBR. As is discussed in an earlier subsection, a comparison of CDT's estimated impacts at various follow-ups suggests that, at least over the first year after it was completed, the program's effects were in fact fairly constant.

ratios, although (as will be discussed shortly) such estimates should be interpreted with a measure of caution, given the assumptions that I was compelled to make in calculating them.[21]

The quantities estimated by these studies differ from one another in a number of other ways. For example, the evaluation of BPBR measures its effect on contraceptive behavior using a five-point qualitative scale, whereas the evaluations of BART, SC, HIVP, and CDT, respectively, measure their effects on the same margin of behavior based on the proportion of sexual encounters in which participants report having used a condom over the previous two months, the proportion of participants who report having used a condom at last sex, the proportion of participants who report having used any contraceptive method at last sex, and the proportion of participants who report having used condoms consistently over the previous three months.

Moreover, while the evaluations of BART and SC found that these programs affected condom use (for BART) and contraceptive use (for SC) in their full-sample analyses, the former found evidence of an effect in gender-disaggregated analyses among girls only, while the latter found a gender-specific effect for boys only. One can identify a number of other instances in which the outcomes tracked by these studies are measured somewhat differently or are not uniform in their basic finding of whether the program had an effect on a particular margin of behavior across demographic groups. I have therefore concluded that there is no sensible way of averaging all of the results reported in Table 2 into a single set of estimates that might credibly be taken to represent a precise quantitative aggregation of them. Instead, I synthesize these results in more of a qualitative fashion.

---

[21] The CDT evaluation also reports odds ratios. In addition, however, the authors report marginal tabulations of the share of treatment-group and control-group members at each follow-up who report using contraception and who report having engaged in sexual intercourse. Although the significance levels reported in the CDT evaluation refer specifically to the odds ratios that the authors estimate using data from all three follow-ups, I make the assumption here that the differences between treatment- and control group-members in these marginal tabulations as measured at twelve months can be taken to represent nonrandom differences between the treatment and control groups twelve months after completion of the intervention. I therefore I focus here on differences between the two groups at the twelve-month follow-up.

Table 2. Impacts of Selected Teen Pregnancy Prevention Programs Found to have Affected Both Sexual Activity and Contraceptive Use*

Among Interventions that Have been Evaluated using Random Assignment

Name of Intervention	Details of Evaluation Design for Initial Study	Estimated Program Effects on Sexual Abstinence / Initiation of Sex	Estimated Program Effects on Frequency of Intercourse	Estimated Program Effects on Male Contraceptive Use	Estimated Program Effects on Female Contraceptive Use	Estimated Program Cost Per Participant (in $2008)	Replication Information
Becoming A Responsible Teen	*Randomized controlled experiment* serving African-American youth. Participants were recruited from a low-income community in Jackson, MS. *Treatment group:* participated in eight sessions in a community-based setting, each one lasting 90 to 120 minutes. Curriculum designed specifically to prevent HIV infection among African-American adolescents. *Control group:* received one-time, two-hour HIV-prevention session. *N = 246 at baseline; 225 at follow-up one year after completion of the intervention.	One year after the end of the intervention, treatment-group members were about **65%** as likely as control-group members to report having had sex during the previous two months	No results reported for sexual frequency in evaluations of this program.	*Two months after the end of the intervention:* About **57%** more sexual occasions from the previous two months were reported to have involved the use of a condom among males in the treatment group than among males in the control group. *One year after the end of the intervention:* No significant difference between treatment-group and control-group males in the proportion of sexual occasions protected by a condom. However, combined-sex analyses showed a significant difference at one year: almost **30%** more sexual occasions from the previous two months were reported to have involved the use of a condom among males and females in the treatment group than among males and females in the control group.	*Two months after the end of the intervention:* About **16%** more sexual occasions from the previous two months were reported to have involved the use of a condom among females in the treatment group than among females in the control group. *One year after the end of the intervention:* About **44%** more sexual occasions from the previous two months were reported to have involved the use of a condom among females in the treatment group than among females in the control group.	≈ **$70**	*One successful replication:* Curriculum fully implemented in drug-rehabilitation facility; increased abstinence and condom use. *One unsuccessful replication:* Curriculum shortened by more than half and implemented in a state juvenile reformatory; no significant program effects on sex or contraceptive use.
HIV Prevention for Adolescents in Low-Income Housing Developments	*Randomized controlled experiment* serving adolescents aged 12 to 17. Participants were recruited from 15 low-income housing communities. *Primary treatment group:* residents of the housing developments that were randomly assigned to receive community treatment. Treatment consisted of distribution of free condoms and brochures, two three-hour workshops on HIV prevention, and a community-wide program with various neighborhood initiatives and workshops for parents. *Control group:* residents of control developments received free condoms and brochures, watched a videotape about HIV prevention, and discussed the video after viewing *N = 1,172 at baseline; 763 at follow-up two months after completion of the intervention.	Among participants who were sexually inexperienced at baseline, treatment-group members were about **88%** as likely as control-group members to report having initiated sex within two months of the end of the intervention.	No results reported for sexual frequency in evaluations of this program.	Self reports indicate that, as of the follow-up two months after the completion of the intervention, a condom was used at last sex about **24%** more often among treatment-group members than among control-group members.		Cost information not available from team that designed, implemented, and evaluated the intervention.	No published evaluations of any attempts to replicate program.
Safer Choices	*Randomized controlled experiment* implemented for freshmen and sophomores in twenty high schools in California and Texas. *Treatment group:* students in the schools that were randomly assigned to receive treatment. Intervention was implemented for all students in each treatment school and consisted of 20 sessions focusing on improving students' knowledge about condom use and sexually-transmitted infections and on changing their perception of abstinence in order to make it a more appealing option. In addition, clubs and councils were created and speaker series and parenting-education initiatives were implemented in order to change the culture within treatment schools. *Control group:* students at control schools received standard, five-session sexual-education curriculum and a few other school-wide activities that varied from school to school. *N = 3,869 at baseline; 3,058 at follow-up about one year after completion of the intervention.	*Among all members of the analysis sample:* No statistically-significant difference one year after completion of the intervention (or at earlier follow-ups) in the self-reported odds of having initiated sex between treatment- and control-group members who were sexually-inexperienced at baseline. *Among Latino members of the analysis sample:* About one year after completion of the intervention, sexually-inexperienced treatment-group members were significantly less likely than control-group members to report that they had initiated sex (odds ratio = **.57**).	About one year after completion of the intervention, no significant differences between treatment- and control-group members in the self-reported frequency of sexual intercourse over the previous three months (nor were such differences observed at earlier follow-ups).	About one year after completion of the intervention, males in the treatment group were significantly more likely to report having used contraception at last sex (odds ratio = **1.64**).	About one year after completion of the intervention, no statistically-significant difference between females in the treatment and control groups in the self-reported use of contraception at last sex (results for female contraceptive use not reported for earlier follow-ups, but evaluators found a significant difference in the self-reported use of contraception at last intercourse for the combined male and female samples while the intervention was ongoing; odds ratio = **1.76**).	≈ **$110**	No published evaluations of any attempts to replicate program.

Table 2, Continued. Impacts of Selected Teen Pregnancy Prevention Programs Found to have Affected Both Sexual Activity and Contraceptive Use*

Name of Intervention	Details of Original Evaluation Design	Estimated Program Effects on Sexual Abstinence / Initiation of Sex	Estimated Program Effects on Frequency of Intercourse	Estimated Program Effects on Male Contraceptive Use	Estimated Program Effects on Female Contraceptive Use	Estimated Program Cost	Replication Information
Be Proud! Be Responsible!	*Randomized controlled experiment* serving urban, African-American males aged 13 to 18 in the Philadelphia, PA metropolitan area. Participants were recruited from a local medical clinic, a neighborhood high school, and a local YMCA. *Treatment group:* participated in five-hour intervention designed to prevent HIV infection. Intervention techniques included small-group discussions, videos, and role-playing. *Control group:* participated in career-planning intervention of similar length. *N = 157 at baseline; 150 at follow-up three months after the intervention.	No statistically-significant difference observed three months after completion of the intervention between treatment- and control-group members in the share of participants who reported having had sex over the previous three months (among boys only).	Three months after the intervention, treatment-group members reported having engaged in about **40%** as much sex as control-group members over the previous three months (among boys only).	Three months after the intervention, a significant difference was observed between average self-reported treatment- and control-group scores (**4.4** vs. **3.5**, respectively) on condom-use scale where 1 = "never" and 5 = "always" (among boys only).	Intervention was for boys only.	≈ **$120**	*One successful replication:* Implemented in different communities from original for boys and girls, rather than just for boys; and was evaluated over six months, rather than over just three months. Found to have reduced the incidence of unprotected sex over the evaluation period. *One unsuccessful replication:* Implemented in high-school classrooms during school day. Not found to have any effect on sexual behavior, perhaps because it was mandatory (original version of the program was optional).
Modified Version of "Be Proud!": **¡Cuídate!**	*Randomized controlled experiment* serving Latino youth aged 13 to 18 in the Philadelphia, PA metropolitan area. Participants were recruited from three local high schools and various community organizations. *Treatment and Control groups:* received interventions similar to the ones described above for "Be Proud", although the intervention here was tailored specifically for Latinos and Latinas rather than for African-Americans. *N = 656 at baseline; 553 at follow-up one year after the intervention.	Using data from follow-ups conducted three months, six months, and one year after the intervention, evaluators concluded that treatment-group members were significantly less likely than control-group members to report having had sexual intercourse in the previous three months. At each of the three follow-ups, treatment-group members were about **85%** as likely as control-group members to report having had sex over the previous three months.	No results reported for sexual frequency in evaluations of this program.	Using data from follow-ups conducted three months, six months, and one year after the intervention, evaluators concluded that treatment-group members were significantly more likely to report using condoms consistently. Across the three follow-ups, treatment-group members were between about **50%** and about **65%** more likely than control-group members to report having used condoms consistently over the previous three months. However, no statistically-significant difference observed using data from the three follow-ups between treatment- and control-group members in the share of participants who reported having used condoms at last sex.			No published evaluations of any attempts to directly replicate program. However, *Making Proud Choices!* (MPC), like *¡Cuídate!*, was based on the *Be Proud!* curriculum. MPC: implemented for black boys and girls aged 11 to 13, found to have reduced self-reported sexual frequency and increased self-reported contraceptive use. See above for information on successful *Be Proud!* implementations.

*Notes: Unless otherwise indicated, all findings listed here are statistically significant at the .05 level. The summaries for all five programs are based in large part on information taken from three overviews: Advocates for Youth (2008), Kirby (2007), and Suellentrop (2009). Additional information on Becoming a Responsible Teen was taken from Child Trends (2009), from St. Lawrence et al. (1995), and from St. Lawrence et al. (2002); additional information on HIV Prevention for Adolescents in Low-Income Housing Developments was taken from Sikkema et al. (2005); additional information on Safer Choices was taken from Coyle et al. (2001), from Kirby et al. (2004), from Olaiya (2006), and from Wang et al. (2000); and additional information on Be Proud! Be Responsible!, on Making Proud Choices!, and on ¡Cuídate! was taken from Jemmott et al. (1992), and from Villaruel et al. (2006).

20

Regarding these programs' estimated effects on contraceptive use, I make a variety of simplifying assumptions in order to draw direct comparison between their evaluation results. First, I assume that estimates of programs' effects on the probability of using contraception at last sex are comparable to estimates of their effects on the proportion of sexual encounters in the recent past that involved the use of contraception. Second, I assume that estimates of programs' effects on condom use are comparable to their estimated effects on contraceptive use more generally. Third, for the one evaluation that was implemented for boys only (BPBR), I assume that its results for contraceptive use and sexual behavior are comparable to those of the other programs that are implemented for boys and girls. And fourth, for studies that present estimates both for their entire samples and for boys and girls separately, I focus on full-sample results, since: 1) two of the five studies present only combined-gender results, and 2) the two studies that present contraceptive-use results that are disaggregated by gender arrive at opposite conclusions (as discussed earlier, BART was found in gender-disaggregated analyses to have significantly affected contraceptive use among girls but not among boys, and SC was found to have significantly affected contraceptive use among boys but not among girls; both studies also found that there was a statistically-significant difference in contraceptive use between the full treatment and control groups).

Additionally, in order to make SC's estimated impacts on contraceptive use more comparable to the results of other studies, I perform a back-of-the-envelope calculation to transform the odds ratio reported by that program's evaluators into a relative risk ratio. I estimate that the full-sample odds ratio of 1.76 is comparable to a relative risk ratio of about 1.2. This estimate should be treated with a measure of caution, as I was required to make an untestable (but, I would argue, reasonable) assumption in calculating it. My confidence in this estimate is enhanced by the fact that it is quite consistent with comparable results from other studies. Moreover, as is discussed in the footnote below, the qualitative implications of the estimate are quite robust to reasonable changes in the aforementioned assumption.[22]

---

[22] The odds ratio described above is calculated as the quotient of treatment and control group members' odds of having used contraception at last sex, where each group's odds are calculated as the proportion of the group that used contraception divided by the proportion that did not. Because the authors of SC's evaluation report enough data to allow me to calculate the treatment and control groups' odds of having used contraception at

After making the assumptions enumerated above and transforming the odds ratio reported in the SC evaluation into a relative risk ratio, I am able to compare directly the estimated effects on contraceptive use of BART, HIVP, and SC. These interventions' evaluation results suggest that they increased contraceptive use by 30 percent, 24 percent, and 20 percent, respectively. A rough average of these programs' effects thus suggests that they collectively increased contraceptive use by about 25 percent. Since the BPBR evaluation reports only the program's effects on participants' self-ratings on a five-point scale measuring the consistency of contraceptive use, I can not directly incorporate these results into my estimates. Interestingly, however, the average score on this scale for the treatment group is almost exactly 25 percent higher than the equivalent control-group average. I thus conclude that the BPBR results are qualitatively consistent with the results from the other three studies.

The results for the evaluation of CDT are somewhat of an outlier. While treatment-group members were more than 50 percent more likely than control-group members to report consistent contraceptive use, there was no significant between-group difference in the likelihood of having used a condom at last sex (the difference between the two groups was, however, in the desired direction). Given the incongruity of the findings that CDT increased consistency of contraceptive use but had no effect on the use of condoms at last sex, and because of the notable similarity in contraceptive-use estimates across the other four studies,

baseline but not at follow-up, I use the odds ratio reported at follow-up to estimate the treatment group's odds of using contraception after completion of the intervention under the assumption that the control group's contraceptive behavior at follow-up was similar to their behavior at baseline. The SC evaluation reports that, at baseline, 59 percent of control-group members reported having used contraception at last sex. Thus, I assume that, at follow-up, control group members' odds of using contraception are $(.59/.41) \approx 1.44$. Since the odds ratio at follow-up (1.76) is simply the treatment group's odds divided by the control group's odds, I calculate the treatment group's odds of having used contraception at last sex to be $(1.76*1.44) \approx 2.53$. I then calculate the proportion of treatment group members who used contraception at last sex to be $(2.53/(2.53+1)) \approx .72$. Given my assumption that, at follow-up, 59 percent of control-group members used contraception at last sex, I estimate that the proportional difference in the probability of having used contraception at last sex between treatment- and control-group members at follow up was $(.72/.59) \approx 1.2$. The credibility of this estimate is determined by the plausibility of my assumption that the proportion of control-group members who used contraception at last sex did not change between baseline and the most recent follow-up. If one were to assume that the proportion of control-group members using contraception at follow-up was actually, say, .5, .55, .65, or .70, the value of the relative risk ratio estimated here would instead be 1.27, 1.24, 1.17, and 1.15, respectively. If I were to assume that the relative risk ratio for SC were any one of these estimates rather than 1.2, my ultimate conclusion as to the rough average of the effects of the programs considered in this review would be unchanged.

