September 2017

APPENDIX

This appendix includes the following sections:

1. Analytic Approach
2. Simulation findings for alternative model assumptions
3. Tables of descriptive statistics

Analytic Approach

Using simulations, we assess the bias-variance tradeoffs between randomization of individuals and randomization of clusters. The simulations model a hypothetical evaluation of a program that reduces an individual's *inclination* to engage in risky sexual behavior (an unobserved continuous variable) which, in turn, reduces the prevalence of *actually* engaging in risky sexual behavior (an observed dichotomous variable). The simulations are based on a model of how dating couples are formed and decide whether to engage in risky behavior. Contamination bias occurs when one member of the couple is in the treatment group and the other is in the control group. Given a specified program impact on the inclination to engage in risky behavior, we use simulations to calculate the magnitude of the impact in percentage points on the prevalence rate of actually engaging in risky behavior. We also calculate the probability that the specified impact will be detected (which is statistical power).

Model of Outcomes

The ultimate outcomes of interest are dichotomous measures of engaging in risky behaviors, such as having unprotected sex. We model these dichotomous outcomes using the latent variable approach (Goldberger 1964; Maddala 1983). That is, we assume that there is an unobserved latent inclination to engage in a risky behavior (y_{latent}), but that the dichotomous risky behavior (y) is only observed when a threshold is crossed on the latent continuous variable. We model y_{latent} and y separately for non-dating students (hereafter, singles) and dating pairs (hereafter, couples).

For singles, we assume that y_{latent} is normally distributed with mean 0 and variance 1 (equation 1). This variance involves both a cluster level component and an individual level component. Variance at the cluster-level is specified by the parameter ρ and at the individual-level is $1-\rho$, where $0<\rho<1$ (thus, ρ is the ICC). We represent the threshold as the inverse of the cumulative density function for a specified prevalence rate (P) of the dichotomous outcome (equation 2, where I is the indicator function and Φ is the standard normal cumulative distribution function). For example, if the prevalence rate is 25 percent then the threshold is approximately 0.67 (that is, individuals must be in the top 25 percent of the latent continuous variable, and 0.67 is the 75th percentile of the standard normal distribution).

$$(1)\ y_{latent} \sim N(0,1)$$

$$(2)\ y = I\left(y_{latent} > \Phi^{-1}(1-P)\right)$$

For couples, we assume that each member of the couple has their own value of y_{latent} but that the decision to engage in risky behavior depends on a latent continuous variable $y_{(latent,couple)}$ that is a function (F) of the two individual values of y_{latent}, (equation 3). We represent the threshold as the inverse of the cumulative density function for a specified prevalence rate (P) of the dichotomous outcome (equation 4, where I is the indicator function and G is the cumulative density function for $y_{latent,couple}$).

$$(3)\ y_{latent,couple} = F\left(y_{latent,1}, y_{latent,2}\right)$$

$$(4)\ y = I\left(y_{latent,couple} > \mathrm{G}^{-1}(1-P)\right)$$

The functions F and G vary depending on the models for how dating couples are formed and how members of a couple influence each other. The findings presented in the main text (hereafter, "benchmark findings") are based on simulations in which we assume that dating couples are formed using a model we call "similars-attract," in which the members of each couple are very similar with respect to their latent inclination to engage in risky behavior. Specifically, we form couples by sorting all dating individuals by y_{latent} (where daters are a subset of the full sample). The first couple consist of the first two individuals in the sorted list, the second couple consists of the next two individuals, and so on. The benchmark findings also depend on the assumption that couples influence each other through a process we call "meet-in-the-middle", in which F is the simple average of the couple's individual values of y_{latent}. We describe three alternate specifications below.

For couple formation and influence models underlying the benchmark findings, the function G is described by equation 5, where n is the number of dating individuals (so n/2 is the number of dating couples) and $E(r,n)$ is the expected value of the r-th order statistic, which is described by equation 6[1], where ϕ is the standard normal probability density function. In words, equation 5 says that the probability that $y_{latent,couple}$ is less than some specified value (say, y^*) is equal to the proportion of couples in which the expected value of $y_{latent,couple}$ is less than y^*. In this case, the expected value for a given couple is the average of the expected order statistics for the two members of the couple.

Sensitivity Analyses

$$(5)\ G\left(y_{latent,couple}\right) = \frac{\sum_{i=1}^{n/2} I\left[\left(E\left(1+2(i-1),\frac{n}{2}\right)+E\left(2i,\frac{n}{2}\right)\right)/2 < y_{latent,couple}\right]}{n/2}$$

$$(6)\ E(r,n) = \frac{n!}{(r-1)!(n-r)!}\int_{-\infty}^{\infty} x\{1-\Phi(x)\}^{r-1}\{\Phi(x)\}^{n-r}\phi(x)dx$$

We also examine the sensitivity of our findings to alternative couple formation and influence models. We call the alternative couple formation model "spin-the-bottle" and the alternative couple influence model "two-to-tango." In the spin-the-bottle model, couples are formed randomly with respect to their latent inclination to engage in risky behavior (as opposed to couples being formed by having similar latent inclinations). In the two-to-tango model, $y_{latent,couple} = \min(y_{latent,1}, y_{latent,2})$, as opposed to averaging the two individual values of y_{latent}.

For the couple formation and influence models similars-attract and two-to-tango, the function G is described by equation 7. In words, equation 7 says that the probability that $y_{latent,couple}$ is less than some specified value (say, y^*) is the proportion of couples in which the expected value of $y_{latent,couple}$ is less than y^*. In this case, the expected value for a given couple is the minimum of the expected order statistics for the two members of the couple.

[1]The distributional properties of order statistics are available in textbooks, for example Hogg and Craig (1978).

$$(7)\ G\left(y_{latent,couple}\right) = \frac{\sum_{i=1}^{n/2} I\left[E\left(1+2(i-1),\frac{n}{2}\right)<y_{latent,couple}\right]}{n/2}$$

For the couple formation and influence models spin-the-bottle and meet-in-the-middle, the function G is described by equation 8, where μ=0 and $\sigma = 1/\sqrt{2}$. In other words, $y_{latent,couple} \sim N(0, 1/\sqrt{2})$, which follows directly from $y_{latent,couple}$ being the average of two independent $N(0,1)$ variables.

$$(8)\ G\left(y_{latent,couple}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{y_{latent,couple}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

For the couple formation and influence models spin-the-bottle and two-to-tango, the function G is described by equation 9, which is the cumulative distribution function for the order statistic corresponding to the minimum of two independent $N(0,1)$ variables.

$$(9)\ G\left(y_{latent,couple}\right) = 2\left(1 - \Phi\left(y_{latent,couple}\right)\right)\phi\left(y_{latent,couple}\right)$$

In sum, the benchmark and sensitivity analyses vary according to the parameters described in the table below:

Model	Couple Formation Mechanism	Influence Model	F	G (equation number)
Benchmark	Similars attract	Meet in the middle	Mean	5
Sensitivity analysis approach 1	Similars attract	Two to tango	Min	7
Sensitivity analysis approach 2	Spin the bottle	Meet in the middle	Mean	8
Sensitivity analysis approach 3	Spin the bottle	Two to tango	Min	9

Modeling Impacts

We model the impact of randomization on outcomes using the potential outcomes framework (Rubin 1974). Every person has a potential outcome that they would experience if assigned to the control group and a (possibly) different potential outcome if assigned to the treatment group. We assume the individual-level difference between treatment and control potential latent outcomes is a constant value δ. Equations 10 and 11 show the latent outcomes for the treatment and control groups for singles in the absence of contamination (equation 10 restates equation 1, but now with a superscript indicating that this is the potential outcome when assigned to the control group).

$$(10)\ y^c_{latent} \sim N(0,1)$$

$$(11)\ y^t_{latent} = y^c_{latent} + \delta$$

For couples, the latent outcome depends on the treatment assignments of both members of the couple. With randomization of clusters, both individuals in a couple are in the same treatment group, yielding equations 12 and 13. With randomization of individuals, some couples will, by chance, have outcomes described by equations 12 and 13. But other couples will have one individual in the treatment group and another individual in the control group (equation 14, where i could be either 1 or 2 and j is 1 if i is 2 or 2 if i is 1).

$$(12)\ y^{c}_{latent,couple} = F\left(y^{c}_{latent,1}, y^{c}_{latent,2}\right)$$

$$(13)\ y^{t}_{latent,couple} = F\left(y^{t}_{latent,1}, y^{t}_{latent,2}\right)$$

$$(14)\ y^{ct}_{latent,couple} = F\left(y^{c}_{latent,i}, y^{t}_{latent,j}\right)$$

The impact on the expected latent outcome for singles ($\Delta_{latent,singles}$) is just δ. The impact on the expected latent outcome for couples is described in equation 15, where η is the fraction of couples that include individuals from both the treatment and control groups. When clusters are randomized, we assume $\eta = 0$. When individuals are randomized, η varies randomly. If half of students are randomized to the treatment group the mean of η is 0.50 (the variance of η depends on sample size).

$$(15)\ \Delta_{latent,couples} = \eta E\left\{y^{ct}_{latent,couple}\right\} + (1-\eta)E\left\{y^{t}_{latent,couple}\right\} - E\left\{y^{c}_{latent,couple}\right\}$$

For singles, the dichotomous outcomes for the treatment and control conditions are described in equations 16 and 17, where $P_{singles}$ is the prevalence rate of the dichotomous outcome for singles in the absence of the intervention. The impact for singles is the difference in expected outcomes between the treatment and control conditions (equation 18).

$$(16)\ y^{t}_{singles} = I\left(y^{t}_{latent} > \Phi^{-1}\left(1 - P_{singles}\right)\right)$$

$$(17)\ y^{c}_{singles} = I\left(y^{c}_{latent} > \Phi^{-1}\left(1 - P_{singles}\right)\right)$$

$$(18)\ \Delta_{singles} = E\left\{y^{t}_{singles}\right\} - E\left\{y^{c}_{singles}\right\}$$

For couples, the dichotomous outcomes for the treatment and control conditions are described in equations 19 and 20, where $P_{couples}$ is the prevalence rate of the dichotomous outcome for couples in the absence of the intervention. The dichotomous outcomes for mixed-assignment couples is described in equation 21. The impact for couples is described in equation 22.

$$(19)\ y^{t}_{couples} = I\left(y^{t}_{latent} > \mathrm{G}^{-1}\left(1 - P_{couples}\right)\right)$$

$$(20)\ y^{c}_{couples} = I\left(y^{c}_{latent} > \mathrm{G}^{-1}\left(1 - P_{couples}\right)\right)$$

$$(21)\ y^{ct}_{couples} = I\left(y^{ct}_{latent} > \mathrm{G}^{-1}\left(1 - P_{couples}\right)\right)$$

$$(22)\ \Delta_{couples} = \eta E\left\{y^{ct}_{couples}\right\} + (1-\eta)E\left\{y^{t}_{couples}\right\} - E\left\{y^{c}_{couples}\right\}$$

Monte Carlo Simulations

Given the complexity of how dating affects outcomes, both with and without contamination, we use Monte Carlo simulations to calculate statistical power of cluster and individual randomization designs under various scenarios. The simulation involves the following steps:

1. Select the dating model, couple influence model, and values for all parameters and functions described above. The dating models are similars-attract and spin-the-bottle. The influence models are meet-in-the-middle and two-to-tango. The parameters are $P_{couples}$, $P_{singles}$, ρ, θ (the proportion of individuals in each cluster who date other sample members), and δ. The functions are F and G.
2. Generate the individual-level potential latent outcome variables described in equations 10 and 11, associated with the set of constraints described in step 1.
3. Randomly sample the specified proportion (θ) of individuals in each cluster who are to date other sample members.
4. Within each cluster, form couples using the specified dating mechanism (for example, similars attract). Note that couple formation is based on the latent outcome variable prior to random assignment, meaning that couple formation is not affected by the intervention. Further note that all couples consist of individuals in the same cluster (there are no cross-cluster couples).
5. Randomly assign either clusters or individuals to treatment and control groups
6. Generate dichotomous outcome variables for singles and couples using equations 16, 17, 19, 20, and 21.
7. Calculate the impact estimate as the difference in the average of the dichotomous outcome for individuals assigned to the treatment and control groups. When schools are randomly assigned, this impact estimate is not affected by contamination. When individuals are randomly assigned, this impact estimate is affected by contamination.
8. Repeat steps 1-7 a large number of times. From each replication, save the estimated impact and accompanying p-value.

The power of a study to detect an impact on the latent outcome variable of a specified magnitude (δ) is the proportion of Monte Carlo replications in which the estimated p-value is less than 0.05. For a given design (randomization of individuals or clusters) and set of parameter and function values (listed in step 1), we iteratively search for the value of δ that yields 80 percent power. In each step of the iteration, a Monte Carlo simulation is run to calculate power for a candidate value of δ. The iteration ends when 80 percent power is attained. The value of δ in the last iteration is the MDES.

Simulation findings for alternative models of couple formation and influence

The benchmark findings presented in the main text are based on the couple formation model similars-attract and the couple influence model meet-in-the-middle. In this section, we present findings for other combinations of couple formation and influence models. The overall conclusion from these sensitivity analyses are that (1) findings are much more sensitive to the couple formation model than to the influence model and (2) the couple formation model similars-attract (the benchmark) is much more favorable to randomization of schools while spin-the-bottle is more favorable to randomization of individuals.

