Non-Monetary Incentives for Tobacco Prevention in Indonesia*

Margaret Triyana† Justin S. White‡

September 25, 2020

Abstract

We provide evidence on the effectiveness of a school-based program that uses non-monetary penalty and regular monitoring to prevent risky behavior among adolescents in Indonesia. The school-based field experiment invited students and their parents to sign a pledge for students to abstain from tobacco use and 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 pledges increase biochemically verified tobacco abstinence by 4 percentage points. This effect is sustained 3 months after the program ends. School competition has no additional impact on tobacco abstinence. Our findings highlight the effectiveness of non-monetary incentives to curb risky behaviors among adolescents with self-control problems.

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 thank Alice Chen, Taryn Dinkleman, Willa Friedman, Felipe Gonzalez, Lakshmi Iyer, Julien Labonne, Jean Lee, Maria Rosales, Jessica Pan, 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). An encouraging trend for public health policy makers is that the youth smoking prevalence has been declining worldwide in recent decades: from 12.0% in 1980 to 6.9% in 2015.[1] Yet, the decline in youth smoking has occurred largely in high-income countries, and nearly 80% of the 41 million smokers aged 15-19 worldwide live in LMICs. Youth tobacco use has recently seen a resurgence in many countries with the rise in popularity of nicotine-containing electronic cigarettes and vaping products among youth. The effects of growing e-cigarette use among youth on long-term health and combustible cigarette use continue to be debated, but there is widespread concern in the public health community that the downward trend in youth tobacco use is slowing and reversing itself.

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. 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 the highest male smoking prevalence globally (almost 70%), has about 15% of deaths attributable to tobacco use, and is the only one of the 10 countries with the largest smoking population to experience an increase in youth smoking from 1990-2015 (World Health Organization, 2012a; 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

[1] Calculations in this paragraph rely on estimates of smoking prevalence for individuals aged 15-19, 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.

investments have been used extensively for a variety of health behaviors in LMICs (Gopalan et al., 2014). Yet, incentive-based programs in LMICs have relied almost exclusively on monetary rewards and have rarely targeted adolescents, at a stage when life-long habits often form. In the broader tobacco prevention literature, only three randomized trials have tested incentives for tobacco prevention among adolescents, all in Europe and none judged to have low risk of bias (Hefler et al., 2017).

We implemented a school-based field experiment of 2,700 students from 72 middle schools in Indonesia to test a tobacco prevention program offering 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 demerit points 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 self-control problems.[2] 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 the 5-month intervention, 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. In a random half of the 36 treated schools, we also introduced school-level competition to test how effectively team incentives that emphasize group identity ('school spirit') mobilize peer support and pressure to improve health behavior. These 18 schools received 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 into the next school year. Using an intent-to-treat approach, we find that biochemically verified tobacco abstinence increased by approximately 4 percentage points. The effects of the individual pledge and school competition remained similar in magnitude until 8 months (3 months after biweekly monitoring had ceased). The school competition intervention tests whether social effects, including emphasizing group identity, change behavior more effectively

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

than the individual pledge alone. Our analysis reveals that the individual pledge and school competition interventions produced similar intent-to-treat effects. Thus, we fail to find evidence in support of this hypothesis since the individual pledge, rather than the school competition, drives the program effect.

We further explore the role of potential mechanisms beyond group identity and regular monitoring in explaining our pattern of findings. Potential mechanisms include aspects of peer influence such as fear of social sanctions, social learning, and changing smoking norms, and incentive effects such as increased motivation and effort, overcoming time-inconsistency, and increased salience. We find some evidence consistent with changing norms in school as proxied by a reduction in tobacco-related discussion with friends among treated students after the intervention period had ended. Similarly, treated students were less likely than control students to discuss tobacco use with their parents. We also find some evidence consistent with incentive effects. Specifically, students believe that the reporting and demerit point system would lower tobacco use, suggesting fear of sanctions as one channel. 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.

Our study contributes to a growing literature in behavioral and health economics on the use of incentives to improve health behaviors. Incentive programs have been shown to be effective in changing a number of health behaviors, according to several recent reviews (Mantzari et al., 2015; Giles et al., 2014; Gopalan et al., 2014). We 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. The adolescent period coincides with the development of executive functions that regulate impulse control, long-term planning, and the ability to resist peer influence (Steinberg, 2007; Zelazo and Carlson, 2012). Psychologists attribute increased risk-taking during adolescence to a heightened sensation seeking (Steinberg et al., 2018), and the same reward processes in the brain that make adolescents susceptible to risky behavior such as smoking may also make them responsive to incentive-based interventions. Adolescents also tend to have less disposable income than adults, and thus would be predicted to be more price-elastic than adults, holding all else equal. A better understanding of the impacts of incentives on adolescent risk-taking may improve policy design and implementation, with an opportunity to intervene at a critical stage in the formation of habits and preferences.

Second, our program employs a non-monetary incentive, whereas demand-side health incentives have largely used monetary rewards. Similarly, health-related commitment contracts have typically been structured to put an individual's own money (Bryan et al., 2010; Giné et al., 2010; Halpern et al., 2015; Royer et al., 2015; Bai et al., Forthcoming;

Schilbach, 2019).[3] In our study, the non-monetary penalty (demerit points and parental notification) lowers the financial cost of the intervention and improves the chances that the intervention is sustainable, culturally acceptable, and scalable.[4]

Third, health-related incentive programs have tended to be individual in nature. Social incentive schemes may be able to mobilize peer influence as a form of social commitment and social support. A small 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). We also extend the degree of social accountability by engaging key actors in the student's social network (parents, teachers, and fellow students) to monitor and enforce the individual pledge to avoid tobacco use. Risky behavior in adolescence is heavily influenced by adolescents' social networks, most notably their parents and youth peers. Research in psychology shows that children and adolescents learn and model the behaviors they observe their parents and peers engaging in (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 in adolescence, 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. Interventions in high-income countries may not be generalizable 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.

[3] 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.

[4] During preliminary discussions, local government officials felt that a monetary intervention was not acceptable.

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

Indonesia has the highest male smoking prevalence in the world (almost 70%) and 20% of deaths among men and 12% of deaths among women (mostly from secondhand smoke) are attributable to tobacco use (World Health Organization, 2012b). In spite of the high burden of disease, Indonesian smokers have shown little interest in smoking cessation (Nichter et al., 2009, 2010). In addition, price regulation has not been a politically feasible policy lever in Indonesia, a major tobacco-producing country. Due to these policy constraints, policymakers in Indonesia have a growing interest in non-price strategies, including prevention programs, to reduce reduce tobacco use and the tobacco-attributable disease burden.

The government's reliance on the tobacco industry as a source of revenue and employment has contributed to Indonesia's tobacco problem. The tobacco industry accounts for about 10% of total tax revenue in Indonesia and is the second largest employer in the country (Barber et al., 2008).[5] As such, the industry wields considerable political and financial clout that has been used to thwart tobacco taxation. Non-price regulations of tobacco use have also been limited in Indonesia, as evidenced by Indonesia's 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 the smoking rate is particularly high among adults in Indonesia, early smoking initiation is still understood to be a risky behavior, even among youth in Indonesia (Ng et al., 2007a). Adoption of addictive habits such as tobacco use in adolescence can have lasting consequences (Rabin, 2013). Research shows that nicotine dependence often begins during adolescence (DiFranza et al., 2007), with an age of initiation typically ranging between 15 and 19 (Thomas and Perera, 2006; World Health Organization, 2012a; Lillard et al., 2013). According to recent national data, adolescents in Indonesia report first smoking between the

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

ages of 12 and 13 (World Health Organization, 2015). These data also show that 34% of all middle-school students in Indonesia are current smokers, and 56% have tried smoking. In comparison, only 7% of middle-schoolers in the US report having used any tobacco products (Jamal et al., 2017). These numbers highlight adolescence as a critical period for tobacco prevention in Indonesia.

The wide-ranging health risks associated with tobacco use tend to be downplayed in Indonesia. Among the Indonesian population, 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). This attitude is also prevalent among health professionals in Indonesia; the fact that more than 80% of physicians believe that smoking fewer than 10 cigarettes per day is not particularly harmful for health reinforces this mentality (Ng et al., 2007b). Because smoking is deeply ingrained in Indonesian culture and national political opposition to tobacco use hardly exists, sub-national policy may offer a more 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 new 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.

Our study was implemented in 19 subdistricts within two districts in Yogyakarta Province: Kulon Progo and Sleman. Kulon Progo has adopted some of the most restrictive tobacco control policies in Indonesia, while Sleman has been less proactive due to its status as a tobacco producing area.[6] In 2012, Sleman implemented a local executive order that restricts smoking in some public areas, while in 2014, Kulon Progo implemented a stricter district-wide regulation that bans smoking at schools and in other public areas and bans tobacco advertising and event sponsorship by tobacco firms. At the start of this 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 the local governments to ensure its acceptability and sustainability.