I choose to rely on information from the latter in developing an assumption about the collective implications of these programs' evaluation results. I conclude, based on the information presented above, that these programs collectively increased contraceptive use at last sex by about 25 percent.

All five program's evaluations also estimated their effects on some measure of sexual activity. The evaluations of BART, BPBR, and CDT measured their programs' effects on the probability of having had sex at all during the previous two months (BART) or the previous three months (BPBR and CDT), and the evaluations of HIVP and SC measured their effects on the initiation of sexual activity among participants who were sexually inexperienced at baseline. Thus, one can relatively straightforwardly compare the results of the BART, BPBR, and CDT studies (ignoring, for purposes of practicality, the distinction between the two-month window used for the BART evaluation and the three-month window used for the BPBR and CDT evaluations), and one can similarly compare the results for sexually-inexperienced participants from the HIVP and SC evaluations. (Unless otherwise noted, I refer to participants as "sexually inexperienced" for the remainder of this discussion if they characterized themselves in this fashion during the baseline evaluation of the study in which they participated.)

With respect to the first group of studies, two of the three relevant evaluations showed that the intervention in question affected the share of treatment-group members who reported having had sex in recent months. The average estimated effect across these three studies on this margin of behavior – including the finding of no significant effect reported in the evaluation of BPBR – is about .83. With respect to the second group of studies, one intervention (HIVP) was found to have reduced the proportion of sexually-inexperienced participants who initiated sex as of the most recent evaluation (estimated effect = .88), and the other (SC) was found to have had no such effect within the full sample, but it *was* found to have had an effect among Hispanic sample members (OR = .57). In their paper in which they found that SC an effect for Hispanics, Kirby et al. (2004) report that this group constituted a little more than a quarter of their sample. However, they do not provide enough information in their evaluation to allow me to transform their odds ratio into a relative risk ratio. One might take the SC and HIVP results to imply jointly that an effective

teen pregnancy prevention program could have an impact – but a perhaps only a small one – on the share of sexually-inexperienced participants who remain abstinent for a period of time after the implementation of a successful teen pregnancy prevention intervention.

This conclusion is roughly consistent with the implications of the collective results for BART, BPBR, and CDT if one assumes that a program's effect on whether an individual has sex in the near term is a function both of the program's effect on the continuation of lifetime abstinence among individuals who were sexually inexperienced at baseline *and* on the continuation/initiation of "temporary abstinence" among individuals who were not. Unfortunately, none of these evaluations provides information specifically on whether individuals were sexually active in the near term as a function of whether they were sexually experienced at baseline. As such, I assume, for purposes of practicality, that HIVP's and SC's evaluation results are in fact roughly consistent with those of BART, BPBR, and CDT, and that these programs collectively caused treatment-group members to be about 15 less likely to engage in sexual intercourse over a three-month period.

Next, I examine these interventions' estimated effects on coital frequency among sexually-active individuals. Only two of the five programs' evaluations measured their effects on the frequency of sexual activity independent of measuring their effects on whether participants engaged in intercourse at all. SC's evaluation measured its effects on coital frequency over the previous three months among all participants who were sexually active at follow-up (without regard to whether they were sexually experienced at baseline), and it found that the program had no such effect. The evaluation of BPBR measured its effects on the number of days during the previous three months on which participants report having had sex, and it found that members of the treatment group reported having had sex on 40 percent as many days as did members of the control group. The quantities measured by these two studies are comparable to the extent that they measure sexual frequency in the prior three months. However, they are not comparable to the extent that SC's evaluation studies this margin of behavior only among individuals who were sexually active at follow-up, while BPBR's evaluation appears to include sexually-inactive individuals in its calculations (which is to say that, if an individual had no sex in the previous three months, his level of coital frequency would have been included as a zero in calculations of the BPBR estimates).

24

Because the BPBR evaluation found that the program had no significant effect on whether individuals had sex in recent months, and since sexually-inactive individuals appear to have been included in the estimate of the program's effect on coital frequency, one might conclude that its estimated effects on the latter margin of behavior are primarily attributable to a reduction in sexual frequency among participants who were recently sexually active. However, SC's results suggest just the opposite – the estimate from that program's evaluation is limited specifically to sexually-active individuals, and the authors find that the program had no significant effect on this margin of behavior. To summarize, then: a) the evaluations of only two of the programs included in this synthesis measured their effects on coital frequency; b) only one of these two programs (BPBR) was found to have had such an effect; c) the relevant evidence is in fact only suggestive that the intervention in question affected the frequency of intercourse among individuals who were sexually active; d) this finding also stands in opposition to the evaluation results for SC, which was found to have had no such effect; and e) I already assume that these programs collectively affected sexual behavior by reducing the number of teens who were sexually active in the near term. I therefore conclude that there is not sufficient evidence that these programs had an effect on coital frequency independent of their impact on sexual inactivity to allow me to draw any definitive conclusions in this regard. As such, I model the simulated program's impact on sexual activity by assuming only that it affects the number of teens who are sexually active over any given three-month period.

Teen Pregnancy Prevention Programs: Alternative Assumptions
To summarize my conclusions from the previous section, I surmise that the five programs reviewed here collectively increased the proportion of participants who used contraception at last sex by about 25 percent, and that they reduced the proportion of individuals who were sexually active in an average three-month period by about 15 percent. These assumptions constitute the basic building blocks of the parameters that I use to model changes in sexual and contraceptive behavior for the initial simulations of the effects of an evidence-based teen pregnancy program. I conduct a second set of simulations, however, using alternative assumptions that were developed based on the work of Lauren Scher, Rebecca Maynard, and Matt Stagner (2006). Scher and her coauthors, in a report for the

Campbell Collaboration, scrutinize estimates of the effects of teen pregnancy programs from a large number of published evaluations. The authors conclude that many of the studies included in their review overestimate the effects of the effects of the programs that they evaluated. Among the most common reasons for such overestimation, according to the authors' analysis, is the fact that some studies failed to adjust their standard errors appropriately for the level of randomization (e.g., in analyzing individual-level data for an intervention that is randomized at the school level, one should calculate standard errors that are clustered at the school level) and that some studies failed to account for selection effects (such as when a study limits its analysis to individuals who completed the intervention or were sexually active at a particular follow-up). In light of these concerns, the authors re-estimate the results reported in these studies using data from the originally-published evaluations in each instance in which the original study provided enough data to allow them to do so. For each study that did not properly account for the level of randomization, they re-estimate the study's results with what they consider to be the proper clustering techniques. For studies that estimated their results only for specific subgroups of their samples (e.g., individuals who were abstinent at baseline or were sexually active at follow-up), they re-estimate the evaluation's results using information for the full sample.[23]

Given the authors' preference for estimating program effects using the full sample in all cases, it is not possible for them to produce contraceptive-use estimates that are comparable to the ones presented in the synthesis above, since the estimates described above are always limited to individuals who were sexually active. Instead, they create a new "pregnancy risk" variable that is set equal to one for individuals who report that they do not always use

---

[23] The study's lead author told me that she does not consider the limitation of the sample to subgroups of individuals on the basis of baseline characteristics to be experimentally unsound, so long as one believes the relevant subgroups within the treatment and control groups to have been randomly assigned. Thus, for example, limiting the sample to Hispanics (as the authors of one of the SC evaluations did) or to individuals who were sexually abstinent at baseline (as the evaluators of both SC and HIVP did) is not problematic, so long as one believes the members of these subgroups within the treatment and control groups to be comparable in terms of their observable and unobservable characteristics. The authors' primary concern in this regard relates instead to instances in which studies focused on specific groups based on participants' behavioral characteristics as measured after the intervention was implemented (e.g., limiting the study sample to individuals who were sexually active at follow-up for the purposes of measuring the program's impact on contraceptive use), since the intervention may have caused participants to select into or out of those groups. Such sample limitations have the potential to confound the program's effects on, say, sexual activity with its effects on, say, contraceptive use. For the sake of consistency, the authors took the general approach of re-estimating all evaluation results using the full sample whenever possible, even when subgroups were created based only using baseline characteristics.

protection when they have sex, or – if data on the consistency of contraception use are not available – for individuals who report that they have had sex but did not use protection at last intercourse. This variable is set equal to zero if none of these conditions is met. Thus, individuals are coded as zeroes for the authors' pregnancy-risk variable if they report that they used contraception consistently (for studies that collect data on consistency of contraceptive use), if they report that they used contraception at last sex (for studies that do not collect consistency-of-use data), *or* if they report that they are sexually abstinent. This approach allows them to include all sample members in their analysis of contraceptive behavior. The authors' pregnancy-risk variable might therefore be thought of as capturing the joint effects (if any) that a given intervention had on sexual inactivity, coital frequency, and contraceptive use. Scher and her coauthors also estimate programs' effects on the proportion of treatment- and control-group members who have ever had sexual intercourse. Thus, if a study estimates an intervention's impact on, say, the probability of having had sex in recent months, the results of the Scher et al.'s analyses may differ from the results of the original evaluation, even if the original evaluation's results were properly estimated.

Some of the evaluations included my synthesis of the literature were also included in Scher et al.'s review. The authors did not consider the results of the CDT evaluation because it was published after they concluded their review.[24] And, although they briefly discuss the SC and BPBR evaluations that are included in my synthesis, they do not re-estimate the results of these studies because they do not have enough information to allow them to re-estimate these evaluations' findings. The report's lead author informed me that the exclusion of these evaluations from their analysis does not necessarily reflect on the quality of their estimates; rather, these studies simply did not present enough data to allow the authors to evaluate their findings properly.[25]

Scher et al. do, however, re-estimate the effects of BART and HIVP. They re-estimate the results of the BART evaluation using the full treatment- and control-group samples for all

---

[24] Based on a conversation with the lead author of the study.
[25] Scher et al. do analyze the results of an evaluation of another version of BPBR that was implemented for boys and girls in the sixth and seventh grades, and they find that the impacts reported in the original study are no longer significant after they re-estimate the program's effects using the full sample. However, I did not include the results of this evaluation in my synthesis because the program in question was implemented for adolescents who were younger than high-school-aged.

analyses, and they adjust the standard-error estimates reported in the HIVP evaluation to account for the fact that participants were randomized at the level of the housing development. In their re-analysis of the HIVP and SC data, the authors find that neither program had a statistically significant effect on the probability of ever having engaged in intercourse, but that both of them had a significant and relatively-substantial effect on the probability of exposure to sexual risk. With respect to the latter outcome, the authors find that BART reduced the pregnancy risk of members of its treatment group by almost 60 percent ($p < .1$) and that HIVP reduced the pregnancy risk of members of its treatment group by nearly 40 percent ($p < .05$).

For a variety of reasons, it is unclear how best to map the results of the Scher et al. re-analysis of these data onto the conclusions of the synthesis described in the previous section. First, three of the five programs included in the synthesis were not considered in Scher et al.'s re-analysis of evaluation findings. Second, Scher et al.'s finding that neither BART nor HIVP had a significant effect on ever having had sex is not necessarily inconsistent with the original evaluations' findings, since both of them reported the program's effects not on lifetime abstinence, but on the probability of having had sex in recent months. And third, the finding that both programs had large and statistically-significant effects on pregnancy risk – which I assume to be a composite measure of these interventions' effects on sexual inactivity, coital frequency, and contraceptive use – appears to be qualitatively consistent with the general conclusions of my synthesis. On the other hand, the authors highlight notable methodological drawbacks of the HIVP evaluation in particular, since that study appears not to have estimated standard errors that were clustered at the level of the housing project. Moreover, given Scher et al.'s finding that neither of these two interventions had a significant effect on ever having had intercourse, it is possible that the original evaluations of these programs overstated their effects on sexual activity. I therefore conduct an additional set of teen pregnancy-prevention simulations under the assumption that the intervention has the same effect on contraceptive use as is assumed under the initial specification, but that it has no effect on sexual behavior.

Teen Pregnancy Prevention Programs: Simulation Parameters

Recall that each of the five interventions described in the previous section was implemented on a relatively small scale. As is shown in Table 2, the baseline sample sizes for the BART, HIVP, SC, BPBR, and CDT evaluations were 246, 1,172, 3,869, 157, and 656, respectively – and about half of the participants in these studies were assigned to control groups. The simulation of an evidence-based teen pregnancy prevention program assumes that the most efficacious components of these programs will be combined to create an intervention that is implemented on a national scale. I have been advised in conversations with numerous individuals who have evaluated interventions of this sort that, if such a program were taken to scale, it would almost certainly have a substantially smaller effect than did these small-scale programs. This assumption is rooted in the fact that, if such a program were taken to a national scale, it would be difficult to maintain a high level of fidelity to the intensity and quality of facilitator training and supervision, instructional practices, and community outreach efforts that were achieved in the initial iterations of these programs, since they were often implemented by individuals who were deeply committed to the success of these interventions.

I have therefore concluded that it is reasonable to assume that, if such a program were implemented on a national scale, its effects on sexual behavior and contraceptive use would be about half as large as the effects of these small-scale programs.[26] Thus, in parameterizing the simulation of an evidence-based teen pregnancy intervention that is implemented on a national scale, I assume that the program's effects would be half the size of the effects of the small-scale programs described earlier. Given that most of the programs included in this synthesis were implemented for what might be called "at-risk youth," (e.g., minority adolescents in urban areas or teens living in public housing developments), I conduct the teen pregnancy prevention simulation only for unmarried, teenaged members of the

---

[26] One program evaluator provided an alternative perspective. She suggested that a program implemented nationally could have *larger* effects than these small-scale programs, since a large-scale program might induce the many adolescents participating in it to affect one another's behavior. She also argued that a well-funded, ongoing, national program might improve over time as the individuals designing and administering it learn how best to refine its curriculum and reinforce its message based on the rich array of programmatic experiences that would be afforded by a large-scale implementation. For the purposes of the present exercise, however, I opt to rely upon the assumption of the bulk of experts with whom I spoke that taking an effective program to scale would likely dilute its effects.

simulation population who are tagged as "low-SES," under the assumption that socioeconomic status serves as a reasonable proxy for "at-risk" status.[27] Table 3 summarizes the parameters that are used to implement the simulation of this program under the initial and alternative assumptions described above.[28]

Table 3. Parameters Used to Simulate
An Evidence-Based Teen Pregnancy Prevention Program
for At-Risk Teens
That is Implemented on a National Scale

	Effects of Intervention on the Proportion of Low-SES Teens Who Have Sex During an Average Three-Month Period	Effects of Intervention on The Proportion Of Low-SES Teenaged Males Who use Condoms During Intercourse	Effects of Intervention on The Proportion of Low-SES Teenaged Females Who use Oral Contraception During Intercourse
Initial Specification	7.5% reduction	12.5% increase	12.5% increase
Alternative Specification	--	12.5% increase	12.5% increase

[27] I do not model behavioral changes among the small group of low-SES teens who are married. I choose not to include married teens in this simulation for several reasons. First, the participants in the programs included in this synthesis were almost all unmarried, and it is unclear whether it would be appropriate to apply to married couples the assumptions described above regarding the effects of these programs on the (overwhelmingly-unmarried) populations that they served. Second, it seems particularly unreasonable to assume that a pregnancy-prevention program would induce temporary abstinence among married couples. Third, although one could imagine that a well-structured program might cause married teen couples to use contraception more effectively than would otherwise have been the case, such a program might differ in structure and tone from the ones studied here, which appear, for all practical purposes, to have been geared specifically towards unmarried teens. Thus, it might not be the case that a common intervention would affect married and unmarried teens alike. And finally, married teens are such a small group within the simulation population (and within the real-world population) that including them in the intervention's target group would have no material effect on the qualitative implications of the simulation's results.

