Non-Monetary Incentives for Tobacco Prevention in Indonesia*

Margaret Triyana† Justin S. White‡

July 26, 2021

Abstract

We provide evidence on the effectiveness of a school-based program that uses a non-monetary penalty and regular monitoring to prevent risky behavior among adolescents in Indonesia. The field experiment invited students to sign a pledge to abstain from tobacco use and a similar pledge for parents to monitor their children. To test group incentives, a subset of treated schools also competed against each other for the highest tobacco abstinence rates. We find that the individual pledge increases biochemically verified tobacco abstinence by 4 percentage points. This effect is sustained 3 months after the program ended. School competition has no additional impact on tobacco abstinence. Our findings highlight the effectiveness of non-monetary incentives to curb risky behaviors among adolescents who face limited self-control and peer pressure.

Keywords: Tobacco prevention, incentives, school-based intervention, risky behavior

JEL codes: C93, D91, I12, O10

*We thank Dr. Hasto Wardoyo, Dr. Baning Rahayujati, and Dr. Fatwa Sari Dewi for their support. Thanks to Gumilang Sahadewo, Astuti Kusumaningrum, Jenna Juwono, Fadjar Wibowo, Aulia Larasati, and Colley Windya for excellent research support. We are grateful for helpful comments from Alice Chen, Taryn Dinkleman, Willa Friedman, Felipe Gonzalez, Lakshmi Iyer, Julien Labonne, Jean Lee, Maria Rosales, Jessica Pan, and seminar participants at the University of Notre Dame, the Population Association of America meeting, Luksburg PUC-ND Conference, Population Health Sciences Research Workshop at Boston University, Southern Economic Association meeting, Universidad del Rosario, Universidad de Los Andes, Pontificia Universidad Catolica de Chile, University of Texas Arlington, Providence College, and Wake Forest University. We acknowledge funding from J-PAL Southeast Asia (Indonesia Research Fund), a Nanyang Technological University Start-up Grant, a Singapore Ministry of Education Tier 1 Grant, and a UCSF Research Evaluation and Allocation Committee Grant through the Simon and Ackerman Funds. Please click here for latest draft.

†Wake Forest University

‡Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco

1 Introduction

Tobacco use is associated with 6.4 million annual deaths worldwide, with nearly three-quarters occurring in low- and middle-income countries (LMICs) (Reitsma et al., 2017). Although youth smoking prevalence has declined worldwide in recent decades,[1] this decrease has occurred largely in high-income countries, and nearly 80% of the world's 41 million smokers aged 15-19 live in LMICs. With the rising popularity of nicotine-containing electronic cigarettes and vaping products, youth tobacco use has recently resurged in many countries—a trend that is causing widespread concern in the public health community.

Higher tobacco taxes are widely viewed as one of the most effective ways to reduce youth smoking (International Agency for Research on Cancer, 2011). Yet, tobacco taxation has faced obstacles to effective implementation in many high-prevalence LMICs. For example, some LMICs have a range of tobacco products, each with different price points and tax rates (e.g., bidis and cigarettes in India), providing ample scope for tax-induced product substitution (White and Ross, 2015). Many tobacco-producing countries, including top producer China, have resisted tobacco tax increases for fear of the adverse economic impacts (Chen, 2007). Both obstacles are present in our study setting, Indonesia, where clove cigarettes (*kreteks*) and higher priced conventional cigarettes are both popular and tobacco excise tax revenue accounts for 10% of total tax revenue. Indonesia also claims among the highest male smoking prevalence globally (almost 70%), has about 15% of its deaths attributable to tobacco use, and is the only country among the ten with the largest smoking population to experience an increase in youth smoking from 1990-2015 (World Health Organization, 2012; Reitsma et al., 2017). Due to policy constraints in Indonesia and many other LMICs, policymakers have a growing interest in non-tax tobacco control strategies, including prevention programs.

Motivated by the high disease burden and political resistance to tax increases in Indonesia, we test whether a non-monetary incentive program effectively prevents tobacco use among adolescents in Indonesia. Personal incentive programs have become increasingly common in high- and low-income settings to offset underinvestments in preventive health behavior (Kenkel, 1994; Dupas, 2011). Demand-side incentives for increasing health investments have been used extensively for a variety of health behaviors in LMICs (Gopalan et al., 2014). Yet, these incentive-based programs have relied largely on monetary rewards and have rarely targeted adolescents, who are at a stage when life-long habits

[1] Calculations in this paragraph rely on estimates of smoking prevalence for individuals aged 15-19, and drawn from the Global Burden of Disease Study 2015. Data are available at `http://ghdx.healthdata.org/record/ihme-data/gbd-2015-smoking-prevalence-1980-2015`. Appendix Figure B.1 shows the time trends in smoking by country income category.

often form. In the broader tobacco prevention literature, three randomized trials (total $N = 1,108$) have tested incentives for tobacco prevention among adolescents (Hefler et al., 2017), although Cochrane judged them to provide low-quality, unreliable evidence.[2] These studies tested the Smokefree Class Competition program, which has been widely implemented in Europe. In the competition, classes commit to being smoke-free and compete to win prizes in a lottery. Yet, all of these programs have been fielded in high-income settings.

To address this gap, we implemented a school-based field experiment of 2,700 students from 72 middle schools in Indonesia to test a tobacco prevention program with non-monetary incentives. In the treated sample of 36 schools, 7th grade students were invited to publicly sign a commitment pledge in which they agreed to abstain from tobacco use, and their parents were invited to privately sign a pledge to monitor their children for tobacco use at home. Students received disciplinary action for each infraction and parental notification for multiple infractions, a severe penalty in this setting. We designed the incentives to counteract social norms and peer pressure to smoke and to serve as a partial commitment contract for adolescents who suffer from poor self-control.[3] The intervention attempted to engage adolescents' various social networks—teachers (school), parents (family), and student peers (friends)—to monitor and enforce the students' actions. Throughout a 5-month intervention period, all 72 schools collected biweekly reports from students and their teachers in 7th grade to identify smokers and create a sense of accountability and salience for the anti-tobacco intervention.

We introduced school-level competition in a randomly selected 18 of the 36 treated schools to test how effectively team incentives that emphasize group identity ('school spirit') mobilize peer pressure and support to improve health behavior. These 18 schools received a combined intervention consisting of the student and parent pledge intervention plus biweekly feedback that congratulated them if their tobacco abstinence rate ranked in the top 10 among participating schools. At the end of 5 months, the local government officially recognized the top-performing schools with a plaque.

We conducted an initial follow-up at 3 months to evaluate the immediate program effects and a longer-term follow-up at 8 months to evaluate the sustainability of program effects

[2] Cochrane reviews follow a systematic approach to rating the certainty of evidence syntheses, assessed by two independent reviewers (Higgins et al., 2011). Rating domains include risk of bias (e.g., selection bias, attrition), imprecision of estimates, inconsistency of evidence, and publication bias. In our case, studies suffered from selection bias (75% of included RCTs), attrition (67%), and wide confidence intervals (Hefler et al., 2017).

[3] The incentives may be viewed as a "partial" commitment contract, because, given the power differential between students and their parents and teachers, we cannot assess the extent to which students entered them voluntarily.

into the next school year. Using an intent-to-treat approach, we find that biochemically verified tobacco abstinence in the treated group increased by 6% (4 percentage points). The intent-to-treat effects of the individual pledge and the pledge plus school competition remain similar in magnitude until 8 months (3 months after biweekly monitoring had ceased). Thus, we conclude that the individual pledge drives the program effect. We fail to find evidence in support of the hypothesis that school competition enhances the pledge intervention by harnessing social effects, including emphasizing group identity and social sanctions from peers. This finding is consistent with randomized trials of class competitions in Europe (Hefler et al., 2017).

We further explore the role of several potential mechanisms to explain our pattern of findings. Potential mechanisms include aspects of peer influence such as fear of social sanctions, social learning, conformity to and perceptions of smoking norms, and group identity, as well as incentive effects including fear of punishment and regular monitoring, increased motivation and effort, overcoming time-inconsistency, and increased salience. We find some evidence consistent with changing smoking norms. We also find heterogeneity in program effects by our measure of time-inconsistency, suggesting that the intervention may help students to follow through on their plans to abstain from smoking.

Our study contributes to a growing literature in behavioral and health economics on the use of incentives to improve health behaviors. According to recent reviews (Mantzari et al., 2015; Giles et al., 2014; Gopalan et al., 2014), incentive programs have been shown to be effective in changing a number of health behaviors. We, however, depart from the existing incentives literature in four important ways. First, the health literature has focused largely on adults, providing little guidance on the effectiveness of incentives among adolescents. Because the adolescent period coincides with the development of executive brain function that regulates impulse control, long-term planning, and the ability to resist peer influence (Steinberg, 2007; Zelazo and Carlson, 2012), healthy habit formation at this stage is crucial for long-term outcomes. Psychologists have attributed increased risk-taking during adolescence to a heightened sensation seeking (Steinberg et al., 2018), and we speculate that the same neural reward processes that make adolescents susceptible to risky behavior like smoking may make them responsive to incentive-based interventions. Studies also find that adolescent smoking participation responds to cigarette tax changes, pointing to the potential for incentives, whether monetary or non-monetary, to influence behavior (Lillard et al., 2013; Carpenter and Cook, 2008).[4] A better understanding of the impacts of

[4] Although the literature is mixed, theory predicts that youth smoking demand is more elastic than for adults due to the addictive nature of tobacco and fewer financial resources in youth (Chaloupka and Warner, 2000; DeCicca et al., Forthcoming).

incentives on adolescent risk-taking may improve policy design and implementation at a critical stage in habit and preference formation.

Second, whereas demand-side health incentives have relied largely on monetary rewards, our program employs a non-monetary incentive. Health-related commitment contracts have also typically been structured to put money at risk (Bryan et al., 2010; Giné et al., 2010; Halpern et al., 2015; Royer et al., 2015; Bai et al., Forthcoming; Schilbach, 2019).[5] Since our study involves adolescents who have limited financial resources, we employ a non-monetary penalty (school disciplinary action and parental notification). This penalty also lowers the financial cost of the intervention and improves the chances that it is sustainable, culturally acceptable, and scalable.[6]

Third, health-related incentive programs tend to be individual in nature. Social incentive schemes, however, may be able to mobilize peer influence as a form of social pressure and social support. The literature has found that team incentives for physical activity and weight loss can outperform individual-based incentives (Kullgren et al., 2013; Babcock et al., 2015; Patel et al., 2016), though the evidence for smoking cessation is less encouraging (Halpern et al., 2015; White et al., 2020). We also extend the degree of social accountability by engaging key actors in the student's social network (parents, teachers, and peers) to monitor and enforce the individual pledge to avoid tobacco use. Risky adolescent behavior is heavily influenced by adolescents' social networks, most notably their parents and youth peers whose modeled behavior is often reproduced (Bandura, 1986). Since parents and peers play an important role in influencing adolescents, they also have the potential to play a key role in mitigating risky behavior, particularly tobacco use (Lundborg, 2006; Trogdon et al., 2008; Ennett et al., 2001; Kremer and Levy, 2008; Card and Giuliano, 2013).

Fourth, we implement the incentive-based tobacco prevention program in a middle-income country. Nearly all studies on risky behavior in adolescence are drawn from high-income settings. A recent meta-analysis of 49 randomized controlled trials for smoking prevention among adolescents demonstrated that school-based interventions reduced smoking initiation by 12% (Thomas et al., 2013). However, only four of the 49 studies included in the analysis took place in lower-income countries: two studies in China, one in South Africa, and one in Thailand. Moreover, all studies of incentives for smoking prevention in adolescents were conducted in high-income countries (Hefler et al., 2017).

[5] Commitment contracts are targeted to individuals who display time-inconsistent preferences. Demand for commitment can be rationalized under several different theories of behavior, including quasi-hyperbolic discounting, choice-set-dependent utility, and dual-self models (Bryan et al., 2010). We remain agnostic about the model underlying the preferences of adolescents in our study.

[6] During preliminary discussions, local Indonesian government officials felt that a monetary intervention was not acceptable for adolescents.

Interventions in high-income countries may not generalize to LMICs due to their limited resources, higher prevalence of smoking, and cultural and institutional differences.

The remainder of the paper is organized as follows. Section 2 provides an overview of the experimental design. Section 3 describes the data and methods used to estimate the intervention's effects. Section 4 presents our findings, followed by robustness in Section 4.5. Section 5 concludes with policy implications.

2 Experimental Design

We implemented a cluster randomized controlled trial, using schools as the unit of randomization, in order to understand the effects of non-monetary incentives and school competition on tobacco use among middle-school children in Indonesia. We used schools as the unit of randomization to minimize contamination effects among classmates and to increase perceived acceptability by school officials. Our pre-analysis plan is available in the AEA RCT Registry (AEARCTR-0001607).

2.1 Study Setting

The Indonesian government's reliance on the tobacco industry as a source of revenue and employment has limited the country's tobacco control strategies. The tobacco industry accounts for about 10% of total tax revenue and is the second largest employer in the country (Barber et al., 2008).[7] As such, the industry wields considerable political and financial clout that has been used to thwart tobacco taxation (Assunta and Dorotheo, 2016). Non-price regulations of tobacco use, such as age restrictions, have also been limited in Indonesia. Indonesia's reluctance to regulate tobacco is illustrated by its status as one of the few nations not to have ratified the Framework Convention on Tobacco Control, a World Health Organization treaty that establishes binding regulatory standards for price and non-price measures on tobacco use.

While Indonesia has among the highest male smoking prevalence globally, early smoking initiation is still understood to be a risky behavior, even among Indonesian youth (Ng et al., 2007a). Research shows that nicotine dependence often begins during adolescence (DiFranza et al., 2007), with initiation typically ranging between the ages of 15 and 19 (Thomas and Perera, 2006; World Health Organization, 2012; Lillard et al., 2013). According to recent national data, adolescents in Indonesia report first smoking between the ages of 12 and 13,

[7] The revenue estimate is from: `https://nasional.kompas.com/read/2018/03/20/23224701/penerimaan-cukai-tembakau-terus-meningkat`, last accessed September 27, 2018.

and most children aged 13-15 who smoke report purchasing cigarettes from vendors despite a legal prohibition (World Health Organization, 2015a). Many vendors sell cigarettes by the stick, which lowers the costs of experimentation and use, and tobacco retailers are often located near schools (Astuti et al., 2019). Data show that 34% of all middle-school students in Indonesia currently smoke, and 56% have tried smoking (World Health Organization, 2015a). In comparison, only 7% of middle-schoolers in the U.S. report having used any tobacco products (Jamal et al., 2017). These numbers highlight adolescence as a critical period for smoking prevention in Indonesia.

The wide-ranging health risks associated with tobacco use tend to be downplayed in Indonesia. For instance, it is widely believed that smoking fewer than 10 cigarettes per day is not harmful insofar as the smoker finds a brand of cigarettes considered ”suitable” for his body (Nichter, 2006). Even among health professionals in Indonesia, more than 80% of physicians believe that smoking fewer than 10 cigarettes per day is not particularly harmful for health (Ng et al., 2007b). Because smoking is deeply ingrained in Indonesian culture and national political opposition to tobacco use is limited, sub-national policy may offer a promising approach to regulating tobacco use. Since 2005, local governments have had the authority to issue district regulations, and district leaders have had the power to issue local executive orders. The devolution of policy making opens up opportunities for local actors to experiment with tobacco control policy. Therefore, Indonesia’s high smoking rates and policy environment provide an ideal setting to test innovative tobacco control strategies.

We targeted more rural communities of Indonesia, where tobacco use tends to be higher (World Health Organization, 2012). Our study was implemented in 19 subdistricts in two adjacent districts in Yogyakarta Province: Kulon Progo and Sleman. Kulon Progo has adopted certain restrictive tobacco control policies, including banning tobacco advertising, while Sleman has been less proactive perhaps because it is a tobacco-growing area. Our intervention was initially designed for Kulon Progo, but we incorporated subdistricts in Sleman that are geographically proximate and similar to Kulon Progo in their socio-economic characteristics to meet the sample size requirement. We excluded the primary tobacco-growing areas in Sleman from our sample because of concerns they would differ in important ways.

At the outset of the study, we partnered with local government agencies in each district to help them implement a tobacco prevention program in schools. Our intervention design was informed by extensive consultations with local government officials to ensure its acceptability and sustainability. Specifically, local stakeholders discouraged the use of monetary rewards because it might be negatively perceived as gambling in school and because parents typically supervise their children’s financial resources.

Female smoking prevalence is low, at about 2%, in Indonesia (World Health Organization, 2017). Among males, in a nationally representative survey, about 36% of smokers in Kulon Progo and 45% in Sleman began smoking between 15 and 19 years of age, while 13% and 9% began smoking between the ages of 10 and 14 years in Kulon Progo and Sleman, respectively (Ministry of Health, Indonesia, 2013). Therefore, our study targeted 12- to 13-year-old male students in the 7th grade (the first year of junior secondary school, *Sekolah Menengah Pertama*). Most students who smoke use kreteks, which are the most commonly found form of tobacco in Indonesia, followed by regular cigarettes and e-cigarettes.[8]

Schools in our sample have an average of four classes per grade, and each class has a maximum of 42 students. The majority of Indonesian schools, including many schools in our sample, employ a disciplinary system wherein teachers issue demerit points to students who misbehave. While administered separately from the existing system, our penalty was framed as being similar in structure to ensure that students understood the nature of the penalty.

2.2 Interventions and Procedures

Figure 1 summarizes the study design. The study included a total of 72 schools: 36 in Kulon Progo and 36 in Sleman. We initially identified 78 schools, but 6 were excluded due to their remote location and small school size.

We used pair matching to randomly assign 36 schools to the control arm, 18 schools to the individual pledge arm, and 18 schools to the pledge plus school competition arm. Studies have shown that pair matching performs well relative to other randomization methods in balancing arms (Imai et al., 2009; Bruhn and McKenzie, 2009). We formed school pairs by minimizing the Mahalanobis distance between covariates values within pairs. The following covariates were used: district, subdistrict, distance from Yogyakarta city, school size, and average national examination score in mathematics as a proxy for school quality. In each pair, one school was randomly assigned to the control and the other to the treated arm. We then paired treated schools and randomly assigned one to each sub-treatment group. We conducted sample size calculations based on pairwise comparisons between equal-sized study arms. Based on previous tobacco prevention studies, our study was powered for a minimum detectable effect of 5 percentage points (Thomas et al., 2013). Further details on the sample size calculations and randomization procedure are available in Appendix A.

We screened a total of 3,031 students, 248 of whom did not return an informed consent form and were excluded. Thus, we randomized 2,783 students and allocated 1,346 students

[8] Kreteks are likely at least as harmful as regular cigarettes, according to a recent systematic review (Nuryunarsih et al., 2021).

to the control arm, 683 students to the individual pledge arm, and 754 students to the pledge plus school competition arm. All students who consented signed the pledge agreement.

Figure 2 describes the study timeline. The study started in January 2017, corresponding to the second half of the academic year 2016-17. We tested participants for cotinine, a metabolite of nicotine, using a urine assay to biochemically verify tobacco abstinence at baseline. The cotinine test has a detection window of about 2 to 3 days after last nicotine use, depending on quantity smoked (Benowitz et al., 2020). During initial focus group discussions, parents and teachers showed their strong support for the urine test and none of the students objected. The urine test is a rapid, 10-minute test that was performed uniformly at each school. While neither the school nor students were told the individual test results, schools were given the aggregate number of positive cotinine tests.

We then collected biweekly reports on smoking status from all treated schools between January and May 2017. Schools were told that there would be at least one unannounced audit, which was conducted in all 72 schools in March 2017 (3 months after the program started). The aim of the audit was to validate the biweekly reports in treated schools. The audit included biochemical verification of smoking status for about 80% of students. After the audit, biweekly reporting continued through May 2017 when the national examination period for all schools began. The academic year ended in June 2017, and the new academic year began in July 2017. Another round of biochemical testing was conducted in August 2017, about 8 months after the intervention had started and 3 months after regular monitoring had ceased.