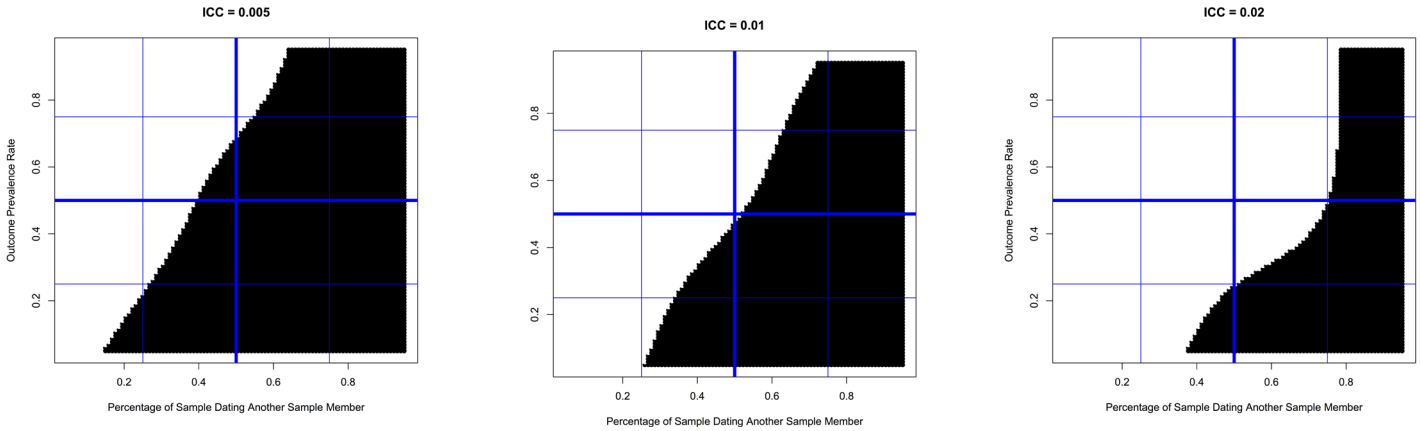
Specifically:

- The combination of similars-attract and two-to-tango models yield findings almost identical to the benchmark findings (figures A1-A4).
- The combination of spin-the-bottle and meet-in-the-middle yields findings that are much more favorable to randomization of individuals than the benchmark findings did (figures A5-A8). The reason spin-the-bottle is more favorable to randomization of individuals is that this model acts as a random shock at the school-level, effectively increasing the ICC. The random shock is due to changes in the small-sample distribution of the $\bar{y}^{c}_{latent,couple}$ within a school.
- The combination of spin-the-bottle and two-to-tango also yields findings that are more favorable to randomization of individuals than the benchmark findings did (figures A9-A12), but not quite as favorable as the combination of spin-the-bottle and meet-in-the-middle.

Simulation findings for alternative sample sizes

The benchmark findings presented in the main text are based on sample sizes of 8 and 30 schools. In this section, we report findings for sample sizes of 16 and 60 schools (Figures A13-A16). As in the main text, these findings are based on the couple formation model similars-attract and the couple influence model meet-in-the-middle.

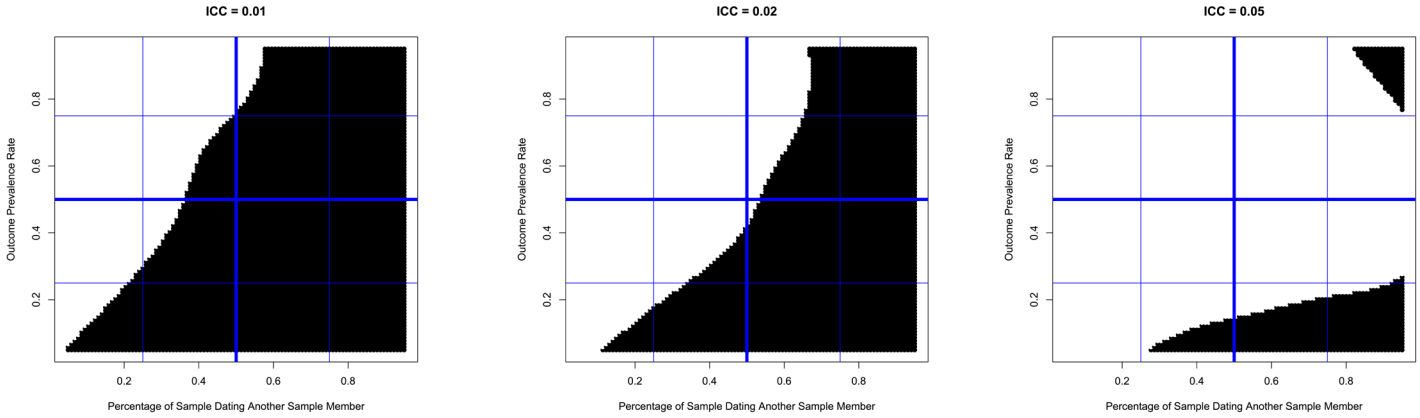
Figure A1: MDES differences between student and school randomization with 8 schools, varying the ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

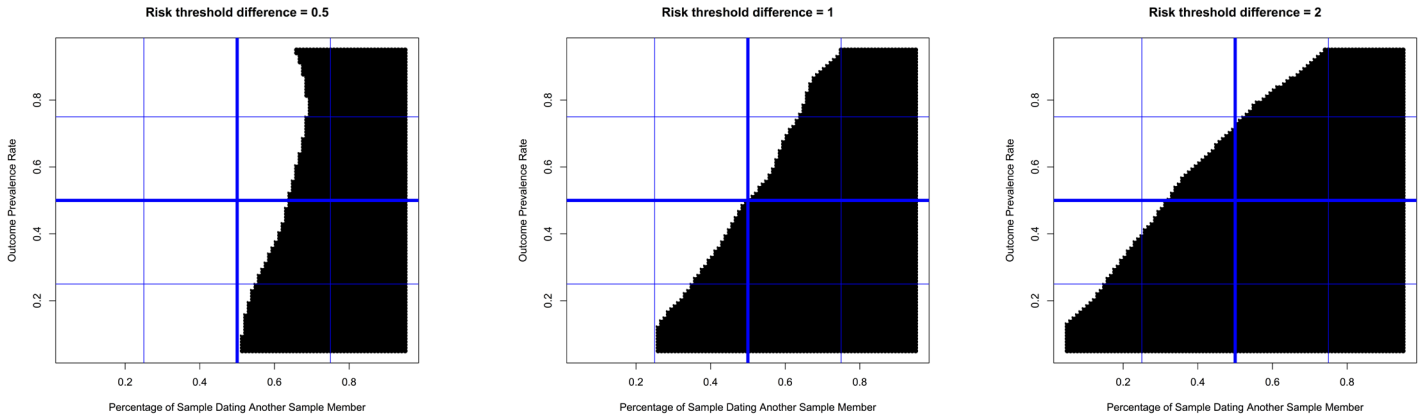
Figure A2: MDES differences between student and school randomization with 30 schools, varying the ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