[28] In Thomas and Monea (2009), my coauthor and I report that surgical sterilization, condoms, and oral contraception (i.e., the pill) are each used by somewhat less than a third of contraceptors, and that the remaining share of contraceptors rely on one of a multitude of alternative options. In constructing FamilyScape, we therefore chose to simulate the use of only these three methods. The sub-set of female contraceptors who report having used other methods are, for the purposes of the simulation, considered to have used the pill. Thus, simulating a 12.5 percent increase in the number of teenaged pill users here is equivalent to simulating a 12.5 percent increase in the number of female teenagers who use any method of contraception other than condoms or sterilization.

Teen Pregnancy Prevention Programs: Estimated Cost

Finally, I discuss the simulation's assumptions regarding the cost of the intervention. As per the discussion in a previous subsection, I take the effects described above to reflect the simulated program's average impacts over a two-year period. Given that none of these programs exceeds two years in duration, I make the simplifying assumption that, in order for these effects to be achieved, teens must therefore participate in this "synthetic intervention" once every two years, and I estimate the annual, per-participant cost of such a program to be equal to half of the average total cost of the program per participant. Embedded in this latter assumption is yet another important assumption: that facilitators for these programs must be newly trained every two years. The cost of training facilitators is a key expense for most programs – it might be thought of as constituting the bulk of such programs' short-run "fixed costs," such as they are – and, in discussions with individuals who have implemented and evaluated interventions of this sort, I have been advised that their ability to staff their programs with well-trained facilitators is critical to their success. In estimating the cost of a program that is implemented somewhat continuously, one must therefore make some kind of assumption as to the frequency with which its short-run fixed costs (i.e., facilitator training) will be renewed. Because of the pivotal importance of having well-trained facilitators, and under the assumption that there will be turnover in facilitator positions over time, I assume that a new facilitator-training fixed cost will be incurred every two years, and that this cost will be the same each time.

I use the cost estimates reported in Table 2 for the programs included in the synthesis of the evaluation literature to formulate an assumption of the simulated program's cost. I express the cost of the program in $2008 and on a per-participant basis. I present three cost estimates in Table 2: BART is estimated to cost $70 per participant, SC is estimated to cost $110 per participant, and BPBR and CDT are estimated to cost $120 per participant. The average of these costs is exactly $100. Thus I assume that this program, if implemented on a national scale, would cost $100 per participant every two years, and I therefore assume that the average annual cost of the intervention per low-SES teen (i.e., per member of the target population) would be $(100/2) = 50. I combine this estimate with my tabulation of the number of low-SES teens to calculate that the annual cost of the program would be about

$145 million.[29]

Expanded Access to Medicaid Family Planning

In this section, I describe the parameters that are used to implement the simulation of an expansion in income eligibility for Medicaid family-planning services. I begin by discussing my assumptions regarding the expansion's effect on contraceptive use. In subsequent subsections, I then compare the results produced by this simulation to those of related studies; I discuss the way in which I account for the fact that individuals would presumably be affected by this expansion only if they live in states that have not yet implemented income-based waivers; I detail my method for imputing income eligibility for take-up of Medicaid family-planning services; and I present estimates of the cost of the expansion.

Medicaid Family Planning Expansion: Estimated Effects

As is discussed in Thomas (2010), women have traditionally been able to take up Medicaid family-planning services only if: a) they are pregnant or have children and b) their incomes fall below a relatively low threshold. Over the last 15 years, however, the federal government has granted waivers to 21 states allowing them to serve all income-eligible women – regardless of whether or not they are pregnant or have children – and, in most cases, to raise their income-eligibility thresholds as well.[30] And, more recently, the newly-enacted Patient Protection and Affordable Care Act grants states the option to increase their income-eligibility thresholds for family-planning services to a level that is less than or equal to the thresholds that they use to determine eligibility for Medicaid pregnancy-related care.[31] In Thomas (2010), implement this simulation under the assumption that income-eligibility

---

[30] Another six states have been granted "duration waivers," which allow them to extend coverage for Medicaid family-planning services to women who would otherwise have lost their Medicaid coverage for any number of reasons. Four of these six states cover women who received pregnancy-related care through Medicaid but who would normally have lost their access to Medicaid-subsidized family-planning services after the standard 60-day postpartum period during which such services are typically offered. The other two states cover such services for women who would otherwise have lost Medicaid coverage for any reason. Women who qualify for coverage in these states generally receive family planning services for two additional years. However, the paper by Kearney and Levine (2009) described below suggests that these waivers have had little effect on contraceptive use or pregnancy rates. Moreover, the state option in the new health care legislation specifically allows states to expand their income-eligibility criteria for Medicaid family-planning services; it does not allow them separately to extend the period of time over which women are eligible for these services. I therefore focus here only on income-based expansions in eligibility for these services and, in the simulation described below, I assume that these expansions would be implemented in states that were never granted any family-planning waivers and in states that were only granted duration waivers. For state-by-state information on Medicaid family-planning waivers, see Guttmacher Institute (2010a).

[31] National Campaign to Prevent Teen and Unplanned Pregnancy (2010), United States Congress (2010).

expansions take place in all states that have not yet implemented them (hereafter, "non-waiver states"). In point of fact, the health-care legislation could affect some states that have already implemented income-eligibility expansions (hereafter, "waiver states"). A comparison of waiver states' income-eligibility thresholds for Medicaid pregnancy care as reported in Kaiser Family Foundation (2010) and of their thresholds for Medicaid family-planning services as reported in Guttmacher Institute (2010a) suggests that, in five of these states, the former is greater than the latter.[32] Thus, it is possible that this small number of states could also incrementally increase their income-eligibility thresholds somewhat further. For purposes of simplicity, however, I ignore this possibility here.

I develop parameters for this simulation using results reported in Kearney and Levine's (2009) excellent paper, in which the authors estimate the impact on women's contraceptive use of previous income-eligibility expansions in waiver states. I use their estimates of the effects of these expansions to parameterize this simulation because: a) according to my calculations, the population-weighted average income-eligibility threshold for Medicaid family-planning services in waiver states is about 190 percent of poverty; and b) I also calculate that the population-weighted average income-eligibility threshold for Medicaid pregnancy care in non-waiver states is about 195 percent of poverty.[33] Given the similarity between these thresholds, I assume that the effects of waiver implementation in waiver states provide a reasonably good indication of what might occur in non-waiver states if they were to avail themselves of the new family-planning option.[34]

---

[32] Specifically, the income-eligibility thresholds for Medicaid pregnancy care and for Medicaid family-planning services in Iowa are 300 percent and 200 percent of poverty, respectively; the corresponding thresholds in Minnesota are 275 percent and 200 percent of poverty; the corresponding thresholds in New Mexico are 235 percent and 185 percent of poverty; the corresponding thresholds in Virginia are 200 percent and 133 percent of poverty; and the corresponding thresholds in Wisconsin are 300 percent and 200 percent of poverty.

[33] On the income-eligibility thresholds for Medicaid family-planning services in waiver states, see Guttmacher Institute (2010a). The above-referenced estimate of the average income-eligibility threshold for Medicaid pregnancy care services across non-waiver states was calculated using data reported in Kaiser Family Foundation (2010). The population weights for these calculations were developed using data from United States Census Bureau (2010).

[34] There is a provision in the new health-care law that could cause the impact of the expansions that I simulate (hereafter, "the simulated expansions") to be somewhat smaller after a period of a few years than is suggested by Kearney and Levine's estimates. Under the new law, states will be required to extend Medicaid coverage to adults with incomes under 133 percent of the poverty line starting in 2014 (hereafter, "the 133-percent expansion"; sources: National Campaign to Prevent Teen and Unplanned Pregnancy, 2010 and United States Congress, 2010). About half of states cover some (but not all) childless adults, and, although all states cover at least some low-income parents, the income-eligibility thresholds for such coverage are below 133 percent of poverty in the majority of states (Kaiser Family Foundation, 2009). Thus, the baseline conditions that prevailed

Among the outcomes that Kearney and Levine consider in their paper, the ones that are most relevant for this discussion are: a) the probability of having failed to use contraception at last intercourse (which is estimated only among women who reported having had sex in the previous three months); b) the incidence of birth; and c) the incidence of abortion. The authors' analyses of the first of these three outcomes were conducted using individual-level data on women in waiver and non-waiver states. For their analyses of the other two outcomes, they use state-level data. In most instances, the authors conduct separate analyses for teens and non-teens, although, for their analyses of the incidence of abortion, they estimate models for teens and for all women. They do not conduct a separate abortion analysis for non-teens because of limitations in the data available to them.

For their analyses of the effects of income-based waivers on contraceptive use among teens, the authors adopt a difference-in-differences strategy in which they identify the effects of the policy by comparing changes in contraceptive use over time between women who do and do not live in waiver states. For their analysis of the effects of these waivers on contraceptive use among non-teens, the authors adopt a triple-difference strategy in which they identify the

in the non-waiver states that essentially formed Kearney and Levine's control group will be somewhat different once the 133-percent expansion is in place. Specifically, there are some women who – absent the 133-percent expansion – would be made eligible for family-planning services under the simulated expansions, but who – once the 133-percent expansion is implemented – will be eligible for these services whether or not the simulated expansions take place, since family planning constitutes a core benefit for Medicaid beneficiaries. I am unaware of any published data that would allow me to estimate precisely the number of women who could eventually be made "redundantly eligible" by the simulated expansions and the 133-percent expansion. However, Dubay et al. (2009) and Cook et al. (2009), in anticipation of its inclusion in the final health-care reform bill, estimate the number of uninsured adults who would be eligible for public insurance under the 133-percent expansion. Based the results of a series of back-of-the-envelope calculations using data reported by Dubay and her coauthors, by Cook and her coauthors, and by Kearney and Levine, I conclude that, depending upon the extent to which the 133-percent expansion crowds out private insurance, between about 25 percent and about 50 percent of the women who would be made eligible for family-planning services by simulated expansions will eventually be made eligible by the 133-percent expansion, even if the simulated expansions do not occur. Thus, my estimates of the simulated expansion's impacts on the national incidence of pregnancy, birth, and abortion should be considered to be reflective of the expected effects of the simulated expansions only over the next few years. The effects of such expansions in later years might be 25 percent to 50 percent smaller than is suggested by my results. On the other hand, my estimates of the simulated expansions' benefit-cost ratios (and of its costs per birth and pregnancy prevented) would likely be unaffected by this consideration, since – as is discussed in a subsequent subsection – I make the simplifying assumption that there is a constant marginal cost for each additional woman served by the program. To summarize, then: a) the results reported here reflect the simulated expansions' projected effects on the number of pregnancies and various pregnancy outcomes in the next few years; b) the simulated expansions' effects on these outcomes in later years might be between 50 percent and 75 percent of the magnitude suggested by my results; and c) the benefit-cost ratios and cost-effectiveness estimates reported here can be considered to reflect my projections of the simulated expansions' estimated effects both over the next few years and in later years.

effect of the policy by comparing changes over time in women's behavior on the basis of whether they live in waiver states and on the basis of whether their incomes make them newly eligible to take up these services. The authors do not adopt a triple-difference strategy for teens due to concerns about their ability to measure income accurately for this group. For their analyses of the waivers' effects on birth and abortion rates, the authors adopt a difference-in-differences strategy in which they identify the effects of the policy by comparing changes in these outcomes over time between states that did and did not implement waivers.[35] For these latter analyses, the authors present results from three different specifications. In one specification, they do not control for state-level trends; in a second specification, they include linear controls for state trends; and, in a third analysis, they include linear and quadratic controls for such trends. In the discussion below, I focus on the results for the third of these specifications.

Table 4 summarizes the key results from Kearney and Levine's analyses. Bolded results are statistically significant at or beyond the .05 level.[36] With respect to contraceptive use, the authors find income-based waivers reduced the probability of having failed to use contraception at last sex among non-teens by a little more than five percent, and that the relationship between waiver implementation and contraceptive use among teens was not statistically significant. Regarding the latter finding, I would note that the relevant coefficient in the contraceptive-use regressions for teens is (barely) correctly-signed, and that its standard error is quite large. Thus, I interpret these results as suggesting not that the waivers had no effect on teenage contraceptive use, but instead that the authors' estimates are not precise enough to allow them to identify any such effect that the waivers might have

---

[35] The authors also present results from a set of simple difference-in-differences analyses of the waivers' effects on contraceptive use and sexual activity among non-teens. The results of these analyses are qualitatively similar the results of their triple-difference analyses. I focus here on the results from the authors' triple-difference analyses, since they were produced using what I consider to be a superior identification strategy.

[36] The authors also estimate the effects of both income-based and duration waivers on sexual activity. They find that neither type of waiver had a significant effect on the sexual activity of non-teens, but their results do suggest that there was a significant and negative relationship between the implementation of both income and duration waivers and sexual activity among teens. However, they write that they are skeptical of this finding because it seems implausible that the implementation of these waivers could have caused teenagers to have less sex, and because their analyses that produced these results do not include a within-state control group (i.e., the authors do not compare the sexual behavior of teenaged women who are and are not newly eligible to take up subsidized family-planning services in states that implemented waivers). I find these arguments to be convincing, and I therefore assume, in the simulation of expanded eligibility for family-planning services, that this expansion does not affect the sexual activity of either teens or non-teens.

had. Moreover, as will be discussed in more detail below, the authors find that income-based waivers did in fact have a statistically-significant effect on teen birth rates in states in which they were implemented. For the purposes of parameterizing the simulation, I therefore interpret the authors' overall results as suggesting that income-based waivers increased contraceptive use among teens as well as non-teens. More specifically, I assume that implementation of these waivers caused five percent fewer sexually-active teenaged and non-teenaged women to fail to use contraception at a given sexual encounter.[37]

The bottom two rows of the table show Kearney and Levine's estimates of the waivers' effects on the incidence of birth and abortion in states within which they were implemented. The authors find that income-based waivers had a significant effect on the birth rates of teens and non-teens but were not significantly related to abortion rates for either group. However, the standard errors for the abortion estimates shown below are both large. The authors therefore write of their results that "this is not conclusive evidence that family planning waivers have little or no effect on abortions, but rather it indicates that we are unable to find any evidence in support of such an effect."[38] Given the strong evidence presented by the authors suggesting that income waivers affected birth rates, I assume that they also affected abortion rates. As is discussed in a subsequent subsection, I use the authors' results regarding the effects of the income waivers on birth rates, in particular, to develop parameters for an alternative specification for this simulation.