Female smoking prevalence is low, about 2%, in Indonesia (World Health Organization, 2017). Therefore, our study targeted 12- to 13-year-old male students in the 7th grade (*Sekolah Menegah Pertama*). In 2013, 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.[7] In our

[6] We excluded the primary tobacco growing areas in Sleman from our sample. We selected subdistricts in Sleman that are geographically proximate and similar to Kulon Progo in their socio-economic characteristics. Details on sample selection are available in Appendix A.

[7] This is according to a nationally representative health survey (Riset Kesehatan Dasar, 2013).

study, most male students report smoking for the first time between ages 11 and 13, indicating that the 7th grade is an apt time for intervention. Most students who smoke choose clove cigarettes (*kretek*), which are the most commonly found form of tobacco in Indonesia, followed by white cigarettes and e-cigarettes.

Schools in our sample have an average of four classes per grade, and each class has a maximum of 42 students. Many Indonesian schools, including many schools in our sample, employ a disciplinary system wherein teachers issue demerit points to students who misbehave. By embedding our penalty within the existing system, we ensure that students understood the nature of the penalty.

2.2 Interventions and Procedures

Figure 1 summarizes the study design. The study involved a total of 72 schools: 36 schools in Kulon Progo and 36 schools in Sleman. 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 school competition arm.[8] 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 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).[9]

A total of 3,031 students were screened. We excluded 248 students who did not return the informed consent form. Thus, we randomized 2,783 students. We allocated 1,346 students to the control arm, 683 students to the individual pledge arm, and 754 students to the school competition arm. All students who consented signed the pledge.

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 for cotinine, a metabolite for nicotine, using a urine assay to biochemically verify tobacco abstinence at baseline. We then collected biweekly reports from all treated schools between January and May 2017. We conducted an unannounced audit in all 72 schools in March 2017, 3 months after the intervention started. We completed biochemical verification of about 80% of students during the 3-month audit. 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.

[8]We initially identified 78 schools, but 6 were excluded due to the remoteness of the area and school size.

[9]Further details on the randomization procedure and power analysis are available in Appendix A.

All participating schools received an information campaign on the risks of smoking, comprised of one 45-minute session per 7th grade class 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 system, which often included a policy that banned smoking on school property.

In the 36 treated schools, the information session ended with an invitation for students to sign a commitment pledge to abstain from tobacco use during the intervention period. After students signed their individual pledge, they were asked to sign an additional 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, but similar, pledge stating that parents would monitor their children to prevent them from using tobacco products.[10] Students bwho violated the terms of the pledge, as measured by a biochemically verified cotinine test or student and teacher reports, were penalized using a school-based demerit system.[11] Upon each of the first two infractions, the student was summoned to the guidance counselor's office so the counselor could issue a warning, reminding the students of their pledge and urging them not to smoke in the future. Upon the third and subsequent infractions, the student was summoned to the principal's office, and the principal 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 student behavior.

Guidance counselors, members of the student council, and classmates 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 in the school. In the study area, student councils are comprised of older students (in 8th and 9th grade), who are widely viewed with respect by other students. Due to their tendency to be academically minded, this group tends to have low smoking rates. We trained members of the student council on how to promote smoking abstinence among participating students.[12] The student council, supervised by the

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

[11] The demerit system for our study was kept separate from the school's pre-existing demerit system, in order to standardize the program. Prior to our intervention, tobacco use typically resulted in a warning from the guidance counselor, and in the schools with pre-existing demerit point systems, the demerit points were generally lower than other disciplinary infractions. All students in 7th grade were eligible to be reported within our demerit system regardless of whether they signed the pledge. However, only those who signed the pledge received demerit points upon failing to comply.

[12] 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 did not always have time to handle the added burden of monitoring and reporting students' tobacco use for the study.

guidance counselor, received, investigated, and recorded reports of students who smoked. These reports were ultimately kept by the guidance counselor to be evaluated, and if necessary, followed up using the demerit system. The reports were kept private (except as part of an appeals process[13]).

Every two weeks, with the assistance of the student council members, 7th grade students in the treated schools were asked to complete a form indicating any students who they observed smoking on or off school premises. These biweekly reports helped create a sense of accountability and salience to the intervention. The reporting mechanism, in which all students provided reports at the same time, was designed 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. Students could also report cases directly to the student council and teachers in school or our program administrators via email, text message, or phone call. The audit conducted 3 months post-intervention served to validate the biweekly reports in treated schools. 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. In larger schools, we randomly selected 48 students to test.[14]

School performance in the school competition arm was scored based on several data sources: the biweekly reports, the results from biochemically verified tests of smoking status during an unannounced study visit (i.e., the audit) to each school at 3 months, and the results from biochemically verified tests at 8 months.[15] The school's score included several components: timeliness of reporting, completeness of reporting, and the proportion of students who were smoke-free.[16] We provided a list of the top 10 ranked schools to

[13]The reporter's identity may be revealed to the guidance counselor as part of an appeals process. An accused student could appeal to the guidance counselor, who would investigate. No appeals were requested during the study period.

[14]Selecting all participants in smaller schools and randomly selecting 48 participants in larger schools were done to obtain an expected sample average of 30 participants per school for the unannounced audit—a target determined by the available budget and staffing capacity.

[15]One potential concern is schools in the competition arm would 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, when biweekly reporting had ended. Figure B.4 shows the average number of reported students in each sub-treatment. Differences in the number of reported students between the individual pledge and school competition arms are not statistically significant.

[16]We used the following formula to calculate a biweekly score for each school:

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

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

student council members on an ongoing basis. The schools that finished in the top 10 received a certificate of recognition from the district's Department of Education.

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 specific classes for the entire academic year. Students remain in the same classroom throughout the day, while subject teachers rotate through. In schools with more than one class per grade, administrators assign students in each cohort to different classes in 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 rise to the 8th grade. Biweekly reporting did not resume at the start of the new academic year because the student council had not yet been formed 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.

We had two key study hypotheses. First, combining social pressure and non-monetary incentives that act as partial commitment contracts 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 behavior, including monitoring from parents and teachers, fear of the in-school remediation for tobacco use, and enhanced motivation and self-efficacy not to smoke. Second, school competition combined with a pledge would reduce tobacco use among students, compared to a pledge alone or receiving information alone. The competition emphasized a spirit of teamwork, and we expected this to increase the salience of the group members' identity as non-smokers by creating peer pressure from classmates not to smoke and engaging students to establish a norm against smoking.

3 Data and Estimation Strategy

3.1 Data Sources

Our study drew on four main data sources: administrative data on schools, school records, participant surveys, and biochemical test results. We conducted self-administered surveys of participants in the 72 schools, including participating male students in 7th grade and their parents, each school's guidance counselor and principal, and members of the student council. Figure 2 describes the timing of data collection. The baseline survey was conducted

council and guidance counselor that we would check the reports during random audits, and there would be a penalty of 50 points if reports were falsified.

in January 2017, an unannounced audit was conducted in March, and a follow-up survey was conducted in August, at the beginning of the new academic year.

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 also highly sensitive at 100 ng/ml. 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 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, the student council, teachers, and classmates 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's self-reported smoking abstinence is verified with the cotinine test. If a student reported abstinence, but failed the cotinine test, we code the student as failing to abstain from smoking. We drop observations with missing outcome data.

The share of students who had a self-report but missing cotinine test results is similar across treated and control participants at baseline and at 8 months.[17] Similarly, the rate of students who self-reported smoking abstinence but failed the cotinine test is similar across treatment and control at each wave of the survey.[18] We test the robustness of our main results by conducting sensitivity analyses by assuming students with missing outcome data were non-smokers unless they were observed or reported to have smoked at some point during the trial (Section 4.5).

Additional outcomes to explore potential mechanisms include tobacco knowledge, the

[17] All students present in the school at 3 months who self-reported their smoking status also took the cotinine test. The difference in the share of missing cotinine test is 0.004 (p-value 0.89) at baseline and 0.02 (p-value 0.47) at 8 months.

[18] The difference across treatment and control is 0.01 (p-value 0.522) at baseline, -0.003 (p-value 0.796) at 3 months, and 0.008 (p-value 0.60) at 8 months.

frequency of being reported as having smoked within the prior 7 days, smoking intentions, and discussion with parents and peers. Tobacco knowledge is based on a score on a series of 9 questions to assess students' awareness of the health risks of tobacco use.[19] The frequency of being reported as having smoked is based on the 8-month survey. A student's smoking intentions are measured by the probability he expects to be smoking at all in 3 months, reported on a scale from 0 to 10. Students were asked whether they had discussions on smoking with their parents or peers in the 8-month survey. We also included a set of statements on program characteristics to explore mechanisms in treated schools, such as "The program made me want to abstain from tobacco use" and "The program made my friends want to abstain from tobacco use." In the survey of student council members, we included statements on the perceived effectiveness of demerit points and reporting in preventing smoking in school.