Each 7th grade class in all participating schools received an information campaign at baseline on the risks of smoking, comprised of one 45-minute session led by a trained facilitator. Schools in the control group received the information session only, with no additional parental or school involvement. The control schools continued to enforce any pre-existing disciplinary policies, which often included banning smoking on school property.

In the 36 treated schools, at the end of the information session, students were invited to sign a commitment pledge to abstain from tobacco use during the intervention period. After students signed their individual pledges, they were asked to sign an additional group pledge as a class. Before the information session, students were asked to give their parents a letter outlining the program and an invitation to sign a separate, similar pledge agreeing to monitor their children to prevent them from using tobacco products.[9] A student who violated the terms of the pledge, as measured by a biochemically verified cotinine test or student and teacher reports, was penalized according to the school-based demerit system.[10]

[9] The student and parent pledges are provided in Appendix Figures B.2 and B.3.

[10] Some schools had a pre-existing demerit system. In such cases, we kept our demerit system separate to

Upon each of the first two infractions, the student was summoned to the guidance counselor, who issued a warning, reminding the student of his pledge and urging him not to smoke in the future. Upon the third and subsequent infractions, the student was summoned to the office of the principal, who then called the student's parents to report the infraction. Teachers and administrators told us during pilot work that this penalty would be sufficiently severe to encourage good behavior.

Guidance counselors, classmates, and members of the student council were all tasked with monitoring and reporting students' smoking behavior in treated schools. Most schools in Indonesia have student councils, which are elected each academic year. Students from the 7th, 8th, and 9th grades elect council members, who are typically required to maintain good academic and disciplinary standing. Student councils are comprised of older (8th and 9th grade) students who are widely respected by the other students. Due to their conscientiousness, this group tends to have low smoking rates. We trained members of the student council to promote smoking abstinence among participating students.[11] Under the supervision of the guidance counselor, the student council received, investigated, and recorded reports of students who smoked. The reports were then kept by the guidance counselor who evaluated the reports, followed up when necessary, and issued the demerit penalty. The reports were kept private except in instances when students appealed.[12]

Every two weeks, student council members assisted the 7th grade students to complete a form indicating any students who they observed smoking on or off school premises. These biweekly reports lent the intervention a sense of accountability and salience. All students provided reports at the same time in order to reduce under-reporting from less motivated students and to avoid bullying of reporters by classmates. Anyone on the school premises was also able to submit a report in a comment box posted at the school, and students could also report cases directly to the student council, teachers, or our program administrators via email, text message, or phone call.

standardize the program. We tracked demerits separately, did not share our data with the administration, and did not punish infractions with school suspension or expulsion, as occurs under some school systems. Prior to our intervention, tobacco use typically resulted in a warning from the guidance counselor, and in schools with a pre-existing demerit system, the demerit points for smoking were typically lower than for other disciplinary infractions. All 7th grade students in treated schools were eligible to be reported within our demerit system regardless of whether they had signed the pledge. However, only those who signed the pledge received a penalty for non-compliance.

[11] Based on our pilot work, providing clear instructions and involving the student council seemed a more feasible approach to tracking student behavior than relying on teacher reports alone. Teachers and guidance counselors are often overwhelmed by other obligations, and do not always have time to handle the added burden of monitoring and reporting students' tobacco use for the study.

[12] An accused student could appeal to the guidance counselor, who would investigate. The guidance counselor could learn the reporter's identity as part of the appeals process. No appeals were made during the study period.

The random audit occurred 3 months after the program started. Schools were given less than 48 hours' notice and were asked not to notify students of the visit. If a school had fewer than 48 participants, we attempted to test all students from that school during the audit. In larger schools, we randomly selected 48 students to test, due to the limited budget and staffing capacity.

School performance in the pledge plus school competition arm was scored based on several data sources: the biweekly reports, the urine test results during the unannounced 3-month audit, and the urine test results at 8 months.[13] A school's score took the following into account: timeliness of reporting, completeness of reporting, and the proportion of students who were nicotine-free.[14] We provided a list of the top 10 ranked schools to student council members on an ongoing basis. The districts' Departments of Education issued certificates of recognition to the schools that finished in the top 10—the same manner in which they recognize top-performing schools in sports. Thus, the recognition was not accompanied by any additional funding, ensuring that the program would be low-cost, which was a priority for local government officials.

When the new academic year began in July 2017, students were quasi-randomly assigned new classmates when they entered 8th grade. Middle schools in Indonesia assign students to a specific homeroom class for the entire academic year. Students remain in the same classroom throughout the day, while the subject teachers rotate through. In schools with more than one class per grade, administrators assign students in each grade to different homerooms every academic year to allow students to meet new peers. We leverage this source of variation in our analysis of peer effects when the students are in 8th grade at the 8-month follow-up. Specifically, we use as an instrument the variation in the number of 7th grade friends who are assigned to the same 8th grade class. Biweekly reporting could not resume at the beginning of the new academic year because the student council had not

[13] One potential concern is students in the pledge plus competition arm have an incentive to under-report smoking. We find similar rates of reporting across the individual pledge arm and the school competition arm at 3 months and 8 months, after biweekly reporting had ended. Figure B.4 shows the time trend in students reported by sub-treatment. Differences in the number of reported students between the individual pledge and pledge plus school competition arms are not statistically significant. Moreover, our main outcomes do not rely on the biweekly reports due to potential misreporting.

[14] We used the following formula to calculate a biweekly score for each school based on biweekly reports:

$$\text{Score} = \%\ \text{non-smokers} - \%\ \text{follow-up} - 5 \times \%\ \text{late follow-up} - 10 \times \%\ \text{no follow-up}$$

where % non-smokers is the percentage of students who did not smoke, % follow-up is the percentage of reported smokers whose cases were adjudicated within a week, % late follow-up is the percentage of reported smokers whose cases were adjudicated between 1 and 2 weeks, and % no follow-up is the percentage of reported smokers whose cases were not adjudicated within two weeks. We also explained to the student council and guidance counselor that we would check the reports during random audits, and there would be a penalty of 50 points for falsified reports.

yet been elected for the new year. We used biochemical verification during the 8-month assessment in August 2017 to evaluate the persistence of the program effects after regular monitoring had ended.

Hypothesized mechanisms Our study proposed two key hypotheses. First, we expected that the individual pledge intervention acting as a partial commitment contract would reduce tobacco use among students compared with receiving information only. We expected that a combination of factors would lead the non-monetary incentives to curb smoking, including monitoring from parents and teachers, fear of the in-school remediation for tobacco use, and greater ability to exercise self-control. Second, we hypothesized that school competition combined with a pledge would reduce tobacco use among students, compared with a pledge only or receiving information alone. The competition emphasized a spirit of teamwork, and we expected this to increase the salience of group members' identity as non-smokers, thereby creating peer pressure not to smoke and engaging students to establish a norm against smoking.

We expected several potential intervention-driven mechanisms to influence participant behavior, while acknowledging that data availability and the experimental design limit our ability to distinguish among mechanisms. It is worth emphasizing that the mechanisms marshal a combination of positive reinforcement (e.g., building a smoke-free group identity) and negative reinforcement (e.g., threat of parental punishment) to motivate the students to abstain from tobacco.[15] We classify the potential mechanisms under the broad categories of peer influence and incentive effects.

Peer influence on adolescent smoking behavior has been well documented, perhaps not surprisingly because smoking is a social activity during adolescence.[16] We examine the reduced-form effect of any peer influence using the quasi-random assignment of participants to new classmates in 8th grade. We then explore several types of peer influence: (i) a fear of social sanctions from peers and pressure to comply, (ii) group identity, (iii) social learning, and (iv) changing perceived social norms and conformity to

[15] The use of positive and negative reinforcement raises concerns of unintended effects. Extrinsic motivators (positive or negative) have been theorized to crowd out intrinsic motivation (Ryan and Deci, 2000), prosocial behavior (Bénabou and Tirole, 2006), and trust (Herold, 2010). The first is unlikely in our context because program effects persist beyond year's end. Our data do not allow us to test crowding out of trust or prosocial behavior, but some research finds that peer monitoring and trust are interlinked for students (Rotsaert et al., 2017). Second, there are concerns that the program leads to student bullying. We did not receive any bullying reports from students or teachers, although it remains a potential risk in such programs.

[16] In the economics literature, studies of peer effects for adolescent smoking include: Ennett et al. (2001); Lundborg (2006); Trogdon et al. (2008); Kremer and Levy (2008); Fletcher (2010); Card and Giuliano (2013); Fletcher and Ross (2018). Many public health studies have also found peer influence to be a determinant of adolescent smoking (e.g., Valente et al., 2003; Huang et al., 2014).

those norms. We posit that fear of social sanctions or pressure to comply and group identity would be differentially activated by the school competition, because the competition may lead students to worry about the social consequences of being caught smoking and to feel a part of a group defined by being smoke-free. Thus, greater abstinence in the pledge plus competition group, compared with the pledge alone, might suggest a role for these mechanisms. Peer influence has been suggested to operate through social learning, in which we expected treated students to gain relatively more knowledge about why smoking is bad and how to avoid smoking (Akers and Lee, 1996). Finally, the treatment is assumed to influence social norms by denormalizing smoking within the peer group, and treated students may perceive smoking to be less acceptable and may be more willing to opportunistically conform to smoking norms (Huang et al., 2014; Charness et al., 2019). Indicators of changing norms include how often students smoke with friends and whether a student's tobacco use is affected by friends' smoking status.

We further expected that smoking may be influenced by several mechanisms within the broad category of incentive effects, i.e., behavioral responses to the experimental payoff structure: (i) fear of punishment from the school or parents and the associated regular monitoring, (ii) increased motivation and effort not to smoke, (iii) increased ability to exercise self-control and overcome time-inconsistency, (iv) increased salience of smoking decisions, and (v) habit formation. First, the threat of being disciplined at school or at home (including due to the biweekly reporting and unannounced audit) may have pressured students into compliance. Any program effects due to monitoring or punishment should fade after the program ends, i.e., by the 8-month follow-up. We can also assess whether students believe that program features like the urine spot checks and parental notification affect their behavior. Second, both the commitment pledge and school competition may enhance motivation and effort not to smoke, for example, leading students to be more likely to intend not to smoke or to report friends who smoke during the biweekly reporting. Third, the commitment pledge may help students who have limited self-control to overcome the temptation to smoke (Bryan et al., 2010). Fourth, the program may raise the salience of not smoking through the incentives and regular monitoring. As with the monitoring effects above, we would expect a salience effect to fade after the program ended. The program also may be more salient if students discuss smoking more frequently with parents due to parents reminding their children to avoid smoking. Conversely, we expected increased salience to be associated with less frequent discussion with friends who would be less likely to invite their friends to smoke (Ng et al., 2007a). Fifth, the program may spur students to develop a habit of not smoking, which might continue even after the incentives are removed (Dupas, 2014).

3 Data and Estimation Strategy

3.1 Data Sources

Our study drew on three main data sources: administrative data on schools, participant surveys, and biochemical test results. The administrative data include several school characteristics, such as the number of classes. We collected self-administered survey data from several types of respondents in the 72 schools: male students in 7th grade and their parents; each school's guidance counselor, 7th grade teacher, and principal; and members of the student council. The biochemical tests of smoking status were collected in 3 waves: at baseline, during the 3-month audit, and at the 8-month follow-up. Figure 2 describes the timing of data collection. The baseline survey was conducted in January 2017 (wave 1), an unannounced audit was conducted in March (wave 2), and a follow-up survey was conducted in August at the beginning of the new academic year (wave 3).

Outcomes Our primary outcome is the proportion of students who abstained from smoking at each follow-up assessment. We relied on several methods to assess the smoking status of students. First, we conducted biochemical verification of smoking status using urine cotinine tests. We used the COT One Step Cotinine Test, an immunoassay that detects cotinine in urine at concentrations $\geq$ 200 ng/ml. According to the package insert, the test is highly sensitive at 100 ng/ml, which can detect consumption of 1 cigarette per day, on average (Shiffman et al., 2014). The window of detection is typically up to 2 to 3 days after nicotine use, depending on the intensity of tobacco use (Benowitz et al., 2020). The cotinine test would detect the use of any products that contain nicotine, including cigarettes, e-cigarettes or vaping products, or smokeless tobacco. The urine tests were conducted at baseline, during the unannounced audit after 3 months of implementation, and at the 8-month follow-up. Second, students provided self-reports of their smoking status as part of each survey round conducted at the same time as each urine test. Third, during the 3- and 8-month follow-ups, student council members, teachers, and classmates in control and treated schools were asked to report the names of students from the school who they had observed smoking within the prior 7 days.

We combine the cotinine results and self-reports of smoking abstinence during the prior 7 days to generate a combined smoking status variable known as "7-day point-prevalence abstinence," as is standard in the smoking literature (West et al., 2005). Specifically, biochemically verified abstinence takes the value one if a student self-reported smoking abstinence and obtained a negative cotinine test. If a student reported smoking abstinence, but obtained a positive cotinine test, we code the student as failing to abstain from smoking. Since the cotinine test only detects smoking up to 2-3 days prior, if a student self-reported smoking during the prior 7 days, but obtained a negative cotinine test, we code the student as failing to abstain from smoking. If a student self-reported smoking during the prior 7 days or obtained a positive cotinine test, we code the student as failing to abstain from smoking. For the main analysis, we use all available outcome data, consistent with common practice in the smoking literature (Hall et al., 2001).

We perform several robustness analyses (Section 4.5). We test for selection into biochemical verification by treatment status as students may employ multiple strategies for hiding their own or their friends' tobacco use. Specifically, we apply our estimation strategy to three outcome measures to assess different forms of selection by treatment status: 1) the share of students who had a self-report but missing cotinine test results, for those who refused or skipped the biochemical verification (13.2% in the control group at baseline), 2) the share who self-reported abstinence but failed the cotinine test (7.8% in the control group at baseline), for those who tried to hide their tobacco use, and 3) the frequency of being reported by peers, for those who under-report their smoking behavior. Students in control and treated schools were asked to list friends who smoked during the 2 weeks prior to the survey in waves 2 and 3 (on average, a smoker in the control group received 6.1 reports in wave 2). We further test the robustness of our main results by conducting sensitivity analyses that use alternative definitions of smoking abstinence. The first alternative outcome assumes students with missing cotinine data were smokers; the second assumes those with missing cotinine were non-smokers unless they were observed or reported to have smoked at some point during the trial; and the third relies on the cotinine test only. The fourth and fifth alternative definitions combine the cotinine test and self-reported smoking status during the prior two weeks and prior one month (Section 4.5).

Covariates The administrative school data include several school characteristics, such as the subdistrict, number of 7th grade classes, and distance from Yogyakarta city. The baseline school survey includes characteristics such as the number of teachers who smoke in the school, the presence of disciplinary action against smoking, and the average national examination scores in mathematics as a proxy for school quality since there is no official school ranking. We also include questions on the size of the student council because these students assisted in the study implementation. The teacher and principal surveys include a series of nine questions to assess their tobacco knowledge and the six component questions of the Fagerström Test of Nicotine Dependence scale among current smokers (Heatherton et al., 1991). Further details are included in Appendix A.

The student survey includes information on the student's mathematics score on the national examination, self-reported closeness to parents, and family's asset ownership, which include: radio, television, cable TV, refrigerator, bicycle, motorcycle, car or motor boat, mobile phone, poultry, and livestock. The student survey also includes time and risk preferences, for which students were given hypothetical waiting times for monetary gifts and hypothetical gambles to elicit their time-inconsistency and preference for certainty. Students were also asked whether they believed their parents would find it (very) unacceptable if the student smoked to capture students' beliefs about their parents' attitude towards smoking. We also include self-reported truancy, defined as the number of days of unexcused absence in the 7 days prior to the survey, as a related disciplinary outcome and risky behavior.[17] As an additional covariate, we use students' verified tobacco abstinence at baseline to calculate the class average after excluding the index student's own outcome.

The parent survey includes household size, home ownership, land ownership, the occupation of the household head, the number of smokers in the household, and whether at least one parent smoked, typically the father. The occupation of the household head is classified as agriculture, non-agricultural self-employment, government employee, blue-collar work, and non-agricultural private employment. We further define skilled occupations as being self-employed outside of agriculture, a government employee, or privately employed outside of agriculture. To measure parents' smoking attitudes, we include an indicator that takes the value one if parents would find it (very) unacceptable if their child smoked.

3.2 Baseline Sample Characteristics

Overall, baseline characteristics are well balanced across randomization groups (Table 1). Panel A describes school characteristics; Panel B describes student characteristics; Panel C describes household characteristics; and Panel D describes smoking abstinence outcomes at baseline. We calculate adjusted differences between control and treated groups by including district and school pair fixed effects (Columns 5 to 7). We compare the control group to the pooled treatment (Column 5) and each sub-treatment (Columns 6 and 7).

Schools in our sample are similar across arms (Panel A). The schools are, on average, about 17 to 18 kilometers from the provincial capital of Yogyakarta. The average student-to-teacher ratio is 14 to 15 in the sample. All of the schools in our sample have student councils, comprised of about 10 students, with a median of 6. For academic performance, schools score an average of 52 to 57% on the national mathematics examination. Schools

[17] We find no effect of the intervention on truancy as a related risky behavior.

in the control group have about 1.4 teachers who smoke, while those in the treated group have about 2.0, a small but statistically significant difference. The average (and median) number of classes in the school is 12, for an average of 4 classes per grade. The differences between the control, treated, and sub-treatment groups on school characteristics are not jointly significant.

Student baseline characteristics are also well balanced across the control and treated groups (Panel B). Nearly three-quarters of students had ever smoked by the time they reached 7th grade, which is higher than the national average in 2015 but consistent with the increasing trend in youth smoking in Indonesia (World Health Organization, 2015c). Students' tobacco knowledge was fairly high; they correctly answered about 6 of the 9 questions about tobacco, on average. Students reported tobacco use among their peers, with an average of 1.5 of their 5 closest friends currently smoking at baseline, with a median of 1. In terms of other risky behaviors, about 5% of students had been truant (missed class) during the prior week. Almost 70% of students exhibited time-inconsistency. About 34% of students in the control and 44% in the treated group had a preference for risk-seeking, and this difference is statistically significant. At baseline, about 25% of students believed their parents would find smoking among children (very) unacceptable. The differences between the control, treated, and sub-treatment arms are small and not statistically significant for most comparisons, and the differences for student characteristics are also not jointly significant. Household characteristics are similar across the control and treated groups (Panel C). The average household has about 5 members, with a small but significant difference between the control and individual pledge arms. About half of households in the sample own land, one proxy for wealth. About one-third of household heads work in a skilled occupation. The average number of smokers in the household is 0.7, and about half of students have at least one household member who smoked, most likely the father. About 40% of parents would find their child smoking to be (very) unacceptable, in contrast to students' belief of about 25% (in Panel B). Households have an average of 6.5 of 10 assets. Differences across arms by household characteristic are significant only for one comparison—household size in the pledge versus control arms, with a difference of modest magnitude. The baseline differences for household characteristics are jointly significant for the pledge versus control arms only, driven by the modest differences in household size.

Students' outcomes are similar across arms at baseline (Panel D). Our main outcome is biochemically verified smoking abstinence during the 7 days prior to the survey, defined as self-reporting smoking abstinence and obtaining a negative cotinine test. About 78% of students abstained from smoking during the 7 days prior to the baseline survey, and the differences across study groups are not significant. This implies 22% of students had smoked, which is alarming, but consistent with national estimates that show 23% of 13- to 15-year-old boys smoke (World Health Organization, 2015b). The rate of smoking abstinence is similar under alternative definitions of smoking abstinence, ranging from 67% to 85%, and the differences across study groups are not significant.