---

[37] The 95 percent confidence interval for the authors' estimate of the effect of income-based waivers on teenage contraceptive use spans from -.254 to .258. Thus, my assumption that the policy caused five percent fewer teens to fail to use contraception is consistent with a point estimate that is reasonably close to the middle of the confidence interval that they report.

[38] Kearney and Levine (2009), p. 143.

Table 4. Estimated Effects of Medicaid Family Planning Income-Based Waivers Among Teenaged and Non-Teenaged Women as Reported in Kearney and Levine (2009)*

	Effects Among Teens	Effects Among Non-Teens or Among All Women
Estimated Effect on the Probability of Not Having Used Birth Control at Last Intercourse	.002 (.128)	*Among non-teens:* **-.053 (.022)**
Estimated Effect on The Number of Births (Expressed in Percentage Terms)	**-.042 (.014)**	*Among non-teens:* **-.020 (.009)**
Estimated Effect on The Number of Abortions (Expressed in Percentage Terms)	.175 (.159)	*Among all women:* .075 (.087)

 * The first of the two estimates reported in each cell are the coefficients from the authors' OLS regression models. The quantities included underneath them in parentheses are the corresponding standard errors. Bolded estimates are statistically significant at or beyond the .05 level. Where the authors present results for the same outcome from multiple analyses, the results reported here are taken from their preferred specifications.

Medicaid Family Planning Expansion: Simulation Parameters for Initial Specification

As is stated above, I interpret Kearney and Levine's results as suggesting that the implementation of income-based waivers caused five percent fewer sexually-active teenaged and non-teenaged women in waiver states to fail to use contraception at a given sexual encounter. The authors define a woman as being sexually active if she has had intercourse during the previous three months. I therefore simulate changes in contraceptive use only among women who have sex at least once every three months, on average. I make the simplifying assumption that all new contraceptors within the simulation use oral contraception. Thus, I model the effects of this policy by switching about five percent of sexually-active women in the simulation population from being non-contraceptors to being pill users. I do not model a comparable change for women who have sex less frequently for three reasons. First, Kearney and Levine's results do not provide information on any effects that the waivers might have had on women who have sex very infrequently. Second, it seems likely that this policy change would have a smaller effect on the behavior of women who only rarely have sex than on women who have sex relatively often. And third, modeling an increase in contraceptive use among women who have sex fewer than four times per year would likely have little effect on the results of the simulation.

I assume that that the simulated policy change would induce an increase in contraceptive use within all non-waiver states. So, if, say, 40 percent of all births are estimated to occur in these states, and if the expansion in income eligibility is estimated to have reduced the number of births by, say, five percent in these states, then the policy would be estimated to have reduced births *nationally* by .(4*.05) = two percent. Because the average eligibility threshold for Medicaid family-planning services in non-waiver states would be about 200 percent of poverty under the assumptions for this simulation – and since, according to my tabulations of 2002 NSFG data, a little more than 70 percent of women who use publicly-subsidized contraception are unmarried – I concentrate all of the simulated increase in contraceptive use among non-contraceptors who are estimated to be under 200 percent of the poverty line, and I model these changes in such a way as to ensure that about 70 percent of the resulting increase in contraceptive use occurs among women who are unmarried.[39] In subsequent subsections, I present estimates of the share of pregnancies and various pregnancy outcomes that occur in non-waiver states, and I discuss the method by which each agent in the simulation is assigned an income-to-needs status relative to twice the poverty line.

Medicaid Family Planning Expansion: Simulation Parameters for Alternate Specification
As is discussed above, Kearney and Levine (2009) find that the implementation of income-based waivers induced five percent fewer sexually-active women to use contraception at last sex and produced a two percent reduction in the non-teen birth rate in states in which those waivers were implemented. However, when I model a five percent reduction in the number of women who use contraception at a given act of intercourse, the resulting reduction in childbearing among non-teenaged women within the simulation is about eleven percent in the simulation that uses what I call my "baseline pregnancy-outcome assumptions" and about seven percent in the simulation that uses what I call my "unintended-pregnancy-

---

[39] Regarding the marriage rate among users of publicly-subsidized contraception, I find that 73 percent of women using subsidized contraception – and that 71 percent of women using contraceptive service subsidized by Medicaid in particular – are unmarried. These tabulations exclude subsidized tubal ligations. If tubal ligations are included in one's tabulations, the equivalent rates of non-marriage for these two groups are 72 percent and 79 percent, respectively. To be clear about my treatment of income-to-needs status within this simulation: I assume that the policy change induces a five percent reduction in the number of *all* sexually-active women who fail to use contraception at a given sexual encounter, and I assume that all of the women who newly use contraception fall below 200 percent of poverty.

outcome assumptions" (see the last section of this report for a description of the differences between these two versions of the simulations). The simulated impact of an equivalent change in contraceptive use among teens on teen births is also larger than the corresponding estimates reported by Kearney and Levine, but to a lesser extent.

One might interpret these results as suggesting that the authors' findings regarding the waivers' effects on contraceptive use and on childbearing are incompatible. However, an inspection of the results reported in Table 4 shows that there is considerable uncertainty around the reported estimates of the effects of the policy on both outcomes. For example, the 95 percent confidence interval around the authors' point estimate of the effect of the implementation of income-based waivers on contraceptive use extends from about .9 percent to about 9.7 percent, and the confidence interval around their estimate of these waivers' effect on the non-teen birth rate extends from about .2 percent to about 3.8 percent. There are a range of values in the former confidence interval that produce reductions in childbearing within the simulation that are consistent with a range of values contained in latter interval. I therefore conduct a second simulation in which I assume that income-based waivers reduce the number of sexually-active women who fail to use contraception by about 2.5 percent, which produces a 3.65 percent reduction in simulated non-teen childbearing in the version of the simulation that uses my unintended-pregnancy-outcomes assumptions.[40] Both of these estimates are contained within the relevant confidence intervals from Kearney and Levine's paper. The alternative simulation's estimates of the policy's effect on teen childbearing and contraceptive use, and on teen and non-teen abortions, are also contained within the corresponding confidence intervals reported by Kearney and Levine.

I also conduct a third simulation (results not reported here) in which I assume that the policy induces a reduction of a little less than two percent in the number of women who fail to use

---

[40] One might alternatively think of this specification as reflecting an assumption that non-waiver states would, if they were to relax their eligibility criteria, not be as generous as the "first-mover" states that applied for (and were granted) income-based waivers. Indeed, there is already some variation across states in the generosity of their Medicaid family-planning programs. For example, a Kaiser Family Foundation (2007) report indicates that a minority of state Medicaid programs (less than one fifth) cover no over-the-counter birth-control methods. Thus, it is possible that first-mover states might, on the margins, have been willing to spend greater proportions of their budgets on family-planning services than would other states.

contraception at a given act of intercourse. This estimate is also contained within the relevant confidence interval from Kearney and Levine's paper, and it produces a reduction of about two percent in simulated non-teen childbearing. Thus, there is in fact an estimate of the policy's effect on contraceptive use that: a) is in the bottom portion of the corresponding confidence interval reported in Kearney and Levine's paper, and b) produces an effect on childbearing within the simulation that is nearly identical to the point estimate that the authors report for the same outcome. However, given that this estimate of the policy's effect on contraceptive use is considerably lower than the equivalent point estimate reported by Kearney and Levine, I opt instead to report in Thomas (2010) my results for the "middle" option, in which I assume that the policy reduces the number of women who fail to use contraception at a given act of intercourse by about 2.5 percent. Since this simulation produces an effect on childbearing that is closer to the estimates reported by Kearney and Levine, I consider it to be the preferred specification for these simulations.

Medicaid Family Planning Expansion: Comparison of Simulation Estimates to the Results of Kearney and Levine (2009) and Frost et al. (2006)

I now summarize differences between the results of the simulations described here and the results reported by Kearney and Levine and by Frost et al. (2006) in terms of the implied share of births prevented by expanded income-eligibility for Medicaid family-planning services. The latter paper presents results from a series of simulations that estimate the impact of expansions in eligibility for Medicaid family-planning services and therefore serves as a useful point of reference against which to compare my findings.[41] Frost and her coauthors estimate the impact of a variety of different Medicaid expansions, and, in what they call their "scenario pregnancy care" option, they assume that states would be required to set the income level that they use to determine eligibility for Medicaid-subsidized family-planning services equal to the income level that they use to determine eligibility for Medicaid pregnancy care. Since this is essentially the option given to states under the new law, and since it is also essentially the policy that I model in my own simulations, I use the authors' results from that particular analysis for the purposes of these comparisons. As is discussed above, Kearney and Levine present separate estimates of the effects of income-based

---

[41] Frost and her coauthors do not simulate the effects of Medicaid expansions on contraceptive use. Thus, it is not possible for me to use their results to parameterize an alternative specification of this simulation.

waivers on births among teens and non-teens. In order to achieve comparability across estimates, I present here a weighted average of the two, which is to say that I use the authors' results to estimate the impact of income-based waivers on teen and non-teen births combined.

The results reported in Thomas (2010) express estimates of this policy's effects in terms of the share of all births nationally that would be prevented if all non-waiver states were to take up the state family-planning option. However, in order to achieve comparability between the results presented in that study, in the Frost et al. report, and in the Kearney and Levine paper, I present estimates here of the impact of these waivers on the incidence of birth only in non-waiver states. Thomas (2010) reports results for four different Medicaid simulations: the version of the initial specification that uses FamilyScape's original pregnancy-outcome assumptions, the version of the initial specification that uses unintended-pregnancy-outcome assumptions, the version of the alternative specification that uses the original pregnancy-outcome assumptions, and the version of the alternative specification that uses unintended-pregnancy assumptions. In these four specifications, the simulation results imply that, in states in which income-eligibility expansions are implemented, they would prevent 9.8 percent of births, 7.3 percent of births, 4.9 percent of births, and 3.6 percent of births, respectively.

According to my analysis of Kearney and Levine's results, their estimates suggest that the policy would prevent 2.2 percent of the births that occur in non-waiver states, and, according to my analysis of the results reported by Frost et al., their estimates suggest that an equivalent policy would prevent 8.4 percent of the births occurring in the same states.[42] For the most part, then, my estimates fall in between those of Kearney and Levine and Frost et

---

[42] I make this calculation for Frost et al.'s results using estimates reported in their Table 3.5, which presents state-by-state estimates of the policy's impact on the number of births under their "scenario pregnancy care" assumption. I tabulate the total number of births prevented in states that have not yet implemented income waivers, and I calculate the ratio of the number of prevented births to the number of all births in these states using state-level birth data from the NVSS system. In order to calculate an equivalent quantity using Kearney and Levine's results, I again use NVSS data to estimate the total number of births that occur to teens and non-teens in the same states. I apply the authors' estimates of the effect of the policy on teen and non-teen births to age-specific tabulations of births in these states to calculate the total number of births that would be prevented, and I then calculate the ratio of the number of prevented births to the total number of births that occur in these states.

al. For example, the estimate from my preferred specification of this simulation (i.e., the version of the alternative specification that uses unintended-pregnancy-outcome assumptions) is a little more than 60 percent above the corresponding estimate reported by Kearney and Levine and a little less than 60 percent below the corresponding estimate reported by Frost et al.

Medicaid Family Planning Expansion: Estimating the Share of Pregnancies and Various Pregnancy Outcomes that Occur in Non-Waiver States

I adjust my initial estimates of the effect of expanding Medicaid family-planning services to reflect the fact that such expansions have already taken place in 21 states. I make these adjustments using estimates of the share of pregnancies and various pregnancy outcomes that occur in non-waiver states. Table 5 reports my estimates of these quantities.

Table 5. Estimated Share of All Pregnancies and Selected Pregnancy Outcomes Occurring in Non-Waiver States*

	Pregnancies	*Births*	*Out-of-Wedlock Births*	*Abortions*
Among all women	37.3%	37.4%	37.2%	36.8%
Among teenaged women	36.4%	36.6%	37.1%	35.9%

* Estimates of the share of births that occur in non-waiver states are calculated using data from the National Vital Statistics System (NVSS); estimates of the share of abortions occurring in non-waiver states are calculated using data from Guttmacher Institute (2010b) and Henshaw and Kost (2008); and estimates of the share of pregnancies occurring in these states are based on the aforementioned, state-specific data on the incidence of abortion and childbearing in the same states.

As an example of the way in which these adjustments were made: I estimate that 37.4 percent of all births occur in non-waiver states. For the initial specification of this simulation, my results suggest that the implementation of income-based waivers reduces the number of births by 9.8 percent in the states in which they are implemented. I therefore estimate that the policy would reduce the number of births nationally by (.374*.098) ≈ 3.7 percent, and I make comparable calculations for the other outcomes for which results are presented in Thomas (2010). I report national-level results for this simulation in that table so that they will be comparable to the estimated effects of other simulated policies.

Medicaid Family Planning Expansion: Imputing Income Eligibility

For reasons that are discussed in a previous subsection, I assume that the expansion in Medicaid family-planning services would only affect women who are below 200 percent of the poverty line. I therefore concentrate all of the increase in contraceptive use for these simulations among women who are estimated to fall below this threshold. I discuss here my method for imputing women's income-to-needs status. In Monea and Thomas (2010), my coauthor and I detail our method for imputing income-to-needs status to pregnant women. As is discussed in that report, we conduct those imputations for the purpose of determining whether pregnant women and their children qualify for publicly-subsidized benefits and services such as Medicaid-subsidized care for pregnant women and infants and a range of other means-tested benefits provided to young children. For the Medicaid simulation, it is necessary that I identify women who are assumed to fall below 200 percent of the poverty line in order to determine eligibility for Medicaid family-planning benefits. The imputation process described in Monea and Thomas (2010) relies on the results of analyses of real-world data on the income-to-needs statuses of pregnant women. For this simulation, however, I use data instead for all women regardless of whether or not they are pregnant, given that the availability of Medicaid family-planning services is not conditioned on pregnancy status. I use CPS data for these imputations, since the CPS is widely accepted as the most reliable source of information on individuals' income-to-needs statuses; and I use 2002 data in particular, since I use data from the NSFG from the same year to impute income-to-needs statuses to pregnant women and since FamilyScape is largely parameterized using data from that year.