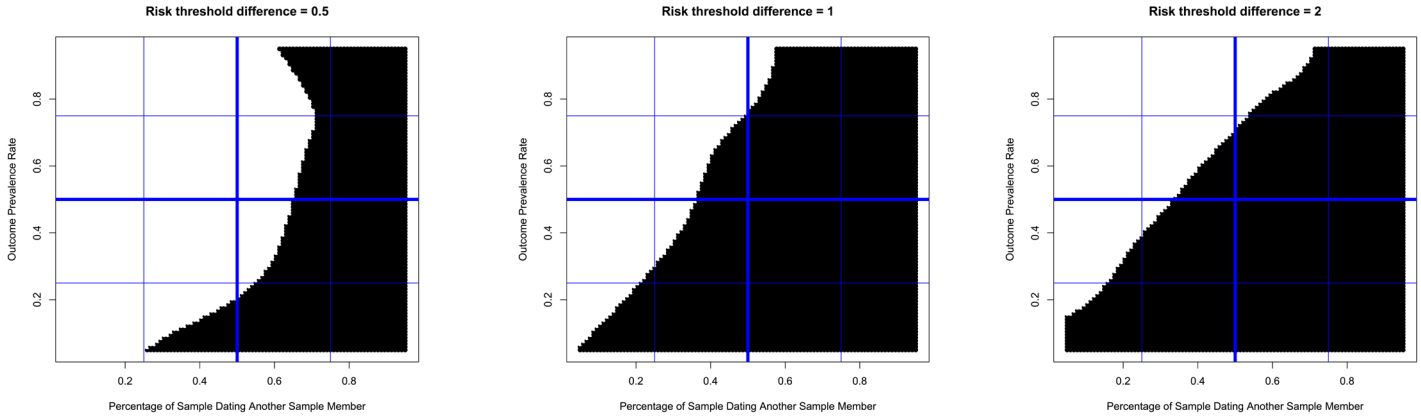
Figure A3: MDES differences between student and school randomization with 8 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
ICC of 0.01

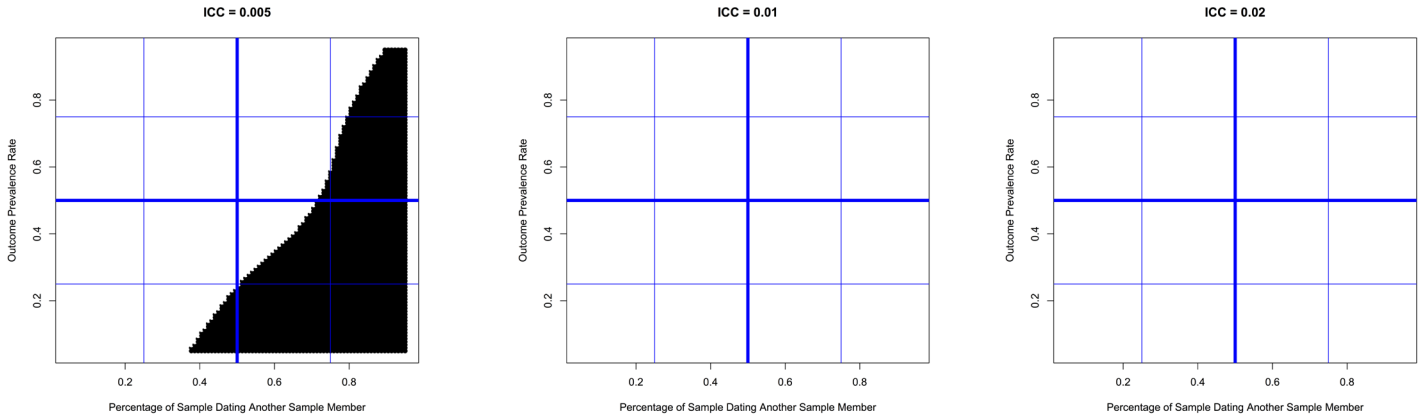
Figure A4: MDES differences between student and school randomization with 30 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
ICC of 0.02

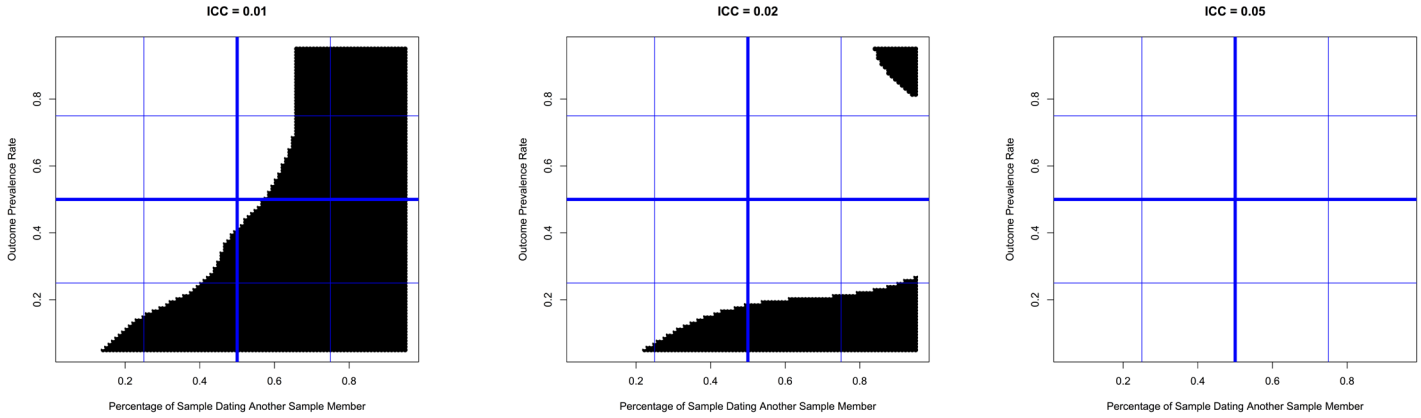
Figure A5: MDES differences between student and school randomization with 8 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

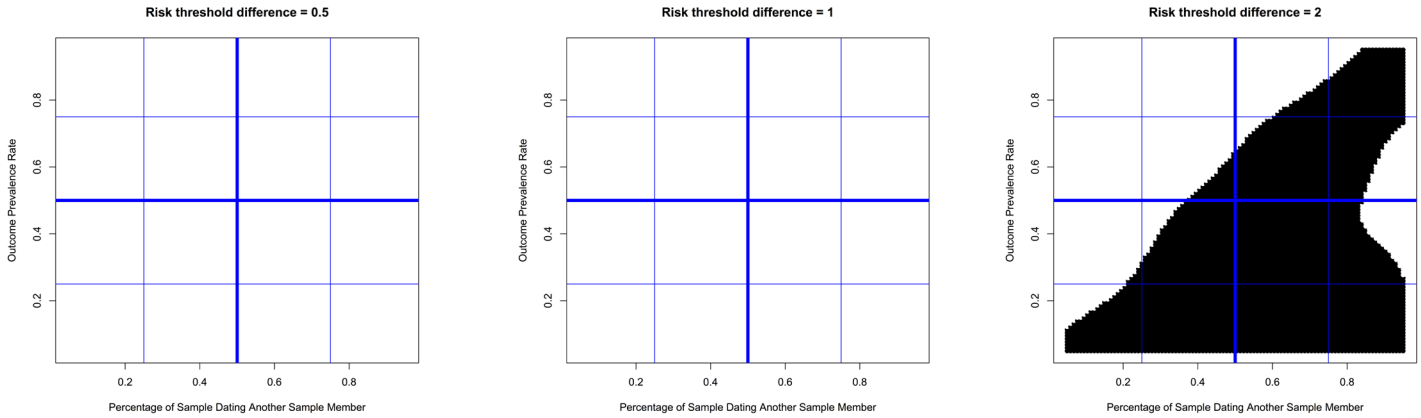
Figure A6: MDES differences between student and school randomization with 30 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