Covariates The administrative school data include several school-level characteristics, such as the district, number of 7th grade classes in the school, and distance from the city of Yogyakarta. The baseline school survey includes characteristics, such as the number of teachers who smoke in the school, the timing of the last parent-teacher conference, the presence of disciplinary action against smoking, and the average national examination scores in mathematics as a proxy for school quality. We also include questions on the size of the student council since 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).[20]

The student survey includes information on the student's mathematics score on the national examination, self-reported closeness to parents, and family's ownership of assets.[21] We also ask students to predict the likelihood of each of their five closest classmates smoking within 3 months, reported on a scale from 0 to 10. We use these reported friendship networks to estimate peer effects in Appendix B.1. The student survey also includes time and risk preferences, where students were given hypothetical waiting times for monetary gifts and hypothetical gambles to elicit their time-inconsistency and

[19]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.

[20]Questions include the timing of the smoker's first cigarette in a day, difficulty refraining from smoking in smoke-free areas, daily consumption, difficulty giving up the first morning cigarette, likelihood to smoke more in the morning, and smoking while sick in bed.

[21]Assets include: radio, television, cable TV, refrigerator, bicycle, motorcycle, car or motor boat, mobile phone, poultry, and livestock (cows, water buffaloes, horses, or goats).

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 past 7 days prior to the survey, as a related disciplinary outcome and risky behavior.[22] As an additional covariate, we also use students' verified tobacco abstinence 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 head of the household 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 capture parents' smoking attitudes, we also 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; and Panel C describes household characteristics. We also calculate adjusted differences between control and treated groups, including district fixed effects (Columns 5 to 7). We compare the pooled treatment to the control group (Column 5) and each sub-treatment to the control group (Columns 6 and 7). We control for baseline characteristics in the main analyses to take these differences into account.

Schools in our sample are similar across arms (Panel A). The schools are, on average, about 17 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%. Schools in the control group have about 1.4 teachers who smoke, while those in the treatment group have about 2; this difference is small although statistically significant. 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 arms are not jointly significant.

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

Student characteristics are also well balanced across the control and treated groups (Panel B). Most students, about 78%, abstained in the 7 days prior to the baseline survey. This implies that at least 22% of students had smoked by the time they reached 7th grade. In addition, more than 70% of students had ever smoked by 7th grade. Students' tobacco knowledge ir fairly high, they correctly answered about 6 of the 9 questions on tobacco at baseline. Students reported tobacco use among their peers and about 1.5 of students' 5 closest friends smoked at baseline, with a median of 1.

In terms of other risky behavior, about 4 to 6% of students had been truant (missed class) during the prior week. About 65 to 70% of students exhibited time-inconsistency and about 34 to 44% reported risk-seeking preferences. About 25% of students believe 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. The baseline 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. About half of households in the sample own land, one proxy for wealth. About one-third of households have their head of household in skilled occupations. The average number of smokers in the household is 0.7, and more than half of students have at least one household member who smoked, most likely the father, because smoking among women in Indonesia is rare. About 40% of parents would find their child smoking to be unacceptable or very unacceptable, in contrast to students' belief of about 25% (in Panel B). Households have an average of 6.5 assets (out of 10), and around 80% of households have above median assets. Differences across arms by household characteristic are significant only for one comparison—household size in the pledge versus control arms—although the magnitude of the difference is modest.

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 verified 7-day smoking abstinence, combining self-reports with the cotinine test results for individual 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:

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

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 treatment group. We include pair and wave fixed effects and all standard errors are clustered at the school level to account for within-cluster correlation of outcomes. Similarly, we conduct the following sub-treatment analysis using the following equation:

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

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 class smoking proportion at baseline, and individual smoking status at baseline. Fully adjusted models add a vector of student, household, and school characteristics. We include the following individual baseline characteristics: truancy, ever smoked status, tobacco knowledge (indicator for being above the median of 6), above-median number of friends who smoke (indicator for >1), an indicator if their parents would find the student smoking unacceptable or 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.

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. Next, we estimate a cost-benefit analysis of the program. Finally, we discuss various robustness checks.

4.1 Effects on Tobacco Abstinence

We examine the effect of the program on verified tobacco abstinence in Table 2. Panel A presents the pooled effect and Panel B compares the effects by sub-treatment. The pooled treatment resulted in a 3.7 to 4.4 percentage point increase in abstinence, equivalent to a 5.6% increase from the baseline mean. The estimated effect remains largely unchanged when we include individual and class baseline outcomes and individual characteristics (Column 3). We then compare the effect on abstinence of the individual pledge and the pledge plus school competition (Panel B). We find that the individual pledge drives the program effect, with a 5.1-point increase in tobacco abstinence in the fully adjusted model (Column 3), compared to an estimated effect of the pledge plus school competition of 3.7 points and not significantly different from 0. We also fail to reject the equality of coefficients (with p-values ranging from 0.4 to 0.6). These results suggest that including school competition has no additional effect beyond the individual pledge on smoking abstinence.

This result may be attributed to several factors. First, in spite of the existing literature that adolescent smoking is susceptible to peer influence, smoking is ultimately carried out by the individual student, and group-based prevention may be less effective in this setting. Indeed, in treated schools, 88% of students (strongly) agreed with the statement that the program made them want to abstain from tobacco use. Conversely, only 68% of students (strongly) agreed with the statement that the program made their friends want to abstain from tobacco use, suggesting that the effect is driven by the individual pledge. Second, the token reward of an acknowledgment from the Department of Education in the school competition arm may not sufficiently motivate individual students to change their behavior and shift smoking norms.[23] We also examine the program effects using cross-sectional data at 3 and 8 months post-implementation (Table B.3). We find that the program led to a 3- to 6-percentage point increase in verified tobacco abstinence at 3 months, and this effect remains at 8 months (Panel A). Similar to the results under difference in differences, we fail to reject the equality of coefficients of the estimated effects of the sub-treatment arms (Panel B). Thus, the program effect using the cross-sectional data is similar to the estimated effect under the difference-in-differences framework. Additionally, the persistence of the effect to 8 months, at which point the class monitoring had ceased, suggests that the program effect is not driven solely by monitoring.

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 using students' social networks, individual pledges, and monitoring to promote

[23] Our results may be related to recent evidence that shows that penalties can be more effective than bonuses for minimizing undesirable behavior (Homonoff, 2018).

tobacco abstinence is feasible. Additionally, our estimates of program effects are similar in magnitude to results from higher income settings, where smoking prevalence is lower (Thomas and Perera, 2006).

4.2 Potential Mechanisms

In Section 4.1, we showed that the commitment pledge increased smoking abstinence while adding school competition did not lead to further improvement. There are several possible underlying mechanisms that might explain this pattern of findings. While we are limited in our ability to distinguish among these mechanisms due to data availability and the experimental design, we describe the suggestive evidence for certain mechanisms in this section. Broadly, we classify the intervention-driven mechanisms under the categories of peer influence and incentive effects. Peer influence may encompass harnessing group identity, fear of social sanctions, social learning, and changing social norms.[24] Incentive effects may encompass fear of parental sanctions, increased motivation and effort, overcoming time-inconsistency, increased salience, and habit formation. The evidence on potential mechanisms is summarized in Table 3.

We begin by considering peer influence mechanisms. First, we can rule out group identity and fear of social sanctions as primary explanations because both would be differentially activated under the commitment pledge plus school competition subtreatment. While research finds these mechanisms to be potent determinants of behavior in other settings, the absence of an additional effect from the commitment pledge plus school competition subtreatment on smoking abstinence means that they are not primary channels here. Indeed, the difference in the fraction of students agreeing with the statement that the program made their friends want to quit is not significant across the subtreatment arms (Table B.4).

Second, students may learn 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 students' tobacco knowledge after the trial (Table 3, Column 1), although this may be due in part to the relatively high baseline knowledge. The fraction of students agreeing to the statement that better knowledge would make them abstain from tobacco use is similar across treatment and control.

Third, the intervention may have changed perceived norms regarding the acceptability of smoking. To determine if smoking norms changed, we assess whether the intervention changed the frequency with which the students discussed smoking with friends. Whereas

[24] We use the 8th grade class assignment to estimate peer effects in Appendix B.1.

during the trial a reduction in the perceived acceptability of smoking could be associated with more discussions (e.g., discussing harms of smoking or the intervention itself) or fewer discussions (e.g., fewer plans to meet after school for a smoke), we postulate that after the trial has ended students would have fewer discussions about smoking if it had become less acceptable. This is, in fact, what we find (Column 2). Compared with control students, treated students are less likely to discuss smoking with their friends at the 8-month follow-up. Perceived smoking norms at home also may have shifted as our intervention asked parents to help their children abstain from tobacco use. First, we analyze whether students discussed tobacco use with their parents and find that treated students are less likely to discuss smoking with their parents (Column 3), which is consistent with the changing norms in school. Second, we examine whether students changed their belief on their parents' approval on smoking. About 25% of students at baseline were not sure whether their parents would find it acceptable if the student smoked. We note that a clearer understanding of parental disapproval could lead students to "fall in line" and smoke less or to rebel and smoke more. However, we find no change in the effect of the intervention on beliefs about parental approval, with a tight confidence interval around 0 (Column 4).[25]

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 the commitment pledge increased students' fear of sanctions. Based on our survey of the members of the student government, we find that they believed having a demerit point system and reporting in school through teachers and student government would make students more likely to abstain from tobacco use (Table B.5). Similarly, students in treated schools are 30% more likely to report their peers who smoke, and the difference across the subtreatment arms is not significant. These results suggest that the incentive effect through reporting is a potential mechanism behind our main findings. Second, the program may have increased students' motivation and effort to stay smoke-free. We explore individual motivation to abstain from tobacco use by analyzing changes in students' own smoking intentions (Table 3, Column 5). We find no significant change in the smoking intentions of students in the treated group compared with students in the control group. Third and relatedly, the non-monetary incentives may have helped students with self-control problems to overcome their time-inconsistency. As we describe in Section 4.3, we find evidence that 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

[25] We also find no significant change in the fraction of students who are unsure whether their parents would be accepting if the student smoked.