3.3 Estimation of Treatment Effects

Our key empirical models test whether the randomized interventions affect tobacco abstinence at each follow-up. Our main outcome variable, Y_{ist}, is biochemically verified 7-day smoking abstinence, which combines self-reports with the cotinine test results for student i in school s at time t. To increase statistical power, we employ a difference-in-differences strategy to estimate the program effects by estimating the following equation:

$$Y_{ist} = \alpha + \rho Post_t + \tau T_s + \delta(T_s \times Post_t) + \gamma X_{is0} + \varepsilon_{ist} \quad (1)$$

where $Post_t$ is an indicator variable equal to 1 for the follow-up period (at 3 or 8 months), T_s is an indicator variable equal to 1 if the school is randomly allocated to a treated group. We include school pair and wave fixed effects. When wave fixed effects are included, the indicator $Post$ is replaced with wave fixed effects. All standard errors are clustered at the school level to account for intra-cluster correlation of outcomes. For outcomes that are available in wave 3 only, we include the treatment indicator and baseline covariates. Similarly, we conduct the following sub-treatment analysis using the following equation:

$$Y_{ist} = \alpha + \rho_1 Post_t + \tau_1 D_s + \tau_2 C_s + \delta_1(D_s \times Post_t) + \delta_2(C_s \times Post_t) + \gamma X_{is0} + \varepsilon_{ist} \quad (2)$$

where D_s is an indicator variable equal to 1 if the school is randomly allocated to the individual commitment pledge only (i.e., with no school competition), C_s is an indicator variable equal to 1 if the school is randomly allocated to the individual commitment pledge plus school competition. We estimate the model using a linear probability model (ordinary least squares), including pair and wave fixed effects and clustering standard errors at the school level.

Baseline covariates X_{is0} include the full sets of student, household, and school characteristics. Partially adjusted models include pair fixed effects, the proportion of the class who smokes at baseline, and individual smoking status at baseline. Fully adjusted models add a vector of student, household, and school characteristics to take into account baseline differences. We include the following individual baseline characteristics: truancy, an indicator for ever smoked cigarettes, tobacco knowledge (indicator for being above the median of 6), above-median number of friends who smoke (indicator for >1), an indicator if students believe their parents would find the student smoking (very) unacceptable, time-inconsistent preferences, and risk-seeking preferences. We include the following household baseline covariates: household size (indicator for being above the median of 4), land ownership, skilled occupation of the household head, any smokers in the household, and an indicator for having an above-median asset count. We also include the number of teachers who smoke.

Experimental Validity There are two primary threats to the empirical design. First, the randomization may produce imbalanced groups either by chance or if the process was somehow corrupted. It is unlikely that the process was corrupted since enumerators were given school-specific materials and schools were not informed of their treatment status prior to the intervention. We test for balance of pre-intervention school characteristics across study groups using a t-test of equality in bivariate comparisons and a regression of each school covariate on the dichotomous treatment variable, adjusted for district fixed effects (Table A.2). The primary independent variable is the intent-to-treat random allocation of schools to the treated and control groups, a dichotomous variable. A secondary trichotomous measure captures whether a school is allocated to the control group, commitment pledge group, or commitment pledge plus competition group. The t-tests show that baseline differences in school characteristics between the treated and control groups are not significant. The adjusted differences in Table 1 are also generally small and not statistically significant. Other than the joint test for household characteristics for the individual pledge and control arms, the joint tests of balance fail to reject the null that either the pooled treatment group or the sub-treatment groups are different from the control group.

4 Results

In this section, we begin by describing the main effects of the program on tobacco abstinence. We then assess potential mechanisms underlying the program effects. We continue by estimating heterogeneous effects of the program for different baseline characteristics. Then, we conduct a cost-benefit analysis of the program. Finally, we discuss various robustness checks.

4.1 Effects on Smoking Abstinence

We examine the effect of the program on biochemically verified smoking abstinence in Table 2. Panel A presents the pooled effect, and Panel B compares the effects by sub-treatment. Column 1 pools the post-intervention survey waves; Column 2 separates the post-intervention survey waves; Column 3 adds student and class baseline outcomes; and Column 4 adds individual baseline covariates. Column 5 estimates the fully adjusted model and separates the treatment effect by wave by interacting the treatment indicator with a wave indicator. The estimated program effect is not statistically significant without covariates (Column 1), but the estimate becomes significant and remains of similar magnitude when covariates are included (Columns 3-5). The pooled treatment resulted in a 3.7 to 4.4 percentage-point increase in smoking abstinence, equivalent to a 5.6% increase from the baseline mean under our preferred specification that includes the full set of controls (Column 4). The relationships between smoking abstinence and baseline covariates are as expected; abstinence at baseline is positively correlated with abstinence at follow-up, while ever smoking and having parents who smoke are negatively correlated with abstinence (Table B.1).

We then compare the effect of the individual pledge and the pledge plus school competition on smoking abstinence (Panel B). Using our preferred, fully adjusted specification (Column 4), we find that the individual pledge drives the program effect, with a 5.1-point increase in smoking abstinence, compared to an estimated effect of the pledge plus school competition of 3.7 points, which is not significantly different from zero. While this may suggest that the addition of the school competition making the individual pledge less effective, we fail to reject the equality of the sub-treatment coefficients, with *p*-values ranging from 0.4 to 0.6. We take these results to imply that adding school competition has no additional effect beyond the individual pledge on smoking abstinence. This finding is consistent with null results from randomized studies of class competitions in Europe (Hefler et al., 2017).

We also examine the program effects in each follow-up survey by interacting the treatment indicator with a wave indicator for the 3- and 8-month follow-up (Column 5). When we pool the sub-treatment arms, we find a 4-point increase in verified tobacco abstinence at 3 months, and this effect remains similar at 8 months. The persistence of the effect to 8 months, several months after the class monitoring had ceased, suggests that monitoring alone does not drive the program effect. While we are unable to reject the equality of treatment coefficients at 3 and 8 months, the effect may fade somewhat since the raw means at 3 and 8 months show that verified smoking abstinence declines from 81 to 78%. Comparing the sub-treatment arms, we find similar effects at 3 months and 8 months, and that the effect is driven by the individual pledge arm (Panel B). We fail to reject the equality of the coefficients of the estimated effects of the sub-treatment arms, and we also fail to reject the equality of the sub-treatment coefficients at 3 months ($p = 0.441$) and 8 months ($p = 0.879$).

These results may be attributed to several factors, as detailed in the next section. Here, we make two broad points. First, whereas the existing literature has largely found that adolescent smoking is susceptible to peer influence, we find that group incentives are not effective in our setting. Our study design does not allow us to disentangle whether this is due to our particular group incentive scheme or our particular setting. At least two smoking cessation studies have found team rewards to be no more effective and potentially less effective than equal-sized individual rewards Halpern et al. (2015); White et al. (2020). Second, the token reward of acknowledgment from the Department of Education in the pledge plus school competition arm may not have sufficiently motivated individual students to change their behavior and to shift smoking norms.

Overall, these results provide evidence for the effectiveness of a non-monetary incentive program for promoting healthy behavior among adolescents. We therefore conclude that our approach of harnessing students' social networks, individual pledges, and monitoring to promote tobacco abstinence is feasible. In addition, our estimates of program effects are similar in magnitude to results from higher income settings, where smoking prevalence is lower (Thomas et al., 2013). The authors of that work conducted a meta-analysis of randomized school-based interventions for smoking prevention. They found that programs with a social influence component ($N = 41$ studies), the program type most similar to ours, lead to an average reduction in smoking of 8%. This relative effect size is slightly larger than our estimated 6% reduction in the probability of smoking.

4.2 Potential Mechanisms

In Section 4.1, we showed that the commitment pledge increased smoking abstinence, whereas the addition of school competition did not lead to further improvement. There are several possible underlying mechanisms that might explain this pattern of findings. In Section 2.2, we enumerate several of these mechanisms under the categories of peer influence and incentive effects. We summarize our evidence on potential mechanisms in Table 3.

We begin by considering the role of peer influence. We estimate the effect of any peer influence on smoking by exploiting quasi-random class assignment in 8th grade. This provides plausibly exogenous variation in friendship networks, namely the number of 7th grade friends who are assigned to the same 8th grade class. Specifically, we estimate how smoking abstinence varies with the number of abstaining friends at baseline who are assigned to the same 8th grade class, and following De Giorgi et al. (2010) we instrument for this endogenous exposure using the average number of friends of friends assigned to the same 8th grade class who abstained. We find that students with more friends who abstain are more likely to abstain themselves (3 to 5 percentage points per additional friend), and the program effect remains similar for the sub-treatment analysis (Table C.1). In further analyses, we assess the role of friends and peers who smoke as sources of heterogeneous treatment effects (Table C.2). We find no differential effect from having a higher portion of friends who abstained at baseline and are in the same 8th grade class, but students who were predicted to smoke by their friends are more likely to abstain by 0.3 points. Further detail can be found in Appendix C.

We further explore several types of peer influence as potential mechanisms. First, we anticipated that the group incentives in the pledge plus school competition arm would differentially activate two mechanisms: (1) students' fear of social sanctions from peers, translating into pressure to comply, and (2) group identity. Because we find no incremental effect from the addition of the school competition to the pledge, we can rule out these mechanisms as primary channels here.

Second, students can learn, including from each other, how to quit. One measure of (social) learning is whether students display increased knowledge of the risks of tobacco use following the intervention. We find no significant improvement in treated students' tobacco knowledge after the trial compared with the control group (Table 3, Panel A, Columns 1-2). This finding may be due in part to the relatively high baseline knowledge. As an additional indicator, we examine the fraction of students who agreed with the statement that better knowledge would make them abstain from tobacco use, and find that the share is similar across the treatment and control groups.

Third, the intervention may have changed perceived norms regarding the acceptability of smoking or changed students' willingness to conform to perceived norms (Table 3, Panel A). We find no significant change in how often treated students smoke with their friends compared with control students (Column 3), or whether treated students agree with the statement that having friends who abstain from smoking would help them abstain (Column 4). Perceived smoking norms at home may have shifted as our intervention asked parents to monitor their children's tobacco use, yet we find no significant effect on treated students changing their belief about whether their parents disapproved of them smoking (Column 5). We note that changing beliefs about parental disapproval could lead students either to obey and smoke less or to rebel and smoke more. However, we find no program effect on beliefs about parental approval, with a tight confidence interval around zero.

Next, we turn to the set of mechanisms that we classify as incentive effects, that is, behavioral responses to the payoff structure in the experiment. First, perhaps the most straightforward explanation for our findings is that both sub-treatment arms increased students' fear of punishment (Table 3, Panel B). Our program effects do not fade after the end of the monitoring period, suggesting this channel is not a primary driver. We also use students' agreement with statements on demerit points, urine testing, and reporting to student council, teachers, and parents. The estimates are noisy, except that students in treated schools are more likely to agree with the statements that regular reporting by the student council helps them abstain from smoking (Column 3).

Second, the program may have increased students' motivation and effort to stay smoke-free (Table 3, Panel C, Columns 1-2). We explore individual motivation to abstain from tobacco use by analyzing changes in students' own smoking intentions and their likelihood to report friends who smoke during the biweekly reporting, and we find no significant change. Third and relatedly, the non-monetary incentives may have helped students with limited self-control to overcome their time-inconsistency. As we describe in Section 4.3, we find evidence that treated students who are time-inconsistent based on a hypothetical monetary choice task are more likely to abstain (Figure B.5). This would suggest that the program helped time-inconsistent students to follow through on their plans to abstain, although we caution that hypothetical elicitations may be confounded by other factors, such as risk preferences (Frederick et al., 2002).

Fourth, the program may have increased the salience of not smoking, notably through the regular monitoring of smoking behavior and the non-monetary incentives. We would expect a salience effect to fade after the program ended. Instead, we find that the program effects persist beyond the school year. We also find some conflicting evidence for increased salience as proxied by a lower probability to discuss smoking with their parents and friends (Panel C, Columns 3-4). The lower probability of peer discussion is likely due to reduced peer pressure to smoke, which suggests increased salience in this context, but fewer discussions with parents is related to lower salience, as parents typically remind their children not to smoke (Ng et al., 2007a). Finally, the short-term incentives may have helped to establish a healthy habit of not smoking. We observe that carry-on effects of the program continue for at least 3 months, in line with habit formation, although we lack data to test directly for habit formation.

Our findings are similar when we consider the sub-treatment arms (Table B.2), suggesting that the aforementioned potential mechanisms are not differentially activated by the pledge plus competition sub-treatment arm. We conclude that there is strongest evidence in favor of changing norms and overcoming time-inconsistency as potential mechanisms that might account for the program effects. More work is needed to distinguish their relative contribution and to test them more formally.

4.3 Heterogeneous Program Effects

We explore heterogeneity by student, household, and school characteristics, estimating difference-in-differences models that interact the treatment variable with each of our full set of baseline covariates. Figure B.5 presents heterogeneity of the pooled estimates, and Figure B.6 presents heterogeneity of the sub-treatment effects.[18] Continuous variables are dichotomized at the median split.

By student characteristics Student characteristics may be one important source of heterogeneity, as different types of students may respond differently to the intervention. We consider the following student characteristics: truancy at baseline as a related risky behavior, ever smoked and current smoking abstinence status at baseline, baseline smoking knowledge, and whether the student's closest friend smoked at baseline (Figure B.5). We also include our measures of time-inconsistency and risk-seeking behavior, since smokers may have higher risk tolerance. The program has larger effects on time-inconsistent students, but otherwise we find no significant heterogeneous treatment effects by student characteristics in the pooled analysis. Because the intervention includes students' belief of their parents' attitude towards youth smoking, we include the indicator for students' belief that parents would find smoking unacceptable as a potential source of heterogeneity and find no significant heterogeneity. Similarly, when we separately consider the sub-treatments, we find no significant heterogeneities by student characteristics, although there is a marginally significant (at the 10% level) decrease in the effect of the pledge plus competition among truant students.

By household characteristics Students with different family backgrounds may respond differently to the intervention. To explore this possibility, we consider the following baseline household characteristics: household size, land ownership, the head of household in skilled labor, parent smokes, any smokers in the household, parents find child smoking to be very unacceptable, and having above-median household assets (Figure B.5). Except for the indicator for the head of household in skilled labor, we find no significant heterogeneous treatment effects under the pooled or sub-treatment analyses.[19]

By school characteristics School quality may play a role in how students respond to the intervention. For baseline school quality, we consider class smoking prevalence, distance to

[18] Appendix Tables B.3, B.4, and B.5 present the standard errors.

[19] Smoking is more prevalent among adults with higher per-capita household expenditure in Indonesia, but it is unclear whether children from these backgrounds would respond differentially to the intervention.

Yogyakarta city, student-to-teacher ratio, above-median number of student council members, the mathematics score on the national examination, whether any teachers smoke, and number of classes (Figure B.5). Schools with a higher mathematics score have more academically minded students, who tend to avoid smoking. We find larger treatment effects among those in higher performing mathematics test schools and schools with more classes, but otherwise we do not detect heterogeneous effects by school characteristics.

4.4 Cost-Benefit Analysis

A potentially important advantage of our tobacco prevention program is that it is relatively low cost. The use of non-monetary incentives are an important low-cost input. To assess the value of our program, we first compare its cost to a cessation program in a similar setting. Using results from the Philippines, a country with a similar GDP per capita as Indonesia, a cessation program that used a monetary commitment contract was associated with a 14% increase in abstinence at 6 months and 5% at 12 months (Giné et al., 2010). The program's cost per quit at 12 months was $700, adjusted for purchasing power parity, in 2010.

The costs associated with our prevention program include the unannounced cotinine tests (assumed to occur twice per year, with 50% of students selected during each visit), printing the individual pledges and posters, and administrative costs associated with the information session and maintaining a phone line and emails for reporting. The cost of the cotinine test kit was $2 per test, and printing costs were $0.10 per respondent. Administrative costs were $6,000, based on an estimated cost of $500 per month. These program costs, spread over our total sample, is approximately $4.40 per student. Our program effect is a 5% increase in abstinence at 8 months, and we conservatively assume that the program effect would be half as large at 12 months, or about 2.5 percent.[20] Based on the expected 12-month effect, the cost per smoker averted would be $220, substantially lower than the cost of the cessation intervention. Our program effect would need to be 0.6 points, one-tenth as large at 12 months as at 6 months, in order for our program's cost-effectiveness to be comparable to the smoking cessation program in the Philippines.

We then perform back-of-the-envelope calculations of the benefit-to-cost difference and ratio of our program (Appendix Table B.6). We assume the program would delay smoking initiation by 1 to 3 years. We choose a maximum window of 3 years to correspond to an average age of smoking initiation of 17, and the majority of smokers in Indonesia begin smoking between 15 and 19 years (Ministry of Health, Indonesia, 2013). We further assume that the benefits would accrue from productivity gains from delayed cardiovascular disease

[20] Extrapolating from the linear trend at 8 months, the expected abstinence effect would be much larger.

(CVD) and increased labor market participation for 1 to 3 years from delaying smoking initiation by 1 to 3 years, as well as savings from delaying the onset of pulmonary disease. All future gains are discounted at 3%. We assume a 40% probability of pulmonary disease and a 40% probability of CVD.[21] To estimate a lower bound, we exclude externalities from secondhand smoke exposure and other health expenditures. Even with these conservative assumptions, benefits of the program exceed costs by $3,000 to $6,500, implying a benefit-to-cost ratio of 5 to 18.

4.5 Robustness

Attrition Almost 96% of students were present in all three waves of the survey. Nonetheless, selective attrition may modestly bias our findings if students who are more or less likely to abstain from smoking remain in the study after 3 and 8 months. We examine the differences across treatment groups and find no significant differences in attrition across treatment groups at the individual level (Table B.7). At the school level, we also find no significant differences in the share of students who attrited across the treated and control groups. In addition, we examine differential attrition across treated and control groups by student and household characteristics and find that the differences are generally small and not statistically significant (Figure B.7). We also restrict the sample to students who were present in all three waves of the survey (Table B.8) and find that the estimated program effect is similar to our earlier findings (Table 2). These results suggest that the treatment effect is unlikely to be driven by selective attrition.

Other selection concerns We use verified tobacco abstinence as our main outcome of interest. Three potential concerns may arise in terms of selection into urine testing. First, students may hide their tobacco use by only providing a self-report and refusing the biochemical verification. We analyze whether students refuse the urine test differentially

[21] The general risk of CVD in Indonesia is about 30%, and smokers face a 25-30% higher risk of CVD (Maharani et al., 2019). Almost 50% of smokers are estimated to eventually get chronic obstructive pulmonary disease (COPD) (Marsh et al., 2006). We assume life expectancy is 67, retirement age at 60, and GDP per capita at USD 3800 (http://www.healthdata.org/indonesia). Costs due to secondhand smoke exposure are excluded. The cost of CVD is about USD 2500 in Indonesia. The estimated cost of noncommunicable diseases (NCDs) is USD 17,800 in Indonesia, and CVD accounts for about one-third of NCDs, and we assume about 40% of CVD cases can be attributed to smoking (http://www3.weforum.org/docs/WEF_The_Economics_of_non_Disease_Indonesia_2015.pdf). COPD has an associated cost of 5 days of missed activities and an annual 12% probability of hospitalization (Wier et al., 2010); this would cost USD 150 (https://www.who.int/choice/country/idn/cost/en/). Under these assumptions, CVD would occur on average between the ages of 58 to 60 (https://www.georgeinstitute.org/sites/default/files/reducing-the-burden-of-cvd-in-indonesia-evidence-review.pdf) and pulmonary disease onset would, on average, be between the ages of 39 to 41 (COPD Foundation).

across the treatment and control. In addition, the commitment pledge plus school competition intervention may provide an incentive for students to refuse the biochemical verification as a way to improve the school's performance. To address this potential concern, we create an indicator for missing biochemical verification (i.e., self-reported tobacco use only) and find no evidence of differential refusal across the treated and control groups, or across the sub-treatment arms (Table 4, Column 1).

Second, students may hide their tobacco use by lying: self-reporting abstinence and failing the biochemical verification (Column 2). To address this potential concern, we create an indicator that takes the value one if a student reports tobacco abstinence but the biochemical verification shows otherwise. We find no significant effect across the treated and control groups on this measure. We also explore the frequency that a student is reported and find that students are less likely to be reported over time, which is consistent with a reduction in smoking.