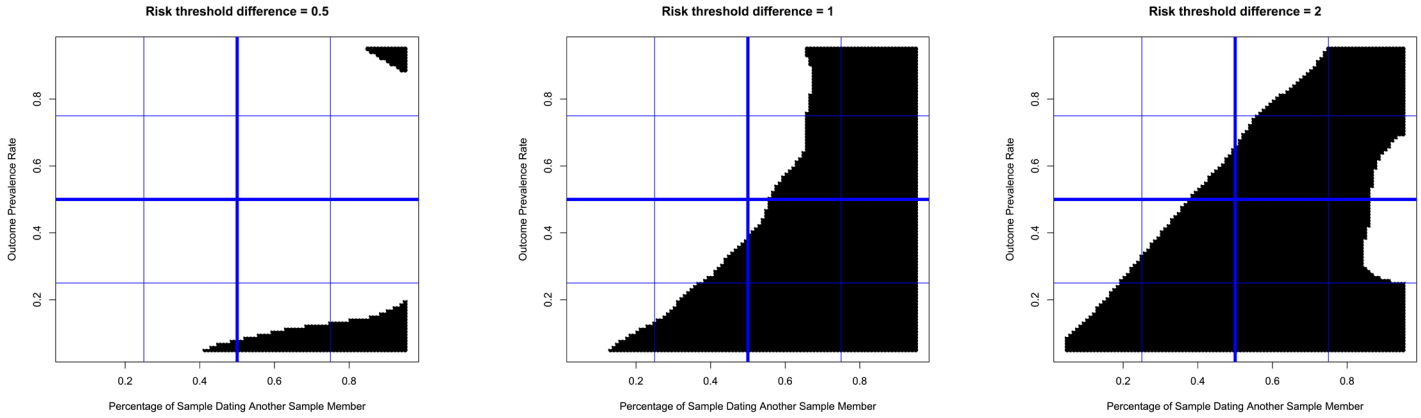
Figure A7: MDES differences between student and school randomization with 8 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
ICC of 0.01

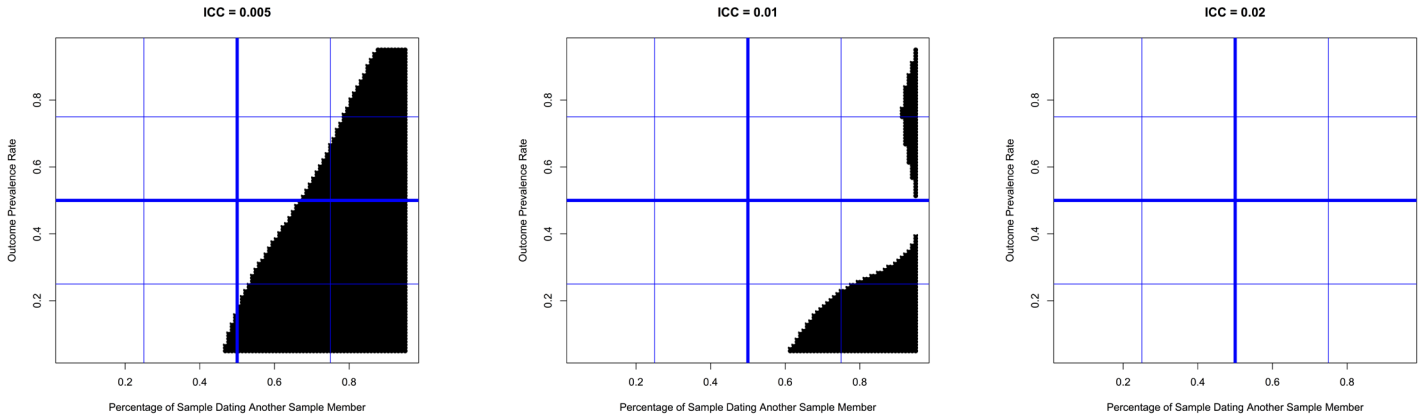
Figure A8: MDES differences between student and school randomization with 30 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
ICC of 0.02

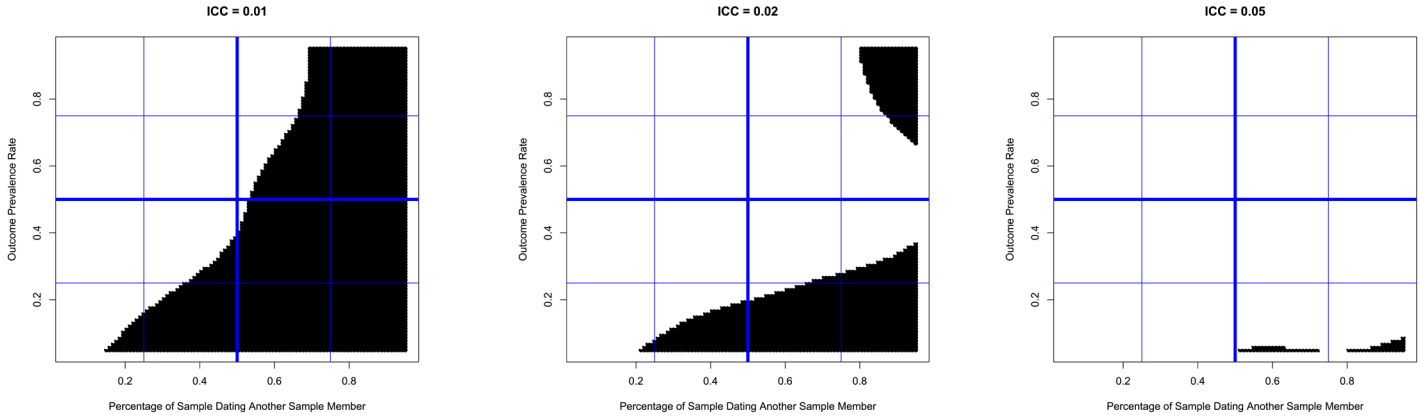
Figure A9: MDES differences between student and school randomization with 8 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