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 acted by increasing the salience of not smoking, notably through the regular monitoring of smoking behavior. A salience effect would be expected to fade after the program ended. Instead, we find that the program effects persist beyond the school year. Finally, the short-term incentives may have helped to establish healthy habits not to smoke. We cannot directly test for habit formation with our data, although we would not expect that it would differ across sub-treatments.

We conclude that there is some evidence in favor of fear of sanctions and changing norms as mechanisms that might account for the program effects. More work is needed to disentangle 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.[26] 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 begin by analyzing the heterogeneity of treatment effects by students' peer characteristics (Table B.11). 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 follow-up, 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 the treatment indicator with the mean predicted probability that the student will smoke as reported by the student's 5 closest friends, to explore how accurately peers can predict students' behavior. We find that students whose peers predict are 10% more likely to smoke are 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. We further explore peer effects in Appendix B.1. We find similar program effects and that students with tobacco-free peers are more likely to abstain.

[26] Appendix Tables B.6, B.7, and B.8 present the standard errors.

We also 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 person'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. Since 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. The program has larger effects on time-inconsistent students, but otherwise we find no significant heterogeneous treatment effects by student characteristics. Similarly, when we separately consider the sub-treatments, we find no significant heterogenities by student characteristics.

By household characteristics Students with different family backgrounds may respond differently to the intervention. To explore that possibility, we consider the following baseline household characteristics: household size, land ownership, the head of household in skilled labor, parent is a smoker, any smokers in the household, parents find child smoking to be very unacceptable, and having above-median household assets. 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.[27]

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 the main city of Yogyakarta, student-to-teacher ratio, above-median number of student council members, national examination math score, whether any teachers smoke, and number of classes. We find larger treatment effects among those in higher performing math 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 the tobacco prevention program is that it is relatively low cost. 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.,

[27] 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.

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 (we assume this would 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 percent reduction 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.[28] Based on the expected 12-month effect, the cost per smoker averted would be \$220, substantially lower than the cost for 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.

We then perform back-of-the-envelope calculations of the benefit-to-cost difference and ratio of our program (Appendix Table B.9). We assume the program would delay initiation by 1 to 3 years. We further assume that the benefits would accrue from productivity gains from cardiovascular disease averted and increased labor market participation for 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 cardiovascular disease .[29] To estimate a lower bound, we exclude externalities from secondhand smoke 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.

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

[29] 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 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 second-hand smoke are excluded. The cost of cardiovascular disease is about USD 2500 in Indonesia. The estimated NCD cost is USD 17,800, and CVD accounts for about a 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). Chronic obstructive pulmonary disease (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, cardiovascular disease would occur 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 be between the ages of 39 to 41 (COPD Foundation).

4.5 Robustness

Experimental Validity There are two primary threats to the empirical design. First, the randomization may produce imbalanced groups either by chance or if the randomization 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. 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. We test for balance along pre-intervention characteristics using a t-test of equality across arms and in bivariate comparisons and a regression of each covariate on the dichotomous treatment variable, adjusted for district fixed effects (Table A.2). The t-tests show that re-intervention 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. The joint test across all covariates fails to reject the null that the two groups are different. Similarly, differences across schools in the sub-treatment arms and the control arm are small and generally not statistically significant. The joint tests of balance also fail to reject the null that either the pooled treatment group or the sub-treatment groups are different from the control group. To take into account these baseline differences, we control for them in our analyses.

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 their self-report and refusing the biochemical verification. We analyze whether students refuse the urine test differentially 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 treatment 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 1 if a student reports tobacco abstinence but the biochemical verification shows otherwise. We find no significant effect across the treatment and control groups on this measure. We also explore alternative outcomes to take into account this potential selection concern below.

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 and pledge plus school competition arms (Column 3). We find no significant reduction in the number of reports under the individual pledge and no statistically significant difference under pledge plus school competition. Moreover, the point estimates for the individual pledge and pledge plus school competition are not statistically different, which suggests that students in the pledge plus school competition arm are not less likely to report their peers.

Alternative Smoking Abstinence Measures We analyze alternative smoking abstinence measures for 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 as an outcome variable (Column 3). Fourth, we combine the verified tobacco abstinence measure and student's self-reported last cigarette: we use the thresholds of two weeks and a month (Columns 4 and 5). Although some of the program effects are imprecisely estimated, the estimates are similar using these alternative outcome measures, suggesting the effectiveness of the treatment in preventing smoking initiation.

5 Conclusion

Non-monetary incentvies 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 6.5% after 3 months, an effect size that is sustained into the next school year (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.

Interestingly, adding school competition did not 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.

We also 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 6 to 10 percentage points for each additional friend in the person's class who abstained. Both social commitment and social norms interventions have proven effective in other settings. There may be opportunities to enhance our commitment intervention by leveraging these social forces.

By focusing on more rural schools in Indonesia, our study also 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 the world leader 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 require local experimentation and action to reduce its substantial tobacco-attributable disease burden. Our study indicates that our tobacco prevention program, implemented in two districts, is relatively low-cost and might be readily scalable elsewhere in Indonesia, and perhaps in other settings too. As evidence of the sustainability of the program, one of the two districts 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.

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.

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 and 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.

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

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.

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.

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.

Homonoff, Tatiana A, "Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use," *American Economic Journal: Economic Policy*, 2018.

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 tobacco 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.

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

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.

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.

Nichter, Mimi, Mark Nichter, Retna Siwi Padmawati, and Nawi Ng, "Developing a smoke free household initiative: an Indonesian case study," *Acta Obstetricia et Gynecologica Scandinavica*, 2010, *89* (4), 578–581.

_, Siwi Padmawati, Mahesto Danardono, Nawi Ng, Yayi Prabandari, and Mark Nichter, "Reading culture from tobacco advertisements in Indonesia," *Tobacco Control*, 2009, *18* (2), 98–107.

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.

Rabin, Matthew, "Healthy habits: Some thoughts on the role of public policy in healthful eating and exercise under limited rationality," in Adam Oliver, ed., *Behavioural Public Policy*, Cambridge University Press, 2013, chapter 5, pp. 115–139.

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.

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.

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.

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.

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.

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.

_ , *World health statistics 2012*, World Health Organization, 2012.

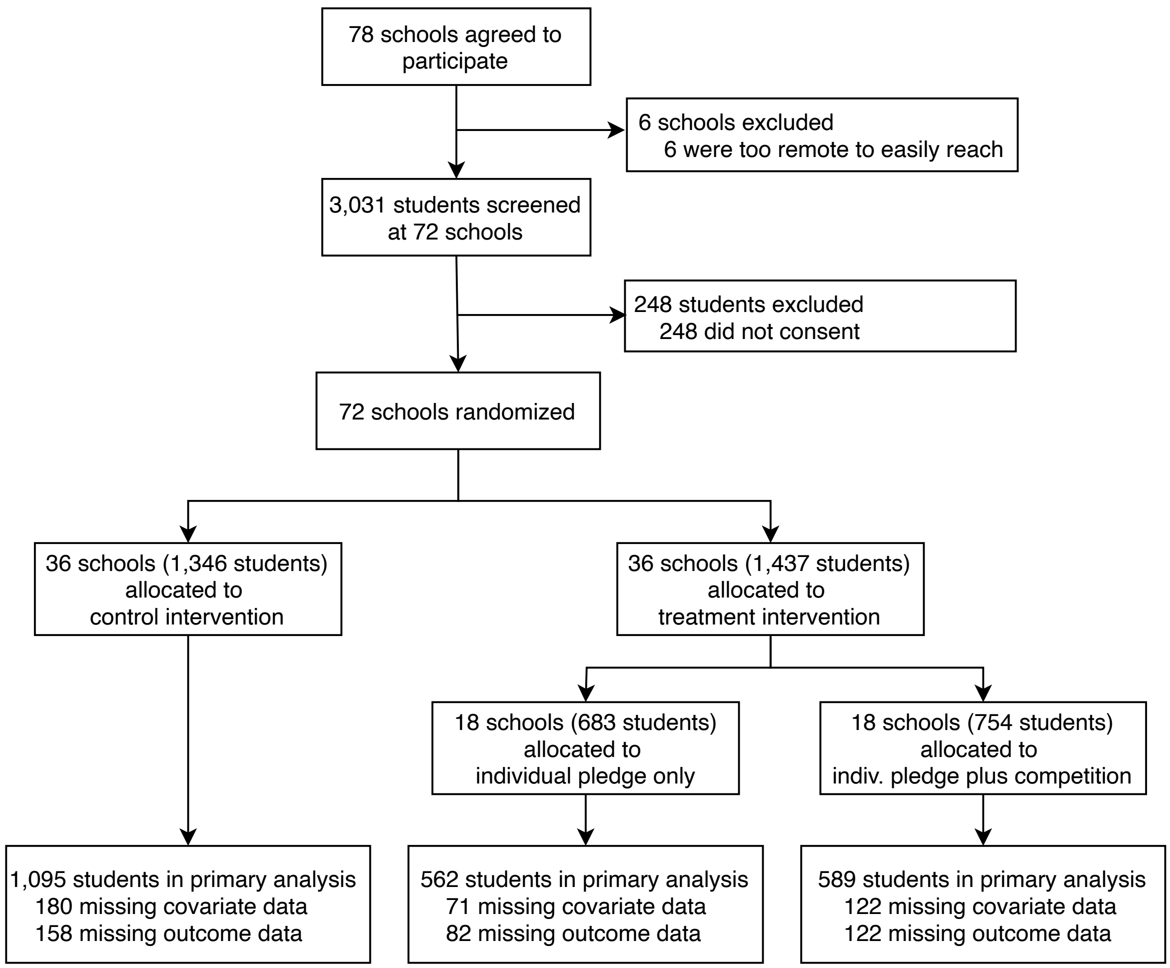
_ , "Global Youth Tobacco Survey: Indonesia Report 2014," Technical Report, World Health Organization Regional Office for Southeast Asia, New Delhi, India 2015.