Third, under the commitment pledge plus school competition, students may be reluctant to report their friends in order to improve their school's performance, and this may lead to under-reporting when school competition is included. We compare the number of friends who report an index student's smoking behavior under the individual pledge versus pledge plus school competition arms (Column 3, Panel B). We find a reduction in the frequency of being reported under both sub-treatment arms, and the point estimates are not statistically different. This suggests that students in the pledge plus school competition arm are not less likely to report their peers compared to students in the pledge only arm.

Alternative measures of smoking abstinence We analyze alternative smoking abstinence measures as a test of robustness (Table 5). First, we assume students who did not take the cotinine test were smokers (Column 1). Second, we assume students who did not take the cotinine test were non-smokers if they were non-smokers at baseline (Column 2). Third, we restrict the sample to students with non-missing cotinine test and only use the test results as an outcome variable (Column 3). Fourth, we investigate alternative time horizons of abstinence—at 2 weeks and 1 month—combining the verified tobacco abstinence measure and self-reported abstinence (Columns 4 and 5). Although some of the program effects are imprecisely estimated, the estimates using the alternative outcome measures are similar, reinforcing the estimated effectiveness of the program for preventing smoking initiation.

5 Conclusion

Non-monetary incentives have been shown to be effective across multiple behaviors, domains, and contexts. Our paper sheds light on whether non-monetary incentives that serve as partial commitment contracts can encourage behavior change among adolescents, who may be especially susceptible to self-control problems and peer influence. We developed a novel penalty-based intervention to prevent tobacco use among middle-school students. We find that the penalty-based intervention reduced the probability of smoking by 5% after 3 months, an effect size that is sustained into the next school year (6% after 8 months, 3 months after the intervention has ended). We find that the program effect is especially large for students who are identified as time-inconsistent, as would be expected for a commitment mechanism. We do not find substantial heterogeneity on most other dimensions of student, household, and school characteristics, indicating that the intervention may be appropriate for a large cross-section of students and schools.

Interestingly, adding school competition to the commitment pledge did not further motivate students in our sample to abstain from smoking. One possibility is that school competition alone, not in combination with the commitment program, might have had more favorable effects. Our pledge plus competition intervention emphasized group identity without any tangible rewards (or penalties); this too may have weakened its effectiveness. Future studies might experiment with different types of rewards to find ones that are particularly motivating. Both social commitment and social norms interventions have proven effective in other settings (e.g., Kast et al., 2018; Allcott and Rogers, 2014). We find that students' smoking behavior is heavily influenced by their friendship network. Using quasi-random variation in class assignment, our best estimates indicate that abstinence increases by 3 to 5 percentage points for each additional friend who is assigned to the person's 8th grade class and abstains. There may be opportunities to enhance our commitment intervention by leveraging these social forces.

By focusing on more rural schools in Indonesia, our study advances the literature on tobacco prevention in low-income settings. School-based studies have focused almost exclusively on high-income settings, despite the fact that the tobacco-attributable disease burden is concentrated in low- and middle-income countries. Indonesia's status as one of the world leaders in male smoking prevalence makes it a particularly important setting for intervention. In the absence of stronger political will and clout to overcome tobacco industry interests in national policy-setting for tobacco control, Indonesia will need local experimentation and action to reduce its substantial tobacco-attributable disease burden. We find that our tobacco prevention program, implemented in two districts in Yogyakarta

Province, is relatively low-cost and might be readily scalable elsewhere in Indonesia, and perhaps in other settings too. It will be especially important to test whether such programs can be scaled in tobacco-producing areas of Indonesia, where local resistance may be stronger. We believe that our intervention, with a favorable cost-benefit ratio, when combined with other tobacco control programs, can lower adolescent smoking prevalence. As evidence of the sustainability of the program, one of the two districts, Kulon Progo, decided to restart the program on its own after our study had ended. Researchers may be able to play an important role in identifying the appropriate mix of interventions for reducing tobacco use in Indonesia and other low-income settings.

References

Advani, Arun and Bansi Malde, "Credibly identifying social effects: Accounting for network formation and measurement error," *Journal of Economic Surveys*, 2018, *32* (4), 1016–1044.

Akers, Ronald L. and Gang Lee, "A Longitudinal Test of Social Learning Theory: Adolescent Smoking," *Journal of Drug Issues*, 1996, *26* (2), 317–343.

Allcott, Hunt and Todd Rogers, "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation," *The American Economic Review*, 2014, *104* (10), 3003–3037.

Assunta, Mary and E Ulysses Dorotheo, "SEATCA Tobacco Industry Interference Index: a tool for measuring implementation of WHO Framework Convention on Tobacco Control Article 5.3," *Tobacco Control*, 2016, *25* (3), 313–318.

Astuti, Putu A. S., Ketut H. Mulyawan, Susy K. Sebayang, Ni Made D. Kurniasari, and Becky Freeman, "Cigarette retailer density around schools and neighbourhoods in Bali, Indonesia: A GIS mapping," *Tobacco Induced Diseases*, 2019, *17* (July), 55.

Babcock, Philip, Kelly Bedard, Gary Charness, John Hartman, and Heather Royer, "Letting Down the Team? Evidence of Social Effects of Team Incentives," *Journal of the European Economic Association*, 2015, *13* (5), 841–870.

Bai, Liang, Benjamin Handel, Edward Miguel, and Gautam Rao, "Self-Control and Demand for Preventive Health: Evidence from Hypertension in India," *Review of Economics and Statistics*, Forthcoming.

Bandura, Albert, *Social foundations of thought and action*, Englewood Cliffs, NJ: Prentice-Hall, 1986.

Barber, Sarah, Sri Moertiningsih Aditomo, Abdillah Ahsan, and Diahhadi Setyonaluri, "Tobacco economics in Indonesia," Technical Report, International Union Against Tuberculosis and Lung Disease, Paris 2008. International Union Against Tuberculosis and Lung Disease.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, December 2006, *96* (5), 1652–1678.

Benowitz, Neal L, John T Bernert, Jonathan Foulds, Stephen S Hecht, Peyton Jacob III, Martin J Jarvis, Anne Joseph, Cheryl Oncken, and Megan E Piper, "Biochemical verification of tobacco use and abstinence: 2019 update," *Nicotine & Tobacco Research*, 2020, *22* (7), 1086–1097.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of Econometrics*, 2009, *150* (1), 41–55.

Bruhn, Miriam and David McKenzie, "In pursuit of balance: randomization in practice in development field experiments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 200–232.

Bryan, Gharad, Dean Karlan, and Scott Nelson, "Commitment devices," *Annual Review of Economics*, 2010, *2* (1), 671–698.

Card, David and Laura Giuliano, "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends," *Review of Economics and Statistics*, 2013/12/10 2013, *95* (4), 1130–1149.

Carpenter, Christopher and Philip J. Cook, "Cigarette taxes and youth smoking: New evidence from national, state, and local Youth Risk Behavior Surveys," *Journal of Health Economics*, 2008, *27* (2), 287 – 299.

Chaloupka, Frank J. and Kenneth E. Warner, "The economics of smoking," in Anthony J. Culyer and Joseph P. Newhouse, eds., *Handbook of Health Economics*, Vol. Volume 1, Part 2, Elsevier, 2000, pp. 1539–1627.

Charness, Gary, Michael Naef, and Alessandro Sontuoso, "Opportunistic conformism," *Journal of Economic Theory*, 2019, *180*, 100–134.

Chen, Margaret Harris, "Economic concerns hamper tobacco control in China," *The Lancet*, 2007, *370* (9589), 729–730.

DeCicca, Philip, Donald S. Kenkel, and Michael F. Lovenheim, "The Economics of Tobacco Regulation: A Comprehensive Review," *Journal of Economic Literature*, Forthcoming.

DiFranza, Joseph R, Judith A Savageau, Kenneth Fletcher, Lori Pbert, Jennifer O'Loughlin, Ann D McNeill, Judith K Ockene, Karen Friedman, Jennifer Hazelton, Connie Wood et al., "Susceptibility to nicotine dependence: the Development and Assessment of Nicotine Dependence in Youth 2 study," *Pediatrics*, 2007, *120* (4), e974–e983.

Dupas, Pascaline, "Health behavior in developing countries," *Annu. Rev. Econ.*, 2011, *3* (1), 425–449.

_ , "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment," *Econometrica*, 2014, *82* (1), 197–228.

Ennett, Susan T, Karl E Bauman, Vangie A Foshee, Michael Pemberton, and Katherine A Hicks, "Parent-Child Communication About Adolescent Tobacco and Alcohol

Use: What Do Parents Say and Does It Affect Youth Behavior?," *Journal of Marriage and Family*, 2001, *63* (1), 48–62.

Fletcher, Jason M., "Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects," *Health Economics*, 2010, *19* (4), 466–484.

Fletcher, Jason M and Stephen L Ross, "Estimating the effects of friends on health behaviors of adolescents," *Health Economics*, 2018, *27* (10), 1450–1483.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Giles, Emma L, Shannon Robalino, Elaine McColl, Falko F Sniehotta, and Jean Adams, "The effectiveness of financial incentives for health behaviour change: systematic review and meta-analysis," *PloS One*, 2014, *9* (3), e90347.

Giné, Xavier, Dean Karlan, and Jonathan Zinman, "Put your money where your butt is: a commitment contract for smoking cessation," *American Economic Journal: Applied Economics*, 2010, *2* (4), 213–235.

Giorgi, Giacomo De, Michele Pellizzari, and Silvia Redaelli, "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2010, *2* (2), 241–75.

Gopalan, Saji S., Ronald Mutasa, Jed Friedman, and Ashis Das, "Health sector demand-side financial incentives in low- and middle-income countries: A systematic review on demand- and supply-side effects," *Social Science & Medicine*, 1 2014, *100* (0), 72–83.

Hall, Sharon M., Kevin L. Delucchi, Janice Y. Tsoh, Wayne F. Velicer, Christopher W. Kahler, James Ranger Moore, Donald Hedeker, and Ray Niaura, "Statistical analysis of randomized trials in tobacco treatment: Longitudinal designs with dichotomous outcome," *Nicotine & Tobacco Research*, 08 2001, *3* (3), 193–202.

Halpern, Scott D., Benjamin French, Dylan S. Small, Kathryn Saulsgiver, Michael O. Harhay, Janet Audrain-McGovern, George Loewenstein, Troyen A. Brennan, David A. Asch, and Kevin G. Volpp, "Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation," *New England Journal of Medicine*, 2015, *372* (22), 2108–2117. PMID: 25970009.

Heatherton, Todd F., Lynn T. Kozlowski, Richard C. Frecker, and Karl-Olov Fagerström, "The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire," *British Journal of Addiction*, 1991, *86* (9), 1119–1127.

Hefler, Marita, Selma C. Liberato, and David P. Thomas, "Incentives for preventing smoking in children and adolescents," *Cochrane Database of Systematic Reviews*, 2017, *6*, CD008645.

Herold, Florian, "Contractual incompleteness as a signal of trust," *Games and Economic Behavior*, 2010, *68* (1), 180–191.

Higgins, Julian P T, Douglas G Altman, Peter C Gøtzsche, Peter Jüni, David Moher,

Andrew D Oxman, Jelena Savović, Kenneth F Schulz, Laura Weeks, and Jonathan A C Sterne, "The Cochrane Collaboration's tool for assessing risk of bias in randomised trials," *BMJ*, 2011, *343*, d5928.

Huang, Grace C., Daniel Soto, Kayo Fujimoto, and Thomas W. Valente, "The Interplay of Friendship Networks and Social Networking Sites: Longitudinal Analysis of Selection and Influence Effects on Adolescent Smoking and Alcohol Use," *American Journal of Public Health*, 2014, *104* (8), e51–e59. PMID: 24922126.

Imai, Kosuke, Gary King, and Clayton Nall, "The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation," *Statistical Science*, 2009, *24* (1), 29–53.

International Agency for Research on Cancer, "Tax, price, and tobaco use among young people," in "Effectiveness of Tax and Price Policies for Tobacco Control," Vol. 14 of *IARC Handbooks of Cancer Prevention* Lyon, France 2011, pp. 201–257.

Jamal, Ahmed, Andrea Gentzke, S Sean Hu, Karen A Cullen, Benjamin J Apelberg, David M Homa, and Brian A King, "Tobacco use among middle and high school students—United States, 2011–2016," *MMWR Morbidity and Mortality Weekly Report*, 2017, *66* (23), 597.

Kast, Felipe, Stephan Meier, and Dina Pomeranz, "Saving more in groups: Field experimental evidence from Chile," *Journal of Development Economics*, 2018, *133*, 275 – 294.

Kenkel, Donald S, "The demand for preventive medical care," *Applied Economics*, 1994, *26* (4), 313–325.

King, Gary and Richard Nielsen, "Why propensity scores should not be used for matching," *Political Analysis*, 2019, *27* (4), 435–454.

Kremer, Michael and Dan Levy, "Peer effects and alcohol use among college students," *The Journal of Economic Perspectives*, 2008, *22* (3), 189.

Kullgren, Jeffrey T., Andrea B. Troxel, George Loewenstein, David A. Asch, Laurie A. Norton, Lisa Wesby, Yuanyuan Tao, Jingsan Zhu, and Kevin G. Volpp, "Individual-Versus Group-Based Financial Incentives for Weight Loss: A Randomized, Controlled Trial," *Annals of Internal Medicine*, 2013, *158* (7), 505–514.

Lillard, Dean R, Eamon Molloy, and Andrew Sfekas, "Smoking initiation and the iron law of demand," *Journal of Health Economics*, 2013, *32* (1), 114–127.

Lundborg, Petter, "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, 2006, *25* (2), 214–233.

Maharani, Asri, Devarsetty Praveen, Delvac Oceandy, Gindo Tampubolon, and Anushka Patel, "Cardiovascular disease risk factor prevalence and estimated 10-year cardiovascular risk scores in Indonesia: The SMARThealth Extend study," *PloS one*, 2019, *14* (4), e0215219.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, July 1993, *60* (3), 531–542.

Mantzari, Eleni, Florian Vogt, Ian Shemilt, Yinghui Wei, Julian P.T. Higgins, and Theresa M. Marteau, "Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis," *Preventive Medicine*, 2015, *75*, 75 – 85.

Marsh, S, S Aldington, P Shirtcliffe, M Weatherall, and R Beasley, "Smoking and COPD: what really are the risks?," *European Respiratory Journal*, 2006, *28* (4), 883–884.

Ministry of Health, Indonesia, "Riset Kesehatan Dasar (Basic Health Research)," 2013.

Ng, Nawi, Lars Weinehall, and Ann Öhman, "'If I don't smoke, I'm not a real man'–Indonesian teenage boys' views about smoking," *Health Education Research*, 2007, *22* (6), 794–804.

_ , Yayi Suryo Prabandari, Retna Siwi Padmawati, Felix Okah, C Keith Haddock, Mark Nichter, Mimi Nichter, Myra Muramoto, Walker S C Poston, Sara A Pyle, Nurazid Mahardinata, and Harry A Lando, "Physician assessment of patient smoking in Indonesia: a public health priority," *Tobacco Control*, June 2007, *16* (3), 190–196.

Nichter, Mark, "Introducing tobacco cessation in developing countries: an overview of Project Quit Tobacco International," *Tobacco Control*, June 2006, *15* (Suppl_1), i12–17.

Nuryunarsih, Desy, Sarah Lewis, and Tessa Langley, "Health Risks of Kretek Cigarettes: A Systematic Review," *Nicotine & Tobacco Research*, 01 2021. ntab016.

Patel, Mitesh S., David A. Asch, Roy Rosin, Dylan S. Small, Scarlett L. Bellamy, Kimberly Eberbach, Karen J. Walters, Nancy Haff, Samantha M. Lee, Lisa Wesby, Karen Hoffer, David Shuttleworth, Devon H. Taylor, Victoria Hilbert, Jingsan Zhu, Lin Yang, Xingmei Wang, and Kevin G. Volpp, "Individual Versus Team-Based Financial Incentives to Increase Physical Activity: A Randomized, Controlled Trial," *Journal of General Internal Medicine*, 2016, pp. 1–9.

Reitsma, Marissa B, Nancy Fullman, Marie Ng, Joseph S Salama, Amanuel Abajobir, Kalkidan Hassen Abate, Cristiana Abbafati, Semaw Ferede Abera, Biju Abraham, Gebre Yitayih Abyu et al., "Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015," *The Lancet*, 2017, *389* (10082), 1885–1906.

Rotsaert, Tijs, Ernesto Panadero, Eduardo Estrada, and Tammy Schellens, "How do students perceive the educational value of peer assessment in relation to its social nature? A survey study in Flanders," *Studies in Educational Evaluation*, 2017, *53*, 29–40.

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a Fortune-500 company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Ryan, Richard M. and Edward L. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.," *American Psychologist*, 2000, *55* (1), 68–78.

Sacerdote, Bruce, "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps

Forward?," *Annual Review of Economics*, 2014/10/16 2014, *6* (1), 253–272.

Schilbach, Frank, "Alcohol and Self-Control: A Field Experiment in India," *American Economic Review*, April 2019, *109* (4), 1290–1322.

Shiffman, Saul, Michael S. Dunbar, and Neal L. Benowitz, "A Comparison of Nicotine Biomarkers and Smoking Patterns in Daily and Nondaily Smokers," *Cancer Epidemiology and Prevention Biomarkers*, 2014, *23* (7), 1264–1272.

Steinberg, Laurence, "Risk taking in adolescence: New perspectives from brain and behavioral science," *Current Directions in Psychological Science*, 2007, *16* (2), 55–59.

_ , Grace Icenogle, Elizabeth P. Shulman, Kaitlyn Breiner, Jason Chein, Dario Bacchini, Lei Chang, Nandita Chaudhary, Laura Di Giunta, Kenneth A. Dodge, Kostas A. Fanti, Jennifer E. Lansford, Patrick S. Malone, Paul Oburu, Concetta Pastorelli, Ann T. Skinner, Emma Sorbring, Sombat Tapanya, Liliana Maria Uribe Tirado, Liane Peña Alampay, Suha M. Al-Hassan, and Hanan M. S. Takash, "Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation," *Developmental Science*, 2018, *21* (2), e12532.

Thomas, R and Rafael Perera, "School-based programmes for preventing smoking," *Cochrane Database Syst Rev*, 2006, *3.*

Thomas, Roger E., Julie McLellan, and Rafael Perera, "School-based programmes for preventing smoking," *Evidence-Based Child Health: A Cochrane Review Journal*, 2013, *8* (5), 1616–2040.

Trogdon, Justin G, James Nonnemaker, and Joanne Pais, "Peer effects in adolescent overweight," *Journal of Health Economics*, 2008, *27* (5), 1388–1399.

Valente, Thomas W., Beth R. Hoffman, Annamara Ritt-Olson, Kara Lichtman, and C. Anderson Johnson, "Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools," *American Journal of Public Health*, 2003, *93* (11), 1837.

West, Robert, Peter Hajek, Lindsay Stead, and John Stapleton, "Outcome criteria in smoking cessation trials: proposal for a common standard," *Addiction*, 2005, *100* (3), 299–303.

White, Justin S. and Hana Ross, "Smokers' Strategic Responses to Sin Taxes: Evidence from Panel Data in Thailand," *Health Economics*, 2015, *24* (2), 127–141.

White, Justin S, Christopher Lowenstein, Nucharee Srivirojana, Aree Jampaklay, and William H Dow, "Incentive programmes for smoking cessation: Cluster randomized trial in workplaces in Thailand," *BMJ*, 2020, *371.*

Wier, Lauren M, BUEP Anne Pfuntner, and Claudia Steiner, "STATISTICAL BRIEF# 103," *Hospital*, 2010.

World Health Organization, "Global Adult Tobacco Survey (GATS): Indonesia report 2011," *Jakarta: WHO Regional Office for South East Asia*, 2012.

_ , "Global Youth Tobacco Survey: Indonesia Report 2014," Technical Report, World Health Organization Regional Office for Southeast Asia, New Delhi, India 2015.

_ , "Indonesia 2015 Global School-based Student Health Survey: Fact Sheet," Technical Report, World Health Organization Regional Office for Southeast Asia 2015.