Other than the fact that I use a different data source and a different subsample for these imputations, I adopt an approach that is almost identical to the one described in Monea and Thomas (2010). First, I use CPS data to conduct separate OLS regressions for married and unmarried women in which the dependent variable is set equal to one if a woman is below 200 percent of the federal line and zero if she is not. The independent variables included in these regressions control for age, race, and educational attainment.[43] These variables are

---

[43] In the income-to-needs regressions among pregnant women that are described in Monea and Thomas (2010), a control variable is also included to account for socioeconomic status (SES). However, FamilyScape's

coded in a manner that is consistent with the coding of the variables used to parameterize FamilyScape. The results of these regressions are reported in Table 6. These coefficients are used to assign to every woman in the simulation population a probability of falling below twice the poverty line. For each woman, the results of a unique random draw from a uniform (0,1) distribution are then compared to her assigned probability in order to impute a binary income-to-needs status for her.

The results reported in Table 19 in Monea and Thomas (2010) show that a larger share of pregnant women than of all women fall below 200 percent of the poverty line. This is especially true for unmarried women. I therefore make the simplifying assumption for the purposes of this simulation that any woman who is imputed to be below 200 percent of the poverty line before becoming pregnant will remain below this threshold if and when she becomes pregnant. Thus, I assume that all pregnant women who were imputed to be below 200 percent of the poverty line before becoming pregnant will qualify for publicly-subsidized benefits and services. In other words, I assume that all pregnancies prevented by the expanded provision of subsidized contraception through Medicaid would, had those pregnancies occurred, have qualified to be publicly subsidized. See Monea and Thomas (2010) for further discussion of the methods that were used to estimate the public cost of subsidized pregnancies and of the way in which income-to-needs statuses are imputed to pregnant women.

---

measure of socioeconomic status is determined by maternal educational attainment, and the CPS does not contain information on this characteristic (the regressions for pregnant women were conducted using the NSFG, which does contain such data). Thus, these regressions do not control for SES.

Table 6. Regressions Results
For Imputing Income Below 200 Percent of Poverty
Among Women Aged 15 – 44 in the 2002 CPS

	Unmarried Women	*Married Women*
Age: 20-24	0.2561***	-0.0297
Age: 25-29	0.2205***	-0.1029**
Age: 30-44	0.2225***	-0.1532***
Education: More Than High School	-0.2892***	-0.4138***
Education: High School	-0.0714***	-0.2474***
Race: Black	0.0965***	0.0754***
Race: Hispanic	0.0977***	0.1133***
Race: White	-0.0998***	-0.0711***
Constant	0.4174***	0.7209***
Mean of Predicted Values	0.411	0.251
P-Value for Joint Test of all Coefficients	0.000	0.000
Adjusted R^2	0.1058	0.1842
N (unweighted)	23,508	24,994
N (weighted)	31,486,654	30,382,384

Note: One asterisk (*) indicates that the parameter estimate is statistically significant at or beyond the .1 level, two asterisks (**) indicate that the estimate is significant at or beyond the .05 level, and three asterisks (***) indicate that the estimate is significant at or beyond the .01 level. The reference categories for the age, education, and race covariates are, respectively, teens aged 15 – 19, individuals with less than a high-school education, and individuals whose race categories are coded as "other."

Medicaid Family Planning Expansion: Estimated Cost

Kearney and Levine use data from reports published by the Guttmacher Institute to estimate that the average annual cost per woman served of publicly subsidizing family-planning services is $188. The authors note that this estimate reflects the average cost per woman currently served, rather than the marginal cost of serving a new Medicaid client. They make the simplifying assumption that there is a constant marginal cost for providing family-planning services to new clients. I make the same assumption here. Kearney and Levine also estimate that, for every 1,000 women of childbearing age in waiver states, 54 more obtain family-planning services through Medicaid as a result of the implementation of income-based waivers. I use these two estimates in conjunction with my tabulation of the number of women of childbearing age living in non-waiver states to estimate that the total cost of the expansion would be about $235 million.[44]

I assume that the cost of the program would be the same under the initial and alternative specifications for this simulation – which is to say that, when I assume under the alternative specification that the program would have a smaller impact on contraceptive use, I do not assume that this smaller impact would be a function of more limited participation in the program. Instead, I assume that the expansion of Medicaid services would crowd out privately-subsidized insurance to a greater extent in the alternative specification than in the initial specification of the simulation. One might therefore consider the amount of crowd-out associated with previous Medicaid expansions in order to gauge the plausibility of these two specifications' assumptions regarding the take-up of – and the behavioral changes induced by – expanded Medicaid family-planning services.

In their landmark work on this topic, Cutler and Gruber (1996) estimate the extent of crowd-out associated with earlier expansions in Medicaid eligibility for low-income pregnant women and children. They find that 49 percent of the increase in Medicaid coverage

---

[44] Specifically, I use data from the 2008 American Community Survey to estimate that there are about 23.3 million women of childbearing age living in non-waiver states (like Kearney and Levine, I define a woman as being of childbearing age if she is between the ages of 15 and 44). Under the assumption that 5.4 percent of these women would take up Medicaid family-planning services as a result of this expansion and that the average annual cost of the expansion per woman served is $188, I estimate the total annual cost of the program to be ($188*.054*23.3 million) ≈ $236.5 million. I round this figure down to $235 million for ease of exposition.

produced by these expansions crowded out private coverage. I can not calculate with any precision the amount of crowd-out that is implied in my simulations, since: a) state Medicaid programs provide a variety of family-planning services other than publicly-subsidized contraception, such as pregnancy tests, testing for and treatment of STDs, pap smears, and, in certain instances, abortions; and b) the only two pieces of relevant information included in this simulation relate to the share of all women who use family-planning services and the share of sexually-active women who use contraception as a result of the Medicaid expansions.[45] In order to produce a reliable estimate of the amount of crowd-out associated with these expansions, I would need to have information on the number of women who use publicly-subsidized services other than those that involve the provision of contraception as a result of these expansions. Since I do not have this information, I instead use the data described above to calculate an upper-bound estimate of the implied amount of crowd-out associated with the expansions. To the extent that these expansions induced some women to claim family-planning benefits other than subsidized contraception – and they undoubtedly did have such an effect, even if I am unable to measure its magnitude – the true extent of crowd-out is lower than is implied by these estimates.

The results of my simulations suggest that about two-thirds of women of childbearing age have intercourse at least once over any given three-month period. Based on other findings reported by Kearney and Levine, I estimate that this level of sexual activity is roughly comparable to the equivalent quantity in their data.[46] Recall that I assume that 5.4 percent of women of childbearing age take up Medicaid family-planning services in states that implement income-based waivers as a result of their implementation. Recall also that I assume in the initial specification of this simulation that about five percent fewer sexually-

---

[45] For examples of the range of services provided under states' Medicaid family-planning programs, see Florida Department of Health (2010) and New York State Department of Health (2010).

[46] FamilyScape's simulation population contains women who are between the ages of 15 and 44, and Kearney and Levine's analyses are limited to women of the same age. In analyses whose results I do not discuss here, the authors find that the implementation of income-based waivers reduced the number of women who had unprotected sex in the prior three months by 3.3 percent. Because FamilyScape produces generally-realistic rates of sexual activity, the effect of modeling an increase in contraceptive use that is consistent with Kearney and Levine's findings produces a similar effect on the frequency of unprotected sex within the simulation. Specifically, simulating a five percent reduction in the number of sexually-active women who fail to use contraception at a given act of intercourse produces a 3.2 percent reduction in the number of women who have unprotected sex during any given three-month period. I take the comparability of these results to suggest that the rate of short-term sexual activity within the simulation is roughly comparable to the equivalent quantity in Kearney and Levine's data.

active women in these states fail to use contraception as a result of these waivers'
implementation. I therefore calculate the ratio of new contraceptors to new claimants of
family-planning services to be ((.05*.67)/.054) ≈ 62 percent. Under the (heroic) assumption
that none of the increase in take-up of these services involves use of any publicly-subsidized
benefits other than contraception, I therefore calculate the implied amount of crowd-out
within this specification of the simulation to be about (1 – .62) = 38 percent. And, for the
alternate specification – in which I assume that about 2.5 percent of sexually-active women
use contraception as a result of the simulation – I calculate that the implied amount of
crowd-out within the simulation is (1 – ((.025*.67)/.054)) ≈ 69 percent. Given that some
women claim Medicaid family-planning benefits other than contraceptive services, the
implied amount of crowd-out within the simulations is actually less than the amount
suggested by the estimates cited here.

To summarize: I assume for both specifications of the simulation that the expansion in
eligibility for Medicaid family-planning services would cost $235 million. The back-of-the-
envelope calculations described above suggest that the amount of crowd-out that is implied
in the initial specification by the estimates used to make this cost calculation is lower than
the amount of crowd-out suggested by Cutler and Gruber's findings, and that the implied
amount of crowd-out implied in the alternate specification may or may not be lower than the
amount suggested by Cutler and Gruber, depending on the number of women who claim
family-planning services other than those involving the provision of publicly-subsidized
contraception.[47]

---

[47] Recall that Kearney and Levine express their estimates of the effect of the waivers on contraceptive use in
terms of the probability that a sexually-active woman used contraception at last intercourse. For practical
reasons, I incorporate this finding into the simulations by making the simplifying assumption that, as a result of
the implementation of these waivers, a certain share of sexually-active women newly use contraception every
time that they have sex. In other words, I assume within the simulations that the share of women who are
induced to use contraception at their most recent sexual encounter is equivalent to the share of women who are
induced to use contraception at all, and that these new contraceptors use contraception at every act of
intercourse Relaxing this assumption – by assuming, say, that the number of women who are induced to use
contraception at all is twice as large as the number of women who used contraception at last sex, and that each
of these new contraceptors uses contraception in half of her sexual encounters – should have little effect on the
outcome of the simulation, since, under either assumption, the same share of all sexual encounters would
involve the use of contraception. However, this assumption *does* have important implications for the implied
amount of take-up, since, the greater the number women who use contraception as a result of the Medicaid
expansion, the lower the extent of crowd-out that is associated with the expansion. Thus, the crowd-out
estimates calculated above can also be thought of as upper bounds because they do not account for the

Mass Media Campaigns

I begin this section by describing my assumptions regarding the effects of a nationally-implemented mass media campaign encouraging contraceptive use. I then present estimates of the costs of such a campaign.

Mass Media Campaigns: Estimated Effects

In their widely-cited meta-analysis, Snyder et al. (2004) estimate the average effects a wide variety of mass media campaigns. As is discussed in Thomas (2010), the evaluations of these campaigns virtually never use random assignment.[48] Rather, the best evaluations tend to compare changes over time between a locality (or localities) in which a campaign was implemented and a demographically-similar locality (or localities) in which it was not. Thus, the results of these evaluations should not be considered to be as reliable as the estimates described elsewhere this report of the effects of teen pregnancy prevention programs that were evaluated using random assignment – or even of the estimates of the effects of expansions in eligibility for Medicaid family-planning services, some of which were calculated using a carefully-designed triple-difference strategy. However, since the results of the studies included in the Snyder et al. meta-analysis represent the best estimates to date of the effects of media campaigns, I use them for the purposes of parameterizing this simulation – although, as is discussed below, I also make adjustments to these estimates under the assumption that the studies in question may over-state the effects of the campaigns that they evaluate.

For all of the campaigns included in their study, Snyder and her coauthors report correlation coefficients reflecting the relationship between implementation of a given campaign and the share of the target population whose behavior is estimated to have been affected. These results are not directly interpretable in a way that would help me to develop parameters for this simulation. However, whenever possible, the authors also report estimates of campaigns' impacts in terms of the proportion of the target population whose members

---

likelihood that the induced increase in contraceptive use is distributed across more individual women than is implied by these calculations.

[48] For an overview of the methodological challenges inherent to the evaluation of media campaigns, see Noar (2009).

were induced to change their behavior. These estimates can be used to develop parameters for this simulation. Snyder and her coauthors find that, as a whole, the 48 studies included in their meta-analysis had an average effect size of .09. Among studies that reported estimates of the share of the population whose behavior was changed by these campaigns, they find that the relevant campaigns changed the behavior of eight percent of the members of their target populations. However, some of these campaigns – those that encourage the use of seat belts are the most prominent examples – were supported by a regime of legal enforcement. The authors calculate a separate set of average effects after excluding these enforcement campaigns from their analysis. They find that, depending on their specific characteristics, non-enforcement campaigns' average effect sizes were between .05 and .06, and that these campaigns changed the behavior of between three percent and five percent of the members of their target populations.

Among the campaigns included in the authors' analysis, four specifically encouraged the use of condoms during sex. The authors report that these sexually-oriented campaigns had an effect size of .04, and that the two campaigns reporting effects in percentage terms changed the behavior of about six percent of the members of their target populations.[49] This percentage-point estimate is roughly comparable to the results of a more recent study by Zimmerman et al. (2007). Zimmerman and his colleagues oversaw and evaluated a saturation television campaign encouraging condom use in Lexington, Kentucky. The authors compared the change in the frequency of condom use in Lexington before and after the campaign to the equivalent change during the same time period in Knoxville, Tennessee, which they took to be the study's control city. Their findings imply that the campaign increased condom use by somewhat more than six percent among members of the overall

---

[49] The evaluations of three of the four sexually-oriented campaigns included in Snyder et al.'s (2004) synthesis adopted a pre-intervention/post-intervention, treatment-community/control-community research design (see CDC AIDS Community Demonstration Projects Research Group, 1999; Keegles et al., 1996; and Santelli et al., 1995). All three of these studies found that the campaigns in question increased condom use. The evaluation of the fourth campaign simply compared individuals who were and were not exposed to the treatment in question after it was implemented (see Snyder, 1991). This fourth study, whose evaluation design was considerably weaker than were the designs of the other three studies, found that the campaign had little or no effect on the behavior of the individuals who were exposed to the treatment. The finding from Snyder et al.'s meta-analysis that such campaigns can change individual behavior is thus a reflection of the fact that the results of the other three (stronger) evaluations outweighed the less-positive results of the weaker study.

target population.[50] I therefore interpret the relevant literature as suggesting that the average media campaign encouraging contraceptive use has the potential to change the behavior of about six percent of the members of its target population. Since (for some studies) all or (for other studies) almost all of the members of the target populations for these campaigns were unmarried, I assume that the simulated campaign would only affect the behavior of unmarried individuals.[51]

In fact, however, the true average effect of such campaigns is probably less than six percent, given that Snyder et al. only consider results reported in published articles, which are presumably less likely to have reported findings of no effect.[52]   Moreover, for reasons discussed above, it is difficult to know with certainty whether the relationships that these studies identify are altogether causal or are partly correlational in nature. I therefore adopt Noar's (2006) assumption that the true average effect of media campaigns may only be about half of the level reported by Snyder and her colleagues, which is to say that I assume that the sexually-oriented campaigns described above altered the behavior of about three percent of the members of their target populations.