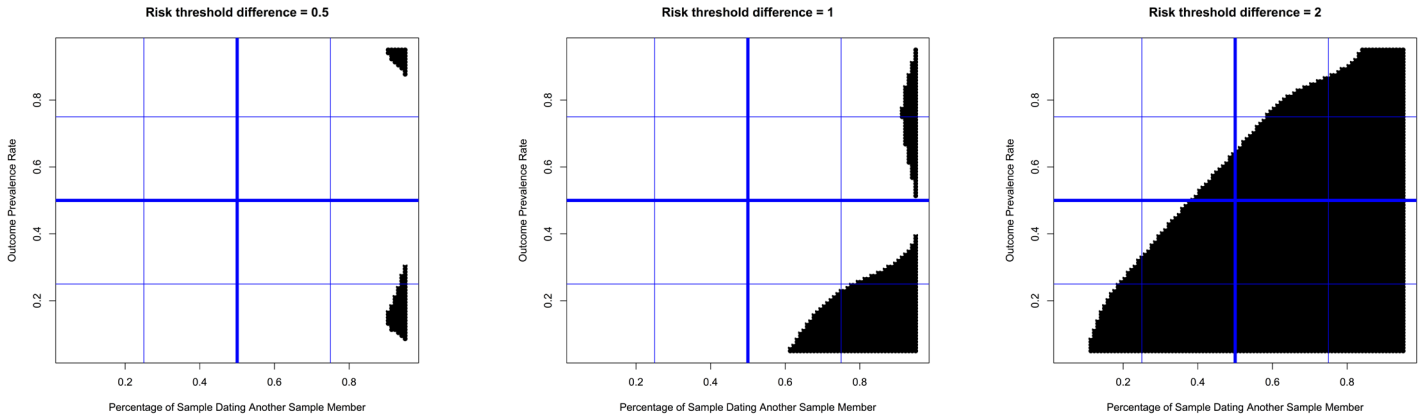
Figure A10: MDES differences between student and school randomization with 8 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

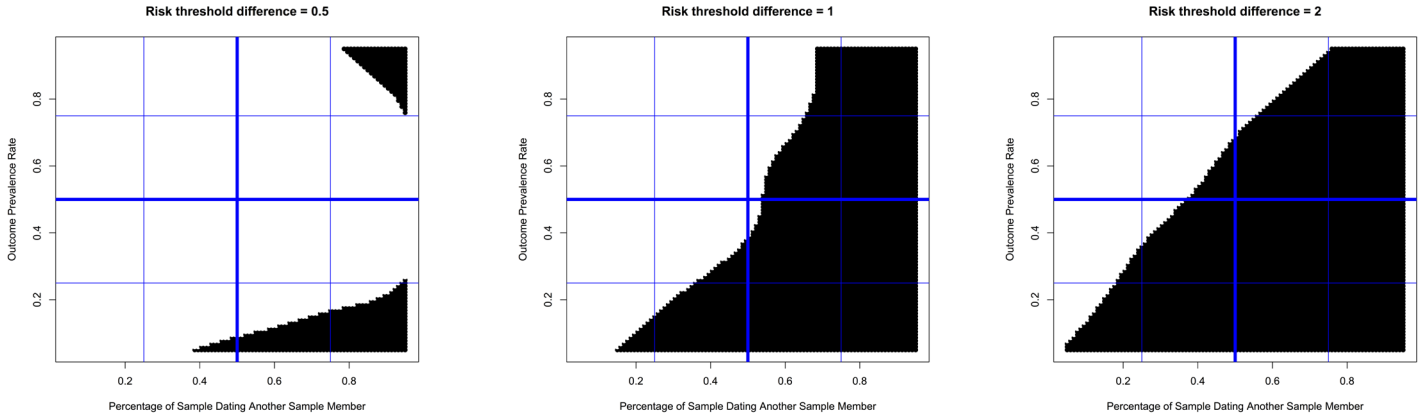
Figure A11: MDES differences between student and school randomization with 8 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
8 schools with 100 students per school
ICC of 0.01

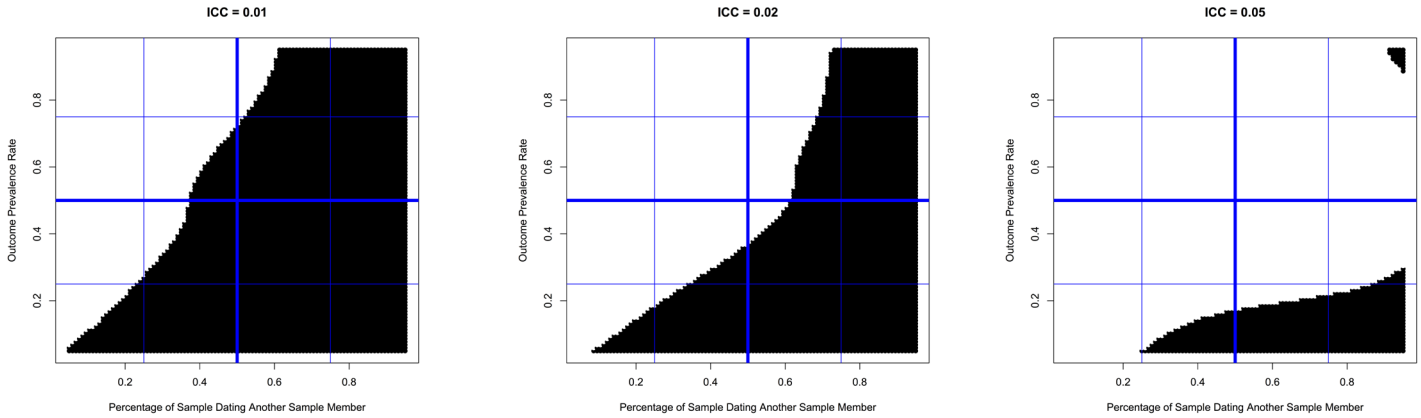
Figure A12: MDES differences between student and school randomization with 30 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
30 schools with 100 students per school
ICC of 0.02