_ , "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.

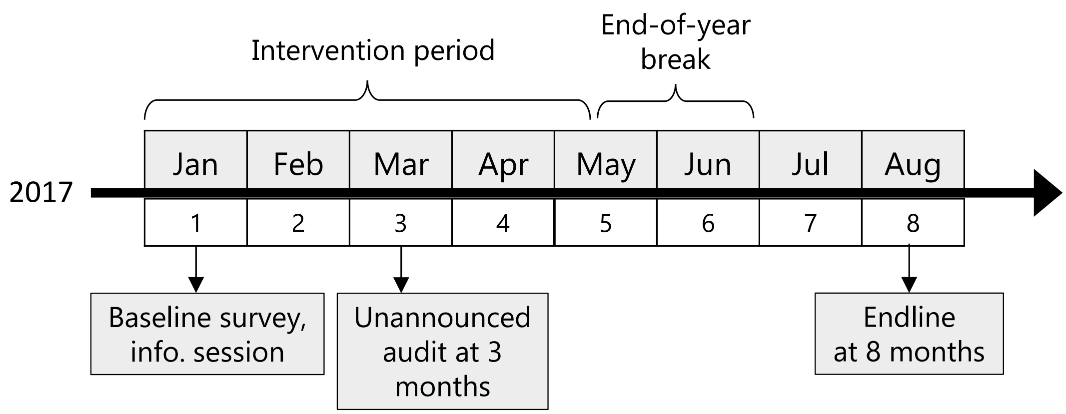
Tables and Figures

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 began in July 2017.

Table 1: Baseline characteristics and balance under pair matching

	(1) Control	(2) Treated (Pooled)	(3) Individual pledge	(4) School competition	(5) Adjusted difference (Pooled)	(6) Adjusted difference (Ind. pledge)	(7) Adjusted difference (Sch. comp)
Panel A. School characteristics							
Distance to Jogjakarta	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)
Number of student	11.372	9.140	9.858	11.363	-2.565	-1.228	-3.744
council members	(10.930)	(10.702)	(9.455)	(12.177)	(2.627)	(3.691)	(2.500)
Avg. mathematics score	56.932	52.516	55.238	12.031	-3.962	-2.791	-2.623
	(14.536)	(13.273)	(13.873)	(55.564)	(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)					0.264	0.531	0.428
Panel B. Student characteristics							
Verified smoking	0.785	0.777	0.753	0.798	-0.006	-0.028	0.018
abstinence	(0.411)	(0.416)	(0.431)	(0.401)	(0.033)	(0.039)	(0.033)
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)
Number of friends 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	0.989	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 household	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)					0.102	0.041	0.167

Notes: The control group has 1153 students and the treated 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 fixed effects. Column 5 compares the control and pooled treatment. Columns 6 and 7 compare the sub-treatment arm separately. Column 8 tests the equality of coefficients in Columns 6 and 7. 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)
Panel A. Pooled estimates			
Treat	-0.008	-0.006	0.008
	(0.027)	(0.013)	(0.014)
Treat × Post	0.037	0.044**	0.044*
	(0.023)	(0.022)	(0.022)
Wave = 2	-0.009	-0.017	-0.017
	(0.018)	(0.018)	(0.018)
Wave = 3	-0.034**	-0.043**	-0.043**
	(0.017)	(0.016)	(0.017)
Pair FE	Yes	Yes	Yes
Student and class baseline outcomes	No	Yes	Yes
Additional controls	No	No	Yes
No. observations		7207	
R-squared	0.07	0.43	0.44
Dep. variable mean for control group		0.786	
Panel B. Sub-treatment estimates			
Pledge	-0.012	-0.016	-0.004
	(0.042)	(0.017)	(0.017)
Competition	-0.004	0.003	0.019
	(0.032)	(0.018)	(0.020)
Pledge × Post	0.051**	0.052**	0.051**
	(0.023)	(0.022)	(0.023)
Competition × Post	0.025	0.036	0.037
	(0.031)	(0.030)	(0.030)
Wave = 2	0.025**	0.026**	0.026**
	(0.012)	(0.012)	(0.012)
Wave = 3	-0.034**	-0.043**	-0.043**
	(0.017)	(0.016)	(0.017)
Pair FE	Yes	Yes	Yes
Student and class baseline outcomes	No	Yes	Yes
Additional controls	No	No	Yes
No. observations		7207	
R-squared	0.07	0.43	0.44
Dep. variable mean for control group		0.786	
Equality of coefficient (p-value)	0.403	0.578	0.621

Notes: *Post* takes the value one for the follow-up period. *Treat* (*Pledge*, *Competition*) takes the value one if the school is randomized into treatment (the commitment pledge arm or the school competition arm). Column 1 includes pair fixed effects. Column 2 includes pair fixed effects and baseline outcome. Column 3 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. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 3: Potential mechanisms

	(1) Tobacco knowledge	(2) Discuss with friends	(3) Discuss with parents	(4) Belief that parents do not accept smoking	(5) Smoking intentions
Panel A. Pooled estimates					
Treatment	0.125	-0.041**	-0.026*	-0.000	0.062
	(0.096)	(0.017)	(0.013)	(0.020)	(0.085)
Treatment × Post	-0.020			0.024	-0.135
	(0.222)			(0.022)	(0.130)
Wave = 3	0.005			-0.026	0.249**
	(0.168)			(0.018)	(0.101)
No. observations	5454	2589	2589	4605	5337
R-squared	0.234	0.069	0.045	0.096	0.209
Dep. variable mean for control group	5.618	0.166	0.087	0.701	1.139
Panel B. Sub-treatment estimates					
Pledge	0.118	-0.040	-0.026	0.019	0.084
	(0.138)	(0.025)	(0.021)	(0.026)	(0.105)
Competition	0.132	-0.042**	-0.025	-0.019	0.042
	(0.132)	(0.020)	(0.017)	(0.025)	(0.112)
Pledge × Post	0.021			0.016	-0.178
	(0.294)			(0.028)	(0.140)
Competition × Post	-0.056			0.031	-0.097
	(0.239)			(0.026)	(0.162)
Wave = 3	0.005			-0.026	0.249**
	(0.168)			(0.018)	(0.101)
No. observations	5454	2589	2589	4605	5337
R-squared	0.234	0.069	0.045	0.096	0.209
Dep. variable mean for control group	5.618	0.166	0.087	0.701	1.139

Notes: Tobacco knowledge is based on a score on a series of 9 questions to assess students' awareness of the health risks of tobacco use. Belief that parents do not accept smoking is based on students' report of whether their parents would find it (very) unacceptable if they smoked. Smoking intentions is measured as the probability that the student expects to be smoking at all in 3 months, reported on a scale from 0 to 10. Discussion with friends and parents in the past month prior to the survey are only asked at the 8-month survey. *Post* takes the value one for the follow-up period. *Treat* (*Pledge*, *Competition*) takes the value one if the school is randomized into treatment (the commitment pledge arm or the school competition arm). The models include 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 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	0.011	0.000	-0.983*
	(0.024)	(0.013)	(0.550)
Treatment × Post	-0.001	-0.002	-0.866
	(0.024)	(0.016)	(0.758)
Wave = 2	-0.132***	-0.021*	
	(0.020)	(0.011)	
Wave = 3	-0.039**	-0.008	2.191***
	(0.019)	(0.013)	(0.662)
No. observations	7829	7211	2698
R-squared	0.091	0.037	0.253
Dependent variable mean	0.133	0.078	7.042
Panel B. Sub-treatment estimates			
Pledge	-0.010	-0.017	-0.616
	(0.023)	(0.016)	(0.670)
Competition	0.031	0.016	-1.336*
	(0.034)	(0.018)	(0.702)
Pledge × Post	0.017	0.003	-0.737
	(0.026)	(0.018)	(0.842)
Competition × Post	-0.016	-0.007	-1.018
	(0.030)	(0.022)	(0.874)
Wave = 2	-0.132***	-0.021*	
	(0.020)	(0.011)	
Wave = 3	-0.039**	-0.008	2.198***
	(0.019)	(0.013)	(0.662)
No. observations	7829	7211	2698
R-squared	0.092	0.037	0.254
Dependent variable mean	0.133	0.078	7.042