---

[50] The authors' analysis indicated that the media campaign increased condom use by about 13 percent among the highest-risk half of the population and had no significant effect on the behavior of the lowest-risk half of the population. Thus, the average effect across these two groups was about 6.5 percent. The authors do not appear to have explored in great depth whether there were other factors that might have contributed to the differential changes in condom use that they observe over time between the two cities. Indeed, they present data suggesting that the aggregate trend in condom use before the start of the intervention was negative in Lexington but was positive in Knoxville. However, they do not discuss in detail the implications of these divergent trajectories for their results.
[51] Among the four sexually-related campaigns whose results are included in Snyder et al.'s analysis, one was implemented exclusively for gay men and therefore presumably had relatively little effect on married individuals (see Keegles et al., 1996); the evaluation of another campaign reports that the marriage rate among members of the evaluation sample was only slightly above ten percent (see Santelli et al., 1995); and the evaluations of the other two do not indicate whether the campaign in question focused primarily on unmarried individuals. One would assume, however, that these latter two campaigns focused disproportionately on the unmarried population, since both of them encouraged the use of condoms as a means of avoiding transmission of STDs (see Fishbein et al. 1996 and Snyder, 1991). Additionally, although the authors of the Lexington study do not present tabulations of the marriage rate within their sample, their evaluation appears to have focused primarily on unmarried men, since they report that the members of their sample were a little less than 22 years old, on average (see Zimmerman et al., 2007). Given that virtually all of these campaigns' messages focus on the importance of using condoms as a means of avoiding contraction of STDs, I assume that their estimated effects can not be taken to suggest what the impact might be of a campaign encouraging contraceptive use among married individuals, since married couples are substantially less likely than unmarried couples to use condoms as their chosen method of contraception, and since STD transmission is presumably much less of a concern for the latter group than for the former.
[52] Noar (2006).

Given that all of the results described above are specific to campaigns that encourage the use of condoms in particular as a means of avoiding the transmission of STDs, I assume that the simulated campaign would only affect the frequency of condom use. The use of condoms is much more common among young males than among older males. For example, in Thomas and Roessel (2008), my coauthor and I find that about three quarters of sexually-active, unmarried teenaged males report having used a condom at last sex, but that well below half of sexually-active, unmarried older males report having done so. Similarly, the results of the baseline FamilyScape simulations indicate that about 78 percent of sexually-active unmarried teenaged males tend to use condoms during intercourse, but that the same is true of only about 33 percent of older males. Thus, simulating a three-percent increase in the number of all teenaged and non-teenaged men who use condoms reduces the total number of non-condom users by about $(3/22) \approx 14$ percent among unmarried, sexually-active, teenaged males, but by only $(3/67) \approx$ five percent among unmarried, sexually-active, non-teenaged males. Moreover, since teenaged girls are less likely than non-teenaged women to rely on methods of contraception other than condoms, a proportionally-equivalent increase in condom use among teens and non-teens would have a larger impact on rates of pregnancy and childbearing in the former group than in the latter group.

For both of these reasons, simulating a change in the behavior of the same share of the teen and non-teen populations reduces rates of pregnancy and childbearing by almost three times as much among teens as among non-teens. It seems unlikely that a media campaign's effect on teens would be so disproportionate. I therefore assume that the simulated campaign induces an increase in condom use that is roughly half as large among teens (about 1.5 percent) as among non-teens (about 3 percent).[53] Even after I make this assumption, the campaign's impacts on simulated rates of pregnancy and childbearing among teens are about 40 percent larger than among non-teens.

---

[53] This assumption is also consistent with diffusion theory, which posits that, as a group's baseline rate of behavior approaches 100 percent, it becomes increasingly difficult to induce the remaining non-compliers to alter their behavior (Snyder et al., 2004).

Mass Media Campaigns: Estimated Costs

I assume that the simulated campaign would be implemented on a national scale. Not surprisingly, there is evidence that a campaign's persistence has implications for the durability of its effects. For example, Zimmerman et al.'s (2007) results suggest that the Lexington campaign's effects began to fade about three months after the cessation of public-service advertisements.[54] Thus, I assume that the simulated campaign must be ongoing in order to produce the effects described above on a persistent basis. I develop a set of assumptions regarding the cost of an ongoing, national-level media campaign using itemized data on the costs of similar, health-related campaigns that have been implemented in the relatively-recent past. Specifically, I use data on the American Legacy Foundation's *Truth Campaign*, the Centers for Disease Control's *VERB* campaign, the Office of National Drug Control Policy's *National Youth Anti-Drug Media Campaign* (NYADMC), and the Lexington condom campaign described in the previous subsection. I discuss the costs of each of these campaigns separately below.

The *Truth Campaign* is an anti-smoking campaign that is funded by United States tobacco companies under the terms of their 1998 settlement with state attorneys general. The plaintiffs brought this lawsuit in an attempt to recover costs to state Medicaid programs for treatment of tobacco-related conditions. The campaign relies heavily on the use of television advertising. Farrelly et al. (2005) find that the *Truth Campaign* has been associated with substantial declines in youth smoking, and Holtgrave et al. (2009) find that it produced positive results in a cost-effective manner. Holtgrave and his coauthors present itemized estimates of the cost of the campaign during the first three full years of its national implementation. According to my tabulations of the data presented in that paper, the campaign cost about $100 million annually during its first three years, and about 70 percent of these expenditures were devoted to media-related activities (most of the rest of the campaign's expenditures were spent on administrative costs, public relations, evaluation costs, and a variety of other miscellaneous expenses).

---

[54] Zimmerman et al. (2007).

VERB was a social-marketing campaign sponsored by the federal government that encouraged physical activity among pre-teens and young teens. The campaign relied primarily on the use of television advertisements aired between 2002 and 2006 on cable networks that are popular with children.[55] The campaign was found to have been successful in increasing physical activity among the members of its target population.[56] Krisberg (2005) reports estimates of annual total funding for *VERB* for the period from 2001 to 2005, and I calculate the average of these annual amounts to be about $68 million. A report from the United States Government Accountability Office (2006b) presents estimates of the campaign's total spending on media contracts for 2003 – 2005, and I use these data to calculate the average annual level of spending on this line item to be about $60 million. Thus, I estimate that media contracts comprise about (60/68) ≈ 88 percent of total spending for this campaign.

NYADMC is funded and administered by the federal government and is the largest anti-drug media campaign in United States history.[57] It was launched in 1998 and has historically relied heavily on the use of television advertising.[58] The program has generally been found to have been ineffective.[59] Nonetheless, there are reliable cost data for the program, so I include it in this discussion. Unlike the *Truth* and *VERB* campaigns, NYADMC's media expenditures were matched, which is to say that, in legislation funding for the campaign, Congress mandated that media organizations accepting purchases of air time from the campaign must match the campaign's spending with a certain amount of in-kind advertising time. Orwin et al. (2006) report estimates of the campaign's own spending and of the value of the matches provided by media outlets for the period from 1999 to 2004. Based on my tabulations of these data, the value of the match was equal to a little less than 30 percent of the total value of the campaign's media activities. I use data from that report to estimate that, during these years, the average annual value of the media component of the campaign (including the value of the media match) was about $200 million. I combine these estimates with information reported by the Government Accountability Office (2006a) on total annual

---

[55] Huhman et al. (2007), Centers for Disease Control and Prevention (2010).
[56] Huhman et al. (2007).
[57] Palmgreen et al. (2007).
[58] Ibid.
[59] For examples of studies finding that the campaign had few positive effects (or none), see Hornik et al. (2008) and Orwin et al. (2006).

appropriations for the campaign to estimate that the average annual value of the campaign (again, inclusive of the value of the match) was about $230 million. I thus calculate that about (200/230) ≈ 87 percent of the campaign's total value was devoted to media-related costs.

The design and estimated effects of the Lexington condom campaign are described briefly in the previous subsection. Seth Noar – who was a member of the team that designed, implemented, and evaluated the campaign – provided me with rough estimates of various components of the campaign's costs. These estimates suggest that the campaign spent a total of about $450,000 to purchase nine months of air time for advertisements.[60] I therefore estimate that the cost of twelve months of air time in the Lexington market would have been ($450,000/.75) = $600,000. The campaign was given a one-to-one match by local network affiliates. I therefore estimate the total annual value of air time for a year-round version of this campaign to be $1.2 million.

I calculate that the Lexington media market constitutes about one half of one percent of the size of the overall national media market.[61] I therefore assume that the television-advertising costs for a national media campaign whose intensity is equivalent to that of the Lexington campaign – and that receives no matching and is implemented on a continuous basis – would be about (200*1.2) = $240 million. Noar also told me that the Lexington campaign spent $100,000 to develop its ads. I assume that, if this campaign were taken to scale nationally, substantially more funding would be required to develop and produce its advertisements, since the campaign would presumably need to appeal to a broader audience. Thus, I assume that a national equivalent of the Lexington campaign would require $1 million to develop its ads.[62] I therefore assume that total media-related costs for such a campaign would be (240 + 1) = $241 million. Recall that I estimate that the *Truth*, *VERB*, and NYADMC campaigns spent about 70, 88, and 87 percent of their total budgets,

[60] The campaign that was evaluated for the analysis described above actually lasted for three months and cost about $150,000. However, Noar told me that the campaign's evaluators also aired two other three-month campaigns for other purposes, and that the air time for these campaigns also cost about $150,000 each.
[61] Based on my tabulations of data reported by Nielsen Media (2010).
[62] In fact, if I had instead assumed that the expense associated with developing these ads was higher or lower by an order of magnitude, my bottom-line conclusions about the cost of the simulated campaign would be the same, since, under any of these scenarios, the cost of airing the campaign overwhelms the cost of developing its advertisements.

respectively, on media-related costs. The average of these three estimates is about 82 percent. As such, I estimate that the total annual cost of a Lexington-style campaign implemented on an ongoing, national basis (and without matching) would be (241/.82) ≈ $295 million.

Table 7 summarizes my estimates of the average annual costs of each of these four campaigns under the assumption that they do not benefit from any matching. I also report estimates of each campaign's intensity, as reflected in its targeted gross-point rating (GRP) per week.[63] A GRP measures the sum of the ratings points per spot for a given television advertisement over a particular period of time. For example, if an advertisement airs twice in a given week, and if the times at which it airs have ratings of 2.0 and 2.5, then that advertisement's GRP for the week is 4.5. The table also indicates whether the bulk of the evaluations for each campaign deemed it to have been effective. I do not inflate the estimates below to $2008 because I only use them in a very qualitative sense to develop assumptions about the cost of a media campaign.

The *Truth* and *VERB* campaigns' estimated costs are quite a bit lower than are those of NYADMC and the Lexington campaign. This cost differential is mirrored by (and is almost certainly a function of) the higher intensity of the latter two campaigns. Given that these data paint two rather different portraits of the cost of such a campaign, I make two different assumptions for the policy simulations regarding the simulated campaign's cost. For the initial specification of the simulations, I assume that the campaign would cost $100 million annually and, in an alternative specification, I assume that it would cost $250 annually. I do not, however, assume that the campaign's effectiveness varies with cost, since – as can be seen below – relatively-more-expensive campaigns are not necessarily more effective. To summarize, then, I assume for both specifications that the campaign would have the effects described in the previous subsection, and I assume its annual costs to be $100 million and

---

[63] I calculated the estimate of the per-week GRP for the *Truth* campaign using data taken from Farrelly et al. (2005) on the campaign's exposure during a period beginning in February 2000 and ending in the second quarter of 2002. The equivalent estimate for the *VERB* campaign was taken from Huhman et al. (2007) and is an average of the authors' estimates of the GRPs per week for the campaign during the first two years of its implementation. The equivalent estimate for NYADMC was taken from Orwin et al. (2006) and is based on data from the 1999 – 2004 time period over which the cost estimates cited above were calculated. The equivalent estimate for the Lexington campaign was taken from a transcript of a forum in which Seth Noar stated that the campaign had a rating of more than 200 GRPs per week (see Kaiser Family Foundation, 2006).

$250 million in the initial and alternative specifications, respectively. One might think of the lesser of these two estimates as corresponding to the cost of a campaign that is able to secure matching from television outlets. On the other hand, the less-expensive campaigns described above are actually the ones whose media expenditures were not matched. Thus, I would simply note that the simulated campaign's costs could be lower than are assumed here to the extent that it is able to secure substantial matching from the stations on which its advertisements are aired.

Table 7. Estimated Costs, Intensity, and Effectiveness
of Four Mass Media Campaigns

Campaign	Estimated Cost *Expressed on an Annualized & Nationalized Basis*	Estimated Campaign Intensity *(GRPs per week)*	Campaign Estimated to be Effective?
Truth	$100 million	117	Yes
VERB	$60 million	147	Yes
NYAMC	$230 million	254	No
Lexington Condom Campaign	$295 million	> 200	Yes

Miscellaneous Issues Related to Simulation Implementation

In this final section of the report, I address a variety of over-arching technical issues, including: a) the time period for the data used to parameterize these simulations; b) the method by which simulation results were corrected to account for the fact that FamilyScape over- or under-simulates pregnancies for certain key demographic groups; c) notable differences across simulations in the way in which policies' behavioral effects are modeled; d) the way in which the alternative unintended-pregnancy simulations were specified; and e) the fact that, for some simulations, the behavioral change simulated differs modestly from the targets specified in earlier sections of this report.

Miscellaneous Issues: Implied Time Period for Policy Simulations

The majority of the baseline parameters for the FamilyScape model were developed using data for calendar year 2002. When data from that year were not available, information from the closest available year was used. Other aspects of the simulations were parameterized using data from several different years. For example, parameters were developed for the simulation of a mass media campaign using findings from evaluations of campaigns that were, for the most part, implemented during the 1980s and 1990s; the simulation of an evidence-based teen pregnancy prevention program was parameterized using findings from evaluations of campaigns that were implemented almost entirely during the 1990's; the Kearney and Levine paper whose results were used to develop parameters for the simulation of expanded Medicaid family-planning services utilized data on existing state-level expansions that were implemented during a period spanning from 1994 to 2006; estimates of the public cost of pregnancy were produced using data from a variety of sources that were primarily gathered between 2001 and 2004; and estimates of policies' costs and of the cost savings that they would produce are typically inflated to $2008 using the CPI-U-RS or the medical component of the CPI.[64]

---

[64] For more information on the data used to parameterize the baseline specification of the FamilyScape model, see Thomas and Monea (2009); for more information on the data used to parameterize the simulation of a mass media campaign, see CDC AIDS Community Demonstration Projects Research Group (1999), Fishbein et al. (1996), Keegles et al. (1996), Noar (2006), Santelli et al. (1995), Snyder (1991), and Snyder et al. (2004); for more information on the data used to parameterize the simulation of a teen pregnancy prevention program, see Coyle et al. (2001), Jemmott et al. (1992), Scher et al. (2006), Sikkema et al. (2005), St. Lawrence et al. (1995) and Villaruel et al. (2006); for more information on the data used to parameterize the simulation of an

Ideally, one would prefer to have developed all of these parameters using data from a common – and relatively-recent – year. For obvious practical reasons, however, this was not a realistic objective. Thus, I make the simplifying assumption that the data used to parameterize the simulations are recent enough to approximate currently-prevailing conditions.