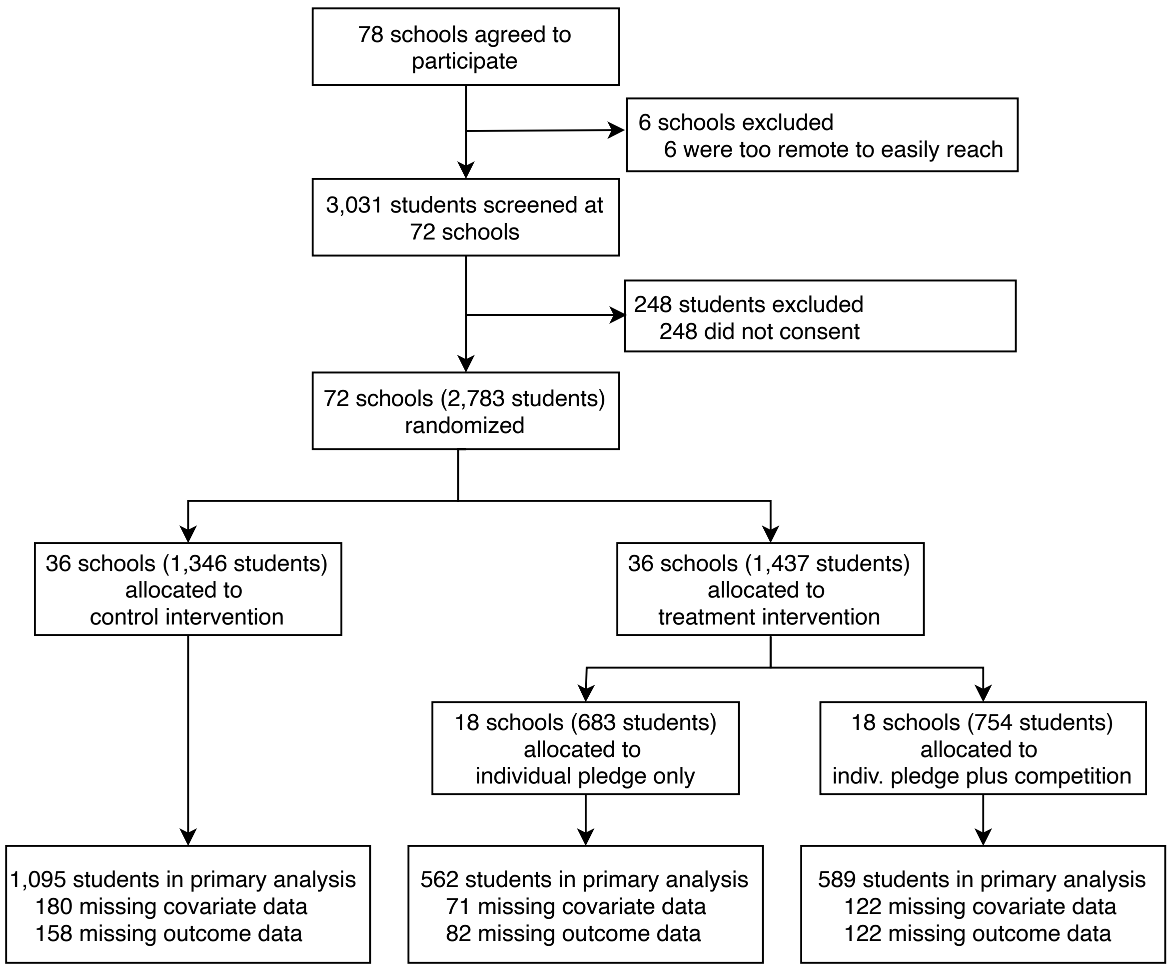
_ , "WHO Global Report on Trends in Prevalence of Tobacco Smoking," Technical Report, World Health Organization, Geneva, Switzerland 2015.

_ , "WHO report on the global tobacco epidemic, 2017: Monitoring tobacco use and prevention policies," Technical Report, World Health Organization, Geneva, Switzerland 2017.

Zelazo, Philip David and Stephanie M Carlson, "Hot and cool executive function in childhood and adolescence: Development and plasticity," *Child Development Perspectives*, 2012, *6* (4), 354–360.

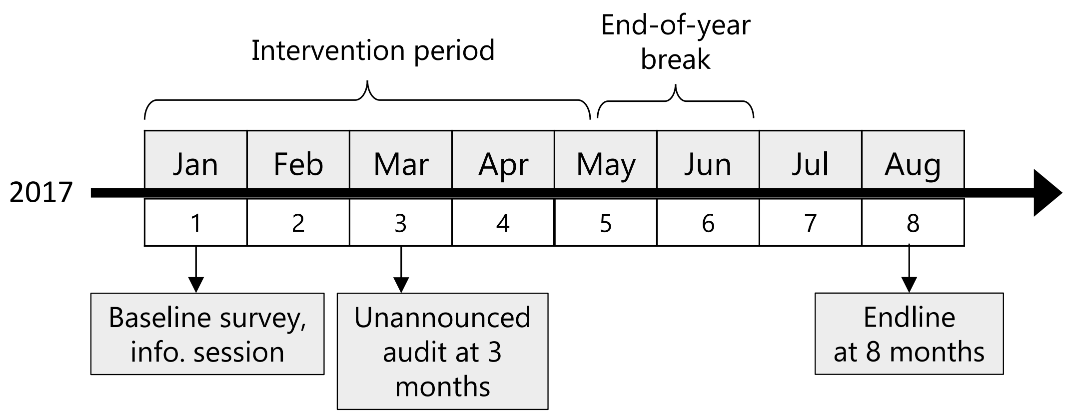
Figures and Tables

Figure 1: Study flowchart



Notes: Observations with missing covariate data and missing outcome data are not mutually exclusive.

Figure 2: Timeline



Notes: The study began in January 2017, which marked the beginning of the second semester of 7th grade. Reporting ended in May 2017 due to final examinations. The new academic year (8th grade) began in July 2017.

Table 1: Baseline characteristics and balance under pair matching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Sub-treatments		Adjusted	Adjusted	Adjusted
		Treated)	Individual	School	difference	difference	difference
	Control	(Pooled)	pledge	competition	(2) - (1)	(3) - (1)	(4) - (1)
Panel A. School characteristics							
Distance to Yogyakarta	17.391	17.595	16.004	19.027	0.657	0.767	0.137
	(10.671)	(9.543)	(9.120)	(9.693)	(1.512)	(1.496)	(1.805)
Student-to-teacher ratio	14.771	14.312	14.483	14.745	-0.380	-0.290	-0.230
	(3.122)	(2.986)	(2.946)	(3.213)	(0.870)	(1.085)	(0.872)
No. student council members	11.372	9.140	9.858	11.363	-2.565	-1.228	-3.744
	(10.930)	(10.702)	(9.455)	(12.177)	(2.627)	(3.691)	(2.500)
Avg. mathematics score	56.932	52.516	55.238	55.564	-3.962	-2.791	-2.623
	(14.536)	(13.273)	(13.873)	(14.658)	(3.467)	(4.327)	(3.752)
No. teachers who smoke	1.429	2.037	1.748	1.529	0.820**	0.282	0.833*
	(1.521)	(1.626)	(1.738)	(1.399)	(0.368)	(0.343)	(0.495)
No. classes	12.481	12.513	12.202	12.793	0.071	1.295	-1.172
	(4.321)	(4.561)	(4.207)	(4.583)	(1.174)	(1.452)	(1.246)
No. observations					72	72	72
Joint significance test (p-value)					0.236	0.500	0.245
Panel B. Student characteristics							
Truancy	0.039	0.059	0.054	0.064	0.019	0.002	0.023
	(0.193)	(0.236)	(0.225)	(0.245)	(0.013)	(0.019)	(0.017)
Ever smoked	0.730	0.737	0.739	0.735	0.006	0.003	0.005
	(0.444)	(0.441)	(0.439)	(0.442)	(0.037)	(0.047)	(0.039)
Tobacco knowledge	5.619	5.758	5.646	5.859	0.137	-0.069	0.238*
	(2.217)	(2.119)	(2.041)	(2.183)	(0.128)	(0.147)	(0.130)
No. friends who smoked	1.538	1.775	1.716	1.829	0.231	0.045	0.253
	(1.704)	(1.761)	(1.756)	(1.765)	(0.176)	(0.191)	(0.195)
Time inconsistent	0.704	0.657	0.645	0.668	-0.045	-0.038	-0.022
	(0.457)	(0.475)	(0.479)	(0.471)	(0.038)	(0.041)	(0.043)
Risk seeking	0.342	0.441	0.390	0.486	0.098**	-0.010	0.134**
	(0.474)	(0.497)	(0.488)	(0.500)	(0.046)	(0.048)	(0.059)
Parents find smoking unacceptable	0.251	0.242	0.238	0.245	-0.009	0.021	-0.035
	(0.434)	(0.428)	(0.426)	(0.431)	(0.022)	(0.020)	(0.031)
Joint significance test (p-value)					0.235	0.978	0.239
Panel C. Household characteristics							
Household size	4.755	4.703	4.499	4.888	-0.055	-0.325***	0.230
	(1.656)	(1.554)	(1.258)	(1.760)	(0.133)	(0.105)	(0.165)
Land ownership	0.551	0.508	0.470	0.543	-0.039	-0.063	0.008
	(0.498)	(0.500)	(0.499)	(0.499)	(0.031)	(0.040)	(0.039)
Parent in skilled occupation	0.348	0.299	0.316	0.284	-0.055	-0.036	-0.037
	(0.477)	(0.458)	(0.465)	(0.451)	(0.033)	(0.046)	(0.029)
Parent smoked	0.442	0.476	0.468	0.483	0.034	0.014	0.031
	(0.497)	(0.500)	(0.499)	(0.500)	(0.025)	(0.027)	(0.029)
No. smokers in househosld	0.726	0.736	0.708	0.761	0.007	-0.044	0.050
	(0.995)	(0.782)	(0.746)	(0.812)	(0.066)	(0.065)	(0.062)
Parents find smoking unacceptable	0.441	0.438	0.453	0.424	-0.002	0.022	-0.023
	(0.497)	(0.496)	(0.498)	(0.495)	(0.026)	(0.025)	(0.030)
Asset count	6.454	6.515	6.399	6.620	0.072	-0.072	0.158
	(1.264)	(1.344)	(1.326)	(1.353)	(0.078)	(0.100)	(0.087)
Joint significance test (p-value)					0.138	0.041	0.244
Panel D. Smoking abstinence outcomes							
Verified abstinence (1° outcome)	0.785	0.777	0.799	0.754	-0.006	-0.024	0.011
	(0.411)	(0.417)	(0.401)	(0.431)	(0.033)	(0.042)	(0.036)
Missing cotinine = smoker	0.680	0.669	0.671	0.667	-0.009	-0.002	-0.015
	(0.467)	(0.471)	(0.470)	(0.472)	(0.042)	(0.053)	(0.049)
Cotinine test only	0.845	0.840	0.856	0.824	-0.003	-0.015	0.009
	(0.362)	(0.367)	(0.351)	(0.381)	(0.029)	(0.038)	(0.032)
Self-report abstained $\geq$ 2 weeks	0.788	0.785	0.807	0.761	-0.001	-0.021	0.016
	(0.409)	(0.411)	(0.395)	(0.427)	(0.032)	(0.043)	(0.033)
Self-report abstained $\geq$ 1 month	0.734	0.717	0.741	0.691	-0.014	0.035	0.004
	(0.442)	(0.450)	(0.438)	(0.462)	(0.036)	(0.048)	(0.038)

Notes: The control group has 1153 students and the treated group has 1191 students. Standard deviations are in parentheses in Columns 1 to 4. Each cell in Columns 5-7 is the coefficient on the treatment variable from a separate OLS regression with district and school pair fixed effects. Column 5 compares the control and pooled treatment. Columns 6 and 7 compare the sub-treatment arm separately. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 2: Difference-in-differences estimates of program effects on verified smoking abstinence

	(1)	(2)	(3)	(4)	(5)
Panel A. Pooled estimates					
Treatment	-0.008	-0.008	-0.006	0.008	0.008
	(0.027)	(0.027)	(0.013)	(0.014)	(0.014)
Treatment × Post	0.038	0.037	0.044**	0.044*	
	(0.023)	(0.023)	(0.022)	(0.022)	
Post	-0.021				
	(0.016)				
Wave=2		-0.009	-0.017	-0.017	-0.016
		(0.018)	(0.018)	(0.018)	(0.019)
Wave=3		-0.034**	-0.043**	-0.043**	-0.044**
		(0.017)	(0.016)	(0.017)	(0.017)
Treatment × Wave=2					0.041*
					(0.025)
Treatment × Wave=3					0.046*
					(0.025)
R-squared	0.07	0.07	0.43	0.44	0.44
Panel B. Sub-treatment estimates					
Pledge	-0.012	-0.012	-0.016	-0.004	-0.003
	(0.042)	(0.042)	(0.017)	(0.017)	(0.017)
Competition	-0.004	-0.004	0.003	0.019	0.019
	(0.032)	(0.032)	(0.018)	(0.020)	(0.020)
Post	-0.021				
	(0.016)				
Pledge × Post	0.051**	0.051**	0.052**	0.051**	
	(0.023)	(0.023)	(0.022)	(0.023)	
Competition × Post	0.025	0.025	0.036	0.037	
	(0.031)	(0.031)	(0.030)	(0.030)	
Wave=2		0.025**	0.026**	0.026**	-0.016
		(0.012)	(0.012)	(0.012)	(0.019)
Wave=3		-0.034**	-0.043**	-0.043**	-0.044**
		(0.017)	(0.016)	(0.017)	(0.017)
Pledge × Wave=2					0.053*
					(0.028)
Pledge × Wave=3					0.049*
					(0.027)
Competition × Wave=2					0.030
					(0.030)
Competition × Wave=3					0.043
					(0.035)
R-squared	0.07	0.07	0.43	0.44	0.44
Equality of sub-treatments (p-value)	0.406	0.403	0.578	0.621	0.823
No. observations			7208		
Dep. variable mean for control group			0.786		
Pair fixed effects	Yes	Yes	Yes	Yes	Yes
Student and class baseline outcomes	No	No	Yes	Yes	Yes
Additional controls	No	No	No	Yes	Yes

Notes: This table presents estimates of Equations 1 (Panel A) and 2 (Panel B). Panel A pools the individual pledge and pledge plus competition arms; Panel B separates the sub-treatment arms. The dependent variable is verified smoking abstinence, defined as self-reporting abstinence and obtaining a negative cotinine test. *Post* equals 1 for the follow-up period. *Treatment* (*Pledge*, *Competition*) equals 1 if the school is randomized into treatment (pledge or pledge plus school competition arm). All columns include school pair fixed effects. Col. 1 pools the follow-up period; Cols. 2 to 4 replace *Post* with indicators for wave 2 (3 months post-intervention) and wave 3 (8 months post-intervention). Col. 3 adds baseline outcomes. Col. 4 adds: baseline ever smoked status, indicator for above median knowledge, indicator for having above-median friends who smoke, time inconsistency, risk seeking, household size, household land ownership, head of household in skilled occupation, any smoker in household, parents find smoking very unacceptable, asset count, and the number of male teachers who smoke in the school. Col. 5 interacts treatment with wave indicators. Standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 3: Potential mechanisms

Panel A.	Social learning: Tobacco knowledge	Social learning: Knowledge will help abstinence†	Social norms: Smoked with friends	Social norms: Friends not not smoking will help†	Social norms: Parents will not accept
Treatment	0.125	0.032	-0.000	0.032	-0.000
	(0.096)	(0.025)	(0.020)	(0.025)	(0.020)
Treatment × Wave=3	-0.020		-0.114***		0.024
	(0.222)		(0.030)		(0.022)
Wave=3	0.005		0.045**		-0.026
	(0.168)		(0.019)		(0.018)
No. observations	5454	2590	4606	2590	4605
R-squared	0.234	0.096	0.254	0.078	0.096
Dependent variable mean	5.618	0.548	0.472	0.377	0.701

Panel B.	Monitoring and fear of punishment: Demerits†	Testing†	Student council reports†	Teacher notified†	Parent notified†
Treatment	-0.021	0.032	0.059**	0.025	0.018
	(0.030)	(0.028)	(0.024)	(0.025)	(0.024)
No. observations	2590	2590	2590	2590	2590
R-squared	0.091	0.087	0.088	0.099	0.104
Dependent variable mean	0.334	0.296	0.258	0.341	0.343

Panel C.	Motivation and effort: Smoking intentions	Motivation and effort: Report friends who smoke†	Salience: Discuss with parents†	Salience: Discuss with friends†
Treatment	0.062	-0.012	-0.026*	-0.041**
	(0.085)	(0.023)	(0.013)	(0.017)
Treatment × Wave=3	-0.135			
	(0.130)			
Wave=3	0.249**			
	(0.101)			
No. observations	5337	2590	2589	2589
R-squared	0.209	0.083	0.045	0.069
Dependent variable mean	1.139	0.254	0.087	0.166

Notes: Outcomes marked with † are available in wave 3 only. *Tobacco knowledge* is the number of correct responses to 9 questions to assess students' tobacco knowledge. *Knowledge will help abstinence* is an indicator for agreeing with the statement that better tobacco knowledge will help them abstain from smoking. *Smoked with friends* is an indicator for reporting smoking with friends. *Friends not smoking will help* is an indicator for agreeing that having friends who do not smoke will help students abstain from smoking. *Demerit*, *testing*, *student council reports*, *teacher notified*, and *parent notified* are indicators for agreeing that introducing each item in school would help students abstain from smoking. *Smoking intentions* in the next 3 months ranges from 0 to 10, with 10 corresponding to "will definitely smoke." *Discuss with parents (friends)* is an indicator for any discussion regarding smoking with parents (friends) in the 6 months prior to the survey. *Treatment* takes the value 1 if the school is randomized into treatment (the pledge arm or plus plus school competition arm). *Post* equals 1 for the follow-up period. Covariates include baseline ever smoked status, indicator for above median knowledge, indicator for having above median friends who smoke, time inconsistency, risk seeking, household size, household land ownership, head of household in skilled occupation, any smoker in household, parents find smoking very unacceptable, asset count, and the number of male teachers who smoke in the school. The baseline outcome included for outcomes that are available at baseline and follow-up. School pair fixed effects included. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 4: Selection into biochemical verification

	(1) Missing cotinine test	(2) Self-reported abstinence and failed cotinine test	(3) Frequency of being reported
Panel A. Pooled estimates			
Treatment	-0.000	0.004	-1.529***
	(0.015)	(0.011)	(0.409)
Treatment × Post	-0.003	-0.003	
	(0.023)	(0.016)	
Wave = 2	-0.127***	-0.015	
	(0.019)	(0.011)	
Wave = 3	-0.035*	-0.005	1.813***
	(0.019)	(0.012)	(0.429)
No. observations	7825	7208	2698
R-squared	0.604	0.124	0.252
Dependent variable mean	0.132	0.078	6.107
Panel B. Sub-treatment estimates			
Pledge	0 -0.008	-0.012	-1.090*
	(0.017)	(0.014)	(0.577)
Competition	0.007	0.019	-1.963***
	(0.019)	(0.016)	(0.616)
Pledge × Post	0.014	0.003	
	(0.025)	(0.017)	
Competition × Post	-0.018	-0.009	
	(0.029)	(0.022)	
Wave = 2	-0.127***	-0.015	
	(0.019)	(0.011)	
Wave = 3	-0.035*	-0.005	1.810***
	(0.019)	(0.012)	(0.429)
No. observations	7825	7208	2698
R-squared	0.605	0.125	0.253
Dependent variable mean	0.132	0.078	6.108
Equality of sub-treatments (p-value)	0.28	0.64	0.31