Notes: Missing cotinine test takes the value one if an individual only has self-reported tobacco abstinence status. Self-reported abstinence and failed cotinine test takes the value one if an individual reported tobacco abstinence but failed the cotinine test. *Post* takes the value one for the follow-up period. *Treat* (*Pledge*, *Competition*) takes the value one if the school is randomized into treatment (the commitment pledge arm or the school competition arm). The models include 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: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

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	0.004	0.001
	(0.015)	(0.016)	(0.012)	(0.015)	(0.014)
Treatment × Post	0.033	0.044*	0.023	0.066***	0.033
	(0.023)	(0.024)	(0.019)	(0.025)	(0.023)
Wave = 2	-0.037**	0.074***	-0.011	-0.032	-0.761***
	(0.017)	(0.023)	(0.013)	(0.021)	(0.023)
Wave = 3	-0.062***	-0.015	-0.046***	-0.056***	-0.062***
	(0.016)	(0.019)	(0.014)	(0.018)	(0.017)
No. observations	7380	8074	7208	6328	6328
R-squared	0.400	0.496	0.444	0.461	0.680
Dep. var. mean for control group	0.786	0.682	0.846	0.734	0.784
Panel B. Sub-treatment estimates					
Pledge	-0.008	0.001	-0.009	-0.007	-0.012
	(0.017)	(0.022)	(0.013)	(0.017)	(0.017)
Competition	0.020	0.013	0.017	0.015	0.014
	(0.021)	(0.019)	(0.019)	(0.021)	(0.018)
Pledge × Post	0.039	0.038	0.033*	0.073***	0.050*
	(0.025)	(0.029)	(0.017)	(0.025)	(0.026)
Competition × Post	0.027	0.050*	0.014	0.060*	0.016
	(0.031)	(0.028)	(0.029)	(0.034)	(0.028)
Wave = 2	-0.037**	0.074***	-0.011	-0.033	-0.761***
	(0.017)	(0.023)	(0.013)	(0.021)	(0.023)
Wave = 3	-0.062***	-0.015	-0.046***	-0.056***	-0.062***
	(0.016)	(0.019)	(0.014)	(0.018)	(0.017)
No. observations	7380	8074	7208	6328	6328
R-squared	0.400	0.496	0.444	0.461	0.680
Dep. variable mean for control group	0.786	0.682	0.846	0.734	0.784

Notes: Column 1 assumes students with missing cotinine test smoked. Column 2 assumes students with missing cotinine test had the same outcome as their baseline outcome. Column 3 uses the cotinine test only. Columns 4 and 5 combine student's biochemically verified tobacco abstinence and self reported the last time he smoked (more than 2 weeks ago and more than 1 month ago respectively). Column 6, the frequency of being reported, is based on the number of times a student is seen smoking by classmates during the 2 weeks prior to the 8-month survey. *Post* takes the value one for the follow-up period. $Treat$ (*Pledge*, *Competition*) takes the value one if the school is randomized into treatment (the commitment pledge arm or the school competition arm). The models include 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

A.1 Sample size calculation

We conducted sample size calculations for a cluster randomized trial with an individual-level binary outcome.[30] 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.[31]

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.

[30] 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.

[31] 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.

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 within pairs. We then randomly assigned one unit to the treatment group and one unit to the control group. Among the treatment group schools, we then paired the two most similar schools by Mahalonobis distance and randomly assigned one to the commitment arm and one to the commitment plus competition arm. Thus, we randomly allocated 36 schools to the control arm, 18 schools to the commitment arm, and 18 schools to the commitment plus competition arm (Table A.1).

Table A.1: Number of schools by subdistrict

District	Control	Commitment	School competition
Sleman	17	12	7
Kulon Progo	19	6	11
N	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 exam scores in mathematics. We combined information on school characteristics and enrollment (age, gender distribution, class size, and students scores on national exams) 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: School characteristics at baseline

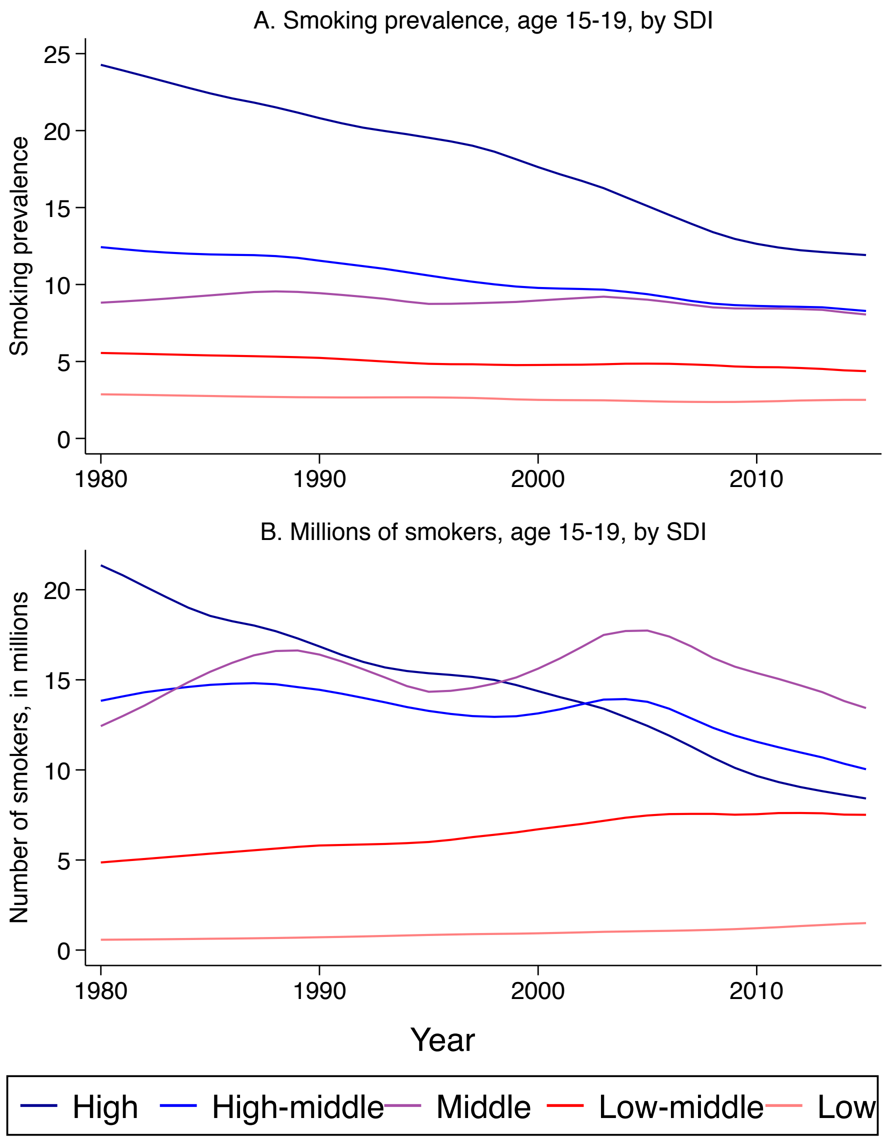
	(1) Pooled	(2) Individual pledge	(3) School competition
Distance to Yogjakarta	0.929	0.325	0.378
Student-to-teacher ratio	0.779	0.829	0.589
Average mathematics score	0.269	0.563	0.488
No. classes	0.924	0.293	0.348
No. observations	72	72	72

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.