Miscellaneous Issues: Demographic Corrections of Simulation Results

Each of the policy simulations is implemented for a subgroup of the overall simulation population that is defined based upon individuals' age and/or marital-status characteristics. Specifically, the simulation of a mass media campaign only affects the behavior of unmarried males, the simulation of a teen pregnancy prevention program only affects unmarried teens, and the simulation of an expansion in Medicaid-subsidized family planning services only affects unmarried females. In its baseline specification, FamilyScape produces rates of pregnancy, birth, abortion, and fetal loss for all of these subgroups that are at least modestly different from their real-world equivalents. For instance, the real-world pregnancy rate among unmarried women is about one percent higher than the simulated rate of pregnancy for the same group. For married couples, the model is less accurate: the real-world rate of pregnancy is about a third lower than the equivalent simulated rate. Within the married and unmarried populations, simulated incidences of pregnancy for teens and non-teens also differ to varying degrees from their real-world equivalents. Among unmarried teens, for example, the real-world pregnancy rate is about 44 percent higher than its simulated equivalent, and, among unmarried non-teens, the real-world rate is about eight percent lower than its simulated equivalent.[65]

Although FamilyScape is, by agent-based modeling standards, quite realistic, these discrepancies may nonetheless pose problems for the benefit-cost simulations for two reasons. First, they suggest the possibility that the model fails to account for some key

expansion in Medicaid family-planning services, see Kearney and Levine (2009); and, for more information on the data used to measure the public cost of pregnancy, see Amaral et al. (2007), Billen et al. (2007), Frost et al. (2006), Macomber et al. (2010), Machlin and Rhode (2007), Merrill and Steiner (2006), and Trussell et al. (2009).

[65] See Thomas and Monea (2009) for a detailed comparison of real-world and simulated rates of pregnancy, childbearing, abortion, and fetal loss for a variety of different demographic groups.

consideration(s) whose exclusion causes the policy simulations to produce distorted results. And second, the findings reported in Thomas (2010) might be misstated because: a) they are expressed in terms of policies' effects on the incidence of pregnancy and various pregnancy outcomes within the simulation population as a whole, and b) they are driven by changes within specific subgroups for which FamilyScape's baseline specification produces results that differ to varying degrees from their real-world equivalents.

Regarding the first of these two considerations, I have conducted sensitivity analyses in which the effects of various policies were re-estimated using alternative (but arguably-plausible) behavioral assumptions that produce baseline rates of simulated pregnancy and childbearing are more closely aligned to their real-world equivalents. The results from these alternative policy simulations were qualitatively similar to the ones that were produced under the model's original specification. In Thomas and Monea (2009), for example, my coauthor and I present results from two earlier versions of the simulation of an expansion in Medicaid family-planning services. One version was conducted using the simulation's baseline assumptions, and – given that FamilyScape over-simulates pregnancies among married couples – the alternative version was conducted under the assumption that married couples over-report their frequency of intercourse. The estimated effects of the policy in terms of the percent reduction in the incidence of pregnancy and childbearing were nearly identical under these differing assumptions. Thus, the results of the policy simulations are generally robust to the adoption of plausible alternatives to the model's baseline assumptions.

The second consideration is arguably more pressing: if a given group's contribution to the overall rate of pregnancy is different in the simulation than in the real world, the effect of a policy differentially affecting that group on the aggregate pregnancy rate could be misstated – and potentially substantially so. For the baseline simulation and for each policy simulation, I therefore separately correct the initially-simulated simulated pregnancy, birth, abortion, and fetal-loss rates for unmarried teens, married teens, unmarried non-teens, and married non-teens to ensure that they match their real-world equivalents. For example, I multiply the initially-simulated pregnancy rate for unmarried teens by $(1 + .44) = 1.44$, and I multiply the initially-simulated rate for unmarried non-teens by $(1 - .08) = 92$. I make similar corrections for married teen and non-teen pregnancy rates. I also make equivalent corrections for the

birth, abortion, and fetal-loss rates for each group, and I make such corrections for the results produced under the model's baseline assumptions and under each set of policy assumptions. Thus, all results reported in Thomas (2010) have been corrected in this fashion.

Miscellaneous Issues: Differences Across Policy Simulations in the Operationalization of Behavioral Effects
As is made clear throughout this report, the simulated effects of each policy on sexual activity and contraceptive are strongly rooted in evidence that is reported in the relevant research literature. The way in which these policies' estimated effects are expressed in the literature therefore helps to determine the manner in which the policy simulations are implemented. For instance, Kearney and Levine (2009) estimate that, as a result of the expansion in eligibility for Medicaid family planning services, about five percent fewer sexually-active women reported that they did not use contraception at last sex in the states in which these expansions occurred. In other words, the authors' estimates suggest that, for every 100 sexually-active women, about five fewer failed to use contraception at last sex as a result of the expansion in eligibility for Medicaid family-planning services. Note that this is quite different from saying that the policy induced a five percent reduction in the number of women who do not use contraception. Assume, for example, that the population in question consists of 1,000 sexually-active women, 500 of whom did not use contraception at last sex. A finding that five percent fewer sexually-active women failed to use contraception implies a behavioral change for (.05*1000) = 50 women, whereas a finding that there was a five percent reduction in the number of sexually-active women who failed to use contraception implies a behavioral change for (.05*500) = 25 women. Thus, it is critically important that one be clear about the population to which a given finding refers.

I describe the relevant population for the Medicaid expansion above. The mass media campaign is parameterized using the findings of studies that similarly express their estimates in terms of campaigns' impacts on the target population as a whole. Recall that I take the relevant literature to suggest that media campaigns encouraging healthy sexual practices alter the behavior of three percent of the members of the target population. I apply this finding within the mass media simulations by changing the contraceptive behavior of about three percent of unmarried males. In the teen pregnancy prevention literature, however, findings

tend to be expressed in terms the proportional increase in the number of individuals who engage in a particular behavior (e.g., who use contraception or who abstain from sex) as a result of the intervention. Thus, I model a 12.5 percent increase in contraceptive use for the teen pregnancy simulations by increasing the number of female pill users and male condom users in the target population by 12.5 percent each (rather than by changing the contraceptive behavior of 12.5 percent of all members of the target population, which would produce a notably larger effect). I model changes in sexual activity for these simulations in the same manner.

To summarize: the estimates used to parameterize the simulations of an expansion in Medicaid family planning services and of a mass media campaign reflect the assumed effects of these initiatives in terms of the share of the overall target population among whom the policy induces a behavioral change, whereas the estimates used to parameterize the simulation of a teen pregnancy prevention program reflect the proportional effects that the program is assumed to have on the number of individuals who engage in a particular behavior(s).

Miscellaneous Issues: Alternative Unintended-Pregnancy Simulation Specifications
FamilyScape does not explicitly account for the pregnancy intentions of members of the simulation population. It does, however, implicitly account for this characteristic to the extent that it is correlated with age, marital status, and the other demographic characteristics that are incorporated into the simulation. According to Finer and Henshaw (2006), unintended pregnancies are substantially less likely than intended pregnancies to result in live births, since unintended pregnancies often result in abortions. It seems reasonable to assume that pregnancies that are prevented by a given policy would have been unintended, had they occurred. One would therefore expect the ratio of prevented abortions to prevented births in the policy simulations to be higher than the ratio of abortions to births for all simulated pregnancies. Although the results of the initial specifications of the policy simulations are in fact consistent with this expectation, the share of prevented pregnancies that would have resulted in births is still higher than one would have expected based on the data described above. I therefore implement an alternative specification for each policy simulation in which I assume that the distribution of pregnancy outcomes that would have

obtained from policy-prevented pregnancies is the same as the equivalent distribution that is observed for real-world unintended pregnancies. I implement these alternative specifications using unpublished tabulations of data that were gathered by the Guttmacher Institute and were provided to the National Campaign to Prevent Teen and Unplanned Pregnancy.

Table 8 reports the results of these tabulations, which are disaggregated by age and marital status. I assume that policies' effects on the incidence of pregnancy are the same for these alternative specifications as for the corresponding base specifications. I then use the tabulations shown below to make back-of-the-envelope calculations of the distribution of outcomes that these prevented pregnancies would have produced, had they occurred. I make these calculations separately for unmarried teens, unmarried non-teens, and married individuals. (There were not enough married teens in Guttmacher's data to allow them to produce separate estimates of the distribution of pregnancy outcomes for this group. I therefore assume that the distributions of outcomes for married teens and married non-teens are identical.). The results of these alternative specifications are presented in Table 3 in Thomas (2010). A quick perusal of these results reveals: a) that the effects of each policy on the incidence of pregnancy are the same for the "baseline-pregnancy-outcome" and "unintended-pregnancy-outcome" specifications; b) that each policy's effect on the incidence of childbearing is larger for the former specification than for the latter; and c) that each policy's effect on the incidence of abortion is larger for the latter than for the former.

Table 8. Estimated Distributions of Unintended-Pregnancy Outcomes, by Age and Marital Status*

Pregnancy Outcome	Unmarried Teens	Unmarried Non-Teens	All Unmarried	All Married	All Teens	All Non-Teens	All
Birth	47.0%	32.5%	36.9%	65.5%	49.3%	44.3%	45.5%
Abortion	35.5%	53.9%	48.3%	21.6%	32.7%	42.6%	40.3%
Fetal Loss	17.6%	13.6%	14.8%	12.9%	17.9%	13.1%	14.2%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

* Based on special tabulations of data gathered by the Guttmacher Institute.

Miscellaneous Issues: Differences Between Behavioral Targets and Simulated Behavioral Changes

FamilyScape models behavior probabilistically. Each individual in the simulation population is assigned a set of behavioral probabilities by applying the results of regression analyses that were conducted using real-world data, and behavioral attributes are then assigned by comparing these probabilities to the results of a series of random draws taken from a uniform (0,1) distribution. For example, regressions were estimated of the probability of using oral contraception using data from the 2002 NSFG. Separate models were estimated for married and unmarried women, and the independent variables in these regressions controlled for race, education, educational attainment, socioeconomic status, and coital frequency. The coefficients from these regressions are imported into FamilyScape and are used to calculate, for each female member of the simulation population, a probability of using oral contraception. Women are then assigned to be oral contraceptors (or not) by comparing their assigned probabilities to the results of separate random draws that are taken for each woman. A similar approach is used to model the use of other types of contraception, the frequency of intercourse, pregnancy outcomes, etc.[66]

For each policy simulation, behavioral changes are modeled by multiplying the relevant population's behavioral probabilities by the appropriate ratio. For example, in order to simulate a three percent increase in condom use among adult males, I multiply the condom-use probabilities for all members of this group by 1.03 before comparing these probabilities to the results of the random draws described above. Thus, no two runs of the model under its baseline specification or under any particular policy specification will produce precisely the same results in terms of the share of the simulation population that uses contraception, the amount of intercourse that occurs, the rates of pregnancy and childbearing that obtain, etc. The results presented in Thomas (2010) are in fact averages that are calculated using data from approximately 500 one-year runs of the model.[67]

---

[66] See Thomas and Monea (2009) for more information on the method by which behavioral profiles are assigned to members of the FamilyScape simulation population.

[67] More precisely, for each simulation, the model is typically run 50 times over in a steady state for ten years of analysis time. Annual pregnancy and birth rates and other outcomes of interest are then calculated by averaging data from these 500 years of analysis time in order to ensure that the results for a given simulation are not simply a function of random outliers in the results for any single year.

Given the model's heavy reliance on the random assignment of individual behavioral attributes, it is difficult to model behavioral changes with perfect precision. This is especially true for sexual activity, which is a function of numerous random dynamics, including individuals' own sexual proclivities, the frequency with which they enter into relationships over the course of the simulation, the durations of those relationships, and the sexual proclivities of the individuals with whom they are paired. Thus, there is sometimes a modest difference between the magnitudes of the behavioral changes that one wishes to model for a given simulation and the changes that actually obtain. Table 9 compares the behavioral targets for each simulation to the behavioral changes that are actually simulated. These results show that the simulated behavioral changes for each simulation are either exactly equal to or are very close to their relevant targets.

Table 9. Comparison of Behavioral Targets and Simulated Behavioral Changes for Benefit-Cost Policy Simulations

		Behavioral Target	Simulated Behavioral Change
Mass Media Campaign	Share of Unmarried Adult Males Who Use Condoms As a Result of the Campaign	3.0%	2.9%
	Share of Unmarried Teenaged Males Who Use Condoms As a Result of the Campaign	1.5%	1.5%
Teen Pregnancy Prevention Program	*Initial Specification:* Increase in Number of Condom Users Among Unmarried, Low-SES Teenaged Males	12.5%	12.5%
	Initial Specification: Increase in Number of Pill Users Among Unmarried, Low-SES Teenaged Females	12.5%	12.1%
	Initial Specification: Increase in Number of Unmarried, low-SES Teens who are Sexually Inactive in the Near Term	7.5%	7.6%
	Alternate Specification: Increase in Number of Condom Users Among Unmarried, Low-SES Teenaged Males	12.5%	12.7%
	Alternate Specification: Increase in Number of Pill Users Among Unmarried, Low-SES Teenaged Females	12.5%	12.8%
Expansion in Subsidized Family Planning Services Under Medicaid *Simulated Change in Contraceptive Use : Concentrated Among Women Under 200 Percent of Poverty*	*Initial Specification:* Share of Sexually-Active Women Who Use Contraception as a Result of the Expansion	5.0%	4.8%
	Initial Specification: Proportion of Increase in Contraceptive Use Occurring Among Unmarried Women	70.0%	70.0%
	Alternate Specification: Share of Sexually-Active Women Who Use Contraception as a Result of the Expansion	2.5%	2.5%
	Alternate Specification: Proportion of Increase in Contraceptive Use Occurring Among Unmarried Women	70.0%	69.9%

REFERENCES

Advocates for Youth. *Science and Success, Second Edition: Sex Education and Other Programs that Work to Prevent Teen Pregnancy, HIV & Sexually Transmitted Infections.* Washington, D.C.: Advocates for Youth, 2008.

Amaral, Gorette, Diane Greene Foster, M. Antonia Biggs, Carolyn Bradner Jasik, Signy Judd, and Claire D. Brindis. "Public Savings from the Prevention of Unintended Pregnancy: A Cost Analysis of Family Planning Services in California." *Health Services Research* 42 (2007): 1960 – 80.

Billen, Patricia, Donald Boyd, Lucy Dadayan, and Thomas Gais. *State Funding for Children: Spending in 2004 and how it Changed from Earlier Years.* Albany, NY: The Nelson A. Rockefeller Institute of Government, 2007.

Borawski, Elaine A., Erika S. Trapl, Kimberly Adams-Tufts, Laura L. Hayman, Meredith A. Goodwin, and Loren D. Lovegreen. "Taking Be Proud! Be Responsible! to the Suburbs: A Replication Study." *Perspectives on Sexual and Reproductive Health* 41 (2009): 12 – 22.

CDC AIDS Community Demonstration Projects Research Group. "Community-Level HIV Intervention in 5 Cities: Final Outcome Data from the CDC AIDS Community Demonstration Projects." *American Journal of Public Health* 89 (1999): 336 – 345.

Centers for Disease Control and Prevention. *Youth Media Campaign: VERB, It's What you Do*, http://www.cdc.gov/youthcampaign/ (accessed April 1, 2010).