Notes: *Missing cotinine test* takes the value 1 if an individual has self-reported tobacco abstinence status only. *Self-reported abstinence and failed cotinine test* takes the value 1 if an individual reported abstinence but failed the cotinine test. *Frequency of being reported* is the number of times a student is reported by friends; this measure is only available in waves 2 and 3. *Post* takes the value one for the follow-up period. *Treatment* (*Pledge*, *Competition*) takes the value 1 if the school is randomized into treatment (the pledge arm or pledge plus school competition arm). All models include school pair fixed effects, student and class baseline outcomes, and baseline ever smoked status, indicator for above median knowledge, indicator for having above median friends who smoke, time inconsistency, risk seeking, household size, household land ownership, head of household in skilled occupation, any smoker in household, parents find smoking very unacceptable, asset count, and the number of male teachers who smoke in the school. Standard errors clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 5: Alternative tobacco abstinence outcomes

	(1) Missing cotinine = smoker	(2) Missing cotinine = baseline	(3) Cotinine test only	(4) Self-report abstained ≥ 2 weeks	(5) Self-report abstained ≥ 1 month
Panel A. Pooled estimates					
Treatment	0.006	0.007	0.005	0.001	0.004
	(0.015)	(0.016)	(0.012)	(0.014)	(0.015)
Treatment × Post	0.033	0.044*	0.023	0.033	0.066***
	(0.023)	(0.024)	(0.019)	(0.023)	(0.025)
Wave = 2	-0.037**	0.074***	-0.011	-0.761***	-0.032
	(0.017)	(0.023)	(0.013)	(0.023)	(0.021)
Wave = 3	-0.062***	-0.015	-0.046***	-0.062***	-0.056***
	(0.016)	(0.019)	(0.014)	(0.017)	(0.018)
No. observations	7380	8074	7208	6328	6328
R-squared	0.400	0.496	0.444	0.680	0.461
Dependent variable mean	0.786	0.682	0.846	0.784	0.734
Panel B. Sub-treatment estimates					
Pledge	-0.008	0.001	-0.009	-0.012	-0.007
	(0.017)	(0.022)	(0.013)	(0.017)	(0.017)
Competition	0.020	0.013	0.017	0.014	0.015
	(0.021)	(0.019)	(0.019)	(0.018)	(0.021)
Pledge × Post	0.039	0.038	0.033*	0.050*	0.073***
	(0.025)	(0.029)	(0.017)	(0.026)	(0.025)
Competition × Post	0.027	0.050*	0.014	0.016	0.060*
	(0.031)	(0.028)	(0.029)	(0.028)	(0.034)
Wave = 2	-0.037**	0.074***	-0.011	-0.761***	-0.033
	(0.017)	(0.023)	(0.013)	(0.023)	(0.021)
Wave = 3	-0.062***	-0.015	-0.046***	-0.062***	-0.056***
	(0.016)	(0.019)	(0.014)	(0.017)	(0.018)
No. observations	7380	8074	7208	6328	6328
R-squared	0.400	0.496	0.444	0.461	0.680
Dependent variable mean	0.400	0.496	0.444	0.680	0.461
Equality of sub-treatments (p-value)	0.733	0.709	0.517	0.249	0.681

Notes: Column 1 assumes students with a missing cotinine test smoked. Column 2 assumes students with a missing cotinine test had the same outcome as their baseline outcome. Column 3 uses the cotinine test only. Columns 4 and 5 combine cotinine test results and self-reported abstinence in the prior 2 weeks and 1 month, respectively. *Post* takes the value 1 for the follow-up period. *Treatment* (*Pledge*, *Competition*) takes the value 1 if the school is randomized into treatment (the pledge arm or the pledge plus school competition arm). All models include school pair fixed effects, student and class baseline outcomes, and baseline ever smoked status, indicator for above median knowledge, indicator for having above median friends who smoke, time inconsistency, risk seeking, household size, household land ownership, head of household in skilled occupation, any smoker in household, parents find smoking very unacceptable, asset count, and the number of male teachers who smoke in the school. Standard errors clustered at the school level.
Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Appendix A Sample Selection and Evaluation

A.1 Sample size calculation

We conducted sample size calculations for a cluster randomized trial with an individual-level binary outcome.[22] The estimates are based on pairwise comparisons between equal-sized study arms using a test of difference in proportions and assuming a two-sided alternative hypothesis. The model includes a random intercept for each school. The pairwise comparisons indicate that the study has sufficient power to detect differences in the probability of smoking of 5 percentage points or greater compared to the control arm. We based our calculations on the the meta-analysis of school-based smoking prevention programs by Thomas and Perera (2006). Combined social competence and social influences curricula are associated with an odds ratio of 0.49 (95% CI 0.28 to 0.87), while social competence curricula are associated with an odds ratio of 0.52 (95% CI 0.30 to 0.88). Using this minimum detectable effect, the study will have 80% power to detect an intervention-related reduction in smoking rates for the treatment group with a sample size of 72 schools.

We identified 78 schools in Kulon Progo and Sleman, but excluded 6 schools due to size and distance from the city of Yogyakarta. The intervention was implemented in 72 schools located in 19 subdistricts. These subdistricts are drawn from two districts: Sleman and Kulon Progo. We included 11 subdistricts from Kulon Progo and 8 from Sleman. We selected subdistricts in Sleman that are geographically proximate to Kulon Progo and have a similar socioeconomic profile. We exclude Kalasan and Ngaglik since they are the primary tobacco producing subdistricts in Sleman.[23]

Class size is typically between 12 and 20 students. We excluded schools with fewer than 12 male students in 7th grade to limit the project budget. We also excluded private schools that tend to draw students from a more privileged background. We identified 78 eligible schools. Due to transportation costs, we further excluded the 6 schools located farthest from Yogyakarta proper. If a school had one to three 7th grade classes, we included all male students, up to 60 students per school. If a school had four of more 7th grade classes, we randomly selected two or three classes to include in the sample, up to 60 students per school.

[22] Calculations were performed using Optimal Design Software for Multi-level and Longitudinal Research, version 3.01. The software is available for free download at: https://sites.google.com/site/optimaldesignsoftware/home.

[23] Some tobacco is also grown in the following subdistricts: Prambanan, Ngemplak, Sleman, Seyegan, and Tempel. Source: http://jogja.tribunnews.com/2015/04/19/dua-kecamatan-jadi-sentra-produksi-tembakau. Sleman borders the province of Central Java, where 25% of tobacco is grown.

A.2 Randomization

We used a pair matching procedure to randomly allocate schools to intervention arms. Studies have shown that pair matching outperforms other randomization methods in balancing arms (Imai et al., 2009; Bruhn and McKenzie, 2009). We formed pairs so as to minimize the Mahalanobis distance between the values of selected covariates so as to obtain pairs that are close in their covariate values. We then randomly assigned one unit to the treatment group and one unit to the control group. Mahalanobis distance is a scale-invariant distance metric that takes the inverse of covariates' variance-covariance matrix, thereby finding the pairs of units that are closest in multi-dimensional covariate space (King and Nielsen, 2019). Similarly, among the treatment group schools, we paired the two most similar schools by Mahalonobis distance and randomly assigned one to the commitment pledge arm and one to the pledge plus school competition arm. Thus, we randomly allocated 36 schools to the control arm, 18 schools to the commitment arm, and 18 schools to the commitment pledge plus school competition arm (Table A.1).

Table A.1: Number of schools by subdistrict

District	Control	Pledge	School competition
Sleman	17	12	7
Kulon Progo	19	6	11
No. schools	36	18	18

Our procedure matched on the following covariates: district, subdistrict, distance from the school to Yogyakarta proper, number of male students and classrooms in 7th grade, total number of students, teachers, classes, and student council members in the school, electricity capacity, and average national examination scores in mathematics. We combined information on school characteristics and enrollment (age, gender distribution, class size, and students' scores on national examination) from the website of the national education ministry (http://sekolah.data.kemdikbud.go.id) and information from the local education departments in Kulon Progo and Sleman. The matching procedure was performed in R.

Table A.2: Differences in baseline school characteristics by study group

	p-value from difference in means by study group		
	(1)	(2)	(3)
	Pooled	Individual pledge	School competition
Distance to Yogjakarta	0.929	0.325	0.378
Student-to-teacher ratio	0.779	0.829	0.589
Average mathematics score	0.269	0.563	0.488
No. classes	0.924	0.293	0.348
No. schools	72	72	72

Notes: Each cell in Column 1 is the p-value of a separate t-test comparison between the control and treatment at the school level, Column 2 compares the individual pledge arm, Column 3 compares the school competition arm.

A study investigator implemented the random allocation sequences using computer-generated random numbers, concealing the sequence from field staff, school personnel, and students until after the baseline survey was completed.

A.3 Survey

The survey includes separate questionnaires for students, parents, teachers, and principals. Students were surveyed at baseline at 8 months post intervention. Parents and teachers were surveyed at baseline only.

The questions on tobacco knowledge were given to students, teachers and principals. The questions include: whether smoking is harmful to smokers' health, whether all cigarettes are equally harmful, whether smoking is dangerous to nonsmokers, and whether smoking causes the following diseases: stroke, impotence for male smokers, premature aging, chronic obstructive pulmonary disease (COPD), heart attacks and heart failure, and miscarriages during pregnancy. The correct answer is 'yes' to all the questions. Each correct response corresponds to one point for the respondent.

Risk preference is based on a hypothetical gamble using a yellow and blue marble. Students were asked to choose one out of six potential gambles with the following payoffs:

Table A.3: Hypothetical gambles

	Yellow	Blue
Option 1	IDR 10,000	IDR 10,000
Option 2	IDR 8,000	IDR 15,000
Option 3	IDR 6,000	IDR 19,000
Option 4	IDR 4,000	IDR 24,000
Option 5	IDR 2,000	IDR 28,500
Option 6	IDR 500	IDR 30,000

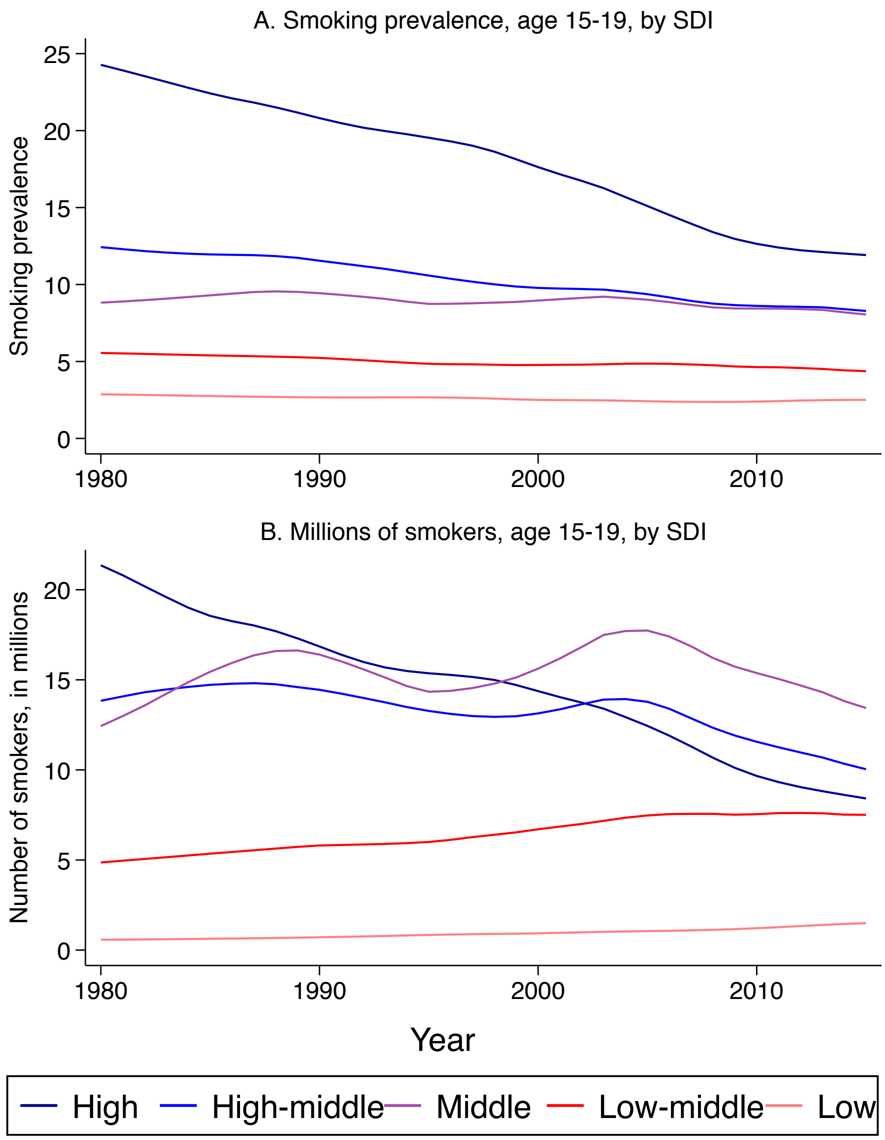
Notes: 1 USD is approximately IDR 14,000. Students were given two examples before they were asked to make their choice.

We code a respondent as risk seeking when he chooses option 4, 5, or 6.

Time inconsistency is based on the hypothetical timing of receiving money. We asked the following two questions: (i) whether students would rather receive IDR 100,000 today or IDR 200,000 in six months, (ii) whether students would rather receive IDR 100,000 in six months or IDR 200,000 in seven months. We code a respondent as time inconsistent when he first chooses to wait for 6 months (in (i)), but not willing to wait 7 months next (in (ii)).

Appendix B Additional Figures and Tables

Figure B.1: Smoking trends among youth



Notes: The data are drawn from the Global Burden of Disease Study 2015, available at http://ghdx.healthdata.org/record/ihme-data/gbd-2015-smoking-prevalence-1980-2015. The Socio-Demographic Index (SDI), created by the GBD Study, is a summary measure of development, comprised of income per capita, educational attainment, and total fertility rate. Here, the SDI is split into quintiles based on 2015 values. Indonesia is in the "middle" grouping.

Figure B.2: Individual commitment pledge for students

At the end of the information session, students in treated schools were invited to sign the document below.

Fill in the column with your private promise:

I promise to abstain from smoking while I am a student of [School name] because smoking can harm my health, including to causing me to risk premature death from serious diseases such as respiratory disease, heart disease, cancer, etc.

Yogyakarta,

( )

PONTAR

Perjanjian untuk Tidak Merokok

After students filled in their private promise, they were invited to sign a separate document with the same statement with the rest of the class. The group document was then presented to the guidance councilor, who was encouraged to display the document in the classroom.

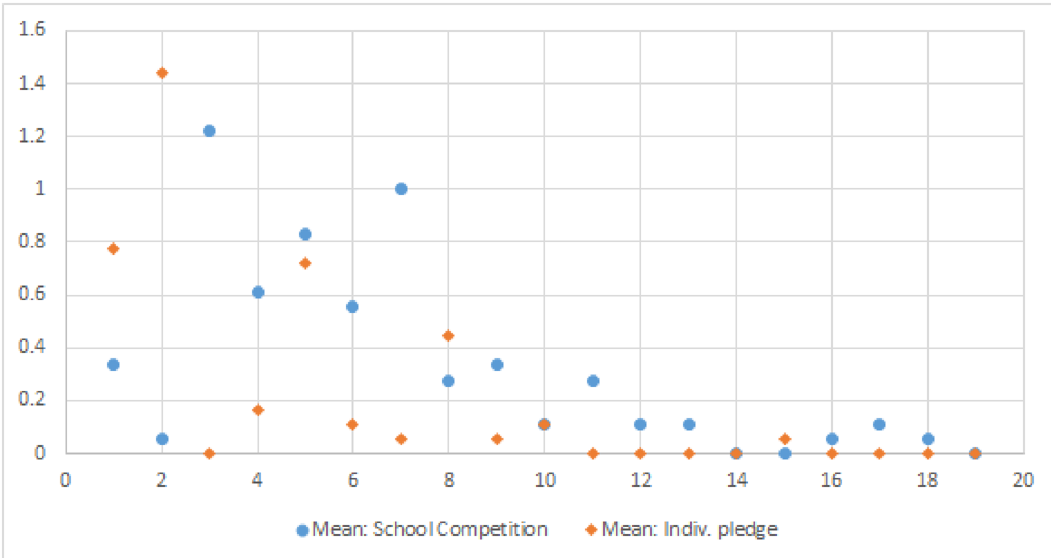
Figure B.3: Commitment pledge for parents

The following statement is attached to the consent forms that parents receive at baseline:

For this program, we invite your and your son's signature as proof of your son's promise to abstain from smoking. If your son smoked, your son would receive 10 demerit points and the school would report this to you. These demerit points would be tallied at the end of the academic year. Your son's success or failure to comply will be discussed when you receive your son's report card. We hope you would provide your support to help your son avoid smoking.

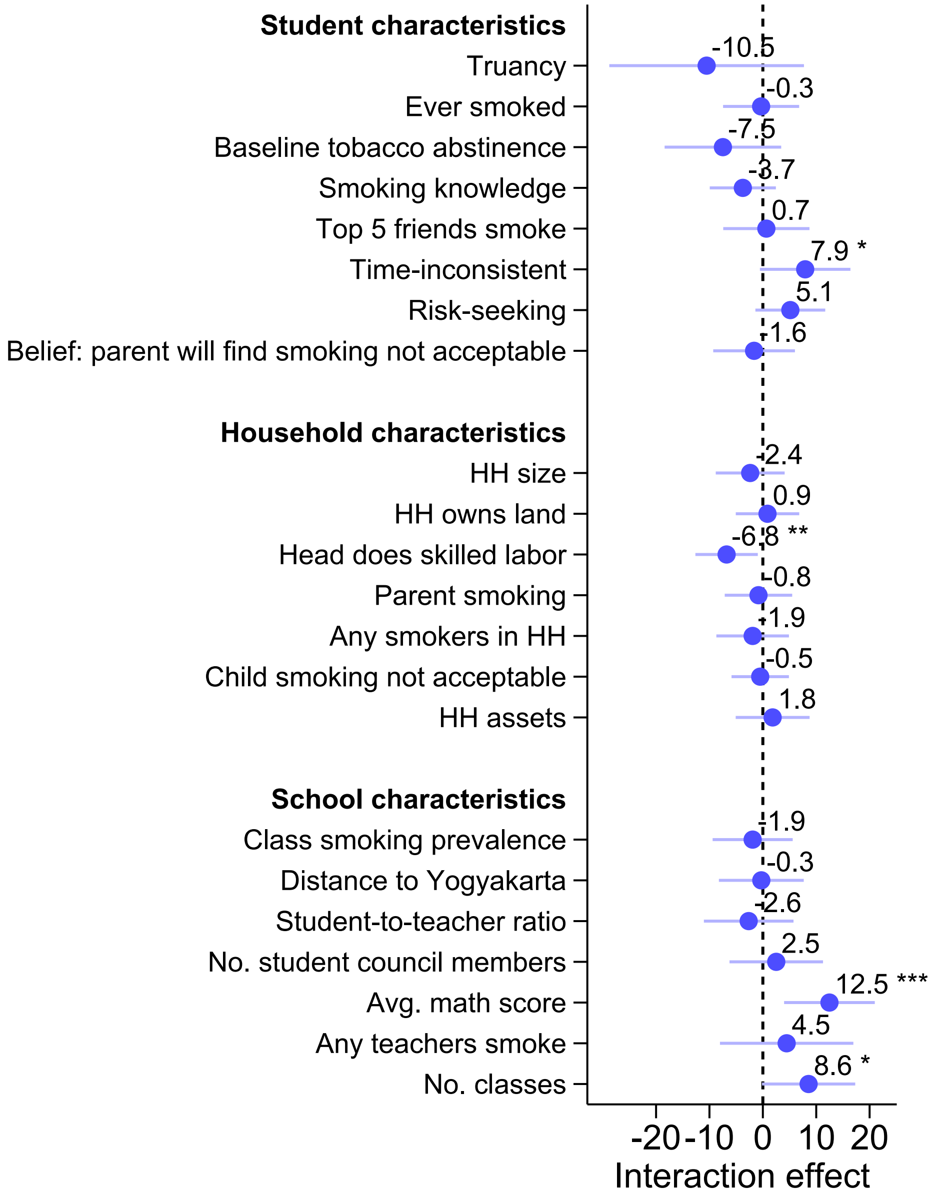
[Parent's signature] [Child's signature]