Appendix B Additional figures and tables

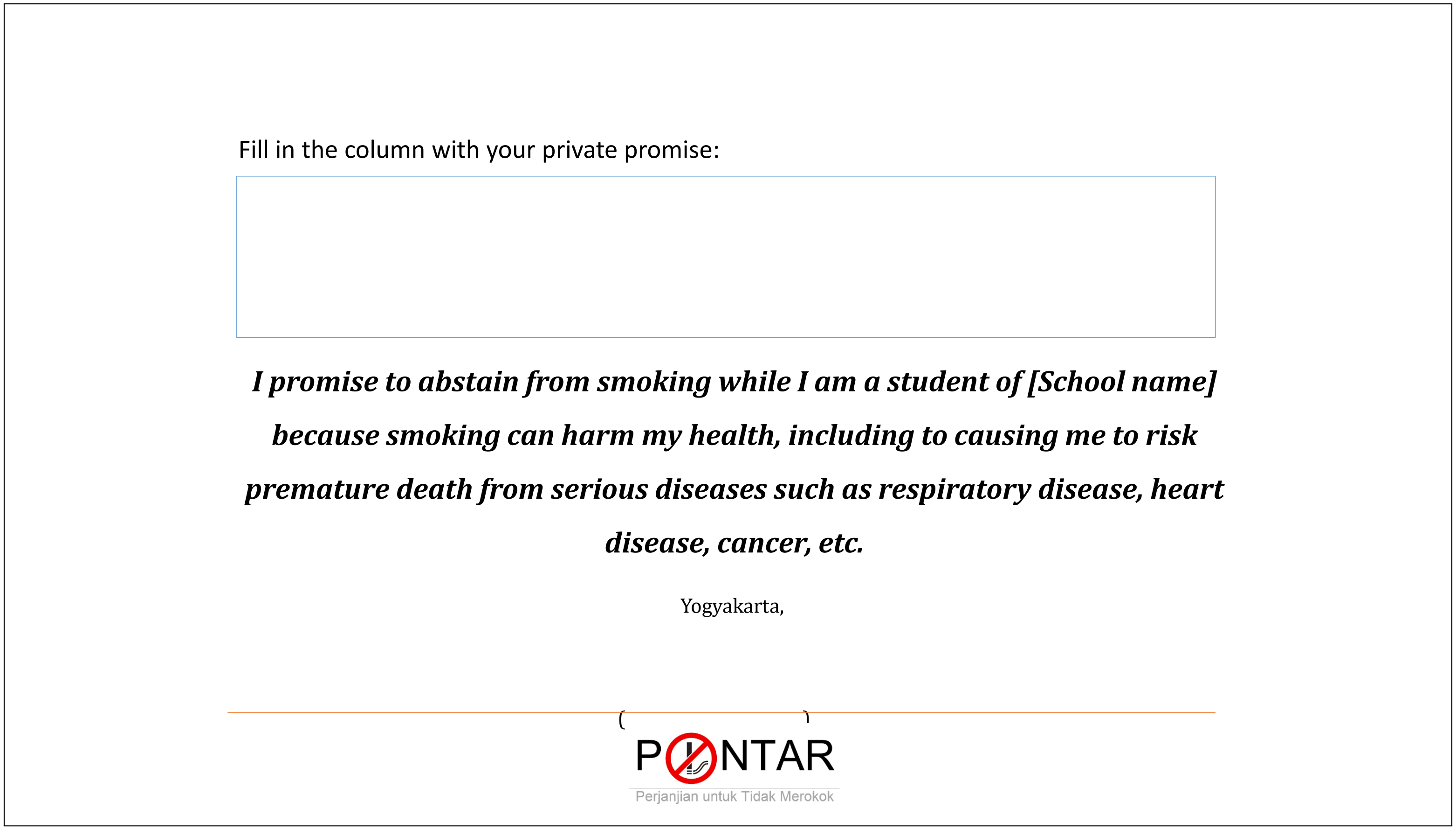
Figure B.1: Smoking trends among youth



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.

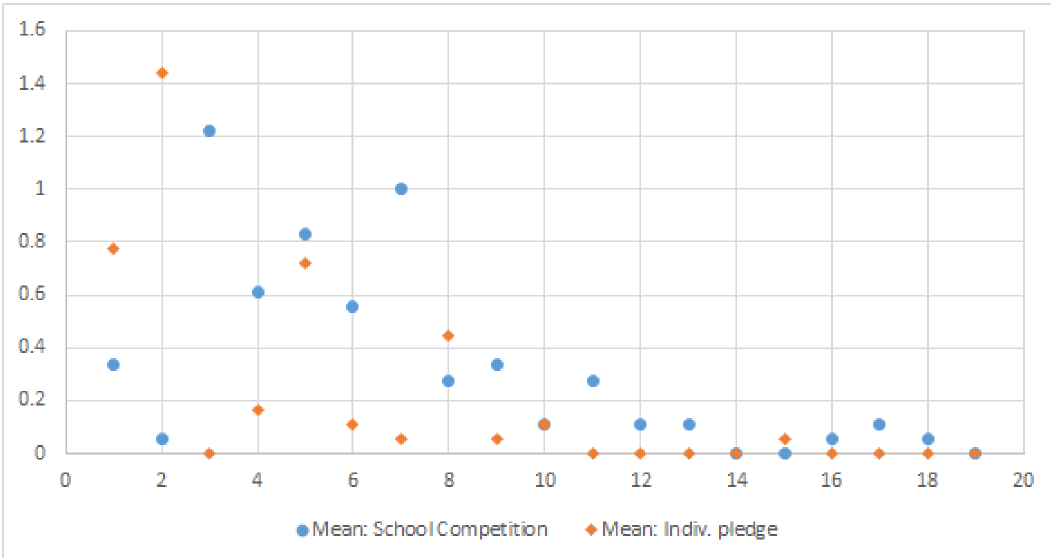
Figure B.3: Commitment pledge for parents

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.

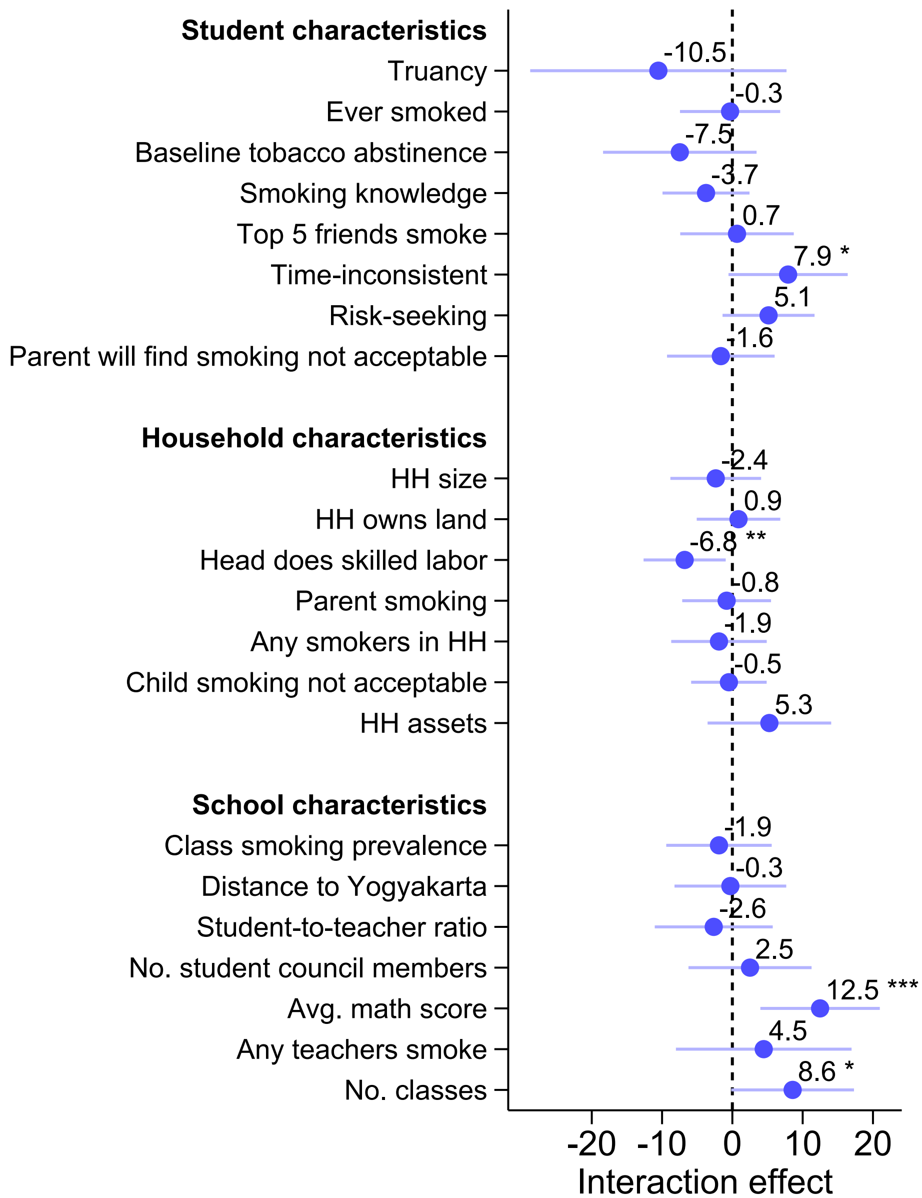
[Parent's signature] [Child's signature]