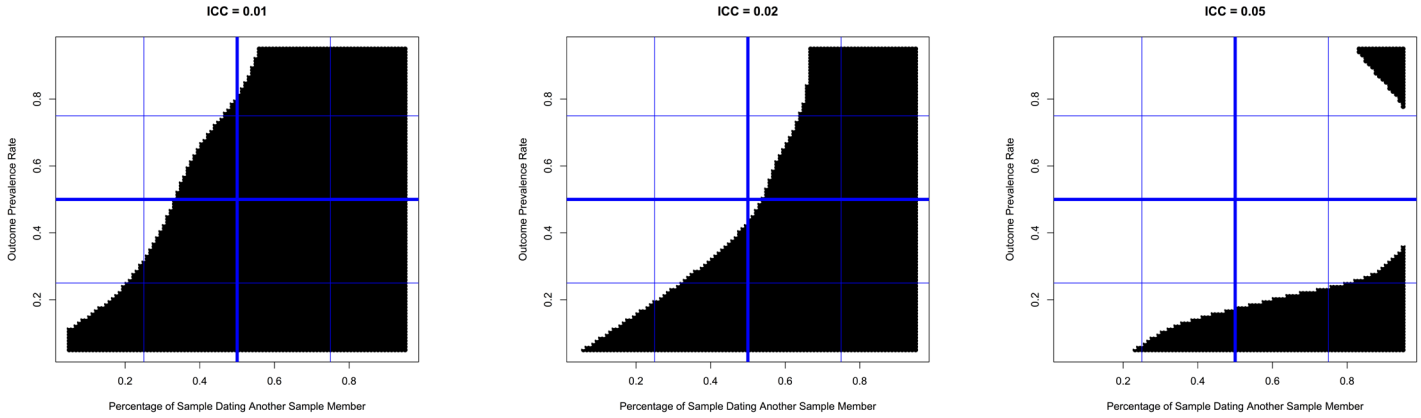
Figure A13: MDES differences between student and school randomization with 16 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
16 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

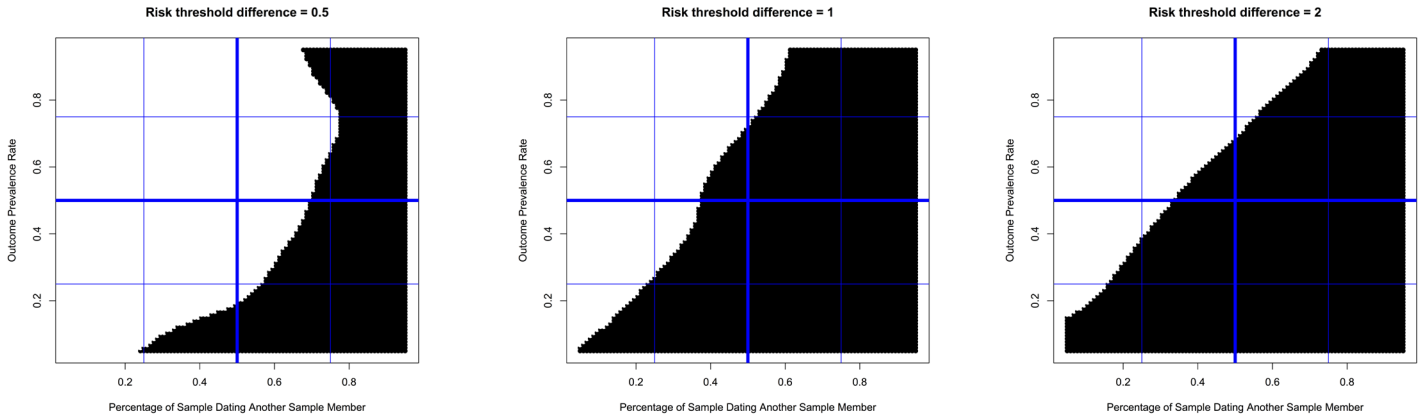
Figure A14: MDES differences between student and school randomization with 60 schools, varying ICC
(White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
60 schools with 100 students per school
There is a 1 standard deviation
difference between daters and non-daters in the threshold for engaging in risky behavior

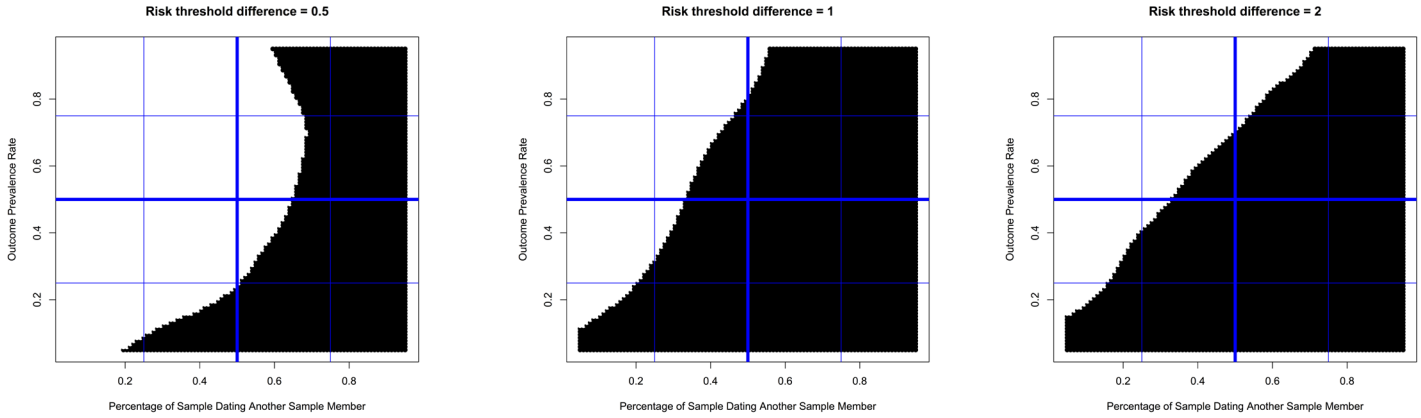
Figure A15: MDES differences between student and school randomization with 16 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of schools)



Plot A **Plot B** **Plot C**

Parameter values held constant:
16 schools with 100 students per school
ICC of 0.02

Figure A16: MDES differences between student and school randomization with 60 schools, varying the relationship between dating and risky sex (White indicates a lower MDES for randomization of individuals; black indicates a lower MDES for randomization of school)



Plot A **Plot B** **Plot C**

Parameter values held constant:
60 schools with 100 students per school
ICC of 0.02

Table A1: Estimates of Dating Prevalence, by Data Source

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Nationally Representative				
Child Trends 2015 (Monitoring the Future)	Cross-sectional survey of 8th, 10th and 12th graders	2013	Never date	60% of 8th graders
				44% of 10th graders
				38% of 12th graders
				Among 8th graders:
				53% of males
				67% of females
				63% of Whites
				54% of Blacks
				55% of Hispanics
				Among 10th graders:
				39% of males
				48% of females
				44% of Whites
				43% of Blacks
				40% of Hispanics
				Among 12th graders:
				36% of males
				39% of females
				35% of Whites
				44% of Blacks
				38% of Hispanics
			Date frequently	5% of 8th graders
				9% of 10th graders
				16% of 12th graders
				Among 8th graders:
				7% of males
				4% of females
				4% of Whites
				7% of Blacks
				7% of Hispanics
				Among 10th graders:
				10% of males
				8% of females
				9% of Whites
				7% of Blacks
				11% of Hispanics
				Among 12th graders:
				16% of males
				16% of females
				17% of Whites
				11% of Blacks
				18% of Hispanics