Figure B.4: Average number of reported smokers in treated group by week



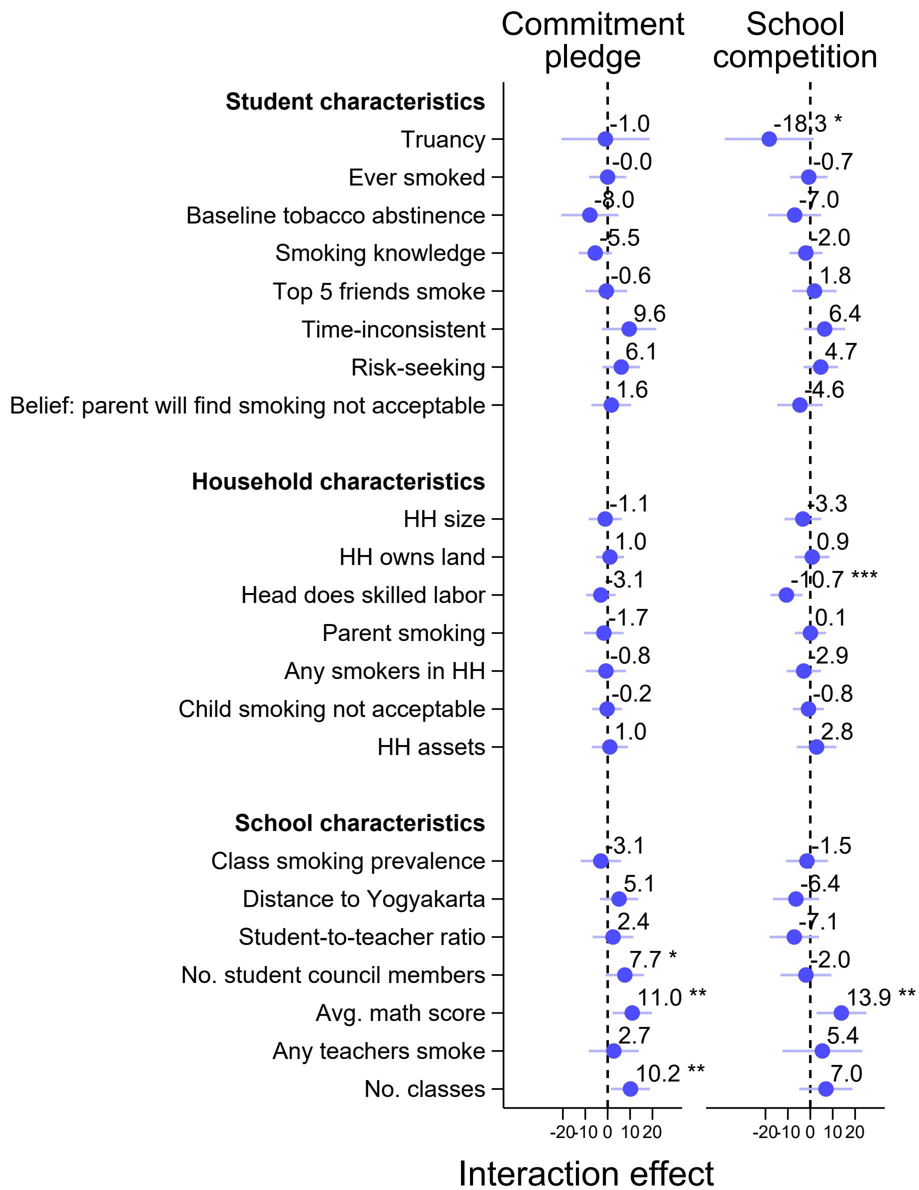
Notes: This figure shows the average number of smokers reported in each sub-treatment arm in each week after the intervention was implemented. Differences between the two sub-treatment arms are not statistically significant.

Figure B.5: Heterogeneous program effects



Notes: This figure shows the effect of the pooled intervention on smoking abstinence by baseline student, household, and school characteristics. Continuous variables are dichotomized into the median split. The plot shows the interaction effects, expressed in percentage points, estimated from a difference-in-differences model that adjusts for the full set of baseline covariates. Error bars are 95% confidence intervals. Robust standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

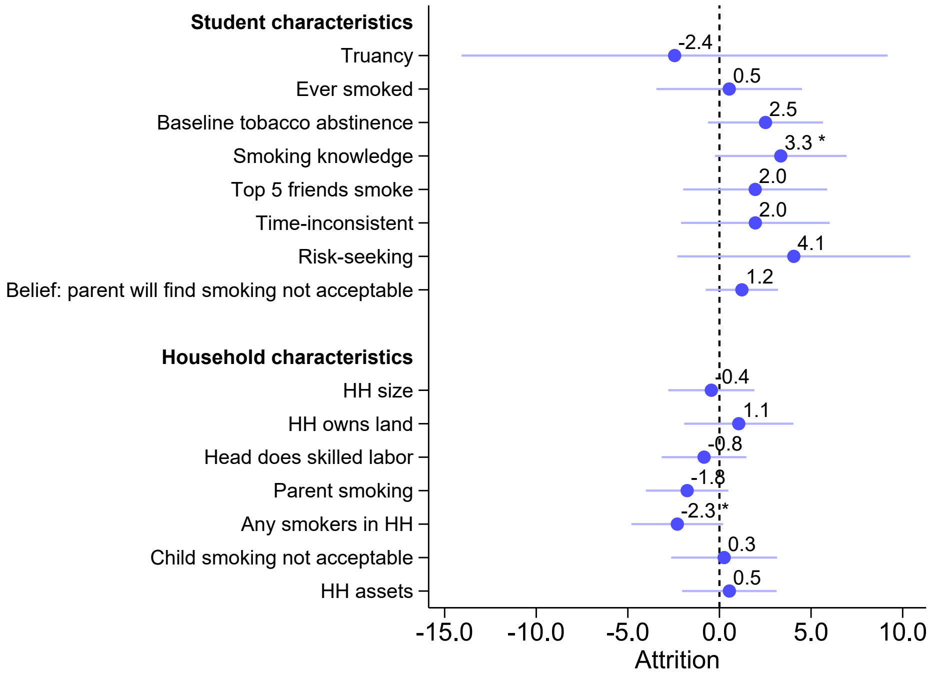
Figure B.6: Heterogeneous program effects: sub-treatment



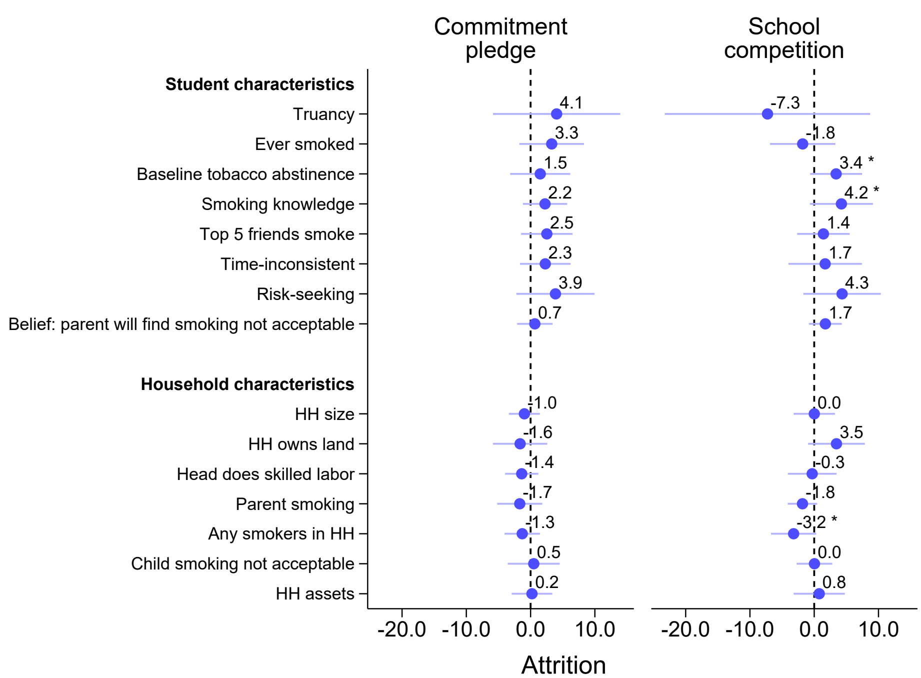
Notes: This figure shows the effect of the sub-treatment arms on smoking abstinence by baseline student, household, and school characteristics. Continuous variables are dichotomized into the median split. The plot shows the interaction effects, expressed in percentage points, estimated from a difference-in-differences model that adjusts for the full set of baseline covariates. Error bars are 95% confidence intervals. Robust standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Figure B.7: Attrition across control and treatment arms

Panel A. Pooled



Panel B. Sub-treatment



Notes: Panel A shows the probability of a student being in all three waves of the survey across the treatment and control groups by baseline student and household characteristics. Panel B shows the interaction term between sub-treatment and the variable of interest at baseline, expressed in percentage points. Continuous variables are dichotomized into the median split. The plot shows the interaction term between treatment and the variable of interest at baseline, expressed in percentage points. Error bars are 95% confidence intervals. Pair FE included, robust standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.1: Difference-in-differences estimates of program effects on verified smoking abstinence (full output)

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.006	0.008	0.008			
	(0.013)	(0.014)	(0.014)			
Treatment × Post	0.044**	0.044*				
	(0.022)	(0.022)				
Treatment × Wave=2			0.041*			
			(0.025)			
Treatment × Wave=3			0.046*			
			(0.025)			
Pledge				-0.016	-0.004	-0.003
				(0.017)	(0.017)	(0.017)
Competition				0.003	0.019	0.019
				(0.018)	(0.020)	(0.020)
Pledge × Post				0.052**	0.051**	
				(0.022)	(0.023)	
Competition × Post				0.036	0.037	
				(0.030)	(0.030)	
Pledge × Wave=2						0.053*
						(0.028)
Pledge × Wave=3						0.049*
						(0.027)
Competition × Wave=2						0.030
						(0.030)
Competition × Wave=3						0.043
						(0.035)
Wave=2	-0.017	-0.017	-0.016	0.026**	0.026**	-0.016
	(0.018)	(0.018)	(0.019)	(0.012)	(0.012)	(0.019)
Wave=3	-0.043**	-0.043**	-0.044**	-0.043**	-0.043**	-0.044**
	(0.016)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)
Own baseline verified abstinence	0.623***	0.587***	0.587***	0.623***	0.587***	0.587***
	(0.019)	(0.018)	(0.018)	(0.019)	(0.018)	(0.018)
Avg. class baseline verified abstinence	0.134***	0.110**	0.110**	0.135***	0.110**	0.110**
	(0.046)	(0.044)	(0.044)	(0.046)	(0.044)	(0.044)
Missing own baseline	0.493***	0.469***	0.469***	0.493***	0.468***	0.469***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Truancy		-0.057**	-0.057**		-0.057**	-0.058**
		(0.023)	(0.023)		(0.023)	(0.023)
Ever smoked (at baseline)		-0.075***	-0.075***		-0.075***	-0.075***
		(0.011)	(0.011)		(0.011)	(0.011)
Above-median tobacco knowledge (at baseline)		0.005	0.005		0.005	0.005
		(0.008)	(0.008)		(0.008)	(0.008)
More than 1 friend smoked (at baseline)		-0.049***	-0.049***		-0.049***	-0.049***
		(0.010)	(0.010)		(0.010)	(0.010)
Time inconsistent		0.002	0.002		0.002	0.002
		(0.010)	(0.010)		(0.010)	(0.010)
Risk seeking		-0.010	-0.010		-0.010	-0.010
		(0.010)	(0.010)		(0.010)	(0.010)
No. teachers who smoked		-0.007*	-0.007*		-0.007*	-0.007*
		(0.004)	(0.004)		(0.004)	(0.004)
Household size		-0.004	-0.004		-0.005	-0.005
		(0.003)	(0.003)		(0.003)	(0.003)
Household own land		0.005	0.005		0.005	0.005
		(0.009)	(0.009)		(0.009)	(0.009)
Household head in skilled occupation		0.003	0.003		0.003	0.003
		(0.010)	(0.010)		(0.010)	(0.010)
Parent smoked		-0.037***	-0.037***		-0.037***	-0.037***
		(0.008)	(0.008)		(0.008)	(0.008)
Belief that smoking is unacceptable		0.016*	0.016*		0.016*	0.016*
		(0.008)	(0.008)		(0.008)	(0.008)
Missing occupation		0.014	0.014		0.014	0.014
		(0.013)	(0.013)		(0.013)	(0.013)
Household asset count		0.002	0.002		0.002	0.002
		(0.004)	(0.004)		(0.004)	(0.004)
No. observations	7208	7208	7208	7208	7208	7208
R-squared	0.43	0.44	0.44	0.43	0.44	0.44
Dep. var. mean for control	0.786	0.786	0.786	0.786	0.786	0.786
Equality of sub-treatments (p-value)				0.578	0.621	0.823

Notes: This table presents estimates of Equation 1 (Columns 1-3, pooled) and Equation 2 (Columns 4-6, by sub-treatment). *Treatment* (*Pledge*, *Competition*) equals 1 if the school is randomized into treatment (pledge or pledge plus school competition arm). All columns include school pair fixed effects. Columns 3 and 6 interact treatment with wave indicators. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.2: Potential mechanisms: Sub-treatment arms

Panel A.	Social learning: Tobacco knowledge	Social learning: Knowledge will help abstinence†	Social norms: Smoked with friends	Social norms: Friends not not smoking will help†	Social norms: Parents will not accept
Pledge	0.118	0.025	0.005	0.002	0.019
	(0.138)	(0.027)	(0.021)	(0.026)	(0.026)
Competition	0.132	0.038	-0.006	0.060	-0.019
	(0.132)	(0.039)	(0.032)	(0.040)	(0.025)
Pledge × Wave=3	0.021		-0.120***		0.016
	(0.294)		(0.028)		(0.028)
Competition × Wave=3	-0.056		-0.109**		0.031
	(0.239)		(0.044)		(0.026)
Wave=3	0.005		0.045**		-0.026
	(0.168)		(0.019)		(0.018)
No. observations	5454	2590	4606	2590	4605
R-squared	0.234	0.096	0.254	0.079	0.096
Dependent variable mean	5.618	0.548	0.472	0.377	0.701
Equality of sub-treatments (p-value)	0.795	0.676	0.814	0.218	0.617

Panel B.	Monitoring and fear of punishment: Demerits†	Testing†	Student council reports†	Teacher notified†	Parent notified†
Pledge	-0.043	0.008	0.033	0.017	0.032
	(0.030)	(0.031)	(0.025)	(0.031)	(0.032)
Competition	-0.000	0.053	0.082**	0.032	0.006
	(0.047)	(0.045)	(0.037)	(0.042)	(0.034)
No. observations	2590	2590	2590	2590	2590
R-squared	0.092	0.087	0.089	0.099	0.104
Dep. var. mean	0.334	0.296	0.258	0.341	0.343
Equality of sub-treatments (p-value)	0.431	0.408	0.265	0.773	0.572

Panel C.	Motivation and effort: Smoking intentions	Motivation and effort: Report friends who smoke†	Salience: Discuss with parents†	Salience: Discuss with friends†
Pledge	0.084	-0.002	-0.026	-0.040
	(0.105)	(0.034)	(0.021)	(0.025)
Competition	0.042	-0.020	-0.025	-0.042**
	(0.112)	(0.033)	(0.017)	(0.020)
Pledge × Wave=3	-0.178			
	(0.140)			
Competition × Wave=3	-0.097			
	(0.162)			
Wave=3	0.249**			
	(0.101)			
No. observations	5337	2590	2589	2589
R-squared	0.209	0.083	0.045	0.069
Dependent variable mean	1.139	0.254	0.087	0.166
Equality of sub-treatments (p-value)	0.608	0.719	0.960	0.954

Notes: Outcomes marked with † are available in wave 3 only. *Tobacco knowledge* is the number of correct responses to 9 questions to assess students' tobacco knowledge. *Knowledge will help abstinence* is an indicator for agreeing with the statement that better tobacco knowledge will help them abstain from smoking. *Smoked with friends* is an indicator for reporting smoking with friends. *Friends not smoking will help* is an indicator for agreeing that having friends who do not smoke will help students abstain from smoking. *Demerit*, *testing*, *student council reports*, *teacher notified*, and *parent notified* are indicators for agreeing that introducing each item in school would help students abstain from smoking. *Smoking intentions* in the next 3 months ranges from 0 to 10, with 10 corresponding to "will definitely smoke." *Discuss with parents (friends)* is an indicator for any discussion regarding smoking with parents (friends) in the 6 months prior to the survey. *Pledge* (*Competition*) takes the value 1 if the school is randomized to the pledge (pledge plus school competition) arm. *Post* equals 1 for the follow-up period. Covariates are the same as those in Table 3. School pair fixed effects included. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.3: Heterogeneous program effects on smoking abstinence by student characteristics

	(1) Pooled	(2) Pledge	(3) Pledge plus competition
Truancy	-0.105	-0.010	-0.183*
	(0.091)	(0.099)	(0.100)
Ever smoked	-0.003	-0.000	-0.007
	(0.036)	(0.042)	(0.042)
Baseline status	-0.075	-0.080	-0.070
	(0.055)	(0.064)	(0.059)
Smoking knowledge	-0.037	-0.055	-0.020
	(0.031)	(0.037)	(0.037)
Top 5 friend smokes	0.007	-0.006	0.018
	(0.041)	(0.047)	(0.049)
Time-inconsistent	0.079*	0.096	0.064
	(0.043)	(0.061)	(0.046)
Risk-seeking	0.051	0.061	0.047
	(0.033)	(0.042)	(0.039)
Parent finds smoking unacceptable	-0.016	0.016	-0.046
	(0.038)	(0.044)	(0.051)

Notes: This table shows the effect of the pooled intervention and sub-treatment arms on smoking abstinence by baseline student characteristics, estimated from difference-in-differences models that interacts each covariate with the treatment indicator and adjusts for the full set of baseline covariates. Continuous variables are dichotomized into the median split. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.4: Heterogeneous program effects on smoking abstinence by household characteristics

	(1) Pooled	(2) Pledge	(3) Pledge plus competition
Household size	-0.024	-0.011	-0.033
	(0.032)	(0.037)	(0.041)
Household owns land	0.009	0.010	0.009
	(0.030)	(0.032)	(0.039)
Household head does skilled labor	-0.068**	-0.031	-0.107***
	(0.029)	(0.033)	(0.036)
Parent a smoker	-0.008	-0.017	0.001
	(0.032)	(0.044)	(0.034)
Any smokers in household	-0.019	-0.008	-0.029
	(0.034)	(0.045)	(0.038)
Parent finds child smoking unacceptable	-0.005	-0.002	-0.008
	(0.027)	(0.034)	(0.035)
Household assets	0.053	0.040	0.065
	(0.043)	(0.060)	(0.051)

Notes: This table shows the effect of the pooled intervention and sub-treatment arms on smoking abstinence by baseline household characteristics, estimated from difference-in-differences models that interacts each covariate with the treatment indicator and adjusts for the full set of baseline covariates. Continuous variables are dichotomized into the median split. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.5: Heterogeneous program effects on smoking abstinence by school characteristics

	(1) Pooled	(2) Pledge	(3) Pledge plus competition
Class smoking prevalence	-0.019	-0.031	-0.015
	(0.038)	(0.045)	(0.047)
Distance to Yogyakarta	-0.003	0.051	-0.064
	(0.040)	(0.043)	(0.052)
Student-to-teacher ratio	-0.026	0.024	-0.071
	(0.042)	(0.045)	(0.055)
No. student council members	0.025	0.077*	-0.020
	(0.044)	(0.043)	(0.057)
Average mathematics score	0.125***	0.110**	0.139**
	(0.043)	(0.044)	(0.055)
Any teachers smoke	0.045	0.027	0.054
	(0.063)	(0.056)	(0.089)
No. classes	0.086*	0.102**	0.070
	(0.044)	(0.044)	(0.060)

Notes: This table shows the effect of the pooled intervention and sub-treatment arms on smoking abstinence by baseline school characteristics, estimated from difference-in-differences models that interacts each covariate with the treatment indicator and adjusts for the full set of baseline covariates. Continuous variables are dichotomized into the median split. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.6: Cost-benefit calculations of program

Delayed initiation	1 year	2 years	3 years
Effectiveness (%-point ↓ smoking)	2.33	0.78	0.26
Cost per smoker averted	184	553	1,662
Benefits:			
Delayed cardiovascular disease	257	249	242
Delayed pulmonary disease	450	437	424
Productivity gain	2629	5183	7661
Benefit-to-cost difference	3151	5315	6665
Benefit-to-cost ratio	18.1	10.6	5.01

Notes: Our program effect at 8 months is approximately 4 percentage points. We assume that the program effect will be one-third as large every 12 months. Based on the expected 12-month effect, the cost per smoker averted would be $220. The cost per smoker increases as program effectiveness falls. We assume life expectancy is 67, retirement age at 60, GDP per capita at USD 3800. Costs due to second hand smoke are excluded. The gains are based on increasing labor market participation for 1 to 3 years, with a 3% discount rate. About 40% of smokers suffer from cardiovascular disease with a cost of $2,500. About 40% suffer from chronic obstructive pulmonary disease (COPD) with an associated cost of 5 days of missed activities and an annual 12% probability of $150 hospitalization cost. Under these assumptions, cardiovascular disease would occur between the ages of 58 to 60 and pulmonary disease onset would be between the ages of 39 to 41.