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



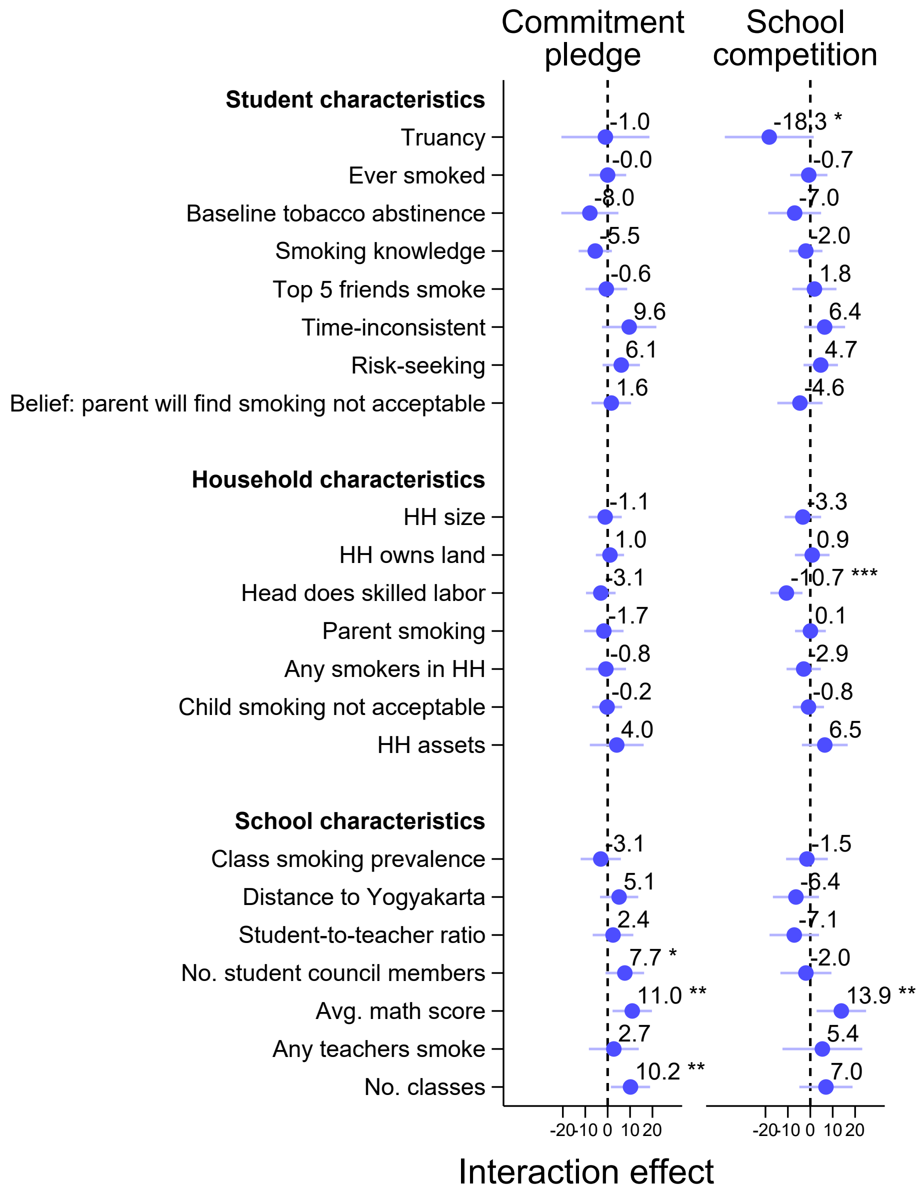
Notes: This figure shows the average number of smokers reported in each sub-treatment arm in each week after the intervention was implemented. Differences between the two sub-treatment arms are not statistically significant.

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: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.3: Program effects on tobacco abstinence by follow-up

	(1)	(2)	(3)	(4)	(5)	(6)
	3 months			8 months		
	Unadjusted	Partially adjusted	Fully adjusted	Unadjusted	Partially adjusted	Fully adjusted
Panel A. Pooled estimates						
Treatment	0.037	0.032*	0.061***	0.038	0.027	0.038**
	(0.038)	(0.019)	(0.017)	(0.039)	(0.017)	(0.017)
No. observations	2491	2491	2491	2344	2344	2344
R-squared	0.002	0.221	0.267	0.002	0.237	0.299
Dep. var. mean for control group			0.786			
Panel B. Sub-treatment estimates						
Pledge	0.025	0.035	0.051**	0.024	0.004	0.020
	(0.046)	(0.026)	(0.024)	(0.046)	(0.027)	(0.022)
Competition	0.047	0.029	0.071***	0.051	0.047**	0.057**
	(0.046)	(0.026)	(0.023)	(0.054)	(0.022)	(0.023)
No. observations	2491	2491	2491	2344	2344	2344
R-squared	0.002	0.221	0.267	0.002	0.238	0.299
Dep. var. mean for control group			0.786			
Equality of coefficients (p-value)	0.674	0.865	0.511	0.659	0.213	0.234

Notes: Partially adjusted includes pair fixed effects, baseline smoking status. Fully adjusted includes: 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 are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.4: Students' opinion of the program

	(1)	(2)	(3)	(4)	(5)	(6)
	Students			Student council		
	↑ own abstinence	↑ friends' abstinence	Discuss with parents	↑ own abstinence	↑ friends' abstinence	Discuss with parents
Competition	-0.012	0.015	0.045	0.042*	0.037	0.033
	(0.036)	(0.039)	(0.036)	(0.021)	(0.093)	(0.077)
No. observations	1411	1411	1411	243	243	243
R-squared	0.003	0.001	0.004	0.028	0.002	0.002

Notes: Subdistrict FE included, standard errors clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.5: Student council's beliefs about mechanisms for tobacco abstinence

	(1) Sign	(2) Demerit points	(3) Student gov reporting	(4) Smoke-free student gov	(5) Friend reporting	(6) Smoke-free friend	(7) Teacher report	(8) Parent report	(9) Random test	(10) Make cigarettes hard to get	(11) Better knowledge
Panel A. Pooled estimates											
Treatment	0.043	-0.004	0.238***	0.044	0.133*	0.016	0.128*	0.056	0.141**	0.028	0.065
	(0.059)	(0.082)	(0.075)	(0.075)	(0.076)	(0.066)	(0.067)	(0.067)	(0.062)	(0.077)	(0.059)
Obs.											
R-squared	397	396	397	397	397	397	397	397	397	397	397
	0.008	0.001	0.058	0.004	0.018	0.000	0.025	0.023	0.039	0.001	0.008
Panel B. Sub-treatment estimates											
Pledge	-0.069	-0.057	0.186*	-0.046	0.034	0.008	0.100	0.019	0.062	-0.082	0.014
	(0.076)	(0.104)	(0.094)	(0.096)	(0.094)	(0.077)	(0.082)	(0.087)	(0.077)	(0.103)	(0.075)
Competition	0.134**	0.038	0.280***	0.116	0.213**	0.023	0.152**	0.086	0.205***	0.116	0.106*
	(0.056)	(0.102)	(0.085)	(0.082)	(0.081)	(0.084)	(0.070)	(0.073)	(0.061)	(0.079)	(0.060)
No. observations	397	396	397	397	397	397	397	397	397	397	397
R-squared	0.046	0.007	0.063	0.023	0.038	0.001	0.027	0.026	0.055	0.026	0.018

Notes: Subdistrict FE included, standard errors clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.6: Heterogeneous program effects on smoking abstinence by student characteristics

	(1) Pooled	(2) Pledge	(3) 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. Continuous variables are dichotomized into the median split. The table shows the interaction effects estimated from a difference-in-differences model that adjusts for the full set of baseline covariates. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.7: Heterogeneous program effects on smoking abstinence by household characteristics

	(1) Pooled	(2) Pledge	(3) 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. Continuous variables are dichotomized into the median split. The table shows the interaction effects estimated from a difference-in-differences model that adjusts for the full set of baseline covariates. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.8: Heterogeneous program effects on smoking abstinence by school characteristics

	(1) Pooled	(2) Pledge	(3) 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. Continuous variables are dichotomized into the median split. The table shows the interaction effects estimated from a difference-in-differences model that adjusts for the full set of baseline covariates. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.9: 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 6 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.

B.1 Peer effects

We explore the role of students' peer group on smoking, due to the nature of smoking as a social activity during adolescence. 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 B.10). Our second IV analysis exploits the quasi-random class assignment in 8th grade to provide plausibly exogenous variation in friendship networks. 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 B.10). 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.[32]

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

[32] 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. 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.

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 B.10, 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 B.10, columns 3-4, Panel A).[33] 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 B.11). 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 peers predict are 10% more likely to smoke are 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 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

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

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 B.12 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 B.10: 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 B.11: Program effects on smoking abstinence by peers' smoking status

	(1) Share of 5 friends who abstained and in same 8th grade class	(2) Share of 7th grade classmates who abstained and in same 8th grade class	(3) 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*	0.006	-0.014
	(0.025)	(0.108)	(0.028)
Treat × Characteristic	0.042***		-0.042***
	(0.014)		(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**	0.027	-0.000
	(0.029)	(0.107)	(0.032)
Competition × Post	0.035	-0.036	-0.026
	(0.033)	(0.164)	(0.038)
Pledge × Characteristic	0.037**		-0.044**
	(0.018)		(0.018)
Competition × Characteristic	0.049**		-0.039**
	(0.024)		(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: Fully adjusted for baseline covariates. Standard errors are clustered at the school level. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table B.12: 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: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.