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Wildsmith 2013 (Monitoring the Future)	Cross-sectional survey of 8th, 10th and 12th graders	2013	Ever date	47% of 8th graders
				62% of 10th graders
				66% of 12th graders
Arcidacono 2010 (National Longitudinal Study of Adolescent to Adult Health)	Panel	1994-1999	Dating someone in their school	46% of youth in a dating relationship
			Dating someone in the same grade	42% of youth in a dating relationship with someone at their school
			Female partner is younger than male partner	40% of youth in a dating relationship with someone at their school
Carver 2003	Panel	1994-1999	Romantic relationship in the past 18 months	55% of full sample
(National Longitudinal Study of Adolescent to Adult Health)				25% of 12 year olds
				75% of 18 year olds
				53% of males
				57% of females
				Among males:
				54% of Whites
				54% of Blacks
				53% of Hispanics
				39% of Asians
				56% of Native Americans
				Among females:
				60% of Whites
				52% of Blacks
				50% of Hispanics
				42% of Asians
				56% of Native Americans
Not Nationally Representative				
Coyle 2013	Survey of urban 7th graders in California	2010-2012	Ever had a boyfriend/girlfriend	69%
			Currently have a boyfriend/girlfriend	46%
			Dated someone their own age	52% of youth that have ever had a boy/girlfriend
			Dated someone 1-2 years older	37% of youth that have ever had a boy/girlfriend
			Dated someone 3+ years older	9% of youth that have ever had a boy/girlfriend
Giordano 2010 (Toledo Adolescent Relationships Study)	Survey of youth in grades 7, 9 and 11 in Toledo	2001-2002	Currently in a relationship or have been in a relationship in the past 12 months	72%

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Furman and Hand 2006 (Toledo Adolescent Relationships Study)	Survey of youth in grades 7, 9 and 11 in Toledo	2001-2002	Have had romantic involvements in the past 12 months	32% of 7th graders 32% of 9th graders 59% of 11th graders
Connolly 2004	Survey of later elementary students in large Canadian city	unknown	Currently have a boyfriend/girlfriend	20%

Table A2: Estimates of Dating Duration, by Data Source

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Nationally Representative				
Carver 2003	Panel	1994-1999	*Among youth in a relationship*	*Among youth in a relationship* 14 months
(National Longitudinal Study of Adolescent to Adult Health)			Median Duration	12 months, males
				16 months, females
				5 months, < 14 year-olds
				8 months, 14-15 year-olds
				21 month, 16 year-olds or older
				12 months, White
				24+ months, Blacks
				15 months, Hispanic
				11 months, Asian
			In same relationship (baseline and 11 month follow-up)	51%
				44% of males
				58% of females
				22% of 13 year-olds and younger
				49% of 14-15 year-olds
				58% of 16 year-olds and older
				56% of Whites
				36% of Blacks
				49% of Hispanics
				57% of Asians
Not Nationally Representative				
Coyle 2013	Survey of urban 7th graders in California	2010-2012	In current relationship for less than one month	40% of youth in a relationship
			In current relationship for 1-3 months	26% of youth in a relationship

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
			In current relationship for 4-12 months	22% of youth in a relationship
			In current relationship for more than 12 months	12% of youth in a relationship
Furman and Hand 2006 (Toledo Adolescent Relationships Study)	Survey of youth in grades 7, 9 and 11 in Toledo	2001-2002	Median length of current relationship	4 months
Connolly 2004	Survey of later elementary students in large Canadian city	unknown	Percentage point change in the rate of dating over one academic year	2% (from 20% to 22%)

Table A3: Estimates of Sexually Risky Behavior Prevalence, by Data Source

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Nationally Representative				
Kann et al 2014 (Youth Risk Behavior Surveillance)	Cross-sectional survey	2013	Ever had sexual intercourse	47% of full sample
				46% of females
				48% of males
				44% of Whites
				61% of Blacks
				49% of Hispanics
				30% of 9th graders
				41% of 10th graders
				54% of 11th graders
				64% of 12th graders
			Had sexual intercourse before age 13	6% of full sample
				3% of females
				8% of males
				3% of Whites
				14% of Blacks
				6% of Hispanics
				6% of 9th graders
				6% of 10th graders
				6% of 11th graders
				5% of 12th graders
			Had sexual intercourse with four or more persons during their life	15% of full sample
				13% of females
				17% of males

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
				13% of Whites
				26% of Blacks
				13% of Hispanics
				7% of 9th graders
				13% of 10th graders
				19% of 11th graders
				23% of 12th graders
			Currently sexually active	34% of full sample
				35% of females
				33% of males
				33% of Whites
				42% of Blacks
				35% of Hispanics
				20% of 9th graders
				29% of 10th graders
				40 of 11th graders
				49% of 12th graders
			Condom use	59% of youth who are currently sexually active
				Among youth who are currently sexually active:
				53% of females
				66% of males
				57% of Whites
				65% of Blacks
				58% of Hispanics
				63% of 9th graders
				62% of 10th graders
				62% of 11th graders
				53% of 12th graders
			Birth control pill or IUD/ implant or Shot or Patch or Ring Use	25% of youth who are currently sexually active
				Among youth who are currently sexually active:
				30% of females
				20% of males
				33% of Whites
				15% of Blacks
				15% of Hispanics
				14% of 9th graders
				22% of 10th graders
				26% of 11th graders
				32% of 12th graders

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
			Condom AND one of Birth control pill or IUD/ implant or Shot or Patch or Ring Use	9% of youth who are currently sexually active
				Among youth who are currently sexually active:
				10% of females
				7% of males
				11% of Whites
				6% of Blacks
				5% of Hispanics
				5% of 9th graders
				7% of 10th graders
				11% of 11th graders
				10% of 12th graders
			Did not use any method to prevent pregnancy	14% of youth who are currently sexually active
				Among youth who are currently sexually active:
				16% of females
				12% of males
				11% of Whites
				16% of Blacks
				20% of Hispanics
				16% of 9th graders
				14% of 10th graders
				12% of 11th graders
				13% of 12th graders
Carver 2003 (National Longitudinal Study of Adolescent to Adult Health)	Panel	1994-1999	Ever had sexual inter-course	41% of youth in a dating relationship
				23% of 14-15 year olds
				50% of 16 year olds
			Ever touched each other's genitals	52% of youth in a dating relationship
				36% of 14-15 year olds
				61% of 16 year olds
			Ever touched each other under clothing	57% of youth in a dating relationship
				42% of 14-15 year olds
				66% of 16 year olds
Not Nationally Representative				
Coyle 2013	Survey of urban 7th graders in California	2010-2012	Ever had vaginal sex	7%
			Touched each other's private parts	10%

Data Source	Description of Data	Year of Data Collection	Measure	Dating Rate Estimate
Giordano 2010 (Toledo Adolescent Relationships Study)	Survey of youth in grades 7, 9 and 11 in Toldeo, Ohio	2001-2002	Had sexual intercourse with partner, among those currently dating	36% of youth who are currently dating
			Sexually intimate (activities beyond kissing and making out) with partner	45% of youth who are currently dating and have not had sexual intercourse with their partner