Table B.7: Attrition across treatment arms

	(1) Pooled	(2) Pledge	(3) Pledge plus competition
Panel A. Individual level			
Probability of being in 3 waves	-0.002	-0.005	0.001
	(0.030)	(0.010)	(0.054)
No. observations	2764	2764	2764
R-squared	0.45	0.45	0.45
Panel B. School level			
Share of students in 3 waves	0.222	0.056	0.167
	(0.233)	(0.169)	(0.164)
No. observations	72	72	72
R-squared	0.02	0.34	0.35

Notes: Panel A examines the probability of a student remaining in all 3 waves of the survey in the pooled treatment (Column 1), individual pledge (Column 2), and pledge plus school competition arms (Column 3). Similarly, Panel B examines the share of students remaining in all 3 waves of the survey. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.8: Panel sample: Difference-in-differences estimates of program effects on verified smoking abstinence

	(1)	(2)	(3)	(4)	(5)
Panel A. Pooled estimates					
Treatment	-0.032	-0.032	-0.014	-0.002	
	(0.026)	(0.026)	(0.012)	(0.013)	
Treatment × Post	0.060***	0.060***	0.059***	0.059***	
	(0.021)	(0.021)	(0.021)	(0.021)	
Post	-0.037**				
	(0.016)				
Wave=2		-0.029*	-0.028	-0.028	-0.027
		(0.017)	(0.017)	(0.017)	(0.019)
Wave=3		-0.046***	-0.044***	-0.044**	-0.044**
		(0.017)	(0.017)	(0.017)	(0.018)
Treatment=1 × Wave=2					0.058**
					(0.023)
Treatment=1 × Wave=3					0.059**
					(0.025)
R-squared	0.08	0.08	0.41	0.43	0.43
Panel B. Sub-treatment estimates					
Pledge	-0.023	-0.023	-0.018	-0.009	-0.009
	(0.041)	(0.041)	(0.018)	(0.016)	(0.016)
Competition	-0.041	-0.041	-0.011	0.005	0.005
	(0.029)	(0.029)	(0.016)	(0.018)	(0.018)
Post	-0.037**				
	(0.016)				
Pledge × Post	0.060***	0.060***	0.059***	0.059***	
	(0.022)	(0.022)	(0.022)	(0.022)	
Competition × Post	0.059**	0.059**	0.059**	0.059**	
	(0.027)	(0.027)	(0.027)	(0.027)	
Wave=2		0.017	0.017	0.017	-0.027
		(0.012)	(0.012)	(0.012)	(0.019)
Wave=3		-0.046***	-0.044***	-0.044**	-0.044**
		(0.017)	(0.017)	(0.017)	(0.018)
Pledge × Wave=2					0.062**
					(0.025)
Pledge × Wave=3					0.055*
					(0.028)
Competition × Wave=2					0.055*
					(0.028)
Competition × Wave=3					0.063*
					(0.032)
R-squared	0.08	0.08	0.41	0.43	0.43
Equality of sub-treatments (p-value)	0.960	0.960	0.986	0.994	0.963
No. observations			6438		
Dep. variable mean for control group			0.807		
Pair fixed effect	Yes	Yes	Yes	Yes	Yes
Student and class baseline outcomes	No	No	Yes	Yes	Yes
Additional controls	No	No	No	Yes	Yes

Notes: This table re-estimates Equations 1 (Panel A) and 2 (Panel B) after restricting the sample to individuals present in all 3 waves of the survey. Panel A pools the pledge and pledge plus competition arms, Panel B separates the sub-treatment arms. The dependent variable is Verified smoking abstinence, defined as self-reporting abstinence and obtaining a negative cotinine test. *Post* equals 1 for the follow-up period. *Treat* (*Pledge*, *Competition*) equals 1 if the school is randomized into treatment (pledge or pledge plus competition arm). All columns include pair fixed effects. Col. 1 pools the follow-up period. Cols. 2-4 replace *Post* with indicators for wave 2 (3 months post-intervention) and wave 3 (8 months post-intervention). Col. 3 adds baseline outcomes. Col. 4 adds: baseline ever smoked status, indicator for above median knowledge, indicator for having above-median friends who smoke, time inconsistency, risk seeking, household size, household land ownership, head of household in skilled occupation, any smoker in household, parents find smoking very unacceptable, asset count, and the number of male teachers who smoke in the school. Col. 5 interacts treatment with wave. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Appendix C Peer Effects

We explore the role of students' peer group on smoking, due to the nature of smoking as a social activity during adolescence and the large literature on peer effects in adolescent smoking (Footnote 16). The identification of peer effects is complicated by two well-known problems (Manski, 1993). First, peers may sort endogenously into groups or face unobserved group-level shocks. Second, a person's outcomes may causally affect the mean of his peer group, rather than the other way around (the reflection problem). Accounting for these identification challenges has been the focus of a growing literature examining peer effects in empirical settings (Bramoullé et al., 2009; Sacerdote, 2014; Advani and Malde, 2018).

We attempt to overcome these identification challenges using a novel identification strategy introduced by De Giorgi et al. (2010). The strategy leverages the fact that peer groups, formed at the individual level, often include a group of excluded, or non-overlapping, friends of friends. Under the assumption that students are randomized to classrooms and not influenced by their excluded peers, variation in the mean outcome of the excluded peer group is sufficient to overcome the reflection problem. Moreover, the mean outcome of the excluded peer group can be used as a valid instrument to address endogeneity due to possible correlated effects.

We perform two sets of instrumental variables (IV) analyses aimed at estimating peer effects. Our first IV analysis uses baseline social network information on friends and excluded friends of friends to estimate the effects of friends' abstinence at baseline on an index student's abstinence at follow-up. This analysis may be biased if sorting into friendship groups is endogenous. The first-stage F-statistic on the instrument is large (> 390 in Table C.1). Our second IV analysis exploits the quasi-random class assignment in 8th grade to provide plausibly exogenous variation in friendship networks as this generates variation in the number of 7th grade friends in the same 8th grade class. This refinement provides an alternative estimate of peer effects on abstinence in our setting. Again, the first-stage F-statistic on the instrument is large (> 140 in Table C.1). In further analyses, we analyze the role of friends and peers who smoke as sources of heterogeneous treatment effects. For all analyses, friendship networks are constructed based on a list of closest friends reported by each student at baseline. We identify the index student's friends and friends of friends using the list of 5 closest classmates at baseline by matching the index student's first and last names when he is listed as a friend by a classmate. [24]

We start with our first IV analyses of peer effects, assessing the role of friends' smoking

[24] In cases where the last name is not given on the friend list, we use the first name only if there are no other students with that same first name in the class. If there are multiple students with the same first name in the class, we code the identity as missing.

abstinence at baseline on an index student's smoking abstinence at follow-up. Our endogenous exposure is the number of friends who abstained at baseline. We instrument for this variable using the mean number of excluded friends of friends who abstained at baseline and estimate the difference-in-differences model with the full set of covariates, along with the size of the index person's friendship network (i.e., number of friends reported). We find that the program effect on the index student's smoking abstinence remains around 4 percentage points when we include the social network measures (Table C.1, columns 1-2, Panel A). Further, each additional friend who abstained at baseline is associated with a 2 to 6 percentage point increase in the probability of abstaining. The estimated effects are similar for the sub-treatment analysis (Panel B).

We next move to our more refined IV model of peer effects that exploit the quasi-random assignment of students to 8th grade classmates (Table C.1, columns 3-4, Panel A).[25] We estimate how abstinence varies with the number of abstaining friends who were assigned to the same 8th grade class. The program effect on smoking abstinence is about 4 percentage points, similar to our earlier estimated effect. In terms of peer effects, while noisy, abstinence in 8th grade increases by 3 to 5 percentage points for each additional 8th grade friend who abstained, similar to our earlier IV estimate. The estimated effects are similar for the sub-treatment analysis (Panel B).

To further explore the role of peers, we analyze the heterogeneity of treatment effects by peer characteristics (Table C.2). First, we interact treatment with the share of 7th grade 5 closest friends who abstained at baseline and stayed in the same 8th grade class. Second, using the 8-month followup, we exploit the quasi random 8th grade assignment (cross-section) and interact treatment with the share of the student's 8th grade peers who abstained 3 months after the intervention in 7th grade. We find no significant heterogeneity based on these measures. Third, we interact treatment with the student's 5 closest friends' predicted probability that he will smoke to explore how accurately peers can predict students' behavior. We find that students whose friends predicted to smoke are 0.3 percentage points more likely to abstain, suggesting the possibility that the intervention is effective for students who appear to be likely smokers to friends.

While some of these estimates are noisy, they support the importance of peer effects in tobacco use among adolescents in a lower income, high prevalence setting. We provide evidence that peer effects in this setting are similar to findings from the peer effects literature that focuses on adolescent tobacco use in high-income countries (Card and Giuliano, 2013; Fletcher, 2010; Fletcher and Ross, 2018).

In our models of peer effects, it is possible that the assignment of students to new

[25] This analysis uses cross-sectional data from the 8 month follow-up only.

classmates in 8th grade was not random. Indonesian schools assign unique number-letter combinations to denote each class within each grade, e.g., 7A to denote class A in the 7th grade. We compare each school's assignment to the probability of assigning a student to the same class in 7th and 8th grades (i.e., moving from class 7A to 8A) based on school size. For example, if a school has 3 7th grade classes, then the probability is one third. We then compare the difference between the school's probability of same class assignment and the probability based on school size. To check for balance, we regress this difference on the treatment indicator, the covariate, and the interaction term (Table C.3 presents the coefficients of the interaction terms). The 8th grade interaction terms are not significant, suggesting similar characteristics across treated and control schools, thus allowing us to use the 8th grade peers of students as an additional source of experimental variation.

Table C.1: IV estimates of peer effects on smoking abstinence

	(1)	(2)	(3)	(4)
	All friends		Friends in the same 8th grade class	
Panel A. Pooled estimates				
No. friends abstained at baseline	0.056***	0.024***	0.050**	0.027
	(0.012)	(0.008)	(0.020)	(0.017)
Treated	0.000	-0.003	0.039	0.038**
	(0.025)	(0.014)	(0.025)	(0.018)
Treated × Post	0.040*	0.045**		
	(0.022)	(0.022)		
Post	0.025**	0.026**		
	(0.012)	(0.012)		
No. friends	-0.032***	-0.014***	-0.031***	-0.019**
	(0.008)	(0.004)	(0.009)	(0.008)
Pair fixed effect	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	No	No
Baseline outcome	No	Yes	No	Yes
No. observations	7234	7234	2352	2352
R-squared	0.10	0.43	0.11	0.27
Dep. variable mean	0.781	0.781	0.766	0.766
F-statistic on instrument	409.5	398.7	148.9	144.5
Panel B. Sub-treatment estimates				
No. friends abstained at baseline	0.056***	0.025***	0.052**	0.029
	(0.012)	(0.008)	(0.021)	(0.018)
Pledge	-0.010	-0.014	0.016	0.013
	(0.037)	(0.019)	(0.039)	(0.029)
Pledge × Post	0.053**	0.054**		
	(0.022)	(0.022)		
Competition	0.011	0.007	0.059*	0.059***
	(0.031)	(0.020)	(0.031)	(0.022)
Competition × Post	0.027	0.036		
	(0.030)	(0.029)		
Wave=2	0.025**	0.026**		
	(0.012)	(0.012)		
No. friends	-0.032***	-0.014***	-0.032***	-0.020**
	(0.008)	(0.004)	(0.010)	(0.008)
Pair fixed effect	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	No	No
Baseline outcome	No	Yes	No	Yes
No. observations	7234	7234	2352	2352
R-squared	0.10	0.43	0.11	0.27
Dep. variable mean	0.781	0.781	0.766	0.766
F-statistic on instrument	411.6	400.0	154.9	149.4

Notes: This table shows the estimated effects of friends' abstinence on the index student's abstinence, using the number of excluded friends of friends who abstained at baseline as an IV for the number of friends who abstained at baseline. Columns 1 and 2 are difference-in-difference 2SLS estimates. Columns 3 and 4 are 2SLS estimates of abstinence in 8th grade (i.e., at 8 months) as a function of friends (and friends of friends) assigned to the same 8th grade class. Columns 1 and 3 include pair fixed effects. Columns 2 and 4 include pair fixed effects and student baseline outcomes. Standard errors are clustered at the school level. Robust standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table C.2: Program effects on smoking abstinence by peers' smoking status

	(1)	(2)	(3)
	Peer characteristic		
	Share of 5 friends who abstained and in same 8th grade class	Share of 7th grade classmates who abstained and in same 8th grade class	Mean belief of 5 friends at baseline
Panel A. Pooled estimates			
Treat	-0.006	0.037	0.076***
	(0.025)	(0.082)	(0.027)
Treat × Post	0.046*		-0.014
	(0.025)		(0.028)
Treat × Characteristic	0.042***	0.006	-0.042***
	(0.014)	(0.108)	(0.012)
Treat × Post × Characteristic	-0.013		0.034***
	(0.012)		(0.012)
Characteristic		0.176**	
		(0.074)	
Wave = 2	-0.010		0.001
	(0.018)		(0.020)
Wave = 3	-0.036**		-0.034*
	(0.017)		(0.019)
No. observations	7211	1985	5319
R-squared	0.16	0.29	0.17
Dep. variable mean for control group	0.781	0.766	0.781
Panel B. Sub-treatment estimates			
Pledge	-0.012	0.015	0.082**
	(0.037)	(0.080)	(0.035)
Competition	-0.001	0.073	0.071*
	(0.033)	(0.128)	(0.037)
Pledge × Post	0.059**		-0.000
	(0.029)		(0.032)
Competition × Post	0.035		-0.026
	(0.033)		(0.038)
Pledge × Characteristic	0.037**	0.027	-0.044**
	(0.018)	(0.107)	(0.018)
Competition × Characteristic	0.049**	-0.036	-0.039**
	(0.024)	(0.164)	(0.017)
Pledge × Post × Characteristic	-0.015		0.032*
	(0.017)		(0.017)
Competition × Post ×Characteristic	-0.012		0.037**
	(0.016)		(0.016)
Characteristic		0.181**	
		(0.077)	
Wave=2	-0.010		0.001
	(0.018)		(0.020)
Wave=3	-0.036**		-0.034*
	(0.017)		(0.019)
No. observations	7211	1985	5319
R-squared	0.16	0.29	0.17
Dep. var. mean for control group	0.781	0.776	0.781

Notes: The column headings represent the peer characteristics (*Characteristic*) interacted in each difference-in-differences model. Fully adjusted for baseline covariates. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table C.3: Balance in 8th grade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					Adjusted difference		
	Control	Treated (Pooled)	Individual pledge	School competition	Pooled	Individual pledge	School Competition
Panel A. School characteristics							
Distance to Jogjakarta	17.391	17.595	16.004	19.027	0.003	-0.011	0.018
	(10.671)	(9.543)	(9.120)	(9.693)	(0.009)	(0.011)	(0.010)
Student-to-teacher ratio	14.771	14.312	3.388	2.525	-0.029	-0.021	-0.045
	(3.122)	(2.986)	3.122	3.122	(0.021)	(0.024)	(0.029)
Average mathematics score	56.932	52.516	14.530	12.031	0.001	0.005	-0.005
	(14.536)	(13.273)	14.536	14.536	(0.006)	(0.007)	(0.008)
No. teachers who smoke	1.429	2.037	1.093	1.946	-0.012	-0.047	0.008
	(1.521)	(1.626)	1.521	1.521	(0.056)	(0.091)	(0.060)
Number of classes	12.481	12.513	5.007	3.950	-0.022	-0.027	-0.020
	(4.321)	(4.561)	4.321	4.321	(0.016)	(0.019)	(0.021)
No. of observations					72	72	72
Panel B. Student characteristics							
Truancy	0.039	0.059	0.054	0.064	-0.068	-0.074	-0.078
	(0.193)	(0.236)	(0.225)	(0.245)	(0.087)	(0.082)	(0.110)
Ever smoked	0.730	0.737	0.739	0.735	0.006	-0.022	0.031
	(0.444)	(0.441)	(0.439)	(0.442)	(0.063)	(0.073)	(0.076)
Tobacco knowledge	5.619	5.758	5.646	5.859	0.030	0.097	-0.032
	(2.217)	(2.119)	(2.041)	(2.183)	(0.041)	(0.051)	(0.042)
No. 5 best friends smoke	1.538	1.775	1.716	1.829	-0.054	-0.005	-0.028
	(1.704)	(1.761)	(1.756)	(1.765)	(0.065)	(0.080)	(0.073)
Time-inconsistent	0.704	0.657	0.645	0.668	-0.009	-0.084	0.026
	(0.457)	(0.475)	(0.479)	(0.471)	(0.064)	(0.069)	(0.071)
Risk seeking	0.342	0.441	0.390	0.486	-0.022	-0.037	0.011
	(0.474)	(0.497)	(0.488)	(0.500)	(0.060)	(0.057)	(0.087)
Parents find smoking unacceptable	0.251	0.242	0.238	0.245	-0.106	0.018	0.000
	(0.434)	(0.428)	(0.426)	(0.431)	(0.095)	(0.050)	(0.064)
Panel C. Household characteristics							
Household size	4.755	4.703	4.499	4.888	0.011	0.009	0.023
	(1.656)	(1.554)	(1.258)	(1.760)	(0.013)	(0.012)	(0.014)
Land ownership	0.551	0.508	0.470	0.543	0.057	-0.010	0.099
	(0.498)	(0.500)	(0.499)	(0.499)	(0.049)	(0.050)	(0.058)
Parent in skilled occupation	0.348	0.299	0.316	0.284	0.036	0.108	-0.072
	(0.477)	(0.458)	(0.465)	(0.451)	(0.055)	(0.051)	(0.067)
Parent smoked	0.442	0.476	0.468	0.483	-0.003	0.025	-0.014
	(0.497)	(0.500)	(0.499)	(0.500)	(0.032)	(0.036)	(0.037)
No. smokers in household	0.726	0.736	0.708	0.761	-0.030	-0.069	-0.032
	(0.995)	(0.782)	(0.746)	(0.812)	(0.040)	(0.045)	(0.051)
Parents find smoking unacceptable	0.441	0.438	0.453	0.424	0.001	0.024	-0.015
	(0.497)	(0.496)	(0.498)	(0.495)	(0.060)	(0.066)	(0.066)
Asset count	6.454	6.515	6.399	6.620	0.023	0.011	0.026
	(1.264)	(1.344)	(1.326)	(1.353)	(0.015)	(0.020)	(0.017)

Notes: Each cell is the coefficient is the interaction term between the treatment variable and the characteristic of interest. The dependent variable is the probability that students are assigned to 8th grade randomly. District fixed effects included. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.