Child Trends. *Programs with Mixed Reviews for Delaying the Initiation of Sexual Intercourse: HIV/AIDS Education*, http://www.childtrends.org/what_works/youth_development/ResponsibleTeen.asp (accessed September 14, 2009).

Cook, Allison, Lisa Dubay, and Bowen Garrett. *How Will Uninsured Parents be Affected by Health Reform?* Washington, D.C.: Kaiser Commission on Medicaid and the Uninsured, 2009.

Coyle, Karin, Karen Basin-Enquist, Douglas Kirby, Guy Parcel, Stephen Banspach, Janet Collins, Elizabeth Baumler, Scott Carvajal, and Ronald Harrist. "Safer Choices: Reducing Teen Pregnancy, HIV, and STDs." *Public Health Reports* 116 (2001): 82 – 93.

Cutler, David and Jonathan Gruber. "Does Public Insurance Crowd Out Private Insurance?" *Quarterly Journal of Economics* 111 (2006): 391 – 430.

Dubay, Lisa, Allison Cook, and Bowen Garrett. *How will Uninsured Childless Adults be Affected by Health Reform?* Washington, D.C.: Kaiser Commission on Medicaid and the Uninsured, 2009.

Farrelly, Matthew C., Kevin C. Davis, Lyndon Haviland, Peter Messeri, and Cheryl G. Healton. "Evidence of a Dose-Dose Relationship Between "truth" Antismoking Ads and Youth Smoking Prevalence." *American Journal of Public Health* 95 (2005): 425 – 431.

Finer, Lawrence B. and Stanley Henshaw. "Disparities in Rates of Unintended Pregnancy in the United States, 1994-2001." *Perspectives on Sexual and Reproductive Health* 38 (2006): 90 – 96.

Fishbein, Martin, Carolyn Guenther-Grey, Wayne D. Johnson, Richard J. Wolitski, Alfred McAlister, Cornelis A. Rietmeijer, and Kevin O'Reilly. "Using a Theory-Based Community Intervention to Reduce AIDS Risk Behaviors: The CDC's AIDS Community Demonstration Projects." In Stuart Oskamp and Suzanne C. Thompson, Eds. *Understanding and Preventing Risk Behavior.* London: Sage Publications (1996), 177 – 206.

Florida Department of Health. *Revised Medicaid Family Planning Waiver Program*, http://www.doh.state.fl.us/family/famplan/waiver.html (accessed March 30, 2010).

Frost, Jennifer J., Adam Sonfield, and Rachel Benson Gold. *Estimating the Impact of Expanding Medicaid Eligibility for Family Planning Services.* New York: Guttmacher Institute Occasional Report #28, 2006.

Government Accountability Office. *Contractor's National Evaluation did not Find that the Youth Anti-Drug Media Campaign was Effective in Reducing Youth Drug Use.* Washington, D.C.: United States Government Accountability Office, 2006a.

Government Accountability Office. *Highlights: Media Contracts, Activities and Financial Obligations for Seven Financial Departments.* Washington, D.C.: United States Government Accountability Office Report #06-818, 2006b.

Guttmacher Institute. *State Policies in Brief: State Medicaid Family Planning Eligibility Expansions.* New York: The Guttmacher Institute, 2010a.

Guttmacher Institute. *U.S. Teenage Pregnancy Statistics: National and State Trends and Trends by Race and Ethnicity.* New York: The Guttmacher Institute, updated January 2010b.

Guttmacher Institute. *U.S. Teenage Pregnancy Statistics: National and State Trends and Trends by Race and Ethnicity.* New York: The Guttmacher Institute, 2006.

Henshaw, Stanley K. and Kathryn Kost. *Trends in the Characteristics of Women Obtaining Abortions, 1974-2004.* New York: Guttmacher Institute, 2008.

Holtgrave, David R., Katherine A. Wunderink, Donna M. Vallone, and Cheryl G. Healton. "Cost-Utility Analysis of the National Truth Campaign to Prevent Youth Smoking." *American Journal of Preventive Medicin*e 36 (2009): 385 – 388.

Hornik, Robert, Lela Jacobsohn, Robert Orwin, Andrea Piesse, and Graham Kalton. "Effects of the National Youth Anti-Drug Media Campaign." *American Journal of Public Health* 98 (2008): 2229 – 2236.

Huhman, Marian E., Lance D. Potter, Jennifer C. Duke, David R. Judkins, Carrie D. Heitzler, and Faye L. Wong. "Evaluation of a National Physical Activity Intervention for Children: VERB Campaign, 2002 – 2004." *American Journal of Preventive Medicine* 32 (2007): 38 – 43.

Jemmott, John B. III, Loretta Sweet Jemmott, and Geoffrey T. Fong. "Reductions in HIV Risk-Associated Sexual Behaviors among Black Male Adolescents: Effects of an AIDS Prevention Intervention." *American Journal of Public Health* 82 (1992): 372 – 377.

Jones, Rachel K., Mia R.S. Zolna, Stanley K. Henshaw and Lawrence B. Finer. "Abortion in the United States: Incidence and Access to Services, 2005." *Perspectives on Sexual and Reproductive Health* 40 (2008): 6 – 16.

Kaiser Family Foundation. *Income Eligibility Levels for Children's Separate CHIP Programs by Annual Incomes and as a Percent of the Federal Poverty Level*, http://www.statehealthfacts.org/comparemaptable.jsp?ind=204&cat=4#notes-1 (accessed March 31, 2010).

Kaiser Family Foundation. *Where are States Today? Medicaid and State-Funded Eligibility Levels for Low-Income Adults*, http://www.kff.org/medicaid/upload/7993.pdf (2009).

Kaiser Family Foundation. *Medicaid's Role in Family Planning*. Menlo Park, CA: The Henry J. Kaiser Family Foundation, 2007.

Kaiser Family Foundation. *Forum to Explore Effectiveness of Public Education Campaigns*. Forum Transcript. Menlo Park, CA: The Henry J. Kaiser Family Foundation, 2006.

Kearney, Melissa S. and Daniel M. Levine. "Subsidized Contraception, Fertility, and Sexual Behavior." *Review of Economics and Statistics* 91 (2009): 137 – 151.

Keegles, Susan M., Robert B. Hays, and Thomas J. Coates. "The MPowerment Project: A Community-Level HIV Prevention Intervention for Young Gay Men." *American Journal of Public Health* 86 (1996): 1129 – 1136.

Kirby, Douglas. *Emerging Answers 2007: Research Findings on Programs to Reduce Teenage Pregnancy and Sexually Transmitted Diseases*. Washington, D.C.: The National Campaign to Prevent Teen and Unplanned Pregnancy, 2007.

Kirby, Douglas, Elizabeth Baumler, Karin Coyle, Karen Basen-Enquist, Guy S. Parcel, Ron Harrist, and Stephen Baspach. "The 'Safer Choices' Intervention: Its Impact on the Sexual Behaviors of Different Subgroups of High School Students." *Journal of Adolescent Health* 35 (2004): 442 – 452.

Krisberg, Kim. "Successful CDC VERB Campaign in Danger of Losing Funding: Budget Cuts Likely." *Nations Health* 35 (2005).

Machlin, Steven R. and Frederick Rohde. 2007. *Health Care Expenditures for Uncomplicated Pregnancies.* Rockville, MD: Agency for Healthcare Research and Quality, MEPS Research Findings Report #27, 2007.

Macomber, Jennifer, Julia Isaacs, Tracy Vericker, and Adam Kent. *Public Investment in Children's Early and Elementary Years (Birth to Age 11).* Washington, D.C.: The Urban Institute and The Brookings Institution, 2010.

Manlove, Jennifer, Kerry Franzetta, Krystal McKinney, Angela Romano Papillo, and Elizabeth Terry-Humen. *A Good Time: After-School Programs to Reduce Teen Pregnancy.* The National Campaign to Prevent Teen and Unplanned Pregnancy, 2004.

Merrill, Chaya and Claudia Steiner. *Hospitalizations Related to Childbirth, 2003.* Rockville, MD: Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project Statistical Brief #11, 2006.

Monea, Emily and Adam Thomas. *The Public Cost of Pregnancy.* Washington, D.C.: The Brookings Institution, 2010.

National Campaign to Prevent Teen and Unplanned Pregnancy. *Summary of Provisions Related to Teen and Unplanned Pregnancy in HR 3590: Patient Protection and Affordable Care Act*, http://www.thenationalcampaign.org/policymakers/PDF/SummaryProvisions_TUO_HealthReform.pdf (accessed April 2, 2010).

New York State Department of Health. *Medicaid Family Planning Services*, http://www.health.state.ny.us/health_care/managed_care/famplan10ques.htm (accessed March 30, 2010).

Nielsen Media. *Local Television Market Universe Estimates*, http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/08/2009-2010-dma-ranks.pdf (accessed March 30, 2010).

Noar, Seth A. "Challenges in Evaluating Health Communication Campaigns: Defining the Issues." *Communication Methods and Measures* 3 (2009): 1 – 11.

Noar, Seth A. "A 10-Year Retrospective of Research in Health Mass Media Campaigns: Where do we go from Here?" *Journal of Health Communication* 11 (2006): 21 – 42.

Olaiya, Samuel Tunde. *Medical Cost Savings Attributable to Comprehensive Sex Education Programs that Delay Coitus and Increase Condom Use Among Adolescents in the United States.* Doctoral Dissertation. Columbus, OH: The Ohio State University, 2006.

Orwin, Robert, Diane Cadell, Adam Chu, Graham Kalton, David Maklan, Carol Morin, Andrea Piesse, Sanjeev Sridharan, Diane Steele, Kristie Taylor, and Elena Tracy. *Evaluation of the National Youth Anti-Drug Media Campaign: 2004 Report of Findings.* Rockville, MD: Westat, 2006.

Palmgreen, Philip, Elizabeth P. Lorch, Michael T. Stephenson, Rick H. Hoyle, and Lewis Donohue. "Effects of the Office of National Drug Control Policy's Marijuana Initiative Campaign on High-Sensation-Seeking Adolescents." *American Journal of Public Health* 97 (2007), 1644 – 1649.

Santelli, John S., David D. Celentano, Christine Rozsenich, Aria D. Crump, Mary V. Davis, Michelle Polacsek, Marycatherine Augustyn, Jon Rolf, Alfred L. McAlister, and LaWanda Burwell. "Interim Outcomes for a Community-Based Program to Prevent Perinatal HIV Transmission." *AIDS Education and Prevention* 7 (1995): 210 – 220.

Scher, Lauren S., Rebecca A. Maynard, and Matthew Stagner. *Interventions Intended to Reduce Pregnancy-Related Outcomes among Adolescents.* Oslo, Norway: Campbell Systematic Reviews Report #2006-12, 2006.

Sikkema, Kathleen J., Eileen S. Anderson, Jeffrey A. Kelly, Richard A. Winett, Cheryl Gore-Felton, Roger A. Roffman, Timothy G. Heckman, Kristi Graves, Raymond G. Hoffman, and Michael J. Brondino. "Outcomes of a Randomized, Controlled Community-Level HIV Prevention Intervention for Adolescents in Low-Income Housing Developments." *AIDS* 19 (2005): 1509 – 1516.

Snyder, Leslie B. "The Impact of the Surgeon General's Understanding AIDS Pamphlet in Connecticut." *Health Communication* 3 (1991): 37 – 57.

Snyder, Leslie B., Mark A. Hamilton, Elizabeth W. Mitchell, James Kiwanuka-Tondo, Fran Fleming-Milici, and Dwayne Proctor. "A Meta-Analysis of Mediated Health Communication Campaigns on Behavior Change in the United States." *Journal of Health Communication* 9 (2004): 71 – 96.

St Lawrence, Janet, Ted L. Brasfield, Kennis W. Jefferson, Edna Allenyne, Robert E. O'Bannon III, and Aaron Shirley. "Cognitive-Behavioral Intervention to Reduce African American Adolescents' Risk for HIV Infection." *Journal of Consulting and Clinical Psychology* 63 (1995): 221 – 237.

St Lawrence, Janet S., Richard A. Crosby, Ted L. Brasfield, and Robert E. O'Bannon III. "Reducing STD and HIV Risk Behavior of Substance-Dependent Adolescents: A Randomized Controlled Trial." *Journal of Consulting and Clinical Psychology* 70 (2002): 1010 – 1021.

Suellentrop, Katherine. *What Works 2009: Curriculum-Based Programs that Prevent Teen Pregnancy.* Washington, D.C.: The National Campaign to Prevent Teen and Unplanned Pregnancy, 2009.

Thomas, Adam. "A Benefit-Cost Assessment of Three Strategies to Prevent Unintended Pregnancy." Revise and Resubmit, *Journal of Policy Analysis and Management*, 2011.

Thomas, Adam and Emily Monea. *FamilyScape: A Simulation Model of Family Formation.* Washington, D.C.: The Brookings Institution, 2009.

Thomas, Adam and Emily Roessel. *Contraceptive Use and Efficacy.* Washington, D.C.: The Brookings Institution, 2008.

Trussell, James, Anjana M. Lalla, Quan V. Doan, Eileen Reyes, Lionel Pinto, and Joseph Gricar. "Cost Effectiveness of Contraceptives in the United States." *Contraception* 79 (2009): 5 – 14.

United States Census Bureau. *The 2010 Statistical Abstract of the United States.* Washington, D.C.: United States Census Bureau, 2010.

United States Congress, 11th Congress, 2nd Session. *HR 3590: The Patient Protection and Affordable Care Act.* Washington, D.C: Government Printing Office, 2010.

United Way of Rochester. *School Aged Evidence-Based Programs,* http://www.uwrochester.org/Data/Documents/YouthProgramWriteUps.pdf (accessed January 26, 2010).

Ventura, Stephanie J., Joyce C. Abma, and William D. Mosher. *Estimated Pregnancy Rates by Outcome for the United States, 1990-2005: An Update.* Hyattsville, MD: National Center for Health Statistics, National Vital Statistics Report #58-4, 2009.

Ventura, Stephanie J., Joyce C. Abma, William D. Mosher, and Stanley Henshaw. "Estimated Pregnancy Rates for the United States, 1990-2000: An Update." *National Vital Statistics Reports* 52 (2004): 1 – 9.

Villaruel, Antonia M., John B. Jemmott III, and Loretta S. Jemmott. "A Randomized Controlled Trial Testing an HIV Prevention Intervention for Latino Youth." *Archives of Pediatrics & Adolescent Medicine* 160 (2006): 772 – 777.

Wang, Li Yan, Margarett Davis, Leah Robin, Janet Collins, Karin Coyle, and Elizabeth Baumler. "Economic Evaluation of Safer Choices." *Archives of Pediatrics & Adolescent Medicine* 154 (2000): 1017 – 1024.

Zimmerman, Rick S., Philip M. Palmgreen, Seth M. Noar, Mia Liza A. Lustria, Hung-Yi Lu, and Mary Lee Horosewski. "Effects of a Televised Two-City Safer Sex Mass Media Campaign Targeting High-Sensation-Seeking and Impulsive-Decision-Making Young Adults." *Health Education & Behavior* 34 (2007): 810 – 826.